UNIVERSITY OF CALIFORNIA SAN DIEGO

Defining the anatomy of the jumbo phage nucleus

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Biology

by

Eray Enustun

Committee in charge:

Professor Joe Pogliano, Chair
Professor Kevin D. Corbett, Co-Chair
Professor Arshad Desai
Professor Susan Golden
Professor Elizabeth Villa

2023

The Dissertation of Eray Enustun is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2023

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First, I would like to acknowledge and thank my advisors Professor Joe Pogliano and Professor Kevin Corbett for their mentorship and their support. Ever since he met me, Dr. Pogliano supported my ideas, believed in my scientific talent and creativity. I am grateful for Dr. Corbett who accepted to be my co-mentor and always found time to discuss my ideas, valued my perspective, and guided me to become a better scientist. As much of my PhD career and my thesis content, I thank them for shaping me to be the scientist I am today.

Similarly, I thank my committee members Dr. Elizabeth Villa, Dr. Susan Golden and Dr. Arshad Desai. They always made me feel supported, shared their invaluable advice, and allowed me to shape my science and my time at UCSD in the best way possible. While she was not officially a part of my thesis committee, I would like to also thank Dr. Kit Pogliano for listening my results and helping me with her expert opinion.

I would like to acknowledge all the current and past members of Pogliano Labs that I have been working with in the last five years. I would like to sincerely thank Eammon Riley and Emily Armburster for their support, discussing ideas and having fun conversations even when we were in the lab late at night or on weekends. I would also like to thank Hannah Tsunemoto and Roland Liu for helping me adjust the lab and the PhD program.

Likewise, I would like to thank every past and current Corbett Lab members I had the pleasure of working with. After the start of my co-mentorship, everybody in the lab was so welcoming, helping and you allowed me to feel as a part of the team. I sincerely

thank Amar Deep and Yajie Gu for teaching me important techniques, helping me in my projects and always finding time to listen my scientific ideas.

Outside the lab, I would like to thank my prior roommates Robert Gallant and Deniz Cakan. In addition to their interests in my research and fun scientific discussions with each of them, I appreciate them both to allow me not to feel lonely especially in the first years after moving from Turkey. I appreciate Olcay Soyalan for our friendship and all her help to before I start at UCSD and my first year in San Diego. I also thank my UCSD BioSci PhD program, my cohort and my BGSE group. I also thank my friends from Turkey, especially my best friend Dila Kakirman for the invaluable support.

Finally, I would like to dedicate this work to my family: Engin and Tulay Enustun and my family in San Diego: Miranda J. Barraza, Luna (the dog) and Luna (the cat). I have been blessed with you in my life and I would not be able to do these work and these accomplishments without your support. Thank you for everything you have done for me.

Chapter 2, in full, has been submitted for publication of the material as it may appear in *NSMB,* 2023. Enustun, E., Deep, A., Gu, Y., Nguyen, K.T., Chaikeeratisak, V., Armbruster, E., Ghassemian, M., Villa, E., Pogliano, J., Corbett, K.D. (2023). Identification of the bacteriophage nucleus protein interaction network. *NSMB*. The dissertation author was the primary investigator and author of this paper.

Chapter 3, in full, is currently being prepared for submission for publication of the material. Enustun, E., Armbruster, E., Lee, J., Zhang, S., Lee, B., Gu, Y., Deep, A., Naritomi, J., Liang, Q., Aigner, S., Chaikeeratisak, V., Cleveland, D., Ghassemian, M., Yeo, G., Pogliano, J., Corbett, K.D. A phage nucleus-associated RNA-binding protein

required for jumbo phage infection. The dissertation author was the primary investigator and author of this material.

VITA

2016	Bachelor of Science in Molecular Biology and Genetics, Koç University, Istanbul, Turkey
2016	Bachelor of Science in Chemical and Biological Engineering, Koç University, Istanbul, Turkey
2018	Master of Science in Molecular Biology and Genetics, Koç University, Istanbul, Turkey
2023	Doctor of Philosophy, Biology, University of California San Diego

PUBLICATIONS

Enustun E, Armbruster E, Lee J, Gu Y, Deep A, Zhang S, Naritomi J, Lee B, Liang Q, Chaikeeratisak V, Ghassemian M, Yeo G, Pogliano J, Corbett K. RNA binding protein promotes mRNA translocation through jumbo phage nuclear shell. (In Preparation)

Tsunemoto H, Sugie J, **Enustun E**, Pogliano K, Pogliano J. Using bacterial cytological profiling to probe the interplay between jumbo phage ϕKZ infection and treatment with cell wall active antibiotics in Pseudomonas aeruginosa. (Accepted by **PLOS One**)

Enustun E, Deep A, Gu Y, Nguyen K, Chaikeeratisak V, Armbruster E, Ghassemian M, Villa E, Pogliano J, Corbett K. Identification of the bacteriophage nucleus protein interaction network. (Accepted by **Nat Struct Mol Biol**)

Lau R, **Enustun E**, Gu Yajie, Nguyen JV, Corbett KD. A conserved signaling pathway activates bacterial CBASS immune signaling in response to DNA damage, **EMBO J**, 26 September 2022.

Laughlin TG, Deep A, Prichard AM, Seitz C, Gu Y, **Enustun E,** Suslov S, Khanna K, Birkholz E, Armbruster E, McCammon JA, Amaro RE, Pogliano J, Corbett KD, Villa E. Architecture and self-assembly of the jumbo bacteriophage nuclear shell, **Nature**, 03 August 2022.

Enustun E, Payzın-Dogru D, Kocak E, Ozel E, Akbari S, Sahin G. N., Karahuseyinoglu S, Erdal E, Onder TT. DOT1L is a barrier to fibroblast- to-hepatocyte direct lineage conversion. (Submitted to **Frontiers in Cell and Developmental Biology**)

Chaikeeratisak V, Khanna K, Nguyen KT, Egan ME, **Enustun E**, Armbruster E, Lee J, Pogliano K, Villa E, Pogliano J. Subcellular Organization of Viral Particles During Maturation of Nucleus-Forming Jumbo Phage, **Science Advances**, 04 May 2022.

Montaño ET, ..., **Enustun E**, Prichard A, Gomez A, Khanna K, Trigg S, Pogliano K, Pogliano J. Isolation and characterization of Streptomyces bacteriophages and Streptomyces strains encoding biosynthetic arsenals: Streptomyces strains and phages for antibiotic discovery, **PLOS One**, 21 January 2022.

Montaño ET, Nideffer JF, Sugie J, **Enustun E**, Shapiro AB, Tsunemoto H, Derman A, Pogliano K, Pogliano J. Bacterial Cytological Profiling Identifies Rhodanine-containing PAINS Analogs as Specific Inhibitors of E. coli Thymidylate Kinase in Vivo, **Journal of Bacteriology**, 08 September 2021.

Ugurlu-Cimen D, Odluyurt D, Sevinc K, Ozkan-Kucuk NE, Ozcimen B, Demirtas D, **Enustun E**, Aztekin C, Philpott M, Oppermann U, Ozlu N, Onder TT. AF10 (MLLT10) prevents somatic cell reprogramming through regulation of DOT1L-mediated H3K79 methylation, **Epigenetics & Chromatin**, 02 July 2021.

Akbari S, Sevinc GG, Ersoy N, Basak O, Kaplan K, Sevinc K, Ozel E, Sengun B, **Enustun E**, Ozcimen B, Bagriyanik A, Arslan N, Onder TT, Erdal E. Robust, Long-Term Culture of Endoderm-Derived Hepatic Organoids for Disease Modeling, **Stem Cell Reports**, 08 October 2019.

Stogsdill JA, Ramirez J, Liu D, Kim YH, Baldwin KT, **Enustun E**, Ejikeme T, Ji RR, Eroglu C. Astrocytic neuroligins control astrocyte morphogenesis and synaptogenesis, **Nature**, 09 November 2017. (Journal Cover)

Singh SK, Stogsdill JA, Pulimood NS, Dingsdale H, Kim YH, Pilaz LJ, Kim IH, Manhaes AC, Rodrigues WS Jr., Pamukcu A, **Enustun E**, Ertuz Z, Scheiffele P, Soderling SH, Silver DL, Ji RR, Medina AE, Eroglu C. Astrocytes Assemble Thalamocortical Synapses by Bridging Neurexin-1α and Neuroligin-1 via Hevin, ***Cell***, 14 January 2016.

ABSTRACT OF THE DISSERTATION

Defining the anatomy of the jumbo phage nucleus

by

Eray Enustun

Doctor of Philosophy in Biology

University of California San Diego, 2023

Professor Joe Pogliano, Chair
Professor Kevin Corbett, Co-Chair

Summary

Recently, a family of bacteriophages has been found to form a nucleus-like replication compartment, called the phage nucleus, which encapsulates the phage DNA and protects it from bacterial host defense systems. Although we have discovered the general replication involves assembly of the phage nucleus, it is still poorly understood at the molecular level, especially the macromolecule translocations and the key proteins that play roles in these functions. The aim of this thesis is to identify the phage nucleus and its associated proteins in greater detail to shed light specifically on how DNA and mRNA export occurs, as well as the selective transport of proteins into this structure.

In Chapter 2, we identify a new set of phage nucleus-associated proteins through comprehensive proteomics and biochemistry. One of the identified proteins, now termed ChmB, is further investigated and found to be directly interacting with ChmA, which forms most of the phage nucleus. In addition, it is found to be directly interacting with the portal protein, suggesting that ChmB might be forming pore-like structures to accommodate DNA packaging. This study provides new insights into the composition and functions of the phage nucleus, particularly in protein-protein interactions.

In Chapter 3, we focus on another phage nucleus-associated protein, ChmC. We found that ChmC has structural homology to known RNA binding proteins. We confirmed that ChmC binds to mRNA through its surface-exposed positively charged residues and that it also has phase separation properties. Investigating the samples collected during the infection confirmed the phage mRNA binding, particularly in 5' regions of the transcripts. When ChmC was knocked down via ddCas13d system, microscopy images

showed that the phage nucleus could form, but the replication was arrested at an early stage. Correspondingly, we found that the knockdowns cause a significant reduction in phage bouquet formation as well as phage titers. These results show that ChmC has a critical role in the phage nucleus system, possibly being a part of the mRNA export mechanism after the switch to non-virion RNA Polymerase for transcription.

Chapter 4 describes the other key proteins for the phage nucleus system, such as the portal protein and its potential docking site of octameric assembly of ChmB. We also investigate the other phage nucleus-associated proteins we found by proximity labeling, gp63 and gp64 of phiPA3. Our observations suggest gp63 has a critical role in selective protein transport to the phage nucleus and that gp64 interacts with it to potentially form a complex. Finally, we identify phage proteins that potentially inhibit host cell division, which is required as the jumbo phages require longer time to complete their replication compared to their hosts. These encoded proteins might provide a more comprehensive understanding of the replication mechanism as it is crucial for phages' ability to propagate in bacterial hosts, while potentially providing underlying mechanisms that could be used in therapeutics and the biotech industry.

In conclusion, this thesis provides important insights into the phage nucleus system and the molecular mechanisms behind it. We believe that identifying and characterizing the phage nucleus and its associated proteins can have important implications for developing host-immune system evading phage therapies and allowing broader abilities to prevent severe bacterial infections.

One Sentence Summary

This thesis defines previously unannotated phage nucleus-associated proteins that have potential roles in the export of DNA, mRNA, and macromolecules, providing insights into the mechanisms underlying phage replication and potential therapeutic targets

Chapter 1: Introduction

1.1 An old, large, and dynamic population: Bacteriophages

Bacteriophages, or phages in short, are viruses that infect bacteria or archaea by hijacking the cells for their own reproduction. Phages have specific host ranges and recognize the cells they infect by binding to specific proteins or sugars on the surface of the host cell[1]. The genetic material, either DNA or RNA in double stranded or single stranded form, is injected into the host[2]. Once the genetic material gets injected, the infection can follow two different routes: lytic or lysogenic. Lytic phages hijack the hosts immediately and make progeny phages within minutes which are released upon cell lysis to repeat the process. Lysogenic phages instead infiltrate the host without being detected and integrate its genomic material into the host genome. This way, their genome is replicated along with the host with each cell division and wait until they are triggered to enter the lytic phase.

Since their discovery by Felix d'Herelle and Frederick Twort[3,4], phages have been one of the most studied phenomena in science. For billions of years, the arms race between them and their hosts has been a major driving force for both of their evolution[5]. Since they have specific host ranges and great diversity, phage isolation and characterization has been an ongoing process[6]. Some phages can be propagated in laboratory conditions to be studied while others can only be identified by metagenomic analysis of materials from the environment[7]. A common method to isolate phages is growing the sample that contains phages in a suspension of the required host and performing an assay, such as spotting on an agar plate or liquid cultures where there will

be a decrease in the bacterial population[8]. If the results have phenotypes that indicate the phages' presence such as a zone of clearing or the appearance of lysed bacteria, the phage culturing is successful.

Phages are the most abundant biological entities in the biosphere and there are estimated to be more than 10^{31} particles in the world[9]. This corresponds to 10^{23} phage infections happening globally every second, which regulates bacterial populations as well as making their population very dynamic and ever evolving. In addition to the wide variety in their genomes and how they store it; morphologic differences of phages can be categorized as: filamentous, cubic, or pleomorphic[10]. The most frequently identified phages so far are tailed phages that have icosahedral capsids that store their genomes as double stranded DNA (Caudovirales)10. Morphologically tails can be categorized as: Siphoviridae which have long flexible non-contractile tails, Myoviridae which have contractile tails and *Podoviridae* with short stubby tails[10]. Their large numbers and diversity make phages one of the top interests of scientists since their discovery.

1.2 Significance of bacteriophages in the world

A great portion of knowledge in biology, medicine, and biotechnology has been gained through the study of phages and their interactions with their environment. The discovery that phages can facilitate horizontal gene transfer between bacteria has led to experiments that form the backbone of our understanding of genomics and cellular processes[11]. Studying the arms race between bacteria and phages allowed us to understand evolution, co-evolution and competition between biological entities[5]. The pressure on bacteria has allowed defense systems to evolve against phages[12], and in

turn, phages have evolved to evade these systems [13]. This has given us many of the molecular biology tools we use today, such as polymerases, ligases, nucleases, and bacterial restriction endonucleases that naturally target phage genomes[6]. By using these tools, researchers have exponentially increased our knowledge in biology.

The Hershey-Chase study used radiolabeled phages to establish the central dogma and the flow of genetic information starting from DNA[14]. Another crucial step in our understanding of DNA was taken with phages, as Sanger sequenced the first complete genome using Phage ΦX174[15]. Most recently, the discovery of CRISPR-Cas systems, which evolved as bacterial defense against phages, has given us the most advanced way to edit DNA in research and potentially cure genetic diseases[16–18].

In addition to CRISPR-Cas systems, phages have greatly aided us in understanding medicine. As lysogenic phages regulate the cellular processes of their host, we have learned about the root of some of the deadliest pathogens, including Escherichia, Salmonella, Staphylococcus, Streptococcus, and Vibrio[19], which have been a great threat to humanity. Similarly, we discovered how bacteria gain resistance to antibiotics, as phages carry some antibiotic resistance cassettes via horizontal gene transfer[20]. Phages and the knowledge we have gained from them have helped us to invent new antibiotics, which are crucial for medicine. In certain parts of the world, using phages to treat bacterial infections was also common, but this practice was given up with the discovery of numerous antibiotics[21]. However, as the rise of multidrug-resistant bacterial infections becomes a greater threat to humanity, phage therapy is becoming one of the forefronts of medical research.

1.3 Jumbo Phages and PhiKZ family

In the span of over a century, a wide variety of phages have been reported to have genome sizes ranging from a few kb[10] to 735 kb[7]. For more than 100 years after their discovery, almost all phage research used small genome versions that have genomes of tens of kb. This was mainly caused by the widely used phage isolation techniques, which favored the phages with genomes smaller than 200 kb[8]. Recent advancements in sequencing technologies have allowed us to collect great metagenomic data on different ecosystems, revealing that jumbo phages (more than 200 kb) widely exist and are worth investigating[7]. The sequences of detected jumbo phages encode more than a hundred genes, while most small genome phages (30-140 kb) encode only dozens of genes. By using recently available bioinformatic tools, researchers have been able to report on genes that were once thought to be only hypothetical by correlating them to proteins that play a significant role in phage replication mechanisms, such as phage-encoded tRNAs, or that allow for evading bacterial defense systems, such as anti-CRISPR proteins[7].

One of the first jumbo phages that was discovered is PhiKZ, which drew researchers' interest due to its large capsid size (120 nm in diameter) and its considerably large genome (280 kb)[22]. Out of the 306 predicted open reading frames, only 59 of them initially were able to match proteins with known functions[22]. This represented a great evolutionary distinction from known genomes, including the encoding of six of its own tRNAs[22]. Moreover, electron microscopy showed a cylindrical structure in the capsids of PhiKZ which was later named the "inner body" and potentially injected into the host alongside the DNA[23,24]. As the host of PhiKZ, Pseudomonas aeruginosa is the common

cause of infections in cystic fibrosis (CF), and several other phages similar to PhiKZ were characterized[25].

While researchers were investigating these phages for potential treatments for Cystic Fibrosis, they noticed that Rifampicin (an antibiotic targets bacterial polymerase) has no effect on PhiKZ replication[26]. This indicated that these phages do not use bacterial polymerase during their replication and instead encode their own RNA polymerases (RNAP)[26,27]. Previous studies reported bacteriophages that encode single-subunit RNAP, while ΦKZ is one of the only few viruses that encode multi-subunit RNAP[28]. In addition, they encode two different versions of multi-subunit RNAP: one that is packed into the capsid to be used for early-expressed genes and is called virion RNAP (vRNAP), and the other is expressed by vRNAP and used for expressing genes during middle/late infection (nvRNAP) [29].

Another interesting biological feature that the PhiKZ family was found to be encoding is a tubulin homolog called PhuZ, which stands for **ph**age-encoded T**u**b**Z**[30]. While viruses are known to be able to hijack cytoskeletal machinery, it was recently discovered that phages could encode their cytoskeletal elements. Initially, a lysogenic phage, c-st, was reported to be encoding a tubulin homolog and reported to be used during the division of Clostridium botulinum[31]. Later, CGP3 prophage was shown to be encoding an actin-like protein to organize viral DNA molecules during replication[32]. Relatedly, PhuZ was shown to be organizing the phage DNA during 201Φ2-1 (a PhiKZ-like phage) infection of *Pseudomonas chlororaphis*[30]. Although PhuZ has little similarity with eukaryotic tubulin, it contains the tubulin fold, polymerizes and displays dynamic

instability[33]. Previous studies have shown that jumbo phages degrade bacterial DNA at an early stage, and phage DNA is concentrated at the midcell[30,33,34]. However, when catalytically inactive PhuZ is expressed during infection, viral DNA remains at the poles instead of midcell[30,35]. Further work has revealed that phages create a nucleoid containing viral DNA, with filaments of PhuZ anchored at one end to the host cell pole[36]. The dynamic instability of PhuZ is responsible for trafficking the nucleoid to the center of the cell [30,33,34]. Interestingly, the subpopulation of jumbo phages that contain PhuZ also contains multi-subunit RNAP, suggesting a more organized replication mechanism than small genome phages.

The literature on the replication of small genome phages indicates that they replicate in the host cytoplasm in a dispersed manner[37]. Subcellular compartmentalization is conserved in all domains of life, but organelles were previously considered unique to eukaryotes. Recently, there have been examples of organelle-like subcellular organizations found in bacteria[38]. Researchers have shown that PhuZ switches to a treadmilling mechanism during mid-infection and traffics the phage inner cell membrane-assembled phage capsids to the viral nucleoid[30,35,36]. This suggests the existence of subcellular organization during prokaryotic viral infection. When mass spectrometry was performed during 201Φ2-1 infection, a previously unannotated non-structural phage protein (now named ChmA) was found to be expressed very early and localized around the viral DNA, forming a proteinaceous shell[35,39]. By fluorescently labeling phage proteins and host proteins belonging to different biological mechanisms, researchers concluded that DNA replication and RNA transcription proteins were exclusively found inside this

structure, while translation and metabolic proteins were excluded[35,40]. As this organization of proteins and DNA resembles the eukaryotic nucleus, researchers have named this structure the "phage nucleus".

1.4 Replication mechanism of phages that encode the phage nucleus

Most of the phages studied so far inject their genomic material into the host cytoplasm and do not have a protective layer around the genomic material[37]. This allows RNA transcription to occur freely and synthesize mRNA as readily available for translation by the host ribosomes. This led bacteria to evolve defenses that target phage DNA, such as Restriction Enzymes or CRISPR-Cas[12]. When Pseudomonas and Serratia jumbo phages that encode phage nucleus were tested with bacteria containing DNA-targeting bacterial defense systems, these phages were found to be able to evade and replicate[41,42]. When EcoRI, a restriction enzyme, and Cas9 were fluorescently labeled, researchers observed that these proteins were excluded from the phage nucleus[42]. Therefore, the protection was due to these proteins physically not being able to interact with the viral DNA. However, when the viral DNA was isolated and tested with these proteins in a cell-free system, phage DNA was able to be cut by these proteins[42]. This suggests that the phage nucleus might have evolved against these DNA targeting anti-phage mechanisms.

As the phage DNA stays inside of the phage nucleus for protection, this suggests the reason why capsids are trafficked to the phage nucleus for DNA packaging. By these mechanisms, phage DNA always stays inside a protective layer. The host ribosomes that are required for translation are not trafficked for translation[35]. Instead, the mRNA that is synthesized inside is exported from the phage nucleus to the host cytoplasm to be

translated. This makes the phage mRNA susceptible to host defense systems[41]. As expected, hosts that express Cas13 that targets phage mRNA were found to be protected against the phage nucleus forming Serratia jumbo phages[41]. These observations raise the questions of how this replication system evolved to selectively import proteins and exporting mechanisms for DNA and mRNA.

From the literature of small genome phages' replication, phage DNA packaging is known to be a complex process that involves different steps and proteins. Briefly, the phage DNA is replicated by DNA polymerases. In the meantime, the capsid proteins are expressed and start to oligomerize with the portal protein to form the empty capsid. The phage genome contains packaging signals that are located near the ends of the phage genomes and are recognized and cut by terminases. Once the pro-capsid is assembled, cut phage DNA passes through the portal protein with the help of packaging ATPases. As the packaging ATPase functions as a molecular motor by using the energy from ATP hydrolysis, DNA moves to the empty capsid to be properly organized.

We do not know the full details of how nucleus-encoding jumbo phages accommodate the required changes arising from its subcellular organization. So far, we have shown that empty capsids are trafficked by PhuZ and dock on the phage nucleus to be filled[36]. In the meantime, PhuZ filament treadmilling rotates the phage nucleus, allowing even capsid distribution on the phage nucleus[36]. Once the phage DNA is packaged, full capsids detach from the phage nucleus surface and localize on either side[43]. In these positions, they become organized in structures called Phage Bouquets in which full capsids assemble with other structural proteins such as tail proteins[43]. When

mature viral particles fully form, the host cell lyses [43]. In comparison to small genome phages, which their replication takes 30-40 minutes [44], one round of PhiKZ-like phages' replication takes around 120 minutes[43].

1.5 What we do not know about the phage nucleus

Since the discovery of the phage nucleus, our understanding of it is constantly evolving with new research being conducted. While the phage nucleus probably is a result of co-evolution between bacteriophage and bacteria, we still do not know its origin. One possibility is, unlike other small genome phages, certain phages' genomes were not replicating in a dispersed manner and have been in a state like phase separation, which a protein related to ChmA coat the viral genome and later evolved to be a more complex structure. A similar idea would be that the phage nucleus did not evolve spontaneously but due to evolutionary pressure caused by defense systems of the host that target phage DNA. Alternatively, phages might have obtained the genes required to have this type of compartmentalization from bacteria that had proto-nucleus like structures and started to adapt it for their use.

Similar to its origin, we still do not know the exact structure of the phage nucleus and how it is controlled. The current model suggests most of the structure is formed by ChmA, however there are indicators that ChmA alone *in vivo* is not enough to form a functional phage nucleus[35]. The model suggests how ChmA monomers can interact with each other and sufficient to form higher oligomeric structures alone *in vitro* [39]. However, it is not clear how the phage nucleus grows and how the growth is regulated. One hypothesis is that the accumulation of DNA due to genome replication creates a pressure on the phage

nucleus to increase in volume, which regulates the growth. Even if this is the case, it is still not clear how viral DNA and pressure is being detected by the structure to accommodate the volume needed.

Relatedly, it is also not known how the phage nucleus is able to distinguish the phage DNA from the bacterial genome. In the PhiKZ-like jumbo phages, it is reported that the bacterial genome is degraded early in the infection, however there are examples such as Goslar (*E.coli*) or RAY (*Erwinia*) phages does not fully degrade the host genome[45,46]. The earliest detectable formed phage nucleus with GFP-tagged ChmA colocalizes with the phage DNA in a space where the bacterial genome is degraded[35]. This raises the question of whether viral DNA gets detected by a certain sequence or the state it is in to be targeted by ChmA for the location to assemble of the phage nucleus. Another possibility is that viral DNA is being injected with proteins that would be forming an earlier stage of the phage nucleus that would allow the genome to be enclosed and protected from host defense throughout the infection. Mass spectrometry of the phages did not report ChmA to be packaged with the viral genome[24]. However, the proteins packaged in the capsids of PhiKZ-like phages might be potentially injected for this purpose. Recent observations of inner-body proteins of phiKZ colocalize with the phage DNA at the earliest detectable time point[47]. Relatedly, the same group found that knockout of some proteins labeled as “inner body proteins” in phiKZ significantly reduced the replication in the affected strains[47]. These findings support the possibility that proteins additional to ChmA might be involved in at least an early stage of phage DNA protection. However, if this theory is true, the initial events of the phage nucleus assembly are still unclear.

Interestingly, the same study that found inner body proteins to be important for PhiKZ replication also showed that knocking out PhuZ did not significantly affect the number of progeny phages[47]. One possible explanation is that, while PhuZ is involved in trafficking the phage nucleus and capsids, its absence may not be essential as the movements of these structures can be compensated for by free diffusion. In the replication of Goslar, PhuZ does not form a bipolar spindle but instead surrounds the phage nucleus in an array-like vortex[45]. These observations suggest that PhuZ and the phage nucleus system may have different evolutionary routes and are not mutually exclusive. This raises the question of which proteins are essential for replication via the phage nucleus forming pathway. It is also unclear if the phage nucleus is essential for replication in these phages. While small genome phages replicate without a phage nucleus, it is possible that the phage nucleus provides an extra layer of protection against bacterial immune systems. Alternatively, the evolution that led to the phage nucleus system might have been so drastic that these phages' whole replication would fail without it. However, to answer this question, we first need to identify the components of the phage nucleus and how it is assembled.

Our previous research showed that the phage nucleus is mostly composed of ChmA protein. When we isolated ChmA alone, we found that it is sufficient to form assemblies of tetramer and 24-mer structures[39]. The current model for the phage nucleus was created by overlaying the structure obtained by single particle CryoEM on the Cryo-FIB ET data collected from infected cells[39,48]. Although these data are consistent with each other, we do not believe we have the full mechanism for phage nucleus formation and composition.

When ChmA is expressed in an uninfected host cell, we do not observe a fully mature phage nucleus, which indicates that other components are required to form the mature version of the phage nucleus[35]. Additionally, the calculated pore size of the current model is not wide enough to accommodate DNA, mRNA, and protein trafficking[39]. The overarching goal of this thesis is to identify the other components and how trafficking is achieved in and out of the phage nucleus.

1.6 References

1. Abedon, S. T. Bacteriophage Ecology: Population Growth, Evolution, and Impact of Bacterial Viruses. (Cambridge University Press, 2008).
2. Guttman, B., Raya, R. & Kutter, E. Basic phage biology. Bacteriophages: Biology and applications 4, 30–63 (2005).
3. d'Herelle, F. An invisible microbe that is antagonistic to the dysentery bacillus. C.R. Acad. Sci. 165, 373–375 (1917).
4. Twort, F. W. An investigation on the nature of ultra-microscopic viruses. Acta Kravsi (1961). at <https://scholar.archive.org/work/wuup4h47hbco7kysfzxn2tz4du/access/wayback/http://sgc.anlis.gob.ar/bitstream/123456789/752/2/ActaKravsi1961_2_37-40.pdf>
5. Durzyńska, J. & Goździcka-Józefiak, A. Viruses and cells intertwined since the dawn of evolution. Virol. J. 12, 169 (2015).
6. Salmond, G. P. C. & Fineran, P. C. A century of the phage: past, present and future. Nat. Rev. Microbiol. 13, 777–786 (2015).
7. Al-Shayeb, B., Sachdeva, R., Chen, L.-X., Ward, F., Munk, P., Devoto, A., Castelle, C. J., Olm, M. R., Bouma-Gregson, K., Amano, Y., He, C., Méheust, R., Brooks, B., Thomas, A., Lavy, A., Matheus-Carnevali, P., Sun, C., Goltsman, D. S. A., Borton, M. A., Sharrar, A., Jaffe, A. L., Nelson, T. C., Kantor, R., Keren, R., Lane, K. R., Farag, I. F., Lei, S., Finstad, K., Amundson, R., Anantharaman, K., Zhou, J., Probst, A. J., Power, M. E., Tringe, S. G., Li, W.-J., Wrighton, K., Harrison, S., Morowitz, M., Relman, D. A., Doudna, J. A., Lehours, A.-C., Warren, L., Cate, J. H. D., Santini, J. M. & Banfield, J. F. Clades of huge phages from across Earth's ecosystems. Nature 578, 425–431 (2020).
8. Serwer, P., Hayes, S. J., Thomas, J. A. & Hardies, S. C. Propagating the missing bacteriophages: a large bacteriophage in a new class. Virol. J. 4, 21 (2007).
9. Hendrix, R. W., Smith, M. C. M., Burns, R. N., Ford, M. E. & Hatfull, G. F. Evolutionary relationships among diverse bacteriophages and prophages: All the world's a phage. Proceedings of the National Academy of Sciences 96, 2192–2197 (1999).
10. Hatfull, G. F. & Hendrix, R. W. Bacteriophages and their genomes. Curr. Opin. Virol. 1, 298–303 (2011).
11. Thomas, C. M. & Nielsen, K. M. Mechanisms of, and barriers to, horizontal gene transfer between bacteria. Nat. Rev. Microbiol. 3, 711–721 (2005).

12. Bernheim, A. & Sorek, R. The pan-immune system of bacteria: antiviral defence as a community resource. Nat. Rev. Microbiol. 18, 113–119 (2020).
13. Samson, J. E., Magadán, A. H., Sabri, M. & Moineau, S. Revenge of the phages: defeating bacterial defences. Nat. Rev. Microbiol. 11, 675–687 (2013).
14. Hershey, A. D. & Chase, M. Independent functions of viral protein and nucleic acid in growth of bacteriophage. J. Gen. Physiol. 36, 39–56 (1952).
15. Sanger, F., Air, G. M., Barrell, B. G., Brown, N. L., Coulson, A. R., Fiddes, J. C., Hutchison, C. A., Slocombe, P. M. & Smith, M. Nucleotide sequence of bacteriophage ϕX174 DNA. Nature 265, 687–695 (1977).
16. Mojica, F. J., Juez, G. & Rodríguez-Valera, F. Transcription at different salinities of Haloferax mediterranei sequences adjacent to partially modified PstI sites. Mol. Microbiol. 9, 613–621 (1993).
17. Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A. & Charpentier, E. A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science 337, 816–821 (2012).
18. Burstein, D., Harrington, L. B., Strutt, S. C., Probst, A. J., Anantharaman, K., Thomas, B. C., Doudna, J. A. & Banfield, J. F. New CRISPR–Cas systems from uncultivated microbes. Nature 542, 237–241 (2016).
19. Lin, D. M., Koskella, B. & Lin, H. C. Phage therapy: An alternative to antibiotics in the age of multi-drug resistance. World J. Gastrointest. Pharmacol. Ther. 8, 162–173 (2017).
20. de la Cruz, F. & Davies, J. Horizontal gene transfer and the origin of species: lessons from bacteria. Trends Microbiol. 8, 128–133 (2000).
21. Lu, T. K. & Koeris, M. S. The next generation of bacteriophage therapy. Curr. Opin. Microbiol. 14, 524–531 (2011).
22. Mesyanzhinov, V. V., Robben, J., Grymonprez, B., Kostyuchenko, V. A., Bourkaltseva, M. V., Sykilinda, N. N., Krylov, V. N. & Volckaert, G. The genome of bacteriophage ϕKZ of Pseudomonas aeruginosa. J. Mol. Biol. 317, 1–19 (2002).
23. Wu, W., Thomas, J. A., Cheng, N., Black, L. W. & Steven, A. C. Bubblegrams Reveal the Inner Body of Bacteriophage ϕKZ. Science 335, 182–182 (2012).
24. Thomas, J. A., Weintraub, S. T., Wu, W., Winkler, D. C., Cheng, N., Steven, A. C. & Black, L. W. Extensive proteolysis of head and inner body proteins by a morphogenetic protease in the giant Pseudomonas aeruginosa phage ϕKZ. Mol. Microbiol. 84, 324–339 (2012).
25. Monson, R., Foulds, I., Foweraker, J., Welch, M. & Salmond, G. P. C. The Pseudomonas aeruginosa generalized transducing phage phiPA3 is a new member of the phiKZ-like group of 'jumbo' phages, and infects model laboratory strains and clinical isolates from cystic fibrosis patients. Microbiology 157, 859–867 (2011).
26. Ceyssens, P.-J., Minakhin, L., Van den Bossche, A., Yakunina, M., Klimuk, E., Blasdel, B., De Smet, J., Noben, J.-P., Bläsi, U., Severinov, K. & Lavigne, R. Development of giant bacteriophage ϕKZ is independent of the host transcription apparatus. J. Virol. 88, 10501–10510 (2014).
27. Lavysh, D., Sokolova, M., Minakhin, L., Yakunina, M., Artamonova, T., Kozyavkin, S., Makarova, K. S., Koonin, E. V. & Severinov, K. The genome of AR9, a giant transducing Bacillus phage encoding two multisubunit RNA polymerases. Virology 495, 185–196 (2016).

28. Weinheimer, A. R. & Aylward, F. O. A distinct lineage of Caudovirales that encodes a deeply branching multi-subunit RNA polymerase. Nat. Commun. 11, 4506 (2020).
29. Wicke, L., Ponath, F., Coppens, L., Gerovac, M., Lavigne, R. & Vogel, J. Introducing differential RNA-seq mapping to track the early infection phase for Pseudomonas phage ϕKZ. RNA Biol. 18, 1099–1110 (2021).
30. Kraemer, J. A., Erb, M. L., Waddling, C. A., Montabana, E. A., Zehr, E. A., Wang, H., Nguyen, K., Pham, D. S. L., Agard, D. A. & Pogliano, J. A phage tubulin assembles dynamic filaments by an atypical mechanism to center viral DNA within the host cell. Cell 149, 1488–1499 (2012).
31. Oliva, M. A., Martin-Galiano, A. J., Sakaguchi, Y. & Andreu, J. M. Tubulin homolog TubZ in a phage-encoded partition system. Proc. Natl. Acad. Sci. U. S. A. 109, 7711–7716 (2012).
32. Donovan, C., Heyer, A., Pfeifer, E., Polen, T., Wittmann, A., Krämer, R., Frunzke, J. & Bramkamp, M. A prophage-encoded actin-like protein required for efficient viral DNA replication in bacteria. Nucleic Acids Res. 43, 5002–5016 (2015).
33. Erb, M. L., Kraemer, J. A., Coker, J. K. C., Chaikeeratisak, V., Nonejuie, P., Agard, D. A. & Pogliano, J. A bacteriophage tubulin harnesses dynamic instability to center DNA in infected cells. Elife 3, (2014).
34. Zehr, E. A., Kraemer, J. A., Erb, M. L., Coker, J. K. C., Montabana, E. A., Pogliano, J. & Agard, D. A. The structure and assembly mechanism of a novel three-stranded tubulin filament that centers phage DNA. Structure 22, 539–548 (2014).
35. Chaikeeratisak, V., Nguyen, K., Khanna, K., Brilot, A. F., Erb, M. L., Coker, J. K. C., Vavilina, A., Newton, G. L., Buschauer, R., Pogliano, K., Villa, E., Agard, D. A. & Pogliano, J. Assembly of a nucleus-like structure during viral replication in bacteria. Science 355, 194–197 (2017).
36. Chaikeeratisak, V., Khanna, K., Nguyen, K. T., Sugie, J., Egan, M. E., Erb, M. L., Vavilina, A., Nonejuie, P., Nieweglowska, E., Pogliano, K., Agard, D. A., Villa, E. & Pogliano, J. Viral Capsid Trafficking along Treadmilling Tubulin Filaments in Bacteria. Cell 177, 1771–1780.e12 (2019).
37. Kasman, L. M. & Porter, L. D. Bacteriophages. (StatPearls Publishing, 2022).
38. Greening, C. & Lithgow, T. Formation and function of bacterial organelles. Nat. Rev. Microbiol. 18, 677–689 (2020).
39. Laughlin, T. G., Deep, A., Prichard, A. M., Seitz, C., Gu, Y., Enustun, E., Suslov, S., Khanna, K., Birkholz, E. A., Armbruster, E., McCammon, J. A., Amaro, R. E., Pogliano, J., Corbett, K. D. & Villa, E. Architecture and self-assembly of the jumbo bacteriophage nuclear shell. Nature 608, 429–435 (2022).
40. Chaikeeratisak, V., Nguyen, K., Egan, M. E., Erb, M. L., Vavilina, A. & Pogliano, J. The Phage Nucleus and Tubulin Spindle Are Conserved among Large Pseudomonas Phages. Cell Rep. 20, 1563–1571 (2017).
41. Malone, L. M., Warring, S. L., Jackson, S. A., Warnecke, C., Gardner, P. P., Gumy, L. F. & Fineran, P. C. A jumbo phage that forms a nucleus-like structure evades CRISPR-Cas DNA targeting but is vulnerable to type III RNA-based immunity. Nat Microbiol 5, 48–55 (2020).

42. Mendoza, S. D., Nieweglowska, E. S., Govindarajan, S., Leon, L. M., Berry, J. D., Tiwari, A., Chaikeeratisak, V., Pogliano, J., Agard, D. A. & Bondy-Denomy, J. A bacteriophage nucleus-like compartment shields DNA from CRISPR nucleases. Nature 577, 244–248 (2020).
43. Chaikeeratisak Vorrapon, Khanna Kanika, Nguyen Katrina T., Egan MacKennon E., Enustun Eray, Armbruster Emily, Lee Jina, Pogliano Kit, Villa Elizabeth & Pogliano Joe. Subcellular organization of viral particles during maturation of nucleus-forming jumbo phage. Science Advances 8, eabj9670 (2022).
44. Brüssow, H., Canchaya, C. & Hardt, W.-D. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. Microbiol. Mol. Biol. Rev. 68, 560–602, table of contents (2004).
45. Birkholz, E. A., Laughlin, T. G., Armbruster, E., Suslov, S., Lee, J., Wittmann, J., Corbett, K. D., Villa, E. & Pogliano, J. A cytoskeletal vortex drives phage nucleus rotation during jumbo phage replication in E. coli. Cell Rep. 40, 111179 (2022).
46. Prichard, A., Lee, J., Laughlin, T. G., Lee, A., Thomas, K. P., Sy, A., Spencer, T., Asavavimol, A., Cafferata, A., Cameron, M., Chiu, N., Davydov, D., Desai, I., Diaz, G., Guereca, M., Hearst, K., Huang, L., Jacobs, E., Johnson, A., Kahn, S., Koch, R., Martinez, A., Norquist, M., Pau, T., Prasad, G., Saam, K., Sandhu, M., Sarabia, A. J., Schumaker, S., Sonin, A., Uyeno, A., Zhao, A., Corbett, K., Pogliano, K., Meyer, J., Grose, J. H., Villa, E., Dutton, R. & Pogliano, J. Identifying the core genome of the nucleus-forming bacteriophage family and characterization of Erwinia phage RAY. bioRxiv (2023). doi:10.1101/2023.02.24.529968
47. Guan, J., Oromí-Bosch, A., Mendoza, S. D., Karambelkar, S., Berry, J. D. & Bondy-Denomy, J. Bacteriophage genome engineering with CRISPR--Cas13a. Nature microbiology 1–11 (2022).
48. Nieweglowska, E. S., Brilot, A. F., Méndez-Moran, M., Kokontis, C., Baek, M., Li, J., Cheng, Y., Baker, D., Bondy-Denomy, J. & Agard, D. A. The ϕPA3 phage nucleus is enclosed by a self-assembling 2D crystalline lattice. Nat. Commun. 14, 927 (2023).

Chapter 2: Identification of the bacteriophage nucleus protein interaction network

Eray Enustun[1], Amar Deep[2], Yajie Gu[2], Katrina T. Nguyen[1], Vorrapon Chaikeeratisak[1,3], Emily Armbruster[1], Majid Ghassemian[4], Elizabeth Villa[1,5], Joe Pogliano[1]‡, Kevin D. Corbett[2]‡

[1]Department of Molecular Biology, University of California San Diego, La Jolla, CA, USA.
[2]Department of Cellular and Molecular Medicine, University of California San Diego, La Jolla, CA, USA.
[3]Department of Biochemistry, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand
[4]Biomolecular and Proteomics Mass Spectrometry Facility, University of California San Diego, La Jolla, CA, USA.
[5]Howard Hughes Medical Institute, La Jolla, California, USA.

2.1 Abstract

In the arms race between bacteria and bacteriophages (phages), some large-genome jumbo phages have evolved a protein shell that encloses their replicating genome to protect it against host immune factors. By segregating the genome from the host cytoplasm, however, the “phage nucleus” introduces the need to specifically translocate mRNA and proteins through the nuclear shell, and to dock capsids on the shell for genome packaging. Here, we use proximity labeling and localization mapping to systematically identify proteins associated with the major nuclear shell protein chimallin (ChmA) and other distinctive structures assembled by these phages. We identify six uncharacterized nuclear shell-associated proteins, one of which directly interacts with self-assembled ChmA. The structure and protein-protein interaction network of this protein, which we term ChmB, suggests that it forms pores in the ChmA lattice that serve as docking sites for capsid genome packaging, and may also participate in mRNA and/or protein translocation.

2.2 Introduction

Since the discovery of phages more than a century ago, research on these remarkable entities has yielded fundamental insights into a broad range of pathways across biology[1]. Historically, most phage studies have focused on small-genome phages (~30-140 kb), leaving larger "jumbo phages" with genomes over ~200 kb much less well understood despite their abundance in nature[2–4]. We previously showed that one family of jumbo phages forms distinctive structures in infected cells, including a nucleus-like compartment bounded by a proteinaceous shell and a spindle-like structure that centers and rotates this compartment within the host cell[5–7]. The phage nucleus encloses the replicating phage genome and excludes most host proteins including CRISPR effectors and restriction enzymes, rendering this family of phages broadly resistant to DNA-targeting bacterial immune systems[8,9].

The jumbo phage nuclear shell provides an important selective advantage by protecting the replicating phage genome from host-encoded defense nucleases, but that protection comes at the cost of significant added complexity in the phage life cycle. The phage nuclear shell is composed primarily of one protein, termed Chimallin (ChmA) or Phage Nuclear Enclosure (PhuN), which forms a single-layer thick flexible lattice that separates the phage genome from the bacterial cytoplasm[10,11]. Pores in the ChmA lattice are less than ~2 nm in width, large enough to pass metabolites but too small for passage of most proteins or mRNAs[10]. As in eukaryotes, mRNAs are transcribed within the phage nucleus but translated in the cytoplasm, meaning that phage mRNAs must be translocated out of the nucleus[6]. At the same time, phage-encoded proteins necessary for genome replication and mRNA transcription must be specifically translocated into the

nucleus[6,12]. Finally, during virion production newly assembled capsids are trafficked along filaments of the tubulin-like protein PhuZ to the nuclear shell, where they dock for genome packaging[6,13]. Following genome packaging, capsids are assembled with virion tails at a pair of structures termed the “phage bouquets” prior to cell lysis and virion release[14,15].

The diverse functions of the jumbo phage nuclear shell – including mRNA and protein translocation through the shell, and capsid docking on the shell – imply that this structure incorporates multiple components in addition to ChmA that mediate these functions. Here, we use proximity labeling (miniTurboID[16]) in *Pseudomonas aeruginosa* cells infected by the jumbo phage ΦPA3 [17] to identify proteins that localize both within the phage nucleus and specifically to the nuclear shell. We identify six new nuclear shell-associated proteins, one of which interacts directly with both ChmA and the putative portal protein[18]. These interactions suggest that this protein, which we term ChmB, forms pores in the ChmA lattice and mediates capsid docking and genome packaging. ChmB’s overall protein-protein interaction network further suggests additional roles in mRNA and/or protein translocation across the phage nuclear shell. More broadly, our data define the protein interaction network of the jumbo phage nuclear shell, and reveal the subcellular localization of dozens of previously uncharacterized jumbo phage proteins.

2.3 Results

2.3.1 Identification of jumbo phage nucleus-associated proteins

The genomes of nucleus-forming jumbo phages are poorly characterized: for instance, 290 of the 378 genes encoded by the *Pseudomonas* jumbo phage ΦPA3 have no annotated function in the NCBI protein database. To overcome this deficit, we used a proximity labeling approach to identify proteins associated with the phage nuclear shell that could endow this structure with additional functionality like mRNA/protein translocation and capsid docking. We fused the promiscuous biotin ligase miniTurboID[16] to the ChmA protein from the jumbo phage ΦPA3 (gp53) and to the phage's nuclear-localized RecA protein (gp175) (**Fig. 1a-b, Extended Data Fig. 1a-b**)[7]. We first verified that fusing either gp53 or gp175 to GFP and miniTurboID did not alter these proteins' localization in ΦPA3-infected cells (**Extended Data Fig. 1a-b**). We then expressed miniTurboID-gp53 or gp175-miniTurboID (lacking GFP to minimize off-target effects) in *Pseudomonas aeruginosa* cells infected with ΦPA3, collected samples at 45 minutes post infection (mpi) which is the earliest time point when the mature nuclear shell is observed with docked capsids[7], and performed streptavidin pulldown and mass spectrometry analysis to identify biotinylated proteins. By focusing on the phage proteins that were biotinylated and normalizing the results to a control miniTurboID-GFP fusion (which remains diffuse in the host cell cytoplasm throughout infection; **Extended Data Fig. 1b**), we identified candidate proteins that preferentially localize in close proximity to ChmA and/or RecA (**Tables 1-2, Extended Data Tables 1-2**).

To validate our interaction data, we generated GFP fusions of the top 25 ChmA- and RecA-interacting proteins from our miniTurboID datasets (42 proteins plus ChmA and RecA; **Table 3**), expressed each in *P. aeruginosa*, and determined their localization in both uninfected and ΦPA3-infected cells. Proteins that interact with RecA are expected to localize inside the phage nucleus, while ChmA-interacting proteins are expected to localize on or near the nuclear shell. Given the architecture of the ChmA lattice, we expect the N-terminal miniTurboID tag to be localized on the outer surface of the nuclear shell[10]. Among the 42 proteins tested, we identified eight that localize within the nucleus-like compartment, six that localize to the nuclear shell itself, and five that localize to the phage bouquets (**Fig. 1c-e**). Most proteins that localize within the phage nucleus are annotated in the NCBI database as putative nucleic acid interacting proteins, including a subunit of the phage-encoded non-virion RNA polymerase (nvRNAP; gp62)[19], two predicted helicases (gp131 and gp134), a predicted DNA ligase (gp257), and two predicted endonucleases (gp78 and gp210) (**Fig. 1c, Table 3**)[18]. Two nuclear-localized proteins (gp108 and gp200) have no annotated or predicted function. Of the five proteins that localize to phage bouquets, three (gp12, gp164, and gp166) are predicted phage tail proteins, one (gp233) is a predicted helicase, and one (gp355) has no annotated or predicted function[18,20,21].

To date, the only known component of the jumbo phage nuclear shell is ChmA[6,7,10]. Among our list of RecA- and ChmA-interacting proteins, we identified six proteins (gp2, gp61, gp63, gp64, gp148, and gp375) that clearly localize to the nuclear shell upon ΦPA3 infection of *P. aeruginosa* cells (**Fig. 1d, Table 3**). One of these proteins, gp148, is

predicted to be the portal protein of the phage capsid[18]. In other phages, the portal protein forms a homododecameric complex that orchestrates capsid assembly[22] and associates with the terminase to translocate genomic DNA into the capsid[23–27]. We previously showed that in this family of jumbo phages, capsids are docked on the nuclear shell for genomic DNA packaging[6] , and our finding that the putative ΦPA3 portal protein associates with the shell suggests that it is directly responsible for capsid docking. Notably, we find that overexpressed GFP-tagged gp148 localizes to the phage nuclear shell as early as 30 minutes post infection, well before capsid assembly and docking begins at around 45 minutes post infection[6] (**Extended Data Fig. 3**). This finding suggests that the portal protein can localize to the nuclear shell on its own, supporting the idea that it directly mediates capsid docking.

Apart from the portal protein, the remaining five nuclear shell-localized proteins have no predicted function and were not found in prior mass spectrometry studies of mature jumbo phage virions[20]. Three of these five proteins (gp61, gp63, gp64) are in a block of genes that is well-conserved across jumbo phages, while homologs of the remaining two proteins (gp2 and gp375) can be identified only in *Pseudomonas*-infecting jumbo phages (**Extended Data Fig. 4**). None of these five proteins show detectable sequence homology to any other known proteins. Moreover, all five proteins are expressed early in infections, with timing similar to the major nuclear shell protein ChmA[6,7]. Thus, we speculate that some or all of these proteins are components of the nuclear shell itself, and that they may mediate translocation of mRNA and/or proteins

through the nuclear shell, and mediate docking of capsids to the shell for genome packaging.

2.3.2 gp2 is an interaction hub at the nuclear shell

Among the identified nuclear shell-associated proteins, gp2 was among the most highly biotinylated proteins in our miniTurboID-ChmA samples (**Table 1**). We verified that GFP fusions of both ΦPA3 gp2 and its homolog from the related jumbo phage 201Φ2-1 (also gp2) colocalize with mCherry-fused ChmA in infected cells (**Fig. 2a, b**). To further define the interaction network of ΦPA3 gp2, we expressed GFP-fused gp2 in ΦPA3-infected *P. aeruginosa* cells, then purified the protein and interacting partners using GFP affinity chromatography. After purification of gp2 with a C-terminal GFP tag, we identified two strong bands at a molecular weight of ~70 kDa on a silver-stained SDS-PAGE gel (**Fig. 2c**). We extracted a gel slice containing these two closely spaced bands and used trypsin mass spectrometry to identify the proteins. In this sample, the strongest signal (149 peptides, 81% sequence coverage) was for ChmA (gp53, 66.7 kDa), and the second strongest (41 peptides, 40% sequence coverage) was for the major capsid protein (gp136, 82.8 kDa; **Table S3**). These data suggest that the observed doublet at ~70 kDa represents these two proteins, and that gp2 interacts strongly with both ChmA and capsids.

To further investigate the interactions between nuclear shell-associated proteins, we next performed mass spectrometry on the full purified samples from GFP-tagged ChmA, gp2 (both N- and C-terminal GFP tags), and the putative nvRNAP subunit gp62 (**Fig. 2d**, **Extended Data Fig. 5a**, **Extended Data Table 4**). We confirmed that this

approach is successful in purifying functional protein complexes, as we successfully identified all five subunits of the nvRNAP complex in the gp62-GFP pulldown[19] (**Fig. 2d**, **Extended Data Table 4**). The most enriched protein in GFP-tagged gp2 samples was ChmA, and we also identified the major capsid protein (gp136), the predicted portal protein (gp148), and 14 other proteins annotated as either tail or virion structural proteins (**Fig. 2d**, **Extended Data Table 4**). Two predicted phage tail proteins (gp213 and gp11) were also identified in pulldowns with gp62, suggesting that these proteins may simply be highly abundant in cell lysates (**Extended Data Table 4**). Nonetheless, the strong enrichment of phage structural proteins in GFP-tagged gp2 pulldowns strongly suggests that gp2 interacts directly with capsids, potentially as they dock on the nuclear shell for genomic DNA packaging. Also identified in the GFP-tagged gp2 pulldowns were two other shell-associated proteins, gp63 and gp64 (**Fig. 2d**, **Extended Data Table 4**), suggesting that these proteins may interact either directly with gp2, or instead interact indirectly through the ChmA lattice. Finally, two additional proteins (gp138 and gp358) were identified in both the ChmA and gp2 GFP pulldowns, as well as having been detected by ChmA-miniTurboID labeling (**Fig. 2d**, **Extended Data Tables 1**, **4**). These two proteins are conserved across jumbo phages infecting *Pseudomonas* but have no annotated or predicted function, leaving their potential roles unknown.

2.3.3 gp2 interacts directly with ChmA and the phage portal protein

We and others have shown that the nuclear shell in diverse jumbo phages including ΦPA3 is composed primarily of ChmA[5–7], which self-assembles into closed structures *in vitro* and forms a flexible lattice that surrounds the phage genome in phage-

infected cells[10,11,28]. To confirm that gp2 interacts directly with ChmA, we co-expressed 201Φ2-1 gp2 and ChmA in *E. coli* and performed Ni^{2+} pulldowns using a His_6-tag on gp2 (ΦPA3 ChmA is poorly-expressed in *E. coli*, precluding analysis in this phage). We found that His_6-tagged 201Φ2-1 gp2 robustly interacts with ChmA in this assay, showing that the two proteins directly interact (**Fig. 2e, Extended Data Fig. 5b**). We next deleted the N- and C-terminal segments of ChmA (NTS and CTS, respectively), which bind to neighboring protomers in the ChmA lattice to mediate nuclear shell assembly. *In vitro*, deletion of either NTS or CTS leads to a loss of ChmA self-assembly[10], and we found that gp2 was unable to interact with ChmA mutants lacking NTS, CTS, or both NTS and CTS (**Fig. 2e**). The loss of gp2 binding when deleting either the ChmA NTS or CTS suggests that gp2 does not simply bind one of these tail segments; rather the data suggest that gp2 interacts specifically with the assembled ChmA lattice.

We next performed a similar coexpression experiment with His_6-tagged ΦPA3 gp2 and the portal protein gp148. We found that gp2 could interact directly with gp148 in this assay (**Fig. 2f**). Combined with our data showing that gp2 interacts directly with self-assembled ChmA, these data suggest that gp2 is an integral component of the phage nuclear shell that is directly involved in the docking and filling of capsids through an interaction with the portal. Based on its localization and likely crucial role in phage nuclear structure and function, we name ΦPA3 gp2 and its homologs in related jumbo phages chimallin B (ChmB).

2.3.4 ChmB forms a homodimer with a novel fold

To determine the structural basis for ChmB interactions with other proteins, we recombinantly purified the protein from several jumbo phages and determined a 2.6 Å resolution crystal structure of ChmB from the related phage PA1C (gp2; 38% identical to ΦPA3 gp2) (**Extended Data Table 5**). ChmB forms a homodimer in solution (**Fig. 3a, Extended Data Fig. 6a-c**), and the structure reveals an intertwined dimeric structure with each protomer's N-terminus forming a short β-strand and an α-helix that pack against the C-terminal globular domain of its dimer mate. Overall, the ChmB dimer adopts a distinctive U shape with dimensions of ~5 x 8 nm (**Fig. 3b**). Searches with the DALI or FoldSeek protein structure comparison tools[29,30] show no known structural relatives.

2.3.5 ChmB point mutants disrupt phage nucleus formation

To determine the roles of ChmB in phage nucleus formation and function, we first overexpressed GFP-tagged wild-type ΦPA3 ChmB in ΦPA3-infected *P. aeruginosa* cells (**Fig. 4a**). In uninfected cells, gp2 overexpression did not cause a significant growth defect, indicating that the protein is not inherently toxic (**Extended Data Fig. 7a-b**). In infected cells, we observed a striking increase in the average size of the phage nucleus and a concomitant increase in the nuclear DNA content (as measured by total DAPI signal within the nucleus) (**Fig. 4b-c, Extended Data Fig. 8a-c**). Further, a significant fraction of cells (42%) showed ChmB localization suggestive of multiple juxtaposed phage nuclei, or a single phage nucleus with an aberrant shell structure (**Fig. 4d**).

We next aligned ChmB homologs from jumbo phages that infect *Pseudomonas* species (**Extended Data Fig. 5d**) and identified two highly conserved surface residues:

Q53 and A159 (ΦPA3 gp2 numbering). We generated point mutations of these residues (Q53A and A159D, respectively) designed to alter the ChmB surface and potentially disrupt specific protein-protein interactions. In *P. aeruginosa* cells infected with ΦPA3, overexpression of either ChmB-Q53A or A159D reversed the large/multiple phage nucleus phenotype observed when overexpressing wild-type ChmB, and caused a significant disruption in nucleus formation and/or DNA content in a large fraction of cells. In around 20% of cells (23% for ChmB-Q53A, 17% for ChmB-A159D), the phage nuclear DNA signal in late infections resembled the puncta usually observed in very early infections, prior to significant phage nuclear DNA replication and nucleus growth (**Fig. 4e, Extended Data Fig. 8d-e**)[6]. In another population of cells (5% for ChmB-Q53A, 7% for ChmB-A159D), the nuclear DNA appeared to be entirely absent despite these cells sometimes showing ChmB localization reminiscent of a phage nuclear shell (**Fig. 4e**). Finally, a third population of cells (14% for ChmB-Q53A, 21% for ChmB-A159D) showed an aberrant nuclear shell structure with nuclear DNA staining that appeared to incompletely fill the nuclear area (**Fig. 4e**). Overall, nearly half of infected cells expressing mutant ChmB (42% for ChmB-Q53A, 45% for ChmB-A159D) showed abnormal nuclear shell and/or nuclear DNA morphology.

To test whether the ChmB point mutants affected oligomerization or binding to known partner proteins, we first purified wild-type or mutant ΦPA3 ChmB and analyzed the purified proteins by size exclusion chromatography. All three proteins showed similar elution profiles, indicating that ChmB oligomerization is not affected by the mutations (**Extended Data Fig. 7c**). Next, we coexpressed His_6-tagged 201Φ2-1 ChmB (wild-type

or mutants equivalent to ΦPA3 ChmB Q53A and A159D) with untagged ChmA. Neither mutation affected ChmA association in our pulldown assay (**Extended Data Fig. 7d**), consistent with our observation that ChmB point mutants localize properly to the phage nuclear shell in infected cells (**Fig. 4a**). Similarly, neither ChmB mutation affected the ability of His_6-tagged ΦPA3 ChmB to associate with the portal protein (gp148) in a coexpression assay (**Extended Data Fig. 7e**). Thus, the observed effects of ChmB point mutants on phage nuclear shell development are likely not caused by a failure to interact with ChmA or the phage capsids. Rather, these effects may involve other putative interaction partners of ChmB, such as the uncharacterized nuclear shell-associated proteins gp61, gp63, and gp64.

2.4 Discussion

In contrast to well-studied small-genome phages, nucleus-forming jumbo phages build several characteristic structures in infected cells including the phage nucleus[6,7,9,15], PhuZ spindle[6,7,13,31], and phage bouquets[14,15]. The distinctive subcellular organization imposed by these phages enabled us to use proximity labeling with a phage for the first time, allowing the identification of dozens of phage proteins that localize to these distinctive structures. By combining proximity labeling with subcellular localization analysis by fluorescence microscopy, we established a subcellular protein localization map comprising 44 phage proteins (**Fig. 5a**). While this map is incomplete, it nonetheless represents a major step in functional understanding of this distinctive family of phages.

Here, we focus on the phage nucleus, which separates the replicating phage genome from the host cytoplasm and protects the phage from DNA-targeting immune factors encoded by the host. We and others have shown that the jumbo phage nuclear shell is predominantly composed of a single layer of the phage-encoded protein chimallin (ChmA), which assembles into a lattice with pores less than 2 nm in width[10,11]. Because these pores are likely too small for the passage of nucleic acids or proteins, we theorized that the ChmA lattice incorporates additional components that mediate mRNA and protein translocation through the phage nuclear shell. Further, based on our prior observation that capsids dock on the nuclear shell for genome packaging[5], this structure must also incorporate components that mediate capsid docking and enable the passage of genomic DNA through the nuclear shell.

Our proximity labeling and localization mapping approach identified six proteins in the jumbo phage ΦPA3 that associate with the phage nuclear shell. We find that one of these proteins (gp2) associates directly with self-assembled ChmA *in vitro*, suggesting that it is an integral phage nuclear shell protein and prompting us to name it ChmB. ChmA self-assembles into a flexible square lattice with extended N- and C-terminal segments (NTS and CTS, respectively) binding to neighboring protomers (**Fig. 5b**)[10,11]. NTS-mediated interactions define ChmA homotetramers that measure ~11.5 x 11.5 nm, while CTS-mediated interactions mediate interactions between neighboring tetramers. The distinctive U shape and overall dimensions of the ChmB dimer, at ~5 x 8 nm, suggest a model in which multiple ChmB dimers could line a hole created by the removal of four or more neighboring ChmA protomers from the lattice (two possibilities shown in **Fig. 5b**). Indeed, size exclusion chromatography of ChmB at high concentration suggests a propensity for higher-order self-assembly that may be reinforced by integration into the ChmA lattice (**Extended Data Fig. 7c**).

In addition to associating directly with ChmA, we find that ChmB is at the center of a large protein-protein interaction network in phage-infected cells. ChmB interacts directly with the phage portal protein, and the portal protein can also localize on its own to the phage nuclear shell in infected cells. Based on these data, we propose that ChmB mediates capsid docking and genome packaging at the nuclear shell through a direct interaction with the portal protein. ChmB may also interact with other nuclear shell-associated proteins such as gp61, gp63, and gp64 to mediate specific translocation of mRNA and protein through the shell.

Overexpression of wild-type ChmB in infected cells resulted in phage nuclei that were significantly larger than normal, often showing aberrant morphology with multiple (up to five) distinct phage nuclei in a single cell. Conversely, overexpression of mutant ChmB proteins (Q53A or A159D) led to the formation of significantly smaller phage nuclei or phage nuclei that appeared to be completely devoid of DNA based on DAPI staining. Importantly, these point mutations did not disrupt ChmA binding or ChmB localization to the phage nuclear shell, nor did they disrupt binding to the phage portal protein. The strong phenotypic effects of ChmB point-mutant expression in infected cells therefore suggests that ChmB has additional binding partners with fundamental roles in the formation and maturation of the phage nucleus.

ChmB is conserved among all known nucleus-forming jumbo phages that infect *Pseudomonas*, but it is not found in distantly related phage such as *E.coli* phage Goslar[15] or *Serratia* phage PCH45[9] (**Extended Data Fig. 4**). Given its apparent role in phage nuclear shell function, why is ChmB not more widely conserved? The most likely explanation is that it is conserved, but that distant homologs of ChmB are too divergent to be recognized by sequence-based searches. This scenario is supported by the high sequence divergence of gp2 homologs relative to other important proteins like ChmA: across five representative jumbo phages infecting *Pseudomonas* (ΦPA3, PA1C, Phabio, 201Φ2-1, and ΦKZ), pairwise sequence identities for ChmA homologs average 51%, while ChmB proteins average 27% identity (**Extended Data Fig. 4**). Alternatively, a different protein might be performing a role equivalent to ChmB in other nucleus-forming jumbo phage families.

