Capsid protein structure, self-assembly, and processing reveal morphogenesis of the marine virophage mavirus

Diana Born1, Lukas Reuter2, Ulrike Mersdorf2, Melanie Mueller2, Matthias G. Fischer1, Anton Meinhart2, and Jochen Reinstein1,3

*Department of Biomolecular Mechanisms, Max Planck Institute for Medical Research, 69120 Heidelberg, Germany

Edited by Michael G. Rossmann, Purdue University, West Lafayette, IN, and approved May 30, 2018 (received for review March 28, 2018)

Virophages have the unique property of parasitizing giant viruses within unicellular hosts. Little is understood about how they form infectious virions in this tripartite interplay. We provide mechanistic insights into assembly and maturation of mavirus, a marine virophage, by combining structural and stability studies on capsomers, virus-like particles (VLPs), and native virions. We found that the mavirus protease processes the double jelly-roll (DRJ) major capsid protein (MCP) at multiple C-terminal sites and that these sites are conserved among virophages. Mavirus MCP assembled in Escherichia coli in the absence and presence of penton protein, forming VLPs with defined size and shape. While quantifying VLPs in E. coli lysates, we found that full-length rather than processed MCP is the competent state for capsid assembly. Full-length MCP was thermally more labile than truncated MCP, and crystal structures of both states indicate that full-length MCP has an expanded DRJ core. Thus, we propose that the MCP C-terminal domain serves as a scaffolding domain by adding strain on MCP to confer assembly competence. Mavirus protease processed MCP more efficiently after capsid assembly, which provides a regulation mechanism for timing capsid maturation. By analogy to Sputnik and adenovirus, we propose that MCP processing renders mavirus particles infection competent by loosening interactions between genome and capsid shell and destabilizing pentons for genome release into host cells. The high structural similarity of mavirus and Sputnik capsid proteins together with conservation of protease and MCP processing suggest that assembly and maturation mechanisms described here are universal for virophages.

Virophages are parasites of giant viruses within protists. They reduce giant virus production and increase host cell survival. They provide a defense system for protists against giant viruses in diverse environments, likely with ecological relevance for protist populations. To understand the remarkable virophage life cycle, it is crucial to investigate how they assemble into infectious particles and which processes require interactions with giant virus and host. We examined the marine virophage mavirus to show that its major and minor capsid proteins assemble into virus-like particles in the absence of specific host or viral factors. Subsequently, the virophage-encoded protease processes the major capsid protein to prepare virions for infection.

Author contributions: D.B. and J.R. designed research; D.B., L.R., U.M., M.M., M.G.F., and A.M. performed research; D.B., U.M., M.G.F., A.M., and J.R. analyzed data, and D.B. and J.R. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

Significance

Virophages are parasites of giant viruses within protists. They reduce giant virus production and increase host cell survival. They provide a defense system for protists against giant viruses in diverse environments, likely with ecological relevance for protist populations. To understand the remarkable virophage life cycle, it is crucial to investigate how they assemble into infectious particles and which processes require interactions with giant virus and host. We examined the marine virophage mavirus to show that its major and minor capsid proteins assemble into virus-like particles in the absence of specific host or viral factors. Subsequently, the virophage-encoded protease processes the major capsid protein to prepare virions for infection.
and to act as a defense system against CoV has been proposed to be of ecological importance for zoonotic populations (14).

Results
Mavirus Protease Processes MCP at Multiple Sites in Vitro. The virophage protease is suggested to be involved in capsid maturation, as observed for other eukaryotic viruses of the PRD1-adenovirus lineage (15, 20). To investigate its function in the virophage mavirus, we recombinantly expressed the mavirus capsid protease (MVP), MCP, and penton protein. Purified MVP was present as a dimer, MCP as a trimeric capsomer, and penton protein as a pentamer (SI Appendix, Fig. S1 A–C). MVP cleaved recombinant MCP at multiple sites but not recombinant penton protein (Fig. 1A and SI Appendix, Fig. S1E). We used electrospray-ionization quadrupole time-of-flight mass spectrometry (MS) to identify MCP cleavage fragments from this in vitro turnover (SI Appendix, Fig. S1F and Table S1). The largest MCP fragment comprised residues (res.) 1–516 of full-length MCP (flMCP), which we termed MCPΔC (Fig. 1B). The C-terminal domain (CTD) of flMCP (res. 517–606) seems to be cleaved from flMCP first in this in vitro experiment and was further processed at three additional sites, generating MCPCT1 (res. 517–567), MCPCT2 (res. 568–580), MCPCT3 (res. 581–592) and MCPCT4 (res. 593–606). Three of four cleavage sites in mavirus MCP are preceded by a dipeptide motif (GG | X, Fig. 1B), the other cleavage occurred within GY | G. A sequence alignment of diverse virophage MCPs revealed at least four GG | X or GX | G motifs in 10 of 12 MCPs in the C terminus (SI Appendix, Fig. S2), suggesting that these cleavage motifs are highly conserved. Since the mavirus MCP is cleaved at C-terminal sites that are preserved in diverse virophage MCPs, we propose that processing by the cysteine protease has a key function in the virophage life cycle.

