This is a repository copy of Cryo-EM structure and in vitro DNA packaging of a thermophilic virus with supersized T=7 capsids.

White Rose Research Online URL for this paper: http://eprints.whiterose.ac.uk/143214/

Version: Published Version

Article:
Bayfield, Oliver W, Klimuk, Evgeny, Winkler, Dennis C et al. (9 more authors) (2019) Cryo-EM structure and in vitro DNA packaging of a thermophilic virus with supersized T=7 capsids. Proceedings of the National Academy of Sciences of the United States of America. pp. 3556-3561. ISSN 1091-6490

https://doi.org/10.1073/pnas.1813204116

Reuse
This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here: https://creativecommons.org/licenses/

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.
Cryo-EM structure and in vitro DNA packaging of a thermophilic virus with supersized T=7 capsids

Oliver W. Bayfield, Evgeny Klimuk, Dennis C. Winkler, Emma L. Hesketh, Maria Chechkina, Naiqian Cheng, Eric C. Dykeman, Leonid Minakhin, Neil A. Ranson, Konstantin Severinov, Alasdair C. Steven, and Alfred A. Antson

*York Structural Biology Laboratory, Department of Chemistry, University of York, York YO10 5DD, United Kingdom; Laboratory of Structural Biology Research, National Institute of Arthritis Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, MD 20892; Center for Life Sciences, Skolkovo Institute of Science and Technology, 143025 Skolkovo, Russia; Institute of Molecular Genetics, Russian Academy of Sciences, 123182 Moscow, Russia; Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds LS2 9JT, United Kingdom; Department of Mathematics, University of York, York YO10 5DD, United Kingdom; and Waksman Institute for Microbiology, Rutgers, The State University of New Jersey, Piscataway, NJ 08854

Edited by Michael G. Rossmann, Purdue University, West Lafayette, IN, and approved January 10, 2019 (received for review August 14, 2018)

Double-stranded DNA viruses, including bacteriophages and herpesviruses, package their genomes into preformed capsids, using ATP-driven motors. Seeking to advance structural and mechanistic understanding, we established an in vitro packaging assay for a thermophilic bacteriophage P23-45 of Thermus thermophilus. Both the unexpanded procapsid and the expanded mature capsid can package DNA in the presence of packaging ATPase over the 20°C to 70°C temperature range, with optimum activity at 50°C to 65°C. Cryo-EM reconstructions for the mature and immature capsids at 3.7-Å and 4.4-Å resolution, respectively, reveal conformational changes during capsid expansion. Caspomer interactions in the expanded capsid are reinforced by formation of intersubunit β-sheets with N-terminal segments of auxiliary protein trimers. Unexpectedly, the capsid has T=7 quasi-symmetry, despite the P23-45 genome being twice as large as those of known T=7 phages, in which the DNA is compacted to near-crystalline density. Our data explain this anomaly, showing how the canonical HK97 fold has adapted to double the volume of the capsid, while maintaining its structural integrity. Reconstructions of the procapsid and the expanded capsid defined the structure of the single vertex containing the portal protein. Together with a 1.95-Å resolution crystal structure of the portal protein and DNA packaging assays, these reconstructions indicate that capsid expansion affects the conformation of the portal protein, while still allowing DNA to be packaged. These observations suggest a mechanism by which structural events inside the capsid can be communicated to the outside.

Key step in the assembly pathway of double-stranded DNA viruses, including bacteriophage and evolutionarily related herpesviruses, is the packaging of viral genomic DNA into the procapsid (1). During this process, the large terminase protein docks onto the portal protein, which is embedded in the capsid wall, and translocates DNA using free energy liberated from ATP hydrolysis. This motor can work against high internal pressure, generating forces reaching 100 pN and translocating DNA at rates reaching 100 bps to 2,000 bps (2, 3). As a result, DNA is packaged inside the virion to near-crystalline density (>500 mg/mL) (4, 5). This makes the motor a suitable tool for biotechnological applications such as gene delivery or sequencing, and a potential target for drug discovery in the case of human and animal viruses (6). The unique features of this motor also make it a useful system for studying fundamental biological processes involving DNA translocation and the coupling between ATP hydrolysis and mechanical work.