In summary, this work reveals new insights into the organizational principles of nucleus-forming jumbo phages, and into the molecular mechanisms of the phage nucleus in particular. Our results identify a protein interaction network centered around the phage nucleus that will form the basis for future research in this area. We identified a key protein, ChmB, at the center of this network that likely plays multiple roles as both a pore for macromolecular translocation through the nuclear shell, and for capsid docking and genomic DNA packaging. Further work will be required to fully understand the composition of the phage nucleus and the myriad proteins that contribute to its remarkable functions.

2.5 Experimental Procedures

2.5.1 Strains, growth condition and phage preparation

P. chlororaphis strain 200-B cells were cultured on solid Hard Agar (HA). *P. aeruginosa* strains PA01 and K2733 (efflux pump knockout; *ΔMexAB-OprMΔMexCD-OprJΔMexEF-OprNΔMexXY-OprM*)[32] were cultured in Luria-Bertani (LB) media. To amplify phages, host strains were cultured in liquid media at 30°C overnight, then 20 µl of high-titer phage lysates were mixed with 100 µL of OD_{600}=0.6 cells (*P. chlororaphis* for 201Φ2-1, *P. aeruginosa* for ΦPA3), incubated for 20 minutes at room temperature, then mixed with 5 mL of HA (for 201Φ2-1 in *P. chlororaphis*) or LB top agar (for phage ΦPA3 in *P. aeruginosa*) and poured over a HA or LB plate. The plates were then incubated upside-down at 30°C overnight. Next day, the plates were incubated with 5 mL of phage buffer for 5 hours at room temperature. Lysates were collected and centrifuged at 15,000 rpm for 10 minutes. Supernatants were stored at 4°C with 0.01% chloroform.

2.5.2 Plasmid construction and bacterial transformation

Genes of interest were PCR-amplified from high-titer phage lysates, then ligated into linearized plasmid backbones using an NEBuilder HiFi DNA Assembly Cloning Kit (New England Biolabs # E5520S). For protein expression in *P. chlororaphis* and *P. aeruginosa*, the pHERD30T vector was used[33]. For overexpression in *E. coli*, UC Berkeley MacroLab vectors 2-BT (Ampicillin resistant, His_6-TEV tag; Addgene #29666) and 13-S (Spectinomycin resistant, no tag; Addgene # 48323) were used (**Extended Data Table 6**). Recombinant plasmids were transformed into *E. coli* DH5α and plated on LB agar containing appropriate antibiotics (25 µg/mL gentamicin sulfate, 100 µg/mL ampicillin, or 100 µg/mL spectinomycin as appropriate). Constructs were confirmed by DNA sequencing and subsequently introduced into indicated organisms of interest and selected on LB supplemented with antibiotics. Selected overnight cultures were stored in 25% glycerol at -80°C.

2.5.3 Fluorescence microscopy of single cell-infection assay

1.2% agarose pads were prepared on concavity slides. Each pad was supplemented with 0.05%-1% arabinose to induce protein expression. In certain experiments, FM4-64 (1 µg/mL) was added to stain cell membranes and DAPI (1 µg/mL) was added to stain the nucleoid. *P. chlororaphis* and *P. aeruginosa* strains were inoculated on the pads and were grown in a humid chamber at 30°C for 2 hours. For phage infections, 10 µl of phages (10^8 pfu/mL) were added on the cells and incubated for an additional 40 minutes at 30°C for infection to proceed. At the desired time points, the pads were sealed with a coverslip and fluorescent microscopy was performed with a

DeltaVision Spectris Deconvolution Microscope (Applied Precision, Issaquah, WA, USA). Cells were imaged using at least 8 images in the Z-axis from the middle focal plane in 0.15 µm increments. For time-lapse imaging, 8 Z points were selected and suspequently imaged using the UltimateFocus mode. Images were further processed by the deconvolution algorithm (DeltaVision SoftWoRx Image Analysis Program) and analyzed in Fiji[34].

2.5.4 Phage nucleus area and DAPI measurements from fixed cells

1.2% agarose pads were prepared on concavity slides containing 1% arabinose plus 1 µg/mL FM4-64. *P. aeruginosa* strains were inoculated on pads and grown for 2 hours in a humid chamber at 30°C. 75 minutes post infection with 10 µl of phage (10^8 pfu/mL), 20 µl of fixation mixture (1.5% Glutaraldehyde, 13.76% PFA, and 0.16 M $NaPO_4$ at pH=7.4) was added to each pad and incubated at room temperature for 20 minutes. For high levels of nucleoid staining for quantifications, 20 µl of 20 µg/mL DAPI was added to the fixed cells and incubated for 20 minutes at room temperature. Pads were sealed with coverslips and fluorescent microscopy was performed as above. For quantitation of DAPI area and intensity in nucleoids or phage nuclei, images prior to deconvolution were analyzed by Fiji with built-in measurement tools. Statistical analysis was performed in GraphPad Prism.

2.5.5 Proximity Labeling with miniTurboID

For proximity labeling, overnight cultures of *P. aeruginosa* were grown in LB media with 25 µg/mL gentamicin sulfate. Cultures were then diluted to OD_{600}=0.1 and supplemented with 500 µM biotin. When cells reached OD_{600}=0.5, they were diluted 1:10

into 50 mL total volume in 250 mL flasks and grown in LB supplemented with 0.1% arabinose, 500 µM biotin, 25 µg/mL gentamicin sulfate, and 0.2 mM $CaCl_2$. When the cells reached OD_{600}=0.3, the cells were infected with phage ΦPA3 at a multiplicity of infection (MOI) of 3. At 45 minutes post infection, cultures were collected and centrifuged at 4000 rpm at 4°C. Cell pellets were stored at -80°C for mass spectrometry.

2.5.6 Mass Spectrometry

To prepare biotinylated samples for IP and mass spectrometry, frozen cell pellets (100 µL) were thawed and resuspended in 100 µL water. Ten µL of resuspended cells were mixed with 200 µL of 6M guanidine-HCl, vortexed, and then subjected to three cycles of incubation at 100°C for 5 minutes, followed by cooling to room temperature. 1.8 mL of pure methanol was added to the boiled cell lysate, vortexed, incubated at -20°C for 20 minutes, then centrifuged at 14K RPM for 10 minutes at 4°C. The tube was inverted and dried to remove any liquid, then the pellet was resuspended in 200 µL of 8 M urea in 0.2 M ammonium bicarbonate. The mixture was incubated for 1 hour at 37°C with constant agitation. Following the incubation, 4 µL of 500 mM TCEP (Tris(2-carboxyethyl) phosphine) and 20 µL 400 mM chloro-acetamide were added. Protein concentration was measured by BCA assay from a 10 µL sample, then 600 µL of 200 mM ammonium bicarbonate was added to bring the urea concentration to 2 M. One µg of sequencing-grade trypsin was added for each 100 µg of protein in the sample, then incubated overnight at 42°C. Following trypsin incubation, 50 µL of 50% formic acid was added (ensuring that the pH dropped to 2 using pH test strips), then samples were desalted using C18 solid phase extraction (Waters Sep-Pak C18 12 cc Vac Cartridge #

WAT036915) as described by the manufacturer protocol. The peptide concentration of each sample was measured using BCA after resuspension in 1 ml PBS buffer.

For biotin IP, 200 µL of 50% slurry of NeutrAvidin beads (Pierce) was washed three times with PBS, then 1 mg of resuspended peptide solution in PBS was added and incubated 1 hour at room temperature. Beads were washed three times with 2 mL PBS plus 2.5% acetonitrile, washed once in ultrapure water, and excess liquid was carefully removed with a micropipette. Biotinylated peptides were eluted twice with 300 µL of elution buffer (0.2% trifluoroacetic acid, 0.1% formic acid, and 80% acetonitrile in water), with the second elution involving two 5-minute incubations at 100°C. Samples were then dried completely prior to mass spectrometry.

2.5.7 LC-MS-MS

Trypsin-digested peptides were analyzed by ultra-high pressure liquid chromatography (UPLC) coupled with tandem mass spectroscopy (LC-MS/MS) using nano-spray ionization. The nanospray ionization experiments were performed using a Orbitrap fusion Lumos hybrid mass spectrometer (Thermo) interfaced with nano-scale reverse-phase UPLC (Thermo Dionex UltiMate 3000 RSLC nano System) using a 25 cm, 75-micron ID glass capillary packed with 1.7-μm C18 (130) BEH beads (Waters corporation). Peptides were eluted from the C18 column into the mass spectrometer using a linear gradient (5–80%) of ACN (Acetonitrile) at a flow rate of 375 µl/min for 3 hours. The buffers used to create the ACN gradient were: Buffer A (98% H_2O, 2% ACN, 0.1% formic acid) and Buffer B (100% ACN, 0.1% formic acid). Mass spectrometer parameters are as follows; an MS1 survey scan using the orbitrap detector (mass range

(m/z): 400-1500 (using quadrupole isolation), 120000 resolution setting, spray voltage of 2200 V, Ion transfer tube temperature of 275°C, AGC target of 400000, and maximum injection time of 50 ms) was followed by data dependent scans (top speed for most intense ions, with charge state set to only include +2-5 ions, and 5 second exclusion time, while selecting ions with minimal intensities of 50000 at in which the collision event was carried out in the high energy collision cell (HCD Collision Energy of 30%), and the fragment masses where analyzed in the ion trap mass analyzer (With ion trap scan rate of turbo, first mass m/z was 100, AGC Target 5000 and maximum injection time of 35ms). Protein identification and label free quantification was carried out using Peaks Studio 8.5 (Bioinformatics solutions Inc.) Variable modification at lysine residues of +226.08 amu was used in the peptide sequencing parameters.

2.5.8 Mass Spectrometry Analysis

For each sample, biotinylated peptides identified by mass spectrometry were divided into host (*P. aeruginosa*) and phage (ΦPA3) peptides. Host peptides that were identified in all samples were used to normalize phage peptide signals across the dataset. Biotinylated phage protein peak areas were calculated by summing the peak areas of each peptide assigned to a given protein. The fold changes of proteins from ChmA (gp53)-miniTurboID and RecA (gp175)-miniTurboID were calculated by comparing to protein peak areas from GFP-miniTurboID samples. The proteins were sorted according to their average normalized fold changes.

The genome sequence of ΦPA3 (NCBI RefSeq NC_028999.1) is misannotated between the coding regions for gp64 (NCBI accession #YP_009217147.1, nucleotides

51942-53267) and gp68 (NCBI accession #YP_009217148.1, nucleotides 58478-60010). We manually annotated this region to identify gp65-66 (which together code for a single protein, separated by an intron spanning nucleotides 55005-55454)[19] and gp67 for identification by mass spectrometry:

>gp65-66 (NC_028999.1 nucleotides 53811-55004 and 55455-56447)

MYEEHNLRRAVREIHAKLLGHAALDPYYGTTSAARGAMFLSHIGQAPVVEGNEPRRV
MTGMEMRYAEYTFDVRLPTDCTILHKVRKYPTGQGYGAIQHNPVTTLIYENYYDEYKTI
GVLHVPEYMSFHQDFGYELVKNKEVWESLQPDQMFAKDTVIAQSSTVKSNGLYGMG
VNANVAFMSVPGTIEDGFVVSDEFLERMSPRTYTTAVCGAGKKAFFLNMYGDDKIYKP
FPDIGEKIREDGVIFAVRDLDDDLAPAEMTPRALRTLDRTFDRAVIGDPGATVKDIKVY
WDERQNPSFTPSGMDGQLRKYYDALCTYYREIIKIYRGLLARRKDKLRISEEFNQLLVE
AMIYLPQAEGQRKLTRMYRLEQLDEWRVELTYESIKVPGGAYKLTDFHGGKGVVCEV
RPKADMPVDEFGNVVDAIIFGGSTMRRSNYGRIYEHGFGAASRDLAQRLRVEAGLPR
HGVVPEQDLNRVCSNREWVTYAFAELQEFYYIIAPTMHEILREHPSPAEYVKTVLRDGF
SYIYSPVDDPVDLMSSLNCIMNSRFCPNHTRVTYRGQDGKMVTTKDKVLVGPLYMML
LEKIGEDWSAVASVKVQQFGLPSKLNNSDRSSTPGRESAIRSFGESETRSYNCTVGPE
ATVELLDQTNNPRAHLAVINSILTADKPSNIERAVDRTKVPFGSSRPVDLLEHLLECRGL
KFEYATTDGVQPVHTAVPIRAQQKVKSEAIEE*

>gp67 (NC_028999.1 nucleotides 56450-58420)

MNQYNARDLLNMSYDDLFAIPNEWHKIIFDDGEILTKDRATKLSILLWHPLKQFPNATLS
VKYHLGDTRVTSKSLVKLLNSVIWGIHAWSNEQVDPEVLARLAIEAKNVLYNEATSRLG
AYVATLSMFEIAEVYNHPKVREANQNIEPTTHGIETIAYGKIKEAFNDPTQFRGNSIIEGL
RSGTQKMEQLLQAFGPRGFPTDINSDIFAEPCLTGYIDGIWGLYENMIESRSGTKALLY
NKELLRVTEYFNRKSQLIAQYVQRLHKGDCGAGYIEFPVIKAYLKSLRGKFYLNEETGK
REILQGNETHLIGKKIKMRSVLGCVHPDPQGICATCYGTLADNIPRGTNIGQVSAVSMG
DKITSSVLSTKHTDATSAVEQYKITGVEAKYLREGQAPETLYLKKELANKGYRLMIGRN
EAQNLADVLMIDNLSAYPPTSASELTRIGLVRTVDGIDEGDVLTVSLYNRKASLSIELLQ
HVKRVRWELDNRDNIVIDLNGFDFSLPFLTLPYKHVNMYEVMKRIQSFLHSGSDTEGS
KLSSDKVGFTSKTYLKNYNDPIDAVAAFASLVNEKIQLPMPHCEVLVYAMMVRSTQQR
DYRLPKPGISGQFEKYNKLMQSRSLAGAMAFEKQHEPLNNPGSFLYTLRNDHPYDLA
VKGGKLY*

2.5.9 GFP pulldowns

For GFP pulldowns, GFP-Trap Magnetic Agarose beads (Chromotek # gtma-20) were used. Overnight cultures of *P. aeruginosa* expressing GFP-tagged proteins of interest were grown in LB plus 25 µg/mL gentamicin sulfate. Cells were diluted to an

OD_{600} of 0.1, then grown further to an OD_{600} of 0.5. Cultured were diluted 1:10 into 50 mL total volume of culture in 250 mL flasks and grown in LB supplemented with 0.1% arabinose, 25 µg/mL gentamicin sulfate, and 0.2 mM $CaCl_2$. When the cells reached OD_{600}=0.3, cells were infected with ΦPA3 at MOI 3. Cultures were collected at 45 minutes post infection and centrifuged at 4000 rpm at 4°C. Cell pellets were stored at -80°C.

For the GFP pulldown, thawed cell pellets were incubated for 1 hour with 500 µL lysis buffer (10% glycerol, 25 mM Tris (pH 7.5), 150 mM NaCl, 4 mg/mL lysozyme, 20 µg/mL DNase I, 2x cOmplete Protease Inhibitor, 0.4 mM PMSF). Cell suspensions were sonicated for 10 rounds x 20 pulses/round (Duty Cycle 40, Output 4). Lysed cells were centrifuged for 30 minutes at 15,000 rpm at 4ºC. 25 µL of bead slurry (prewashed into dilution buffer: 10 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5 mM EDTA) was used for each sample. 500 µL of cell lysate was added to the beads and rotated end-to-end for 1 hour at 4ºC. Beads were washed 5 times with wash buffer (10 mM Tris/Cl pH 7.5, 150 mM NaCl, 0.05 % NP-40 substitute, 0.5 mM EDTA). For SDS-PAGE, cells were resuspended with 2x SDS buffer (120 mM Tris-HCl pH 6.8, 20% glycerol, 4% SDS, 0.04% bromophenol blue, 10% β-mercaptoethanol) and boiled at 100ºC for 5 minutes. 10 µL samples of each elution were run on two separate SDS-PAGE gels and visualized by either silver staining or Coomassie blue staining. For tryptic mass spectrometry of gel bands, bands were cut out of Coomassie blue-stained gels. The remaining 80% of each elution was used for mass spectrometry identification of proteins as described above.

2.5.10 Protein purification

Codon-optimized sequences encoding full-length phage gp2 homologs were synthesized (Invitrogen/GeneArt) and cloned into UC Berkeley Macrolab vector 2-BT (Addgene #29666) to generate constructs with N-terminal TEV protease-cleavable His_6-tags. Proteins were expressed in *E. coli* strain Rosetta 2 (DE3) pLysS (EMD Millipore). Cultures were grown at 37°C to OD_{600}=0.7, then induced with 0.25 mM IPTG and shifted to 20°C for 16 hours. Cells were harvested by centrifugation and resuspended in buffer A (25 mM Tris pH 7.5, 10% glycerol, and 1 mM NaN_3) plus 300 mM NaCl, 5 mM imidazole, 5 mM β-mercaptoethanol. Proteins were purified by Ni^{2+}-affinity (Ni-NTA agarose, Qiagen) then passed over an anion-exchange column (Hitrap Q HP, Cytiva) in Buffer A plus 100 mM NaCl and 5 mM β-mercaptoethanol, collecting flow-through fractions. Tags were cleaved with TEV protease[35], and cleaved protein was passed over another Ni^{2+} column (collecting flow-through fractions) to remove uncleaved protein, cleaved tags, and tagged TEV protease. The protein was passed over a size exclusion column (Superdex 200, Cytiva) in buffer GF (buffer A plus 300 mM NaCl and 1 mM dithiothreitol (DTT)), then concentrated by ultrafiltration (Amicon Ultra, EMD Millipore) to 10 mg/ml and stored at 4°C.

For characterization of oligomeric state by size exclusion chromatography coupled to multi-angle light scattering (SEC-MALS), 100 μL of purified proteins at 5 mg/mL were injected onto a Superdex 200 Increase 10/300 GL column (Cytiva) in buffer GF. Light scattering and refractive index profiles were collected by miniDAWN TREOS and Optilab

T-rEX detectors (Wyatt Technology), respectively, and molecular weight was calculated using ASTRA version 8 software (Wyatt Technology).

2.5.11 Crystallization and structure determination

Purified PA1C gp2 in a buffer containing 20 mM Tris pH 8.5, 1 mM DTT, and 100 mM NaCl (14 mg/mL) was mixed 1:1 with well solution containing 0.1 M Tris pH 8.5, and 1.5 M Lithium sulfate in hanging drop format. Crystals were cryoprotected by the addition of 24% glycerol and flash-frozen in liquid nitrogen. Diffraction data were collected at the Advanced Photon Source NE-CAT beamline 24ID-C (see support statement below) and processed with the RAPD data-processing pipeline (https://github.com/RAPD/RAPD), which uses XDS[36] for data indexing and reduction, AIMLESS[37] for scaling, and TRUNCATE[38] for conversion to structure factors. We determined the structure by molecular replacement in PHASER[39], using a predicted structure from AlphaFold2[40] as a search model. We manually rebuilt the initial model in COOT[41] , and refined in phenix.refine[42] using positional and individual B-factor refinement (**Extended Data Table 5**).

2.6 APS Support Statement

This work is based upon research conducted at the Northeastern Collaborative Access Team beamlines, which are funded by the National Institute of General Medical Sciences from the National Institutes of Health (P30 GM124165). This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357.

2.7 Data Availability

Final refined coordinates and reduced diffraction data for the structure of PA1C gp2 is available at the RCSB Protein Data Bank (www.rcsb.org) under accession ID 7UYX. Raw diffraction data for the structure of PA1C gp2 is available at the SBGrid Data Bank (data.sbgrid.org) under accession ID 908.

2.8 Tables

Table 2.1 Top 25 identified proteins from ΦPA3 ChmA (gp53) miniTurboID

Protein	NCBI Accession number	Times detected from 3 trials	log_2(fold change)[1]	Normalized peak area[2]	Annotation	Found in RecA miniTurbo ID?
gp53 (ChmA)	YP_009217136.1	3	4.00	3.25E-02	Major nuclear shell protein	
gp52	YP_009217135.1	3	3.27	2.09E-02	hypothetical protein	
gp355	YP_009217434.1	3	3.79	6.49E-03	hypothetical protein	
gp164	YP_009217199.1	3	3.93	1.04E-03	tail protein	
gp119	YP_009217199.1	3	∞	8.81E-04	hypothetical protein	Yes
gp166	YP_009217245.1	2	∞	8.75E-04	tail protein	Yes
gp358	YP_009217437.1	3	8.51	6.90E-04	hypothetical protein	Yes
gp2	YP_009217084.1	2	∞	5.87E-04	hypothetical protein	
gp134	YP_009217213.1	3	∞	5.00E-04	putative helicase	Yes
gp131	YP_009217211.1	2	∞	3.45E-04	putative helicase	Yes
gp63	YP_009217146.1	2	∞	2.85E-04	hypothetical protein	Yes
gp62	YP_009217145.1	2	∞	2.51E-04	non-virion RNAP subunit	Yes
gp222	YP_009217301.1	2	∞	2.13E-04	hypothetical protein	Yes
gp106	YP_009217186.1	2	6.10	1.57E-04	virion structural protein	
gp18	YP_009217100.1	1	∞	1.57E-04	hypothetical protein	Yes
gp202	YP_009217281.1	1	∞	1.55E-04	hypothetical protein	
gp148	YP_009217227.1	1	∞	1.20E-04	portal protein	
gp14	YP_009217096.1	1	4.95	1.08E-04	hypothetical protein	
gp12	YP_009217094.1	1	∞	9.40E-05	tail protein	Yes
gp378	YP_009217457.1	2	∞	8.48E-05	NrdA	
gp247	YP_009217326.1	1	∞	4.45E-05	hypothetical protein	Yes
gp257	YP_009217336.1	2	3.21	3.90E-05	DNA ligase	Yes
gp219	YP_009217298.1	1	∞	3.87E-05	hypothetical protein	
gp335	YP_009217414.1	1	∞	3.83E-05	hypothetical protein	Yes
gp85	YP_009217165.1	1	∞	3.67E-05	hypothetical protein	
gp370	YP_009217449.1	1	∞	2.79E-05	hypothetical protein	

[1]Fold change calculated as the fold-change in average normalized peak area when comparing three trials with three negative-control trials

[2]Normalized peak area calculated as the fraction of the total peak area for ΦPA3 proteins (per dataset) assigned to a given protein

Table 2.2. Top 25 identified proteins from ΦPA3 RecA (gp175) miniTurboID

Protein	NCBI Accession number	Times detected from 3 trials	log_2(fold change)[1]	Normalized peak area[2]	Annotation	Found in ChmA miniTurbo ID?
gp175	YP_009217254.1	3	6.99	1.09E-02	RecA	
gp210	YP_009217289.1	3	8.38	2.96E-03	putative endonuclease	
gp253	YP_009217332.1	1	7.47	1.59E-03	hypothetical protein	
gp222	YP_009217301.1	1	∞	1.46E-03	hypothetical protein	Yes
gp313	YP_009217392.1	2	3.31	7.11E-04	hypothetical protein	
gp49	YP_009217132.1	1	3.73	6.16E-04	hypothetical protein	
gp358	YP_009217437.1	1	8.02	4.91E-04	hypothetical protein	Yes
gp166	YP_009217245.1	2	∞	2.12E-04	tail protein	Yes
gp78	YP_009217158.1	2	∞	1.71E-04	endonuclease	Yes
gp271	YP_009217350.1	1	5.18	1.41E-04	hypothetical protein	Yes
gp12	YP_009217094.1	1	∞	1.27E-04	tail protein	Yes
gp200	YP_009217279.1	1	∞	1.19E-04	hypothetical protein	Yes
gp233	YP_009217312.1	1	6.33	9.00E-05	putative SNF2 domain helicase	Yes
gp131	YP_009217211.1	1	∞	7.80E-05	putative helicase	Yes
gp308	YP_009217387.1	1	∞	7.22E-05	hypothetical protein	Yes
gp64	YP_009217147.1	1	∞	7.05E-05	hypothetical protein	Yes
gp257	YP_009217336.1	1	4.06	7.04E-05	DNA ligase	Yes
gp144	YP_009217223.1	1	∞	6.75E-05	hypothetical protein	
gp350	YP_009217429.1	1	∞	5.58E-05	hypothetical protein	
gp239	YP_009217318.1	1	∞	4.77E-05	hypothetical protein	Yes
gp108	YP_009217188.1	1	∞	4.48E-05	hypothetical protein	
gp217	YP_009217296.1	1	∞	3.10E-05	hypothetical protein	Yes
gp61	YP_009217144.1	1	∞	2.42E-05	hypothetical protein	Yes
gp247	YP_009217326.1	1	∞	1.76E-05	hypothetical protein	Yes
gp375	YP_009217454.1	1	∞	1.70E-05	hypothetical protein	
gp62	YP_009217145.1	1	∞	1.69E-05	non-viral RNA polymerase subunit	Yes

[1]Fold change calculated as the fold-change in average normalized peak area when comparing three trials with three negative-control trials

[2]Normalized peak area calculated as the fraction of the total peak area for ΦPA3 proteins (per dataset) assigned to a given protein

Table 2.3. Localization of miniTurboID hits in infected cells

Protein	NCBI Accession Number	Annotation	Detected with ChmA or RecA?	Localization in infected cells
gp2	YP_009217084.1	hypothetical protein	ChmA	shell
gp12	YP_009217094.1	tail protein	both	phage bouquet
gp14	YP_009217096.1	hypothetical protein	ChmA	cytoplasmic puncta
gp18	YP_009217100.1	hypothetical protein	both	nuclear
gp49	YP_009217132.1	hypothetical protein	RecA	diffuse cytoplasmic
gp52	YP_009217135.1	hypothetical protein	ChmA	cytoplasmic puncta
gp53	YP_009217136.1	Major nuclear shell protein (ChmA)	ChmA	shell
gp61	YP_009217144.1	hypothetical protein	both	shell
gp62	YP_009217145.1	non-virion RNA polymerase	both	nuclear
gp63	YP_009217146.1	hypothetical protein	both	shell
gp64	YP_009217147.1	hypothetical protein	both	shell
gp78	YP_009217158.1	endonuclease	both	nuclear
gp85	YP_009217165.1	hypothetical protein	ChmA	diffuse cytoplasmic
gp106	YP_009217186.1	virion structural protein	ChmA	cytoplasmic puncta
gp108	YP_009217188.1	hypothetical protein	RecA	nuclear
gp119	YP_009217199.1	hypothetical protein	both	cytoplasmic puncta
gp131	YP_009217211.1	putative helicase	both	cytoplasmic puncta
gp134	YP_009217213.1	putative helicase	both	nuclear
gp144	YP_009217223.1	hypothetical protein	RecA	cytoplasmic
gp148	YP_009217227.1	portal protein	ChmA	shell
gp164	YP_009217243.1	tail protein	ChmA	phage bouquet
gp166	YP_009217245.1	tail protein	both	phage bouquet
gp175	YP_009217254.1	putative UvsX protein (RecA)	RecA	nuclear
gp200	YP_009217279.1	hypothetical protein	both	nuclear
gp202	YP_009217281.1	hypothetical protein	ChmA	diffuse cytoplasmic
gp210	YP_009217289.1	putative endonuclease	RecA	nuclear
gp217	YP_009217296.1	hypothetical protein	both	cytoplasmic puncta
gp219	YP_009217298.1	hypothetical protein	ChmA	cytoplasmic puncta
gp222	YP_009217301.1	hypothetical protein	both	diffuse cytoplasmic
gp233	YP_009217312.1	putative SNF2 domain helicase	both	phage bouquet
gp239	YP_009217318.1	hypothetical protein	both	diffuse cytoplasmic
gp247	YP_009217326.1	hypothetical protein	both	diffuse cytoplasmic
gp253	YP_009217332.1	hypothetical protein	RecA	cytoplasmic puncta
gp257	YP_009217336.1	DNA ligase	both	nuclear
gp271	YP_009217350.1	hypothetical protein	RecA	diffuse cytoplasmic
gp308	YP_009217387.1	hypothetical protein	both	cytoplasmic puncta
gp313	YP_009217392.1	hypothetical protein	RecA	diffuse cytoplasmic
gp335	YP_009217414.1	hypothetical protein	both	diffuse cytoplasmic
gp350	YP_009217429.1	hypothetical protein	RecA	diffuse cytoplasmic
gp355	YP_009217434.1	hypothetical protein	ChmA	phage bouquet/cytoplasmic puncta
gp358	YP_009217437.1	hypothetical protein	both	cytoplasmic puncta
gp370	YP_009217449.1	hypothetical protein	ChmA	diffuse cytoplasmic
gp375	YP_009217454.1	hypothetical protein	RecA	shell
gp378	YP_009217457.1	NrdA	ChmA	diffuse cytoplasmic

2.9 Figures

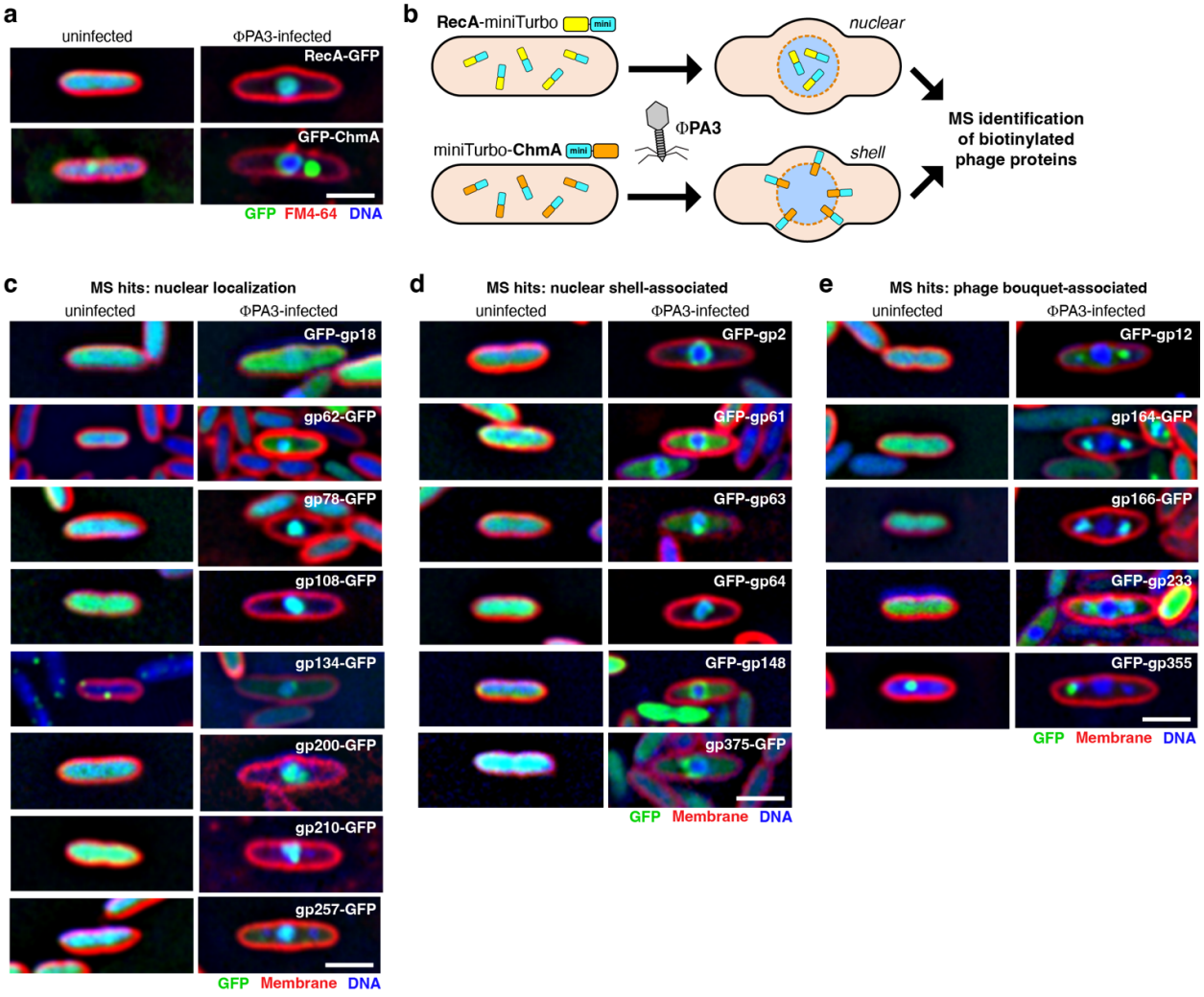


Figure 2.1. Identification of jumbo phage nuclear shell-associated proteins
(a) Subcellular localization of GFP-tagged ΦPA3 RecA (gp175) and ChmA (gp53) in uninfected (left) and ΦPA3-infected (right) *P. aeruginosa* cells. GFP is shown in green, FM4-64 (to visualize membranes) in red, and DAPI (to visualize nucleic acids) in blue. Scale bar = 2 μm. (b) Experimental schematic for identification of jumbo phage nuclear or nuclear-shell-associated genes, by proximity labeling with miniTurboID-fused RecA (C-terminal miniTurboID) or ChmA (N-terminal miniTurboID) in ΦPA3-infected *P. aeruginosa* cells. See **Extended Data Fig. 1a** for fusion construct design, and **Extended Data Fig. 1b** for localization of miniTurboID-fused proteins. See **Tables 1-2** for top 25 identified proteins, **Extended Data Tables 1-2** for full protein lists, and **Extended Data Fig. 1c-d** for diagrams showing overlap between independent mass spectrometry datasets. (c-e) Subcellular localization of selected proteins identified by proximity labeling, with panel (c) showing nuclear-localized proteins, (d) showing nuclear shell-associated proteins, and (e) showing phage bouquet-associated proteins. See **Extended Data Fig. 2** for further data, and **Table 3** for a collated list of localizations. Scale bar = 2 μm.

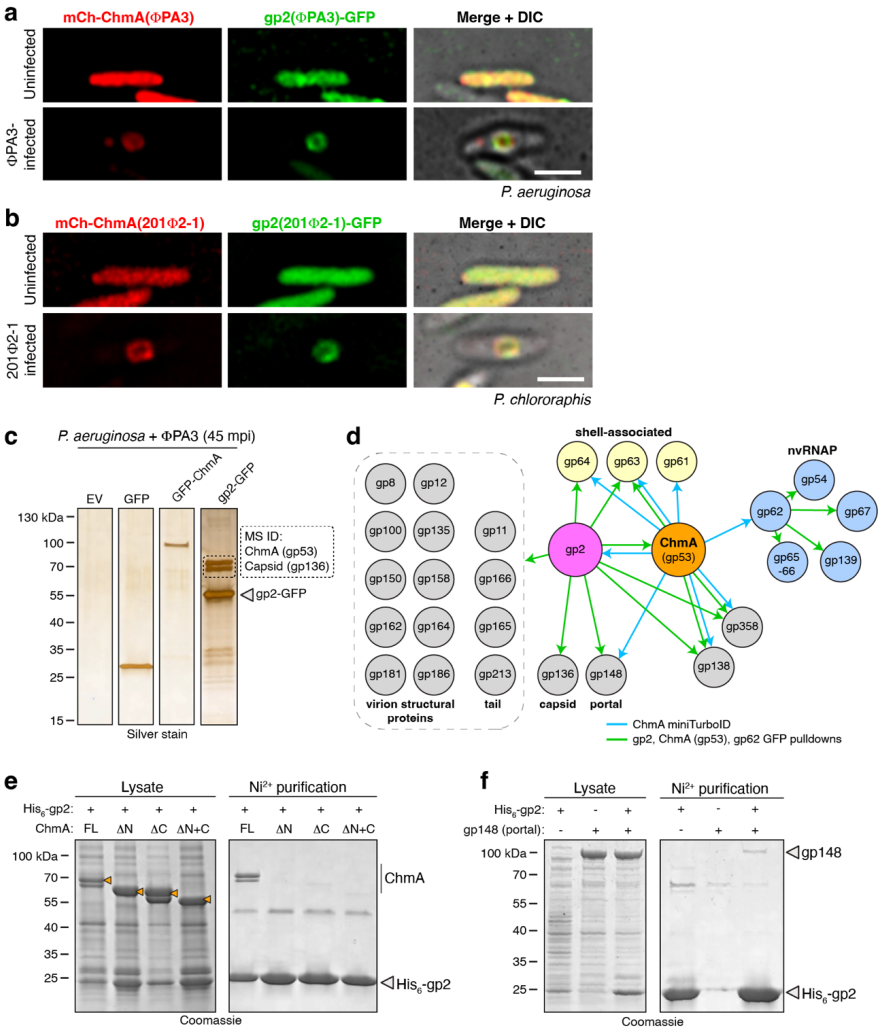


Figure 2.2. gp2 is an interaction hub in the jumbo phage nuclear shell
(a) Colocalization of mCherry-fused ΦPA3 ChmA (red) and GFP-fused ΦPA3 gp2 (green) in *P. aeruginosa* cells. Scale bar = 2 μm. (b) Colocalization of mCherry-fused 201Φ2-1 ChmA (gp105; red) and GFP-fused 201Φ2-1 gp2 (green) in *P. chlororaphis* cells. Scale bar = 2 μm. (c) Silver stain SDS-PAGE analysis of GFP pulldown experiments. EV: empty vector. Dotted box indicated the gel slice that was cut out (of the same bands in a Coomassie blue-stained gel) for tryptic mass spectrometry protein identification (see **Extended Data Table 3**). (d) Interaction network of the jumbo phage nuclear shell, with blue arrows indicating interactions identified by ChmA miniTurboID and green arrows indicating interactions identified in GFP pulldowns (see **Extended Data Fig. 4a** for SDS-PAGE gels of all analyzed GFP pulldown samples, and **Extended Data Table 4** for full data). nvRNAP: non-virion RNA polymerase. (e) Ni^{2+} pulldown analysis of *E. coli*-coexpressed 201Φ2-1 gp2 (His_6-tagged) and ChmA (full-length or truncated: ΔN missing residues 1-63, ΔC missing residues 583-631, and ΔN+C missing residues 1-63 and 583-631). Doublet bands for ChmA arise from a methionine codon at position 33 of the annotated gene. Orange marks show the presence of ChmA in the lysates. See **Extended Data Fig. 4b** for control pulldown. (f) Ni^{2+} pulldown analysis of *E. coli*-coexpressed ΦPA3 gp2 (His_6-tagged) and portal (gp148).

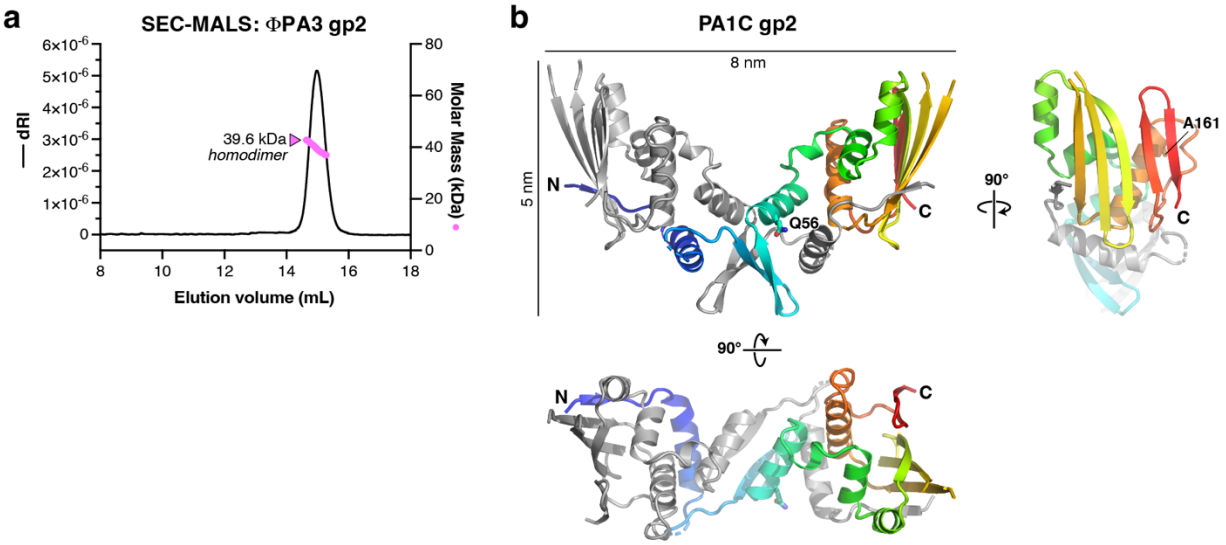


Figure 2.3. Structure of gp2
(a) Size exclusion chromatography coupled to multi-angle light scattering (SEC-MALS) of ΦPA3 gp2, showing that it is homodimeric in solution (monomer molecular weight = 22.5 kDa). See **Extended Data Figure 5a-c** for SEC-MALS analysis of other jumbo phage gp2 proteins. (b) Structure of the PA1C gp2 homodimer, with one protomer colored gray and the other colored as a rainbow from N-terminus (blue) to C-terminus (red).

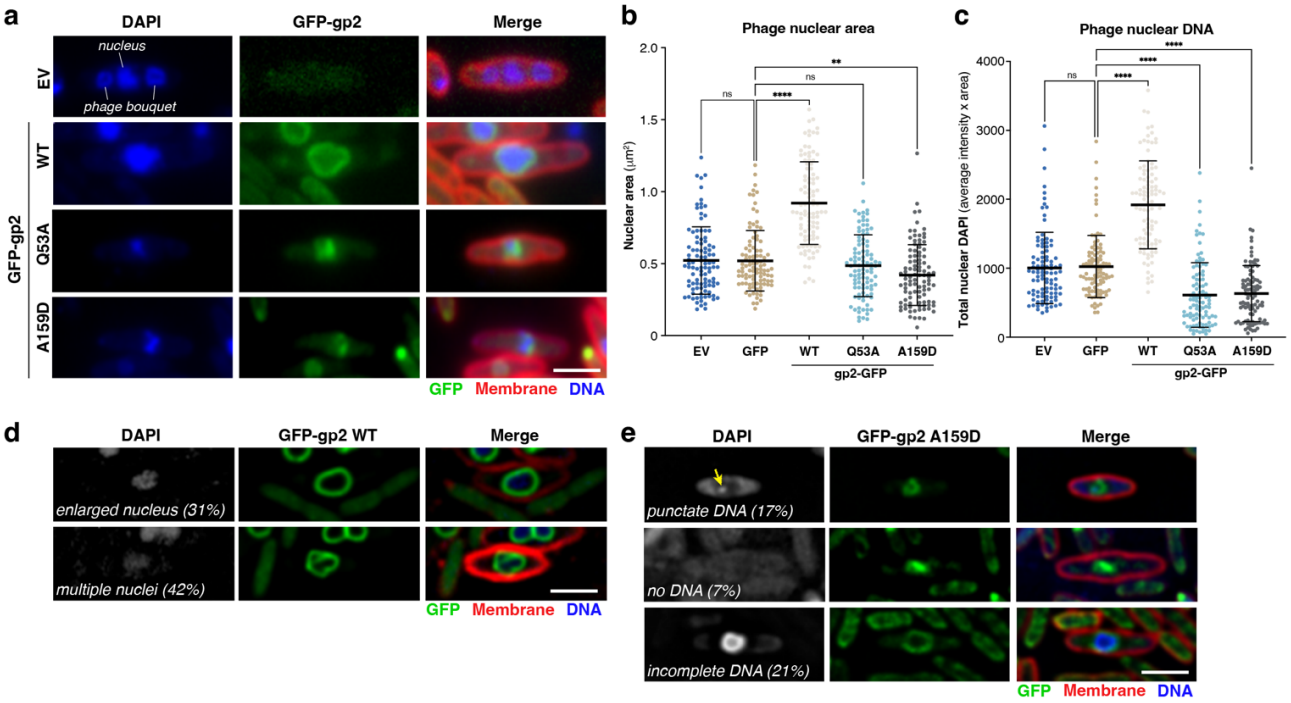


Figure 2.4. gp2 mutations cause defects in phage nucleus formation and morphology

(a) Fluorescence imaging of ΦPA3-infected *P. aeruginosa* cells expressing no additional proteins (EV: empty vector), GFP-tagged wild-type gp2, or GFP-tagged gp2 point mutants. Un-deconvolved images that were used for DAPI quantitation (panels b and c) are shown. GFP is shown in green, FM4-64 (to visualize membranes) in red, and DAPI (to visualize nucleic acids) in blue. Scale bar = 2 μm. (b) Phage nuclear area of ΦPA3-infected *P. aeruginosa* cells expressing no additional proteins (EV: empty vector), GFP-tagged wild-type gp2, or GFP-tagged gp2 point mutants. n=100 for all samples; error bars represent mean +/- standard deviation. P-values: ns: p>0.05 (not significant); **:p<0.01; ****:p<0.0001. (c) Total nuclear DNA in ΦPA3-infected *P. aeruginosa* cells expressing no additional proteins (EV: empty vector), GFP-tagged wild-type gp2, or GFP-tagged gp2 point mutants, calculated by multiplying each cell's average DAPI signal within the nucleus by that cell's nuclear area (panel (b)). (d) Visual phenotypes observed in ΦPA3-infected *P. aeruginosa* cells expressing GFP-tagged wild-type gp2 (n=100 cells). See **Extended Data Fig. 6c** for additional examples. (e) Visual phenotypes observed in ΦPA3-infected *P. aeruginosa* cells expressing GFP-tagged gp2 A159D (n=100 cells). See **Extended Data Fig. 6d** for additional examples, and **Extended Data Fig. 6e** for examples of similar phenotypes from gp2 Q53A.

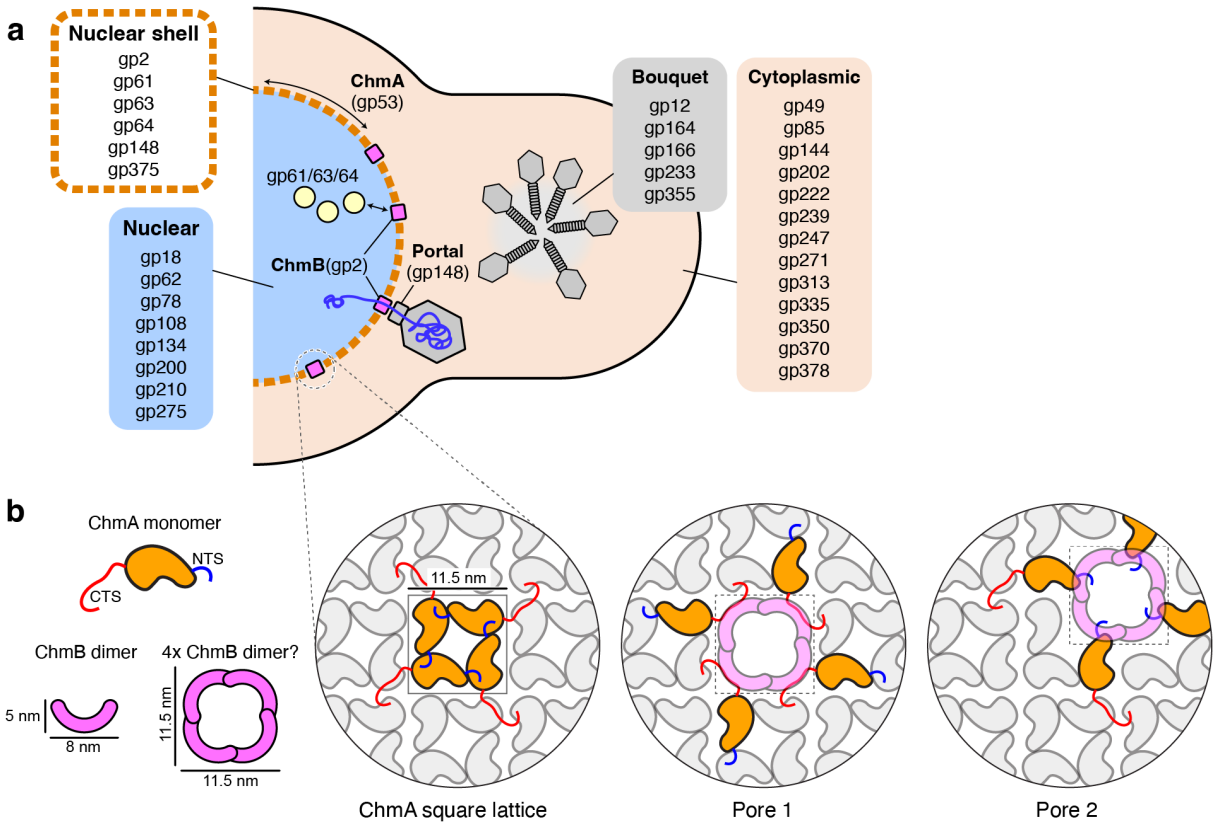


Figure 2.5. Model for jumbo phage protein localization and nuclear shell architecture and function.

(a) Schematic of a ΦPA3-infected *P. aeruginosa* cell with assembled phage nucleus (blue) bounded by a ChmA lattice (orange). Proteins that we find localize to the nucleus, nuclear shell, phage bouquet, or cytoplasm are listed. ChmB (pink) is shown integrated into the ChmA lattice, where it may mediate the docking of phage capsids by binding the portal protein, for genomic packaging. Further interactions with gp61/gp63/gp64 (light yellow) or other shell-associated proteins could accommodate mRNA export or specific protein import. (b) Schematic of the ChmA lattice derived from cryoET analysis of intact 201Φ2-1 and Goslar nuclear shells. ChmA is shown in orange with N-terminal and C-terminal tails shown blue and red, respectively. Removal of four contiguous ChmA protomers from the lattice would leave a cavity ~11.5 x 11.5 nm (two possibilities shown), which could be filled by an assembly of ChmB to generate a pore.

2.10 Extended Data Tables

Extended Table 2.1. Phage proteins enriched >3-fold in ΦPA3 ChmA (gp53) miniTurboID

Gene	Accession	log2(FC)	- LOG10(P)	Average Peak	Annotation
gp2	YP_009217084.1	∞	0.580797161	0.000587274	hypothetical protein
gp9	YP_009217091.1	∞	0.427243412	3.36734E-06	virion structural protein
gp12	YP_009217094.1	∞	0.446180573	9.39527E-05	virion structure
gp18	YP_009217100.1	∞	0.427243412	0.000156877	hypothetical protein
gp19	YP_009217101.1	∞	0.427243412	7.17898E-06	hypothetical protein
gp31	YP_009217114.1	∞	0.427243412	1.32121E-05	hypothetical protein
gp36	YP_009217119.1	∞	0.427243412	2.63775E-05	hypothetical protein
gp37	YP_009217120.1	∞	0.427243412	1.20429E-06	hypothetical protein
gp39	YP_009217122.1	∞	0.427243412	7.76359E-06	hypothetical protein
gp44	YP_009217127.1	∞	0.427243412	3.06334E-06	hypothetical protein
gp47	YP_009217130.1	∞	0.427243412	1.25106E-05	hypothetical protein
gp54	YP_009217137.1	∞	0.427243412	5.6707E-07	hypothetical protein
gp59	YP_009217142.1	∞	0.427243412	1.0254E-05	nuclease SbcCD, D subunit
gp61	YP_009217144.1	∞	0.427243412	2.0487E-05	hypothetical protein
gp62	YP_009217145.1	∞	0.475501761	0.000250661	hypothetical protein
gp63	YP_009217146.1	∞	0.559175677	0.000285223	hypothetical protein
gp64	YP_009217147.1	∞	0.427243412	4.7704E-06	hypothetical protein
gp75	YP_009217155.1	∞	0.427243412	2.49043E-05	hypothetical protein
gp77	YP_009217157.1	∞	0.427243412	8.26635E-06	putative RNA polymerase beta subunit
gp78	YP_009217158.1	∞	0.427243412	3.25042E-07	endonuclease
gp85	YP_009217165.1	∞	0.427243412	3.66993E-05	hypothetical protein
gp90	YP_009217170.1	∞	0.427243412	2.7969E-06	virion structural protein
gp93	YP_009217173.1	∞	0.427243412	1.77721E-06	virion structural protein
gp94	YP_009217174.1	∞	0.427243412	1.96428E-05	virion structural protein
gp103	YP_009217183.1	∞	0.427243412	5.30824E-06	hypothetical protein
gp119	YP_009217199.1	∞	0.599563001	0.000881381	hypothetical protein
gp122	YP_009217202.1	∞	0.427243412	7.29591E-06	hypothetical protein
gp124	YP_009217204.1	∞	0.427243412	2.25659E-06	hypothetical protein
gp131	YP_009217211.1	∞	0.509791975	0.000344554	putative helicase
gp134	YP_009217213.1	∞	1.301334868	0.000500235	putative helicase
gp137	YP_009217216.1	∞	0.427243412	3.66631E-06	hypothetical protein
gp138	YP_009217217.1	∞	0.427243412	1.01792E-05	hypothetical protein
gp147	YP_009217226.1	∞	0.427243412	4.48979E-06	virion structural protein
gp148	YP_009217227.1	∞	0.427243412	0.000119552	virion structural protein

Extended Table 2.1. Phage proteins enriched >3-fold in ΦPA3 ChmA (gp53) miniTurboID (continued)

gp150	YP_009217229.1	∞	0.427243412	1.09789E-06	virion structural protein
gp153	YP_009217232.1	∞	0.427243412	9.71618E-07	virion structural protein
gp158	YP_009217237.1	∞	0.427243412	6.91006E-07	virion structural protein
gp165	YP_009217244.1	∞	0.427243412	9.82141E-07	tail protein
gp166	YP_009217245.1	∞	0.65949239	0.00087454	tail protein
gp168	YP_009217247.1	∞	0.427243412	2.53168E-07	tail fibre protein
gp170	YP_009217249.1	∞	0.427243412	1.08269E-05	hypothetical protein
gp189	YP_009217268.1	∞	0.427243412	1.22171E-05	hypothetical protein
gp198	YP_009217277.1	∞	0.427243412	1.01488E-05	hypothetical protein
gp200	YP_009217279.1	∞	0.427243412	1.07217E-05	hypothetical protein
gp202	YP_009217281.1	∞	0.427243412	0.000154944	hypothetical protein
gp209	YP_009217288.1	∞	0.427243412	2.46704E-07	hypothetical protein
gp217	YP_009217296.1	∞	0.427243412	1.05814E-06	hypothetical protein
gp218	YP_009217297.1	∞	0.427243412	1.17284E-05	hypothetical protein
gp219	YP_009217298.1	∞	0.427243412	3.87127E-05	hypothetical protein
gp222	YP_009217301.1	∞	0.692394617	0.000213331	hypothetical protein
gp225	YP_009217304.1	∞	0.427243412	1.44983E-06	hypothetical protein
gp239	YP_009217318.1	∞	0.427243412	2.27997E-06	hypothetical protein
gp247	YP_009217326.1	∞	0.427243412	4.44653E-05	hypothetical protein
gp251	YP_009217330.1	∞	0.427243412	2.20046E-05	hypothetical protein
gp254	YP_009217333.1	∞	0.427243412	2.71258E-05	hypothetical protein
gp267	YP_009217346.1	∞	0.427243412	1.62521E-06	hypothetical protein
gp268	YP_009217347.1	∞	0.427243412	1.75382E-06	hypothetical protein
gp275	YP_009217354.1	∞	0.427243412	5.96066E-06	hypothetical protein
gp276	YP_009217355.1	∞	0.427243412	2.1233E-05	hypothetical protein
gp278	YP_009217357.1	∞	0.427243412	4.89901E-06	hypothetical protein
gp280	YP_009217359.1	∞	0.427243412	8.74573E-06	hypothetical protein
gp281	YP_009217360.1	∞	0.427243412	8.26635E-06	hypothetical protein
gp298	YP_009217377.1	∞	0.427243412	1.62521E-05	hypothetical protein
gp299	YP_009217378.1	∞	0.427243412	3.91687E-07	hypothetical protein
gp300	YP_009217379.1	∞	0.427243412	1.15518E-05	hypothetical protein
gp304	YP_009217383.1	∞	0.427243412	1.42644E-05	hypothetical protein
gp305	YP_009217384.1	∞	0.427243412	5.64731E-06	hypothetical protein
gp308	YP_009217387.1	∞	0.427243412	2.17943E-05	hypothetical protein
gp320	YP_009217399.1	∞	0.427243412	1.3446E-06	hypothetical protein
gp321	YP_009217400.1	∞	0.427243412	8.92111E-07	hypothetical protein
gp323	YP_009217402.1	∞	0.427243412	2.94759E-06	hypothetical protein

Extended Table 2.1. Phage proteins enriched >3-fold in ΦPA3 ChmA (gp53) miniTurboID (continued)

gp334	YP_009217413.1	∞	0.427243412	7.12052E-06	hypothetical protein
gp335	YP_009217414.1	∞	0.427243412	3.83316E-05	hypothetical protein
gp341	YP_009217420.1	∞	0.427243412	3.17559E-06	hypothetical protein
gp343	YP_009217422.1	∞	0.427243412	1.0406E-05	hypothetical protein
gp344	YP_009217423.1	∞	0.427243412	1.25808E-05	hypothetical protein
gp345	YP_009217424.1	∞	0.427243412	1.65327E-05	hypothetical protein
gp347	YP_009217426.1	∞	0.427243412	4.75871E-07	hypothetical protein
gp351	YP_009217430.1	∞	0.427243412	2.93473E-06	hypothetical protein
gp356	YP_009217435.1	∞	0.427243412	1.32121E-05	hypothetical protein
gp360	YP_009217439.1	∞	0.427243412	1.38669E-05	hypothetical protein
gp367	YP_009217446.1	∞	0.427243412	1.00669E-05	hypothetical protein
gp370	YP_009217449.1	∞	0.427243412	2.78858E-05	hypothetical protein
gp376	YP_009217455.1	∞	0.427243412	7.14391E-07	hypothetical protein
gp378	YP_009217457.1	∞	0.61182646	8.48101E-05	NrdA [Pseudomonas phage PhiPA3]
gp358	YP_009217437.1	8.509889478	0.489259407	0.000689784	hypothetical protein
gp106	YP_009217186.1	6.095050782	0.546215811	0.000157237	virion structural protein
gp14	YP_009217096.1	4.94982771	0.410897208	0.000108036	hypothetical protein
gp233	YP_009217312.1	4.326278772	0.401898111	2.23893E-05	putative SNF2 domain helicase
gp53	YP_009217136.1	3.998973103	2.986709393	0.032450428	shell
gp164	YP_009217243.1	3.927108325	1.224285589	0.001044067	virion structural protein
gp355	YP_009217434.1	3.786417991	0.831207171	0.006488366	hypothetical protein
gp52	YP_009217135.1	3.274960008	1.981598946	0.020900228	hypothetical protein
gp257	YP_009217336.1	3.20807031	0.442878384	3.90047E-05	DNA ligase

Extended Table 2.2. Phage proteins enriched >3-fold in ΦPA3 RecA (gp175) miniTurboID

Gene	Accession	log2(FC)	- LOG10(P)	Average Peak	Annotation
gp12	YP_009217094.1	∞	0.427243412	0.000127007	virion structure
gp18	YP_009217100.1	∞	0.427243412	3.55654E-06	hypothetical protein
gp19	YP_009217101.1	∞	0.427243412	3.48306E-06	hypothetical protein
gp43	YP_009217126.1	∞	0.427243412	2.70415E-07	hypothetical protein
gp44	YP_009217127.1	∞	0.427243412	7.81851E-06	hypothetical protein
gp47	YP_009217130.1	∞	0.427243412	1.09342E-05	hypothetical protein
gp55	YP_009217138.1	∞	0.427243412	1.50786E-05	hypothetical protein
gp61	YP_009217144.1	∞	0.427243412	2.41904E-05	hypothetical protein
gp62	YP_009217145.1	∞	0.427243412	1.69009E-05	hypothetical protein
gp63	YP_009217146.1	∞	0.427243412	4.43539E-06	hypothetical protein
gp64	YP_009217147.1	∞	0.427243412	7.04695E-05	hypothetical protein
gp78	YP_009217158.1	∞	0.434415134	0.000170803	endonuclease
gp108	YP_009217188.1	∞	0.427243412	4.48242E-05	hypothetical protein
gp119	YP_009217199.1	∞	0.427243412	3.49776E-06	hypothetical protein
gp122	YP_009217202.1	∞	0.427243412	1.63425E-05	hypothetical protein
gp131	YP_009217211.1	∞	0.427243412	7.80073E-05	putative helicase
gp134	YP_009217213.1	∞	0.427243412	2.26325E-06	putative helicase
gp144	YP_009217223.1	∞	0.427243412	6.74861E-05	hypothetical protein
gp166	YP_009217245.1	∞	0.838237102	0.000211819	tail protein
gp198	YP_009217277.1	∞	0.427243412	3.89897E-06	hypothetical protein
gp200	YP_009217279.1	∞	0.427243412	0.000119041	hypothetical protein
gp217	YP_009217296.1	∞	0.427243412	3.10433E-05	hypothetical protein
gp222	YP_009217301.1	∞	0.427243412	0.001462444	hypothetical protein
gp239	YP_009217318.1	∞	0.427243412	4.77488E-05	hypothetical protein
gp247	YP_009217326.1	∞	0.427243412	1.76357E-05	hypothetical protein
gp276	YP_009217355.1	∞	0.427243412	4.9821E-06	hypothetical protein
gp298	YP_009217377.1	∞	0.427243412	6.76037E-06	hypothetical protein
gp304	YP_009217383.1	∞	0.427243412	1.83706E-06	hypothetical protein
gp305	YP_009217384.1	∞	0.427243412	6.2019E-07	hypothetical protein
gp308	YP_009217387.1	∞	0.427243412	7.22184E-05	hypothetical protein
gp323	YP_009217402.1	∞	0.427243412	1.69009E-06	hypothetical protein
gp335	YP_009217414.1	∞	0.427243412	5.20255E-07	hypothetical protein
gp344	YP_009217423.1	∞	0.427243412	1.48434E-06	hypothetical protein
gp350	YP_009217429.1	∞	0.427243412	5.5773E-05	hypothetical protein

Extended Table 2.2. Phage proteins enriched >3-fold in ΦPA3 RecA (gp175) miniTurboID (continued)

gp367	YP_009217446.1	∞	0.427243412	6.93673E-07	hypothetical protein
gp369	YP_009217448.1	∞	0.427243412	1.05227E-06	hypothetical protein
gp375	YP_009217454.1	∞	0.427243412	1.70479E-05	hypothetical protein
gp210	YP_009217289.1	8.378497561	0.762137902	0.002960149	putative endonuclease
gp358	YP_009217437.1	8.020010698	0.425319178	0.000491185	hypothetical protein
gp253	YP_009217332.1	7.467502947	0.424419141	0.001594316	hypothetical protein
gp175	YP_009217254.1	6.992819348	1.133408442	0.010889519	putative UvsX protein
gp233	YP_009217312.1	6.334117859	0.42103046	9.00452E-05	putative SNF2 domain helicase
gp271	YP_009217350.1	5.175341463	0.413287617	0.000140645	hypothetical protein
gp257	YP_009217336.1	4.059375626	0.39664078	7.03696E-05	DNA ligase
gp49	YP_009217132.1	3.726835891	0.388512353	0.000616487	hypothetical protein
gp313	YP_009217392.1	3.309539196	0.544752432	0.000711422	hypothetical protein

Extended Table 2.3. Mass spectrometry of ChmB-GFP 70 kDa bands

Protein	-10lgP	Coverage (%)	Peak Area	# Peptides	#Spectra	MW (Da)	Annotation	Accession Number
gp53	450.71	81	7780000000	149	715	66612	ChmA	YP_009217136.1
gp136	328.73	40	418000000	41	139	82847	Major capsid protein	YP_009217215.1
gp166	232.86	21	22600000	21	35	115675	tail protein	YP_009217196.1
gp11	215.69	26	8490000	19	27	77793	tail sheath protein	YP_009217093.1
gp2	128.52	16	5160000	2	7	22501	hypothetical protein	YP_009217084.1
gp8	110	7	727000	3	4	61839	virion structural protein	YP_009217090.1
gp208	59.21	3	210000	2	3	66318	hypothetical protein	YP_009217287.1
gp113	23.78	13	173000	1	1	14283	hypothetical protein	YP_009217193.1
gp214	21.51	1	57400	1	1	72021	hypothetical protein	YP_009217293.1
gp190	21.12	1	0	1	1	92001	SbcC protein	YP_009217269.1