MCP and MV13 Are Processed in Native Mavirus Virions. We analyzed native mavirus particles by MS peptide-mass fingerprinting (PMF) on excised SDS/PAGE bands to investigate MCP processing (Fig. 1C). MVP was detected as truncated ∼55-kDa fragment, and the presence of MCPΔC was confirmed by MS PMF (Fig. 1C, Left). Notably, no MCP C-terminal fragments could be identified by MS PMF in native mavirus virions, potentially due to both the specific MS method used and the low concentration of available native mavirus virions, but we can also not exclude that MCP C-terminal fragments are absent in these viruses. In addition, five further native mavirus gene products were identified: MV02 (integrase), two cleavage fragments of MV13 (predicted α/β hydrolase), MV14 (unknown function), MV15 (FtsK/HerA-type ATPase), and MV17 (penton protein) (9). The SDS/PAGE fragments at ∼20 kDa and ∼60 kDa correspond to the N- and C-terminal part, respectively, of the 80.6-kDa MV13. Interestingly, the MV13 sequence between MV13N and MV13C that was not covered by tryptic peptide masses (res. 476–499) contains a potential MVP cleavage site, GK | G. We also identified MVP in native mavirus virions by immunoblotting (Fig. 1C, Right), supporting the hypothesis that it is the relevant protease for processing not only MCP but also MV13 in mavirus virions.

MCP Processing Results in Jelly-Roll Core Compaction and Tower Expansion. To elucidate the effect of processing on mavirus MCP regarding structure and stability, we recombinantly generated MCPΔC and compared it with flMCP. Both proteins form trimers in solution (SI Appendix, Fig. S1 A and D), but MCPΔC showed higher melting temperatures (Tm) than flMCP (SI Appendix, Fig. S3A), indicating higher thermal stability of MCPΔC. To answer whether processing alters the MCP conformation, we determined crystal structures of both MCP states (SI Appendix, Table S2). Although the crystal density map at 2.5-Å resolution was unambiguous only for res. 2–504, the presence of the entire protein was confirmed by MS on washed and dissolved crystals. Thus, the C-terminal domain is likely flexible. The flMCP dataset contains a trimeric capsomer in the asymmetric unit that shows a central pore along the quasi-fivefold axis (SI Appendix, Fig. S3 C and D). Protomers exhibit the DJR fold (Fig. 2A and B and SI Appendix, Fig. S3B) and are triply symmetric with the highest root-mean-square deviation on Cα positions (Cα rmsd) of 0.135 Å. The capsomer includes a base and a tower, which would face the capsid interior and exterior, respectively, (Fig. 2A) by analogy to the Sputnik capsid structure (16). The base contains six three-stranded β-sheets, three of which are “pedestal connectors” (PCs) as described for the adenoviral hexon (21), which link JR1 and JR2 within each protomer, and the other are “basal interprotomer sheets” (BISs), which stabilize the trimer (SI Appendix, Fig. S3 C and D).

Similar to flMCP, the crystal structure of MCPΔC at 1.5-Å resolution had a trimeric capsomer located in the asymmetric unit. Superposed flMCP and MCPΔC structures are highly similar (Cα rmsd = 1.21 Å), and both show intact DJR core, tower, and base (Fig. 2C). Both datasets allowed model building up to T504 close to the MVP cleavage site. In contrast to flMCP, the C-terminal region (res. 509–516) could be modeled for MCPΔC, which forms an additional β-strand (M2) at each BIS (SI Appendix, Fig. S3E). The C-terminal carboxylate group of MCPΔC is buried inside the capsomer pore and tightly bound by hydrogen bonds, suggesting that the observed conformation at the BIS was adopted after cleavage. noteworthy, poor and ambiguous electron density for an additional BIS strand was also observed for MCPΔC but could not be traced (SI Appendix, Fig. S3E). Therefore, this strand could be replaced by a dMCP. Even in flMCP than in MCPΔC, suggesting that the MCP C terminus undergoes conformational changes at the BIS upon processing.