A number of mesophilic bacterial viruses have been employed to study DNA packaging in vitro (7–13). Initial studies on bacteriophage λ demonstrated that procapsids were precursors to the expanded head (14) and that purified empty capsids can be packaged with double-stranded DNA in the presence of the ATPase that drives genome packaging—the large terminase—and ATP (15). The in vitro packaging reaction enables this key stage in virus assembly to be studied under controlled conditions, mimicking events inside the host cell. ATP consumption as well as packaging rates and forces have been measured using in vitro systems established for bacteriophages λ, T4, and λ (2, 16, 17). In parallel, cryo-EM 3D reconstructions have been reported for capsid–motor complexes of bacteriophages λ29 and T4, to resolutions of 12 Å and 34 Å, respectively (18, 19). Combined with crystal structures of individual motor components, pseudotatomic models have begun to show how the motor engages with DNA and how translocation is achieved. However, the precise molecular details for the packaging mechanism have yet to be determined. This research would benefit from the availability of a robust packaging system with greater stability.

To advance structural and mechanistic studies, we have established an in vitro DNA packaging system for the thermophilic virus P23-45. P23-45 and P74-26 belong to the genus P23virus, and are closely related to bacteriophage G20k, all of which infect Thermus thermophilus (20). Individual components of the DNA packaging motor of these viruses have been

Author contributions: O.W.B., A.C.S., and A.A.A. designed research; O.W.B., E.K., D.C.W., E.L.H., M.C., N.C., N.A.R., and A.A.A. performed research; E.K., L.M., and K.S. contributed new reagents/analytic tools; O.W.B., E.C.D., A.C.S., and A.A.A. analyzed data; and O.W.B., E.L.H., E.C.D., N.A.R., K.S., C.S., and A.A.A. 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 License 4.0 (CC BY).

Data deposition: Electron Microscopy Data Bank, www.emdbank.org (accession nos. EMD-4443, EMD-4447, EMD-4445, and EMD-4446) and the Protein Data Bank, www.rcsb.org (PDB ID codes 6I9E, 6IB8, and 6I8G).

1To whom correspondence should be addressed. Email: fred.antson@york.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1813204116/-/DCSupplemental.

Published online February 8, 2019.
characterized biochemically and structurally (21–24). We isolated procapsids and expanded capsids of P23-45 and demonstrated DNA packaging in vitro in the presence of cognate large terminase gp85. Despite its large genome, which is twice as big as that of HK97, P23-45 utilizes a similar capsid protein fold, and forms capsids with the same T=7 quasi-symmetry as HK97 and similar phages. Cryo-EM reconstructions explain the structural basis for this anomalous size, showing how the larger capsid lattice is accomplished by modifying the canonical HK97 fold, and how the conformation of the capsid protein changes during capsid expansion. One of the differences between the procapsid and expanded capsid is the presence of trimers of an auxiliary protein on the outer surface of the expanded capsid, a property held in common with λ and TW1 (25–27). Furthermore, reconstructions of the capsid which resolve the unique portal vertex, combined with a 1.95-Å resolution crystal structure of the portal protein, allowed structural characterization of the unique portal-capsid interface, where the respective symmetries of interacting proteins do not match. These data show how the portal protein structure is affected by capsid expansion: Combined with normal mode analysis calculations, they suggest the existence of cross-talk between parts of the portal protein that are respectively inside and outside the capsid.

**Results**

**Cryo-EM Icosahedral Capsid Reconstructions.** Spherical capsids with thick serrated walls (Fig. 1A) were present alongside large, faceted capsids with thinner walls (Fig. 1B) in lysates of P23-45-infected *T. thermophilus* cells. The smaller capsids, hereafter referred to as procapsids, were in higher abundance than the mature expanded empty capsid. Empty and DNA-filled expanded capsids appeared to be of the same size (Fig. 1B), with the diameter of the circumsphere at ~82 nm, compared with ~66 nm for the procapsid. Single-particle reconstructions imposing icosahedral symmetry (Movies S1 and S2) were calculated at 4.4-Å resolution for the procapsid and 3.7-Å resolution for the expanded capsids all exhibited T = 7 quasi-symmetry as HK97, irrespective of the T number (30, 31). In the P23-45 auxiliary protein, however, capsomers are repeated at a ~25% larger interval of ~17 nm (Fig. 24).