Extended Table 2.4. Mass spectrometry of GFP fusion pulldowns

ChmA (gp53)-GFP

Protein	-10lgP	Coverage (%)	Peak Area	# Peptides	#Spectra	MW (Da)	Annotation	Accession Numbers
gp53	315.92	43	4410000000	37	301	66612	ChmA	YP_009217136.1
gp52	155.86	24	23600000	11	39	62495	hypothetical protein	YP_009217135.1
gp358	147.15	40	14400000	9	36	30858	hypothetical protein	YP_009217437.1
gp213	49.83	4	6330000	6	6	252377	tail fibre protein	YP_009217292.1
gp138	36.79	6	220000	1	1	39328	hypothetical protein	YP_009217217.1
gp54	21.97	4	0	1	1	55227	non-virion RNA polymerase subunit	YP_009217137.1
gp216	20.63	9	0	1	1	17417	hypothetical protein	YP_009217295.1

Extended Table 2.4. Mass spectrometry of GFP fusion pulldowns (continued)

GFP-ChmB (gp2) (N-terminal GFP fusion)

Protein	-10lgP	Coverage (%)	Peak Area	# Peptides	#Spectra	MW (Da)	Annotation	Accession Numbers
gp53	329.46	60	2690000000	52	362	66612	ChmA	YP_009217136.1
gp2	298.81	49	7240000000	36	604	22501	hypothetical protein (ChmB)	YP_009217084.1
gp166	291.32	42	398000000	41	148	115675	tail protein	YP_009217245.1
gp136	240.3	48	145000000	24	41	82847	major capsid protein	YP_009217215.1
gp358	196	55	72300000	12	38	30858	hypothetical protein	YP_009217437.1
gp11	184.29	25	21500000	12	28	77793	tail sheath protein	YP_009217093.1
gp164	115.84	35	6840000	5	13	24691	virion structural protein	YP_009217243.1
gp148	110.26	8	16000000	6	10	92750	portal protein	YP_009217227.1
gp12	80.7	8	597000	2	4	32114	virion structural protein	YP_009217094.1
gp158	76.75	9	4320000	2	4	32534	virion structural protein	YP_009217237.1
gp165	63.69	12	27100000	6	7	96202	tail sheath protein	YP_009217244.1
gp208	48.35	6	57400	2	3	66318	hypothetical protein	YP_009217287.1
gp293	48.27	8	45400	1	3	20641	hypothetical protein	YP_009217372.1
gp100	46.35	4	0	1	1	53843	virion structural protein	YP_009217180.1
gp28	43.66	3	1060000	1	1	34453	hypothetical protein	YP_009217111.1
gp175	39.57	3	4150000	2	4	53606	putative UvsX protein (RecA)	YP_009217254.1
gp138	37.68	6	4200000	1	1	39328	hypothetical protein	YP_009217217.1
gp64	37.57	3	92400	1	1	48010	hypothetical protein (ChmE)	YP_009217147.1
gp186	37.29	2	569000	1	1	62420	virion structural protein	YP_009217265.1
gp236	25.34	3	0	1	1	35027	hypothetical protein	YP_009217315.1
gp83	25.29	2	45000000	1	11	54744	hypothetical protein	YP_009217163.1
gp135	22.59	6	411000	1	1	20231	virion structural protein	YP_009217214.1
gp289	22.56	6	0	1	1	52788	hypothetical protein	YP_009217368.1

Extended Table 2.4. Mass spectrometry of GFP fusion pulldowns (continued)

ChmB (gp2)-GFP (C-terminal GFP fusion)

Protein	-10log(P)	Coverage (%)	Peak Area	# Peptides	#Spectra	MW (Da)	Annotation	Accession Numbers
gp53	349.45	59	2280000000	59	410	66612	ChmA	YP_009217136.1
gp2	342.87	64	3980000000	49	580	22501	ChmB	YP_009217084.1
gp166	338.05	54	344000000	56	212	115675	tail protein	YP_009217245.1
gp136	255.06	40	332000000	28	84	82847	capsid	YP_009217215.1
gp11	227.67	34	134000000	18	56	77793	tail sheath	YP_009217093.1
gp358	195.05	40	53000000	11	37	30858	hypothetical	YP_009217437.1
gp164	155.07	41	3770000	6	12	24691	virion structural protein	YP_009217243.1
gp148	149.77	10	19000000	8	16	92750	portal	YP_009217227.1
gp158	126.08	25	9280000	9	13	32534	virion structural protein	YP_009217237.1
gp165	123.26	25	1120000000	14	29	96202	tail sheath	YP_009217244.1
gp236	103.64	29	166000000	8	14	35027	hypothetical	YP_009217315.1
gp12	89.11	17	4920000	3	9	32114	virion structural protein	YP_009217094.1
gp181	86.37	7	53500000	5	6	49664	virion structural protein	YP_009217260.1
gp138	82.24	24	3580000	5	6	39328	hypothetical	YP_009217217.1
gp8	56.68	11	1340000	4	4	61839	virion structural protein	YP_009217090.1
gp135	44	6	107000	1	1	20231	virion structural protein	YP_009217214.1
gp186	43.76	2	657000	1	1	62420	virion structural protein	YP_009217265.1
gp213	31.99	2	18000000	3	3	252377	tail fiber	YP_009217292.1
gp293	31.57	8	0	1	1	20641	hypothetical	YP_009217372.1
gp150	31.46	3	759000	1	3	78941	virion structural protein	YP_009217229.1
gp64	29.77	3	69500	1	1	48010	hypothetical protein	YP_009217147.1
gp18	28.6	4	26400	1	1	38478	hypothetical	YP_009217100.1
gp269	27.73	39	348000	2	2	13468	hypothetical	YP_009217348.1
gp315	27.02	25	1420000	1	2	12632	hypothetical	YP_009217394.1
gp257	24.48	2	191000	1	1	76108	DNA ligase	YP_009217336.1
gp162	23.94	7	2180000	1	1	19889	virion structural protein	YP_009217241.1
gp175	21.9	1	1990000	1	1	53606	RecA	YP_009217254.1
gp212	20.98	2	0	1	1	63067	RNAP beta subunit	YP_009217291.1
gp103	20.72	3	125000	1	1	65004	hypothetical	YP_009217183.1

Extended Table 2.4. Mass spectrometry of GFP fusion pulldowns (continued)

gp62-GFP

Protein	-10log(P)	Coverage (%)	Peak Area	# Peptides	#Spectra	MW (Da)	Annotation	Accession Numbers
gp67	329.56	48	266000000	29	137	73779	non-virion RNA polymerase subunit	N/A
gp65	295.65	43	212000000	23	115	45123	non-virion RNA polymerase subunit	N/A
gp62	265.56	32	251000000	17	96	59703	hypothetical protein	YP_009217145.1
gp139	261.03	39	130000000	21	87	62710	non-virion RNA polymerase subunit	YP_009217218.1
gp66	218.35	40	106000000	10	83	35614	non-virion RNA polymerase subunit	N/A
gp54	206.37	23	133000000	10	70	55227	non-virion RNA polymerase subunit	YP_009217137.1
gp85	42.9	11	78500	4	4	70830	hypothetical protein	YP_009217165.1
gp213	35.25	2	0	3	3	252377	tail fibre protein	YP_009217292.1
gp200	32.84	6	252000	1	1	29433	hypothetical protein	YP_009217279.1
gp11	32.33	2	0	1	1	77793	tail sheath protein	YP_009217093.1
gp17	22.93	5	31900	1	1	48156	hypothetical protein	YP_009217099.1
gp214	20.6	2	435000	1	4	72021	hypothetical protein	YP_009217293.1

Extended Table 2.5. Crystallographic data collection and refinement

Data collection	PA1C gp2
Synchrotron/Beamline	APS 24ID-C
Date collected	44541
Resolution (Å)	67.1 - 2.63
Wavelength (Å)	0.97911
Space Group	P21
Unit Cell Dimensions (a,b,c) Å	60.96, 86.56, 71.88
Unit Cell Angles (α,β,γ)°	90, 110.94, 90
I/σ (last shell)	4.8 (0.9)
a Rsym (last shell)	0.184 (1.101)
b Rmeas (last shell)	0.220 (1.306)
c CC1/2, last shell	0.981 (0.462)
Completeness (last shell) %	98.5 (94.0)
Number of reflections	68907
unique	20612
Multiplicity (last shell)	3.3 (3.4)
Refinement	
Resolution (Å)	67.1 - 2.63
No. of reflections	20421
working	19514
free	907
d Rwork (%)	27.10 (36.48)
d Rfree (%)	30.01 (37.18)
Structure/Stereochemistry	
Number of atoms	10615
hydrogen	5259
r.m.s.d. Bond lengths (Å)	0.0029
r.m.s.d. Bond angles (°)	0.55
Rotamer outliers (%)	0.0067
Ramachandran (%)	
Favored	0.9875
Allowed	0.0125
Outliers	0
MolProbity score	0.92
MolProbity Clashscore	1.7
PDB ID	7UYX
SBGrid Data Bank ID	908

Extended Table 2.6. Plasmid constructs used in this study

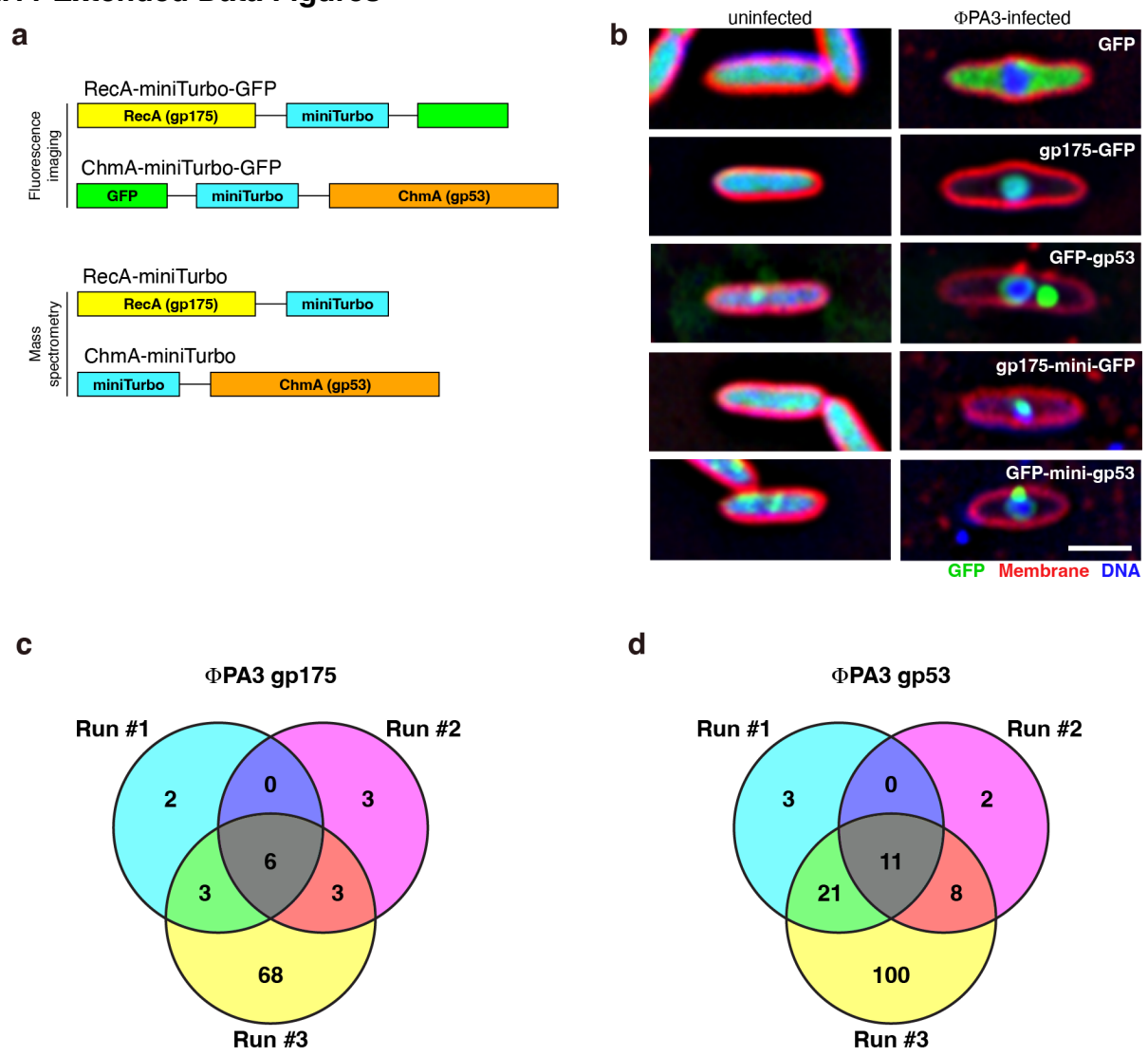
Construct Name	Vector backbone	Fusion tag (terminus)	Inserted protein	Inserted protein accession #	Mutation(s)
	pHERD30T				
pEE1	pHERD30T	GFPmut1/eGFP		AEP13835.1	
pEE2	pHERD30T	sfGFP		ASL68970.1	S2A, D190H & D197H
VC_591	pHERD30T	GFPmut1 (C-terminal)	PhiPA3 gp175		
VC_585	pHERD30T	GFPmut1 (N-terminal)	PhiPA3 gp53	YP_009217136.1	
pEE3	pHERD30T	mini-GFP (C-terminal)	PhiPA3 gp175		
pEE4	pHERD30T	GFP-mini (N-terminal)	PhiPA3 gp53	YP_009217136.1	
pEE5	pHERD30T	mini (C-terminal)	PhiPA3 gp175		
pEE6	pHERD30T	mini (N-terminal)	PhiPA3 gp53	YP_009217136.1	
pEE7	pHERD30T	mini-GFP	N/A		
pEE8	pHERD30T	sfGFP (C-terminal)	PhiPA3 gp18	YP_009217100.1	
pEE9	pHERD30T	sfGFP (C-terminal)	PhiPA3 gp62	YP_009217145.1	
pEE10	pHERD30T	sfGFP (C-terminal)	PhiPA3 gp78	YP_009217158.1	
pEE11	pHERD30T	sfGFP (C-terminal)	PhiPA3 gp108	YP_009217188.1	
pEE12	pHERD30T	sfGFP (C-terminal)	PhiPA3 gp134	YP_009217213.1	
pEE13	pHERD30T	sfGFP (C-terminal)	PhiPA3 gp200	YP_009217279.1	
pEE14	pHERD30T	sfGFP (C-terminal)	PhiPA3 gp210	YP_009217289.1	
pEE15	pHERD30T	sfGFP (C-terminal)	PhiPA3 gp257	YP_009217336.1	
pEE16	pHERD30T	sfGFP (N-terminal)	PhiPA3 gp2	YP_009217084.1	
pEE17	pHERD30T	sfGFP (N-terminal)	PhiPA3 gp61	YP_009217144.1	
pEE18	pHERD30T	sfGFP (N-terminal)	PhiPA3 gp63	YP_009217146.1	
pEE19	pHERD30T	sfGFP (N-terminal)	PhiPA3 gp64	YP_009217147.1	
pEE20	pHERD30T	sfGFP (N-terminal)	PhiPA3 gp148	YP_009217227.1	
pEE21	pHERD30T	sfGFP (C-terminal)	PhiPA3 gp375	YP_009217454.1	
pEE22	pHERD30T	sfGFP (N-terminal)	PhiPA3 gp12	YP_009217094.1	
pEE23	pHERD30T	sfGFP (C-terminal)	PhiPA3 gp164	YP_009217243.1	
pEE24	pHERD30T	sfGFP (C-terminal)	PhiPA3 gp165	YP_009217245.1	
pEE25	pHERD30T	sfGFP (N-terminal)	PhiPA3 gp233	YP_009217312.1	
pEE26	pHERD30T	sfGFP (N-terminal)	PhiPA3 gp355	YP_009217434.1	
pEE27	pHERD30T	sfGFP (C-terminal)	PhiPA3 gp14	YP_009217096.1	
pEE28	pHERD30T	sfGFP (C-terminal)	PhiPA3 gp52	YP_009217135.1	
pEE29	pHERD30T	sfGFP (C-terminal)	PhiPA3 gp106	YP_009217186.1	
pEE30	pHERD30T	sfGFP (C-terminal)	PhiPA3 gp119	YP_009217199.1	
pEE31	pHERD30T	sfGFP (N-terminal)	PhiPA3 gp131	YP_009217211.1	

Extended Table 2.6. Plasmid constructs used in this study

pEE32	pHERD30T	sfGFP (C-terminal)	PhiPA3 gp217	YP_009217296.1	
pEE33	pHERD30T	sfGFP (C-terminal)	PhiPA3 gp219	YP_009217298.1	
pEE34	pHERD30T	sfGFP (C-terminal)	PhiPA3 gp253	YP_009217332.1	
pEE35	pHERD30T	sfGFP (C-terminal)	PhiPA3 gp308	YP_009217387.1	
pEE36	pHERD30T	sfGFP (C-terminal)	PhiPA3 gp358	YP_009217437.1	
pEE37	pHERD30T	sfGFP (C-terminal)	PhiPA3 gp49	YP_009217132.1	
pEE38	pHERD30T	sfGFP (C-terminal)	PhiPA3 gp85	YP_009217165.1	
pEE39	pHERD30T	sfGFP (N-terminal)	PhiPA3 gp144	YP_009217223.1	
pEE40	pHERD30T	sfGFP (N-terminal)	PhiPA3 gp202	YP_009217281.1	
pEE41	pHERD30T	sfGFP (N-terminal)	PhiPA3 gp222	YP_009217301.1	
pEE42	pHERD30T	sfGFP (N-terminal)	PhiPA3 gp239	YP_009217318.1	
pEE43	pHERD30T	sfGFP (N-terminal)	PhiPA3 gp247	YP_009217326.1	
pEE44	pHERD30T	sfGFP (N-terminal)	PhiPA3 gp271	YP_009217350.1	
pEE45	pHERD30T	sfGFP (C-terminal)	PhiPA3 gp313	YP_009217392.1	
pEE46	pHERD30T	sfGFP (C-terminal)	PhiPA3 gp335	YP_009217414.1	
pEE47	pHERD30T	sfGFP (N-terminal)	PhiPA3 gp350	YP_009217429.1	
pEE48	pHERD30T	sfGFP (N-terminal)	PhiPA3 gp370	YP_009217449.1	
pEE49	pHERD30T	sfGFP (C-terminal)	PhiPA3 gp378	YP_009217457.1	
KN483	pHERD30T	mCherry (N-terminal) & GFP (C-terminal)	PhiPA3 gp53 & PhiPA3 gp2	YP_009217136.1 & YP_009217084.1	
KN457	pHERD30T	mCherry (N-terminal) & GFP (C-terminal)	201Phi-2 gp105 & 201Phi-2 gp2	YP_001956829.1 & YP_001956728.1	
pEE50	pHERD30T	sfGFP (C-terminal)	PhiPA3 gp2	YP_009217084.1	Q53A
pEE51	pHERD30T	sfGFP (C-terminal)	PhiPA3 gp2	YP_009217084.1	A159D
pEE52	UCB Macrolab 2-BT	His6-[TEV]- (N-terminal)	PhiPA3 gp2	YP_009217084.1	
pEE53	UCB Macrolab 2-BT	His6-[TEV]- (N-terminal)	201Phi2-1 gp2	YP_001956728.1	
pEE54	UCB Macrolab 2-BT	His6-[TEV]- (N-terminal)	Psa21 gp2	YP_010347551.1	
pEE55	UCB Macrolab 2-BT	His6-[TEV]- (N-terminal)	PA1C gp3	QBX32150.1	
pEE56	UCB Macrolab 13S-A		201Phi-2 gp105	YP_001956829.1	
pEE57	UCB Macrolab 2-BT	His6-[TEV]- (N-terminal)	PhiPA3 gp2	YP_009217084.1	Δ1-63
pEE58	UCB Macrolab 2-BT	His6-[TEV]- (N-terminal)	PhiPA3 gp2	YP_009217084.1	Δ583-631
pEE59	UCB Macrolab 2-BT	His6-[TEV]- (N-terminal)	PhiPA3 gp2	YP_009217084.1	Δ1-63 & Δ583-631
pEE60	UCB Macrolab 13S-A		PhiPA3 gp148	YP_009217227.1	
pEE61	UCB Macrolab 2-BT	His6-[TEV]- (N-terminal)	PhiPA3 gp2	YP_009217084.1	Q53A
pEE62	UCB Macrolab 2-BT	His6-[TEV]- (N-terminal)	PhiPA3 gp2	YP_009217084.1	A159D
pEE63	UCB Macrolab 2-BT	His6-[TEV]- (N-terminal)	201Phi2-1 gp2	YP_001956728.1	Q53A

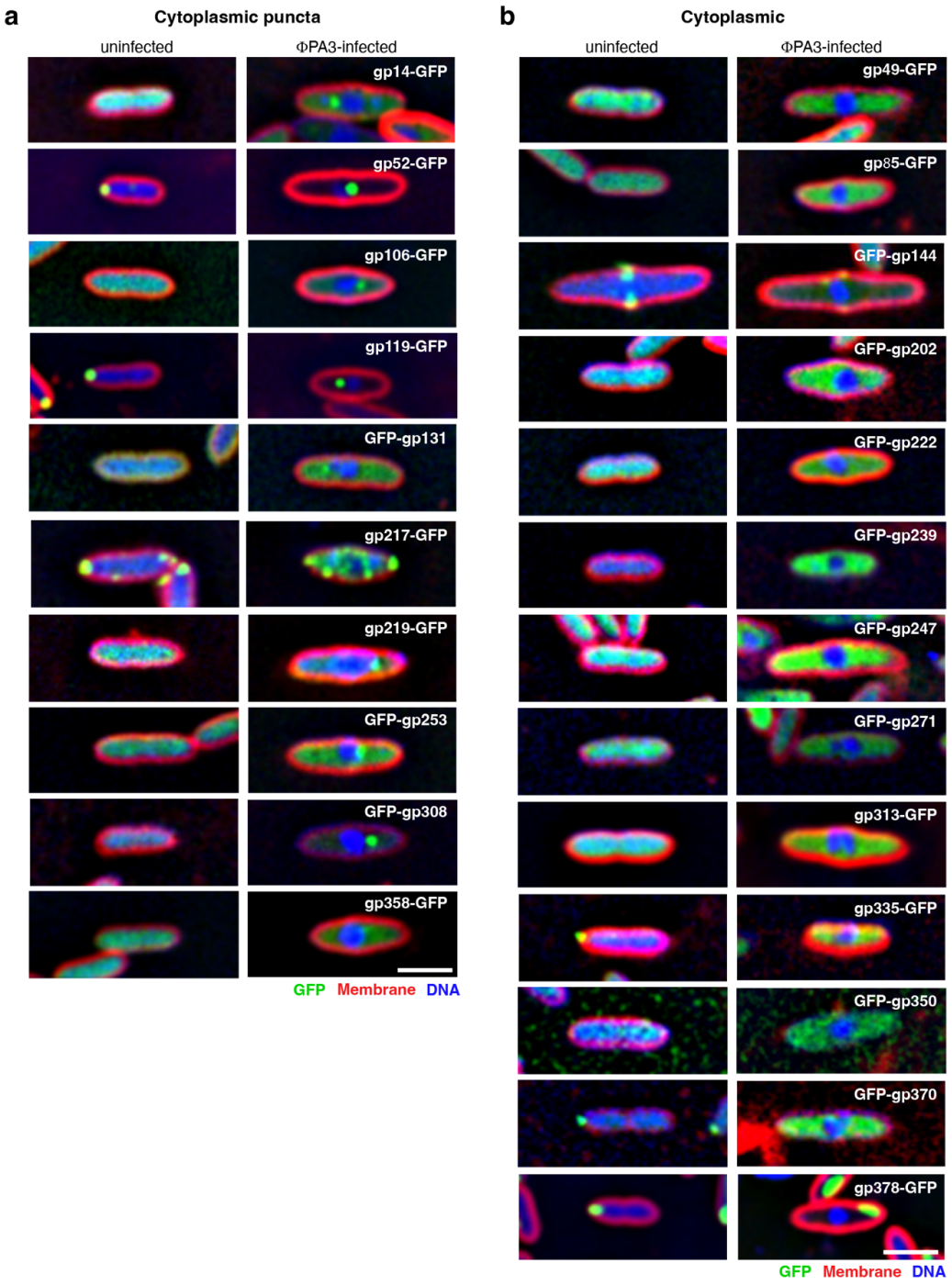
pEE64	UCB Macrolab 2-BT	His6-[TEV]- (N-terminal)	201Phi2-1 gp2	YP_001956728.1	A145D

2.11 Extended Data Figures



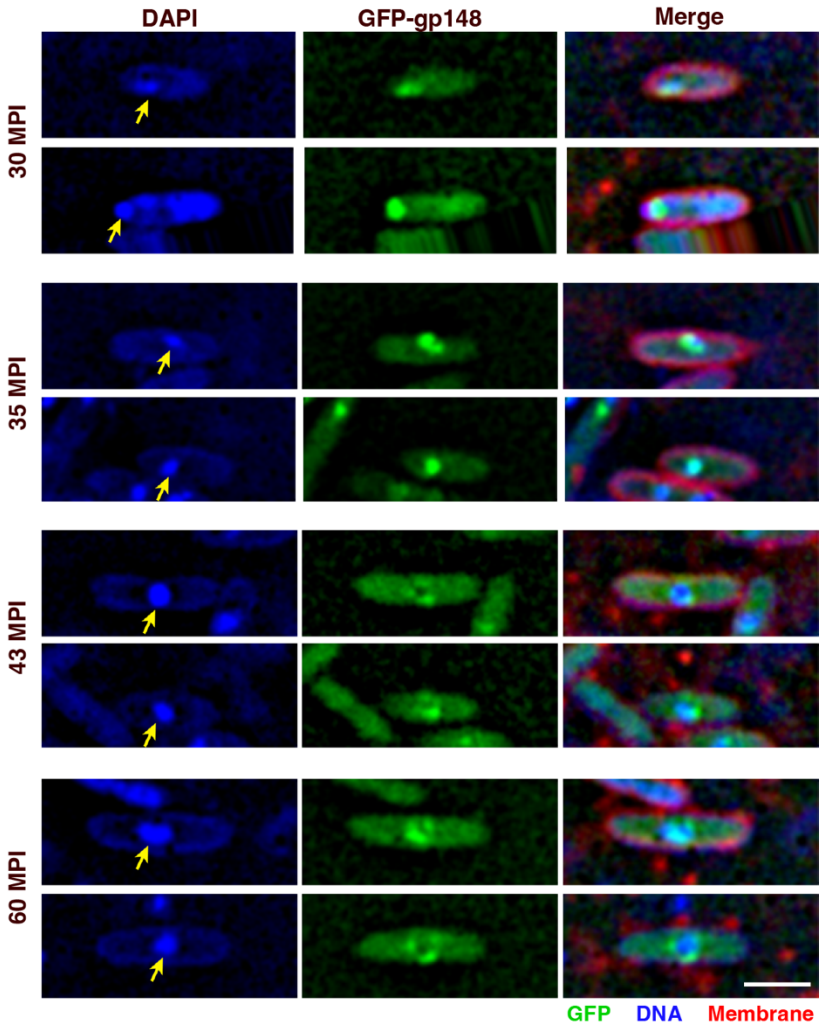
Extended Figure 2.1. miniTurboID proximity labeling of phage nucleus-associated proteins

(a) Construct design for localization and proximity labeling of ΦPA3 RecA (gp175) and ChmA (gp53) associated proteins. (b) Localization of GFP control and GFP- and GFP-miniTurboID tagged RecA (gp175) and ChmA (gp53) in ΦPA3-infected *P. aeruginosa* cells. GFP is shown in green, FM4-64 (to visualize membranes) in red, and DAPI (to visualize nucleic acids) in blue. Scale bar = 2 µm. (c-d) Venn diagrams showing RecA (panel c) and ChmA (panel d) interacting proteins identified by miniTurboID labeling in three independent runs. See **Tables 1-2** and **Extended Data Tables 1-2**.

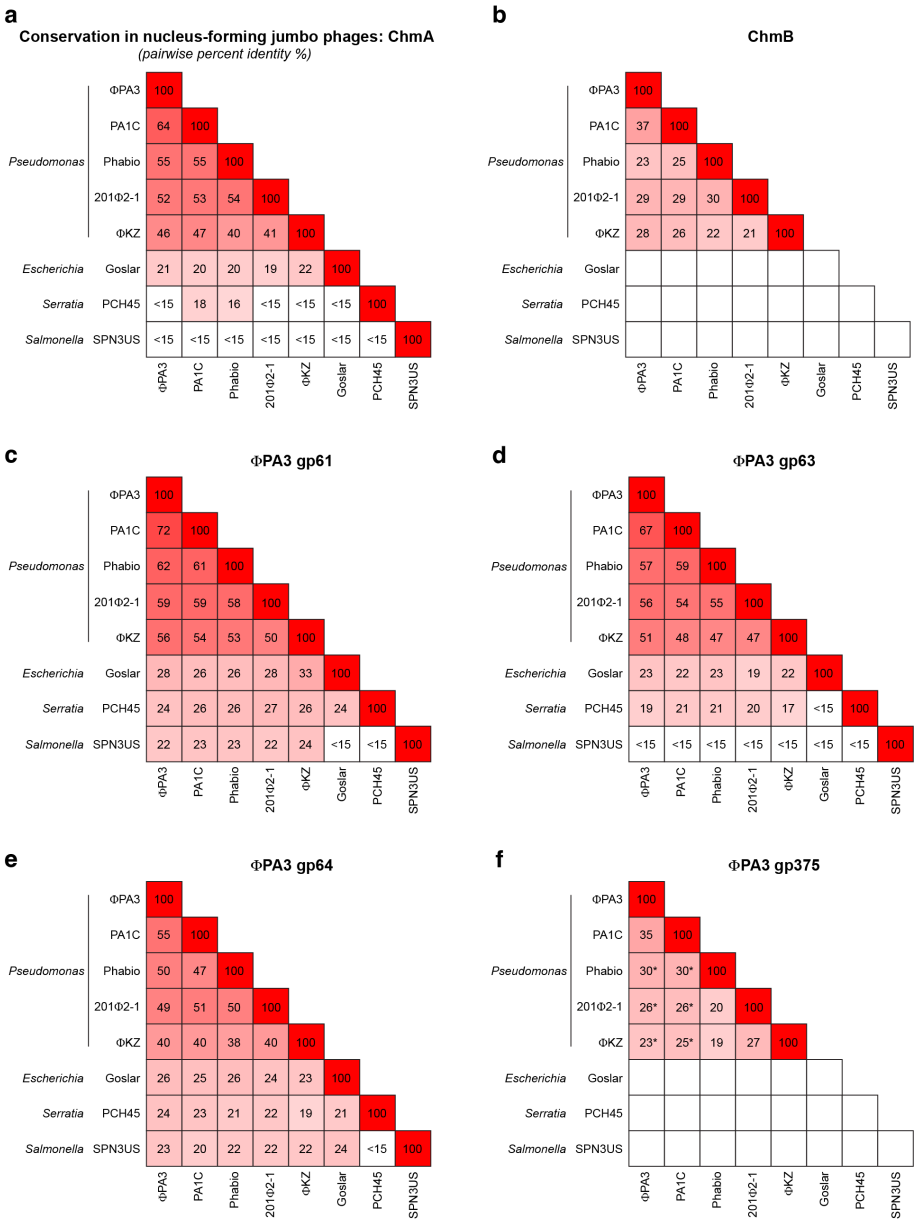


Extended Figure 2.2. Localization analysis of RecA- and ChmA-interacting ΦPA3 proteins

Subcellular localization of selected proteins identified by proximity labeling, with panel (a) showing proteins that localize as cytoplasmic puncta, and (b) showing proteins with diffuse cytoplasmic localization. See **Table 3** for a collated list of localizations. GFP is shown in green, FM4-64 (to visualize membranes) in red, and DAPI (to visualize nucleic acids) in blue. Scale bar = 2 μm.

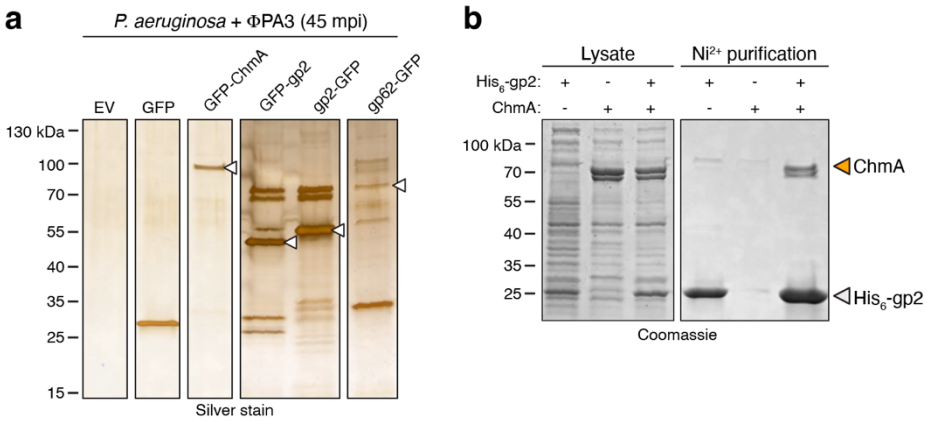


Extended Figure 2.3. Early localization of gp148 to the phage nuclear shell
Subcellular localization of GFP-gp148 in *P. aeruginosa* cells infected with phage ΦPA3, at the indicated times post infection (MPI: minutes post infection). Yellow arrows indicate the position of the phage nucleus. GFP is shown in green, FM4-64 (to visualize membranes) in red, and DAPI (to visualize nucleic acids) in blue. Scale bar = 2 µm.

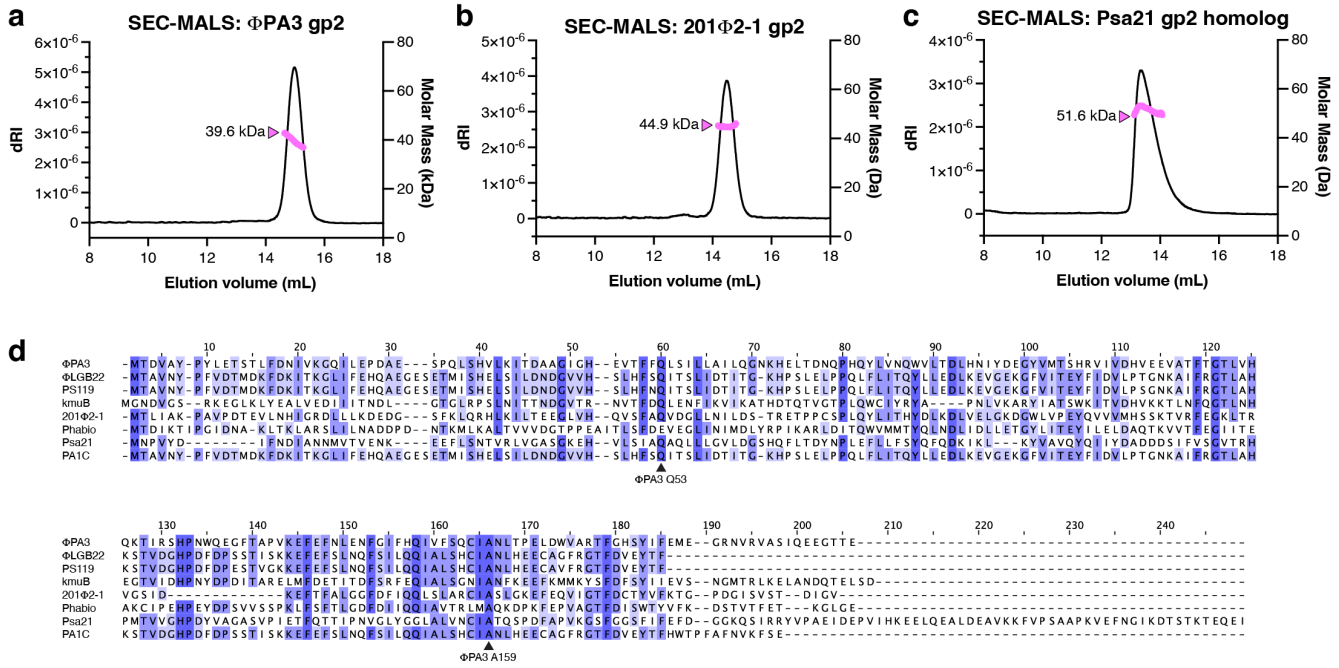


Extended Figure 2.4. Sequence analysis of nuclear-localized jumbo phage proteins

Sequence identity (expressed as percent identity) between selected jumbo phage ChmA proteins: ΦPA3 (NCBI Accession # YP_009217136.1), PA1C (QBX32206.1), Phabio (YP_010348051.1), 201Φ2-1 (YP_001956829.1), ΦKZ (NP_803620.1), Goslar (YP_009820873.1), PCH45 (QFP93061.1), SPN3US (YP_009153316.1). For all pairs of sequences, the "Pairwise Alignment" tool in JalView was used to calculate % identity over the homologous region of each protein. For those identities reported as "<15%", this tool was unable to generate an alignment over a significant portion of the sequences. Asterisks indicate that these identities represent the homologous regions of dramatically different-length proteins.

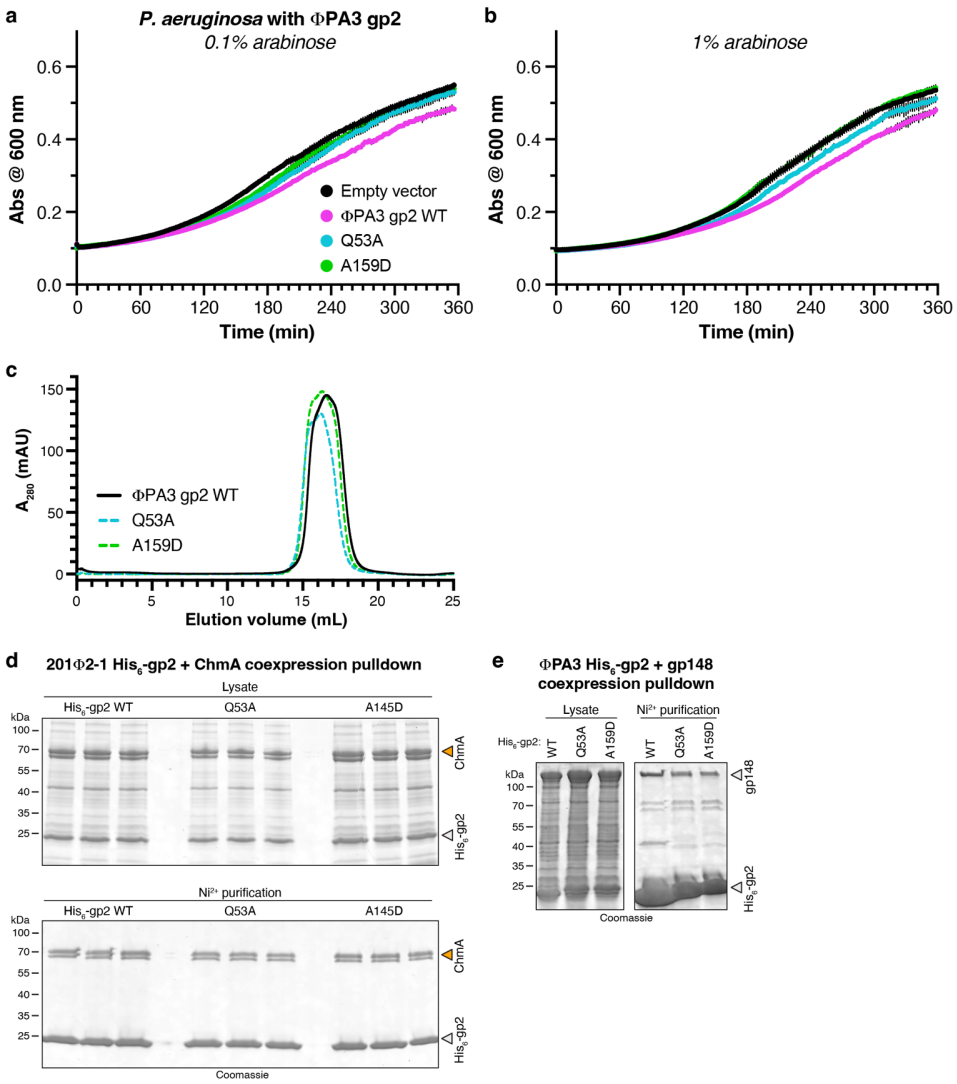


Extended Figure 2.5. Protein-protein interaction analysis
(a) Silver stained SDS-PAGE gels showing GFP pulldown results from GFP-tagged proteins expressed in ΦPA3-infected *P. aeruginosa* (45 minutes post infection). White arrowheads indicate the bait protein for each sample. (b) Coomassie blue-stained SDS-PAGE gel showing Ni^{2+} pulldown results from coexpression of 201Φ2-1 gp2 (His_6-tagged) and ChmA (gp105; untagged). ChmA appears as a doublet because of a second start codon at codon 33 of the gp105 gene.



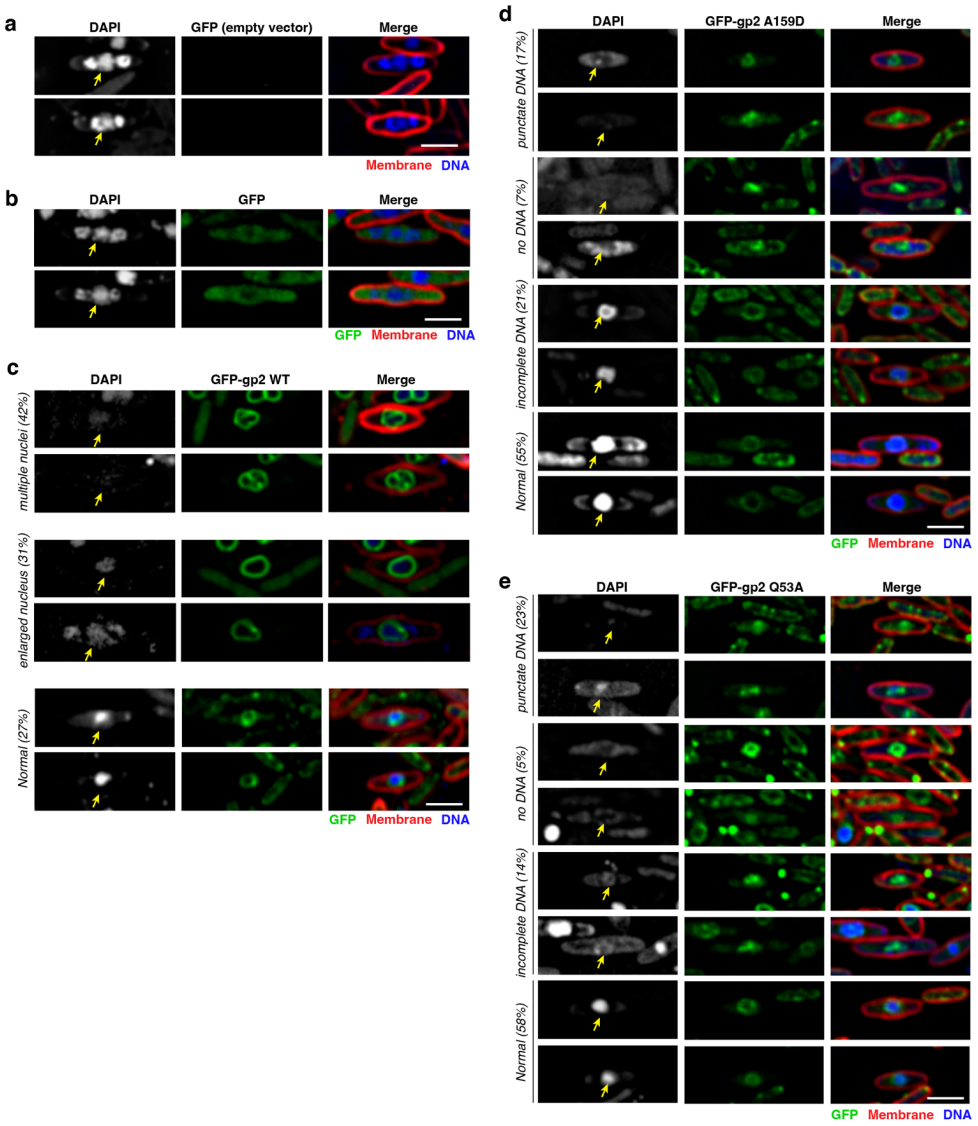
Extended Figure 2.6. Biochemical and sequence analysis of jumbo phage gp2 proteins

(a) Size exclusion chromatography coupled to multi-angle light scattering (SEC-MALS) analysis of ΦPA3 gp2. Measured molecular weight = 39.6 kDa; dimer molecular weight = 45 kDa. (b) SEC-MALS analysis of 201Φ201 gp2. Measured molecular weight = 44.9 kDa; dimer molecular weight = 39.8 kDa. (c) SEC-MALS analysis of the gp2 homolog in phage Psa21 (gp2). Measured molecular weight = 51.6 kDa; dimer molecular weight = 50.6 kDa. (d) Sequence alignment of gp2 homologs in jumbo phage infecting *Pseudomonas*, showing the position of the highly conserved Q53 and A159 (ΦPA3 numbering) residues.



Extended Figure 2.7. Effects of gp2 Q53A and A159D mutations

(a) Growth curves of *P. aeruginosa* cells transformed with pHERD30T empty vector (black) or encoding ΦPA3 gp2 wild-type (violet), Q53A (cyan), or A159D (green). Growth media contained 0.1% arabinose for low-level expression. Datapoints represent the average of three independent replicates, and error bars (black) represent standard deviation. (b) Growth curves of *P. aeruginosa* cells transformed as in panel (a), with growth media containing 1% arabinose for high-level expression. (c) Size exclusion chromatography elution profiles for ΦPA3 ΦPA3 gp2 wild-type (black solid line), Q53A (cyan dotted line), or A159D (green dotted line). (d) Ni^{2+} pulldown analysis of *E. coli*-coexpressed His_6-tagged 201Φ2-1 gp2 (wild-type (WT), Q53A, or A145D (equivalent to ΦPA3 gp2 A159D)) and full-length ChmA. Doublet bands for ChmA arise from a methionine codon at position 33 of the annotated gene. (e) Ni^{2+} pulldown analysis of *E. coli*-coexpressed His_6-tagged ΦPA3 gp2 (wild-type (WT), Q53A, or A159D) and portal (gp148). The ~40 kDa band in pulldown in the Ni^{2+} purification lane for WT gp2 is an unknown contaminant.



Extended Figure 2.8. Overexpression of mutant gp2 proteins affect nuclear shell formation and morphology

(a) Microscopy of ΦPA3-infected *P. aeruginosa* cells transformed with empty vector. Yellow arrow indicates the position of the phage nucleus. The two DAPI staining bodies bracketing the phage nucleus are the phage bouquets. GFP is shown in green, FM4-64 (to visualize membranes) in red, and DAPI (to visualize nucleic acids) in blue. Scale bar = 2 µm. (b) Microscopy of ΦPA3-infected *P. aeruginosa* cells expressing GFP. (c) Microscopy of ΦPA3-infected *P. aeruginosa* cells expressing GFP-tagged wild-type gp2. Common phenotypes and their percentages in the analyzed sample (n=100 cells) are shown. (d) Microscopy of ΦPA3-infected *P. aeruginosa* cells expressing GFP-tagged gp2 A159D. Common phenotypes and their percentages in the analyzed sample (n=100 cells) are shown. (e) Microscopy of ΦPA3-infected *P. aeruginosa* cells expressing GFP-tagged gp2 Q53A. Common phenotypes and their percentages in the analyzed sample (n=100 cells) are shown.

2.12 References

1. Salmond, G. P. C. & Fineran, P. C. A century of the phage: past, present and future. Nat. Rev. Microbiol. 13, 777–786 (2015).
2. Serwer, P., Hayes, S. J., Thomas, J. A. & Hardies, S. C. Propagating the missing bacteriophages: a large bacteriophage in a new class. Virol. J. 4, 21 (2007).
3. Al-Shayeb, B., Sachdeva, R., Chen, L.-X., Ward, F., Munk, P., Devoto, A., Castelle, C. J., Olm, M. R., Bouma-Gregson, K., Amano, Y., He, C., Méheust, R., Brooks, B., Thomas, A., Lavy, A., Matheus-Carnevali, P., Sun, C., Goltsman, D. S. A., Borton, M. A., Sharrar, A., Jaffe, A. L., Nelson, T. C., Kantor, R., Keren, R., Lane, K. R., Farag, I. F., Lei, S., Finstad, K., Amundson, R., Anantharaman, K., Zhou, J., Probst, A. J., Power, M. E., Tringe, S. G., Li, W.-J., Wrighton, K., Harrison, S., Morowitz, M., Relman, D. A., Doudna, J. A., Lehours, A.-C., Warren, L., Cate, J. H. D., Santini, J. M. & Banfield, J. F. Clades of huge phages from across Earth's ecosystems. Nature 578, 425–431 (2020).
4. M Iyer, L., Anantharaman, V., Krishnan, A., Burroughs, A. M. & Aravind, L. Jumbo Phages: A Comparative Genomic Overview of Core Functions and Adaptions for Biological Conflicts. Viruses 13, (2021).
5. Chaikeeratisak, V., Birkholz, E. A. & Pogliano, J. The Phage Nucleus and PhuZ Spindle: Defining Features of the Subcellular Organization and Speciation of Nucleus-Forming Jumbo Phages. Front. Microbiol. 12, 641317 (2021).
6. Chaikeeratisak, V., Nguyen, K., Khanna, K., Brilot, A. F., Erb, M. L., Coker, J. K. C., Vavilina, A., Newton, G. L., Buschauer, R., Pogliano, K., Villa, E., Agard, D. A. & Pogliano, J. Assembly of a nucleus-like structure during viral replication in bacteria. Science 355, 194–197 (2017).
7. Chaikeeratisak, V., Nguyen, K., Egan, M. E., Erb, M. L., Vavilina, A. & Pogliano, J. The Phage Nucleus and Tubulin Spindle Are Conserved among Large Pseudomonas Phages. Cell Rep. 20, 1563–1571 (2017).
8. Mendoza, S. D., Nieweglowska, E. S., Govindarajan, S., Leon, L. M., Berry, J. D., Tiwari, A., Chaikeeratisak, V., Pogliano, J., Agard, D. A. & Bondy-Denomy, J. A bacteriophage nucleus-like compartment shields DNA from CRISPR nucleases. Nature 577, 244–248 (2020).
9. Malone, L. M., Warring, S. L., Jackson, S. A., Warnecke, C., Gardner, P. P., Gumy, L. F. & Fineran, P. C. A jumbo phage that forms a nucleus-like structure evades CRISPR-Cas DNA targeting but is vulnerable to type III RNA-based immunity. Nat Microbiol 5, 48–55 (2020).
10. Laughlin, T. G., Deep, A., Prichard, A. M., Seitz, C., Gu, Y., Enustun, E., Suslov, S., Khanna, K., Birkholz, E. A., Armbruster, E., McCammon, J. A., Amaro, R. E., Pogliano, J., Corbett, K. D. & Villa, E. Architecture and self-assembly of the jumbo bacteriophage nuclear shell. Nature 608, 429–435 (2022).
11. Nieweglowska, E. S., Brilot, A. F., Méndez-Moran, M., Kokontis, C., Baek, M., Li, J., Cheng, Y., Baker, D., Bondy-Denomy, J. & Agard, D. A. The ϕPA3 phage nucleus is enclosed by a self-assembling 2D crystalline lattice. Nat. Commun. 14, 927 (2023).
12. Nguyen, K. T., Sugie, J., Khanna, K., Egan, M. E., Birkholz, E. A., Lee, J., Beierschmitt, C., Villa, E. & Pogliano, J. Selective transport of fluorescent proteins

into the phage nucleus. PLoS One 16, e0251429 (2021).
13. Chaikeeratisak, V., Khanna, K., Nguyen, K. T., Sugie, J., Egan, M. E., Erb, M. L., Vavilina, A., Nonejuie, P., Nieweglowska, E., Pogliano, K., Agard, D. A., Villa, E. & Pogliano, J. Viral Capsid Trafficking along Treadmilling Tubulin Filaments in Bacteria. Cell 177, 1771-1780.e12 (2019).
14. Chaikeeratisak Vorrapon, Khanna Kanika, Nguyen Katrina T., Egan MacKennon E., Enustun Eray, Armbruster Emily, Lee Jina, Pogliano Kit, Villa Elizabeth & Pogliano Joe. Subcellular organization of viral particles during maturation of nucleus-forming jumbo phage. Science Advances 8, eabj9670 (2022).
15. Birkholz, E. A., Laughlin, T. G., Armbruster, E., Suslov, S., Lee, J., Wittmann, J., Corbett, K. D., Villa, E. & Pogliano, J. A cytoskeletal vortex drives phage nucleus rotation during jumbo phage replication in E. coli. Cell Rep. 40, 111179 (2022).
16. Branon, T. C., Bosch, J. A., Sanchez, A. D., Udeshi, N. D., Svinkina, T., Carr, S. A., Feldman, J. L., Perrimon, N. & Ting, A. Y. Efficient proximity labeling in living cells and organisms with TurboID. Nat. Biotechnol. 36, 880–887 (2018).
17. Monson, R., Foulds, I., Foweraker, J., Welch, M. & Salmond, G. P. C. The Pseudomonas aeruginosa generalized transducing phage phiPA3 is a new member of the phiKZ-like group of “jumbo” phages, and infects model laboratory strains and clinical isolates from cystic fibrosis patients. Microbiology 157, 859–867 (2011).
18. Thomas, J. A., Weintraub, S. T., Wu, W., Winkler, D. C., Cheng, N., Steven, A. C. & Black, L. W. Extensive proteolysis of head and inner body proteins by a morphogenetic protease in the giant Pseudomonas aeruginosa phage φKZ. Mol. Microbiol. 84, 324–339 (2012).
19. deYMartín Garrido, N., Orekhova, M., Lai Wan Loong, Y. T. E., Litvinova, A., Ramlaul, K., Artamonova, T., Melnikov, A. S., Serdobintsev, P., Aylett, C. H. S. & Yakunina, M. Structure of the bacteriophage PhiKZ non-virion RNA polymerase. Nucleic Acids Res. 49, 7732–7739 (2021).
20. Thomas, J. A., Rolando, M. R., Carroll, C. A., Shen, P. S., Belnap, D. M., Weintraub, S. T., Serwer, P. & Hardies, S. C. Characterization of Pseudomonas chlororaphis myovirus 201φphi2-1 via genomic sequencing, mass spectrometry, and electron microscopy. Virology 376, 330–338 (2008).
21. Thomas, J. A., Weintraub, S. T., Hakala, K., Serwer, P. & Hardies, S. C. Proteome of the large Pseudomonas myovirus 201φ2-1: delineation of proteolytically processed virion proteins. Mol. Cell. Proteomics 9, 940–951 (2010).
22. Reilly, E. R., Abajorga, M. K., Kiser, C., Mohd Redzuan, N. H., Haidar, Z., Adams, L. E., Diaz, R., Pinzon, J. A., Hudson, A. O., Black, L. W., Hsia, R.-C., Weintraub, S. T. & Thomas, J. A. A Cut above the Rest: Characterization of the Assembly of a Large Viral Icosahedral Capsid. Viruses 12, (2020).
23. Sun, L., Zhang, X., Gao, S., Rao, P. A., Padilla-Sanchez, V., Chen, Z., Sun, S., Xiang, Y., Subramaniam, S., Rao, V. B. & Rossmann, M. G. Cryo-EM structure of the bacteriophage T4 portal protein assembly at near-atomic resolution. Nat. Commun. 6, 7548 (2015).
24. Mitchell, M. S. & Rao, V. B. Functional analysis of the bacteriophage T4 DNA-packaging ATPase motor. J. Biol. Chem. 281, 518–527 (2006).

25. Morita, M., Tasaka, M. & Fujisawa, H. Analysis of the fine structure of the prohead binding domain of the packaging protein of bacteriophage T3 using a hexapeptide, an analog of a prohead binding site. Virology 211, 516–524 (1995).
26. Valpuesta, J. M. & Carrascosa, J. L. Structure of viral connectors and their function in bacteriophage assembly and DNA packaging. Q. Rev. Biophys. 27, 107–155 (1994).
27. Yeo, A. & Feiss, M. Specific Interaction of Terminase, the DNA Packaging Enzyme of Bacteriophage γ, with the Portal Protein of theeh Prohead. J. Mol. Biol. 245, 141–150 (1995).
28. Heymann, J. B., Wang, B., Newcomb, W. W., Wu, W., Winkler, D. C., Cheng, N., Reilly, E. R., Hsia, R.-C., Thomas, J. A. & Steven, A. C. The Mottled Capsid of the Salmonella Giant Phage SPN3US, a Likely Maturation Intermediate with a Novel Internal Shell. Viruses 12, (2020).
29. Holm, L. Using Dali for Protein Structure Comparison. Methods Mol. Biol. 2112, 29–42 (2020).
30. van Kempen, M., Kim, S. S., Tumescheit, C., Mirdita, M., Söding, J. & Steinegger, M. Foldseek: fast and accurate protein structure search. bioRxiv 2022.02.07.479398 (2022). doi:10.1101/2022.02.07.479398
31. Kraemer, J. A., Erb, M. L., Waddling, C. A., Montabana, E. A., Zehr, E. A., Wang, H., Nguyen, K., Pham, D. S. L., Agard, D. A. & Pogliano, J. A phage tubulin assembles dynamic filaments by an atypical mechanism to center viral DNA within the host cell. Cell 149, 1488–1499 (2012).
32. Balemans, W., Vranckx, L., Lounis, N., Pop, O., Guillemont, J., Vergauwen, K., Mol, S., Gilissen, R., Motte, M., Lançois, D., De Bolle, M., Bonroy, K., Lill, H., Andries, K., Bald, D. & Koul, A. Novel antibiotics targeting respiratory ATP synthesis in Gram-positive pathogenic bacteria. Antimicrob. Agents Chemother. 56, 4131–4139 (2012).
33. Qiu, D., Damron, F. H., Mima, T., Schweizer, H. P. & Yu, H. D. PBAD-based shuttle vectors for functional analysis of toxic and highly regulated genes in Pseudomonas and Burkholderia spp. and other bacteria. Appl. Environ. Microbiol. 74, 7422–7426 (2008).
34. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P. & Cardona, A. Fiji: an open-source platform for biological-image analysis. Nat. Methods 9, 676–682 (2012).
35. Tropea, J. E., Cherry, S. & Waugh, D. S. in High Throughput Protein Expression and Purification: Methods and Protocols (ed. Doyle, S. A.) 297–307 (Humana Press, 2009).
36. Kabsch, W. Integration, scaling, space-group assignment and post-refinement. Acta Crystallogr. D Biol. Crystallogr. 66, 133–144 (2010).
37. Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? Acta Crystallogr. D Biol. Crystallogr. 69, 1204–1214 (2013).
38. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A. & Others. Overview of the CCP4 suite and current developments. Acta Crystallogr. D Biol. Crystallogr. 67,

235–242 (2011).
39. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. Phaser crystallographic software. J. Appl. Crystallogr. 40, 658–674 (2007).
40. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A. W., Kavukcuoglu, K., Kohli, P. & Hassabis, D. Highly accurate protein structure prediction with AlphaFold. Nature 596, 583–589 (2021).
41. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. 66, 486–501 (2010).
42. Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H. & Adams, P. D. Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr. D Biol. Crystallogr. 68, 352–367 (2012).

2.13 Acknowledgements

The authors acknowledge support from the National Institutes of Health (R01 GM129245 to J.P. and E.V.; R35 GM144121 to K.D.C.; and NIH shared instrumentation grant S10 OD021724), and the Howard Hughes Medical Institute Emerging Pathogens Initiative (to E.V., J.P., and K.D.C.).

Chapter 2, in full, has been submitted for publication of the material as it may appear in *NSMB,* 2023. Enustun, E., Deep, A., Gu, Y., Nguyen, K.T., Chaikeeratisak, V., Armbruster, E., Ghassemian, M., Villa, E., Pogliano, J., Corbett, K.D. (2023). Identification of the bacteriophage nucleus protein interaction network. *NSMB*. The dissertation author was the primary investigator and author of this paper.

Chapter 3: A phage nucleus-associated RNA-binding protein required for jumbo phage infection

Eray Enustun[1], Emily Armbruster[1], Jina Lee[1], Sitao Zhang[2], Brian Lee[2], Yajie Gu[2], Amar Deep[2], Jack Naritomi[2], Qishan Liang[2], Stefan Aigner[2], Vorrapon Chaikeeratisak[3], Don Cleveland[2], Majid Ghassemian[4], Gene Yeo, Joe Pogliano[1]‡, Kevin D. Corbett[1,2]‡

[1]Department of Molecular Biology, University of California San Diego, La Jolla, CA, USA.

[2]Department of Cellular and Molecular Medicine, University of California San Diego, La Jolla, CA, USA.

[3]Department of Biochemistry, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand

[4]Biomolecular and Proteomics Mass Spectrometry Facility, University of California San Diego, La Jolla, CA, USA.

‡Correspondence to: JP: jpogliano@ucsd.edu, KDC: kcorbett@ucsd.edu

3.1 Abstract

Large-genome bacteriophages (jumbo phages) of the *Chimalliviriadae* family assemble a nucleus-like compartment bounded by a protein shell that protects the replicating phage genome from host-encoded restriction enzymes and CRISPR/Cas nucleases. While the nuclear shell provides broad protection against host nucleases, it necessitates transport of mRNA out of the nucleus-like compartment for translation by host ribosomes, and transport of specific proteins into the nucleus-like compartment to support DNA replication and mRNA transcription. Here we identify a conserved phage nuclear shell-associated protein that we term chimallin C (ChmC), which adopts a nucleic acid-binding fold, binds RNA with high affinity *in vitro* and binds phage mRNAs in infected cells. ChmC also forms phase-separated condensates with RNA. Targeted knockdown of ChmC using mRNA-targeting Cas13d halts infections at an early stage. Taken

together, our data suggest that the conserved ChmC protein acts as a chaperone for phage mRNAs, potentially stabilizing these mRNAs and driving their translocation through the nuclear shell to promote translation and infection progression.

3.2 Introduction

The continual arms race between bacteria and bacteriophages (phages) has driven the development of myriad immune systems in bacteria, along with an equally complex set of phage-encoded immune countermeasures [1,2]. A striking example of these immune countermeasures is the nucleus-like compartment assembled by a family of large-genome "jumbo" phages (defined as phages with genomes >200 kb in size) now termed Chimalliviridae [3–7]. This compartment shields the phages' replicating genomes from host-encoded defenses including restriction enzymes and CRISPR/Cas nucleases [8,9] . The phage nuclear boundary or shell primarily comprises a single protein, termed Chimallin (ChmA), which assembles into a lattice with pores less than ~2 nm in size, which can allow the passage of metabolites but restricts the passage of most proteins [7,10]. Most Chimalliviridae also encode a tubulin homolog called PhuZ, which assembles into dynamic filaments that center and rotate the phage nucleus within the infected cell while also trafficking pro-capsids to the phage nucleus for genome packaging (refs).