Although the largest local differences between flMCP and MCPΔC were observed in loops closing the central capsomer pore (SI Appendix, Fig. S3F), global changes were observed within capsomer domains (Fig. 2 C and D). The JR1 core is 0.9 Å closer at the capsomer center of mass (COM) in MCPΔC than in flMCP. In contrast, the JR2 tower (T2) is 1.5 Å further away from the capsomer COM in MCPΔC than in flMCP. Thus, the DJR core contracted; whereas T2 expanded, in MCPΔC compared with flMCP (Movie S1). These differences are not due to crystal artifacts, since residues involved in crystal contacts (identified by the PISA server; ref. 22) are almost identical for flMCP and MCPΔC.
Using the DALI server (23), the highest structural similarity between mavirus (colors as in SI Appendix, Table S3) and Sputnik penton proteins (PDB ID code 3J26; ref. 16) have high structural similarity (DALI Z score: 22.1; highest Cα rmsd of protomers: 5.07 Å). Superposing only the SJR of a mavirus and Sputnik penton protomer (highest Cα rmsd: 3.70 Å, Fig. 3C) revealed that the ID position relative to the SJR is the most distinctive feature, suggesting a certain degree of flexibility in the connecting loops. Notably, individual superposition of the mavirus and Sputnik IDs revealed a striking fold conservation (highest Cα rmsd: 2.39 Å, Fig. 3D). Such strong structural conservation combined with high sequence variability of the surface-exposed ID among virophages (12) supports the penton’s putative role in host cell or giant virus recognition.

MCP Spontaneously Assembles into Mavirus-Sized Virus-Like Particles in Escherichia coli. Based on the conservation of the cysteine protease (12) and the structural similarity of truncated mavirus and Sputnik MCP, we hypothesize that MCP processing is important for the formation of infectious virophage particles. To study mavirus capsid assembly, we expressed flMCP and MCPΔC in E. coli and analyzed lysates. Surprisingly, virus-like particles (VLPs) with diameters of 60–75 nm were observed independently of whether flMCP or MCPΔC was expressed (SI Appendix, Fig. S5B). These particles were comparable in size to native mavirus (Fig. 4A, d = 70–80 nm), indicating that exclusively MCP but no specific factors from mavirus, CroV, or host cell are required to form capsids. To compare the assembly efficiency of flMCP and MCPΔC, we used E. coli lysates containing the same amounts of expressed MCP and supplemented gold particles (GPs) as an internal VLP quantification standard for negative stain EM. For flMCP, 61 times more VLPs were present per micrograph than for MCPΔC (SI Appendix, Fig. S5A). In contrast to native mavirus virions, both flMCP- and MCPΔC-VLPs appeared more spherical (Fig. 4B), like “waffle ball” capsids that vertex protrude (29). Notably, purified MCPΔC-VLPs showed a high number of isolated capsomers (Fig. 4C), although these should have been removed during purification, and most particles appeared damaged. These observations indicate that MCPΔC-VLPs assembled less efficiently than flMCP-VLPs and are less stable under harsh staining conditions for EM. We compared the protein melting temperature Tm of capsomers and VLPs by calculating a difference Tm (ΔTm = Tm(VLP) – Tm(capsermer)). We used ΔTm to assess how MCP stability is affected when capsomers assemble into VLPs. Interestingly, Tm1 and Tm2 of flMCP shifted to higher temperatures in assembled VLPs compared with capsosomes by ΔTm of at least +4.7 ± 0.1 °C and +2.1 ± 0.1 °C, respectively (SI Appendix, Fig. S6A). This indicates that flMCP capsomers gained stability through VLP assembly. In contrast, MCPΔC stability marginally rose upon VLP assembly.

These results suggest that DJR core contraction and tower expansion in MCP derive from processing by MVP.

High Structural Conservation of Virophage MCPs Extends to Processed C Termini. Using the DALI server (23), the highest structural similarity to mavirus flMCP was detected for the Sputnik virophage MCP, followed by the DJR proteins of members of the proposed order Megavirales (24), whereas the virophage MCP strongly diverged from both the adenoviral hexon and the MCP of bacteriophage PRD1 (SI Appendix, Table S3), as already proposed previously (25). Similar to mavirus MCP, Sputnik MCP was described as C-terminally truncated in mature capsids (16) and is called Sputnik MCPΔC in the following. The overall architecture including tower, DJR core, and base is conserved between mavirus and Sputnik MCPΔC (Fig. 3A, Cα rmsd: 4.36 Å). Interestingly, the MCPΔC C terminus is buried in the capsomer pore in both virophages (SI Appendix, Fig. S3G). The high structural similarity of mavirus and Sputnik MCPΔC including the position of processed C termini suggests that the implications of MCP processing on capsomer and capsid are conserved in virophages.