Analysis of the structure and comparison with the canonical HK97 fold shows that two key factors contribute to the increased capsid lattice spacing. First, the A domains of the P23-45 capsid protein are extended. Compared with HK97, the two conserved helices A1 (213 to 224) and A2 (242 to 247) (Fig. 2B), corresponding to residues 254 to 268 and 279 to 286 in HK97, respectively, are shifted away by ~10 Å from the P domain in P23-45 owing to extensions to loops 202 to 212 and 318 to 321. In addition, the A domain of P23-45 has two further α-helices, A15 (295–298) and A16 (272 to 287), which protrude into the middle of the capsomer apex. Second, the P23-45 capsid protein has a longer E loop than other HK97-like viruses, extending ~12 Å farther. This complements the increased capsomer diameter, allowing the E loop to reach far enough to make interactions with neighboring subunits. The P23-45 capsomer–capsomer interactions are multilayered (Figs. 3 and 4). At the lowest, internal layer of the capsid, the T loop, N arm, and P domain interact at the threefold and twofold icosahedral and local symmetry axes (Fig. 3A). The T loop consists of ~14 amino acids and extends from the local threefold symmetry axis, the N arm and P domain interact at the threefold and twofold icosahedral axes, and interdigitates with the P domain of an apposing subunit. The extended N arm reaches toward an adjacent capsomer, forming a “crook” in the expanded capsid, and contributing one strand to a four-stranded β-sheet (Figs. 2B and 4). This crook also interacts with the S loop and an apposed crook at the icosahedral and local twofold axes positions. Moving upward through the capsid wall, two further antiparallel strands of the β-sheet are contributed by the E loop that overlays the P domain (Fig. 3B). The outermost strand of the sheet is contributed by the N-terminal region of the auxiliary protein (Figs. 5C and 4). This interaction is observed only in the expanded capsid, as the auxiliary proteins are only present in the expanded capsid, at the local and icosahedral twofold axes (Figs. 1D and 3D). Auxiliary proteins of this type were first characterized for bacteriophage λ (25, 26). In the crystal structure of the auxiliary protein from phage P74-26 (32), a close homolog of P23-45, the N-terminal segment, residues 1 to 16, is disordered (Fig. 3E), as in the crystal structure of phage λ gpD (25). In contrast, in the context of the expanded P23-45 capsid, the N-terminal segment of the auxiliary protein adopts a well-defined conformation, forming main-chain hydrogen bonds with the E loop of the capsid protein. In the procapsid, the E loop has a relaxed conformation, with its middle section being partially disordered (Fig. 4 and Movies S3 and S4) but with its end locked in a G-loop/E-loop “trap” (Fig. 4), whereby the tip of the E loop is pinned beneath the G loop of a neighboring subunit. Notably,
such a G-loop/E-loop interaction is observed in both the procapsid and the expanded capsid states. The G loops complete the outermost layer of the capsid wall and buttress the auxiliary trimers (Fig. 3C). Only in the expanded state do the E loops become stretched, adopting a well-defined β-strand conformation and creating the binding site for the N-terminal strand of the auxiliary protein (Figs. 3 and 4). Other significant differences between the expanded and procapsid states are in the N-terminal and C segments of the capsid protein, residues 1 to 23 and 398 to 409, which engage in subunit interactions in the expanded state but are disordered in the procapsid. Notably, the unusual PPFPP motif at the C terminus, residues 401 to 405 (Fig. 2B, top left), reinforces subunit interactions via a phenylalanine–tyrosine “lock” between Y226 and F403 of the neighboring C arm (Fig. 2B, top right).