The physical barrier erected by ChmA-encoding jumbo phages between their replicating genomic DNA and the host cytoplasm effectively counters DNA-targeting host defenses [8,9], but introduces a number of challenges to the replicating phage that mirror challenges faced by eukaryotic cells and their nuclei. In particular, since mRNAs are produced within the phage nucleus but the translation machinery is located in the host

cytoplasm, mRNAs must be translocated out of the phage nucleus for translation. Similarly, any phage protein whose function requires it to be localized within the phage nucleus must be specifically translocated from the cytoplasm into the phage nucleus after translation. Finally, replicated phage genomic DNA must be translocated through the nuclear shell for packaging into pro-capsids that are docked onto the exterior of the nuclear shell.

The requirement for translocation of mRNAs, proteins, and genomic DNA through the phage nuclear shell implies the existence of minor shell components embedded within or associated with the ChmA lattice that mediate these activities. In prior work, we used proximity labeling and localization analysis in the *Chimalliviridae* family phage PhiPA3 to identify proteins that physically associate with ChmA and localize to the nuclear shell (Enustun et al, submitted). One of these proteins, termed ChmB, interacts directly with ChmA both *in vitro* and *in vivo*, and associates with the virion portal protein *in vitro*. ChmB's network of protein-protein interactions and its distinctive 3D structure suggest that it may form pores in the phage nuclear shell that enable the docking of pro-capsids for genome packaging. Further data showing that the expression of dominant-negative ChmB mutants compromises early steps in phage nucleus growth and maturation further suggests that ChmB may participate in mRNA and/or protein translocation through the nuclear shell (Enustun et al, submitted).

Here, we show that another conserved phage nuclear shell-associated protein, which we term ChmC, adopts a nucleic acid-binding fold and binds RNA *in vitro*, and forms RNA-protein condensates through a conserved asparagine-rich C-terminal region.

In phage-infected cells, ChmC specifically binds phage mRNAs. Targeting ChmC using Cas13-based mRNA knockdown reveals that the protein plays critical roles in assembly, growth, and maturation of the phage nucleus. Together, these data suggest that ChmC acts as a chaperone for phage-encoded mRNAs, likely aiding their translocation through the nuclear shell to promote translation and infection progression.

3.3 Results

3.3.1. Identification of the abundant and early-expressed jumbo phage protein ChmC

We previously used proximity ligation to identify proteins in the nucleus-forming jumbo phage PhiPA3 that physically associate with the major nuclear shell protein ChmA and with a phage nucleus-localized protein, UvsX (gp175) (Enustun et al. submitted). Through this analysis, we identified a minor nuclear shell component (ChmB) and several uncharacterized proteins with no known function (Enustun et al. submitted). One of these proteins was PhiPA3 gp61, which is located in a highly conserved cluster of genes that includes *chmA* (gp53), several subunits of the phage-encoded "non-virion RNA polymerase" (nvRNAP: gp62, gp65-66, and gp67), and two additional phage nucleus-associated proteins, gp63 and gp64 (**Figure 1A**). gp61 is of particular interest as its homolog is the second highest-expressed non-structural protein in *Pseudomonas chlororaphis* cells infected with the related jumbo phage 201Phi2-1 (gp123) [4]. To test whether PhiPA3 gp61 is also highly expressed, we infected *P. aeruginosa* cells with PhiPA3 and performed mass spectrometry proteomics to identify the timing and expression of phage proteins. We confirmed that ChmA is the most abundantly expressed non-structural protein, and that gp61 is the second-most abundant (**Figure 1B, Table S1-**

S2). Based on its conservation, abundance, and association with the jumbo phage nuclear shell, we term this protein Chimallin C (ChmC).

Confirming our earlier microscopic observations (Enustun et al. submitted), we find that PhiPA3 ChmC fused to GFP localizes both to the cytoplasm and the phage nuclear shell in late-stage PhiPA3 infections of *P. aeruginosa* (**Figure 1C**). Fluorescent microscopy is unable to determine whether ChmC is localized within the nucleus or just outside the ChmA shell, but our prior identification of ChmC through proximity-labeling with the nuclear-localized protein UvsX suggests that the protein is at least partially localized within the phage nucleus (Enustun et al. submitted). We next purified GFP-tagged ChmC from *P. aeruginosa* cells infected with PhiPA3 and used mass spectrometry to identify associated proteins (**Figure S1, Table S3**). In this analysis, we identified the putative phage nucleus pore protein ChmB (gp2) and one subunit of each phage-encoded RNA polymerase: gp62 is part of the non-virion RNA polymerase, and gp77 is part of the virion RNA polymerase. We also observed association with 19 ribosomal and ribosome-associated proteins plus the host RNA polymerase α, β, and β′ subunits. The enrichment of ribosomal proteins and RNA polymerase subunits in this experiment suggests that ChmC may be a non-specific RNA binding protein. The observed association of ChmC with the putative pore-forming protein ChmB, however, suggests that these two proteins may functionally cooperate at the nuclear shell.

3.3.2. ChmC adopts a nucleic acid-binding fold and binds RNA

ChmC is conserved across jumbo phages but shows no identifiable sequence similarity to known proteins. To gain insight into ChmC's structure and potential function,

we used AlphaFold 2 [11] to predict its 3D structure with high confidence. (**Figure 2A-B, Figure S2A-B**). Analysis of the resulting model using the DALI protein structure comparison server revealed a predicted Whirly domain fold (also termed a PUR domain) common to multiple families of single-stranded RNA and DNA binding proteins[12–14]. The Whirly fold typically comprises a tandem repeat of β-β-β-β-α secondary structure elements and is exemplified by the *Trypanosoma brucei* MRP1 protein (**Figure 2A**) (PDB ID 2GJE;[13]). The predicted structure of PhiPA3 ChmC shows a tandem repeat of β-β-β-α elements, with an overall 3D structure highly reminiscent of the Whirly domain (Cα r.m.s.d. of 4.7 Å for gp61 versus *T. brucei* MRP1 over 120 residues) (**Figure 2A**).

Whirly domain proteins typically form higher-order complexes including homo- and heterotetramers, exemplified by the 2:2 heteroteramers formed by *T. brucei* MRP1 and MRP2 [13]. We used AlphaFold 2 to predict the structure of PhiPA3 ChmC oligomers, and obtained a confident prediction of a ChmC homotetramer that is strikingly similar to the *T. brucei* MRP1:MRP2 heterotetramer structure (**Figure 2B, Figure S2A, C**). We expressed and purified full-length PhiPA3 ChmC in *E. coli* and analyzed its oligomeric state by size exclusion chromatography coupled to multi-angle light scattering (SEC-MALS). Supporting our structure prediction, we found that ChmC forms a stable homotetramer in solution (**Figure 2C**).

Given the known roles of Whirly domain proteins in nucleic acid binding, we tested the ability of purified PhiPA3 ChmC to bind single-stranded DNA or RNA *in vitro*. Using a fluorescence polarization assay, we found that ChmC binds both DNA and RNA, but that the protein binds RNA with a higher affinity (K_d = 58 +/- 4 nM) than it binds DNA (184 +/-

11 nM) (**Figure 2D**). We modeled a ChmC-RNA complex by overlaying the PhiPA3 ChmC tetramer model onto the structure of *T. brucei* MRP1:MRP2 bound to RNA[13] (**Figure 2A, S2C**). Based on this model, we generated two multi-site mutants designed to disrupt nucleic acid binding, termed Kmut (K30A/K48A/K59A) and Rmut (R35A/R44A/R63A) (**Figure 2E-F, S2D**), and found that both mutants completely disrupt RNA and DNA binding (**Figure 2G, S2E**). Based on these data, we conclude that gp61 adopts a homotetramer of Whirly domain folds and binds RNA.

3.3.3. ChmC forms condensates with RNA

AlphaFold2 structure predictions of ChmC from different jumbo phages consistently showed low confidence scores (pLDDT) for the C-terminal ~50-70 residues of the protein, indicating that this region is likely disordered in solution. Across several ChmC orthologs, this region contains a ~40-residue region that is highly enriched in asparagine and glycine residues (N/G-rich), followed by a ~15-residue region enriched in serine, aspartate, and glutamate residues (**Figure 3A**). Disordered regions rich in asparagine and glutamine have been shown to promote macromolecular condensate formation, likely through dipole-dipole interactions [15,16]. Disordered regions rich in serine, aspartate, and glutamate, meanwhile, have been termed “electronegative clusters” (ENCs) which can stabilize RNA binding proteins in solution and suppress non-specific RNA binding [17] . Indeed, both the PSPredictor and catGRANULE algorithms strongly predicted that ChmC forms condensates, and that this propensity is driven by the protein’s C-terminal disordered region (**Figure S3A**). Together with our finding that ChmC binds RNA *in vitro*, these predictions suggested that ChmC might form macromolecular

condensates with RNA through multivalent RNA binding and low-affinity interactions through its C-terminal disordered region, similar to many other RNA binding proteins in both prokaryotes [18,19] and eukaryotes[20–22].

To directly test the propensity for PhiPA3 ChmC to form condensates, we engineered a cysteine residue at the protein's N-terminus to enable fluorescent labeling using a maleimide-linked Cy5 dye. In a low-salt buffer (50 mM NaCl), we observed that purified ChmC forms uniform small droplets, and that the addition of a 2.3 kb RNA with the major capsid protein (gp136) gene sequence stimulates formation of large droplets that showed hallmarks of liquid-liquid phase separation, including dynamic growth and fusion of droplets (**Figure 3B, S3B**). Disrupting ChmC's ability to bind RNA using the Kmut or Rmut multisite mutations dramatically reduced, but did not eliminate formation of condensates in the presence of RNA (**Figure 3B**). Since both the Kmut and Rmut proteins retain some positively-charged residues on the RNA binding surface, it is likely that these proteins retain some ability to bind RNA, albeit with lower affinity than wild-type ChmC. Removal of the C-terminal disordered region (residues 204-251 removed; ChmC ΔC), meanwhile, completely eliminated formation of condensates in both the absence and presence of RNA (**Figure 3B**). Importantly, the ChmC ΔC protein retained the ability to bind RNA in vitro (**Figure S3C-D**).

We next examined localization of ChmC mutants in PhiPA3-infected *P. aeruginosa* cells. The RNA binding mutant proteins did not lose the ability to localize to both the nuclear shell and the host cytoplasm in infected *P. aeruginosa* cells. A limitation of this assay is that the mutant proteins are expressed alongside wild-type phage-encoded

ChmC, and may form hetero-oligomers with wild-type ChmC in infected cells. Despite this limitation, both the Kmut and Rmut showed a localization pattern distinct from wild-type ChmC, retaining nuclear shell localization but showing more uniform cytoplasmic distribution than wild-type protein (**Figure 3C**). Strikingly, removal of the ChmC C-terminus completely disrupted nuclear shell binding activity, suggesting that the protein's ability to form condensates (with or without RNA) is important for the protein's localization to the nuclear shell. Together with our in vitro RNA binding and RNA-mediated condensation data, these localization data suggest that ChmC forms RNA-protein condensates associated with the nuclear shell, and potentially also in the host cytoplasm.

3.3.4. ChmC associates with viral mRNAs in infected cells

To test whether ChmC associates with RNA, particularly phage mRNAs, in infected cells, we performed eCLIP-seq (enhanced UV crosslinking and immunoprecipitation, followed by deep sequencing) in PhiPA3-infected *P. aeruginosa* cells expressing GFP-tagged ChmC. We obtained ~1.6M uniquely-mapping sequence reads for GFP-tagged PhiPA3 ChmC across two independent replicates (1,161,658 + 423,427 reads), and 1.5M reads for control GFP samples (1,070,428 + 423,357 reads). We mapped all reads to the host (*P. aeruginosa* PA01) and phage genomes (**Figure 4A, S4A**), then calculated the enrichment for each gene for immunoprecipitated samples compared to matched input samples. All samples were collected 45 minutes after phage infection, after the host genome is fully degraded and the phage genome has begun replicating [4,5,23].

Nucleus-forming jumbo phages encode two multi-subunit RNA polymerases: a "virion RNA polymerase" (vRNAP) that is packaged in the virion and is responsible for

transcription immediately after infection, and a “non-virion RNA polymerase” (nvRNAP) that is responsible for transcription at the middle and late stages of infection[24–26]. A prior RNA-seq analysis of the jumbo phage PhiKZ defined the operon structure of this phage and found that while vRNAP-transcribed early genes have a defined promoter sequence, genes transcribed by the nvRNAP do not show a reproducible promoter sequence [27]. This analysis also identified a number of noncoding antisense RNAs, which were proposed to regulate translation of phage genes [27]. Since a similar RNA-seq analysis has not been performed for PhiPA3, our analysis of gp61 eCLIP data is limited to annotated protein- and tRNA-encoding genes. We first analyzed the relative enrichment of phage mRNAs versus host mRNAs in the ChmC eCLIP immunoprecipitates versus input RNA. We found that phage mRNAs were slightly enriched (median $\log_2$(fold enrichment) of 0.186) compared to input samples, while host-encoded mRNAs were significantly depleted (median $\log_2$(fold enrichment) of -1.54) compared to input (**Figure 4B**). Overall, these data suggest that ChmC preferentially associates with phage mRNAs compared to host mRNAs, but that within phage mRNAs, ChmC shows little to no specificity.

We next visually inspected ChmC eCLIP sequence coverage on the PhiPA3 genome. While some highly expressed genes (judging from the abundance of sequence reads in eCLIP input samples) were bound across the entire open reading frame (e.g. the *chmC* gene itself; **Figure 4A**), we noticed that the majority of binding occurred in defined peaks near the start codons of genes. While the operon structure of PhiPA3 is not annotated, many peaks occurred near the start codons of genes that appeared to be within polycistronic mRNAs. That is, ChmC binding occurs not only near the 5’ end of an

mRNA (e.g. *chmA*; **Figure 4A**), but also likely occurs near internal start codons in mRNAs that encode multiple genes (e.g. gp57, gp59, and gp60; **Figure 4A**). We performed a metagene analysis and found that across annotated PhiPA3 genes, ChmC shows enriched binding in a ~50-bp region centered 25-30 bp downstream of the start codon (**Figure 4C**). We performed motif analysis to identify any defined sequences bound by ChmC, but were unable to unambiguously identify a preferred binding motif. Finally, we also observed high sequence coverage in short regions that did not correlate with annotated genes, including peaks between the *ChmA* (gp53) and gp54 genes (**Figure 4A**) and between gp203 and gp204 (**Figure S4A**). These data suggest that PhiPA3 encodes small regulatory RNAs similarly to the related jumbo phage PhiKZ [28], and that ChmC binds many of these RNAs in addition to binding protein-encoding mRNAs.

3.3.5. ChmC knockdown halts jumbo phage infections and decreases translation

The jumbo phage nucleus prevents Cas9-based targeting of the phage genome due to the physical barrier between the nucleus-enclosed phage DNA and host-encoded CRISPR/Cas enzymes [8,9]. Because phage mRNAs are transported into the host cytoplasm for translation, however, jumbo phages are susceptible to CRISPR-based mRNA knockdown strategies [8,9]. In a related work, we demonstrated that catalytically-dead Cas13d (ddCas13d) and a guide RNA targeting the 5' end of the *E. coli* jumbo phage Goslar ChmA gene efficiently prevents translation of ChmA and results in a near-total loss of phage nucleus formation (Armbruster et al, in preparation). To determine the biological roles of ChmC, we used a similar strategy to target the Goslar *ChmC* gene (gp176) (**Figure S5A**). Infection of *E. coli* MC1000 cells carrying a plasmid encoding

ddCas13d plus a guide RNA targeting *chmC* significantly reduced phage titer, with the most effective guide (guide 3) reducing the efficiency of plaquing to ~4% of the efficiency observed in cells encoding a non-targeting guide RNA (**Figure S5B-C**). Infected cells expressing ddCas13d and ChmC-targeting guide RNAs also showed significant reductions in phage bouquet formation, indicative of a failure to assemble new virions (**Figure S5D**), and showed smaller phage nuclei with reduced DNA content compared to control cells expressing a non-targeting guide RNA (**Figure 5A, S5E-G**). We could rescue these phenotypes by overexpressing ChmC with the gene's 5' end recorded to avoid ddCas13d-mediated knockdown (**Figure 5B, S6**). While we were unable to determine the extent of ChmC knockdown due to the lack of a specific anti-ChmC antibody, these data suggest that ChmC is efficiently knocked down and that this knockdown results in a strong defect in phase nucleus maturation and virion production.

We next imaged Goslar-infected *E. coli* cells expressing GFP-tagged ChmA (gp246) or the major capsid protein (gp41). Infected cells expressing ddCas13d and a non-targeting guide RNA showed characteristic expansion of the cell diameter around the developing phage nucleus, which was labeled by GFP-tagged ChmA (**Figure 5C, S7A**). Infected cells expressing ddCas13d and a *ChmC*-targeting guide RNA showed little to no expansion of the cell diameter, and phage nuclei were markedly smaller than in control cells (**Figure 5C, S7A**). Similarly, knockdown of ChmC caused infected cells to fail to form phage bouquets, and the major capsid protein was visibly attached to the nuclear shell, typical of early-stage infections when capsids localize to the phage nuclear shell for genomic DNA packaging (**Figure 5D, S7B**) [4,23]. Together, these data suggest that

ddCas13d-mediated ChmC knockdown significantly slows or halts Goslar infections at an early stage, preventing full maturation of the phage nucleus and production of viable phage progeny.

We next imaged GFP-tagged Goslar ChmC, both in the context of an unperturbed Goslar infection (**Figure 5E**) and with ChmC knocked down by ddCas13d (**Figure 5F**). In both cases, we observed that ChmC localizes to the phage nucleus. In contrast to PhiPA3 ChmC, which localizes across the nuclear shell, Goslar ChmC appears to form discrete puncta both within the nucleus and along the nuclear shell. These puncta could represent RNA-protein condensates, or alternatively they could represent hubs for translocation of phage mRNAs through the nuclear shell. We generated mutants of Goslar ChmC analogous to the PhiPA3 ChmC Rmut and ΔC mutants (**Figure S8A-B**), and imaged both in infected cells also expressing ddCas13d and either non-targeting or *ChmC*-targeting guide RNAs. Both ChmC Rmut and ΔC failed to rescue Goslar phage titer when overexpressed in cells alongside ddCas13d and a ChmC-targeting guide RNA, indicating that they cannot support the phage life cycle on their own (**Figure S7C-E**). In infected cells, both mutants localized as puncta within the phage nucleus in cells expressing a non-targeting guide RNA (**Figure 5G-H**). In cells expressing a *ChmC*-targeting guide RNA, however, ChmC Rmut localized as a single punctum that lacked any DNA (**Figure 5G**). We interpret this as indicative of a failure in nuclear shell assembly around the initially-injected phage genome, resulting in an aborted nuclear shell with no DNA content. In contrast, ChmC ΔC localized as puncta on the perimeter of the phage nucleus, rather than within the phage nucleus (**Figure 5H**). Since ChmC ΔC does not form RNA-protein

condensates in vitro, its ability to form puncta at the nuclear shell suggests that these puncta may represent hubs for mRNA translocation across the nuclear shell. Overall, these data support a model in which ChmC is important across *Chimalliviridae* to support proper assembly and maturation of the phage nuclear shell, and to support assembly of new virions.

3.3.6. ChmC knockdown results in a global reduction of phage protein levels

Our data on ChmC mRNA binding and localization are consistent with roles in stabilizing phage mRNAs and/or promoting their translocation through the phage nuclear shell for translation. We tested the effects of ChmC knockdown on global phage protein levels using TMT-tagged mass spectrometry to quantitatively compare phage protein levels in the presence of ddCas13d and either a non-targeting guide RNA or one of two *chmC*-targeting guide RNAs. Overall, we detected 180 of 247 annotated Goslar proteins in all three samples, 143 of which showed a reduction in protein levels after ChmC knockdown (**Figure 6A, Table S4**). Using a separate time-course mass spectrometry dataset of Goslar-infected cells (**Figure 6B, Table S5**), we divided the Goslar proteome into groups of proteins that first appear in early (30 minutes post infection), middle (60 minutes), or late (90 minutes) infections. We found that while ChmC knockdown did not strongly affect early-expressed genes, middle- and late-expressed genes were significantly suppressed (**Figure 6C**). This observation is consistent with our finding that ChmC knockdown significantly slows or halts Goslar infections at an early stage.

3.3.7. Discussion

The nucleus-like compartment assembled by *Chimalliviridae* to protect their genomes from DNA-targeting host immune systems like restriction enzymes and CRISPR/Cas nucleases introduces major challenges to the phage life cycle, principally the need to translocate mRNAs out of the phage nucleus and to translocate certain phage proteins into the nucleus. Here, we identify an abundant, early-expressed, nuclear shell-associated protein, ChmC, that is conserved across *Chimalliviridae* and is encoded in a conserved block of genes alongside several subunits of the phage's non-virion RNA polymerase. ChmC adopts a nucleic acid binding fold, binds phage mRNAs, and forms condensates with RNA in vitro. PhiPA3 ChmC localizes across the phage nuclear shell, while Goslar ChmC forms puncta that localize both within the phage nucleus and along the nuclear shell. Targeted knockdown of Goslar ChmC results in reduced phage nucleus size and a failure to form phage bouquets in infected cells, and results in a dramatic reduction in plaque formation. Together, these data show that ChmC plays crucial roles in the life cycle of *Chimalliviridae*, potentially aiding mRNA transcription within the phage nucleus, promoting mRNA translocation through the nuclear shell, and/or directly promoting translation in the cytoplasm of infected cells.

Our structure predictions and biochemical characterization indicate that ChmC adopts a Whirly domain fold, first identified in single-stranded DNA/RNA binding proteins in plants where they are involved in transcriptional responses to stress [14,29] . In bacteria, Whirly-related proteins are primarily involved in binding and compacting the nucleoid through their non-specific DNA binding activity [30–33]. ChmC preferentially binds RNA over

DNA *in vitro*, and shows a distinctive pattern of mRNA binding with enriched binding to many phage mRNAs near their start codons. While we do not detect any sequence motifs specifically recognized by ChmC, the protein may specifically recognize particular mRNA sequences or structures that determine this binding pattern. Overall, these findings demonstrate that while ChmC is structurally related to other Whirly domain proteins, it has adopted distinct regulatory roles in *Chimalliviridae*.

While bacteriophages generally do not encode RNA-binding proteins, diverse eukaryotic viruses encode RNA binding proteins with diverse roles in the viral life cycle. Many such proteins also form RNA-protein condensates like ChmC [34]. Coronaviruses like SARS, MERS, and SARS-CoV-2, encode an RNA-binding nucleocapsid (N) protein that promotes viral RNA production and suppresses host responses through its ability to form RNA-protein condensates, in addition to packaging viral genomic RNA into virions [35]. Influenza NEP (nuclear export protein) and HIV Rev are both RNA binding proteins that promote export of viral RNAs produced in the host-cell nucleus into the cytoplasm [36].

Our data show that ChmC is crucial for the proper progression of infection in *Chimalliviridae*, likely by aiding the translocation of phage mRNAs through the nuclear shell to promote translation. In Goslar-infected cells, ChmC forms puncta both within the phage nucleus and along the nuclear shell itself. These puncta are not simply RNA-protein condensates, since shell-associated puncta are also formed by ChmC ΔC, which cannot form condensates with RNA. We hypothesize that these puncta may represent hubs for the translocation of mRNAs through the nuclear shell. These hubs are likely built around shell-penetrating pores, potentially the same ChmB pores that mediate capsid

docking and genome packaging in late-stage infections (Enustun et al., submitted). The location of the *chmC* gene within a conserved block of genes encoding nvRNAP subunits further suggests a direct functional link between mRNA transcription and translocation through the nuclear shell, mediated by ChmC. Further work will be required to establish whether ChmB or other proteins are required for this process.

The phage nucleus of *Chimalliviridae* is a fascinating example of convergent evolution, representing a functional analog of the eukaryotic nucleus complete with physical segregation of the genome from the cytoplasm and specific mechanisms for mRNA export and protein import. Our work reveals the first identified RNA binding protein necessary for bacteriophage replication, with potential roles in mRNA production, export, and translation to support the unique life cycle of nucleus-forming jumbo phages.

3.4 Experimental Procedures

3.4.1 Bacterial strains, growth conditions and phage preparations

For PhiPA3 phage, *P. aeruginosa* K2733 (PA01 efflux pump knockout; (*ΔMexAB-OprMΔMexCD-OprJΔMexEF-OprNΔMexXY-OprM*)) was used as the host. For Goslar phage, *E. coli* MC1000 (derived from *E.coli* MG1655) was used as the host. Both bacterial strains were cultured in Luria-Bertani (LB) media or LB top agar (0.35% agar) at 30°C (*P. aeruginosa*) or 37°C (*E. coli*). To amplify phages, 100 µL of liquid culture at OD_{600}=0.6 was mixed with 20 µL of high-titer phage lysate and incubated at room temperature for 20 minutes, then the mixture was added to 5 mL of warm LB top agar, poured onto LB plates, and incubated overnight. The next day, 5 mL of Phage Buffer (10 mM Tris-HCl pH 7.5, 10 mM $MgSO_4$, 68 mM NaCl, and 1 mM $CaCl_2$) was added to each plate and

incubated at room temperature for 5 hours. The phage buffers were then collected, and lysates were centrifuged at 15,000 rpm for 10 minutes. Supernatants were stored and stored at 4°C with 0.01% chloroform.

3.4.2 Plasmid constructions and transformation

Genes of interests were PCR-amplified with 25 bp homology arms from high-titer phage lysates and ligated into respective plasmid backbones using NEBuilder HiFi DNA Assembly Cloning Kit (New England Biolabs # E5520S). Recombinant plasmids were transformed into *E. coli* DH5α and plated on LB agar containing appropriate antibiotics (25 μg/mL gentamicin sulfate, 100 μg/mL ampicillin, 100 μg/mL spectinomycin, or 100 μg/mL chloramphenicol). After plasmids were confirmed by DNA sequencing, chemically competent organisms of interest were transformed and selected on LB plates with relevant antibiotics. Selected colonies were grown in LB media with the antibiotics and stored in 25% glycerol at -80°C.

3.4.3 Fluorescence microscopy of single cell infections

1% agarose pads were prepared on concavity slides with desired inducing reagents of arabinose or IPTG. For imaging *P. aeruginosa*, pad mixes also contained FM4-64 (1 μg/mL) to stain cell membranes and DAPI (1 μg/mL) to stain DNA. Strains of interest were resuspended from overnight incubated LB plates into 25% LB to an OD_{600}= 0.3. 5 uL of resuspensions was spotted on concave slides and incubated in a humidor for 2 hours in 37°C (*E. coli* stains) or 30°C (*P. aeruginosa* strains). For phage infections, 10 μl of phages (10^{10} pfu/mL) were added to cells (resulting in a multiplicity of infection (MOI) of ~7) and incubated until desired time point. For imaging *E. coli*, dyes were added by

spotting 7 µl of the mix containing (2 µg/mL DAPI, 4 µg/mL FM4-64, 25% LB). The slides were sealed with a coverslip and fluorescent microscopy was performed using a DeltaVision Spectris Deconvolution Microscope (Applied Precision). Regions of interests were imaged using at least 8 Z-axis stacks from the middle focal plane in 0.15 µm increments. Final images were created by DeltaVision SoftWoRx Image Analysis Program and its deconvolution algorithm and analyzed by Fiji ImageJ.

3.4.4 Protein structure prediction

To model the structure of PhiPA3 or Goslar ChmC tetramers or PhiPA3 nvRNAP, we used AlphaFold multimer [11,37] using ColabFold [38] installed locally on a Linux workstation with NVIDIA RTX 3090GPU (https://github.com/YoshitakaMo/localcolabfold). Because of the large size of the PhiPA3 nvRNAP complex (2937 amino acids; 334 kDa), we could not model the entire complex in a single prediction due to GPU memory limits. Instead, we performed multiple predictions with different combinations of three subunits, then overlaid the resulting models on common subunits using PyMOL (Schrödinger, LLC) to generate the full complex. The final model agrees with the structures of the PhiKZ and *Bacillus* phage AR9 nvRNAP complexes [24,39].

3.4.5 Protein purification and characterization

For protein expression, *E. coli* strain Rosetta 2 (DE3) pLysS (EMD Millipore) cells were transformed with plasmids and grown overnight in LB plus appropriate antibiotics. The next day, cultures (1L 2XYT media plus antibiotics in 2L shaker flasks) were started and grown at 37°C until they reached an OD_{600} of 0.7, then induced with 0.25 mM IPTG and moved to 20°C for 16 hours. The cells were collected by centrifugation and

resuspended in a buffer containing 25 mM Tris-HCl pH 7.5, 10% glycerol, 1 mM NaN_3, 300 mM NaCl, 5 mM imidazole, and 5 mM β-mercaptoethanol. The proteins were purified using Ni^{2+} affinity chromatography (Ni-NTA agarose, Qiagen) and then passed over an anion-exchange column (Hitrap Q HP, Cytiva). Eluted fractions were concentrated and passed over a size exclusion column (Superdex 200, Cytiva) in GF buffer (Buffer A and 300 mM NaCl and 1 mM dithiothreitol). Fractions corresponding to the peak of interest were concentrated using ultrafiltration (Amicon Ultra, EMD Millipore) to reach a concentration of 10 mg/ml and stored at 4°C.

For analysis of molecular weight in solution using size exclusion chromatography coupled to multi-angle light scattering (SEC-MALS), 100 μl of purified protein at a concentration of 5 mg/ml was injected onto a size exclusion column (Superdex 200 Increase 10/300 GL, Cytiva) in GF buffer, then light scattering and refractive index profiles were collected using miniDAWN TREOS and Optilab T-rEX detectors (Wyatt Technology). SEC-MALS data were analyzed using ASTRA software version 8.

3.4.6 DNA and RNA binding assays

For measurement of DNA and RNA binding affinity by fluorescence polarization, 30 nM of a 22-base DNA (sequence ATTGTACCACTATTCCGAACAA) or RNA (sequence AUUGUACCACUAUUCCGAACAA) was mixed with the indicated concentration of purified PhiPA3 gp61 in FP buffer (20 mM HEPES pH 7.5, 75 mM KCl, 2 mM DTT, 5% glycerol, 0.02% NP40 substitute, 0.15 mg/mL BSA), and incubated for 10 minutes at room temperature. For RNA binding, reactions were supplemented with 1.25 mM RNaseOUT (ThermoFisher Scientific #10777019) Fluorescence polarization was

measured with a Tecan Infinite M1000 PRO fluorescence reader, and data was analyzed by GraphPad Prism using a cooperative binding model.

3.4.7 Knockdown of Phage proteins with dCas13d

31 nucleotides long guides that target the near ribosome binding side or translational start of the gene of interest were designed and cloned in the entry vectors with dCas13d (specifically RfxCas13d (*Ruminococcus flavefaciens* origin) with mutations R239A, H244A, R858A, H863A) (Provided by Adler et al., Doudna Lab). Cas13 variant is under tetR/tetA promoters while guides are expressed with J23119 promoter. Plasmid selection was performed with chloramphenicol while appropriate aTc concentration allows the expression of dCas13d. Plasmids were transformed to host MC1000 by electroporation and selected with 100 µg/mL chloramphenicol. For knockdown of ChmC at Goslar infections, strains were induced with 50 nM aTc and incubated at 37°C for 2 hours before the infection.

3.4.8 Mass spectrometry of phage infections

For *P. aeruginosa*, overnight bacterial cultures in LB media were diluted to OD_{600}=0.1 in fresh media, then further grown to OD_{600}=0.5. Regrown cultures were diluted 1:10 into 50 mL total volume in 250 mL flasks and grown in LB supplemented with 0.2 mM $CaCl_2$. Cells were infected with phage PhiPA3 at a multiplicity of infection (MOI) of 3 when they reached OD_{600}=0.3, then collected at the indicated time points. Cultures were pelleted by centrifugation at 4000 rpm at 4°C. Cell pellets were washed with 25% LB to remove any free phage. After the last wash, pellets were snap-frozen in liquid nitrogen and stored at -80°C.

For *E. coli*, overnight cultures were diluted to OD_{600}=0.1 in fresh media, then further grown to OD_{600}=0.6. 10 mL of pad mix (1% agarose, 25% LB, 50 nM anhydrotetracycline (aTc) and 30 ug/mL chloramphenicol) was poured into a 6 cm petri dish for each infection. When cultures reached desired OD_{600}, each strain was prepared as 200 μL of OD_{600}=0.1 and spread on the prepared pads. Petri dishes were placed in a 37°C humidor and incubated for 2 hours, then 100 μL of Goslar (titer 10^9 PFU/ml) was spread on the pad and incubated at 37°C for the indicated times. At the time of collection, 1 mL of 25% LB media was added and cells were carefully resuspended. Cells were pelleted by centrifugation at 4000 rpm at 4°C, washed with 25% LB, then snap-frozen and stored at -80°C for mass spectrometry.

3.4.9 Plaque Assays for Phage Infectivity

For plaque assays, 500 μl of saturated overnight bacterial cultures were mixed with 4.5 ml of 0.35% LB top agar. This mixture was poured on an LB plate that contains 100 µg/mL chloramphenicol and 50 nM aTc. After 30 minutes of incubation at room temperature for the mixture to dry, 3 μl of 10-fold serial dilutions of Goslar lysates (10^9 pfu/mL) were spotted on each plate. Plates were incubated at 37°C for 16 hours, then plaques were counted.

3.4.10 GFP pulldowns

GFP pulldowns were performed with GFP-Trap Magnetic Agarose beads (Proteintech #gtma-20). *P. aeruginosa* strains carrying pHERD30T plasmids expressing GFP-tagged proteins of interest were grown in LB with 25 µg/mL gentamicin sulfate. Saturated overnight cultures were diluted to OD_{600}=0.1 in fresh media, then grown at 30°C

until they reached OD_{600}=0.5. Cultures were diluted 1:10 in 50 mL LB supplemented with arabinose, gentamicin sulfate, and calcium chloride. Once the cells reached at OD_{600}=0.3, they were infected with PhiPA3 at MOI 3. The cultures were collected at 45 minutes post-infection, centrifuged, and the resulting cell pellets were stored at -80°C.

To perform the GFP pulldown, frozen cell pellets were thawed and incubated for 1 hour with 500 µL lysis buffer (10% glycerol, 25 mM Tris-HCl pH 7.5, 150 mM NaCl, 4 mg/mL lysozyme, 20 µg/mL DNase I, 2x cOmplete Protease Inhibitor, 0.4 mM PMSF). The cell suspensions were then sonicated (10 rounds x 20 pulses/round, Duty Cycle 40, Output 4), and the resulting lysed cells were centrifuged (30 minutes at 15,000 rpm at 4ºC). Beads were washed 3 times with a wash buffer (10 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5 mM EDTA) and added to the supernatant of the cell lysate. The mixture was rotated end-to-end for 1 hour at 4°C. The beads incubated in the cell lysate were washed 3 times with 1 mL wash buffer. After the last wash, beads were stored at -80°C for later analysis by SDS-PAGE and mass spectrometry.

3.4.11 eCLIP-Seq

Overnight bacterial cultures in LB media were diluted to OD_{600}=0.1 and grown to OD_{600}=0.5 at 30C. The cultures were diluted 1:10 into 50 mL total volume in 250 mL flasks and grown in LB supplemented with 0.2 mM $CaCl_2$ and 0.1% Arabinose. Cells were infected with phage PhiPA3 at a multiplicity of infection (MOI) of 3 when they reached OD_{600}=0.3. After 45 minutes of infection, cultures were collected and centrifuged with 4000 rpm at 4C for 8 minutes. Cells were washed with PBS 2 times and resuspended in 10mL fresh PBS. The samples were spread on 10cm petri-dish to coat the surface and

UV crosslinking was performed with 400 mJ/cm2 at 254 nm. Samples were collected from the petri dish and pelleted at 4C at 4000 rpm for 10 minutes and snap-frozen in liquid nitrogen to be stored at -80°C.

Each sample was treated with immunoprecipitation protocol with GFP-Trap beads. After the RNA-bound proteins were collected, cells were treated with FastAP (ThermoFisher) and T4 PNK (NEB), then barcoded RNA adapters were ligated to the 3′ end (T4 RNA Ligase, NEB). Samples were separated by SDS-PAGE and transferred to nitrocellulose membranes. The regions corresponding to the approximately expected size of sfGFP-alone and gp61-sfGFP were excised, and the membrane was suspended in a buffer with proteinase K (NEB). RNA isolation was performed with phenol/chloroform extraction and purified on spin columns (Zymo Research). Reverse-transcription was performed with AffinityScript (Agilent). cDNAs were treated with ExoSAP-IT (Affymetrix) to remove the excess oligonucleotides. Second DNA adapters (containing 5 [N5] or 10 [N10] random bases at the 5′-end) were ligated to the 5′-end of the cDNA (T4 RNA Ligase, NEB). The DNA was amplified by PCR and putified with PippinPrep system (Sage Science) and sequenced with Illumina HiSeq 4000. Libraries were analyzed for fragment size distribution on a D1000 Screentape (Agilent). Reads were processed and mapped to PhiPA3 genome. Normalizations of the eCLIP data were performed with the inputs without pulldown.

Reads were processed according to the protocol as previously described [40]. Briefly, sequenced reads from both IP and corresponding size-matched input (SMInput) were trimmed of adapters and mapped to repeat elements first to remove non-uniquely

mapped reads, then to the genome composed of both Pa_PA01 and PhiPA3 assemblies. PCR collapsing was then performed, and CLIPper [41] was used to call peak clusters on each set of these uniquely mapped, deduplicated (usable) reads. Annotations from [[gff3 annotation version]] were used to construct a custom CLIPper index. Reads within these clusters were normalized against the SMinput sample using scripts (available at: https://github.com/yeolab/eclip), and peaks found to be enriched above a $\log_2$(fold change) of 3 and $-\log_{10}$(Fisher Exact or Chi-square p-value) threshold of 3 were deemed significant and merged with IDR [42] to produce a set of reproducible peaks from replicates. Metagene plots were generated from usable reads, using a strategy laid forth by [43]. Briefly, read densities from SMInput samples were subtracted from corresponding IP signals across the set of phage genes, using SMInput data as a proxy for gene expression.

3.4.12 Mass Spectrometry

Frozen cell pellets were thawed and resuspended in 100 µL water. 10 µL of resuspended cells were mixed with 200 µL of 6M guanidine-HCl, vortexed and subjected to 3 cycles of 100°C for 5 minutes followed by cooling to room temperature. Boiled cell lysates were mixed with 1.8 mL of pure methanol and incubated at -20°C for 20 minutes. The mixture was centrifuged at 14000 rpm for 10 minutes at 4°C. All liquid was removed, and pellet was resuspended in 200 µL of 8 M urea in 0.2 M ammonium bicarbonate and incubated at 37°C for 1 hour with constant agitation. 4 µL of 500 mM TCEP (Tris(2-carboxyethyl) phosphine) and 20 µL 400 mM chloro-acetamide were added to the samples.

Protein concentration was measured by BCA assay and 600 µL of 200 mM ammonium bicarbonate was added to bring the urea concentration to 2 M. 1 µg of sequencing-grade trypsin was added for each 100 µg of protein in the sample and incubated at 42°C for overnight. Next day, 50 µL of 50% formic acid was added (tested for pH dropped to 2), then samples were desalted with C18 solid phase extraction (Waters Sep-Pak C18 12 cc Vac Cartridge # WAT036915) according to the manufacturer's protocol. 1 ml PBS buffer used for resuspension and peptide concentration of each sample was measured with BCA.

Trypsin-digested peptides were analyzed by ultra-high pressure liquid chromatography (UPLC) coupled with tandem mass spectroscopy (LC-MS/MS) using nano-spray ionization. Nanospray ionization was performed with Orbitrap fusion Lumos hybrid mass spectrometer (Thermo) interfaced with nano-scale reverse-phase UPLC (Thermo Dionex UltiMate 3000 RSLC nano System) using a 25 cm, 75-micron ID glass capillary packed with 1.7-μm C18 (130) BEH beads (Waters corporation). Peptides transferred from C18 column into the mass spectrometer by a linear gradient (5–80%) of ACN (Acetonitrile) From (Buffer A (98% H2O, 2% ACN, 0.1% formic acid) to Buffer B (100% ACN, 0.1% formic acid)) at a flow rate of 375 µl/min for 3 hours. Mass spectrometer parameters were; MS1 survey scan using the orbitrap detector (mass range (m/z): 400-1500 (using quadrupole isolation), 120000 resolution setting, spray voltage of 2200 V, Ion transfer tube temperature of 275°C, AGC target of 400000, and maximum injection time of 50 ms) which was followed by a data dependent scans (top speed for most intense ions, with charge state set to only include +2-5 ions, and 5 second exclusion time, while

selecting ions with minimal intensities of 50000 at in which the collision event was carried out in the high energy collision cell (HCD Collision Energy of 30%), and the fragment masses were analyzed in the ion trap mass analyzer (With ion trap scan rate of turbo, first mass m/z was 100, AGC Target 5000 and maximum injection time of 35 ms). Protein identification quantifications were carried out using Peaks Studio X (Bioinformatics solutions Inc).

3.4.13 TMT-tag Mass Spectrometry

Cell pellets from 200uL of culture were resuspended in 600 µl of 8M Urea in 100 mM Tris-HCl pH 8.0 and vortexed for 10 minutes. TCEP (tris(2-carboxyethyl)phosphine) was added to a final concentration of 10 mM, and samples were incubated at -20°C overnight to solubilize proteins. The next day, samples were vortexed until the solution was clear. Chloro-acetamide was added to the final concentration of 40 mM, and the mixture was vortexed for 5 minutes. Urea concentration was reduced to 4 M by adding an equal volume of 50 mM Tris-HCl pH 8.0 to the samples, then protein concentration was measured. LysC was added at 1:500 (w:w) LysC:protein ratio, and mixtures were incubated at 37°C on a roller for 6 hours. Urea concentration was further reduced to 2 M by adding 50 mM Tris-HCl pH 8.0. A 1:50 (w:w) ratio of trypsin was next added, and samples were incubated overnight. The next day, samples were acidified by adding TFA (trifluoroacetic acid) to 0.5% final concentration, and mixtures were vortexed for 5 minutes. Samples were centrifuged at 15,000 x g for 5 minutes to obtain aqueous and organic phases. The lower aqueous phase was collected and desalted with 100 mg C18-StageTips (Thermo Scientific) according to the manufacturer protocol. Samples were

resuspended in Thermo Fisher iTRAQ dissolution buffer and peptide concentrations were measured by BCA. For TMT labeling, TMTpro 16plex (Thermo Scientific, A44520) was used according to the manufacturer protocol. For high pH fractionation, High pH Reversed-Phase Peptide Fractionation Kit (Pierce #84868) was used according to the manufacturer protocol.

Each fraction was analyzed by ultra-high-pressure liquid chromatography (UPLC) coupled with tandem mass spectroscopy (LC-MS/MS) and nano-spray ionization with an Orbitrap fusion Lumos hybrid mass spectrometer (Thermo Fisher Scientific) interfaced with nano-scale reversed-phase UPLC (Thermo Fisher Scientific Dionex UltiMate 3000 RSLC nano System). In these experiments, a 25 cm, 75-micron ID glass capillary packed with 1.7-μm C18 (130) BEH beads (Waters) were used. From C18 into the column, a linear gradient (5–80%) of ACN (Acetonitrile) at a flow rate of 375 µl/min was used to elute peptides into the mass spectrometer for 180 minutes. The ACN gradient was created from Buffer A (98% H_2O, 2% ACN, 0.1% formic acid) to Buffer B (100% ACN, 0.1% formic acid). Parameters of MS1 survey scan using the orbitrap detector (mass range (m/z): 400-1500 (using quadrupole isolation), 60000 resolution setting, spray voltage of 2200 V, Ion transfer tube temperature of 275 C, AGC target of 400000, and maximum injection time of 50 ms) were used for mass spectrometer. This was followed by data dependent scans (top speed for most intense ions, with charge state set to only include +2-5 ions, and 5 second exclusion time). Ions with minimal 50000 intensity were selected with a high energy collision cell (HCD Collision Energy of 38%) and the first quadrupole isolation window was set at 0.7 (m/z). Orbi-trap mass analyzers were used

for analyzing the fragment masses (Ion trap scan rate of turbo, first mass m/z was 100, AGC Target 20000 and maximum injection time of 22ms). Peaks Studio X (Bioinformatic Solutions Inc.) was used for protein identification and quantification.

3.4.14 Condensate analysis

Macromolecular condensation assays were conducted in vitro using phase separation buffer (20 mM HEPES pH 7.4, 50 mM NaCl) at a temperature of 25°C, a protein concentration of 30 μM, and an RNA concentration of 2 μM (for samples including RNA). To prepare samples, unlabeled gp61 was pre-mixed with Cy5-labeled gp61 (linked to an engineered N-terminal cysteine using maleimide linkage) at a ratio of 1:10 and diluted to 60 μM in phase separation buffer. For samples without RNA co-incubation, an equal volume of phase separation buffer was added to each protein sample to reach the final working concentration of 30 μM. For samples with RNA co-incubation, an equal volume of RNA at 2X final concentration (4 μM for 40 base RNA; 166 nM for 2.3 kb RNA) in phase separation buffer was added and mixed gently. Samples were mixed in protein LoBind tubes (Eppendorf #022431064) and then immediately transferred into a 96-well non-binding plate (Greiner Bio-one #655906). Samples were imaged immediately after transfer into the 96-well plate using a CQ1 confocal quantitative image microscope (Yokogawa) with a 20x-PH objective. The fluorescent signal was captured under a laser at 640 nm. For live imaging, the entire field was automatically captured every 10 minutes. For quantitation, condensates were identified by particle analysis in ImageJ [44] . After thresholding, individual particles were counted, and their areas measured. For each

sample showing particles, the coefficient of variation was calculated as the standard deviation of particle area divided by the mean particle area.

3.5 Supplementary Tables

Table 3.S1. Mass spectrometry proteomics of PhiPA3-infected P. aeruginosa (all proteins; 30, 60, 90 MPI)

	Protein	Accession	% Coverage	# Peptides	Annotation
30 MPI	gp53	334738048	36.38	19	ChmA
	gp136	334738127	23.03	10	major capsid protein [Pseudomonas phage PhiPA3]
	gp222	334738213	44.12	6	hypothetical protein [Pseudomonas phage PhiPA3]
	gp61	334738056	13.49	3	hypothetical protein [Pseudomonas phage PhiPA3]
	gp28	334738023	13.33	3	PhuZ
	gp52	334738047	5.82	3	hypothetical protein [Pseudomonas phage PhiPA3]
	gp311	334738302	53.85	3	hypothetical protein [Pseudomonas phage PhiPA3]
	gp310	334738301	16.10	3	hypothetical protein [Pseudomonas phage PhiPA3]
	gp215	334738206	27.05	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp377	334738368	7.61	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp64	334738059	8.39	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp63	334738058	7.85	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp221	334738212	13.56	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp293	334738284	7.73	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp124	334738116	2.81	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp321	334738312	9.80	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp261	334738252	9.42	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp58	334738053	5.66	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp62	334738057	3.81	1	non-virion RNA Polymerase subunit [Pseudomonas phage PhiPA3]
	gp186	334738177	2.41	1	virion structural protein [Pseudomonas phage PhiPA3]
	gp142	334738133	7.10	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp122	334738114	3.69	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp378	334738369	1.58	1	NrdA [Pseudomonas phage PhiPA3]
	gp93	334738085	4.71	1	virion structural protein [Pseudomonas phage PhiPA3]
	gp75	334738067	9.22	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp1	334737995	11.11	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp2	334737996	9.14	1	ChmB
	gp274	334738265	2.31	1	putative thymidylate synthase [Pseudomonas phage PhiPA3]
	gp175	334738166	1.89	1	putative UvsX protein [Pseudomonas phage PhiPA3]
	gp185	334738176	4.29	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp223	334738214	4.70	1	thymidylate kinase [Pseudomonas phage PhiPA3]
	gp101	334738093	1.86	1	virion structural protein [Pseudomonas phage PhiPA3]

Table 3.S1. Mass spectrometry proteomics of PhiPA3-infected P. aeruginosa (all proteins; 30, 60, 90 MPI) (continued)

	Protein	Accession	% Coverage	# Peptides	Annotation
60 MPI	gp136	334738127	71.08	54	major capsid protein [Pseudomonas phage PhiPA3]
	gp53	334738048	59.30	39	ChmA
	gp186	334738177	48.28	23	virion structural protein [Pseudomonas phage PhiPA3]
	gp11	334738005	42.69	20	tail sheath protein [Pseudomonas phage PhiPA3]
	gp165	334738156	34.58	18	tail sheath protein [Pseudomonas phage PhiPA3]
	gp166	334738157	24.70	17	tail protein [Pseudomonas phage PhiPA3]
	gp61	334738056	61.90	15	hypothetical protein [Pseudomonas phage PhiPA3]
	gp79	334738071	35.50	14	hypothetical protein [Pseudomonas phage PhiPA3]
	gp208	334738199	28.07	12	hypothetical protein [Pseudomonas phage PhiPA3]
	gp222	334738213	86.03	11	hypothetical protein [Pseudomonas phage PhiPA3]
	gp168	334738159	52.17	10	putative tail fibre protein [Pseudomonas phage PhiPA3]
	gp377	334738368	33.86	8	hypothetical protein [Pseudomonas phage PhiPA3]
	gp28	334738023	33.65	8	PhuZ
	gp64	334738059	26.76	8	hypothetical protein [Pseudomonas phage PhiPA3]
	gp378	334738369	13.91	8	putative ribonucleotide-diphosphate reductase subunit alpha, NrdA [Pseudomonas phage PhiPA3]
	gp63	334738058	15.51	8	hypothetical protein [Pseudomonas phage PhiPA3]
	gp57	334738052	55.87	8	hypothetical protein [Pseudomonas phage PhiPA3]
	gp223	334738214	22.93	7	thymidylate kinase [Pseudomonas phage PhiPA3]
	gp176	334738167	31.56	7	virion structural protein [Pseudomonas phage PhiPA3]
	gp289	334738280	19.22	7	hypothetical protein [Pseudomonas phage PhiPA3]
	gp310	334738301	47.80	7	hypothetical protein [Pseudomonas phage PhiPA3]
	gp158	334738149	27.00	6	virion structural protein [Pseudomonas phage PhiPA3]
	gp93	334738085	24.24	6	virion structural protein [Pseudomonas phage PhiPA3]
	gp98	334738090	17.84	6	virion structural protein [Pseudomonas phage PhiPA3]
	gp74	334738066	17.28	6	hypothetical protein [Pseudomonas phage PhiPA3]
	gp221	334738212	53.39	5	hypothetical protein [Pseudomonas phage PhiPA3]
	gp100	334738092	14.34	5	virion structural protein [Pseudomonas phage PhiPA3]
	gp175	334738166	15.13	5	putative UvsX protein [Pseudomonas phage PhiPA3]
	gp267	334738258	33.04	4	hypothetical protein [Pseudomonas phage PhiPA3]
	gp293	334738284	33.70	4	hypothetical protein [Pseudomonas phage PhiPA3]
	gp52	334738047	11.45	4	hypothetical protein [Pseudomonas phage PhiPA3]
	gp89	334738081	16.46	4	virion structural protein [Pseudomonas phage PhiPA3]

Table 3.S1. Mass spectrometry proteomics of PhiPA3-infected P. aeruginosa (all proteins; 30, 60, 90 MPI) (continued)

	gp215	334738206	48.36	4	hypothetical protein [Pseudomonas phage PhiPA3]
	gp122	334738114	23.62	4	hypothetical protein [Pseudomonas phage PhiPA3]
	gp12	334738006	15.81	4	virion structural protein [Pseudomonas phage PhiPA3]
	gp164	334738155	39.22	4	virion structural protein [Pseudomonas phage PhiPA3]
	gp311	334738302	62.50	4	hypothetical protein [Pseudomonas phage PhiPA3]
	gp92	334738084	12.44	4	virion structural protein [Pseudomonas phage PhiPA3]
	gp102	334738094	5.86	4	virion structural protein [Pseudomonas phage PhiPA3]
	gp8	334738002	14.13	4	virion structural protein [Pseudomonas phage PhiPA3]
	gp150	334738141	9.49	4	virion structural protein [Pseudomonas phage PhiPA3]
	gp217	334738208	23.28	4	hypothetical protein [Pseudomonas phage PhiPA3]
	gp345	334738336	14.62	3	hypothetical protein [Pseudomonas phage PhiPA3]
	gp232	334738223	14.80	3	hypothetical protein [Pseudomonas phage PhiPA3]
	gp54	334738049	15.16	3	non-virion RNA Polymerase subunit [Pseudomonas phage PhiPA3]
	gp255	334738246	20.83	3	deoxycytidine triphosphate deaminase [Pseudomonas phage PhiPA3]
	gp2	334737996	28.93	3	hypothetical protein [Pseudomonas phage PhiPA3]
	gp239	334738230	19.64	3	hypothetical protein [Pseudomonas phage PhiPA3]
	gp58	334738053	22.17	3	hypothetical protein [Pseudomonas phage PhiPA3]
	gp49	334738044	11.23	3	hypothetical protein [Pseudomonas phage PhiPA3]
	gp274	334738265	7.14	3	putative thymidylate synthase [Pseudomonas phage PhiPA3]
	gp18	334738012	12.65	3	hypothetical protein [Pseudomonas phage PhiPA3]
	gp148	334738139	6.65	3	virion structural protein [Pseudomonas phage PhiPA3]
	gp142	334738133	16.57	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp300	334738291	10.53	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp233	334738224	1.80	2	putative SNF2 domain helicase [Pseudomonas phage PhiPA3]
	gp313	334738304	22.22	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp181	334738172	7.09	2	virion structural protein [Pseudomonas phage PhiPA3]
	gp321	334738312	18.95	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp135	334738126	13.33	2	virion structural protein [Pseudomonas phage PhiPA3]
	gp119	334738111	13.25	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp276	334738267	25.33	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp153	334738144	5.24	2	virion structural protein [Pseudomonas phage PhiPA3]
	gp187	334738178	7.91	2	virion structural protein [Pseudomonas phage PhiPA3]

Table 3.S1. Mass spectrometry proteomics of PhiPA3-infected P. aeruginosa (all proteins; 30, 60, 90 MPI) (continued)

	gp32	334738027	43.68	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp288	334738279	18.49	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp62	334738057	5.52	2	non-virion RNA Polymerase subunit [Pseudomonas phage PhiPA3]
	gp338	334738329	12.82	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp331	334738322	11.80	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp26	334738021	4.20	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp27	334738022	12.84	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp309	334738300	11.11	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp39	334738034	5.34	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp198	334738189	9.88	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp200	334738191	4.38	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp271	334738262	5.20	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp75	334738067	9.22	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp162	334738153	17.71	1	virion structural protein [Pseudomonas phage PhiPA3]
	gp256	334738247	11.87	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp326	334738317	10.27	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp247	334738238	11.41	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp108	334738100	7.50	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp95	334738087	3.65	1	virion structural protein [Pseudomonas phage PhiPA3]
	gp82	334738074	9.52	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp76	334738068	4.26	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp319	334738310	16.46	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp312	334738303	16.67	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp237	334738228	4.86	1	virion structural protein [Pseudomonas phage PhiPA3]
	gp226	334738217	5.15	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp195	334738186	15.27	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp155	334738146	6.06	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp127	334738119	19.42	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp121	334738113	12.79	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp87	334738079	12.24	1	virion structural protein [Pseudomonas phage PhiPA3]
	gp38	334738033	7.38	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp110	334738102	11.00	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp280	334738271	13.73	1	hypothetical protein [Pseudomonas phage PhiPA3]

Table 3.S1. Mass spectrometry proteomics of PhiPA3-infected P. aeruginosa (all proteins; 30, 60, 90 MPI) (continued)

	gp218	334738209	7.34	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp157	334738148	10.56	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp85	334738077	1.97	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp242	334738233	13.85	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp277	334738268	16.83	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp225	334738216	24.69	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp178	334738169	3.32	1	ribonuclease H [Pseudomonas phage PhiPA3]
	gp10	334738004	6.14	1	virion structural protein [Pseudomonas phage PhiPA3]
	gp370	334738361	7.05	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp86	334738078	3.38	1	virion structural protein [Pseudomonas phage PhiPA3]
	gp211	334738202	1.28	1	putative RNA polymerase beta subunit [Pseudomonas phage PhiPA3]
	gp351	334738342	5.00	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp272	334738263	9.92	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp306	334738297	26.50	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp240	334738231	7.32	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp131	334738123	7.69	1	putative helicase [Pseudomonas phage PhiPA3]
	gp281	334738272	5.09	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp22	334738016	4.69	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp205	334738196	4.69	1	virion structural protein [Pseudomonas phage PhiPA3]
	gp315	334738306	24.58	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp139	334738130	1.30	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp287	334738278	17.14	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp90	334738082	1.05	1	virion structural protein [Pseudomonas phage PhiPA3]
	gp301	334738292	6.19	1	hypothetical protein [Pseudomonas phage PhiPA3]

Table 3.S1. Mass spectrometry proteomics of PhiPA3-infected P. aeruginosa (all proteins; 30, 60, 90 MPI) (continued)

	Protein	Accession	% Coverage	# Peptides	Annotation
90 MPI	gp136	334738127	67.34	55	major capsid protein [Pseudomonas phage PhiPA3]
	gp53	334738048	59.63	52	ChmA
	gp186	334738177	41.03	21	virion structural protein [Pseudomonas phage PhiPA3]
	gp63	334738058	44.89	20	hypothetical protein [Pseudomonas phage PhiPA3]
	gp61	334738056	64.68	20	hypothetical protein [Pseudomonas phage PhiPA3]
	gp166	334738157	24.98	18	tail protein [Pseudomonas phage PhiPA3]
	gp378	334738369	30.45	17	putative ribonucleotide-diphosphate reductase subunit alpha, NrdA [Pseudomonas phage PhiPA3]
	gp11	334738005	34.38	15	tail sheath protein [Pseudomonas phage PhiPA3]
	gp175	334738166	42.86	14	putative UvsX protein [Pseudomonas phage PhiPA3]
	gp165	334738156	27.66	14	tail sheath protein [Pseudomonas phage PhiPA3]
	gp222	334738213	93.38	14	hypothetical protein [Pseudomonas phage PhiPA3]
	gp289	334738280	36.93	13	hypothetical protein [Pseudomonas phage PhiPA3]
	gp223	334738214	38.67	13	thymidylate kinase [Pseudomonas phage PhiPA3]
	gp168	334738159	60.00	13	putative tail fibre protein [Pseudomonas phage PhiPA3]
	gp28	334738023	45.71	13	PhuZ
	gp79	334738071	33.64	13	hypothetical protein [Pseudomonas phage PhiPA3]
	gp62	334738057	37.52	12	non-virion RNA Polymerase subunit [Pseudomonas phage PhiPA3]
	gp377	334738368	39.11	12	hypothetical protein [Pseudomonas phage PhiPA3]
	gp64	334738059	34.24	11	hypothetical protein [Pseudomonas phage PhiPA3]
	gp310	334738301	60.49	11	hypothetical protein [Pseudomonas phage PhiPA3]
	gp208	334738199	20.17	10	hypothetical protein [Pseudomonas phage PhiPA3]
	gp57	334738052	72.63	10	hypothetical protein [Pseudomonas phage PhiPA3]
	gp92	334738084	29.79	10	virion structural protein [Pseudomonas phage PhiPA3]
	gp345	334738336	50.77	10	hypothetical protein [Pseudomonas phage PhiPA3]
	gp74	334738066	16.91	10	hypothetical protein [Pseudomonas phage PhiPA3]
	gp102	334738094	12.90	9	virion structural protein [Pseudomonas phage PhiPA3]
	gp52	334738047	23.45	9	hypothetical protein [Pseudomonas phage PhiPA3]
	gp18	334738012	32.65	8	hypothetical protein [Pseudomonas phage PhiPA3]
	gp139	334738130	23.15	8	hypothetical protein [Pseudomonas phage PhiPA3]
	gp58	334738053	51.42	7	hypothetical protein [Pseudomonas phage PhiPA3]
	gp293	334738284	45.86	7	hypothetical protein [Pseudomonas phage PhiPA3]
	gp122	334738114	36.90	7	hypothetical protein [Pseudomonas phage PhiPA3]

Table 3.S1. Mass spectrometry proteomics of PhiPA3-infected P. aeruginosa (all proteins; 30, 60, 90 MPI) (continued)

	gp358	334738349	34.93	7	hypothetical protein [Pseudomonas phage PhiPA3]
	gp239	334738230	38.18	7	hypothetical protein [Pseudomonas phage PhiPA3]
	gp100	334738092	20.92	7	virion structural protein [Pseudomonas phage PhiPA3]
	gp54	334738049	21.52	6	non-virion RNA Polymerase subunit [Pseudomonas phage PhiPA3]
	gp93	334738085	24.24	6	virion structural protein [Pseudomonas phage PhiPA3]
	gp49	334738044	24.60	6	hypothetical protein [Pseudomonas phage PhiPA3]
	gp89	334738081	25.18	6	virion structural protein [Pseudomonas phage PhiPA3]
	gp158	334738149	27.67	6	virion structural protein [Pseudomonas phage PhiPA3]
	gp300	334738291	26.97	6	hypothetical protein [Pseudomonas phage PhiPA3]
	gp98	334738090	17.84	6	virion structural protein [Pseudomonas phage PhiPA3]
	gp355	334738346	30.74	6	hypothetical protein [Pseudomonas phage PhiPA3]
	gp176	334738167	23.75	6	virion structural protein [Pseudomonas phage PhiPA3]
	gp255	334738246	36.46	6	deoxycytidine triphosphate deaminase [Pseudomonas phage PhiPA3]
	gp200	334738191	25.90	5	hypothetical protein [Pseudomonas phage PhiPA3]
	gp346	334738337	27.97	5	hypothetical protein [Pseudomonas phage PhiPA3]
	gp8	334738002	17.39	5	virion structural protein [Pseudomonas phage PhiPA3]
	gp221	334738212	61.86	5	hypothetical protein [Pseudomonas phage PhiPA3]
	gp164	334738155	39.22	5	virion structural protein [Pseudomonas phage PhiPA3]
	gp124	334738116	17.42	5	hypothetical protein [Pseudomonas phage PhiPA3]
	gp215	334738206	54.92	5	hypothetical protein [Pseudomonas phage PhiPA3]
	gp306	334738297	71.79	5	hypothetical protein [Pseudomonas phage PhiPA3]
	gp217	334738208	32.80	5	hypothetical protein [Pseudomonas phage PhiPA3]
	gp195	334738186	45.04	5	hypothetical protein [Pseudomonas phage PhiPA3]
	gp1	334737995	53.85	5	hypothetical protein [Pseudomonas phage PhiPA3]
	gp47	334738042	9.36	5	hypothetical protein [Pseudomonas phage PhiPA3]
	gp181	334738172	13.96	4	virion structural protein [Pseudomonas phage PhiPA3]
	gp155	334738146	29.29	4	hypothetical protein [Pseudomonas phage PhiPA3]
	gp311	334738302	33.65	4	hypothetical protein [Pseudomonas phage PhiPA3]
	gp85	334738077	7.90	4	hypothetical protein [Pseudomonas phage PhiPA3]
	gp142	334738133	26.63	4	hypothetical protein [Pseudomonas phage PhiPA3]
	gp233	334738224	4.54	4	putative SNF2 domain helicase [Pseudomonas phage PhiPA3]
	gp232	334738223	19.28	4	hypothetical protein [Pseudomonas phage PhiPA3]
	gp119	334738111	31.79	4	hypothetical protein [Pseudomonas phage PhiPA3]