Penton Protein Structures Are Preserved in Virophages, Especially the Variable Insertion Domain. We explored whether structural and functional conservation can be extended from MCP to the second relevant virophage capsid protein, the penton protein, which could potentially be involved in host cell or giant virus recognition (12, 16). The crystal structure of mavirus penton protein at 2.7 Å resolution (SI Appendix, Table S2) comprises one pentamer in the asymmetric unit. Rs = 2.30 Å could be modeled into the electron density map for all protomers, which contain a typical SJR domain and an insertion domain (ID). The 139-residue ID is located between SJR strands D and E (Fig. 3B and SI Appendix, Fig. S4A and B), as observed for the Sputnik penton protein (16), and mainly consists of a β-sandwich attached to helix IA. With ID β-strands labeled IA to IH, the β-sandwich is formed by a five- and a three-stranded antiparallel β-sheet, IB-IA-IE-IF and IH-ID-IC, respectively. Although low in sequence identity (19%), mavirus and Sputnik penton proteins (PDB ID code 3J26; ref. 16) have high structural similarity (DALI Z score: 22.1; highest Cα rmsd of protomers: 5.07 Å). Supersuperposing only the SJR of a mavirus and Sputnik penton protomer (highest Cα rmsd: 3.70 Å, Fig. 3C) revealed that the ID position relative to the SJR is the most distinctive feature, suggesting a certain degree of flexibility in the connecting loops. Notably, individual superposition of the mavirus and Sputnik IDs revealed a striking fold conservation (highest Cα rmsd: 2.39 Å, Fig. 3D). Such strong structural conservation combined with high sequence variability of the surface-exposed ID among virophages (12) supports the penton’s putative role in host cell or giant virus recognition.

Fig. 2. Crystal structures of mavirus flMCP and MCPΔC. (A) flMCP domains. JR1 and JR2 are orange and light orange, respectively. PC, pedestal connector. Secondary structure elements are labeled. (B) Scheme of flMCP JR strands and connecting loops. Starting and ending residues of JR strands are indicated. (C) The superposition of flMCP and MCPΔC structures was performed on trimeric casposmers but one protomer is illustrated here. (D) Core compaction and tower expansion after MCP processing. Side view of MCPΔC casposmer with the front protomer shown as solid surface and rear protomers transparently. DJR core compaction (dark green) and tower expansion (light green) are illustrated by arrows. Gray line, quasi-sixfold symmetry axis. Also see Movie S1.

Fig. 3. High structural similarity of mavirus and Sputnik capsid proteins. The Sputnik structure PDB ID code 3J26 used in A, C, and D was published in ref. 16. (A) Superposition of mavirus (green) and Sputnik (purple) MCPΔC. (B) Mavirus penton protein. SJR and ID are shown in light and dark yellow, respectively. Secondary structure elements are labeled. (C) and (D) Structural similarity between mavirus (colors as in B) and Sputnik (purple) penton protein shown by SJR superposition (C) and ID superposition (D).
could be determined for fIMCP/penton-VLPs of only 136 lysates, damaged appearance and small gain in protein stability (SD, ±Δ3 for all given t = n ± X and GX lysate n(VLP)/n(GP) fIMCP X motif or slight variations thereof. X and (M/I/L/N/Q)XGX 17 min) and 650-±= SI Appendix that fIMCP is the assembly relevant state for mavirus capsids. through capsid assembly compared with fIMCP-VLPs, we propose stability, suggesting that the gained free energy is used to induce of penton protein at the vertices only marginally increased VLP which is consistent with the low pressions did not show VLPs in negative stain EM (Fig. 4 pressures did not show VLPs in negative stain EM (Fig. 4 S Appendix, Fig. S6A). Considering the low MCPAC-VLP counts in E. coli lysates, damaged appearance and small gain in protein stability through capsid assembly compared with fIMCP-VLPs, we propose that fIMCP is the assembly relevant state for mavirus capsids. Penton Protein Confers icosahedral Shape on VLPs. To investigate whether mavirus MCP coassembles with penton proteins, we performed coexpression experiments in E. coli. We could observe and extract a VLP-containing light scattering band only for fIMCP/penton but not for MCPAC/penton coexpressions after sucrose gradient centrifugation (SI Appendix, Fig. S7A). Moreover, sucrose gradient fractions from MCPAC/penton coexpressions did not show VLPs in negative stain EM (Fig. 4A), which is consistent with the low E. coli lysate n(VLP)/n(GP) count ratio of 0.016 ± 0.012 per micrograph (SI Appendix, Fig. S5A). These results imply that only fIMCP but not MCPAC is able to stably assemble into VLPs in the presence of penton protein. Purified fIMCP/penton-VLPs indeed contained both capsid proteins (SI Appendix, Fig. S7B) and were icosahedral, resembling mavirus viorn cores rather than spherical MCP-VLPs (Fig. 4A). Only VLPs that included both fIMCP and penton protein appeared to be filled like native mavirus particles in negative stain EM images. Since the nucleic acid content, judged by the absorbance ratio A260nm/A230nm (SI Appendix, Table S4), was comparable for all VLPs and native mavirus particles, it is hypothesized that penton protein adds stiffness to fIMCP-containing VLPs. We compared the thermal stability of fIMCP-VLPs and fIMCP/penton-VLPs and found only slightly higher thermal stability for fIMCP/penton-VLPs with a maximum ΔTm of 1.3 ± 0.2 °C (SI Appendix, Fig. S6B). Thus, the incorporation of penton protein at the vertices only marginally increased VLP stability, suggesting that the gained free energy is used to induce strain or conformational changes.