**Capsid Volume.** A feature of many T=7 HK97-fold bacteriophages is a skewed dimer-of-trimers arrangement of subunits within procapsid hexons, which adopt a near-sx-fold symmetrical arrangement upon capsid expansion. There was no such skew observed in the hexons of P23-45 procapsids, although the ratio of its internal volume versus the expanded capsid, of 0.49, is consistent with observations for other T=7 bacteriophages (Table 1). Strikingly, the P23-45 genome is twice as large as in other T=7 phage. Bacteriophage SIO-2, which has the canonical HK97 capsid fold, but a genome size similar to that of P23-45, instead utilizes a larger (T=12) capsid (33).

**Crystal Structure of the Portal Protein.** A crystal structure was determined for the portal protein from the closely related bacteriophage G20c, which shares 99.3% sequence identity with the P23-45 portal. As the previously determined structure of this protein, we determined the structure for the portal protein, residues 1 to 409, which adopt an ordered conformation. This structure was refined at 1.95-Å resolution (Fig. 5A and SI Appendix, Table S1). With three subunits in the asymmetric unit, the 12-subunit oligomer is generated by the crystallographic fourfold axis. The portal protein forms a canonical 12-subunit oligomer with an overall shape consistent with those of portal proteins of other viruses (34). We refer to its domains as Clip, Stem, Wing, and Auxiliary protein with an intact N terminus. One goal was to test whether this protein can change its conformation, and that there is a potential cross-talk between the Wing and Clip; these domains of the portal protein are respectively inside and outside the capsid.

**In Vitro DNA Packaging.** An in vitro DNA packaging assay was established for empty P23-45 capsids. Isolated empty capsids

---

**Table 1. Internal volumes of bacteriophage capsids**

| T-number | \(V_p\) \(10^4\) nm\(^3\) | \(V_e\) \(10^4\) nm\(^3\) | Genome size, kb |
|----------|-----------------|-----------------|----------------|
| HK97     | 7               | 3.88 (0.47)     | 8.25           | 39.7           |
| T7       | 7               | 4.79 (0.55)     | 8.64           | 39.9           |
| P22      | 7               | 5.25 (0.58)     | 9.09           | 41.7           |
| SIO-2    | 12              | 9.26 (0.59)     | 15.7           | 81.2           |
| P23-45   | 7               | 8.90 (0.49)     | 18.2           | 84.2           |

Volume calculations for procapsids (\(V_p\)) and expanded capsids (\(V_e\)) were performed for the following structures: HK97 = EMD 5828, EMD 2112, PDB 1OHG; T7 = EMD 6034, EMD 6037; P22 = EMD 1824, EMD 1826; \(\lambda\) = EMD 1507, EMD 5012; and SIO-2 = EMD 5383, EMD 5382.

\(V_p\) as a fraction of \(V_e\) is given in parentheses.
were tested for their ability to protect plasmid DNA (pUC18) from DNase I digestion in the presence of components required for DNA packaging (Fig. 6). Packaging activity was minimal at 20 °C but increased with rising temperature until ∼50 °C. Optimal activity was observed in the temperature range of 50 °C to 65 °C. At higher temperatures (70 °C to 75 °C), packaging activity decreased (Fig. 6A and SI Appendix, Fig. S6). No packaged DNA was detected after 15 min of packaging at 50 °C. The intensity of the DNA band after 30 min of packaging was similar to the intensity of bands after 45 and 60 min of packaging. To test the selectivity of the packaging motor, the substrate contained a mixture of linearized, circular supercoiled, and circular relaxed plasmid DNA. Only the linear double-stranded DNA was protected from DNase I digestion in the presence of components required for DNA packaging (Fig. 6). Packaging activity was minimal at 20 °C but increased with rising temperature until ∼50 °C. Optimal activity was observed in the temperature range of 50 °C to 65 °C. At higher temperatures (70 °C to 75 °C), packaging activity decreased (Fig. 6A and SI Appendix, Fig. S6). No packaged DNA was detected after 15 min of packaging at 50 °C. The intensity of the DNA band after 30 min of packaging was similar to the intensity of bands after 45 and 60 min of packaging. To test the selectivity of the packaging motor, the substrate contained a mixture of linearized, circular supercoiled, and circular relaxed plasmid DNA. Only the linear double-stranded DNA was protected from DNase I digestion (Fig. 6A), and was therefore packaged. Further experiments compared packaging activity of the procapsid and the expanded capsid. Homogeneity of procapsid samples was monitored by negative staining EM (SI Appendix, Fig. S3), and, with mass spectrometry, confirming the absence of auxiliary protein in procapsids and its presence in expanded capsids (SI Appendix, Fig. S4). The packaging activity of procapsids was low at 50 °C, with minimal activity observed at 20 °C (Fig. 6C). Expanded capsids exhibited high packaging activity at 50 °C, but had no detectable activity at 20 °C. Control experiments carried out at 50 °C in which the large terminase was excluded, or where ATP was substituted with ATPγS, detected no packaging activity. Assessment of procapsids after 30 min of packaging at 50 °C by negative stain EM did not indicate capsid disassembly or expansion (SI Appendix, Fig. S5).