Table 3.S1. Mass spectrometry proteomics of PhiPA3-infected P. aeruginosa (all proteins; 30, 60, 90 MPI) (continued)

	gp178	334738169	12.45	4	ribonuclease H [Pseudomonas phage PhiPA3]
	gp313	334738304	37.96	4	hypothetical protein [Pseudomonas phage PhiPA3]
	gp288	334738279	48.74	4	hypothetical protein [Pseudomonas phage PhiPA3]
	gp10	334738004	15.35	4	virion structural protein [Pseudomonas phage PhiPA3]
	gp96	334738088	28.66	4	virion structural protein [Pseudomonas phage PhiPA3]
	gp347	334738338	7.68	3	hypothetical protein [Pseudomonas phage PhiPA3]
	gp12	334738006	15.46	3	virion structural protein [Pseudomonas phage PhiPA3]
	gp26	334738021	15.55	3	hypothetical protein [Pseudomonas phage PhiPA3]
	gp326	334738317	26.03	3	hypothetical protein [Pseudomonas phage PhiPA3]
	gp268	334738259	10.05	3	hypothetical protein [Pseudomonas phage PhiPA3]
	gp22	334738016	19.14	3	hypothetical protein [Pseudomonas phage PhiPA3]
	gp135	334738126	24.44	3	virion structural protein [Pseudomonas phage PhiPA3]
	gp356	334738347	31.71	3	hypothetical protein [Pseudomonas phage PhiPA3]
	gp351	334738342	21.82	3	hypothetical protein [Pseudomonas phage PhiPA3]
	gp338	334738329	21.79	3	hypothetical protein [Pseudomonas phage PhiPA3]
	gp315	334738306	44.92	3	hypothetical protein [Pseudomonas phage PhiPA3]
	gp276	334738267	48.00	3	hypothetical protein [Pseudomonas phage PhiPA3]
	gp2	334737996	28.93	3	ChmB
	gp172	334738163	21.88	3	virion structural protein [Pseudomonas phage PhiPA3]
	gp274	334738265	7.56	3	putative thymidylate synthase [Pseudomonas phage PhiPA3]
	gp225	334738216	46.91	3	hypothetical protein [Pseudomonas phage PhiPA3]
	gp150	334738141	7.05	3	virion structural protein [Pseudomonas phage PhiPA3]
	gp301	334738292	20.95	3	hypothetical protein [Pseudomonas phage PhiPA3]
	gp278	334738269	12.72	3	hypothetical protein [Pseudomonas phage PhiPA3]
	gp270	334738261	28.63	3	hypothetical protein [Pseudomonas phage PhiPA3]
	gp267	334738258	25.22	3	hypothetical protein [Pseudomonas phage PhiPA3]
	gp187	334738178	9.95	3	virion structural protein [Pseudomonas phage PhiPA3]
	gp39	334738034	13.06	3	hypothetical protein [Pseudomonas phage PhiPA3]
	gp331	334738322	26.09	3	hypothetical protein [Pseudomonas phage PhiPA3]
	gp277	334738268	39.60	3	hypothetical protein [Pseudomonas phage PhiPA3]
	gp153	334738144	8.81	3	virion structural protein [Pseudomonas phage PhiPA3]
	gp245	334738236	28.47	3	hypothetical protein [Pseudomonas phage PhiPA3]
	gp110	334738102	26.00	3	hypothetical protein [Pseudomonas phage PhiPA3]

Table 3.S1. Mass spectrometry proteomics of PhiPA3-infected P. aeruginosa (all proteins; 30, 60, 90 MPI) (continued)

	gp271	334738262	15.20	3	hypothetical protein [Pseudomonas phage PhiPA3]
	gp201	334738192	32.00	3	hypothetical protein [Pseudomonas phage PhiPA3]
	gp108	334738100	14.37	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp259	334738250	5.75	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp341	334738332	8.19	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp281	334738272	15.25	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp312	334738303	31.25	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp216	334738207	16.23	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp32	334738027	43.68	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp342	334738333	25.19	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp321	334738312	18.95	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp169	334738160	6.33	2	tail protein [Pseudomonas phage PhiPA3]
	gp257	334738248	4.26	2	putative DNA ligase [Pseudomonas phage PhiPA3]
	gp218	334738209	15.25	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp261	334738252	19.57	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp33	334738028	36.92	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp340	334738331	9.61	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp143	334738134	13.24	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp376	334738367	14.38	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp171	334738162	11.87	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp319	334738310	26.58	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp335	334738326	11.56	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp163	334738154	8.65	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp307	334738298	30.49	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp25	334738020	13.46	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp87	334738079	27.21	2	virion structural protein [Pseudomonas phage PhiPA3]
	gp370	334738361	13.46	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp78	334738070	4.79	2	putative endonuclease [Pseudomonas phage PhiPA3]
	gp308	334738299	10.00	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp272	334738263	19.08	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp205	334738196	7.22	2	virion structural protein [Pseudomonas phage PhiPA3]
	gp353	334738344	14.81	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp369	334738360	21.19	2	hypothetical protein [Pseudomonas phage PhiPA3]

Table 3.S1. Mass spectrometry proteomics of PhiPA3-infected P. aeruginosa (all proteins; 30, 60, 90 MPI) (continued)

	gp367	334738358	15.48	2	hypothetical protein [Pseudomonas phage PhiPA3]
	gp90	334738082	2.31	2	virion structural protein [Pseudomonas phage PhiPA3]
	gp280	334738271	13.73	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp296	334738287	16.67	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp309	334738300	11.11	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp38	334738033	7.38	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp75	334738067	9.22	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp104	334738096	2.97	1	virion structural protein [Pseudomonas phage PhiPA3]
	gp173	334738164	11.32	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp242	334738233	13.85	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp162	334738153	17.71	1	virion structural protein [Pseudomonas phage PhiPA3]
	gp213	334738204	0.56	1	tail fibre protein [Pseudomonas phage PhiPA3]
	gp256	334738247	11.87	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp207	334738198	6.14	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp9	334738003	2.43	1	virion structural protein [Pseudomonas phage PhiPA3]
	gp121	334738113	12.79	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp27	334738022	12.84	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp266	334738257	21.43	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp248	334738239	18.60	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp247	334738238	11.41	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp237	334738228	4.86	1	virion structural protein [Pseudomonas phage PhiPA3]
	gp234	334738225	9.38	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp198	334738189	9.88	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp97	334738089	4.31	1	virion structural protein [Pseudomonas phage PhiPA3]
	gp55	334738050	6.90	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp48	334738043	12.17	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp24	334738018	17.50	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp253	334738244	6.60	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp236	334738227	3.11	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp220	334738211	18.02	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp127	334738119	19.42	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp16	334738010	15.65	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp240	334738231	7.32	1	hypothetical protein [Pseudomonas phage PhiPA3]

Table 3.S1. Mass spectrometry proteomics of PhiPA3-infected P. aeruginosa (all proteins; 30, 60, 90 MPI) (continued)

	gp134	334738125	4.11	1	putative helicase [Pseudomonas phage PhiPA3]
	gp144	334738135	8.52	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp349	334738340	11.21	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp292	334738283	10.00	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp211	334738202	1.19	1	putative RNA polymerase beta subunit [Pseudomonas phage PhiPA3]
	gp336	334738327	9.66	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp76	334738068	4.26	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp329	334738320	3.85	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp40	334738035	8.03	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp154	334738145	3.83	1	virion structural protein [Pseudomonas phage PhiPA3]
	gp170	334738161	5.49	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp327	334738318	11.00	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp94	334738086	11.11	1	virion structural protein [Pseudomonas phage PhiPA3]
	gp273	334738264	10.84	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp360	334738351	2.26	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp190	334738181	1.75	1	SbcC protein [Pseudomonas phage PhiPA3]
	gp328	334738319	18.58	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp37	334738032	3.76	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp7	334738001	1.23	1	Terminase
	gp82	334738074	9.52	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp115	334738107	15.46	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp317	334738308	6.40	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp350	334738341	3.49	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp334	334738325	6.25	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp235	334738226	1.09	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp197	334738188	9.21	1	hypothetical protein [Pseudomonas phage PhiPA3]
	gp126	334738118	8.26	1	hypothetical protein [Pseudomonas phage PhiPA3]

Table 3.S2. Mass spectrometry proteomics of PhiPA3-infected P. aeruginosa (non-structural proteins)

Non-structural protein	Annotation	30MPI	60MPI	90MPI
gp1	hypothetical protein [Pseudomonas phage PhiPA3]	1	0	5
gp2	ChmB	1	3	3
gp7	Terminase	0	0	1
gp16	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	1
gp18	hypothetical protein [Pseudomonas phage PhiPA3]	0	3	8
gp22	hypothetical protein [Pseudomonas phage PhiPA3]	0	1	3
gp24	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	1
gp25	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	2
gp26	hypothetical protein [Pseudomonas phage PhiPA3]	0	1	3
gp27	hypothetical protein [Pseudomonas phage PhiPA3]	0	1	1
gp28	PhuZ	3	8	13
gp32	hypothetical protein [Pseudomonas phage PhiPA3]	0	2	2
gp33	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	2
gp37	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	1
gp38	hypothetical protein [Pseudomonas phage PhiPA3]	0	1	1
gp39	hypothetical protein [Pseudomonas phage PhiPA3]	0	1	3
gp40	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	1
gp47	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	5
gp48	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	1
gp49	hypothetical protein [Pseudomonas phage PhiPA3]	0	3	6
gp52	hypothetical protein [Pseudomonas phage PhiPA3]	3	4	9
gp53	ChmA	19	39	52
gp54	non-virion RNA Polymerase subunit [Pseudomonas phage PhiPA3]	0	3	6
gp55	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	1
gp57	hypothetical protein [Pseudomonas phage PhiPA3]	0	8	10
gp58	hypothetical protein [Pseudomonas phage PhiPA3]	1	3	7
gp61	hypothetical protein [Pseudomonas phage PhiPA3]	3	15	20
gp62	non-virion RNA Polymerase subunit [Pseudomonas phage PhiPA3]	1	2	12
gp63	hypothetical protein [Pseudomonas phage PhiPA3]	2	8	20
gp64	hypothetical protein [Pseudomonas phage PhiPA3]	2	8	11
gp74	hypothetical protein [Pseudomonas phage PhiPA3]	0	6	10
gp75	hypothetical protein [Pseudomonas phage PhiPA3]	1	1	1
gp76	hypothetical protein [Pseudomonas phage PhiPA3]	0	1	1

Table 3.S2. Mass spectrometry proteomics of PhiPA3-infected P. aeruginosa (non-structural proteins) (continued)

gp78	putative endonuclease [Pseudomonas phage PhiPA3]	0	0	2
gp79	hypothetical protein [Pseudomonas phage PhiPA3]	0	14	13
gp82	hypothetical protein [Pseudomonas phage PhiPA3]	0	1	1
gp85	hypothetical protein [Pseudomonas phage PhiPA3]	0	1	4
gp108	hypothetical protein [Pseudomonas phage PhiPA3]	0	1	2
gp110	hypothetical protein [Pseudomonas phage PhiPA3]	0	1	3
gp115	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	1
gp119	hypothetical protein [Pseudomonas phage PhiPA3]	0	2	4
gp121	hypothetical protein [Pseudomonas phage PhiPA3]	0	1	1
gp122	hypothetical protein [Pseudomonas phage PhiPA3]	1	4	7
gp124	hypothetical protein [Pseudomonas phage PhiPA3]	1	0	5
gp126	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	1
gp127	hypothetical protein [Pseudomonas phage PhiPA3]	0	1	1
gp131	putative helicase [Pseudomonas phage PhiPA3]	0	1	0
gp134	putative helicase [Pseudomonas phage PhiPA3]	0	0	1
gp139	non-virion RNA Polymerase subunit [Pseudomonas phage PhiPA3]	0	1	8
gp142	hypothetical protein [Pseudomonas phage PhiPA3]	1	2	4
gp143	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	2
gp144	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	1
gp155	hypothetical protein [Pseudomonas phage PhiPA3]	0	1	4
gp157	hypothetical protein [Pseudomonas phage PhiPA3]	0	1	0
gp163	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	2
gp170	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	1
gp171	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	2
gp173	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	1
gp175	putative UvsX protein [Pseudomonas phage PhiPA3]	1	5	14
gp178	ribonuclease H [Pseudomonas phage PhiPA3]	0	1	4
gp185	hypothetical protein [Pseudomonas phage PhiPA3]	1	0	0
gp190	SbcC protein [Pseudomonas phage PhiPA3]	0	0	1
gp195	hypothetical protein [Pseudomonas phage PhiPA3]	0	1	5
gp197	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	1
gp198	hypothetical protein [Pseudomonas phage PhiPA3]	0	1	1
gp200	hypothetical protein [Pseudomonas phage PhiPA3]	0	1	5
gp201	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	3
gp207	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	1
gp208	hypothetical protein [Pseudomonas phage PhiPA3]	0	12	10

Table 3.S2. Mass spectrometry proteomics of PhiPA3-infected P. aeruginosa (non-structural proteins) (continued)

gp211	putative RNA polymerase beta subunit [Pseudomonas phage PhiPA3]	0	1	1
gp215	hypothetical protein [Pseudomonas phage PhiPA3]	2	4	5
gp216	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	2
gp217	hypothetical protein [Pseudomonas phage PhiPA3]	0	4	5
gp218	hypothetical protein [Pseudomonas phage PhiPA3]	0	1	2
gp220	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	1
gp221	hypothetical protein [Pseudomonas phage PhiPA3]	1	5	5
gp222	hypothetical protein [Pseudomonas phage PhiPA3]	6	11	14
gp223	thymidylate kinase [Pseudomonas phage PhiPA3]	1	7	13
gp225	hypothetical protein [Pseudomonas phage PhiPA3]	0	1	3
gp232	hypothetical protein [Pseudomonas phage PhiPA3]	0	3	4
gp233	putative SNF2 domain helicase [Pseudomonas phage PhiPA3]	0	2	4
gp234	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	1
gp235	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	1
gp236	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	1
gp239	hypothetical protein [Pseudomonas phage PhiPA3]	0	3	7
gp240	hypothetical protein [Pseudomonas phage PhiPA3]	0	1	1
gp242	hypothetical protein [Pseudomonas phage PhiPA3]	0	1	1
gp245	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	3
gp247	hypothetical protein [Pseudomonas phage PhiPA3]	0	1	1
gp248	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	1
gp253	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	1
gp255	deoxycytidine triphosphate deaminase [Pseudomonas phage PhiPA3]	0	3	6
gp256	hypothetical protein [Pseudomonas phage PhiPA3]	0	1	1
gp257	putative DNA ligase [Pseudomonas phage PhiPA3]	0	0	2
gp259	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	2
gp261	hypothetical protein [Pseudomonas phage PhiPA3]	1	0	2
gp266	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	1
gp267	hypothetical protein [Pseudomonas phage PhiPA3]	0	4	3
gp268	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	3
gp270	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	3
gp271	hypothetical protein [Pseudomonas phage PhiPA3]	0	1	3
gp272	hypothetical protein [Pseudomonas phage PhiPA3]	0	1	2
gp273	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	1
gp274	putative thymidylate synthase [Pseudomonas phage PhiPA3]	1	3	3
gp276	hypothetical protein [Pseudomonas phage PhiPA3]	0	2	3

Table 3.S2. Mass spectrometry proteomics of PhiPA3-infected P. aeruginosa (non-structural proteins) (continued)

gp277	hypothetical protein [Pseudomonas phage PhiPA3]	0	1	3
gp278	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	3
gp280	hypothetical protein [Pseudomonas phage PhiPA3]	0	1	1
gp281	hypothetical protein [Pseudomonas phage PhiPA3]	0	1	2
gp287	hypothetical protein [Pseudomonas phage PhiPA3]	0	1	0
gp288	hypothetical protein [Pseudomonas phage PhiPA3]	0	2	4
gp289	hypothetical protein [Pseudomonas phage PhiPA3]	0	7	13
gp292	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	1
gp293	hypothetical protein [Pseudomonas phage PhiPA3]	1	4	7
gp296	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	1
gp300	hypothetical protein [Pseudomonas phage PhiPA3]	0	2	6
gp301	hypothetical protein [Pseudomonas phage PhiPA3]	0	1	3
gp306	hypothetical protein [Pseudomonas phage PhiPA3]	0	1	5
gp307	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	2
gp308	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	2
gp309	hypothetical protein [Pseudomonas phage PhiPA3]	0	1	1
gp310	hypothetical protein [Pseudomonas phage PhiPA3]	3	7	11
gp311	hypothetical protein [Pseudomonas phage PhiPA3]	3	4	4
gp312	hypothetical protein [Pseudomonas phage PhiPA3]	0	1	2
gp313	hypothetical protein [Pseudomonas phage PhiPA3]	0	2	4
gp315	hypothetical protein [Pseudomonas phage PhiPA3]	0	1	3
gp317	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	1
gp319	hypothetical protein [Pseudomonas phage PhiPA3]	0	1	2
gp321	hypothetical protein [Pseudomonas phage PhiPA3]	1	2	2
gp326	hypothetical protein [Pseudomonas phage PhiPA3]	0	1	3
gp327	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	1
gp328	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	1
gp329	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	1
gp331	hypothetical protein [Pseudomonas phage PhiPA3]	0	2	3
gp334	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	1
gp335	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	2
gp336	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	1
gp338	hypothetical protein [Pseudomonas phage PhiPA3]	0	2	3
gp340	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	2
gp341	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	2
gp342	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	2

Table 3.S2. Mass spectrometry proteomics of PhiPA3-infected P. aeruginosa (non-structural proteins) (continued)

gp345	hypothetical protein [Pseudomonas phage PhiPA3]	0	3	10
gp346	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	5
gp347	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	3
gp349	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	1
gp350	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	1
gp351	hypothetical protein [Pseudomonas phage PhiPA3]	0	1	3
gp353	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	2
gp355	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	6
gp356	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	3
gp358	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	7
gp360	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	1
gp367	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	2
gp369	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	2
gp370	hypothetical protein [Pseudomonas phage PhiPA3]	0	1	2
gp376	hypothetical protein [Pseudomonas phage PhiPA3]	0	0	2
gp377	hypothetical protein [Pseudomonas phage PhiPA3]	2	8	12
gp378	NrdA [Pseudomonas phage PhiPA3]	1	8	17

Table 3.S3. Mass spectrometry analysis of GFP-tagged gp61 in PhiPA3-infected P. aeruginosa

	Accession	-10lgP	Coverage (%)	Peak area	#Peptides	#Unique	#Spec Sample 6	Description
GFP-gp61	334738056	273.14	68	1730000000	28	28	394	gp61 [Pseudomonas phage PhiPA3]
	9950492	189.23	72	86200000	14	14	33	50S ribosomal protein L1 [Pseudomonas aeruginosa PAO1]
	9950488	180.55	25	62400000	27	27	33	DNA-directed RNA polymerase beta chain [Pseudomonas aeruginosa PAO1]
	9950477	178.09	61	664000000	22	22	43	50S ribosomal protein L2 [Pseudomonas aeruginosa PAO1]
	9949910	164.29	57	9130000	8	8	16	50S ribosomal protein L19 [Pseudomonas aeruginosa PAO1]
	9950479	152.45	52	32300000	11	11	20	50S ribosomal protein L4 [Pseudomonas aeruginosa PAO1]
	9950487	145.42	24	79400000	29	29	44	DNA-directed RNA polymerase beta* chain [Pseudomonas aeruginosa PAO1]
	9950814	140.33	47	12600000	6	6	12	50S ribosomal protein L21 [Pseudomonas aeruginosa PAO1]

Table 3.S3. Mass spectrometry analysis of GFP-tagged gp61 in PhiPA3-infected P. aeruginosa (continued)

	9950480	139.92	55	14100000	10	9	15	50S ribosomal protein L3 [Pseudomonas aeruginosa PAO1]
	9950493	139.73	64	496000000	11	11	48	50S ribosomal protein L11 [Pseudomonas aeruginosa PAO1]
	9946571	137.55	36	1350000000	129	129	225	hypothetical protein PA0690 [Pseudomonas aeruginosa PAO1]
	9945862	133.14	31	172000000	77	42	104	probable hemagglutinin [Pseudomonas aeruginosa PAO1]
	9948511	133.13	30	404000000	119	84	164	hypothetical protein PA2462 [Pseudomonas aeruginosa PAO1]
	9950496	132.54	29	16900000	10	10	13	elongation factor Tu [Pseudomonas aeruginosa PAO1]
	9950483	132.54	29	16900000	10	10	13	elongation factor Tu [Pseudomonas aeruginosa PAO1]
	9951114	116.98	39	2220000000	28	28	107	probable chemotaxis transducer [Pseudomonas aeruginosa PAO1]
	9950461	114.76	40	12900000	7	7	9	50S ribosomal protein L15 [Pseudomonas aeruginosa PAO1]
	9950454	106.62	38	60500000	6	6	14	50S ribosomal protein L17 [Pseudomonas aeruginosa PAO1]
	9950490	104.62	23	33400000	3	3	6	50S ribosomal protein L7 / L12 [Pseudomonas aeruginosa PAO1]
	9950455	102.45	18	82700000	6	6	19	DNA-directed RNA polymerase alpha chain [Pseudomonas aeruginosa PAO1]
	9950448	97.54	11	3880000	2	2	5	single-stranded DNA-binding protein [Pseudomonas aeruginosa PAO1]
	9947019	87.81	42	51200000	12	12	14	flagellar motor switch protein FliG [Pseudomonas aeruginosa PAO1]
	9946316	75.64	62	1910000	3	3	3	probable cold-shock protein [Pseudomonas aeruginosa PAO1]
	9950503	66.08	22	265000000	9	9	11	exodeoxyribonuclease V alpha chain [Pseudomonas aeruginosa PAO1]
	334737996	65.56	10	1410000	2	2	2	ChmB protein gp2 [Pseudomonas phage PhiPA3]
	9950456	64.52	13	988000	2	2	4	30S ribosomal protein S4 [Pseudomonas aeruginosa PAO1]
	9948941	56.06	12	341000	1	1	1	Outer membrane lipoprotein Oprl precursor [Pseudomonas aeruginosa PAO1]
	9950473	55.06	7	2860000	1	1	3	50S ribosomal protein L16 [Pseudomonas aeruginosa PAO1]
	9949775	53.09	5	139000	1	1	1	RecA protein [Pseudomonas aeruginosa PAO1]

Table 3.S3. Mass spectrometry analysis of GFP-tagged gp61 in PhiPA3-infected P. aeruginosa (continued)

	9950491	50.99	7	1200000	1	1	1	50S ribosomal protein L10 [Pseudomonas aeruginosa PAO1]
	9951633	50.2	13	1580000	1	1	5	50S ribosomal protein L28 [Pseudomonas aeruginosa PAO1]
	9951006	49.31	9	1890000	4	4	4	IF-2 [Pseudomonas aeruginosa PAO1]
	9949627	48.36	3	139000000	2	2	31	methionyl-tRNA synthetase [Pseudomonas aeruginosa PAO1]
	9948171	47.17	5	51400000	2	2	12	1 4-alpha-glucan branching enzyme [Pseudomonas aeruginosa PAO1]
	9950239	46.35	27	1200000	4	4	5	geranyltranstransferase [Pseudomonas aeruginosa PAO1]
	9948393	46.16	20	2300000	4	4	5	methanesulfonate sulfonatase MsuD [Pseudomonas aeruginosa PAO1]
	9951081	45.95	7	18000000	3	3	3	2 4-dienoyl-CoA reductase FadH2 [Pseudomonas aeruginosa PAO1]
	9951384	45.83	3	29900000	2	2	4	conserved hypothetical protein [Pseudomonas aeruginosa PAO1]
	9950465	44.65	8	557000	1	1	2	50S ribosomal protein L6 [Pseudomonas aeruginosa PAO1]
	9950615	43.16	8	862000	3	3	3	GroEL protein [Pseudomonas aeruginosa PAO1]
	9949598	42.11	3	5760000	1	1	5	hypothetical protein PA3456 [Pseudomonas aeruginosa PAO1]
	9950906	42.11	4	11500000	3	3	5	hypothetical protein PA4652 [Pseudomonas aeruginosa PAO1]
	9947552	41.06	6	36200	1	1	1	succinyl-CoA synthetase beta chain [Pseudomonas aeruginosa PAO1]
	9948410	40.61	4	93300	3	3	3	ClpA/B-type protease [Pseudomonas aeruginosa PAO1]
	9949113	40.56	4	13800000	3	3	6	DNA topoisomerase I [Pseudomonas aeruginosa PAO1]
	9947007	40.47	5	24400000	5	5	5	FgtA [Pseudomonas aeruginosa PAO1]
	9949064	40.45	21	0	1	1	1	acyl carrier protein [Pseudomonas aeruginosa PAO1]
	9948756	39.64	2	4750000	2	2	2	conserved hypothetical protein [Pseudomonas aeruginosa PAO1]
	9948081	39.54	2	245000	1	1	1	elongation factor G [Pseudomonas aeruginosa PAO1]
	9947487	39.28	6	38900000	3	3	4	DNA ligase [Pseudomonas aeruginosa PAO1]
	9949423	38.01	9	2200000	3	3	3	alkaline phosphatase [Pseudomonas aeruginosa PAO1]

Table 3.S3. Mass spectrometry analysis of GFP-tagged gp61 in PhiPA3-infected P. aeruginosa (continued)

	9946472	37.93	9	11700000	3	3	3	hypothetical protein PA0599 [Pseudomonas aeruginosa PAO1]
	9950505	36.99	4	6680000	3	3	3	exodeoxyribonuclease V gamma chain [Pseudomonas aeruginosa PAO1]
	9948819	36.81	8	0	1	1	1	50S ribosomal protein L20 [Pseudomonas aeruginosa PAO1]
	9950723	36.8	34	0	2	2	2	conserved hypothetical protein [Pseudomonas aeruginosa PAO1]
	9950519	36.77	10	140000000	3	3	3	TadG [Pseudomonas aeruginosa PAO1]
	9948899	36.64	3	3930000	2	2	4	probable acyl-CoA dehydrogenase [Pseudomonas aeruginosa PAO1]
	9950468	36.63	9	441000	1	1	1	50S ribosomal protein L5 [Pseudomonas aeruginosa PAO1]
	9948547	36.59	8	14200000	3	3	6	Multidrug efflux outer membrane protein OprN precursor [Pseudomonas aeruginosa PAO1]
	9945874	36.55	12	4610000	4	4	5	potential phenazine-modifying enzyme [Pseudomonas aeruginosa PAO1]
	9950485	36.14	8	211000	1	1	2	30S ribosomal protein S7 [Pseudomonas aeruginosa PAO1]
	9948236	36.07	6	2380000	2	2	2	probable porin [Pseudomonas aeruginosa PAO1]
	9951270	35.95	6	173000000	2	2	4	probable oxidoreductase [Pseudomonas aeruginosa PAO1]
	9949894	35.66	1	22700000	3	3	6	hypothetical protein PA3728 [Pseudomonas aeruginosa PAO1]
	9948399	35.51	4	18900000	3	3	3	hypothetical protein PA2361 [Pseudomonas aeruginosa PAO1]
	9948118	35.38	9	294000000	2	2	7	probable cysteine synthase [Pseudomonas aeruginosa PAO1]
	9950605	35.27	8	4630000	3	3	3	nicotinate phosphoribosyltransferase [Pseudomonas aeruginosa PAO1]
	9950371	35.02	3	124000000	1	1	10	hypothetical protein PA4163 [Pseudomonas aeruginosa PAO1]
	9948724	34.73	2	21500000	1	1	3	probable chemotaxis transducer [Pseudomonas aeruginosa PAO1]
	9947419	34.73	9	0	1	1	1	hypothetical protein PA1466 [Pseudomonas aeruginosa PAO1]
	9948445	34.57	1	0	4	2	4	probable non-ribosomal peptide synthetase [Pseudomonas aeruginosa PAO1]

Table 3.S3. Mass spectrometry analysis of GFP-tagged gp61 in PhiPA3-infected P. aeruginosa (continued)

	9949844	34.41	7	0	1	1	1	hypothetical protein PA3681 [Pseudomonas aeruginosa PAO1]
	9949226	34.4	5	8460000	3	3	4	Motility protein FimV [Pseudomonas aeruginosa PAO1]
	9951585	34.37	5	789000	3	3	4	adenylate cyclase [Pseudomonas aeruginosa PAO1]
	9947346	34.35	4	1080000	3	3	3	probable pyruvate carboxylase [Pseudomonas aeruginosa PAO1]
	334738057	34.22	23	2130000	1	1	1	non-virion RNA Polymerase subunit gp62 [Pseudomonas phage PhiPA3]
	9947602	33.68	3	122000	2	2	3	potassium-transporting ATPase B chain [Pseudomonas aeruginosa PAO1]
	9946565	33.27	5	114000	3	3	3	probable type II secretion system protein [Pseudomonas aeruginosa PAO1]
	9951491	33.22	7	7080000	3	2	3	probable acyl-CoA dehydrogenase [Pseudomonas aeruginosa PAO1]
	9950440	32.76	4	7940000	3	3	3	pyochelin synthetase [Pseudomonas aeruginosa PAO1]
	9947914	32.61	7	330000	3	3	3	class III (anaerobic) ribonucleoside-triphosphate reductase subunit NrdD [Pseudomonas aeruginosa PAO1]
	9949872	32.6	8	62600	2	2	2	probable chemotaxis transducer [Pseudomonas aeruginosa PAO1]
	9951828	32.52	6	3690000	3	3	3	DNA polymerase I [Pseudomonas aeruginosa PAO1]
	9945901	32.51	4	5570000	2	2	2	IcmF1 [Pseudomonas aeruginosa PAO1]
	9946450	32.5	17	143000	1	1	1	30S ribosomal protein S21 [Pseudomonas aeruginosa PAO1]
	9945965	32.2	5	53000000	2	2	2	probable ATP-binding component of ABC transporter [Pseudomonas aeruginosa PAO1]
	334738204	32.12	1	2730000	2	2	2	tail fibre protein [Pseudomonas phage PhiPA3]
	9947497	32.11	6	261000000	2	2	6	probable flavin-containing monooxygenase [Pseudomonas aeruginosa PAO1]
	9951539	31.99	1	1740000	1	1	2	probable ATP-binding/permease fusion ABC transporter [Pseudomonas aeruginosa PAO1]
	9949397	31.83	2	2940000	2	2	2	probable ATP-dependent DNA helicase [Pseudomonas aeruginosa PAO1]

Table 3.S3. Mass spectrometry analysis of GFP-tagged gp61 in PhiPA3-infected P. aeruginosa (continued)

	9949281	31.7	10	747000	2	2	2	histidinol-phosphate aminotransferase [Pseudomonas aeruginosa PAO1]
	9948469	31.65	1	237000	3	3	3	PvdL [Pseudomonas aeruginosa PAO1]
	9945820	31.5	3	1060000000	1	1	20	DNA polymerase III beta chain [Pseudomonas aeruginosa PAO1]
	9947830	31.3	1	9370000	2	2	4	methionine synthase [Pseudomonas aeruginosa PAO1]
	9950471	31.08	9	0	1	1	1	30S ribosomal protein S17 [Pseudomonas aeruginosa PAO1]
	9951844	31.07	5	27300000	1	1	3	probable glutamine synthetase [Pseudomonas aeruginosa PAO1]
	9946843	30.84	9	217000	2	2	2	hypothetical protein PA0938 [Pseudomonas aeruginosa PAO1]
	9950568	30.81	6	4080000	2	2	3	probable amidase [Pseudomonas aeruginosa PAO1]
	9949928	30.77	5	17400000	1	1	1	probable aminotransferase [Pseudomonas aeruginosa PAO1]
	9947479	30.57	4	136000000	1	1	6	hypothetical protein PA1522 [Pseudomonas aeruginosa PAO1]
	9947749	30.51	6	46900000	2	2	3	phosphoenolpyruvate synthase [Pseudomonas aeruginosa PAO1]
	9948572	30.16	9	3790000	2	2	4	toluate 1 2-dioxygenase alpha subunit [Pseudomonas aeruginosa PAO1]
	9950133	30.14	2	21600000	2	2	3	probable two-component sensor [Pseudomonas aeruginosa PAO1]
	334738069	30.05	5	110000	1	1	1	virion RNA polymerase subunit gp77 [Pseudomonas phage PhiPA3]
	9951907	30.05	13	0	1	1	1	PA5566 [Pseudomonas aeruginosa PAO1]
	9947858	29.74	4	1910000	2	2	3	secretion protein XqhA [Pseudomonas aeruginosa PAO1]
	110227056	29.47	2	36400	3	1	3	PvdJ [Pseudomonas aeruginosa PAO1]
	9947761	29.35	4	124000000	2	2	5	assimilatory nitrite reductase large subunit [Pseudomonas aeruginosa PAO1]
	9950137	29.34	5	18000000	1	1	3	PA3949 [Pseudomonas aeruginosa PAO1]
	9950504	29.28	2	959000	2	2	2	exodeoxyribonuclease V beta chain [Pseudomonas aeruginosa PAO1]
	334738369	29.16	6	368000	2	2	2	NrdA [Pseudomonas phage PhiPA3]

Table 3.S3. Mass spectrometry analysis of GFP-tagged gp61 in PhiPA3-infected P. aeruginosa (continued)

	9948691	29.08	3	1260000	2	2	2	isocitrate dehydrogenase [Pseudomonas aeruginosa PAO1]
	9948379	29.03	5	1670000	2	2	2	xylulose kinase [Pseudomonas aeruginosa PAO1]
	9946330	29.01	10	179000	1	1	1	hypothetical protein PA0469 [Pseudomonas aeruginosa PAO1]
	9949683	28.93	2	0	1	1	1	probable oxidoreductase [Pseudomonas aeruginosa PAO1]
	9946268	28.77	1	65800000	2	2	2	component of chemotactic signal transduction system [Pseudomonas aeruginosa PAO1]
	9946399	28.66	10	26800	1	1	1	hypothetical protein PA0532 [Pseudomonas aeruginosa PAO1]
	9949700	28.57	7	3060000	2	2	2	alginate o-acetyltransferase AlgJ [Pseudomonas aeruginosa PAO1]
	9948517	28.41	4	500000	1	1	1	Anti-sigma factor FoxR [Pseudomonas aeruginosa PAO1]
	9947759	28.36	4	1850000	2	2	2	assimilatory nitrate reductase [Pseudomonas aeruginosa PAO1]
	9949900	28.3	4	113000	1	1	1	hypothetical protein PA3733 [Pseudomonas aeruginosa PAO1]
	9951242	28.25	7	23100000	2	2	6	probable phosphoserine phosphatase [Pseudomonas aeruginosa PAO1]
	9947866	28.14	5	1630000	2	2	2	probable ATP-binding/permease fusion ABC transporter [Pseudomonas aeruginosa PAO1]
	9948583	28	2	1070000	2	2	2	RND efflux transporter [Pseudomonas aeruginosa PAO1]
	9947410	27.79	4	10600000	2	2	2	probable two-component sensor [Pseudomonas aeruginosa PAO1]
	9947854	27.77	6	0	2	2	2	hypothetical protein PA1865 [Pseudomonas aeruginosa PAO1]
	9951896	27.6	3	274000	1	1	1	ATP synthase alpha chain [Pseudomonas aeruginosa PAO1]
	9950160	27.4	4	0	1	1	1	AMP nucleosidase [Pseudomonas aeruginosa PAO1]
	9946915	27.28	6	7700000	2	2	2	conserved hypothetical protein [Pseudomonas aeruginosa PAO1]
	9948212	27.18	9	48800000	2	2	5	adenylate cyclase ExoY [Pseudomonas aeruginosa PAO1]
	9951291	27.01	5	1520000	2	2	2	probable carbamoyl transferase [Pseudomonas aeruginosa PAO1]

Table 3.S3. Mass spectrometry analysis of GFP-tagged gp61 in PhiPA3-infected P. aeruginosa (continued)

	9950799	26.88	2	271000	2	2	2	PilY1 [Pseudomonas aeruginosa PAO1]
	9951654	26.83	5	0	1	1	4	conserved hypothetical protein [Pseudomonas aeruginosa PAO1]
	9951723	26.75	2	0	1	1	1	DgcA Dimethylglycine catabolism [Pseudomonas aeruginosa PAO1]
	9947203	26.71	1	0	1	1	1	PA1270 [Pseudomonas aeruginosa PAO1]
	9950376	26.7	5	5010000	2	2	2	second ferric pyoverdine receptor FpvB [Pseudomonas aeruginosa PAO1]
	9951832	26.53	1	4410000	1	1	3	NrdJa [Pseudomonas aeruginosa PAO1]
	9950673	26.33	4	117000	1	1	1	tryptophanyl-tRNA synthetase [Pseudomonas aeruginosa PAO1]
	9950741	26.3	8	422000	2	2	2	Glycine-glutamate dipeptide porin OpdP [Pseudomonas aeruginosa PAO1]
	9950713	26.25	2	3260000	2	2	3	hypothetical protein PA4476 [Pseudomonas aeruginosa PAO1]
	9945871	26.21	4	35300000	2	2	3	PA0049 [Pseudomonas aeruginosa PAO1]
	9947807	26.17	6	18900000	2	2	2	PA1822 [Pseudomonas aeruginosa PAO1]
	9947697	26.16	18	0	1	1	1	type III export protein PscI [Pseudomonas aeruginosa PAO1]
	9947090	26.09	4	8110000	1	1	3	hypothetical protein PA1166 [Pseudomonas aeruginosa PAO1]
	9947803	26.05	1	43000000	1	1	10	lysine-specific pyridoxal 5'-phosphate-dependent carboxylase LdcA [Pseudomonas aeruginosa PAO1]
	9948556	25.89	5	6480000	1	1	1	hypothetical protein PA2503 [Pseudomonas aeruginosa PAO1]
	9949819	25.87	3	614000000	2	2	4	protein-PII uridylyltransferase [Pseudomonas aeruginosa PAO1]
	9947072	25.84	3	0	1	1	1	pyocin S2 [Pseudomonas aeruginosa PAO1]
	9951560	25.77	8	177000	1	1	1	conserved hypothetical protein [Pseudomonas aeruginosa PAO1]
	9950958	25.74	4	0	2	2	2	penicillin-binding protein 1B [Pseudomonas aeruginosa PAO1]
	9949103	25.64	1	7860000	1	1	1	transcription-repair coupling protein Mfd [Pseudomonas aeruginosa PAO1]
	9950289	25.57	8	0	1	1	1	probable short-chain dehydrogenase [Pseudomonas aeruginosa PAO1]

Table 3.S3. Mass spectrometry analysis of GFP-tagged gp61 in PhiPA3-infected P. aeruginosa (continued)

	9948609	25.57	5	20700000	2	2	3	probable acyl-CoA dehydrogenase [Pseudomonas aeruginosa PAO1]
	9948954	25.5	8	1710000	2	2	2	probable glycosylase [Pseudomonas aeruginosa PAO1]
	9947670	25.45	2	52000	1	1	1	ATP synthase in type III secretion system [Pseudomonas aeruginosa PAO1]
	9946027	25.39	4	0	2	2	2	probable TonB-dependent receptor [Pseudomonas aeruginosa PAO1]
	9950629	25.39	3	0	1	1	1	probable two-component sensor [Pseudomonas aeruginosa PAO1]
	9948869	25.18	8	24700000	2	2	2	carboxypeptidase G2 precursor [Pseudomonas aeruginosa PAO1]
	9947191	25.11	4	80800	1	1	1	hypothetical protein PA1259 [Pseudomonas aeruginosa PAO1]
	9950931	25.06	3	0	1	1	1	probable TonB-dependent receptor [Pseudomonas aeruginosa PAO1]
	9950104	24.91	2	103000	1	1	2	probable metal transporting P-type ATPase [Pseudomonas aeruginosa PAO1]
	9947317	24.84	6	234000	1	1	3	3-oxoacyl-acyl carrier protein synthase II [Pseudomonas aeruginosa PAO1]
	9946088	24.8	7	2090	1	1	2	probable transcriptional regulator [Pseudomonas aeruginosa PAO1]
	9950993	24.73	4	34900	1	1	1	glucose-6-phosphate isomerase [Pseudomonas aeruginosa PAO1]
	9950788	24.62	9	4420000	2	2	2	pseudouridine synthase [Pseudomonas aeruginosa PAO1]
	9951098	24.62	9	3460000	2	2	2	dihydrolipoamide dehydrogenase 3 [Pseudomonas aeruginosa PAO1]
	9946207	24.62	3	1690000	1	1	2	formamidopyrimidine-DNA glycosylase [Pseudomonas aeruginosa PAO1]
	9947523	24.55	3	7020000	2	2	2	aconitate hydratase 1 [Pseudomonas aeruginosa PAO1]
	9946538	24.55	5	11500000	1	1	2	PA0660 [Pseudomonas aeruginosa PAO1]
	9947169	24.52	4	220000	1	1	1	hypothetical protein PA1239 [Pseudomonas aeruginosa PAO1]
	9948709	24.47	2	0	1	1	1	NADH dehydrogenase I chain F [Pseudomonas aeruginosa PAO1]
	9947752	24.46	11	276000	1	1	1	probable methyltransferase [Pseudomonas aeruginosa PAO1]

Table 3.S3. Mass spectrometry analysis of GFP-tagged gp61 in PhiPA3-infected P. aeruginosa (continued)

	9949670	24.41	4	740000	2	2	2	probable Resistance-Nodulation-Cell Division (RND) efflux transporter [Pseudomonas aeruginosa PAO1]
	9951226	24.37	1	6660000	1	1	4	DNA mismatch repair protein MutL [Pseudomonas aeruginosa PAO1]
	9950342	24.25	4	62200	1	1	1	probable porin [Pseudomonas aeruginosa PAO1]
	9946820	24.23	6	20200000	2	2	2	potassium uptake protein Kup [Pseudomonas aeruginosa PAO1]
	9948726	24.2	4	1360000	1	1	1	probable two-component sensor [Pseudomonas aeruginosa PAO1]
	9948211	24.19	23	5780000	2	2	2	conserved hypothetical protein [Pseudomonas aeruginosa PAO1]
	9949715	24.11	1	42400	1	1	1	phosphotransferase system transporter enzyme I FruI [Pseudomonas aeruginosa PAO1]
	9948295	24.11	3	6580000	1	1	3	probable transcriptional regulator [Pseudomonas aeruginosa PAO1]
	9948626	24.08	4	26400000	1	1	2	conserved hypothetical protein [Pseudomonas aeruginosa PAO1]
	9950762	23.85	3	0	1	1	1	probable chemotaxis transducer [Pseudomonas aeruginosa PAO1]
	9950817	23.79	2	235000	1	1	1	probable cytochrome c [Pseudomonas aeruginosa PAO1]
	9950353	23.74	3	2230	1	1	1	transcriptional regulator AcoR [Pseudomonas aeruginosa PAO1]
	9946477	23.71	9	65500000	2	2	2	probable binding protein component of ABC transporter [Pseudomonas aeruginosa PAO1]
	9948804	23.68	1	6780000	1	1	2	PA2728 [Pseudomonas aeruginosa PAO1]
	9946696	23.65	6	228000	1	1	1	probable oxidoreductase [Pseudomonas aeruginosa PAO1]
	9949573	23.61	15	0	2	2	2	probable transcriptional regulator [Pseudomonas aeruginosa PAO1]
	9949685	23.59	3	255000	2	2	2	probable serine protease [Pseudomonas aeruginosa PAO1]
	9951674	23.58	4	2750000	1	1	2	glycolate oxidase subunit GlcE [Pseudomonas aeruginosa PAO1]
	9947572	23.51	14	4420000	1	1	2	conserved hypothetical protein [Pseudomonas aeruginosa PAO1]
	9951191	23.5	5	84900000	1	1	3	probable transcriptional regulator [Pseudomonas aeruginosa PAO1]

Table 3.S3. Mass spectrometry analysis of GFP-tagged gp61 in PhiPA3-infected P. aeruginosa (continued)

	9949605	23.49	4	831000	2	2	2	probable sensor/response regulator hybrid [Pseudomonas aeruginosa PAO1]
	9945895	23.45	5	9490000	2	2	2	hypothetical protein PA0071 [Pseudomonas aeruginosa PAO1]
	9948942	23.44	3	4270000	1	1	1	conserved hypothetical protein [Pseudomonas aeruginosa PAO1]
	9946185	23.4	2	3310000	1	1	3	phosphoenolpyruvate-protein phosphotransferase PtsP [Pseudomonas aeruginosa PAO1]
	9948217	23.37	4	57900	1	1	1	hydrogen cyanide synthase HcnC [Pseudomonas aeruginosa PAO1]
	9946752	23.28	7	0	1	1	1	hypothetical protein PA0855 [Pseudomonas aeruginosa PAO1]
	9948960	23.19	6	180000	1	1	1	hypothetical protein PA2871 [Pseudomonas aeruginosa PAO1]
	9948924	23.07	6	0	1	1	1	probable transcriptional regulator [Pseudomonas aeruginosa PAO1]
	9948368	22.99	3	8740000	1	1	1	probable sulfatase [Pseudomonas aeruginosa PAO1]
	9946150	22.96	2	112000	1	1	1	PA0305 [Pseudomonas aeruginosa PAO1]
	9946375	22.65	6	2040000	1	1	1	probable uroporphyrin-III c-methyltransferase [Pseudomonas aeruginosa PAO1]
	9950996	22.62	2	936000	1	1	1	PA4735 [Pseudomonas aeruginosa PAO1]
	9948142	22.58	5	51000000	1	1	3	conserved hypothetical protein [Pseudomonas aeruginosa PAO1]
	9951302	22.52	1	653000	1	1	1	pyruvate dehydrogenase [Pseudomonas aeruginosa PAO1]
	9947059	22.49	4	10000000	1	1	3	hypothetical protein PA1139 [Pseudomonas aeruginosa PAO1]
	9945909	22.38	3	66800	1	1	1	conserved hypothetical protein [Pseudomonas aeruginosa PAO1]
	9949781	22.37	3	12700000	1	1	2	conserved hypothetical protein [Pseudomonas aeruginosa PAO1]
	9949896	22.33	2	0	1	1	1	conserved hypothetical protein [Pseudomonas aeruginosa PAO1]
	9948127	22.3	4	85400000	1	1	2	pyroglutatmate porin OpdO [Pseudomonas aeruginosa PAO1]
	9951562	22.27	2	323000	1	1	1	probable ATP-binding component of ABC transporter [Pseudomonas aeruginosa PAO1]

Table 3.S3. Mass spectrometry analysis of GFP-tagged gp61 in PhiPA3-infected P. aeruginosa (continued)

	9950030	22.15	7	1610000	1	1	4	probable transferase [Pseudomonas aeruginosa PAO1]
	9951104	22.12	5	153	1	1	1	hypothetical protein PA4835 [Pseudomonas aeruginosa PAO1]
	9949197	22.09	5	0	1	1	1	conserved hypothetical protein [Pseudomonas aeruginosa PAO1]
	9950266	22.09	8	13800	1	1	1	probable epimerase [Pseudomonas aeruginosa PAO1]
	9946378	22.09	13	0	1	1	1	probable transcriptional regulator [Pseudomonas aeruginosa PAO1]
	9951746	21.99	5	9020000	1	1	1	sarcosine oxidase gamma subunit [Pseudomonas aeruginosa PAO1]
	9946495	21.99	2	0	1	1	1	probable bacteriophage protein [Pseudomonas aeruginosa PAO1]
	9946959	21.94	3	1170000	1	1	1	hypothetical protein PA1046 [Pseudomonas aeruginosa PAO1]
	9950348	21.87	5	30500000	1	1	3	probable secretion protein [Pseudomonas aeruginosa PAO1]
	9948461	21.85	3	346000	1	1	2	periplasmic trehalase precursor [Pseudomonas aeruginosa PAO1]
	9947352	21.68	2	1950000	1	1	1	probable helicase [Pseudomonas aeruginosa PAO1]
	9947968	21.66	10	0	1	1	1	PA1969 [Pseudomonas aeruginosa PAO1]
	9949609	21.56	2	0	1	1	1	probable ATP-dependent RNA helicase [Pseudomonas aeruginosa PAO1]
	9951796	21.34	3	1270000	1	1	1	hypothetical protein PA5464 [Pseudomonas aeruginosa PAO1]
	9949150	21.3	6	0	1	1	1	probable two-component response regulator [Pseudomonas aeruginosa PAO1]
	9948758	21.28	5	379000	1	1	1	two-component response regulator PfeR [Pseudomonas aeruginosa PAO1]
	9949610	21.26	3	7320000	1	1	4	probable major facilitator superfamily (MFS) transporter [Pseudomonas aeruginosa PAO1]
	9951325	21.2	1	112000	1	1	1	glutamate synthase large chain precursor [Pseudomonas aeruginosa PAO1]
	9948708	21.2	5	2470000	1	1	1	NADH dehydrogenase I chain E [Pseudomonas aeruginosa PAO1]
	9947449	21.19	3	81000	1	1	1	conserved hypothetical protein [Pseudomonas aeruginosa PAO1]

Table 3.S3. Mass spectrometry analysis of GFP-tagged gp61 in PhiPA3-infected P. aeruginosa (continued)

	9945866	21.18	3	25600000	1	1	1	exoenzyme T [Pseudomonas aeruginosa PAO1]
	9950273	21.15	9	101000	1	1	1	hypothetical protein PA4075 [Pseudomonas aeruginosa PAO1]
	9946282	21.14	1	15500000	1	1	1	Resistance-Nodulation-Cell Division (RND) multidrug efflux transporter MexB [Pseudomonas aeruginosa PAO1]
	9948222	21.13	2	3710000	1	1	1	conserved hypothetical protein [Pseudomonas aeruginosa PAO1]
	9950420	21.13	4	0	1	1	1	probable outer membrane protein precursor [Pseudomonas aeruginosa PAO1]
	9951092	21.12	4	429000	1	1	1	hypothetical protein PA4824 [Pseudomonas aeruginosa PAO1]
	334738139	21.11	1	9190000	1	1	2	virion structural protein [Pseudomonas phage PhiPA3]
	9947126	21.09	10	3630000	1	1	4	probable lipoprotein [Pseudomonas aeruginosa PAO1]
	9947136	20.99	2	31400000	1	1	3	conserved hypothetical protein [Pseudomonas aeruginosa PAO1]
	9949545	20.93	12	412000	1	1	1	heme acquisition protein HasAp [Pseudomonas aeruginosa PAO1]
	9947011	20.87	5	542000	1	1	1	flagellar capping protein FliD [Pseudomonas aeruginosa PAO1]
	9949652	20.84	3	347000000	1	1	7	PA3505 [Pseudomonas aeruginosa PAO1]
	9946649	20.83	3	0	1	1	1	L-aspartate oxidase [Pseudomonas aeruginosa PAO1]
	9948523	20.72	5	74700	1	1	1	probable major facilitator superfamily (MFS) transporter [Pseudomonas aeruginosa PAO1]
	9951199	20.71	4	347000	1	1	1	hypothetical protein PA4921 [Pseudomonas aeruginosa PAO1]
	9950234	20.71	3	67600000	1	1	1	hypothetical protein PA4039 [Pseudomonas aeruginosa PAO1]
	9946751	20.69	3	66800	1	1	1	fumarate hydratase [Pseudomonas aeruginosa PAO1]
	9949123	20.69	1	272000	1	1	1	probable soluble lytic transglycosylase [Pseudomonas aeruginosa PAO1]
	9948327	20.68	4	139000	1	1	1	hypothetical protein PA2296 [Pseudomonas aeruginosa PAO1]
	9947407	20.67	8	529000	1	1	1	sigma factor FliA [Pseudomonas aeruginosa PAO1]

Table 3.S3. Mass spectrometry analysis of GFP-tagged gp61 in PhiPA3-infected P. aeruginosa (continued)

	9948930	20.67	4	7880000	1	1	3	probable aldolase [Pseudomonas aeruginosa PAO1]
	9948682	20.66	6	159000	1	1	1	thioredoxin reductase 1 [Pseudomonas aeruginosa PAO1]
	9948013	20.65	5	164000	1	1	1	probable transcriptional regulator [Pseudomonas aeruginosa PAO1]
	9951084	20.62	2	813000	1	1	1	PA4816 [Pseudomonas aeruginosa PAO1]
	9948181	20.6	3	1360000	1	1	1	probable glycosyl hydrolase [Pseudomonas aeruginosa PAO1]
	9949093	20.59	6	9720000	1	1	1	conserved hypothetical protein [Pseudomonas aeruginosa PAO1]
	9949854	20.57	1	56900	1	1	1	probable metal-transporting P-type ATPase [Pseudomonas aeruginosa PAO1]
	9949637	20.49	2	299000000	1	1	3	probable ferredoxin [Pseudomonas aeruginosa PAO1]
	9948290	20.44	2	12700000	1	1	2	probable 2-ketogluconate transporter [Pseudomonas aeruginosa PAO1]
	9946683	20.42	3	4410000	1	1	2	propionate catabolic protein PrpD [Pseudomonas aeruginosa PAO1]
	9947450	20.42	8	553000	1	1	1	hypothetical protein PA1495 [Pseudomonas aeruginosa PAO1]
	9950719	20.41	6	600000	1	1	1	rod shape-determining protein MreB [Pseudomonas aeruginosa PAO1]
	9946497	20.41	6	28400	1	1	1	probable bacteriophage protein [Pseudomonas aeruginosa PAO1]
	9950427	20.4	1	3880000	1	1	1	phenazine biosynthesis protein PhzE [Pseudomonas aeruginosa PAO1]
	9947896	20.4	1	3880000	1	1	1	phenazine biosynthesis protein PhzE [Pseudomonas aeruginosa PAO1]
	9951591	20.34	8	0	1	1	1	diaminopimelate epimerase [Pseudomonas aeruginosa PAO1]
	334738205	20.27	2	2010000	1	1	3	hypothetical protein [Pseudomonas phage PhiPA3]
	9947733	20.23	9	177000	1	1	1	hypothetical protein PA1755 [Pseudomonas aeruginosa PAO1]
	9947882	20.2	8	3990000	1	1	1	probable glutathione S-transferase [Pseudomonas aeruginosa PAO1]
	9948161	20.16	1	140000000	1	1	1	glycogen phosphorylase [Pseudomonas aeruginosa PAO1]
	9946540	20.09	6	13700000	1	1	1	N-acetyl-gamma-glutamyl-phosphate reductase [Pseudomonas aeruginosa PAO1]

Table 3.S3. Mass spectrometry analysis of GFP-tagged gp61 in PhiPA3-infected P. aeruginosa (continued)

	9948398	20.09	4	1400000	1	1	1	hypothetical protein PA2360 [Pseudomonas aeruginosa PAO1]
	9949527	20.08	2	29400000	1	1	1	regulatory protein NosR [Pseudomonas aeruginosa PAO1]
	9946178	20.07	9	3760000	1	1	2	ribose 5-phosphate isomerase [Pseudomonas aeruginosa PAO1]
	9951485	20.05	2	0	1	1	1	probable oxidoreductase [Pseudomonas aeruginosa PAO1]
	9948951	20.05	4	421000000	1	1	1	lipase modulator protein [Pseudomonas aeruginosa PAO1]
	9946498	20.04	6	2070000	1	1	1	probable bacteriophage protein [Pseudomonas aeruginosa PAO1]
	9949047	20.04	6	0	1	1	1	hypothetical protein PA2950 [Pseudomonas aeruginosa PAO1]
	9950648	20.03	5	0	1	1	1	UDP-N-acetylmuramoylalanyl-D-glutamyl-2 6-diaminopimelate--D-alanyl-D-alanyl ligase [Pseudomonas aeruginosa PAO1]
	9948293	20.03	3	34100000	1	1	1	gluconate dehydrogenase [Pseudomonas aeruginosa PAO1]
	9949541	20.01	2	658000	1	1	1	probable outer membrane protein precursor [Pseudomonas aeruginosa PAO1]
	9948416	20.01	5	1750000	1	1	1	probable transcriptional regulator [Pseudomonas aeruginosa PAO1]

Table 3.S3. Mass spectrometry analysis of GFP-tagged gp61 in PhiPA3-infected P. aeruginosa (continued)

	Accession	-10lgP	Coverage (%)	Peak area	#Peptides	#Unique	#Spec Sample 2	Description
GFP alone	9950490	46.73	19	9300000	2	2	3	50S ribosomal protein L7 / L12 [Pseudomonas aeruginosa PAO1]
	9950477	41.66	11	1250000	2	2	5	50S ribosomal protein L2 [Pseudomonas aeruginosa PAO1]
	9950492	37.2	5	1850000	1	1	4	50S ribosomal protein L1 [Pseudomonas aeruginosa PAO1]
	9950479	36.12	12	1300000	2	2	4	50S ribosomal protein L4 [Pseudomonas aeruginosa PAO1]
	9946571	34.6	1	50800000	3	3	6	hypothetical protein PA0690 [Pseudomonas aeruginosa PAO1]
	9951114	32.21	3	88400000	2	2	12	probable chemotaxis transducer [Pseudomonas aeruginosa PAO1]
	9950133	31.58	2	171000000	2	2	2	probable two-component sensor [Pseudomonas aeruginosa PAO1]
	9950454	29.62	20	291000	1	1	3	50S ribosomal protein L17 [Pseudomonas aeruginosa PAO1]
	9948337	29.21	3	10700000	2	2	2	AmbB [Pseudomonas aeruginosa PAO1]
	9949627	27.01	3	517000000	2	2	6	methionyl-tRNA synthetase [Pseudomonas aeruginosa PAO1]
	9951270	26.89	2	86100000	2	2	2	probable oxidoreductase [Pseudomonas aeruginosa PAO1]
	334738204	26.43	2	3870000	2	2	2	tail fibre protein [Pseudomonas phage PhiPA3]
	9948469	25.98	1	9470000	2	2	2	PvdL [Pseudomonas aeruginosa PAO1]
	9951365	24.72	5	1410000	2	2	2	probable chemotaxis transducer [Pseudomonas aeruginosa PAO1]
	9947007	23.62	2	40600000	2	2	2	flagellar glycosyl transferase FgtA [Pseudomonas aeruginosa PAO1]
	9949894	20.89	1	687000	1	1	1	hypothetical protein PA3728 [Pseudomonas aeruginosa PAO1]
	9951242	20.69	2	8430000	1	1	4	probable phosphoserine phosphatase [Pseudomonas aeruginosa PAO1]
	9951229	20.54	2	62200000	1	1	1	conserved hypothetical protein [Pseudomonas aeruginosa PAO1]
	9948118	20.17	4	3750000	1	1	3	probable cysteine synthase [Pseudomonas aeruginosa PAO1]

Table 3.S4. TMT-tagged mass spectrometry analysis of Goslar protein level in ChmC knockdown infections. (Knockdown/Nontargeting ratio n=3 for each gRNA)

NCBI accession	guide 2 - 1	guide 2- 2	guide 2-3	guide 3 - 1	guide 3 - 2	guide 3 - 3	Description
YP_009820861.1	0.48	0.37	0.31	0.4	0.34	0.28	ChmC [Escherichia phage vB_EcoM_Goslar]
YP_009820901.1	0.71	0.75	0.7	0.55	0.56	0.55	hypothetical protein HOV27_gp218 [Escherichia phage vB_EcoM_Goslar]
YP_009820708.1	0.85	0.79	0.64	0.65	0.53	0.5	hypothetical protein HOV27_gp023 [Escherichia phage vB_EcoM_Goslar]
YP_009820846.1	0.74	0.7	0.6	0.59	0.57	0.56	BspA family leucine-rich repeat surface protein [Escherichia phage vB_EcoM_Goslar]
YP_009820776.1	0.6	0.6	0.54	0.6	0.62	0.52	hypothetical protein HOV27_gp091 [Escherichia phage vB_EcoM_Goslar]
YP_009820761.1	0.67	0.56	0.58	0.61	0.57	0.57	hypothetical protein HOV27_gp076 [Escherichia phage vB_EcoM_Goslar]
YP_009820771.1	0.72	0.71	0.71	0.62	0.57	0.56	hypothetical protein HOV27_gp086 [Escherichia phage vB_EcoM_Goslar]
YP_009820714.1	0.74	0.68	0.68	0.64	0.56	0.56	hypothetical protein HOV27_gp029 [Escherichia phage vB_EcoM_Goslar]
YP_009820752.1	0.69	0.69	0.68	0.62	0.62	0.64	hypothetical protein HOV27_gp067 [Escherichia phage vB_EcoM_Goslar]
YP_009820773.1	0.84	0.81	0.76	0.65	0.62	0.61	hypothetical protein HOV27_gp088 [Escherichia phage vB_EcoM_Goslar]
YP_009820857.1	0.82	0.59	0.58	0.68	0.61	0.59	hypothetical protein HOV27_gp172 [Escherichia phage vB_EcoM_Goslar]
YP_009820726.1	0.78	0.72	0.68	0.63	0.63	0.63	hypothetical protein HOV27_gp041 [Escherichia phage vB_EcoM_Goslar]
YP_009820744.1	0.67	0.72	0.57	0.73	0.6	0.56	hypothetical protein HOV27_gp059 [Escherichia phage vB_EcoM_Goslar]
YP_009820886.1	0.61	0.86	0.56	0.62	0.78	0.53	hypothetical protein HOV27_gp233 [Escherichia phage vB_EcoM_Goslar]
YP_009820729.1	0.67	0.71	0.65	0.69	0.64	0.6	hypothetical protein HOV27_gp044 [Escherichia phage vB_EcoM_Goslar]
YP_009820902.1	0.89	0.81	0.74	0.66	0.64	0.65	hypothetical protein HOV27_gp217 [Escherichia phage vB_EcoM_Goslar]
YP_009820745.1	1.03	0.82	0.78	0.72	0.6	0.64	hypothetical protein HOV27_gp060 [Escherichia phage vB_EcoM_Goslar]
YP_009820707.1	0.87	0.72	0.73	0.68	0.64	0.65	hypothetical protein HOV27_gp022 [Escherichia phage vB_EcoM_Goslar]
YP_009820782.1	0.68	0.68	0.73	0.69	0.65	0.65	hypothetical protein HOV27_gp097 [Escherichia phage vB_EcoM_Goslar]
YP_009820775.1	0.7	0.65	0.63	0.69	0.67	0.64	tail fiber repeat protein [Escherichia phage vB_EcoM_Goslar]
YP_009820740.1	0.73	0.64	0.67	0.69	0.64	0.68	hypothetical protein HOV27_gp055 [Escherichia phage vB_EcoM_Goslar]
YP_009820764.1	0.8	0.64	0.6	0.79	0.66	0.58	hypothetical protein HOV27_gp079 [Escherichia phage vB_EcoM_Goslar]
YP_009820842.1	0.85	0.61	0.73	0.75	0.6	0.68	metallophosphoesterase family protein [Escherichia phage vB_EcoM_Goslar]
YP_009820746.1	0.84	0.88	0.83	0.69	0.67	0.68	hypothetical protein HOV27_gp061 [Escherichia phage vB_EcoM_Goslar]
YP_009820751.1	0.68	0.7	0.65	0.72	0.7	0.62	hypothetical protein HOV27_gp066 [Escherichia phage vB_EcoM_Goslar]
YP_009820770.1	0.8	0.63	0.7	0.78	0.62	0.64	hypothetical protein HOV27_gp085 [Escherichia phage vB_EcoM_Goslar]
YP_009820702.1	0.65	0.68	0.69	0.64	0.68	0.73	hypothetical protein HOV27_gp017 [Escherichia phage vB_EcoM_Goslar]
YP_009820874.1	0.78	0.66	0.65	0.74	0.66	0.65	hypothetical protein HOV27_gp245 [Escherichia phage vB_EcoM_Goslar]
YP_009820887.1	0.78	0.86	0.88	0.7	0.63	0.72	hypothetical protein HOV27_gp232 [Escherichia phage vB_EcoM_Goslar]