MCP Processing Takes Place After Capsid Assembly. We next aimed at generating mavirus-like VLPs containing both processed MCPAC and penton protein by coexpressing fIMCP with penton protein and MVP in E. coli. We indeed extracted mavirus-sized, icosahedral VLPs (Fig. 4A). They mainly contained MCPAC but also residual fIMCP, penton protein, and MCP but no fragments of the fIMCP CTD (SI Appendix, Fig. S7B). We termed these particles fIMCP/penton/MVP-VLPs since they derive from fIMCP expressions. Interestingly, coexpression of MCPAC, penton protein, and MVP resulted in VLPs that resembled fIMCP/penton/ MVP-VLPs (Fig. 4A) and contained all three expressed proteins (MCPAC/penton/MVP-VLPs, SI Appendix, Fig. S7B). These particles showed a migration behavior comparable to fIMCP/ penton/MVP-VLPs in rate zonal sucrose gradient centrifugation (SI Appendix, Fig. S7A) and similar thermal stability (SI Appendix, Fig. S6C), suggesting that these VLPs are highly similar. To investigate whether capsid processing is regulated, we incubated fIMCP-containing capsomers and VLPs with MVP and analyzed processing by SDS/PAGE (SI Appendix, Fig. S8A). Normally, fIMCP-VLPs were processed with a half-life (t1/2) of 24 ± 3 min (average ± SD, n = 3 for all given t1/2), which was sixfold faster than for fIMCP capsomers (t1/2 = 136 ± 17 min) and 650-fold faster than the processing of fIMCP/penton-VLPs (t1/2 = 15,600 ± 9,300 min). These results suggest two points: First, MVP can enter fIMCP-VLPs to process MCP within the capsid. Second, MVP processes capsid-associated fIMCP faster than free fIMCP capsomers, providing a mechanism for timing capsid maturation after capsid assembly. The slow turnover of fIMCP/penton-VLPs implies that penton-containing VLPs present a closed, protease-resistant shell. MCP Processing Increases Capsid Stability at Acidic pH. To study how processing affects VLP stability, we measured thermal disassembly of VLPs and native mavirus particles through light scattering, fIMCP/penton-VLPs and processed fIMCP/penton/MVP-VLPs showed similar disassembly temperatures (Tm1/2 at pH 7.0, indicating that processing does not alter VLP stability at neutral pH) (Fig. 4B). However, VLP exposure to acidic environments revealed that capsids of processed fIMCP/penton/MVP-VLPs have a higher thermal stability than fIMCP/penton-VLPs. Interestingly, the stability of mavirus viorn cores increased with decreasing pH and showed lower thermal stability (except for pH 4.5) than processed IMCP/ penton/MVP-VLPs (Fig. 4B). Thus, despite the presence of genomic DNA and the complete set of viorn proteins, native mavirus particles were less stable than artificially generated VLPs. VLP and mavirus capsid stability data suggest that MCP processing results in viorn stabilization specifically in acidic environments, with potential implications for host cell infection. Discussion Virion Morphogenesis Is Conserved Among Virophages. We found that mavirus MCP is processed at four C-terminal sites by MVP. These sites are conserved in diverse virophages, which is particularly striking since the MCP CTD of the mavirus virion contains very low sequence identity are highly similar, including the arrangement of N and processed C termini (Fig. 3 and SI Appendix, Fig. S2). MVP recognized two motifs, GGXX and GXG, which are similar to cleavage motifs of the related adenoherpesviral protease, (M/L/L/N/Q)XGG and (M/L/L/N/Q)XG (G27). Most clan CE proteases, such as African swine fever virus (ASFV) p273R (28), recognize the GGXX motif or slight variations thereof. MVP recognition motifs could be more complex since the CTD of mavirus fIMCP contains two more GGX sites for which we could not identify cleavage products in vitro. We show that native mavirus viorn cores contain processed MCPAC at 1°C (Fig. 4C), and it was previously shown that Sputnik MCP is processed at the same site in mature viorn cores (16). The presence of MCPAC in two virophages and detection of MVP in mavirus viorn cores indicates that the cysteine protease has a conserved function among virophages. This hypothesis is further supported by the mavirus and Sputnik MCPAC structures, which despite low sequence identity are highly similar, including the arrangement of N and processed C termini (Fig. 3 and SI Appendix, Fig. S3G). We also solved the crystal structure of the mavirus penton protein and found a striking similarity to the Sputnik equivalent (16) not only regarding the SIR domain but notably also the ID (Fig. 3). Interestingly, the ID fold has so far only been detected in the penton of virophages and adenoviruses, suggesting an evolutionary and potentially functional connection.
between these nonenveloped DJR viruses. In contrast, the virophage MCP appears to be more closely related to enveloped viruses of the proposed order Megavirales than to the adenoviral hexon (SI Appendix, Table S3). Therefore, virophages could provide an evolutionary link between these groups of eukaryotic DJR viruses as proposed by Krupovic and Bamford (18).