**Discussion**

**How Can Capsid Size Be Increased Without Changing the T Number?**

Virus evolution generally involves the acquisition of genes coding for additional functions, which results in an enlargement of the genome. If the packaging capacity limit of the capsid has been reached, further space for accommodating a larger genome can be achieved in several ways. One way is to adopt a larger T number (that is, use a larger number of capsid protein monomers), with only minor adjustment to their fold. This has occurred throughout evolution, and is demonstrated by the different T numbers of HK97-fold capsids and capsid-like compartments, from the smallest T=1 and T=3 encapsulins (36, 37), to the T=16 herpesviruses (38) and T=52 jumbo bacteriophages (39). Another strategy is to pack the DNA more densely, and, in the herpesviruses (T=16), this appears to have occurred with cytomegalovirus, which has a ∼225-kb genome, compared with the ∼155-kb genome of the herpes simplex virus (40). A third option is capsid expansion, whereby a relatively small procapsid is converted to a larger mature capsid, thereby increasing the storage volume for DNA. These options appear to have an upper limit of DNA packaging densities, which is considered to be near-crystalline in most tailed bacteriophages (4, 5); thus, for many T=7
capsids, there is little potential for packaging additional DNA. A fourth possibility, previously unobserved to our knowledge, is to retain the T number but to increase the capsid volume by increasing the size of the capsomers. This is the strategy employed to package DNA in vitro is consistent with observations for other bacteriophage systems (17, 42). The observation that both procapsid and expanded capsid can package DNA when large terminase is added, and that the in vitro reaction is most efficient at 50 °C to 65 °C. Lanes from left to right: size marker, input DNA (no DNase), negative control with no added ATP, and packaging reactions with no added large terminase (control), with added large terminase at 50 °C and 20 °C, and with ATP substituted by ATPγS (γ).

Can DNA Be Packaged into both Capsid Expansion States? In addition to its enlarged capsid lattice spacing, the fold of the P23-45 capsid protein must be capable of maintaining capsid integrity at high temperatures (~70 °C) and in the presence of high internal pressures exerted by packaged DNA. DNA protection assays show that the procapsid and expanded capsids are both capable of initiating packaging with DNA when large terminase is added, and that the in vitro reaction is most efficient at 50 °C to 65 °C. The observation that both procapsid and expanded capsid can package DNA in vitro is consistent with observations for other systems (17, 42–44).

P23-45 expanded capsids exhibit enhanced DNase protection compared with the procapsids. In common, early observations on bacteriophage λ by thin section EM identified expanded capsids with a “grizzled” appearance (45), indicating partial packaging of DNA. Later studies, including those on bacteriophage T4, demonstrated that packaging is not necessarily coupled to expansion (43), and that expanded capsids even display enhanced packaging activity over unexpanded capsids (43, 44). The ability of expanded capsids to package DNA has been further confirmed by single-molecule experiments (17, 42) and structural data (18). Our observations on P23-45 indicate that both capsid expansion states are packaging-competent, but the expanded capsid may be better able to protect DNA either due to greater stability of the capsid or due to enhanced packaging activity which would likely originate from the conformation of the portal.