Table 3.S5. TMT-tagged mass spectrometry analysis of Goslar protein level in ChmC knockdown infections. (Knockdown/Nontargeting ratio n=3 for each gRNA)

YP_009820739.1	0.67	0.59	0.61	0.83	0.56	0.67	hypothetical protein HOV27_gp054 [Escherichia phage vB_EcoM_Goslar]
YP_009820772.1	0.76	0.68	0.72	0.75	0.66	0.67	hypothetical protein HOV27_gp087 [Escherichia phage vB_EcoM_Goslar]
YP_009820727.1	0.99	0.71	0.69	0.81	0.6	0.68	hypothetical protein HOV27_gp042 [Escherichia phage vB_EcoM_Goslar]
YP_009820742.1	0.86	0.77	0.77	0.72	0.71	0.67	hypothetical protein HOV27_gp057 [Escherichia phage vB_EcoM_Goslar]
YP_009820899.1	0.7	0.7	0.65	0.73	0.71	0.66	hypothetical protein HOV27_gp220 [Escherichia phage vB_EcoM_Goslar]
YP_009820811.1	0.96	0.7	0.7	0.78	0.6	0.73	hypothetical protein HOV27_gp126 [Escherichia phage vB_EcoM_Goslar]
YP_009820847.1	0.67	0.62	1	0.72	0.6	0.79	hypothetical protein HOV27_gp162 [Escherichia phage vB_EcoM_Goslar]
YP_009820900.1	0.79	0.72	0.73	0.7	0.69	0.72	hypothetical protein HOV27_gp219 [Escherichia phage vB_EcoM_Goslar]
YP_009820713.1	0.83	0.76	0.78	0.69	0.69	0.74	hypothetical protein HOV27_gp028 [Escherichia phage vB_EcoM_Goslar]
YP_009820710.1	0.81	0.71	0.72	0.78	0.65	0.7	tail fiber domain [Escherichia phage vB_EcoM_Goslar]
YP_009820853.1	1.13	0.93	0.69	0.88	0.69	0.57	hypothetical protein HOV27_gp168 [Escherichia phage vB_EcoM_Goslar]
YP_009820896.1	0.77	0.68	0.79	0.79	0.62	0.73	hypothetical protein HOV27_gp223 [Escherichia phage vB_EcoM_Goslar]
YP_009820895.1	0.82	0.76	0.67	0.77	0.72	0.66	hypothetical protein HOV27_gp224 [Escherichia phage vB_EcoM_Goslar]
YP_009820932.1	0.99	0.81	0.8	0.67	0.74	0.74	hypothetical protein HOV27_gp187 [Escherichia phage vB_EcoM_Goslar]
YP_009820760.1	0.7	0.59	0.77	0.75	0.8	0.61	hypothetical protein HOV27_gp075 [Escherichia phage vB_EcoM_Goslar]
YP_009820763.1	0.75	0.75	0.68	0.72	0.74	0.7	hypothetical protein HOV27_gp078 [Escherichia phage vB_EcoM_Goslar]
YP_009820703.1	0.73	0.69	0.66	0.77	0.7	0.7	hypothetical protein HOV27_gp018 [Escherichia phage vB_EcoM_Goslar]
YP_009820748.1	0.78	0.74	0.73	0.72	0.73	0.72	hypothetical protein HOV27_gp063 [Escherichia phage vB_EcoM_Goslar]
YP_009820756.1	0.86	0.65	0.77	0.86	0.56	0.76	hypothetical protein HOV27_gp071 [Escherichia phage vB_EcoM_Goslar]
YP_009820789.1	0.72	0.74	0.72	0.77	0.7	0.71	hypothetical protein HOV27_gp104 [Escherichia phage vB_EcoM_Goslar]
YP_009820701.1	0.68	0.6	0.64	0.69	0.64	0.87	hypothetical protein HOV27_gp016 [Escherichia phage vB_EcoM_Goslar]
YP_009820783.1	0.83	0.64	0.93	0.69	0.53	0.98	cell wall hydrolase [Escherichia phage vB_EcoM_Goslar]
YP_009820915.1	0.76	0.69	0.72	0.79	0.7	0.71	chitinase [Escherichia phage vB_EcoM_Goslar]
YP_009820808.1	0.82	0.66	0.65	0.85	0.68	0.68	hypothetical protein HOV27_gp123 [Escherichia phage vB_EcoM_Goslar]
YP_009820711.1	0.75	0.75	0.73	0.76	0.73	0.74	hypothetical protein HOV27_gp026 [Escherichia phage vB_EcoM_Goslar]
YP_009820743.1	0.9	0.86	0.8	0.72	0.73	0.78	hypothetical protein HOV27_gp058 [Escherichia phage vB_EcoM_Goslar]
YP_009820747.1	0.74	0.7	0.67	0.72	0.77	0.75	hypothetical protein HOV27_gp062 [Escherichia phage vB_EcoM_Goslar]
YP_009820778.1	0.8	0.75	0.73	0.76	0.73	0.75	hypothetical protein HOV27_gp093 [Escherichia phage vB_EcoM_Goslar]
YP_009820917.1	0.77	0.77	0.81	0.75	0.75	0.74	hypothetical protein HOV27_gp202 [Escherichia phage vB_EcoM_Goslar]
YP_009820843.1	0.81	0.85	0.79	0.71	0.8	0.74	hypothetical protein HOV27_gp158 [Escherichia phage vB_EcoM_Goslar]
YP_009820777.1	0.79	0.72	0.7	0.83	0.75	0.71	tail protein [Escherichia phage vB_EcoM_Goslar]

Table 3.S6. TMT-tagged mass spectrometry analysis of Goslar protein level in ChmC knockdown infections. (Knockdown/Nontargeting ratio n=3 for each gRNA)

YP_009820831.1	0.92	0.84	0.7	0.75	0.76	0.78	hypothetical protein HOV27_gp146 [Escherichia phage vB_EcoM_Goslar]
YP_009820706.1	0.73	0.74	0.68	0.75	0.84	0.71	hypothetical protein HOV27_gp021 [Escherichia phage vB_EcoM_Goslar]
YP_009820741.1	0.89	0.84	0.78	0.79	0.78	0.73	hypothetical protein HOV27_gp056 [Escherichia phage vB_EcoM_Goslar]
YP_009820903.1	0.85	0.8	0.77	0.8	0.73	0.77	hypothetical protein HOV27_gp216 [Escherichia phage vB_EcoM_Goslar]
YP_009820927.1	0.91	0.73	0.73	0.77	0.81	0.72	hypothetical protein HOV27_gp192 [Escherichia phage vB_EcoM_Goslar]
YP_009820716.1	1.08	0.83	0.66	0.88	0.74	0.68	hypothetical protein HOV27_gp031 [Escherichia phage vB_EcoM_Goslar]
YP_009820922.1	0.85	0.67	0.84	0.82	0.75	0.74	hypothetical protein HOV27_gp197 [Escherichia phage vB_EcoM_Goslar]
YP_009820689.1	0.94	0.83	0.76	0.81	0.82	0.68	hypothetical protein HOV27_gp004 [Escherichia phage vB_EcoM_Goslar]
YP_009820712.1	0.79	0.74	0.72	0.77	0.85	0.71	hypothetical protein HOV27_gp027 [Escherichia phage vB_EcoM_Goslar]
YP_009820860.1	0.66	0.61	0.63	0.81	0.77	0.75	hypothetical protein HOV27_gp175 [Escherichia phage vB_EcoM_Goslar]
YP_009820790.1	0.77	0.66	0.71	0.96	0.76	0.62	hypothetical protein HOV27_gp105 [Escherichia phage vB_EcoM_Goslar]
YP_009820909.1	0.86	0.7	0.78	0.81	0.71	0.82	hypothetical protein HOV27_gp210 [Escherichia phage vB_EcoM_Goslar]
YP_009820845.1	0.8	0.75	0.71	0.78	0.82	0.75	helicase [Escherichia phage vB_EcoM_Goslar]
YP_009820762.1	0.9	0.78	0.91	0.79	0.71	0.86	hypothetical protein HOV27_gp077 [Escherichia phage vB_EcoM_Goslar]
YP_009820721.1	0.9	0.99	0.84	0.79	0.8	0.77	hypothetical protein HOV27_gp036 [Escherichia phage vB_EcoM_Goslar]
YP_009820731.1	0.9	0.7	0.8	0.93	0.62	0.81	hypothetical protein HOV27_gp046 [Escherichia phage vB_EcoM_Goslar]
YP_009820908.1	0.76	0.81	0.73	0.78	0.81	0.77	hypothetical protein HOV27_gp211 [Escherichia phage vB_EcoM_Goslar]
YP_009820720.1	0.78	0.76	0.8	0.79	0.76	0.82	hypothetical protein HOV27_gp035 [Escherichia phage vB_EcoM_Goslar]
YP_009820907.1	0.88	0.74	0.75	0.9	0.73	0.74	hypothetical protein HOV27_gp212 [Escherichia phage vB_EcoM_Goslar]
YP_009820898.1	0.79	0.69	0.77	0.81	0.76	0.82	hypothetical protein HOV27_gp221 [Escherichia phage vB_EcoM_Goslar]
YP_009820884.1	0.72	0.66	0.83	0.94	0.77	0.69	hypothetical protein HOV27_gp235 [Escherichia phage vB_EcoM_Goslar]
YP_009820913.1	0.78	0.76	0.76	0.76	0.73	0.91	hypothetical protein HOV27_gp206 [Escherichia phage vB_EcoM_Goslar]
YP_009820718.1	0.83	0.82	0.83	0.81	0.83	0.76	hypothetical protein HOV27_gp033 [Escherichia phage vB_EcoM_Goslar]
YP_009820755.1	0.99	0.74	0.72	0.83	0.73	0.85	hypothetical protein HOV27_gp070 [Escherichia phage vB_EcoM_Goslar]
YP_009820868.1	0.73	0.68	0.63	0.88	0.78	0.77	hypothetical protein HOV27_gp183 [Escherichia phage vB_EcoM_Goslar]
YP_009820779.1	0.71	0.74	0.85	0.9	0.83	0.71	hypothetical protein HOV27_gp094 [Escherichia phage vB_EcoM_Goslar]
YP_009820734.1	1.36	0.94	0.77	0.94	0.72	0.8	hypothetical protein HOV27_gp049 [Escherichia phage vB_EcoM_Goslar]
YP_009820925.1	0.79	0.86	0.77	0.9	0.83	0.74	hypothetical protein HOV27_gp194 [Escherichia phage vB_EcoM_Goslar]
YP_009820698.1	0.78	0.74	0.72	0.81	0.79	0.88	hypothetical protein HOV27_gp013 [Escherichia phage vB_EcoM_Goslar]
YP_009820733.1	0.81	0.66	0.69	0.86	0.84	0.78	hypothetical protein HOV27_gp048 [Escherichia phage vB_EcoM_Goslar]
YP_009820794.1	1.02	0.77	0.81	1	0.7	0.78	hypothetical protein HOV27_gp109 [Escherichia phage vB_EcoM_Goslar]

Table 3.S7. TMT-tagged mass spectrometry analysis of Goslar protein level in ChmC knockdown infections. (Knockdown/Nontargeting ratio n=3 for each gRNA)

YP_009820717.1	0.83	0.82	0.79	0.83	0.83	0.83	hypothetical protein HOV27_gp032 [Escherichia phage vB_EcoM_Goslar]
YP_009820719.1	0.94	0.87	0.8	0.92	0.87	0.71	hypothetical protein HOV27_gp034 [Escherichia phage vB_EcoM_Goslar]
YP_009820738.1	0.88	0.87	0.8	0.81	0.89	0.8	hypothetical protein HOV27_gp053 [Escherichia phage vB_EcoM_Goslar]
YP_009820796.1	0.79	0.64	0.75	0.91	0.74	0.85	class I SAM-dependent methyltransferase [Escherichia phage vB_EcoM_Goslar]
YP_009820715.1	0.83	0.85	0.88	0.76	0.82	0.93	hypothetical protein HOV27_gp030 [Escherichia phage vB_EcoM_Goslar]
YP_009820875.1	0.71	0.73	0.63	0.93	0.9	0.69	hypothetical protein HOV27_gp244 [Escherichia phage vB_EcoM_Goslar]
YP_009820802.1	0.75	0.74	0.73	0.84	0.79	0.91	hypothetical protein HOV27_gp117 [Escherichia phage vB_EcoM_Goslar]
YP_009820834.1	0.84	0.7	0.74	0.84	0.82	0.9	hypothetical protein HOV27_gp149 [Escherichia phage vB_EcoM_Goslar]
YP_009820687.1	0.88	0.86	0.78	0.88	0.86	0.87	hypothetical protein HOV27_gp002 [Escherichia phage vB_EcoM_Goslar]
YP_009820923.1	0.95	1.04	0.81	0.84	0.93	0.84	hypothetical protein HOV27_gp196 [Escherichia phage vB_EcoM_Goslar]
YP_009820695.1	0.8	0.74	0.79	0.9	0.85	0.92	crossover junction endodeoxyribonuclease RuvC [Escherichia phage vB_EcoM_Goslar]
YP_009820767.1	0.8	0.7	0.85	0.87	0.78	1.03	hypothetical protein HOV27_gp082 [Escherichia phage vB_EcoM_Goslar]
YP_009820735.1	0.86	0.9	0.78	0.82	0.95	0.91	hypothetical protein HOV27_gp050 [Escherichia phage vB_EcoM_Goslar]
YP_009820690.1	0.76	0.68	0.9	0.88	0.84	0.98	hypothetical protein HOV27_gp005 [Escherichia phage vB_EcoM_Goslar]
YP_009820736.1	0.8	0.82	0.84	0.81	0.91	0.98	hypothetical protein HOV27_gp051 [Escherichia phage vB_EcoM_Goslar]
YP_009820877.1	0.83	0.71	0.69	0.9	0.92	0.88	hypothetical protein HOV27_gp242 [Escherichia phage vB_EcoM_Goslar]
YP_009820809.1	0.91	0.79	0.7	0.93	0.92	0.88	hypothetical protein HOV27_gp124 [Escherichia phage vB_EcoM_Goslar]
YP_009820709.1	1.15	1.16	0.96	0.96	0.91	0.87	putative peptidoglycan domain protein [Escherichia phage vB_EcoM_Goslar]
YP_009820871.1	0.99	0.71	0.81	1.04	0.79	0.91	HNH endonuclease [Escherichia phage vB_EcoM_Goslar]
YP_009820920.1	1.28	0.8	0.92	1.09	0.8	0.85	hypothetical protein HOV27_gp199 [Escherichia phage vB_EcoM_Goslar]
YP_009820897.1	0.82	0.69	0.69	0.93	0.86	0.96	hypothetical protein HOV27_gp222 [Escherichia phage vB_EcoM_Goslar]
YP_009820916.1	0.88	0.76	0.73	0.95	0.89	0.91	hypothetical protein HOV27_gp203 [Escherichia phage vB_EcoM_Goslar]
YP_009820686.1	0.84	0.87	0.93	0.87	0.85	1.04	hypothetical protein HOV27_gp001 [Escherichia phage vB_EcoM_Goslar]
YP_009820791.1	0.71	0.82	0.82	0.81	0.96	0.99	hypothetical protein HOV27_gp106 [Escherichia phage vB_EcoM_Goslar]
YP_009820931.1	0.91	0.91	0.82	0.88	0.98	0.9	hypothetical protein HOV27_gp188 [Escherichia phage vB_EcoM_Goslar]
YP_009820788.1	0.91	0.84	0.81	1.03	0.86	0.87	hypothetical protein HOV27_gp103 [Escherichia phage vB_EcoM_Goslar]
YP_009820856.1	0.77	0.67	0.69	0.93	0.92	0.91	hypothetical protein HOV27_gp171 [Escherichia phage vB_EcoM_Goslar]
YP_009820730.1	0.78	0.83	0.81	0.86	0.95	0.98	hypothetical protein HOV27_gp045 [Escherichia phage vB_EcoM_Goslar]
YP_009820924.1	0.95	0.71	0.75	1.01	0.85	0.93	hypothetical protein HOV27_gp195 [Escherichia phage vB_EcoM_Goslar]
YP_009820836.1	0.85	0.73	0.73	0.96	0.92	0.94	hypothetical protein HOV27_gp151 [Escherichia phage vB_EcoM_Goslar]
YP_009820863.1	0.83	0.72	0.71	1	0.91	0.91	hypothetical protein HOV27_gp178 [Escherichia phage vB_EcoM_Goslar]

Table 3.S8. TMT-tagged mass spectrometry analysis of Goslar protein level in ChmC knockdown infections. (Knockdown/Nontargeting ratio n=3 for each gRNA)

YP_009820873.1	0.98	0.96	0.91	0.91	0.96	0.96	hypothetical protein HOV27_gp246 [Escherichia phage vB_EcoM_Goslar]
YP_009820753.1	0.88	0.83	0.67	1.02	1	0.84	hypothetical protein HOV27_gp068 [Escherichia phage vB_EcoM_Goslar]
YP_009820893.1	0.78	0.76	0.75	0.97	0.93	0.97	hypothetical protein HOV27_gp226 [Escherichia phage vB_EcoM_Goslar]
YP_009820728.1	0.84	0.71	0.78	1.04	0.77	1.07	hypothetical protein HOV27_gp043 [Escherichia phage vB_EcoM_Goslar]
YP_009820750.1	0.94	0.88	0.88	0.95	0.96	0.97	hypothetical protein HOV27_gp065 [Escherichia phage vB_EcoM_Goslar]
YP_009820849.1	0.77	0.88	0.77	0.87	1.16	0.85	hypothetical protein HOV27_gp164 [Escherichia phage vB_EcoM_Goslar]
YP_009820912.1	0.89	0.85	0.87	0.96	0.96	0.96	hypothetical protein HOV27_gp207 [Escherichia phage vB_EcoM_Goslar]
YP_009820891.1	0.79	0.78	0.54	1.03	1.06	0.8	hypothetical protein HOV27_gp228 [Escherichia phage vB_EcoM_Goslar]
YP_009820805.1	0.84	0.97	0.77	0.89	1.08	0.94	hypothetical protein HOV27_gp120 [Escherichia phage vB_EcoM_Goslar]
YP_009820865.1	0.84	0.7	0.84	1	0.85	1.06	hypothetical protein HOV27_gp180 [Escherichia phage vB_EcoM_Goslar]
YP_009820754.1	1	0.91	0.91	1.01	0.94	0.97	NUMOD1 domain protein [Escherichia phage vB_EcoM_Goslar]
YP_009820820.1	0.92	0.8	0.95	1.04	0.95	0.94	hypothetical protein HOV27_gp135 [Escherichia phage vB_EcoM_Goslar]
YP_009820838.1	0.84	0.76	0.96	0.91	0.82	1.2	hypothetical protein HOV27_gp153 [Escherichia phage vB_EcoM_Goslar]
YP_009820795.1	0.78	0.79	0.72	0.96	1.04	0.93	hypothetical protein HOV27_gp110 [Escherichia phage vB_EcoM_Goslar]
YP_009820862.1	0.92	0.65	0.66	1.28	0.72	0.94	hypothetical protein HOV27_gp177 [Escherichia phage vB_EcoM_Goslar]
YP_009820724.1	0.77	0.78	0.84	0.93	0.98	1.04	hypothetical protein HOV27_gp039 [Escherichia phage vB_EcoM_Goslar]
YP_009820759.1	0.9	0.76	0.76	1.05	0.9	1	hypothetical protein HOV27_gp074 [Escherichia phage vB_EcoM_Goslar]
YP_009820828.1	0.82	0.8	0.83	0.97	0.99	1	hypothetical protein HOV27_gp143 [Escherichia phage vB_EcoM_Goslar]
YP_009820693.1	0.91	0.86	0.81	0.99	1.02	0.96	dTMP kinase [Escherichia phage vB_EcoM_Goslar]
YP_009820869.1	0.82	0.75	0.82	1	0.96	1.02	hypothetical protein HOV27_gp184 [Escherichia phage vB_EcoM_Goslar]
YP_009820882.1	0.88	0.71	0.79	1.1	0.88	1.06	hypothetical protein HOV27_gp237 [Escherichia phage vB_EcoM_Goslar]
YP_009820757.1	0.86	0.8	0.9	0.99	0.95	1.14	hypothetical protein HOV27_gp072 [Escherichia phage vB_EcoM_Goslar]
YP_009820848.1	0.93	0.93	0.89	0.95	1.08	1.06	hypothetical protein HOV27_gp163 [Escherichia phage vB_EcoM_Goslar]
YP_009820892.1	1.01	0.69	0.83	1.18	0.69	1.24	hypothetical protein HOV27_gp227 [Escherichia phage vB_EcoM_Goslar]
YP_009820854.1	0.82	0.67	0.95	0.9	0.74	1.49	hypothetical protein HOV27_gp169 [Escherichia phage vB_EcoM_Goslar]
YP_009820766.1	0.85	0.82	0.89	1	1.03	1.13	hypothetical protein HOV27_gp081 [Escherichia phage vB_EcoM_Goslar]
YP_009820829.1	0.85	0.83	0.88	1.04	1.13	0.99	hypothetical protein HOV27_gp144 [Escherichia phage vB_EcoM_Goslar]
YP_009820929.1	0.89	0.89	0.76	1.05	1.1	1.01	putative ribonuclease HI [Escherichia phage vB_EcoM_Goslar]
YP_009820837.1	0.8	0.78	0.79	1.03	1.13	1.03	hypothetical protein HOV27_gp152 [Escherichia phage vB_EcoM_Goslar]
YP_009820798.1	0.74	0.83	0.76	0.97	1.29	0.95	GNAT family N-acetyltransferase [Escherichia phage vB_EcoM_Goslar]
YP_009820864.1	0.91	0.92	0.79	1.06	1.15	1	hypothetical protein HOV27_gp179 [Escherichia phage vB_EcoM_Goslar]

Table 3.S9. TMT-tagged mass spectrometry analysis of Goslar protein level in ChmC knockdown infections. (Knockdown/Nontargeting ratio n=3 for each gRNA)

YP_009820914.1	1.04	0.87	0.92	1.14	1.02	1.06	hypothetical protein HOV27_gp205 [Escherichia phage vB_EcoM_Goslar]
YP_009820872.1	0.94	0.84	0.8	1.13	1.1	1	hypothetical protein HOV27_gp247 [Escherichia phage vB_EcoM_Goslar]
YP_009820780.1	0.86	0.83	0.84	1.1	1.15	1.03	hypothetical protein HOV27_gp095 [Escherichia phage vB_EcoM_Goslar]
YP_009820792.1	0.94	0.75	0.83	1.09	1.02	1.17	hypothetical protein HOV27_gp107 [Escherichia phage vB_EcoM_Goslar]
YP_009820859.1	0.79	0.79	0.74	1.05	1.18	1.05	hypothetical protein HOV27_gp174 [Escherichia phage vB_EcoM_Goslar]
YP_009820858.1	0.91	0.86	0.88	1.06	1.11	1.13	hypothetical protein HOV27_gp173 [Escherichia phage vB_EcoM_Goslar]
YP_009820806.1	0.84	0.86	0.76	1.09	1.21	1.02	NAD-dependent DNA ligase LigA [Escherichia phage vB_EcoM_Goslar]
YP_009820850.1	0.93	0.87	0.8	1.17	1.17	1	hypothetical protein HOV27_gp165 [Escherichia phage vB_EcoM_Goslar]
YP_009820885.1	0.95	0.96	0.87	1.06	1.18	1.15	hypothetical protein HOV27_gp234 [Escherichia phage vB_EcoM_Goslar]
YP_009820906.1	1.27	0.99	0.9	1.25	1.04	1.11	hypothetical protein HOV27_gp213 [Escherichia phage vB_EcoM_Goslar]
YP_009820890.1	0.89	0.84	0.89	1.11	1.15	1.18	hypothetical protein HOV27_gp229 [Escherichia phage vB_EcoM_Goslar]
YP_009820839.1	0.8	0.84	0.82	1.11	1.18	1.17	hypothetical protein HOV27_gp154 [Escherichia phage vB_EcoM_Goslar]
YP_009820832.1	1.06	1.12	0.95	1.14	1.32	1.08	3'-5' exoribonuclease [Escherichia phage vB_EcoM_Goslar]
YP_009820697.1	0.8	0.76	0.73	1.16	1.28	1.11	putative SbcC-like protein [Escherichia phage vB_EcoM_Goslar]
YP_009820816.1	1.15	0.84	0.86	1.44	1.02	1.09	DUF4262 domain-containing protein [Escherichia phage vB_EcoM_Goslar]
YP_009820910.1	1.27	1.02	1.09	1.29	1.07	1.25	hypothetical protein HOV27_gp209 [Escherichia phage vB_EcoM_Goslar]
YP_009820876.1	1.02	0.9	0.91	1.26	1.14	1.23	hypothetical protein HOV27_gp243 [Escherichia phage vB_EcoM_Goslar]
YP_009820723.1	1.02	0.97	0.91	1.24	1.28	1.27	hypothetical protein HOV27_gp038 [Escherichia phage vB_EcoM_Goslar]
YP_009820749.1	1.14	0.93	0.95	1.29	1.2	1.31	hypothetical protein HOV27_gp064 [Escherichia phage vB_EcoM_Goslar]
YP_009820826.1	0.97	0.93	0.99	1.56	1.22	1.14	thymidylate synthase [Escherichia phage vB_EcoM_Goslar]
YP_009820692.1	1.1	1.24	1.02	1.25	1.33	1.42	hypothetical protein HOV27_gp007 [Escherichia phage vB_EcoM_Goslar]
YP_009820905.1	1.07	1.06	1.02	1.23	1.29	1.55	hypothetical protein HOV27_gp214 [Escherichia phage vB_EcoM_Goslar]
YP_009820694.1	1.15	0.93	0.89	1.58	1.35	1.29	hypothetical protein HOV27_gp009 [Escherichia phage vB_EcoM_Goslar]
YP_009820696.1	1.31	1.09	1.13	1.39	1.23	1.7	hypothetical protein HOV27_gp011 [Escherichia phage vB_EcoM_Goslar]
YP_009820926.1	1.22	1.02	0.96	1.71	1.51	1.44	hypothetical protein HOV27_gp193 [Escherichia phage vB_EcoM_Goslar]
YP_009820840.1	1.02	0.92	0.9	1.61	1.68	1.7	hypothetical protein HOV27_gp155 [Escherichia phage vB_EcoM_Goslar]
YP_009820851.1	0.94	1.47	1.2	1.35	2.67	2.15	hypothetical protein HOV27_gp166 [Escherichia phage vB_EcoM_Goslar]

Table 3.S10. Goslar protein levels in infected E. coli MC1000 cells. (Spectral count at time points)

NCBI accession number	0 MPI	30 MPI	60 MPI	90 MPI	110 MPI
YP_009820686.1	0	0	2	1	1
YP_009820687.1	0	3	5	0	11
YP_009820688.1	0	7	3	6	4
YP_009820689.1	0	17	17	27	39
YP_009820690.1	0	0	1	3	10
YP_009820692.1	0	0	13	30	49
YP_009820693.1	0	6	40	42	48
YP_009820694.1	0	4	72	84	102
YP_009820695.1	0	0	2	5	8
YP_009820696.1	0	7	76	78	118
YP_009820697.1	0	0	30	25	34
YP_009820698.1	0	0	3	20	36
YP_009820701.1	0	5	6	8	7
YP_009820702.1	0	24	23	20	37
YP_009820703.1	0	0	0	0	5
YP_009820706.1	0	26	28	18	29
YP_009820707.1	0	16	49	192	371
YP_009820708.1	0	18	38	105	313
YP_009820709.1	0	0	6	22	56
YP_009820710.1	0	6	57	149	223
YP_009820711.1	0	0	15	47	79
YP_009820712.1	0	0	3	9	14
YP_009820713.1	0	0	13	12	33
YP_009820714.1	0	5	50	119	180
YP_009820715.1	0	0	27	49	88
YP_009820716.1	0	0	4	11	17
YP_009820717.1	0	3	33	66	94
YP_009820718.1	0	0	12	44	36
YP_009820719.1	0	0	16	22	50
YP_009820720.1	0	0	12	31	64
YP_009820721.1	0	0	45	62	110
YP_009820723.1	0	115	169	146	235

Table 3.S5. Goslar protein levels in infected E. coli MC1000 cells. (Spectral count at time points) (continued)

YP_009820724.1	0	41	80	48	86
YP_009820726.1	0	109	732	1279	2469
YP_009820727.1	0	2	21	44	84
YP_009820728.1	0	0	13	14	21
YP_009820730.1	0	35	52	30	56
YP_009820731.1	0	3	3	2	6
YP_009820733.1	0	0	2	15	24
YP_009820734.1	0	0	4	11	21
YP_009820735.1	0	0	14	23	50
YP_009820736.1	0	0	7	24	41
YP_009820737.1	0	0	2	6	8
YP_009820738.1	0	12	89	216	407
YP_009820739.1	0	0	6	12	25
YP_009820740.1	0	0	5	22	42
YP_009820741.1	0	0	24	55	110
YP_009820742.1	0	0	17	43	112
YP_009820743.1	0	38	221	376	723
YP_009820744.1	0	0	0	13	18
YP_009820745.1	0	0	3	19	30
YP_009820746.1	0	0	41	62	77
YP_009820747.1	0	0	8	6	13
YP_009820748.1	0	0	37	79	111
YP_009820749.1	0	0	22	48	71
YP_009820750.1	0	5	40	84	149
YP_009820751.1	0	0	0	16	14
YP_009820752.1	0	0	17	46	68
YP_009820753.1	0	0	7	12	19
YP_009820754.1	0	0	20	28	57
YP_009820755.1	0	15	20	26	21
YP_009820756.1	0	12	18	14	23
YP_009820757.1	0	11	26	15	19
YP_009820758.1	0	1	2	2	3
YP_009820759.1	0	64	82	82	83

Table 3.S5. Goslar protein levels in infected E. coli MC1000 cells. (Spectral count at time points) (continued)

YP_009820760.1	0	19	18	25	14
YP_009820761.1	0	0	0	3	6
YP_009820762.1	0	0	62	124	254
YP_009820763.1	0	0	16	40	70
YP_009820764.1	0	0	0	2	14
YP_009820765.1	0	0	1	0	1
YP_009820766.1	0	7	47	85	112
YP_009820767.1	0	0	4	5	8
YP_009820768.1	0	0	0	2	4
YP_009820770.1	0	0	0	2	3
YP_009820771.1	0	0	0	11	20
YP_009820772.1	0	0	0	7	16
YP_009820773.1	0	1	4	15	49
YP_009820774.1	0	0	0	1	2
YP_009820775.1	0	4	22	30	166
YP_009820776.1	0	0	0	1	0
YP_009820777.1	0	0	5	42	54
YP_009820778.1	0	0	2	9	24
YP_009820779.1	0	0	3	13	22
YP_009820780.1	0	22	33	42	48
YP_009820781.1	0	2	2	2	2
YP_009820782.1	0	0	1	0	2
YP_009820783.1	0	0	1	1	4
YP_009820788.1	0	15	25	22	24
YP_009820789.1	0	19	23	12	22
YP_009820790.1	0	11	11	11	9
YP_009820791.1	0	0	6	10	11
YP_009820792.1	0	2	4	6	7
YP_009820793.1	0	0	0	2	0
YP_009820794.1	0	5	19	19	23
YP_009820795.1	0	9	28	26	35
YP_009820796.1	0	10	9	6	10
YP_009820797.1	0	0	5	3	8
YP_009820798.1	0	0	3	4	2

Table 3.S5. Goslar protein levels in infected E. coli MC1000 cells. (Spectral count at time points) (continued)

YP_009820800.1	0	0	0	1	2
YP_009820804.1	0	0	0	2	0
YP_009820805.1	0	0	5	6	5
YP_009820806.1	0	30	53	54	76
YP_009820808.1	0	0	0	1	0
YP_009820809.1	0	0	0	1	2
YP_009820810.1	0	0	0	0	1
YP_009820811.1	0	8	13	28	13
YP_009820815.1	0	3	2	0	0
YP_009820816.1	0	6	16	8	10
YP_009820820.1	0	3	15	6	21
YP_009820822.1	0	0	0	0	4
YP_009820824.1	0	1	1	5	2
YP_009820826.1	0	12	37	33	46
YP_009820828.1	0	18	30	21	30
YP_009820829.1	0	14	19	9	19
YP_009820831.1	0	0	2	11	13
YP_009820832.1	0	35	61	75	74
YP_009820834.1	0	0	2	8	3
YP_009820836.1	0	2	10	4	7
YP_009820837.1	0	9	15	15	16
YP_009820838.1	0	1	2	2	2
YP_009820839.1	0	13	16	15	16
YP_009820840.1	0	4	8	8	10
YP_009820842.1	0	3	14	19	35
YP_009820843.1	0	12	44	89	163
YP_009820844.1	0	4	2	5	5
YP_009820845.1	0	0	10	26	59
YP_009820846.1	0	5	65	190	385
YP_009820847.1	1	1	0	0	0
YP_009820848.1	0	18	50	50	60
YP_009820849.1	0	0	10	15	13
YP_009820850.1	0	39	77	81	106
YP_009820851.1	0	0	8	10	12

Table 3.S5. Goslar protein levels in infected E. coli MC1000 cells. (Spectral count at time points) (continued)

YP_009820853.1	0	0	5	11	15
YP_009820854.1	0	4	5	3	7
YP_009820856.1	0	20	46	39	56
YP_009820857.1	0	1	0	3	4
YP_009820858.1	0	85	214	203	246
YP_009820859.1	0	66	104	95	109
YP_009820860.1	0	33	71	53	65
YP_009820861.1	0	14	70	65	97
YP_009820862.1	0	0	1	2	7
YP_009820863.1	0	0	7	10	16
YP_009820864.1	0	1	23	32	34
YP_009820865.1	0	0	2	7	13
YP_009820866.1	0	0	1	2	5
YP_009820868.1	0	19	13	14	16
YP_009820869.1	0	1	3	6	4
YP_009820870.1	0	2	0	2	4
YP_009820872.1	0	9	32	34	42
YP_009820873.1	0	140	299	312	429
YP_009820874.1	0	0	0	7	14
YP_009820875.1	0	4	16	12	18
YP_009820876.1	0	4	30	19	53
YP_009820877.1	0	0	1	1	2
YP_009820880.1	0	0	1	0	0
YP_009820882.1	0	4	24	25	28
YP_009820884.1	0	0	0	0	1
YP_009820885.1	0	30	147	191	277
YP_009820886.1	0	0	8	9	25
YP_009820887.1	0	0	6	9	10
YP_009820890.1	0	2	9	11	15
YP_009820891.1	0	1	8	14	17
YP_009820892.1	0	0	2	1	4
YP_009820893.1	0	8	11	18	15
YP_009820894.1	0	0	1	0	1
YP_009820895.1	0	0	17	29	65

Table 3.S5. Goslar protein levels in infected E. coli MC1000 cells. (Spectral count at time points) (continued)

YP_009820896.1	0	0	0	0	3
YP_009820897.1	0	0	3	8	19
YP_009820898.1	0	0	41	99	170
YP_009820899.1	0	0	30	89	134
YP_009820900.1	0	0	26	46	69
YP_009820901.1	0	24	120	276	565
YP_009820902.1	0	9	41	137	288
YP_009820903.1	0	4	13	18	43
YP_009820905.1	0	104	339	311	543
YP_009820906.1	0	7	82	94	95
YP_009820907.1	0	0	20	56	128
YP_009820908.1	0	0	0	6	10
YP_009820909.1	0	0	5	14	34
YP_009820910.1	0	10	82	124	223
YP_009820911.1	0	0	0	0	4
YP_009820912.1	0	0	94	159	233
YP_009820913.1	0	0	2	4	8
YP_009820914.1	0	0	26	53	61
YP_009820915.1	0	0	15	29	128
YP_009820916.1	0	0	23	24	47
YP_009820917.1	0	0	14	14	30
YP_009820920.1	0	0	3	14	19
YP_009820921.1	0	0	3	2	6
YP_009820922.1	0	0	6	16	23
YP_009820923.1	0	1	8	13	31
YP_009820925.1	0	0	11	6	11
YP_009820926.1	0	9	66	69	107
YP_009820927.1	0	0	13	25	26
YP_009820929.1	0	2	52	50	81
YP_009820930.1	0	0	0	2	5
YP_009820931.1	25	23	96	142	145
YP_009820932.1	0	0	0	1	5

3.6 Figures

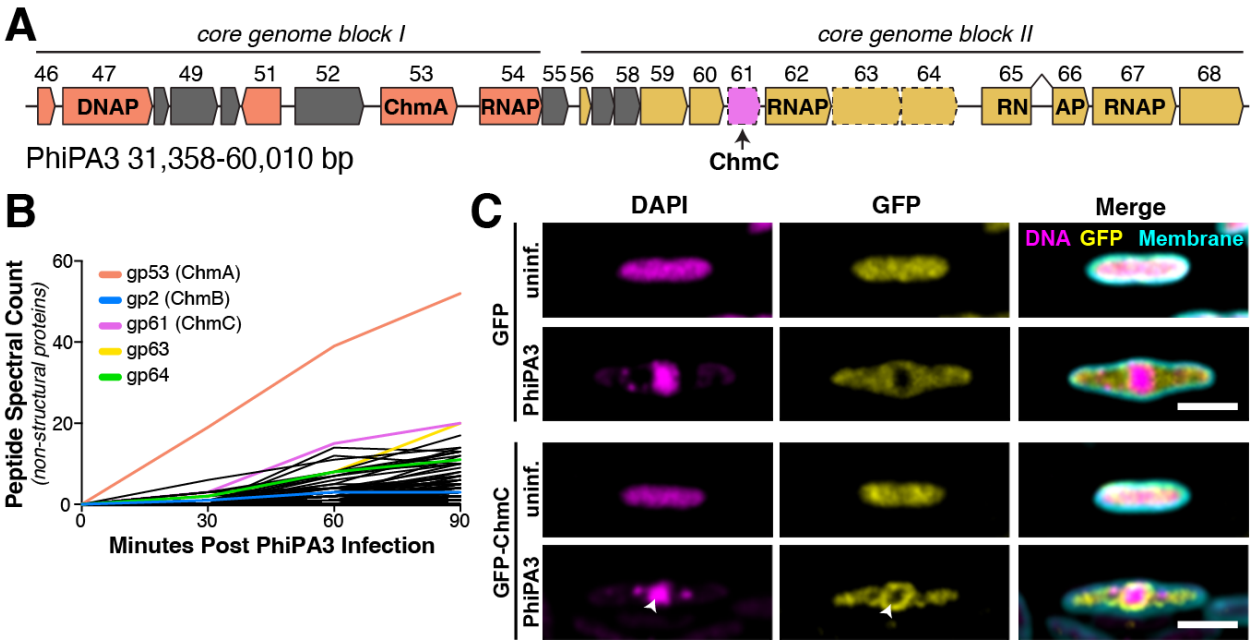


Figure 3.1. PhiPA3 gp61 is associated with the phage nucleus and host ribosomes

(A) Map of the PhiPA3 genome spanning jumbo phage conserved blocks I and II [3]. Genes in conserved block I are colored salmon, and genes in conserved block II are colored goldenrod. Genes conserved with the closely-related jumbo phage PhiKZ but not all jumbo phages are colored gray, and unconserved genes are colored white. gp47 is a putative DNA polymerase (DNAP), and gp53 is the major nuclear shell protein Chimallin (ChmA). Gp54, gp64, gp65-66 (interrupted by a self-splicing intron), and gp67 are subunits of the phage-encoded non-virion RNA polymerase (RNAP). Dotted outlines indicate three proteins (gp61, gp63, and gp64) shown to be associated with the phage nuclear shell (Enustun et al, submitted).

(B) Mass spectrometry proteomics analysis of PhiPA3-infected *P. aeruginosa*, showing spectral counts of non-structural phage proteins. ChmA (gp53), ChmB (gp2), gp61, gp63, and gp64 are shown in colors and labeled. See **Table S1-S2** for full mass spectrometry results.

(C) Localization of sfGFP (top) or sfGFP-fused PhiPA3 gp61 (bottom) in PhiPA3-infected *P. aeruginosa* cells at 75 minutes post infection. Magenta: DAPI nucleic acid dye; cyan: FM4-64 membrane dye; yellow: GFP. Arrowheads indicate phage nuclei, and asterisks indicate phage bouquets. Scale bar = 2 µm.

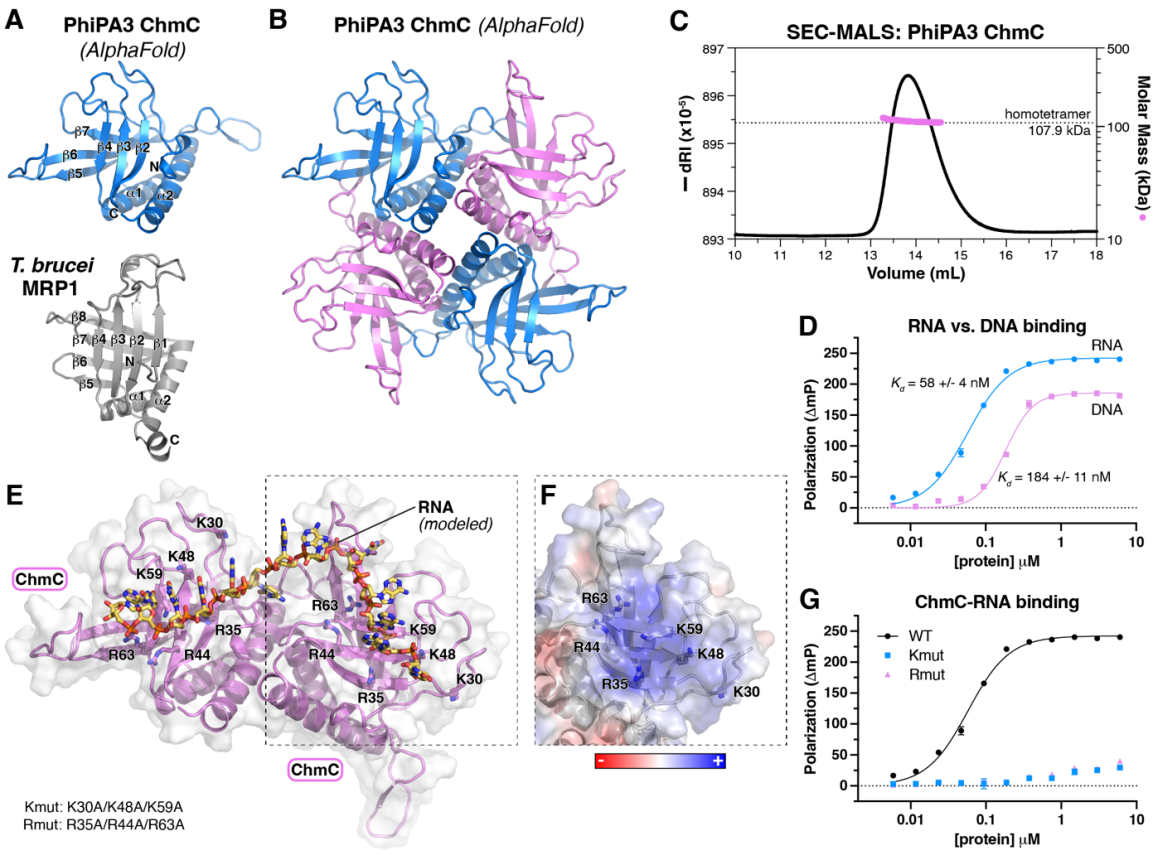


Figure 3.2. ChmC adopts an RNA-binding Whirly fold

(A) AlphaFold predicted structure of PhiPA3 ChmC (gp61; blue) at top oriented equivalently to the structure of T. brucei MRP1 (PDB ID 2GJE; gray) [13] at bottom, with secondary structure elements labeled. (B) AlphaFold predicted structure of a PhiPA3 ChmC homotetramer, with alternating subunits colored blue and pink. See **Figure S2A-B** for details of AlphaFold predictions. (C) Size exclusion chromatography coupled to multi-angle light scattering (SEC-MALS) analysis of purified PhiPA3 ChmC. (D) Single-stranded RNA (blue circles) and DNA (pink squares) binding of PhiPA3 ChmC as measured by fluorescence polarization. Data points are shown as average +/- standard deviation of triplicate technical replicates, and curves are fit with a cooperative binding model (Hill coefficient for RNA binding 1.6 +/- 0.1; Hill coefficient for DNA binding 2.4 +/- 0.3). (E) Structural model of PhiPA3 ChmC binding single-stranded RNA, generated by overlaying two adjacent subunits of the ChmC AlphaFold 2 model shown in panel (B) with structure of a *T. brucei* MRP1:MRP2 heterotetramer bound to RNA (PDB ID 2GJE; see **Figure S2C**) [13] . Lysine and arginine residues mutated to generate the Kmut (K30A/K48A/K59A) and Rmut (R35A/R44A/R63A) are shown as sticks and labeled. (F) Surface charge representation of one ChmC subunit, oriented equivalently to panel (E) (dotted line). See **Figure S2D**. (G) Single-stranded RNA binding of PhiPA3 ChmC wild-type (black circles, reproduced from panel (D)), Kmut (blue squares), and Rmut (pink triangles), as measured by fluorescence polarization. Data points are shown as average +/- standard deviation of triplicate technical replicates. See **Figure S2E** for DNA binding.

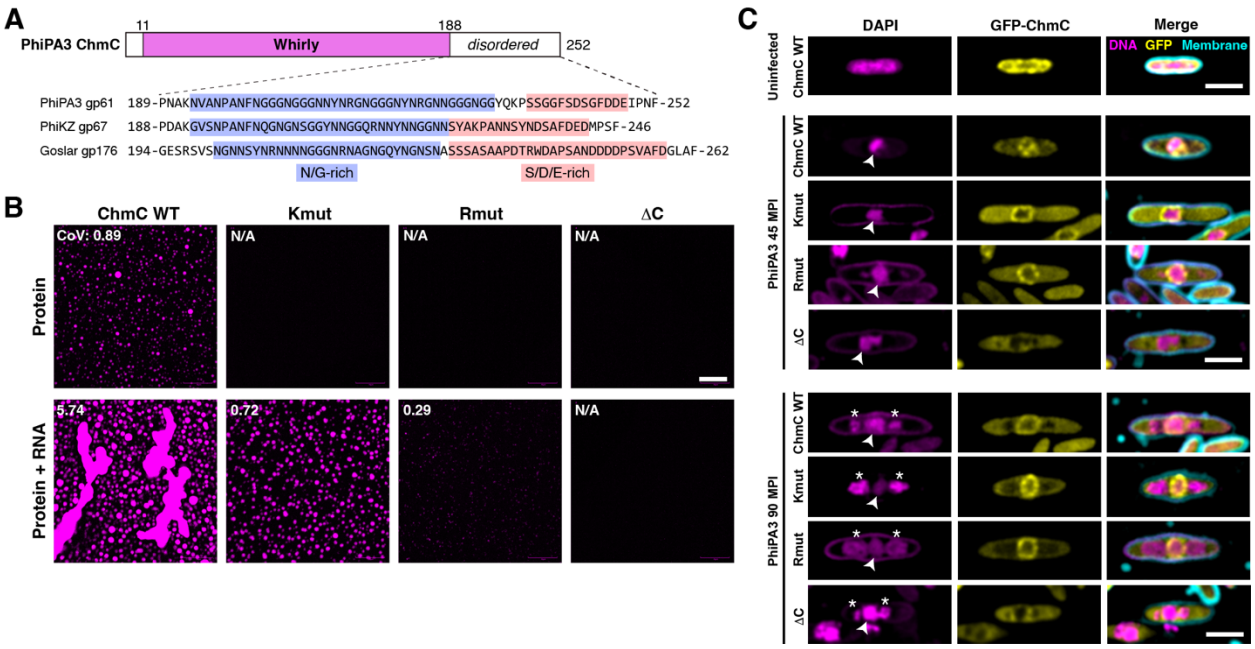


Figure 3.3. ChmC forms phase-separated condensates with RNA

(A) *Top:* Domain structure of PhiPA3 ChmC, with Whirly domain colored magenta and regions predicted to be disordered in white. Bottom: Sequence of the C-terminal predicted disordered domain in ChmC from PhiPA3 (gp61), PhiKZ (gp67), and Goslar (gp176). Asparagine/glycine (N/G) rich regions are highlighted in blue, and serine/aspartate/glutamate (S/D/E) rich regions are highlighted in salmon. See **Figure S3A** for catGRANULE analysis of all three proteins, and **Figure S3C-D** for analysis of ChmC ΔC binding DNA and RNA.

(B) Fluorescence microscopy imaging of PhiPA3 ChmC (wild type, Kmut, Rmut, or ΔC; 10% Cy5-labeled) at 30 μM protein concentration, either alone (top row) or with 2 μM of a 40-base RNA (2.8 μg/mL; middle row) or 83 nM of a 2.3 kb RNA (5.8 μg/mL; bottom row). All images were taken 30 minutes after final dilution and mixing with RNA. For all conditions that showed condensate formation, the coefficient of variation (CoV) was calculated as the standard deviation of particle area divided by the mean particle area (WT protein alone n=423; WT + 40 base RNA n=222; Kmut + 40 base RNA n=247; Rmut + 40 base RNA n=5; WT + 2.3 kb RNA n=552; Kmut + 2.3 kb RNA n=368; Rmut + 2.3 kb RNA n=147). Scale bar = 30 μm. See **Figure S3B** for DIC imaging.

(C) Localization of GFP-tagged PhiPA3 ChmC (wild type, Kmut, Rmut, or ΔC in PhiPA3-infected *P. aeruginosa* cells at 45 and 110 minutes post infection (MPI). Yellow: GFP; magenta: DAPI nucleic acid; cyan: FM4-64 membrane dye. Arrowheads indicate phage nuclei, and asterisks indicate phage bouquets. Scale bar = 2 μm.

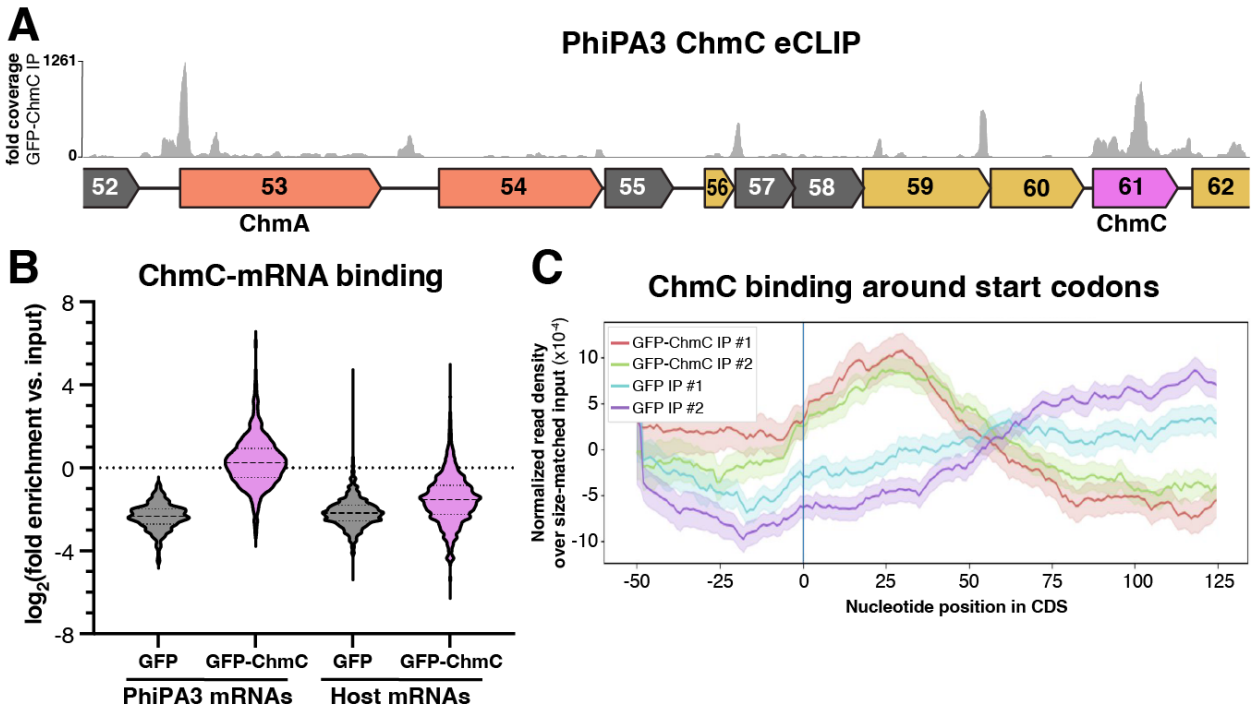


Figure 3.4. ChmC binds phage mRNAs

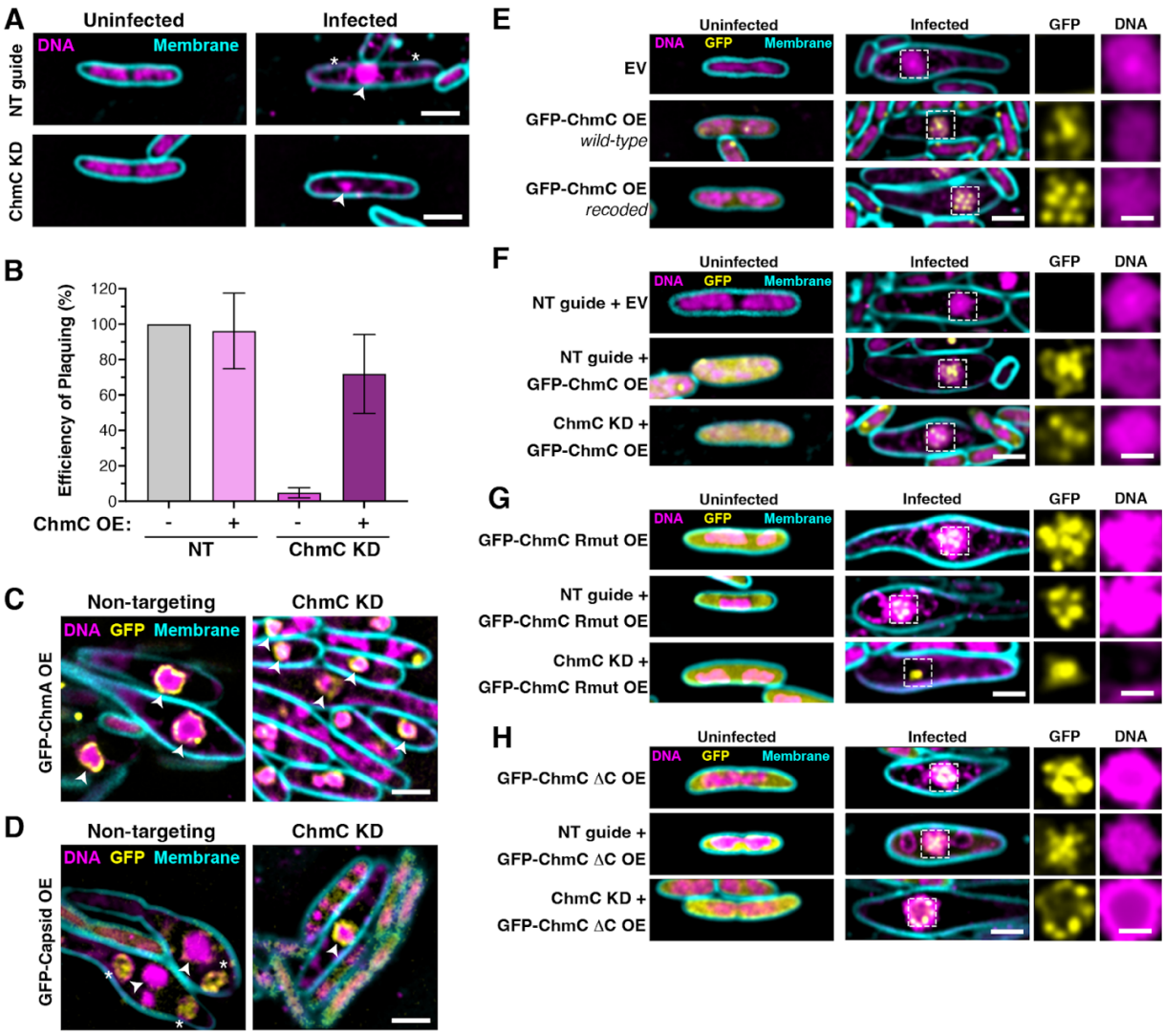
(A) Fold sequence coverage from eCLIP analysis of GFP-ChmC, in the region of the PhiPA3 genome encoding ChmA and ChmC. See **Figure S4** for eCLIP validation.

(B) Log_2(fold enrichment versus input for all PhiPA3 and host (P. aeruginosa) genes detected in eCLIP analysis of GFP (gray) and GFP-ChmC (pink).

(C) Metagene analysis showing enrichment of GFP-ChmC binding near start codons of PhiPA3 genes (replicate #1 red, replicate #2 green).

Figure 3.5. ChmC knockdown impairs phage nucleus development and infection progression

(A) Microscopy of *E. coli* MC1000 cells, either uninfected (left) or infected with Goslar (110 MPI; right).(B) Efficiency of plaquing of Goslar phage infecting *E. coli* MC1000 cells expressing ddCas13d and either a non-targeting guide RNA (NT) or a guide RNA targeting the 5' end of the ChmC (gp176) gene (guide 3; see **Figure S5** for analysis of three guide RNAs). Data are presented as mean +/- standard error of the mean of four biological replicates, calculated as a percentage of Goslar phage plaque forming units with no ddCas13d/guide expression. See **Figure S6** for example plaque images and microscopy of knockdown/rescue cells. (C) Microscopy of *E. coli* MC1000 expressing GFP-tagged ChmA plus ddCas13d and either a non-targeting guide RNA (left) or ChmC-targeting guide 3 (ChmC KD; right). DNA is stained with DAPI and shown in magenta; membranes are stained with FM4-64 and shown in cyan; GFP is shown in yellow. Arrowheads indicate the phage nucleus. Scale bar = 2 μm. See **Figure S7A** for further images.(D) Microscopy of *E. coli* MC1000 expressing GFP-tagged capsid protein plus ddCas13d and either a non-targeting guide RNA (left) or ChmC-targeting guide 3 (ChmC KD; right). DNA is stained with DAPI and shown in magenta; membranes are stained with FM4-64 and shown in cyan; GFP is shown in yellow. Arrowheads indicate the phage nucleus, and asterisks indicate phage bouquets. See **Figure S7B** for further images.(E) Microscopy of *E. coli* MC1000 cells expressing no protein (EV, top row) or GFP-tagged ChmC (gp176, wild-type sequence in middle row and recoded sequence in bottom row), infected with Goslar (110 MPI). Yellow: GFP; magenta: DAPI nucleic acid; cyan: FM4-64 membrane dye. Scale bar = 2 μm for main panels, 1 μm for individual GFP and DNA channels representing zoomed views of the boxed regions.(F) Microscopy of *E. coli* MC1000 cells expressing ddCas13d plus a non-targeting guide RNA (top and middle rows) or a ChmC-targeting guide (bottom roe), expressing either no protein (EV, top row) or GFP-tagged recoded ChmC (middle and bottom rows), and infected with Goslar (110 MPI). Yellow: GFP; magenta: DAPI nucleic acid; cyan: FM4-64 membrane dye. Scale bar = 2 μm for main panels, 1 μm for individual GFP and DNA channels representing zoomed views of the boxed regions.(G) As panel (E), except cells in the middle and bottom row are expressing ChmC Rmut (see **Figure S8** for Rmut design and analysis by Goslar plaquing assay). (H) As panel (E), except cells in the middle and bottom row are expressing ChmC ΔC (see **Figure S8** for ΔC design and analysis by Goslar plaquing assay)


A
Uninfected
Infected
DNA
Membrane
NT guide
ChmC KD
B
Efficiency of Plaquing (%)
120
100
80
60
40
20
0
ChmC OE:
-
+
-
+
NT
ChmC KD
C
Non-targeting
ChmC KD
GFP-ChmA OE
DNA GFP Membrane
D
Non-targeting
ChmC KD
GFP-Capsid OE
DNA GFP Membrane
E
Uninfected
Infected
GFP
DNA
DNA GFP Membrane
EV
GFP-ChmC OE
wild-type
GFP-ChmC OE
recoded
F
NT guide + EV
NT guide +
GFP-ChmC OE
ChmC KD +
GFP-ChmC OE
G
GFP-ChmC Rmut OE
NT guide +
GFP-ChmC Rmut OE
ChmC KD +
GFP-ChmC Rmut OE
H
GFP-ChmC ΔC OE
NT guide +
GFP-ChmC ΔC OE
ChmC KD +
GFP-ChmC ΔC OE

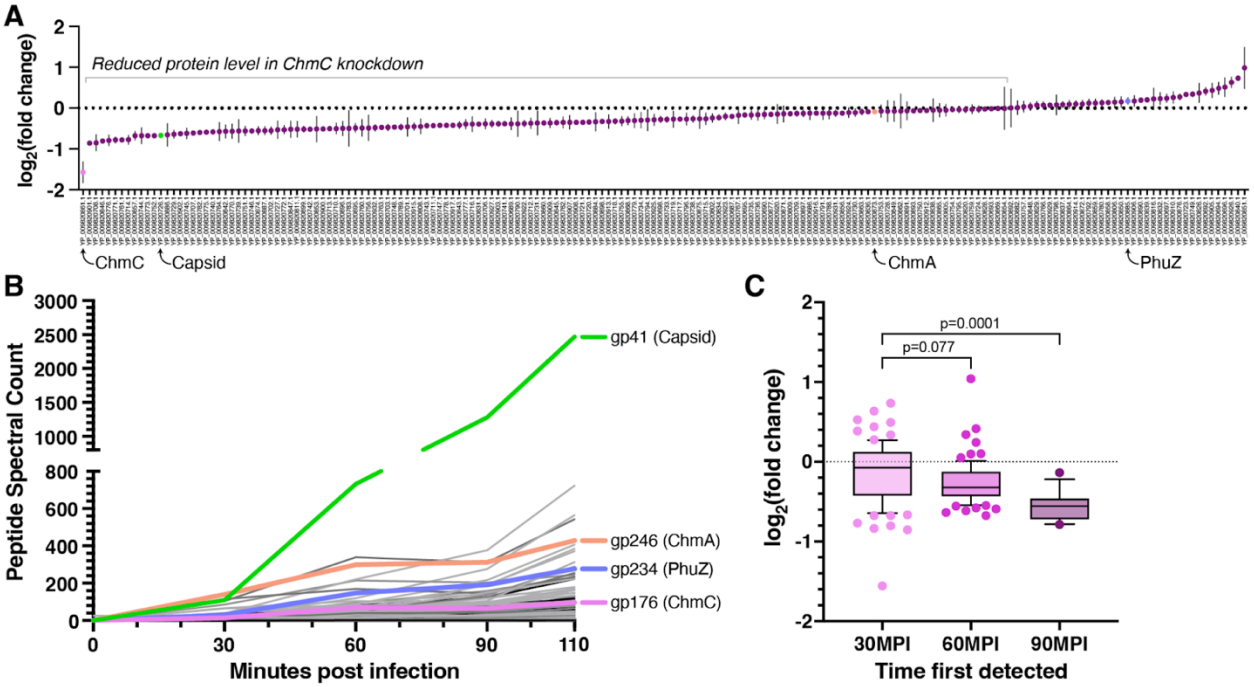


Figure 3.6. ChmC knockdown causes a global reduction in phage protein levels
(A) Log_2(fold change) in expression of 180 Goslar proteins detected by TMT-tagged mass spectrometry analysis of Goslar-infected *E. coli* MC1000 cells (110 MPI) expressing dCas13d and *chmC*-targeting guide RNA 3 (see **Table S4** for full data, including for *chmC*-targeting guide RNA 2). Proteins are labeled by their NCBI accession numbers. ChmC (gp176), Capsid (gp41), ChmA (gp246), and PhuZ (gp234) are marked.
(B) Spectral count of Goslar proteins from mass spectrometry of Goslar-infected *E. coli* MC1000 cell lysate at 0, 30, 60, 90, and 110 minutes post infection (MPI). ChmC (gp176) is shown in violet; Capsid (gp41) in green; ChmA (gp246) in salmon; and PhuZ (gp234) in blue. See **Table S5** for full data.
(C) Aggregate log_2(fold change) in expression upon ChmC knockdown (guide 3) for proteins first detected at 30 minutes (pink, n=84), 60 minutes (violet, n=74), and 90 minutes (purple, n=13) post Goslar infection. Box shows 25th to 75th percentile with mean value marked; whiskers show 10th to 90th percentile; and outlier points are shown individually. P values from one-way ANOVA comparisons are shown.