Considering the strong homology of virophage capsid proteins, we suggest that the capsid architectures of mavirus and Sputnik are similar. Based on the high structural conservation of the MCP, its processing sites and protease activity, we propose that our findings on mavirus capsid assembly and maturation processing are relevant for virophages in general.

**Capsid Assembly Is Enabled by Strain Through the fIMCP C-Terminal Domain.** We found that mavirus MCP, both full-length and truncated, assembled into mavirus-sized VLPs in E. coli, implying that no specific virus or host protein is required to initiate assembly or determine size. This finding is in contrast to other DJR viruses, e.g., PRD1 (29) and adenovirus (30), for which the presence of specific proteins is required to form virions. Although mavirus VLPs readily formed in E. coli, in vitro capsid assembly (including crystallization batches) was not pronounced under our experimental conditions. We assume that high local protein concentrations and the presence of molecular chaperones to avoid aggregation, among other factors, assist capsid formation in E. coli.

Our crystal structures of fIMCP and MCPΔC show that MCP processing results in DJR core compaction and tower expansion (Fig. 2 and Movie S1). In addition, our VLP assembly (Fig. 4) and thermal stability data (SI Appendix, Fig. S6) suggest that fIMCP rather than MCPΔC is relevant for capsid assembly (Fig. 5). We can only speculate how structural and stability changes in capsomers affect mavirus capsid assembly. The seemingly flexible fIMCP CTD could transfer its high free energy to the DJR genome through the C-terminal scaffolding domain (Fig. S1A). This interaction could affect the ability to form capsids. Similarly, the tail dpDNA bacteriophage HK97 uses the C-terminal scaffolding domain by generating strain on capsomers, rendering them assembly competent (31). Thus, we propose that the CTD of fIMCP serves as scaffolding domain during mavirus capsid assembly. We cannot rule out that the fIMCP CTD directly supports capsid assembly, e.g., by contacting neighboring capsomers, as proposed for PRD1 (29).

The penton protein confersicosahedral shape on mavirus VLPs and renders them nearly inaccessible for external MVP (SI Appendix, Fig. S8). Thus, MVP must be packaged during assembly, which we indeed detected MVP in mavirus virions (Fig. 1C). Since all mavirus genes are preceded by the same late CroV promoter motif (9), mavirus MCP and penton protein likely coassemble in virion factories (Fig. 5), which we propose to be mediated through strain on capsomers induced by the fIMCP CTD.

**Virophage Processing Prepares Virions for Infection.** We hypothesize that virophage processing is a prerequisite to generate infectious virions, as described for, e.g., adenoviruses (27, 28) and poxviruses (33). We observed that fIMCP-VLPs are processed faster in E. coli, compared to mavirus MCPΔC (36) and Sputnik (16), which release pentons under stress conditions, e.g., acidification. Penton release is important during adenoviral infection (39) and could be relevant for virophages as well.