How Does the Capsid Influence Portal Protein Conformation? The portal protein is crucial for the assembly of infectious virions: It assists in scaffold-mediated assembly of the procapsid (46) and serves as the docking site for the DNA packaging motor (18), as a gatekeeper for DNA exit (47), and as the tail attachment site (48). The portal protein’s multifunctionality probably reflects its ability to adopt different conformations. It has been suggested that the P22 portal protein adopts a pseudo-fivefold symmetry to facilitate interaction with the large terminase (49), and a similar mismatch with the capsid vertex must also be satisfied (50). It is therefore not unexpected that our reconstructions of the P23-45 portal vertex indicate differences between the portal protein conformation in the procapsid and expanded capsid (Fig. 5 B–E), which may result in their differential packaging activities. Likewise, normal mode analysis shows that the lowest-frequency mode with 12-fold symmetry corresponds to a synchronized movement of the Wing and Clip, so that the portal protein oscillates between stretched and compressed conformations (Movie S5). Our analysis of the P23-45 genome termini in sequencing data indicated that this class of viruses encodes a pac site in its genomes and uses a headful packaging mechanism (SI Appendix, Fig. S7). The portal protein from another pac phage, T4 (35), exhibits a similar low-frequency mode (Movie S6), further reinforcing the notion that the portal protein can adopt different conformations, which depend on the stage of virus assembly. Two conclusions can be drawn from these observations. First, the capsid alone can affect the portal protein’s conformation, while the ability of both the procapsid and expanded capsid to package DNA may be influenced by the size of the capsid vertex.
DNA indicates that conformations of the portal protein required for packaging are accessible in both expansion states of the capsid. Second, synchronized movements of the Wing and Clip may serve as a cross-talk mechanism between the two portal domains, suggesting a functional pathway by which structural events inside the capsid can be communicated to factors outside. In this way, the capsid could (i) present the portal Clip in a compatible conformation to interact with external factors, such as the large terminase, and (ii) convey a signal at the termination of packaging.

Future work can exploit the enhanced stability of the thermophase system for eliciting high-resolution structural data for different stages of the DNA packaging motor. Single-molecule experiments would complement structural studies, potentially benefiting from reduced packaging rate at ambient temperatures.

Materials and Methods

Bacteriophage P22-45 particles, and the portal and large terminase proteins, were produced and purified (SI Appendix, SI Materials and Methods). Cryo-EM structures of capsids as well as the X-ray structure of the portal protein were determined as described in SI Appendix, SI Materials and Methods. Details of the normal mode analysis, in vitro DNA packaging assays, the genome termini analysis, mass spectrometry of capsids, and denatometric analysis of agarose gels are provided in SI Appendix, SI Materials and Methods.

Acknowledgments

We thank Brian Kelch (University of Massachusetts) and Lindsay Black (University of Maryland) for helpful discussions, and Daniel Hurdis and Rebecca Thompson (University of Leeds) for suggestions. Image processing and assistance in data collection, Tim Kirk, Meg Stark, and Adam Dowle (University of York) are acknowledged for assistance with computational resources, EM, and mass spectrometry, respectively. This work was supported by Wellcome Trust Grant 206377 (to A.A.A.), by Wellcome Trust-National Institutes of Health Studentship 103460 (to O.W.B.), by the Intramural Research Program of National Institute of Arthritis and Musculoskeletal and Skin Diseases (A.C.S.), by Russian Foundation for Basic Research Grant 16-34-01371 (to E.K.), by Biotechnology and Biological Sciences Research Council Grant G0212501/1 (to E.L.H.), and by National Institutes of Health Grant R01GM10407 (to K.S.). The Titan Krios microscopes were funded by a University of Leeds Arbury BioStructure Laboratory award and Wellcome Trust Grant 108466 (to N.A.R.).