3.7 Supplementary Figures

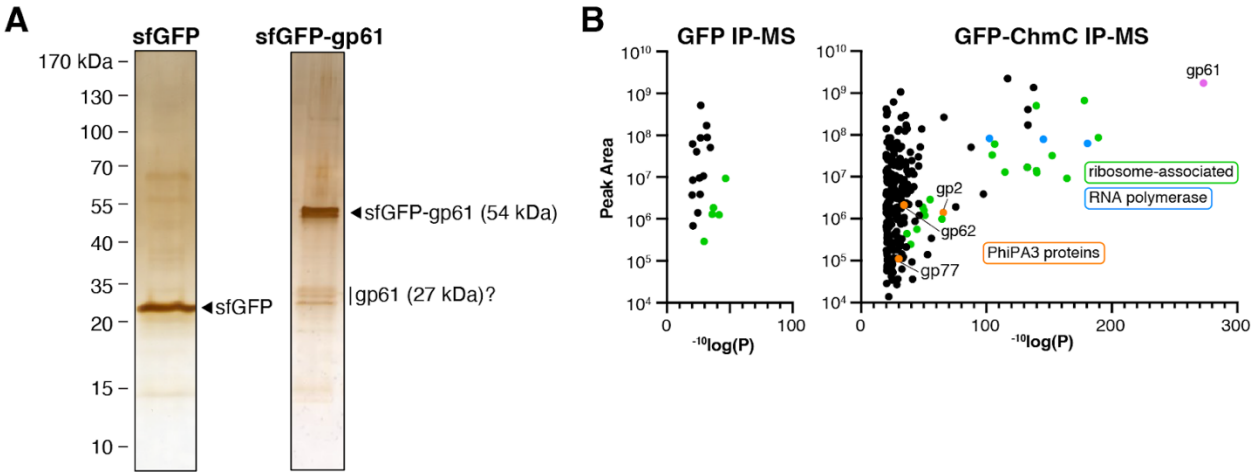


Figure 3.S1. IP-MS of PhiPA3 gp61
(A) Silver-stained SDS-PAGE analysis of sfGFP (left) and sfGFP-gp61 (right) pulldowns from PhiPA3-infected *P. aeruginosa* cells. Tentative assignments of prominent bands are shown for each.
(B) Mass spectrometry proteomics analysis (total peak area for assigned peptides per protein vs. $^{-10}$log(P) for sfGFP (left panel) versus sfGFP-fused gp61 (right panel) immunopurified from PhiPA3-infected *P. aeruginosa* (45 minutes post infection). gp61 is labeled and shown in pink, other PhiPA3 proteins are shown in orange and individually labeled, host ribosomal or ribosome-associated proteins are shown in green, and host RNA polymerase subunits are shown in blue. See **Table S3** for full mass spectrometry data.

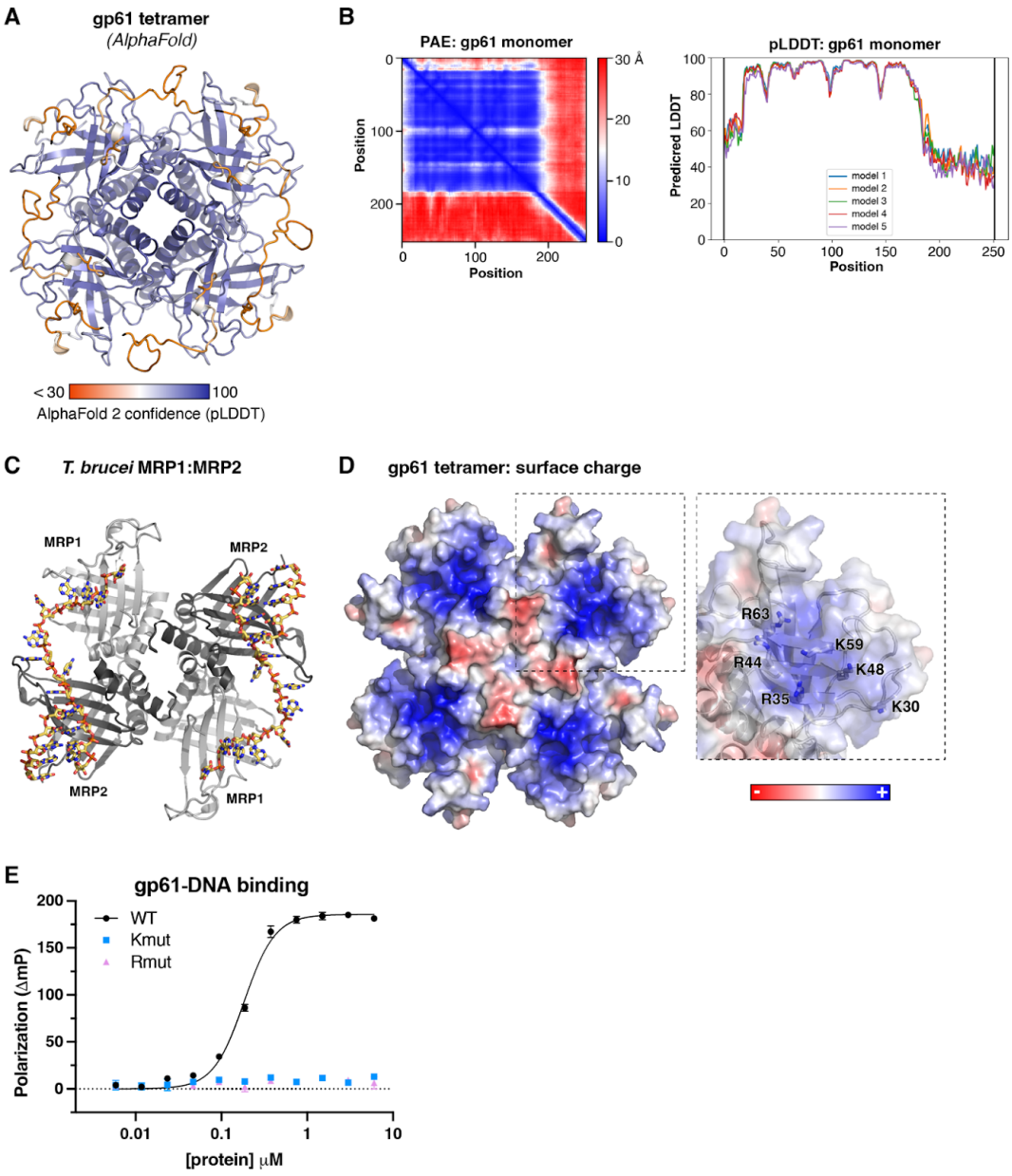


Figure 3.S2. ChmC adopts an RNA-binding Whirly domain fold

(A) AlphaFold 2 predicted structure of a homotetramer of PhiPA3 ChmC (gp61), colored by confidence (pLDDT). (B) Predicted Aligned Error (PAE) and predicted local distance difference test (pLDDT) plots for one PhiPA3 ChmC monomer, as output by AlphaFold 2. (C) Structure of *T. brucei* MRP1:MRP2 bound to RNA (PDB ID 2GJE) [13]. (D) Surface charge distribution of the AlphaFold 2 predicted structure of a homotetramer of PhiPA3 ChmC, calculated by APBS [45] in PyMOL version 2. (E) Single-stranded RNA binding of PhiPA3 ChmC wild-type (black circles, reproduced from **Figure 2D**), Kmut (blue squares), and Rmut (pink triangles), as measured by fluorescence polarization. Data points are shown as average +/- standard deviation of triplicate technical replicates.

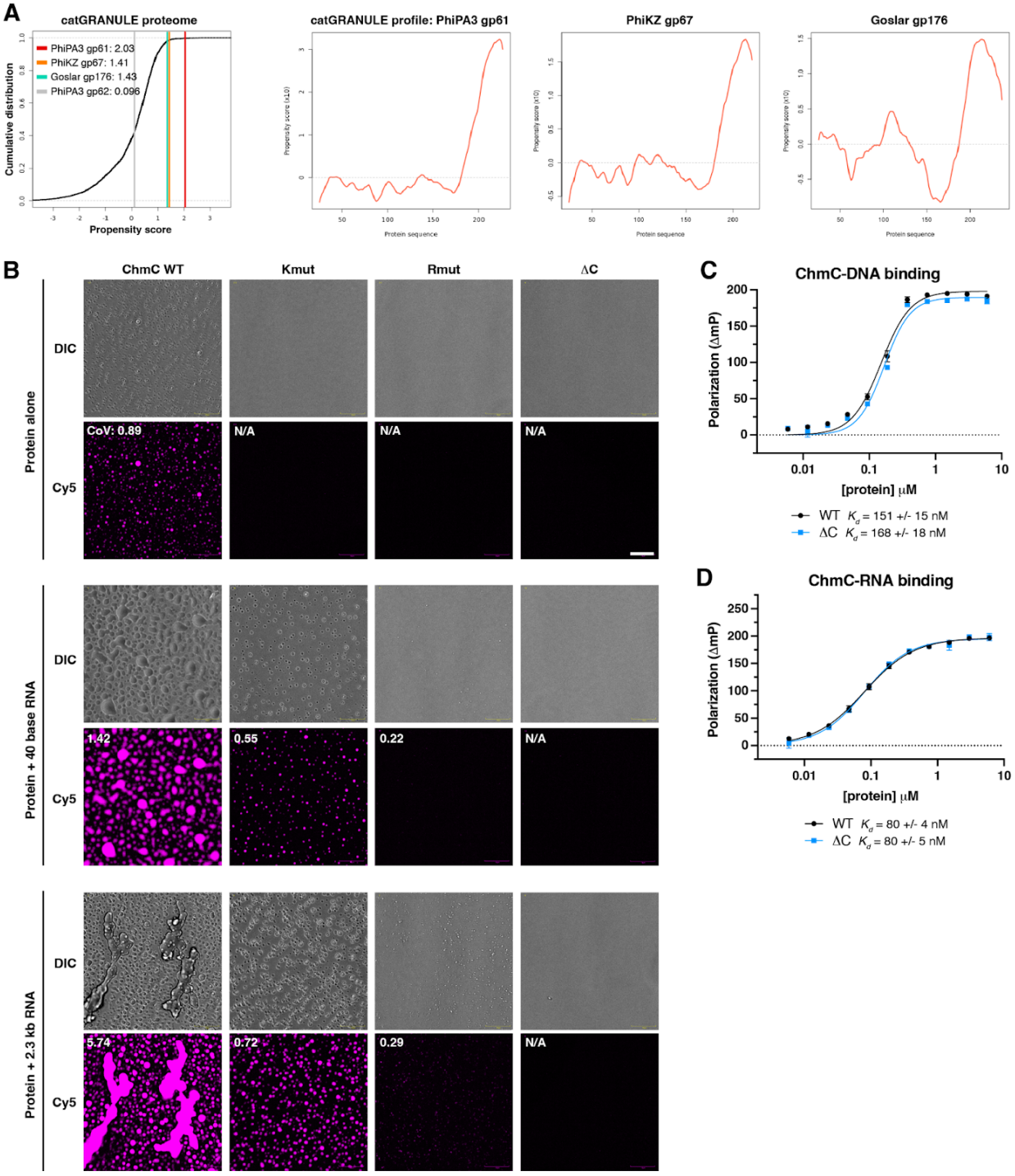
Figure 3.S3. ChmC forms phase-separated condensates with RNA

(A) *Left:* catGRANULE output for ChmC from PhiPA3 (gp61), PhiKZ (gp67; 56% identical to PhiPA3 gp61), and Goslar (gp176; 30% identical to PhiPA3 gp61), plus a soluble protein (PhiPA3 gp62, a nvRNAP subunit) as a negative control. *Right:* catGRANULE profiles for ChmC from PhiPA3 (gp61), PhiKZ (gp67), and Goslar (gp176), showing that the disordered C-terminus of each protein likely drives phase separation. In agreement with catGRANULE output, the machine learning based PSPredictor server [46] strongly predicted condensation behavior for all three proteins (scores of 0.8706 for PhiPA3 gp61; 0.6531 for PhiKZ gp67; and 0.8461 for Goslar gp176 on a scale of 0 (no condensation predicted) to 1 (highly confident condensation behavior predicted)). These predictions were completely dependent on the C-terminal disordered regions, with removal of these regions resulting in low scores for all three proteins (0.0048 for PhiPA3 gp61 residues 1-188; 0.0024 for PhiKZ gp67 residues 1-187; and 0.0226 for Goslar gp176 residues 1-193).

(B) Differential interference contrast (DIC) and fluorescence microscopy (Cy5) imaging of PhiPA3 ChmC (wild type, Kmut, Rmut, or ΔC; 10% Cy5-labeled) at 30 μM protein concentration, either alone (top row) or with 2 μM of a 40-base RNA (2.8 μg/mL; middle row) or 83 nM of a 2.3 kb RNA (5.8 μg/mL; bottom row). All images were taken 30 minutes after final dilution and mixing with RNA. For all conditions that showed condensate formation, the coefficient of variation (CoV) was calculated as the standard deviation of particle area divided by the mean particle area (WT protein alone n=423; WT + 40 base RNA n=222; Kmut + 40 base RNA n=247; Rmut + 40 base RNA n=5; WT + 2.3 kb RNA n=552; Kmut + 2.3 kb RNA n=368; Rmut + 2.3 kb RNA n=147). Scale bar = 30 μm. See Figure S2E for DIC imaging.

(C) Single-stranded DNA binding of PhiPA3 ChmC WT (black squares) and ΔC (blue circles) as measured by fluorescence polarization. Data points are shown as average +/- standard deviation of triplicate technical replicates, and curves are fit with a cooperative binding model (Hill coefficient for WT ChmC binding DNA 2.0 +/- 0.3; Hill coefficient for ChmC ΔC binding DNA 2.2 +/- 0.4).

(D) Single-stranded RNA binding of PhiPA3 ChmC WT (black squares) and ΔC (blue circles) as measured by fluorescence polarization. Data points are shown as average +/- standard deviation of triplicate technical replicates, and curves are fit with a cooperative binding model (Hill coefficient for WT ChmC binding RNA 1.2 +/- 0.1; Hill coefficient for ChmC ΔC binding RNA 1.3 +/- 0.1).


A
catGRANULE proteome
PhiPA3 gp61: 2.03
PhiKZ gp67: 1.41
Goslar gp176: 1.43
PhiPA3 gp62: 0.096
Cumulative distribution
Propensity score
catGRANULE profile: PhiPA3 gp61
PhiKZ gp67
Goslar gp176
Protein sequence
B
ChmC WT
Kmut
Rmut
ΔC
Protein alone
Protein + 40 base RNA
Protein + 2.3 kb RNA
DIC
Cy5
CoV: 0.89
N/A
1.42
0.55
0.22
5.74
0.72
0.29
C
ChmC-DNA binding
Polarization (ΔmP)
[protein] μM
WT K_d = 151 +/- 15 nM
ΔC K_d = 168 +/- 18 nM
D
ChmC-RNA binding
WT K_d = 80 +/- 4 nM
ΔC K_d = 80 +/- 5 nM

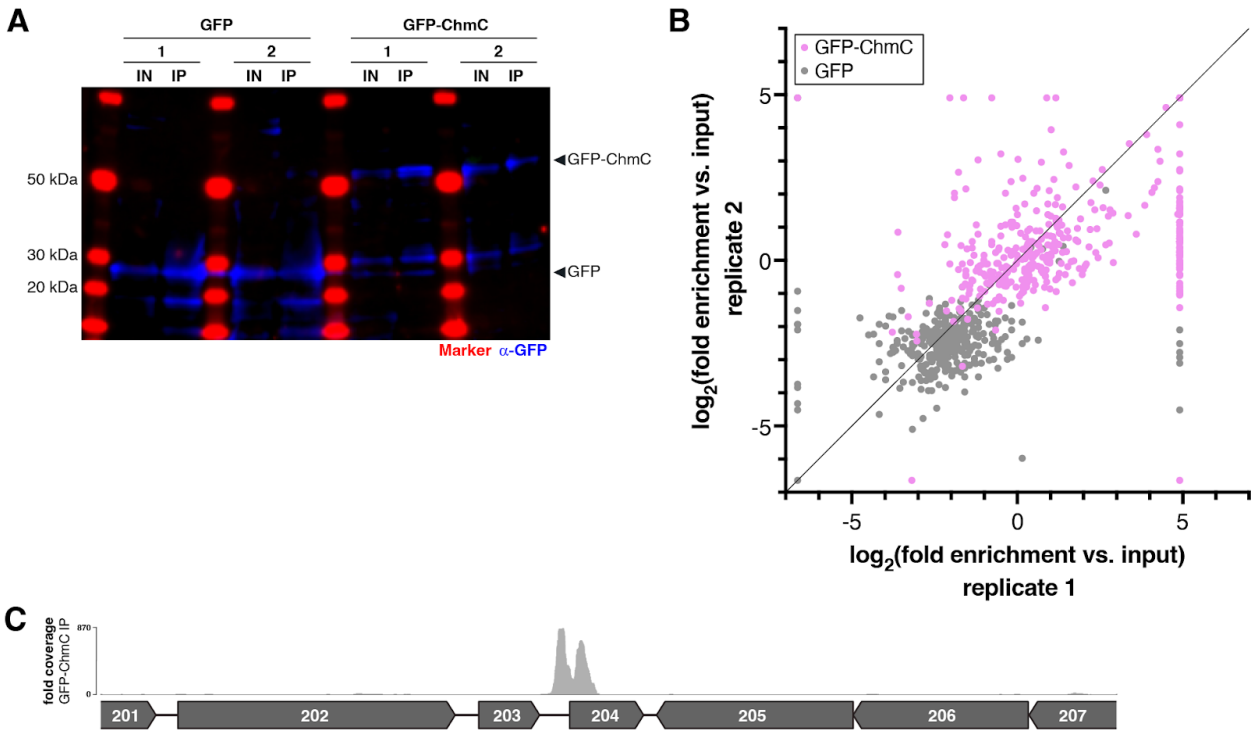


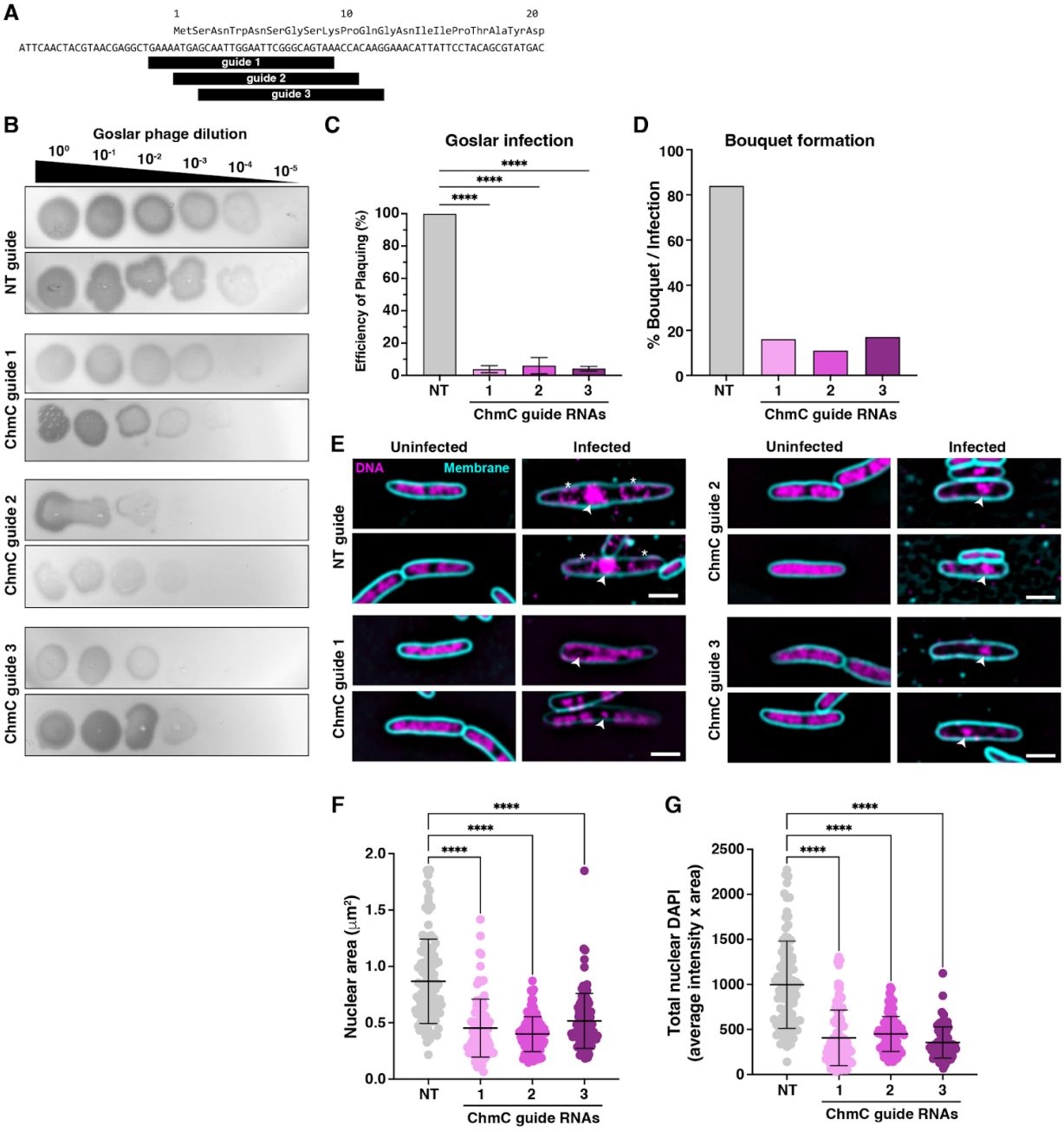
Figure 3.S4. eCLIP analysis of ChmC-RNA interactions

(A) Western blot showing levels of GFP (blue) in eCLIP input (IN) and immunopurified (IP) samples. Replicates 1 and 2 for each sample (GFP or GFP-ChmC) are labeled. Molecular weight markers are shown in red and labeled.

(B) Reproducibility of two replicates for eCLIP analysis of GFP (gray dots) and GFP-ChmC (pink dots), shown as $\log_2$(fold enrichment vs. input) for each phage gene.

(C) Fold sequence coverage from eCLIP analysis of GFP-ChmC, showing large peaks in the region upstream of PhiPA3 gp204 and downstream of the gp204 start codon.

Figure 3.S5. ChmC knockdown disrupts Goslar phage infections
(A) Design of three guide RNAs targeting the start codon of phage Goslar ChmC (gp176). (B) Plaque assays with phage Goslar on *E. coli* MC1000 cells expressing ddCas13d and either a non-targeting guide RNA (NT) or one of three guide RNAs targeting the 5' end of the ChmC (gp176) gene (guide 1, guide 2, or guide 3). (C) Efficiency of plaquing of phage Goslar on *E. coli* MC1000 cells ddCas13d and either a non-targeting guide RNA (NT) or one of three guide RNAs targeting the 5' end of the ChmC (gp176) gene (guide 1, guide 2, or guide 3). *** indicates a P value < 0.0002 in a one-way ANOVA test of statistical significance. (D) Percent of Goslar-infected *E. coli* MC1000 cells forming phage bouquets at 110 MPI. (E) Microscopy of *E. coli* MC1000 (uninfected at left; 110 MPI Goslar-infected at right) expressing ddCas13d and either a non-targeting guide RNA (top) or one of three guide RNAs targeting the 5' end of the ChmC (gp176) gene (guide 1, guide 2, or guide 3). DNA is stained with DAPI and shown in magenta; membranes are stained with FM4-64 and shown in cyan. Arrowheads indicate the phage nucleus, and asterisks indicate phage bouquets. Scale bar = 2 µm. (F) Phage nucleus area from 110 MPI Goslar-infected *E. coli* MC1000 cells expressing dCas13d and either a non-targeting guide RNA (NT; gray) or a guide RNA targeting the 5' end of the ChmC (gp176) gene (guide 1 pink, guide 2 violet, guide 3 purple). Individual data points are shown, and error bars indicate mean +/- standard deviation. **** indicates a P value of <0.0001 in a one-way ANOVA test of statistical significance. (G) Total DAPI signal within the phage nucleus (calculated as the phage nuclear area multiplied by the average DAPI intensity in the phage nucleus) from 110 MPI Goslar-infected *E. coli* MC1000 cells expressing dCas13d and either a non-targeting guide RNA (NT; gray) or a guide RNA targeting the 5' end of the ChmC (gp176) gene (guide 1 pink, guide 2 violet, guide 3 purple). Individual data points are shown, and error bars indicate mean +/- standard deviation. **** indicates a P value of <0.0001 in a one-way ANOVA test of statistical significance.


A
1
10
20
MetSerAsnTrpAsnSerGlySerLysProGlnGlyAsnIleIleProThrAlaTyrAsp
ATTCAACTACGTAACGAGGCTGAAAATGAGCAATTGGAATTCGGGCAGTAAACCACAAGGAAACATTATTCCTACAGCGTATGAC
guide 1
guide 2
guide 3
B
Goslar phage dilution
10^0
10^-1
10^-2
10^-3
10^-4
10^-5
NT guide
ChmC guide 1
ChmC guide 2
ChmC guide 3
C
Goslar infection
Efficiency of Plaquing (%)
NT
1
2
3
ChmC guide RNAs
****
D
Bouquet formation
% Bouquet / Infection
NT
1
2
3
ChmC guide RNAs
E
Uninfected
Infected
DNA
Membrane
NT guide
ChmC guide 1
ChmC guide 2
ChmC guide 3
F
Nuclear area (μm²)
NT
1
2
3
ChmC guide RNAs
****
G
Total nuclear DAPI (average intensity x area)
NT
1
2
3
ChmC guide RNAs
****

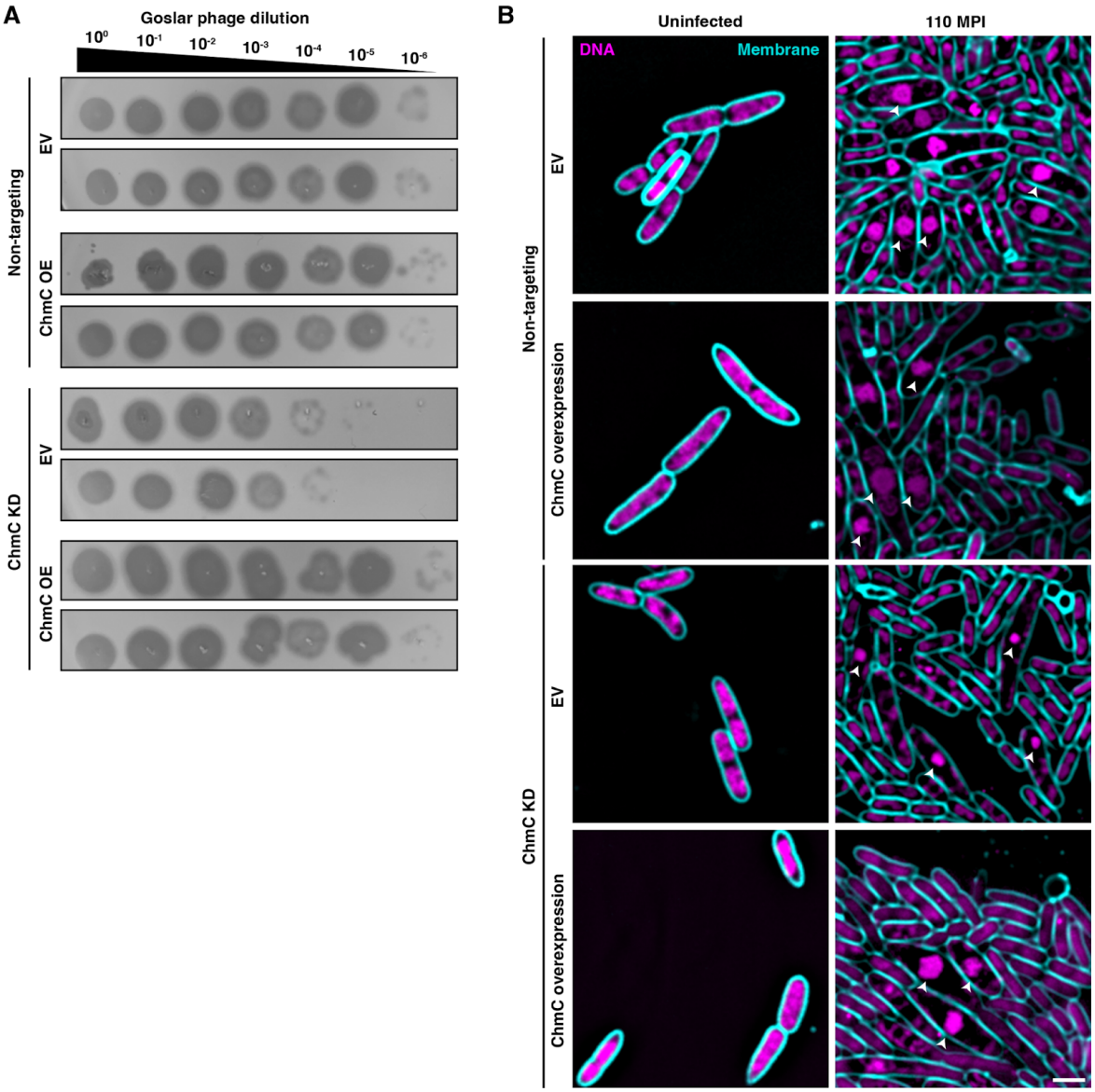


Figure 3.S6. ChmC knockdown and complementation

(A) Plaque assays with phage Goslar on *E. coli* MC1000 cells expressing ddCas13d and either a non-targeting guide RNA (NT) or a *chmC*-targeting guide RNA (ChmC KD). Host cells were additionally transformed with either an empty vector (EV) or a vector expressing recoded (ddCas13d-resistant) ChmC (ChmC OE). See quantitation in **Figure 5B**. (B) Microscopy of *E. coli* MC1000 (uninfected at left; 110 MPI Goslar-infected at right) expressing ddCas13d and either a non-targeting guide RNA (Non-targeting) or a *chmC*-targeting guide RNA (ChmC KD). Host cells were additionally transformed with either an empty vector (EV) or a vector expressing recoded (ddCas13d-resistant) ChmC (ChmC oeverexpression). DNA is stained with DAPI and shown in magenta; membranes are stained with FM4-64 and shown in cyan. Arrowheads indicate the phage nucleus. Scale bar = 2 μm.

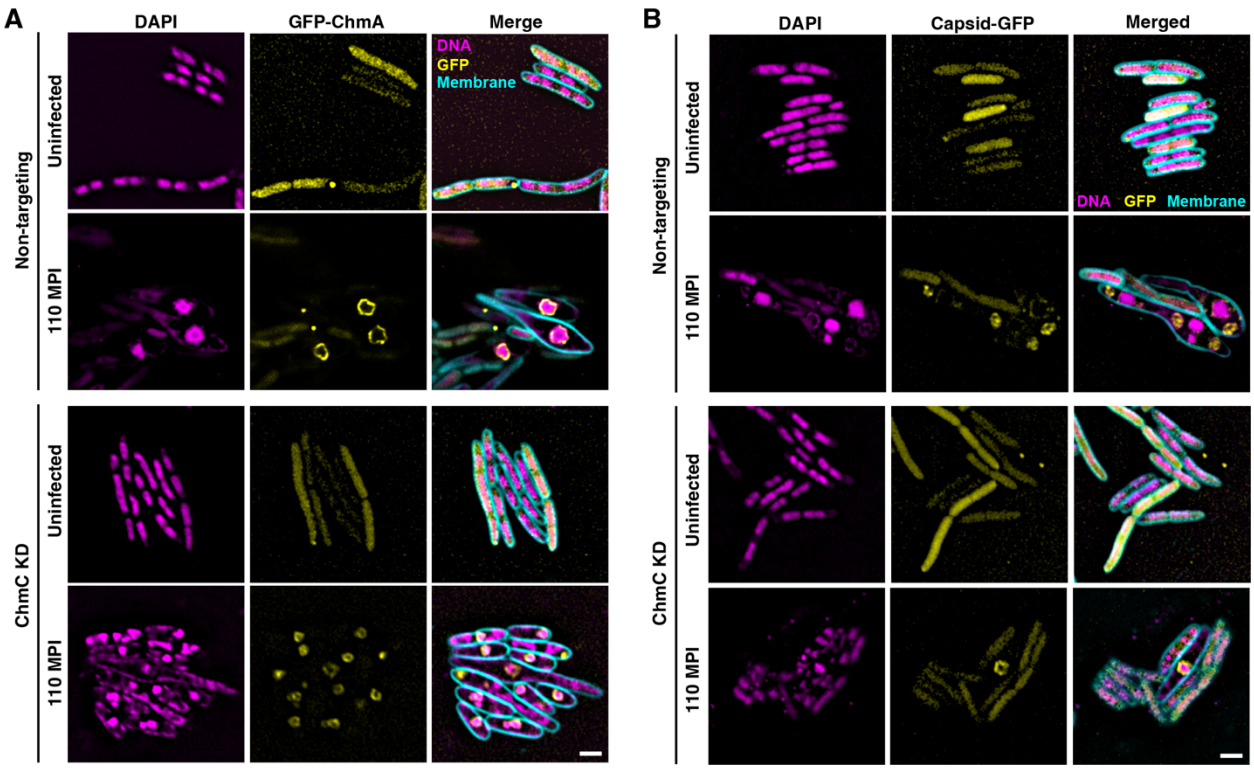


Figure 3.S7. ChmC knockdown disrupts Goslar infections and causes a global protein knockdown

(A) Microscopy of *E. coli* MC1000 cells infected with Goslar (110 MPI) expressing GFP-tagged ChmA (gp246, yellow), dCas13d, and either a non-targeting guide RNA (top) or ChmC guide 3 (bottom). Arrowheads indicate phage nuclei, and asterisks indicate phage bouquets.

(B) Microscopy of *E. coli* MC1000 cells infected with Goslar (110 MPI) expressing GFP-tagged phage capsid protein (gp41, yellow), dCas13d, and either a non-targeting guide RNA (top) or ChmC guide 3 (bottom). Arrowheads indicate phage nuclei, and asterisks indicate phage bouquets.

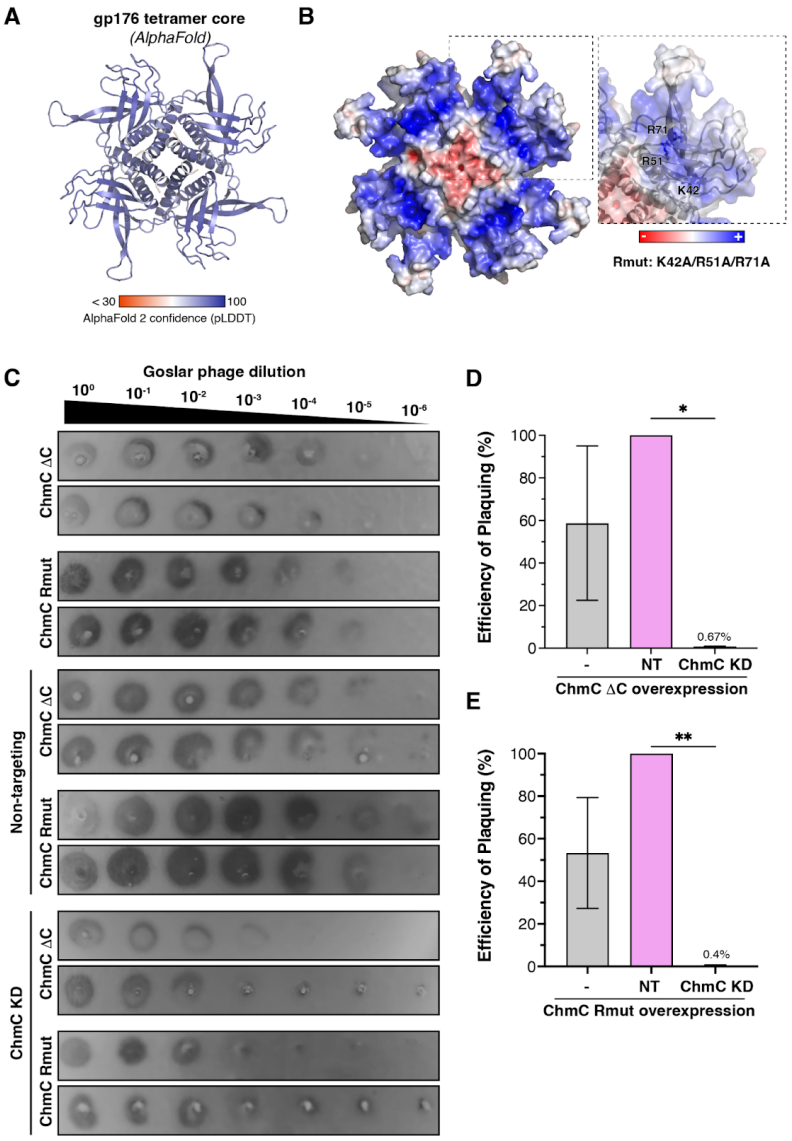


Figure 3.S8. ChmC Rmut or ΔC overexpression does not rescue the knockdown phenotype

(A) AlphaFold predicted structure of a tetramer of Goslar ChmC (gp176), colored by prediction confidence (pLDDT score). Residues 16-192 of each chain are shown; N- and C-terminal tails with high predicted disorder are not shown. (B) View equivalent to panel (A) of a tetramer of Goslar ChmC (gp176), shown as a molecular surface colored by charge as calculated by APBS [45] in PyMOL. Closeup panel shows one monomer with residues equivalent to the PhiPA3 Rmut residues: K42, R51, and R71. (C) Plaque assays with phage Goslar on *E. coli* MC1000 cells overexpressing ChmC ΔC (residues 1-203) or Rmut (K42A/R51A/R71A), or expressing ddCas13d and either a non-targeting guide RNA (NT) or a ChmC-targeting guide in addition to overexpressed ChmC ΔC or Rmut. (D-E) Efficiency of plaquing for the conditions shown in panel (C). * and ** indicate P values < 0.0332 and 0.0021 in a one-way ANOVA test of statistical significance.

3.8 References

1. Hampton, H. G., Watson, B. N. J. & Fineran, P. C. The arms race between bacteria and their phage foes. *Nature* 577, 327–336 (2020).
2. Stanley, S. Y. & Maxwell, K. L. Phage-Encoded Anti-CRISPR Defenses. *Annu. Rev. Genet.* 52, 445–464 (2018).
3. Prichard, A., Lee, J., Laughlin, T. G., Lee, A., Thomas, K. P., Sy, A., Spencer, T., Asavavimol, A., Cafferata, A., Cameron, M., Chiu, N., Davydov, D., Desai, I., Diaz, G., Guereca, M., Hearst, K., Huang, L., Jacobs, E., Johnson, A., Kahn, S., Koch, R., Martinez, A., Norquist, M., Pau, T., Prasad, G., Saam, K., Sandhu, M., Sarabia, A. J., Schumaker, S., Sonin, A., Uyeno, A., Zhao, A., Corbett, K., Pogliano, K., Meyer, J., Grose, J. H., Villa, E., Dutton, R. & Pogliano, J. Identifying the core genome of the nucleus-forming bacteriophage family and characterization of Erwinia phage RAY. *bioRxiv* (2023). doi:10.1101/2023.02.24.529968
4. Chaikeeratisak, V., Nguyen, K., Khanna, K., Brilot, A. F., Erb, M. L., Coker, J. K. C., Vavilina, A., Newton, G. L., Buschauer, R., Pogliano, K., Villa, E., Agard, D. A. & Pogliano, J. Assembly of a nucleus-like structure during viral replication in bacteria. *Science* 355, 194–197 (2017).
5. Chaikeeratisak, V., Nguyen, K., Egan, M. E., Erb, M. L., Vavilina, A. & Pogliano, J. The Phage Nucleus and Tubulin Spindle Are Conserved among Large Pseudomonas Phages. *Cell Rep.* 20, 1563–1571 (2017).
6. Knipe, D. M., Prichard, A., Sharma, S. & Pogliano, J. Replication Compartments of Eukaryotic and Bacterial DNA Viruses: Common Themes Between Different Domains of Host Cells. *Annu Rev Virol* 9, 307–327 (2022).
7. Laughlin, T. G., Deep, A., Prichard, A. M., Seitz, C., Gu, Y., Enustun, E., Suslov, S., Khanna, K., Birkholz, E. A., Armbruster, E., McCammon, J. A., Amaro, R. E., Pogliano, J., Corbett, K. D. & Villa, E. Architecture and self-assembly of the jumbo bacteriophage nuclear shell. *Nature* 608, 429–435 (2022).
8. Mendoza, S. D., Nieweglowska, E. S., Govindarajan, S., Leon, L. M., Berry, J. D., Tiwari, A., Chaikeeratisak, V., Pogliano, J., Agard, D. A. & Bondy-Denomy, J. A bacteriophage nucleus-like compartment shields DNA from CRISPR nucleases. *Nature* 577, 244–248 (2020).
9. Malone, L. M., Warring, S. L., Jackson, S. A., Warnecke, C., Gardner, P. P., Gumy, L. F. & Fineran, P. C. A jumbo phage that forms a nucleus-like structure evades CRISPR-Cas DNA targeting but is vulnerable to type III RNA-based immunity. *Nat Microbiol* 5, 48–55 (2020).
10. Nieweglowska, E. S., Brilot, A. F., Méndez-Moran, M., Kokontis, C., Baek, M., Li, J., Cheng, Y., Baker, D., Bondy-Denomy, J. & Agard, D. A. The ϕPA3 phage nucleus is enclosed by a self-assembling 2D crystalline lattice. *Nat. Commun.* 14, 927 (2023).
11. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A. W., Kavukcuoglu,

K., Kohli, P. & Hassabis, D. Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589 (2021).
12. Graebsch, A., Roche, S. & Niessing, D. X-ray structure of Pur-α reveals a Whirly-like fold and an unusual nucleic-acid binding surface. *Proceedings of the National Academy of Sciences* 106, 18521–18526 (2009).
13. Schumacher, M. A., Karamooz, E., Zíková, A., Trantírek, L. & Lukes, J. Crystal structures of T. brucei MRP1/MRP2 guide-RNA binding complex reveal RNA matchmaking mechanism. *Cell* 126, 701–711 (2006).
14. Desveaux, D., Allard, J., Brisson, N. & Sygusch, J. A new family of plant transcription factors displays a novel ssDNA-binding surface. *Nat. Struct. Biol.* 9, 512–517 (2002).
15. Mitrea, D. M., Cika, J. A., Guy, C. S., Ban, D., Banerjee, P. R., Stanley, C. B., Nourse, A., Deniz, A. A. & Kriwacki, R. W. Nucleophosmin integrates within the nucleolus via multi-modal interactions with proteins displaying R-rich linear motifs and rRNA. *Elife* 5, (2016).
16. Pak, C. W., Kosno, M., Holehouse, A. S., Padrick, S. B., Mittal, A., Ali, R., Yunus, A. A., Liu, D. R., Pappu, R. V. & Rosen, M. K. Sequence Determinants of Intracellular Phase Separation by Complex Coacervation of a Disordered Protein. *Mol. Cell* 63, 72–85 (2016).
17. Zaharias, S., Zhang, Z., Davis, K., Fargason, T., Cashman, D., Yu, T. & Zhang, J. Intrinsically disordered electronegative clusters improve stability and binding specificity of RNA-binding proteins. *J. Biol. Chem.* 297, 100945 (2021).
18. Azaldegui, C. A., Vecchiarelli, A. G. & Biteen, J. S. The emergence of phase separation as an organizing principle in bacteria. *Biophys. J.* 120, 1123–1138 (2021).
19. Nandana, V. & Schrader, J. M. Roles of liquid–liquid phase separation in bacterial RNA metabolism. *Curr. Opin. Microbiol.* 61, 91–98 (2021).
20. Lin, Y., Protter, D. S. W., Rosen, M. K. & Parker, R. Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. *Mol. Cell* 60, 208–219 (2015).
21. Schmidt, H. B. & Görlich, D. Transport Selectivity of Nuclear Pores, Phase Separation, and Membraneless Organelles. *Trends Biochem. Sci.* 41, 46–61 (2016).
22. Chong, P. A., Vernon, R. M. & Forman-Kay, J. D. RGG/RG Motif Regions in RNA Binding and Phase Separation. *J. Mol. Biol.* 430, 4650–4665 (2018).
23. Chaikeeratisak Vorrapon, Khanna Kanika, Nguyen Katrina T., Egan MacKennon E., Enustun Eray, Armbruster Emily, Lee Jina, Pogliano Kit, Villa Elizabeth & Pogliano Joe. Subcellular organization of viral particles during maturation of nucleus-forming jumbo phage. *Science Advances* 8, eabj9670 (2022).
24. deYMartín Garrido, N., Orekhova, M., Lai Wan Loong, Y. T. E., Litvinova, A., Ramlaul, K., Artamonova, T., Melnikov, A. S., Serdobintsev, P., Aylett, C. H. S. & Yakunina, M. Structure of the bacteriophage PhiKZ non-virion RNA polymerase. *Nucleic Acids Res.* 49, 7732–7739 (2021).
25. Yakunina, M., Artamonova, T., Borukhov, S., Makarova, K. S., Severinov, K. & Minakhin, L. A non-canonical multisubunit RNA polymerase encoded by a giant bacteriophage. *Nucleic Acids Res.* 43, 10411–10420 (2015).
26. M Iyer, L., Anantharaman, V., Krishnan, A., Burroughs, A. M. & Aravind, L. Jumbo Phages: A Comparative Genomic Overview of Core Functions and Adaptions for Biological Conflicts. *Viruses* 13, (2021).

27. Ceyssens, P.-J., Minakhin, L., Van den Bossche, A., Yakunina, M., Klimuk, E., Blasdel, B., De Smet, J., Noben, J.-P., Bläsi, U., Severinov, K. & Lavigne, R. Development of giant bacteriophage ϕKZ is independent of the host transcription apparatus. *J. Virol.* 88, 10501–10510 (2014).
28. Wicke, L., Ponath, F., Coppens, L., Gerovac, M., Lavigne, R. & Vogel, J. Introducing differential RNA-seq mapping to track the early infection phase for Pseudomonas phage ϕKZ. *RNA Biol.* 18, 1099–1110 (2021).
29. Krupinska, K., Desel, C., Frank, S. & Hensel, G. WHIRLIES Are Multifunctional DNA-Binding Proteins With Impact on Plant Development and Stress Resistance. *Front. Plant Sci.* 13, 880423 (2022).
30. Luijsterburg, M. S., Noom, M. C., Wuite, G. J. L. & Dame, R. T. The architectural role of nucleoid-associated proteins in the organization of bacterial chromatin: a molecular perspective. *J. Struct. Biol.* 156, 262–272 (2006).
31. Dillon, S. C. & Dorman, C. J. Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nat. Rev. Microbiol.* 8, 185–195 (2010).
32. Hołówka, J. & Zakrzewska-Czerwińska, J. Nucleoid Associated Proteins: The Small Organizers That Help to Cope With Stress. *Front. Microbiol.* 11, 590 (2020).
33. Prikryl, J., Watkins, K. P., Friso, G., van Wijk, K. J. & Barkan, A. A member of the Whirly family is a multifunctional RNA- and DNA-binding protein that is essential for chloroplast biogenesis. *Nucleic Acids Res.* 36, 5152–5165 (2008).
34. Brocca, S., Grandori, R., Longhi, S. & Uversky, V. Liquid–Liquid Phase Separation by Intrinsically Disordered Protein Regions of Viruses: Roles in Viral Life Cycle and Control of Virus–Host Interactions. *Int. J. Mol. Sci.* 21, 9045 (2020).
35. Ye, Q., Lu, S. & Corbett, K. D. Structural Basis for SARS-CoV-2 Nucleocapsid Protein Recognition by Single-Domain Antibodies. *Front. Immunol.* 12, 719037 (2021).
36. Gales, J. P., Kubina, J., Geldreich, A. & Dimitrova, M. Strength in Diversity: Nuclear Export of Viral RNAs. *Viruses* 12, (2020).
37. Richard Evans, Michael O'Neill, Alexander Pritzel, Natasha Antropova, Andrew Senior, Tim Green, Augustin Žídek, Russ Bates, Sam Blackwell, Jason Yim, Olaf Ronneberger, Sebastian Bodenstein, Michal Zielinski, Alex Bridgland, Anna Potapenko, Andrew Cowie, Kathryn Tunyasuvunakool, Rishub Jain, Ellen Clancy, Pushmeet Kohli, John Jumper & Demis Hassabis. Protein complex prediction with AlphaFold-Multimer. *bioRxiv* 2021.10.04.463034 (2022).
38. Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S. & Steinegger, M. ColabFold: making protein folding accessible to all. *Nat. Methods* 19, 679–682 (2022).
39. Fraser, A., Sokolova, M. L., Drobysheva, A. V., Gordeeva, J. V., Borukhov, S., Jumper, J., Severinov, K. V. & Leiman, P. G. Structural basis of template strand deoxyuridine promoter recognition by a viral RNA polymerase. *Nat. Commun.* 13, 3526 (2022).
40. Blue, S. M., Yee, B. A., Pratt, G. A., Mueller, J. R., Park, S. S., Shishkin, A. A., Starner, A. C., Van Nostrand, E. L. & Yeo, G. W. Transcriptome-wide identification of RNA-binding protein binding sites using seCLIP-seq. *Nat. Protoc.* 17, 1223–1265 (2022).
41. Lovci, M. T., Ghanem, D., Marr, H., Arnold, J., Gee, S., Parra, M., Liang, T. Y., Stark, T. J., Gehman, L. T., Hoon, S., Massirer, K. B., Pratt, G. A., Black, D. L., Gray, J. W.,

Conboy, J. G. & Yeo, G. W. Rbfox proteins regulate alternative mRNA splicing through evolutionarily conserved RNA bridges. *Nat. Struct. Mol. Biol.* 20, 1434–1442 (2013).
42. Li, Q., Brown, J. B., Huang, H. & Bickel, P. J. Measuring reproducibility of high-throughput experiments. *aoas* 5, 1752–1779 (2011).
43. Yee, B. A., Pratt, G. A., Graveley, B. R., Van Nostrand, E. L. & Yeo, G. W. RBP-Maps enables robust generation of splicing regulatory maps. *RNA* 25, 193–204 (2019).
44. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P. & Cardona, A. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* 9, 676–682 (2012).
45. Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U. S. A.* 98, 10037–10041 (2001).
46. Chu, X., Sun, T., Li, Q., Xu, Y., Zhang, Z., Lai, L. & Pei, J. Prediction of liquid–liquid phase separating proteins using machine learning. *BMC Bioinformatics* 23, 1–13 (2022).

3.9 Acknowledgments

The authors thank Elizabeth Villa and members of the Corbett and Pogliano labs for helpful discussions. The authors acknowledge support from the National Institutes of Health (R01 GM129245 to J.P.; R35 GM144121 to K.D.C.; S10 OD021724 for shared mass spectrometry resources). This research was supported in part by the Howard Hughes Medical Institute Emerging Pathogens Initiative (to J.P. and K.D.C.).

Chapter 3, in full, is currently being prepared for submission for publication of the material. Enustun, E., Armbruster, E., Lee, J., Zhang, S., Lee, B., Gu, Y., Deep, A., Naritomi, J., Liang, Q., Aigner, S., Chaikeeratisak, V., Cleveland, D., Ghassemian, M., Yeo, G., Pogliano, J., Corbett, K.D. A phage nucleus-associated RNA-binding protein required for jumbo phage infection. The dissertation author was the primary investigator and author of this material.

Chapter 4: Other key proteins and their roles in replication system with phage nucleus

4.1 Abstract

The infection mechanism involving the phage nucleus requires various proteins that are yet to be identified. Chapter 4 of this thesis presents the findings on several phage proteins that are key to the successful replication of phages. First, we investigated the portal protein, which is known to be a crucial part of the phage body and involved in DNA packaging. Our research revealed that the portal protein may have evolved to accommodate the subcellular localization of the phage DNA. While it was previously shown that empty capsids are trafficked by PhuZ to be docked on the phage nucleus, we believe the portal protein evolved to bind the pore-like structures formed by ChmB, while conserving the main features of portal proteins from small genome phages. Next, we investigated the phage nucleus-associated proteins gp63 and gp64 of PhiPA3, which potentially have key roles in selective protein import. Our results so far suggest that gp63 is crucial for interacting with proteins that get imported and interact with ChmB, where proteins potentially enter the phage nucleus. Finally, we examined the phage-encoded proteins that potentially interact with host cell division proteins. Our findings so far suggest these proteins play a role in the phages' ability to hijack the host cell division to inhibit it, allowing for the successful completion of phage replication. Taken together, our results provide valuable insights into the phage nucleus system and complex interactions of jumbo phage proteins. We believe these results will be crucial for future studies and understanding of the replications with phage nucleus.

4.2 Introduction

4.2.1 Portal

Likely due to their larger genomes, jumbo phages are morphologically larger than small genome phages, with larger capsid size and elongated tail[1]. The phage portal protein is a virion structural protein that forms a ring-shaped channel and forms the base of the phage capsid[2]. It has a key role of allowing genomic material in and out of the capsid and has been shown to be important for capsid assembly as well as DNA packaging. Due to their importance for phage replication, portal proteins' structures and function have been extensively studied in small genome phages, including T4, P22, and phi29[2].

Although it has been heavily investigated to understand DNA packaging and has been a target for antiviral therapies, we do not know how portal proteins evolve, especially for our system to accommodate the phage nucleus system. It has been found that the capsid proteins are trafficked to the phage nucleus by the tubulin homolog PhuZ[3–5]. As portal proteins are important for capsid formation, it is likely that the portal protein is also carried by PhuZ to dock on the phage nucleus for DNA packaging. It is believed that the DNA is likely exported through channels or pores in the protein shell of the structure. Alternatively, the DNA may be exported from the nucleus-like structure in a process similar to the way mRNA is exported from the nucleus of eukaryotic cells.

In eukaryotic cells, specific proteins form the nuclear pore complex that allows for the selective transport of molecules such as RNA and proteins between the nucleus and the cytoplasm[6]. It is possible that jumbo phage proteins might have evolved to form

channels or pores that allow for the selective transport of DNA and other molecules. Alternatively, DNA may be exported with phage proteins that interact with DNA through specific interactions to open the pores on the phage nucleus and allow DNA to pass through. When investigated, portal proteins of the phage nucleus forming virions are generally structurally conserved but longer (See Chapter 1). This is likely to accommodate the larger capsid size needed to contain the larger genome. However, the extra domains of these portal proteins might also be important to interact with other proteins and dock on the phage nucleus. Further research is needed to fully understand the exact mechanism of DNA export from the phage nucleus to capsid through portal proteins without exposing the DNA in the host cytoplasm.

4.2.2 Protein import

Protein translocation is a fundamental process for all domains of life, enabling the transport of selected proteins into a cell or a specific compartment such as mitochondria or the endoplasmic reticulum (ER)[7–9]. It is essential for replication, cell growth, differentiation, and function[10]. The process is usually complex and involves several steps, including the recognition of the protein by the translocator and the translocation of the protein across layers of the compartment [11]. In the target compartment, the translocated protein can be folded or processed to perform its function(ref).

In eukaryotic cells, protein import is heavily studied and found to be mediated by specific transporters such as Importin[11,12]. These transporters can be detected in the plasma, cytoplasm, or at the membranes of the compartments. The process is regulated and can be influenced by various factors, such as the cellular environment, the nature of

the protein, including folding and post-translational modification or signaling[13–16]. These processes are extensively studied to understand the molecular mechanisms as it is crucial for cellular functions or developing new therapies caused by defects in protein import or trafficking.

In eukaryotic cells, protein import to the nucleus occurs through the recognition of nuclear localization signals (NLS)[17]. The Importin-α protein recognizes these sequences and binds to them[18]. Importin-β then binds to this complex and interacts with the nuclear pore complex to facilitate import through nucleoporins [19]. While the observations share similarities with eukaryotic nuclear import, the import mechanism to the phage nucleus has not been fully identified or characterized. Previous studies discovered that proteins involved in DNA replication or RNA transcription localize exclusively to the phage nucleus [20]. However, proteins involved in translation or metabolic functions, as well as host defense proteins, are excluded from the nucleus-like structure [20,21]. It is argued that these translocated proteins are recognized and allowed to pass through channels located on the nucleus-like structure [22]. Certain phage proteins likely evolved for this detection and interact with these pores to allow them to enter or pass through these pores with them to the interior side.

We previously showed that the GFPmut1 localizes inside of the PhiKZ phage nucleus, while the sfGFP version does not get imported[22]. Interestingly, while they are closely related, during the PhiPA3 infection neither of the GFP versions localize in the phage nucleus when expressed alone[22]. This observation suggests these phages might have evolved to have differences in their import systems, possibly to avoid misregulations

of protein trafficking as both of these phages use *Pseudomonas Aeruginosa* cells as the host. Further research is needed to identify and characterize these proteins and the full mechanism.

4.2.3 Cell division inhibition

Phages have evolved to replicate by using some of the host cell's structures and resources[23]. For lytic phages, the time window of replication is often correlated with the host cell's division cycle[24]. This is possibly how the phages ensure that the host does not divide during the formation of progeny phage particles and maximize the chances of successful phage spread.

Bacterial cell division is a highly regulated process that ensures reproduction while maintaining genomic material, shape, and size[25]. Once DNA is replicated and chromosomes are segregated by cytoskeletal elements, the cell membrane begins to constrict to separate the cells[26]. While studies are still ongoing to describe the full composition, we know the constriction is driven by the Z-ring, which is mainly composed of FtsZ[27].

The regulation of cell division depends on several processes, as it is crucial for cell survival. One of these regulators is the Min system, a pathway that prevents Z-ring formation at incorrect positions[28]. The Min system is mainly composed of MinC, MinD, and MinE. MinC binds to FtsZ at the division site, and MinD is a membrane-associated ATPase that forms oscillatory waves[28]. MinE regulates MinD binding at the poles of the membrane and allows it to move towards the cell's cente[28]r. This cross-check allows for concentration gradients to be either at the cell poles, where Z-ring formation is prevented,

or at the midcell, where FtsZ allows Z-ring formation[28]. It was reported that MinD oscillation is important to determine the center of the cell before Z-ring formation[28].

Bacterial cell division mechanism and defects in it has importance in the context of diseases and disorders, such as bacterial infections and antibiotic resistance[29]. The study of bacterial cell division is an active and crucial area of research. It has potential implications for developing new treatments and therapies for these conditions[30]. Therefore, understanding how jumbo phages that encode phage nuclei regulate and inhibit host cells is essential.

4.3 Results

4.3.1 Evolution of the phage portal protein and its docking on the phage nucleus

When the portal protein of phiPA3 (gp148) is isolated, we confirmed the structure is generally conserved compared to small genome phages. It forms a dodecameric ring-shaped structure where one end is the base of the phage capsid, and the other end is the gateway for genetic material. In the phage nucleus system, since the genomic material is packaged through the phage nucleus, it is likely that the portal protein co-evolved. When the monomers of PhiPA3 portal and T4 portal proteins [31] are superimposed, the extra domains of phage nucleus forming phages' portal proteins are likely to be responsible for interaction with the phage nucleus and the docking for DNA packaging.

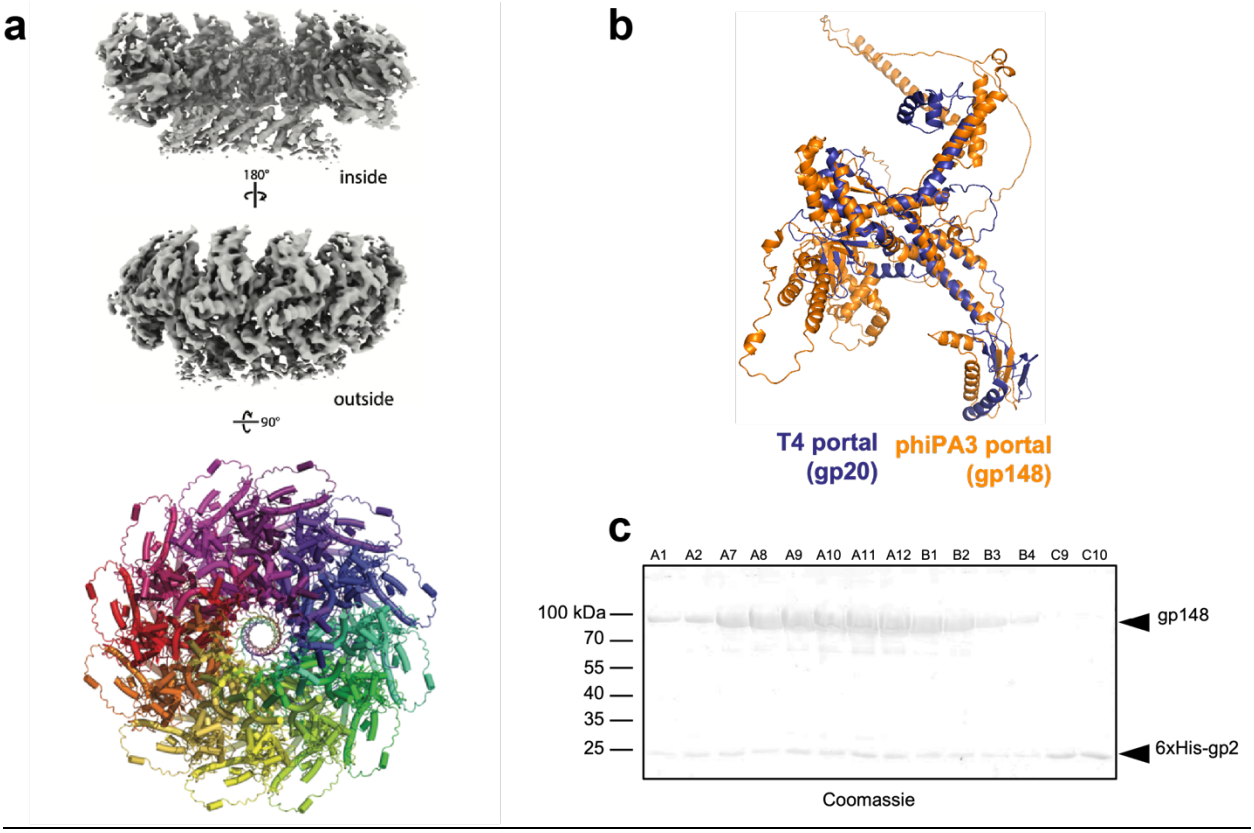


Figure 4.1: Portal proteins in phages replicate with phage nucleus systems

(a) CryoEM structure of gp148 PhiPA3 in a dodecameric assembly (b) Comparison of portal proteins from phiPA3 and T4 by superimposing them in monomeric forms (c) Coomassie stained SDS-PAGE gel from the fractions of His-tagged gp2 and no tag gp148 after size-exclusion chromatography.