Beside penton destabilization, we propose that MCP processing prepares the mavirus genome for infection. Due to the strong positive charge of the fIMCP CTD [theoretical isoelectric point (pI): 10.15], especially in fragments MCPΔC, MCPΔC, and MCPΔC, it could cause an interaction between fIMCP and packed dpDNA, tightly attaching the genome to the capsid interior (Fig. 5). fIMCP processing would detach the genome–CTD complex from the capsid interior, preparing it for cell entry. Using PSIPRED v3.3 (40) and WHAT 2.0 (41), we identified a potential amphipathic α helix in fragment MCPΔC that could be involved in endosomal membrane rupture to allow entering the cytoplasm. We could not detect C-terminal fragments of fIMCP in native mavirus virions, either due to low concentration or intense fragmentation by trypsin, which was used for MS PMF. Thus, we cannot rule out that CTD fragments exit mavirus particles after processing.

Apart from MCPΔC, penton protein, and MVP, four further mavirus gene products were identified in mavirus virions: MV02, two cleavage fragments of MV13, MV14, and MV15 (Fig. 1C). The FtsK/HerA-type ATPase MV15 (9) could be associated with the C-terminal MCP to package the mavirus genome during virion morphogenesis as observed for Vaccinia virus (42). Apart from MV15, we presume that the detected virion proteins play a role during mavirus infection. The MV02 retroviral integrase likely interacts with the mavirus genome to allow host genome integration as observed recently (14). Besides C-terminal MCP fragments, MV13C and MV14 could also interact with the mavirus genome due to their high theoretical pIs (≥9.6). Interestingly, two nuclear localization signals (NLSs) were predicted for MV13 and one for the fIMCP CTD stretching across MCPΔC (43). One of these NLSs could enable nuclear import of the proposed mavirus genome/integrase complex during C. roenbergensis infection. The cleavage into MCPΔC, MCPΔC, and MCPΔC could act in a regulatory way on the proposed genome-interacting and NLS functions of MCPΔC. Further studies are needed to understand the processing and function of these domains for the maturation of the mavirus genome.
In conclusion, this study provides mechanistic insights into virion assembly, stability, and processing in virophages. Our findings have direct implications for virophage maturation and infection, enriching our understanding of these unique parasites of giant viruses. Parallels could be identified not only to other virophages, but also to more distantly related DNA viruses such as adenovirus and poxvirus, which share morphogenetic proteins with virophages. These homologies extend to the even larger Mimiviridae, the actual host site of virophages. Virophages thus appear to be the smallest and least complex member of the PRO1-adenovirus lineage described yet, which makes them particularly exciting for further characterization with implications for the entire viral lineage and potential biotechnological applications as nanocarriers.

Materials and Methods

Detailed materials and methods are described in SI Appendix.

Cloning. Mavirus genes (GenBank: HQ712116.1) were cloned into plasmids pETM-11 (provided by Gunter Stier) or pASK-IBA3C (IBA). For VLP preparation, capsid protein genes were cloned into pRSFDuet-1.

Protein Expression and Purification. Mavirus genes were expressed in E. coli BL21-CodonPlus(DE3)-RIL or BL21(DE3). Hexahistidine-tagged penton protein, flMCP, and MCPaL were mainly purified by Ni²⁺ affinity chromatography.

Size exclusion chromatography (SEC) was the final step for all proteins. MVP was purified by ion exchange chromatography and SEC. Mavirus was prepared as described previously (14). Mavirus and VLPs were purified by sucrose gradient centrifugation.

Crystal Structures. Experimental phases of MCPaL and penton protein from senecomyxomivirus single-wavelength anomalous dispersion data were used for molecular replacement (MR) to phase native datasets. The fIMCP dataset was phased by MR on the native MCPaL model. Protein Data Bank ID codes are 6G41–6G45 (SI Appendix, Table S2).

Negative Stain EM. Samples on glow discharged Formvar-carbon-coated 100 mesh Cu grids were stained by 0.5% uranyl acetate for 90 s and imaged by a Tecno T20 electron microscope.

ACKNOWLEDGMENTS. We thank the Heidelberg and Villigen X-ray crystallography teams for measurements at beamline X10SA (Swiss Light Source, Villigen, Switzerland); lime Schlichting and Miroslaw Tarasewicz for support with crystallization; Sabine Zimmermann and Susanne Eisel for technical support; Martine Haehsler for help with VLP preparations; Robert L. Shoeman and Sebastian Fabritz for help with mass spectrometry; and Dennis H. Bamford, Modesto Redrojo-Rodriguez, and David Stuart for information in the early stages of the project. We acknowledge financial support by the Heidelberg Biosciences International Graduate School (D.B.) and by the Max Planck Society.

1. Raoult D, Forterre P (2008) Redefining viruses: Lessons from mimivirus. Nat Rev Microbiol 6:315–319.
2. Claverie JM (2006) Viruses take center stage in cellular evolution. Genome Biol 7:110.
3. La Scala B, et al. (2008) The virophage as a unique parasite of the giant mimivirus. Nature 455:100–104.
4. Abrasho JS, et al. (2014) Anachthamabda polyphaga mimivirus and other giant viruses: An open field to outstanding discoveries. Viral J 11:120.
5. Krupovic M, Kuhn JH, Fischer MG (2015) A classification system for virophages and proviruses. Nat Rev Microbiol 13:105–115.
6. Desnues C, et al. (2012) Provirrophages and transpovirons as the diverse mobilome of cytoplasmic large DNA viruses.
7. Gaia M, et al. (2014) Zamilon, a novel virophage with Mimiviridae host specificity. Proc Natl Acad Sci USA 111:6375–6380.
8. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. J Mol Biol 372:774–797.
9. Perez-Bernal M, et al. (2015) A new virophage of Mimiviridae host specifity. PLoS One 9:e94923.
10. Shcherbakov S, et al. (2011) A virophage at the origin of large DNA transposons. Science 332:231–234.
11. Yau S, et al. (2011) Virophage control of antarctic algal host-virus dynamics. Proc Natl Acad Sci USA 108:6163–6168.
12. Zhou J, et al. (2013) Diversity of virophages in metagenomic data sets. J Virol 87: 4225–4236.
13. Roux S, et al. (2017) Ecogeoncomics of virophages and their giant virus hosts assessed through time series metagenomics. Nat Commun 8:858.
14. Blanc G, Gallot-Lavallée L, Maumus F (2015) Proviruses in the Bilogexiella genome bear testimony to past encounters with giant viruses. Proc Natl Acad Sci USA 112:8533–8536.
15. Fischer MG, Hackl T (2016) Host genome integration and giant virus-induced reactivation of the virophage mavirus. Nat Micro 2:88–291.
16. Krupovic M, Kuhn JH, Fischer MG (2015) Proviruses of the Bacterial genome that are driven by eukaryotic virus transposon and plasmid evolution. Nat Rev Microbiol 13:105–115.
17. Zhang X, et al. (2012) Structure of Sputnik: A virophage, at 3.5 Å resolution. Proc Natl Acad Sci USA 109:18431–18436.
18. Sun S, et al. (2010) Structural studies of the Sputnik virophage. J Virol 84:894–897.
19. Krupovic M, Bamford DH (2008) Virus evolution: How far does the double beta-barrel viral lineage extend? Nat Rev Microbiol 6:941–948.
20. Rossmann MG, Johnson JE (1989) Isohexahedral RNA virus structure. Annu Rev Biophys 18:533–573.
21. Yutin N, Raoult D, Koonin EV (2013) Virophages, polintons, and transpovirons: A complex evolutionary network of diverse selfish genetic elements with different reproduction strategies. Virol J 10:158.
22. Athappilly FK, Murali R, Rux Ji, Cai Z, Burnett RM (1994) The refined crystal structure of haem, the major coat protein of adenovirus type 2, at 2.9 Å resolution. J Mol Biol 242:430–455.
23. Ksinskel E, Hervick K (2007) Inference of macromolecular assemblies from crystalline state. J Mol Biol 372:774–797.
24. Holm L, Laakso LM (2016) Dali server update. Nucleic Acids Res 44:W351–W355.
25. Colson P, et al. (2013) “Megavirales”, a proposed new order for eukaryotic nucleocyttoplasmic large DNA viruses. Arch Virol 158:2517–2521.
26. Li Y, et al. (2005) Control of virus assembly: HK97 “Whiffleball” mutant capsids without pentons. J Mol Biol 348:167–182.
27. Mangel WF, San Martin C (2014) Structure, function and dynamics in adenovirus maturation. Viruses 6:4536–4570.
28. Krupovic M, Kuhn JH, Fischer MG (2015) Proviruses of the Bacterial genome that are driven by eukaryotic virus transposon and plasmid evolution. Nat Rev Microbiol 13:105–115.