Our previous work found the phage nucleus component ChmB interacts with phage portal protein (Chapter 1). Our interaction networks suggest ChmB might be where the portal protein is docking on the phage nucleus. Co-expression of His-tagged ChmB (gp2) with untagged portal protein (gp148) shows the interaction of the proteins after Ni^{2+} pulldown. The interaction of the proteins is stable as the higher fractions contain both proteins after 2 times of size-exclusion chromatography. However, due to the large size of dodecameric portal protein and its flexible loops, we could not detect the gp2 in high resolution from CryoEM.

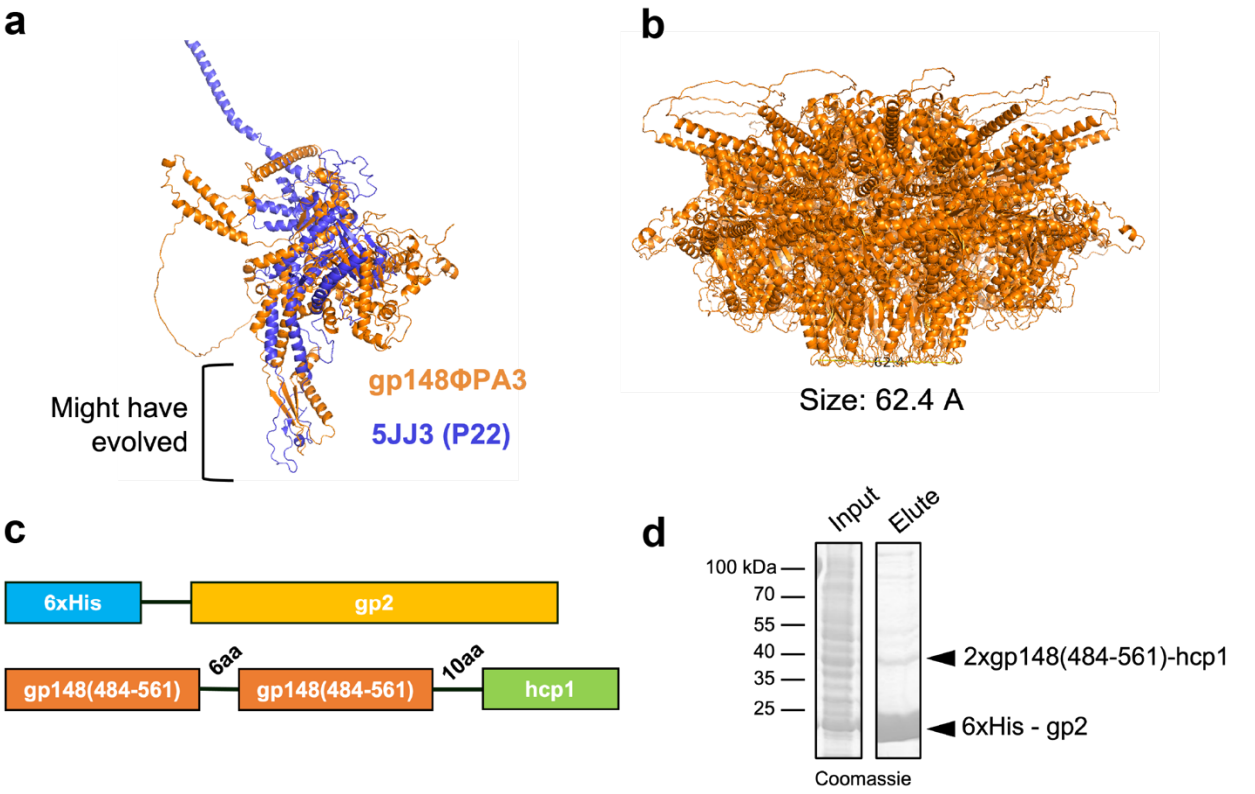


Figure 4.2. Evolution of Portal protein to accommodate the phage nucleus

(a) Comparison of domains for similar length Portal protein (gp148) of phiPA3 and P22 portal protein by superimposition (b) Dodecameric model of gp148 which shows the base part predicted to be evolved to dock on the phage nucleus has size of 62.4A (c) plasmid maps to test the interaction of phiPA3's gp2 and base part of portal fused to hcp1 that hexamerize (e) Coomassie stained gel showing His-tagged gp2 co-purifies portal base in dodecamer form.

As an alternative approach, we decided to investigate the part of the protein that might have evolved to interact with the phage nucleus. In the literature, although it is not a jumbo phage, the portal protein from P22 [2] has similar length and folding to phiPA3. When the monomers of these proteins are superimposed, the base ends of the proteins appear distinctively different from each other. In a dodecameric formation, this part of PhiPA3 portal would form a ring shape with a size of 62.4A. To test if this region is responsible for the interaction, we cloned a plasmid that expresses two regions of interest of the portal

protein (484-561) with a 6 aa linker that allows the space that these two pieces would normally have in a dodecameric form. These two portal parts were fused with the hcp1 protein, which hexametrizes to imitate the original dodecamer structure. Co-purification of this construct with His-tagged gp2 showed that these parts can interact with ChmB.

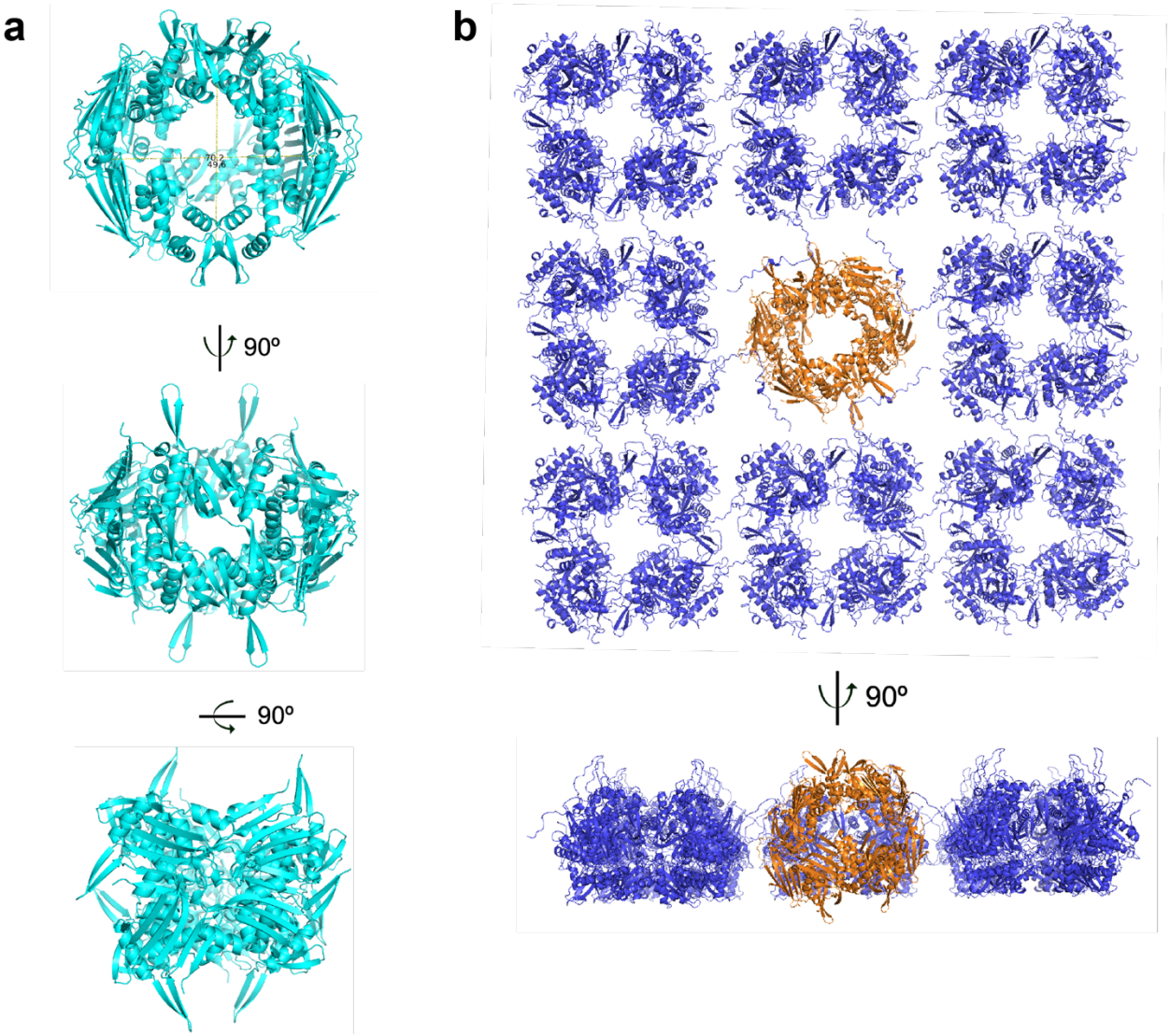


Figure 4.3. Potential docking site of Portal protein on the phage nucleus

(a) Octamer model of phiPA3 gp2 from its solved structure in dimer state. The pore size is predicted to be 70.2A (b) Octamer model of gp2 (orange) placed in a sheet of ChmA (blue) that forms the phage nucleus. Portal protein is predicted to be docking on orange site for DNA packaging.

The calculated size of the portal base may indicate a higher oligomeric assembly than the ChmB structure solved by X-ray crystallography. When we model an octameric state from the solved structure of ChmB, we estimate the pore size can reach 70.2A, which can accommodate the portal protein's docking to our pore hypothesis. Similarly, when this prediction is placed on the current model of the phage nucleus with ChmA, we see that it fits well as a volume and can interact with the free N and C tails of ChmA.

Recently available AlphaFold v2.3 successfully completed multimeric prediction with 8 monomer ChmB input. The PAE graph shows that all 8 subunits are confidently predicted to be interacting with each other. The model created by the predicted interactions has similar geometry to the solved structure of ChmB but has a smaller pore size than the hypothetical "open pore" model. Potentially, the prediction model might be representing a natural "closed" state that turns into an alternative state to allow translocations of molecules.

When we calculate the electrostatic charge distribution of the AlphaFold predicted octamer model, we observe the sides around the pores are negatively charged with patches of positively charged residues. However, the inside of the octamer model shows that the inside surface-exposed region has a significant positive charge, which might be important in trafficking the negatively charged DNA for packaging. Future studies are required to prove the existence of the octamer model and conclude how the DNA is translocated from the phage nucleus.

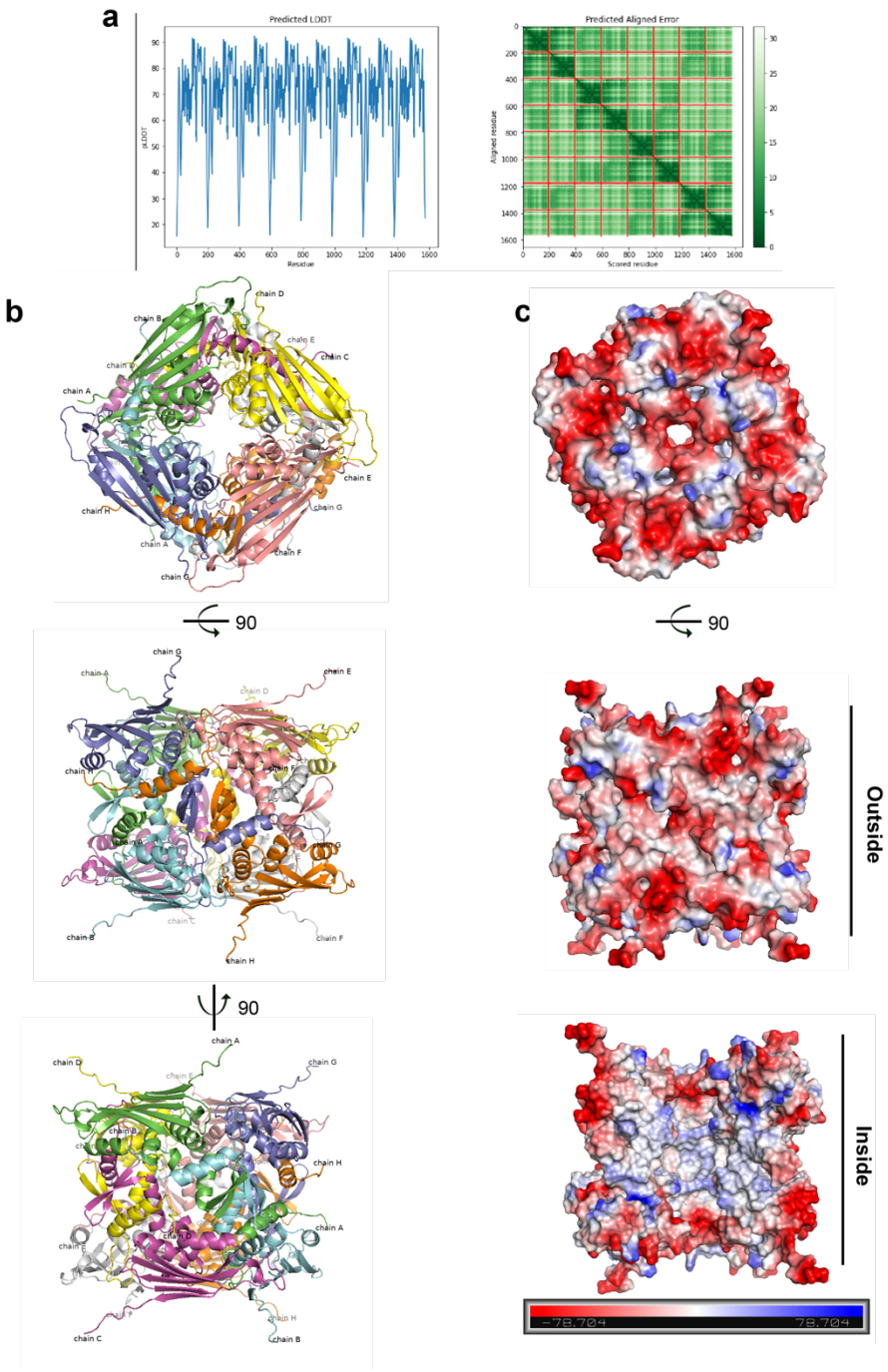


Figure 4.4. AlphaFold predicted octamer formation of ChmB

(a) AlphaFold[34] prediction parameters of 8 monomer ChmB interacting to form an octomer (b) Predicted structure of a ChmB octomer (c) Electrostatic charge distribution of the ChmB octomer prediction. (Top) the side view and (middle and bottom) 90 degrees rotated version to show outside and inside charge distribution.

4.3.2 Components of selective protein translocation system of the phage nucleus

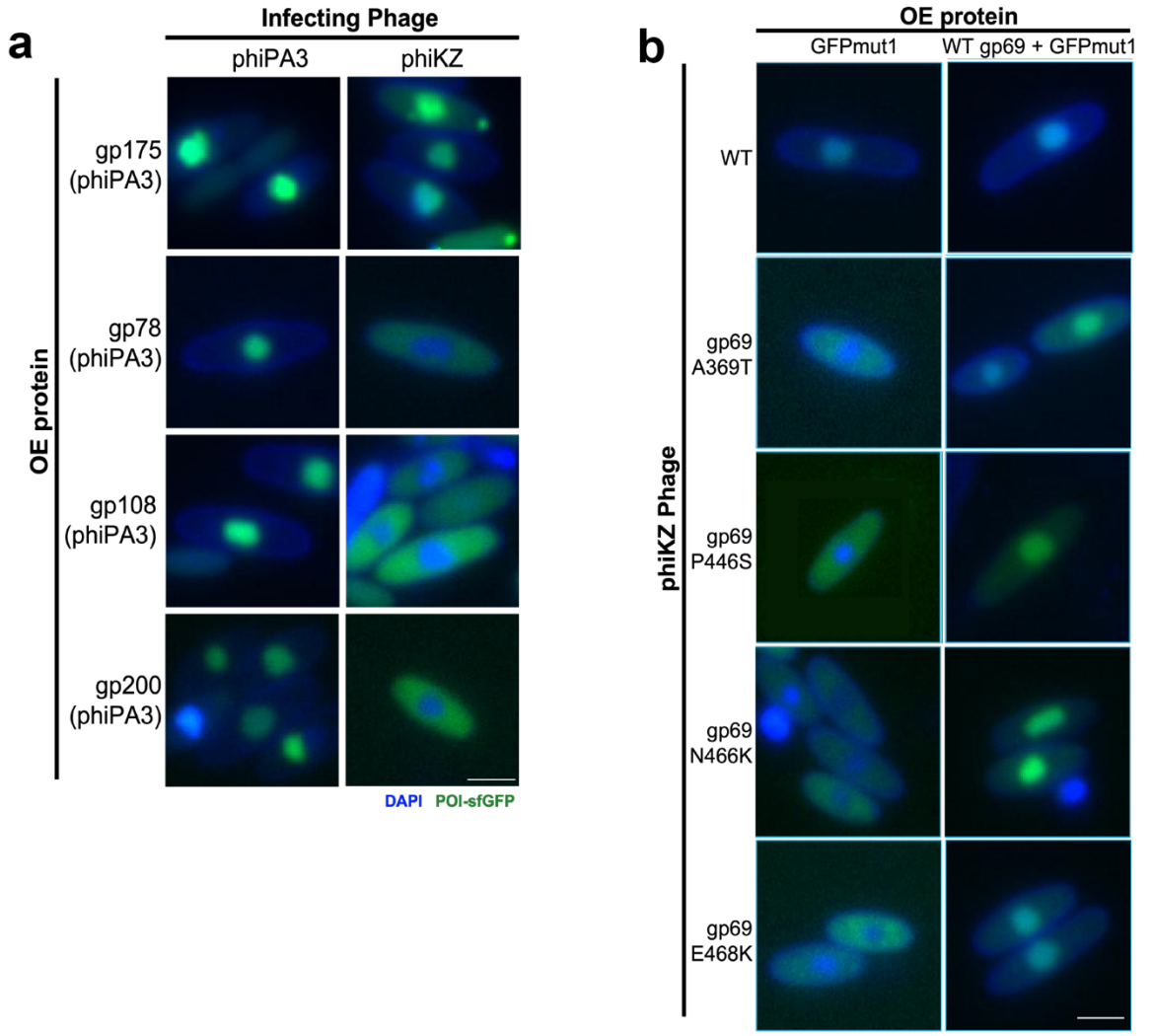


Figure 4.5. Phage nucleus system has an elaborate protein import mechanism

(a) Fluorescent microscopy that shows the difference of localizations during phiPA3 and phiKZ infections for GFP-tagged phiPA3 proteins that previously identified to be inside of the phage nucleus during infection. GFP is shown in green and DAPI (to visualize nucleic acids) in blue. Scale bar = 2 μm. (b) Scale bar = 2 μm.

Previous studies, including miniTurboID with RecA and ChmA, have identified certain phage proteins found to be inside of the phage nucleus (Chapter 1). When tested using fluorescent microscopy, most of the phiPA3 proteins that localize in the phage nucleus remained in the cytoplasm during phiKZ infection. Although these two phages are both in the phiKZ-like family and infect Pseudomonas Aeruginosa, this suggests that

they have evolved different selective protein translocation systems to enter the phage nucleus.

Interestingly, previous work has established that the GFPmut1 version of GFP is imported into the phage nucleus of phiKZ but not phiPA3 [22]. Using this, endonucleases were forced into phiKZ, targeting the phage DNA and leading to the isolation of escaper phages with mutations in gp69 of phiKZ, which is the homolog of miniTurboID hit gp63 of phiPA3. When GFPmut1 was tested for importing into the phage nucleus with the mutant phages, it remained in the cytoplasm. However, when the WT of gp69 is overexpressed from a plasmid while these phages are infecting, the GFPmut1 import is rescued. Our hypothesis is the gp63phiPA3 and its homologs function in selective protein translocation to the phage nucleus.

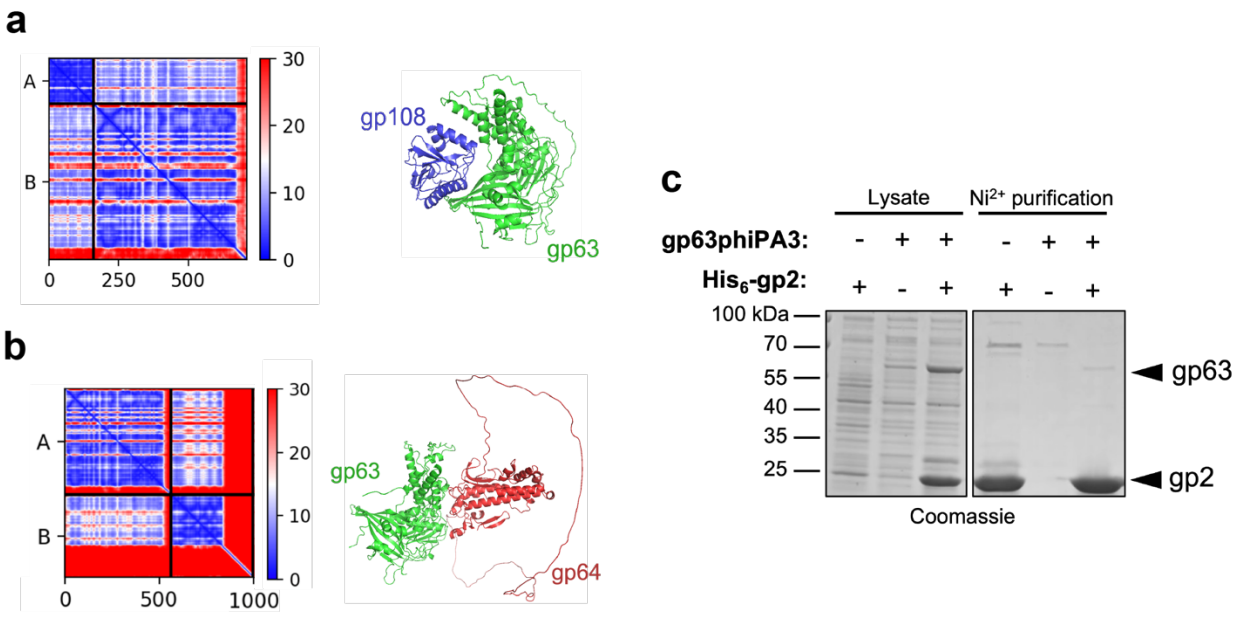


Figure 4.6. Interactions of gp63 with other phiPA3 proteins
Alphafold[34] interaction predictions between (a) phiPA3 protein that gets imported to the phage nucleus (gp108) and gp63 which potentially plays a role in selective protein importation (b) 2 proteins that localize on the phage nucleus with unknown functions of phiPA3 proteins gp63 and gp64 (c) Coomassie stain gel that shows the co-purification of phiPA3 proteins His tagged gp2 and untagged gp63 after Ni2+ pulldown.

One way to determine if gp63 is involved in protein import would be to test if it interacts with other proteins involved in this mechanism. AlphaFold confidently predicted a potential interaction between gp63 and gp108, a protein known to be imported during phiPA3 infection. Additionally, gp63 is strongly predicted to interact with another phage nucleus-localizing protein, gp64 of phiPA3. However, gp64 has a long and disorganized tail, making it difficult to predict all its residues confidently. It is possible that gp64 may also play a role in selective protein import to the phage nucleus, and the long and disordered C-terminus could be involved in phase separation or protein-protein interactions necessary for detecting which proteins to import. Finally, co-precipitation experiments demonstrate that his-tagged gp2 is directly interacting with gp63. This result suggests that the pores formed by gp2 on the phage nucleus might serve a multipurpose role, including protein translocation.

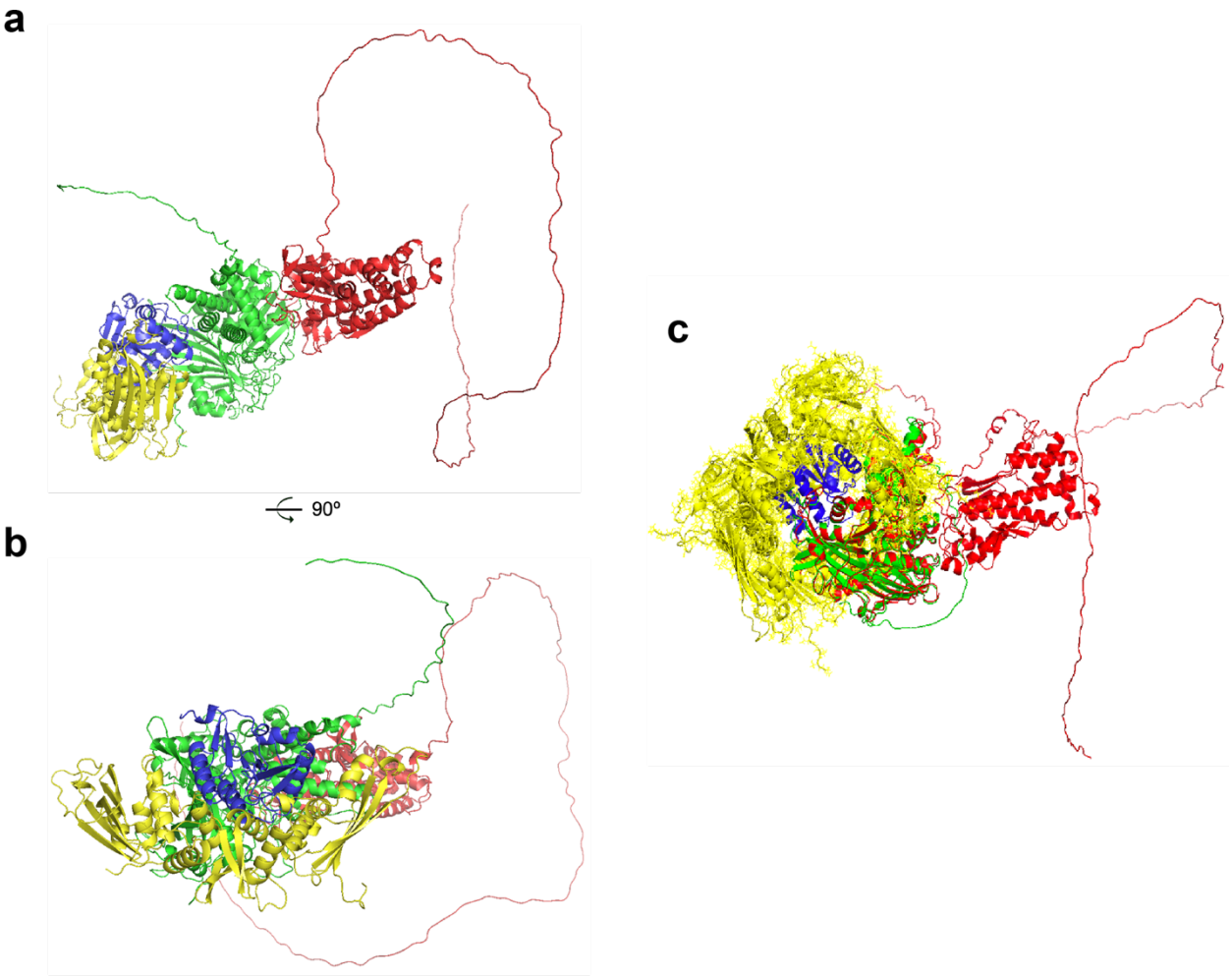


Figure 4.7. Prediction model for protein import to phage nucleus
Model represents potentially how a protein might be selectively translocating to inside of the phage nucleus (a) as a side view (b) from the point that is inside of the phage nucleus (c) with a model of gp2 that has a higher oligomeric assembly predicted by AlphaFold[34]. Yellow is gp2 (pore), blue is gp108 (imported protein), green is gp63 and red is gp64 (functions in the selective protein translocation).

By using the solved structure of ChmB and AlphaFold interaction predictions, we created a model that visualizes our hypothesis regarding protein import to the phage nucleus. In this model, the imported protein gp108 interacts with gp63 and gp64, which bind to the phage nucleus by interacting with ChmB, allowing gp108 to pass through the curvature of ChmB. While an octamer of ChmB in the model exceeds the limitations of AlphaFold, it was able to predict a hexameric assembly. When the solved structure of

ChmB is replaced with a higher oligomeric assembly, the model shows that the pore with the higher oligomeric state is wide enough to accommodate gp108. Future directions of this project would involve confirming these interactions through in vitro experiments and investigating how gp63 can detect proteins such as gp108 to selectively import them.

4.3.3 Phage proteins to control host cell division during infection

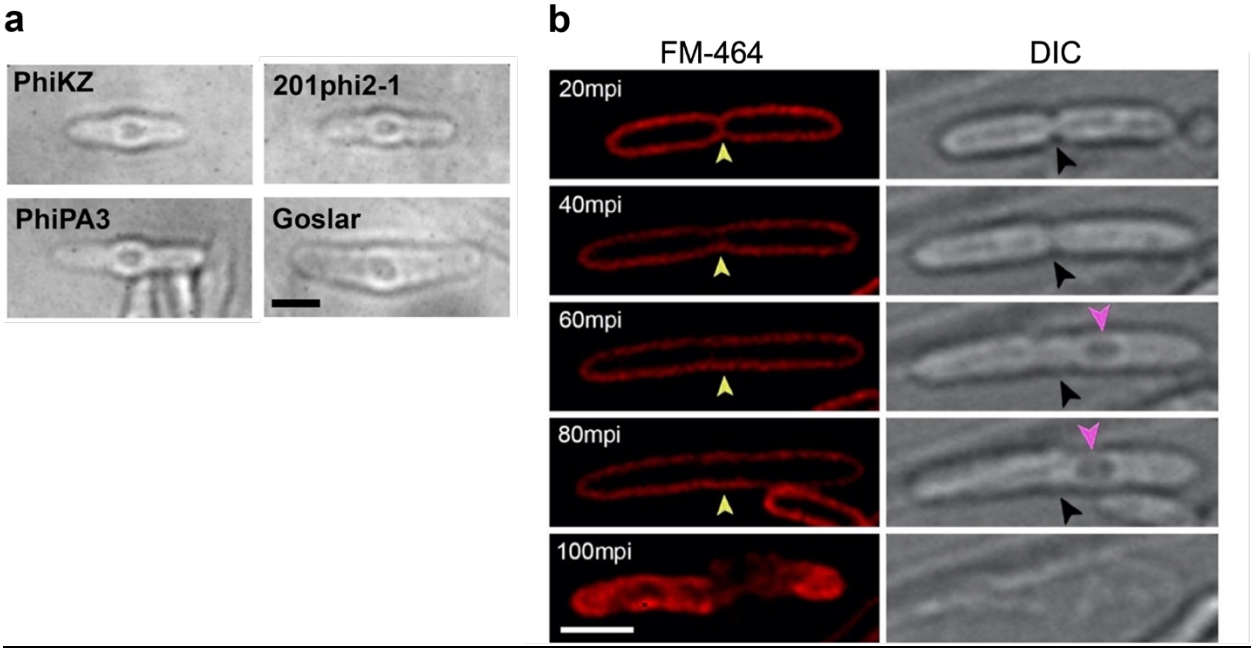


Figure 4.8. Bulge formation and host cell division inhibition at phiKZ-like phage infection
(a) DIC images that show the bulge formation in the middle of the host cell during the infections of phages that replicate with phage nucleus formation. Scale bar = 2 μm. (b) Septal reversal that shown yellow arrow at FM-464 images and black arrow with DIC images during a phiKZ infection. Pink arrow shows the phage nucleus. FM4-64 (to visualize membranes) in red. Scale bar = 2 μm.

During the infection of phage nucleus-encoding phages, we observe a bulge at the midcell with a variety of hosts including *P. chlororaphis*, *P. aeruginosa*, and *E. coli* [3,5,20]. However, we believe this is directly the result of the phage nucleus location as the same phenotype was observed during the infections with catalytically mutant PhuZ was expressed and phage nucleus was mislocalized[20]. Due to the bulge formation occurring

at the midcell, our working hypothesis is these phages disrupt the host cell division machinery.

Jumbo phages' infections are reported to take significantly longer than small genome phage, potentially correlated to their larger genome size[3,5,20]. While small genome phages are reported to be in line with host cells' division timescale, phage nucleus forming jumbo phages lyse the hosts around 2 hours which is 4 times longer than average host division cycle. This suggests jumbo phages might be encoding proteins to inhibit host cell division machinery to make sure successful formation of their progeny. In line with this, we observed some host cells had reversal of septal formation and the cell stays intact although longer than usual.

When we check the localization of FtsZ of PAO1 before phiPA3 infection, it forms a ring at the midcell, where bulging later occurs. At 45MPI, the GFP-tagged FtsZ colocalizes with the bulge, suggesting the rearrangement of the cell division machinery. During our miniTurboID to find potential interaction partners of ChmA, we identified 2 phiPA3 proteins gp259 and gp144 that localize similarly to the FtsZ before and after the infection. Interestingly, the overexpression of gp144 is sufficient for the bulge formation in the uninfected cell. The expression of gp144 homolog from phage Churi confirms this observation. Potentially, these proteins might be directly interacting with the host cell division or allow hijacking other host proteins to disrupt the division. Another protein we identified in the same analysis, gp294, causes cells to be significantly longer, suggesting that the overexpression of this protein is sufficient to block the cell division.

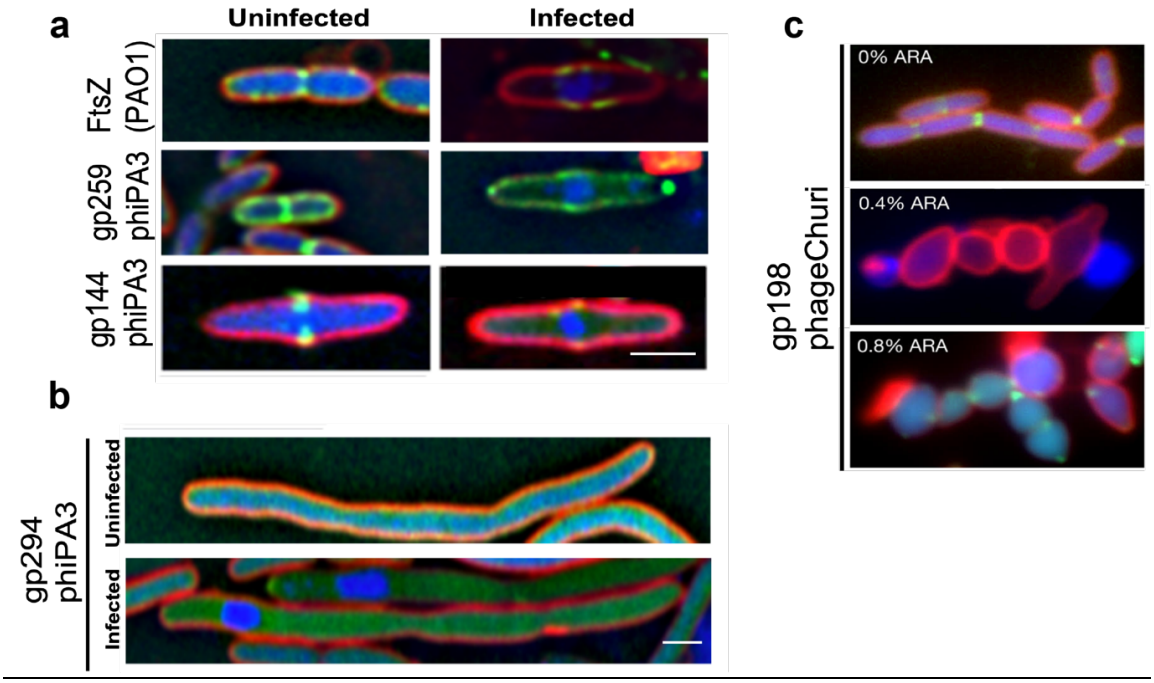


Figure 4.9. phiPA3 proteins interact with host cell division mechanism
(a) Uninfected and infected images show subcellular localization of (top) GFP-tagged FtsZ of PAO1, host of phiPA3. PhiPA3 proteins identified from miniTurboID gp259 (middle) and gp144 (bottom) similar subcellular localization to FtsZ (b) Fluorescence microscopy of gp294 phiPA3 expressing K2733 (c) subcellular localization of gp144phiPA3 homolog of gp198 found in phage Churi. GFP is shown in green, FM4-64 (to visualize membranes) in red, and DAPI (to visualize nucleic acids) in blue. Scale bar = 2 μm.

To identify which host proteins they were interacting with, we performed immunoprecipitation of GFP-tagged phiPA3 proteins gp144, gp259, and gp294. Coomassie staining and silver staining of the eluates after the immunoprecipitation showed multiple intense bands, especially at ~30 kDa for gp144, ~65 kDa for gp259, and ~55 kDa for gp294. We excised the corresponding bands from the polyacrylamide gel and identified the proteins with mass spectrometry. Based on these results, we believe that MinD is the interaction partner of gp144, GroEL may be interacting with gp259, and gp294 is interacting with FtsZ.

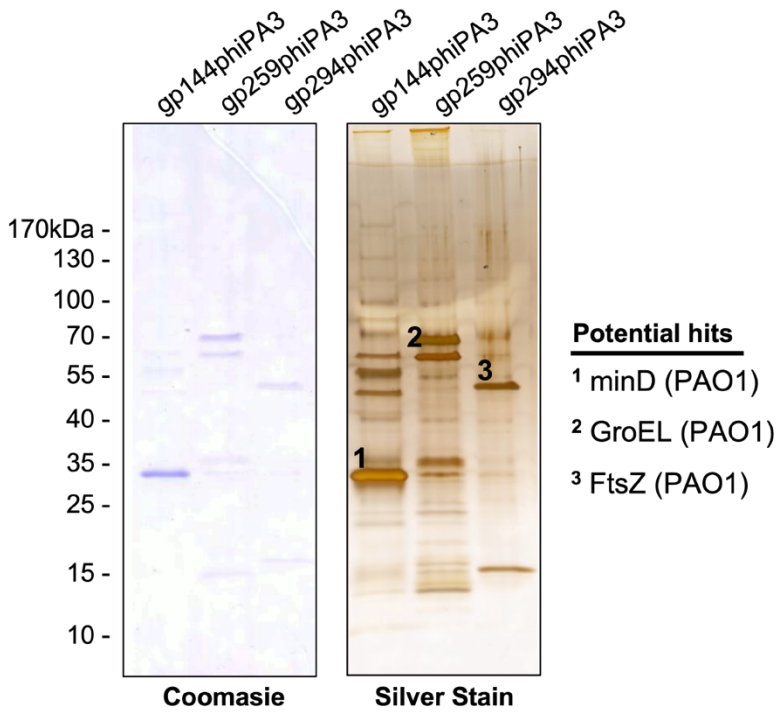


Figure 4.10. Identification of host cell division proteins as potential interaction partners

Protein gels stained with (left) Coomassie (right) Silver Stain shows the GFP-Trap pulldown of GFP-tagged gp144, gp259 and gp294 of phiPA3. Potential hits were identified from mass spectrometry of bands excised after PAGE.

4.4 Discussion

Our results suggest that while the structure of the dodecameric portal protein is generally conserved, it may have coevolved with the phage nucleus system. As ChmA-encoding phages have longer portal proteins, it is likely that some of these domains are responsible for interacting with the phage nucleus-associated protein ChmB. The region from 484-561, which forms the base part of the portal, might be involved in the interactions. Further studies are needed to characterize these interactions and the octameric oligomerization of ChmB. Overall, our data provides new insights into the interaction between phage portal proteins and the phage nucleus, laying the groundwork

for future studies that will explain how the DNA packaging mechanism changed to accommodate compartmentalization.

While previous work had identified certain proteins found in the phage nucleus, and potentially two phage nucleus systems evolved differences in selecting proteins to be imported, such as GFPmut1[22], we still did not know how the import mechanism works. For the first time, our results show that potentially one of the proteins directly plays a role in selective import to the phage nucleus. By tagging with GFPmut1 and forcing endonucleases into the phiKZ phage nucleus, we found that gp69 of phiKZ was consistently mutated in the escaper mutants, which is the homolog of the miniTurboID hit gp63 of phiPA3 that localizes the phage nucleus. Our results show gp69 mutant phages have defects in import, but phenotype gets rescued by overexpressing the wild-type. We also observed gp63 of phiPA3 is directly interacting with the potentially pore-forming ChmB. AlphaFold also confidently predicted interactions between gp63 and imported protein gp108 or phage nucleus associated gp64 (phiPA3).

When all considered, our current model suggests that the pores formed by ChmB on the phage nucleus might have a multipurpose and be used for protein translocation while gp63 and gp64 of phiPA3 homologs play a role in interacting with the cargo to be imported. While these require further confirmation by in vitro experiments, we believe they can have important implications for understanding the selective protein translocation for the phage nucleus and can lead to the development of novel strategies for phage therapy.

The interaction of phage proteins with the host cell machinery can have a significant impact on the host's biology, and in particular, the cell division process. The

proteins MinD, GroEL, and FtsZ are crucial players in bacterial cell division, and their disruption could cause halt or abnormalities in division. Our results suggest that jumbo phages might be encoding proteins to target these host proteins to stop the host's division to give more time for their replication of longer genome and more elaborate systems to escape host defense systems.

If the interaction partners of gp144, gp259, and gp294 of phiPA3 are indeed MinD, GroEL, and FtsZ; these can explain the bulging and septal reversal that we observe. MinD is known to be localizing on the poles and important for selecting the division site in the mid-cell. FtsZ, on the other hand, is one of the most crucial components of the cell division machinery [32]. There are also studies that report GroEL can localize to mid-cell in an FtsZ-dependent manner at ampicillin treatment and the sequestering of host GroEL can disrupt the host folding machinery [33]. Overall, these studies suggest that phage proteins can indeed interact with the host cell division machinery and potentially disrupt normal cellular processes. While this could be a promising target for phage-based antibacterial therapies, it is also important to consider the potential consequences of disrupting the host cell machinery. Further research is needed to understand the mechanisms by which phage proteins interact with host proteins and to explore the potential consequences of these interactions.

4.5 Experimental Procedures

4.5.1 Strains and Cell Culture

For PhiPA3, PhiKZ and Churi phage, *P. aeruginosa* K2733 (PA01 efflux pump knockout; (*ΔMexAB-OprMΔMexCD-OprJΔMexEF-OprNΔMexXY-OprM*)) was used as the host. Host strains were cultured in Luria-Bertani (LB) media or LB top agar (0.35%

agar) at 30°C (*P. aeruginosa*). To amplify phages, 100 µL overnight liquid culture was mixed with 20 µL of high-titer phage lysate. Mixture was incubated at room temperature for 20 minutes, 5 mL of warm LB top agar was added, mixture was poured onto LB plates and incubated overnight. After overnight, 5 mL of Phage Buffer (10 mM Tris-HCl pH 7.5, 10 mM $MgSO_4$, 68 mM NaCl, and 1 mM $CaCl_2$) was added and incubated at room temperature for 5 hours. Buffer was collected and centrifuged at 15,000 rpm for 10 minutes. Lysates were stored at 4°C with 0.01% chloroform.

For cloning the plasmids to express genes of interests, genes were PCR-amplified with 25 bp homology arms from the phage lysates and ligated to the respective plasmid backbones with NEBuilder HiFi DNA Assembly Cloning Kit (New England Biolabs # E5520S). Plasmids were transformed into *E. coli* DH5α. For selection, LB agar plates were used with appropriate antibiotics (25 µg/mL gentamicin sulfate, 100 µg/mL ampicillin, 100 µg/mL spectinomycin, or 100 µg/mL chloramphenicol). After plasmids were confirmed by Sanger sequencing, plasmids were transformed to host bacteria and selected with LB agar plates containing relevant antibiotics. Selected colonies were overnight in liquid culture with the antibiotics and stored in 25% glycerol at -80°C.

4.5.2 Protein Purification

For protein purification, plasmids of interests were transformed to *E. coli* strain Rosetta 2 (DE3) pLysS (EMD Millipore) and grown overnight in liquid culture with appropriate antibiotics. Next day, cultures were backdiluted in 1L 2XYT media plus antibiotics in 2L shaker flasks and grown at 37°C until they reached an OD_{600} of 0.7. The plasmids were induced with 0.25 mM IPTG and moved to 20°C for 16 hours. Next day, cells were collected by centrifugation and resuspended in a buffer containing 25 mM Tris-

HCl pH 7.5, 10% glycerol, 1 mM NaN_3, 300 mM NaCl, 5 mM imidazole, and 5 mM β-mercaptoethanol. The proteins were purified with Ni^{2+} affinity chromatography (Ni-NTA agarose, Qiagen) and then passed over an anion-exchange column (Hitrap Q HP, Cytiva). Elutions were concentrated and separated with size exclusion column (Superdex 200, Cytiva) in GF buffer (Buffer A and 300 mM NaCl and 1 mM dithiothreitol). Fractions of interests with the selected peak were concentrated (Amicon Ultra, EMD Millipore) to reach a concentration of 10 mg/ml and stored at 4°C.

To analyze the molecular weight of the proteins, 100 μl of purified protein at a concentration of 5 mg/ml was injected to size exclusion chromatography coupled to multi-angle light scattering (SEC-MALS). By using miniDAWN TREOS and Optilab T-rEX detectors (Wyatt Technology), light scattering and refractive index profiles were collected and analyzed by ASTRA software version 8.

4.5.3 CryoEM

For preparing the grids, freshly purified protein was collected from size-exclusion chromatography and diluted to 1 mg/ml. Quantifoil Cu 1.2/1.3 300 grids were glow-discharged for 10 s in a preset program using a Solarus II plasma cleaner (Gatan). Vitrobot Mark IV (Thermo Fisher Scientific) held at 4 °C and 100% humidity and 3.5 μl sample was applied to grid. Grid was immediately blotted with filter paper for 5 seconds and plunged into liquid ethane cooled by liquid nitrogen. Grids were mounted onto AutoGrids (Thermo Fisher Scientific) for imaging.

For chameleon-prepared grids, proteins were adjusted with a final concentration around 2.4 mg/ml. The grids were glow-discharged for 40 seconds, and blotted with one strike mode and of 200 ms setting. Imaging was performed with Titan Krios G3

transmission electron microscope (Thermo Fisher Scientific) with the settings of 300 kV configuration of fringe-free illumination including K2 direct electron detector (Gatan) mounted post Quantum 968 LS imaging filter (Gatan). Microscope was used in EFTEM mode with a slit-width of 20 eV and using a 100 μm objective aperture. Data acquisition was performed as automated by using EPU (Thermo Fisher Scientific) and K2 in counting mode. Ten-second movies were collected at the magnification of 165,000×. The pixel size was selected as 0.84 Å, with a total dose of 61.82 e– Å–2 which was distributed uniformly over 40 frames.

In total, movies were acquired with a realized defocus range of –0.5 to –2.5 μm. For crosslinked samples movies were collected with the same defocus range. Cryo-EM data analysis was performed in cryoSPARC version 3.2. Motion correction was applied with patch motion correction (multi) and contrast transfer function (CTF)-estimated using patch CTF estimation (multi). 2D classifications and heterogeneous refinements used to prepare the final version.

4.5.4 Fluorescent microscopy

For imaging with fluorescent microscopy, 1% agarose pads were prepared on concavity slides. For inducing the protein expression, the pads were supplemented with arabinose or IPTG according to the plasmid. For staining cell membrane, FM4-64 (1 µg/mL) and DAPI (1 µg/mL) dyes were used for staining the cell membrane and DNA, respectively. Strains were resuspended from overnight cultures and spotted on concave slides and incubated in a humidor for 2 hours in 30°C (*P. aeruginosa* strains). For, 10 µl of phages (10^{10} pfu/mL) were added to the pad mixes (resulting in a multiplicity of infection (MOI) of ~7) and incubated in 30°C until desired time point. The slides were sealed with

a coverslip before imaging. DeltaVision Spectris Deconvolution Microscope (Applied Precision) was used for image acquisition by collecting at least 8 Z-axis stacks from the middle focal plane in 0.15 µm increments. DeltaVision SoftWoRx Image Analysis Program and its deconvolution algorithm was used for final image preparation and analyzed by Fiji ImageJ.

4.5.5 GFP Pulldown, SDS-PAGE and Silver Stain

For immunopulldown analysis, overnight cultures of strains expressing GFP-tagged proteins were grown in LB with 25 µg/mL gentamicin sulfate. Cells were back diluted to OD_{600} of 0.1, then grown further to an OD_{600} of 0.5. Cultures were transferred to 250 mL flasks with 50mL total volume of LB supplemented with 0.1% arabinose, 25 µg/mL gentamicin sulfate, and 0.2 mM $CaCl_2$. The strains were induced with Arabinose for 2 hours and collected to be centrifuged at 4000 rpm at 4°C. Resulting pellets were stored at -80°C.

For the GFP-Trap pulldown, cell pellets were incubated in 500 µL lysis buffer (10% glycerol, 25 mM Tris (pH 7.5), 150 mM NaCl, 4 mg/mL lysozyme, 20 µg/mL DNase I, 2x cOmplete Protease Inhibitor, 0.4 mM PMSF) for an hour. Suspension was sonicated with 10 rounds x 20 pulses/round (Duty Cycle 40, Output 4) settings. Lysed cultures were centrifuged at 15,000 rpm for 30 minutes at 4ºC. 25 µL of GFP-Trap bead (Proteintech #gtma-20) (prewashed into dilution buffer: 10 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5 mM EDTA) was used for each sample. 500 µL of supernatant from the centrifuged lysate was added to the beads and rotated end-to-end for 1 hour at 4ºC. Beads were washed 3 times

with wash buffer (10 mM Tris/Cl pH 7.5, 150 mM NaCl, 0.05 % NP-40 substitute, 0.5 mM EDTA).

For staining of PAGE, elutes were mixed with 2x SDS buffer (120 mM Tris-HCl pH 6.8, 20% glycerol, 4% SDS, 0.04% bromophenol blue, 10% β-mercaptoethanol) and boiled at 100ºC for 5 minutes. 10 µL of each sample was run on two separate SDS-PAGE gels and visualized by either silver staining or Coomassie blue staining. For tryptic mass spectrometry of gel bands, bands were cut out of Coomassie blue-stained gels. Remaining samples were used for mass spectrometry identification.

4.5.6. Mass Spectrometry

Samples were trypsin-digested and resulting peptides were analyzed by ultra-high pressure liquid chromatography (UPLC) coupled with tandem mass spectroscopy (LC-MS/MS) using nano-spray ionization. The nanospray ionization was performed using a Orbitrap fusion Lumos hybrid mass spectrometer (Thermo) interfaced with nano-scale reverse-phase UPLC (Thermo Dionex UltiMate 3000 RSLC nano System) using a 25 cm, 75-micron ID glass capillary packed with 1.7-μm C18 (130) BEH beads (Waters corporation). Eluting peptides from the C18 column injected to the mass spectrometer using a linear gradient (5–80%) of ACN (Acetonitrile) at a flow rate of 375 µl/min for 3 hours. ACN gradient was created from the buffers: Buffer A (98% H_2O, 2% ACN, 0.1% formic acid) and Buffer B (100% ACN, 0.1% formic acid). Mass spectrometer parameters were; MS1 survey scan using the orbitrap detector (mass range (m/z): 400-1500 (using quadrupole isolation), 120000 resolution setting, spray voltage of 2200 V, Ion transfer tube temperature of 275°C, AGC target of 400000, and maximum injection time of 50 ms). Data dependent scans were performed at top speed for most intense ions, with charge

state set to only include +2-5 ions, and 5 second exclusion time. Ions were selected with minimal intensities of 50000. Collusion event was carried out in the high energy collision cell (HCD Collision Energy of 30%).Fragment masses where analyzed in the ion trap mass analyzer (With ion trap scan rate of turbo, first mass m/z was 100, AGC Target 5000 and maximum injection time of 35ms). Peaks Studio 8.5 (Bioinformatics solutions Inc.) was used for protein identification and label free quantification.

4.6 References

1. Krylov, V., Bourkaltseva, M., Pleteneva, E., Shaburova, O., Krylov, S., Karaulov, A., Zhavoronok, S., Svitich, O. & Zverev, V. Phage phiKZ—The First of Giants. *Viruses* 13, 149 (2021).
2. Prevelige, P. E., Jr & Cortines, J. R. Phage assembly and the special role of the portal protein. *Curr. Opin. Virol.* 31, 66–73 (2018).
3. Chaikeeratisak, V., Nguyen, K., Egan, M. E., Erb, M. L., Vavilina, A. & Pogliano, J. The Phage Nucleus and Tubulin Spindle Are Conserved among Large Pseudomonas Phages. *Cell Rep.* 20, 1563–1571 (2017).
4. Chaikeeratisak, V., Khanna, K., Nguyen, K. T., Sugie, J., Egan, M. E., Erb, M. L., Vavilina, A., Nonejuie, P., Nieweglowska, E., Pogliano, K., Agard, D. A., Villa, E. & Pogliano, J. Viral Capsid Trafficking along Treadmilling Tubulin Filaments in Bacteria. *Cell* 177, 1771–1780.e12 (2019).
5. Birkholz, E. A., Laughlin, T. G., Armbruster, E., Suslov, S., Lee, J., Wittmann, J., Corbett, K. D., Villa, E. & Pogliano, J. A cytoskeletal vortex drives phage nucleus rotation during jumbo phage replication in E. coli. *Cell Rep.* 40, 111179 (2022).
6. Wente, S. R. & Rout, M. P. The nuclear pore complex and nuclear transport. *Cold Spring Harb. Perspect. Biol.* 2, a000562 (2010).
7. Gould, S. J., Keller, G. A., Schneider, M., Howell, S. H., Garrard, L. J., Goodman, J. M., Distel, B., Tabak, H. & Subramani, S. Peroxisomal protein import is conserved between yeast, plants, insects and mammals. *EMBO J.* 9, 85–90 (1990).
8. Reumann, S., Inoue, K. & Keegstra, K. Evolution of the general protein import pathway of plastids (review). *Mol. Membr. Biol.* 22, 73–86 (2005).
9. Dolezal, P., Likic, V., Tachezy, J. & Lithgow, T. Evolution of the molecular machines for protein import into mitochondria. *Science* 313, 314–318 (2006).
10. Yoneda, Y. Nucleocytoplasmic protein traffic and its significance to cell function. *Genes Cells* 5, 777–787 (2000).
11. Stewart, M. Molecular mechanism of the nuclear protein import cycle. *Nat. Rev. Mol. Cell Biol.* 8, 195–208 (2007).
12. Görlich, D., Prehn, S., Laskey, R. A. & Hartmann, E. Isolation of a protein that is essential for the first step of nuclear protein import. *Cell* 79, 767–778 (1994).

13. Payne, A. S., Kelly, E. J. & Gitlin, J. D. Functional expression of the Wilson disease protein reveals mislocalization and impaired copper-dependent trafficking of the common H1069Q mutation. *Proc. Natl. Acad. Sci. U. S. A.* 95, 10854–10859 (1998).
14. Dobson, C. M. Protein folding and misfolding. *Nature* 426, 884–890 (2003).
15. Boslem, E., MacIntosh, G., Preston, A. M., Bartley, C., Busch, A. K., Fuller, M., Laybutt, D. R., Meikle, P. J. & Biden, T. J. A lipidomic screen of palmitate-treated MIN6 β-cells links sphingolipid metabolites with endoplasmic reticulum (ER) stress and impaired protein trafficking. *Biochem. J* 435, 267–276 (2011).
16. Aasen, T., Johnstone, S., Vidal-Brime, L., Lynn, K. S. & Koval, M. Connexins: Synthesis, Post-Translational Modifications, and Trafficking in Health and Disease. *Int. J. Mol. Sci.* 19, (2018).
17. Cokol, M., Nair, R. & Rost, B. Finding nuclear localization signals. *EMBO Rep.* 1, 411–415 (2000).
18. Kosugi, S., Hasebe, M., Matsumura, N., Takashima, H., Miyamoto-Sato, E., Tomita, M. & Yanagawa, H. Six classes of nuclear localization signals specific to different binding grooves of importin α. *J. Biol. Chem.* 284, 478–485 (2009).
19. Lange, A., Mills, R. E., Lange, C. J., Stewart, M., Devine, S. E. & Corbett, A. H. Classical nuclear localization signals: definition, function, and interaction with importin α. *J. Biol. Chem.* 282, 5101–5105 (2007).
20. Chaikeeratisak, V., Nguyen, K., Khanna, K., Brilot, A. F., Erb, M. L., Coker, J. K. C., Vavilina, A., Newton, G. L., Buschauer, R., Pogliano, K., Villa, E., Agard, D. A. & Pogliano, J. Assembly of a nucleus-like structure during viral replication in bacteria. *Science* 355, 194–197 (2017).
21. Mendoza, S. D., Nieweglowska, E. S., Govindarajan, S., Leon, L. M., Berry, J. D., Tiwari, A., Chaikeeratisak, V., Pogliano, J., Agard, D. A. & Bondy-Denomy, J. A bacteriophage nucleus-like compartment shields DNA from CRISPR nucleases. *Nature* 577, 244–248 (2020).
22. Nguyen, K. T., Sugie, J., Khanna, K., Egan, M. E., Birkholz, E. A., Lee, J., Beierschmitt, C., Villa, E. & Pogliano, J. Selective transport of fluorescent proteins into the phage nucleus. *PLoS One* 16, e0251429 (2021).
23. Weinbauer, M. G. Ecology of prokaryotic viruses. *FEMS Microbiol. Rev.* 28, 127–181 (2004).
24. Salim, O., Skilton, R. J., Lambden, P. R., Fane, B. A. & Clarke, I. N. Behind the chlamydial cloak: the replication cycle of chlamydiaphage Chp2, revealed. *Virology* 377, 440–445 (2008).
25. Chien, A.-C., Hill, N. S. & Levin, P. A. Cell size control in bacteria. *Curr. Biol.* 22, R340–9 (2012).
26. Coltharp, C., Buss, J., Plumer, T. M. & Xiao, J. Defining the rate-limiting processes of bacterial cytokinesis. *Proc. Natl. Acad. Sci. U. S. A.* 113, E1044–53 (2016).
27. Li, Z., Trimble, M. J., Brun, Y. V. & Jensen, G. J. The structure of FtsZ filaments in vivo suggests a force-generating role in cell division. *EMBO J.* 26, 4694–4708 (2007).
28. Howard, M. & Kruse, K. Cellular organization by self-organization: mechanisms and models for Min protein dynamics. *J. Cell Biol.* 168, 533–536 (2005).

29. Macheboeuf, P., Contreras-Martel, C., Job, V., Dideberg, O. & Dessen, A. Penicillin binding proteins: key players in bacterial cell cycle and drug resistance processes. *FEMS Microbiol. Rev.* 30, 673–691 (2006).
30. Lin, D. M., Koskella, B. & Lin, H. C. Phage therapy: An alternative to antibiotics in the age of multi-drug resistance. *World J. Gastrointest. Pharmacol. Ther.* 8, 162–173 (2017).
31. Sun, L., Zhang, X., Gao, S., Rao, P. A., Padilla-Sanchez, V., Chen, Z., Sun, S., Xiang, Y., Subramaniam, S., Rao, V. B. & Rossmann, M. G. Cryo-EM structure of the bacteriophage T4 portal protein assembly at near-atomic resolution. *Nat. Commun.* 6, 7548 (2015).
32. Marston, A. L., Thomaides, H. B., Edwards, D. H., Sharpe, M. E. & Errington, J. Polar localization of the MinD protein of Bacillus subtilis and its role in selection of the mid-cell division site. *Genes Dev.* 12, 3419–3430 (1998).
33. Immaneni, A. Localization and biochemical characterization of the cell division protein FtsA in Escherichia coli. (1998).
34. Richard Evans, Michael O'Neill, Alexander Pritzel, Natasha Antropova, Andrew Senior, Tim Green, Augustin Žídek, Russ Bates, Sam Blackwell, Jason Yim, Olaf Ronneberger, Sebastian Bodenstein, Michal Zielinski, Alex Bridgland, Anna Potapenko, Andrew Cowie, Kathryn Tunyasuvunakool, Rishub Jain, Ellen Clancy, Pushmeet Kohli, John Jumper & Demis Hassabis. Protein complex prediction with AlphaFold-Multimer. bioRxiv 2021.10.04.463034 (2022).

4.7 Acknowledgments

Chapter 4 is preliminary results for 3 projects to be prepared as submission for publication. Portal and ChmB part is currently investigated with collaboration of Basu, D., Gu, Y., Deep, A., Pogliano, J., Corbett, K.D.

Protein import section is currently co-investigated with Morgan, C., and includes collaboration with Armbruster,E., Gu, Y., Deep, A., Pogliano, J. Corbett, K.D.

Phage proteins that interact with host cell division chapter is the preliminary work that will continue in collaboration with Basu, D., Pogliano, J., Corbett, K.D. and Chaikeeratisak lab. The dissertation author is the primary investigator and author of this material, including Figure 4.5 from Morgan, C. work and Figure 4.9c from Chaikeeratisak lab's work.

Chapter 5: Conclusions and Perspectives

5.1 Phage and Phage Nucleus Diversity

Phages are one of the most crucial biological entities for biosphere, ecology and evolution. They exhibit incredible diversity in terms of morphology, genome, and host range. The diversity in the host range of phages creates significant driving force in bacterial evolution and co-evolution with phages in their arms race. Every defense mechanism that bacteria develop results in strong selection pressure on phages to overcome. In the context of this thesis topic, phages have evolved the phage nucleus system to counter the bacterial DNA targeting mechanisms against phages. From our current understanding, there are shared mechanisms required to replicate with the phage nucleus. Specifically, exporting DNA replication and mRNA transcription from the phage nucleus require these molecules to be exported. On the other hand, proteins that will function in the phage nucleus need to be selectively translocated inside while keeping other proteins out. These phage proteins potentially are found in the “core genome” of the phage nucleus, such as ChmC homologs. However, we also see differences in these systems, especially for phages that infect different host species. One example is the PhuZ homologs, where this tubulin forms a bipolar spindle to traffic the phage nucleus and capsid to midcell in phiKZ-like pseudomonas phages while it forms vortex-like assemblies from multiple angles in E. coli infecting Goslar. Although analyses show a strong correlation of encoding ChmA homologs with PhuZ homologs, PhuZ is not a part of the core genome. This is further suggested by the finding that the knockout of the PhuZ gene does not inhibit phage replication. Although it is unclear how the systems are different and can accommodate not having PhuZ, it highlights the importance of understanding the

diversities in the replication systems as much as finding the core proteins for the phage nucleus.

5.2 Importance of the phages and the phage nucleus system

Studying phages since they were discovered has led to significant advances in molecular biology, biotechnology, and medicine. It has allowed us to gain a better understanding of biology, ourselves, and our world. These studies have also enabled the development of many tools that we use in biotechnology today. In recent years, there has been renewed interest in phages in medicine due to their therapeutic applications against the rise of bacterial infections. This has placed the study of the phage nucleus in a crucial position.

Understanding these jumbo phages will allow us to gain insights into the ecology and evolution of microbial communities. Moreover, it will enable us to better control bacterial populations and prevent life-threatening infections. By characterizing the phage nucleus and its components, we can obtain important tools for genetic engineering and create the ultimate natural predators to combat pathogenic bacteria. Therefore, I believe that the content of this thesis will be a valuable source for the upcoming scientific discoveries and technological innovations.