Published in final edited form as:

J Mol Biol. 2013 July 24; 425(14): 2450–2462. doi:10.1016/j.jmb.2013.03.032.

Tail tip proteins related to bacteriophage λ gpL coordinate an iron-sulphur cluster

William Tam¹, Lisa G. Pell², Diane Bona³, Alex Tsai¹, Xiao Xian Dai⁴, Aled M. Edwards², Roger W. Hendrix⁴, Karen L. Maxwell^{2,3}, and Alan R. Davidson^{1,2}

¹Department of Biochemistry, Faculty of Medicine, University of Toronto, Medical Sciences Building, Toronto, ON, Canada, M5S 1A8

²Department of Molecular Genetics, University of Toronto, Medical Sciences Building, Toronto, ON, Canada, M5S 1A8

³Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, ON, Canada, M5S 3E1

⁴Department of Biological Sciences, University of Pittsburgh, Pittsburgh, PA 15260

Abstract

The assembly of long non-contractile phage tails begins with the formation of the tail tip complex. Tail tip complexes are multi-functional protein structures that mediate host cell adsorption and genome injection. The tail tip complex of phage λ is assembled from multiple copies of eight different proteins, including gpL. Purified preparations of gpL and several homologues all displayed a distinct reddish colour, suggesting the binding of iron by these proteins. Further characterization the gpL homologue from phage N15, which was most amenable to *in vitro* analyses, showed that it contains two domains. The C-terminal domain was demonstrated to coordinate an iron-sulphur cluster, providing the first example of a viral structural protein binding to this type of metal group. We characterized the iron-sulphur cluster using inductively coupled plasma-atomic emission spectroscopy, absorbance spectroscopy, and electron paramagnetic resonance spectroscopy and found that it is an oxygen-sensitive $[4\text{Fe-4S}]^{2+}$ cluster. Four highly conserved cysteine residues were shown to be required for coordinating the iron-sulphur cluster, and substitution of any of these Cys residues with Ser or Ala within the context of λ gpL abolished biological activity. These data imply that the intact iron-sulphur cluster is required for function. The presence of four conserved Cys residues in the C-terminal regions of very diverse gpL homologues suggest that utilization of an iron-sulphur cluster is a widespread feature of non-contractile tailed phages that infect Gram-negative bacteria. In addition, this is the first example of a viral structural protein that binds an iron-sulphur cluster.

© 2013 Elsevier Ltd. All rights reserved.

*Correspondence should be addressed to: alan.davidson@utoronto.ca, Phone: 416-978-0332, Fax: 416-978-6885.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Introduction

Bacteriophages with long tails account for greater than 80% of the world's phage population¹. The majority of these phages possess tails that are non-contractile, as is found in the *Siphoviridae* or siphophages, while approximately 40% of these phages possess contractile tails, as is found in the *Myoviridae* or myophages¹. These tails are complicated multi-component structures that assemble in precisely coordinated pathways. The conserved architecture of long tails is comprised of a tail terminator protein, which is located at the 'top' or proximal end of the tail where it can interact with the phage head, a long tubular structure through which DNA passes during infection, and a tail tip complex that is located at the 'bottom' or distal end of the tail. Tail tip complexes (TTCs) perform a variety of essential functions. They form the platform upon which the tail tube protein self-assembles², and are also the attachment point for fibres or receptor binding proteins, which mediate phage adsorption to the surface of the host cell. The TTC also plays key roles in penetration of the host cell membrane(s) and injection of the phage genome to the interior of the cell^{3,4}.

In myophages, the TTC is generally referred to as a baseplate, and this component of the *E. coli* myophage T4 has been intensively studied. Three-dimensional reconstructions of this complex in multiple conformations have been derived using cryo-electron microscopy^{3,5,6} and crystal structures have been solved for seven of its structural proteins^{5,7-12}. These studies have revealed intricate conformational changes within the baseplate that are coupled to the contraction of the tail and movement of the tail fibers. Recent crystallographic and EM studies have also provided detailed views of the TTCs of the siphophages p2, TP901-1, and SPP1, which all infect Gram-positive bacteria¹³⁻¹⁶. Although the structures of the characterized myophage and siphophage TTCs are mostly distinct from each other, a baseplate hub protein (gp27 in phage T4, and Tal in the Gram-positive infecting siphophages) with very similar structure is found in both TTCs, and this structure is also present in the phage-related Type VI secretion system¹⁷. These data show that some structural features are common among all TTCs.

In contrast to the cases described above, no structural information is available for the TTC or the component TTC proteins of any Gram-negative infecting siphophages. Furthermore, there is no detectable sequence similarity between the TTC proteins of Gram-negative and Gram-positive infecting siphophages so that no conserved features can be identified. The assembly pathway of the TTC of *E. coli* phage λ has been well characterized. Eight of the eleven different proteins involved in the assembly of the λ tail are required for the assembly of the TTC², which is also referred to as the initiator complex because it initiates the polymerization of the tail tube protein, gpV. TTC assembly begins with the product of gene *J*, or gpJ, also known as the central tail fiber. Three copies of gpJ are positioned at the tip of the tail¹⁸⁻²¹ where the C-terminal portion of the protein can interact with the lamB receptor on host cells²². During tail morphogenesis, proteins gpI, gpL, and gpK interact with gpJ in a step-wise manner, either directly or indirectly, to form a distinct intermediate structure²³. Approximately six copies of the tape measure protein, gpH,¹⁸ are then thought to anchor their C-termini into this complex²⁴. The remaining portion of gpH remains extended and the tail tube is polymerized around it. The formation of the TTC is completed by the addition of gpM²³, which is followed by tail tube polymerization. At present, little is known about the

regulatory mechanisms involved in controlling the step-wise incorporation of proteins into the TTC. In addition, with the exception of gpJ, which is involved in host cell adsorption, and gpH, which plays a role in both tail length regulation^{26,27} and genome injection^{28–30}, the functional roles of the remaining tail tip proteins are undetermined.

To gain further insight into the structure and function of the TTCs from Gram-negative infecting siphophages, we have investigated gpL of phage λ and several of its homologues. We present evidence that these proteins contain two domains and that the C-terminal domain coordinates an iron sulphur (Fe-S) cluster. We have also identified the residues involved in metal ion coordination and have shown that these residues are required for the biological activity of gpL. Finally, we have discovered very diverse gpL homologues that also likely coordinate Fe-S clusters.

Results

λ gpL and its homologues are reddish in colour and comprise two domains

To characterize gpL, an N-terminally 6-His tagged version was overexpressed from a pET-based plasmid and purified by Ni-NTA affinity chromatography. Surprisingly, the purified protein solution displayed a distinct reddish colour. To evaluate the significance of this unusual colouration, we also purified gpL homologues from phages N15, HK022, and ϕ 80, and from the F-type pyocin of *Pseudomonas aeruginosa* PAO1 (Table 1). We are confident that these proteins are gpL homologues because they were identified with high significance after only one iteration of PSI-BLAST³² using λ gpL as a query (E-value < $1e^{-25}$). They also display at least 25% sequence identity to gpL, and are all members of one sequence family defined in the Pfam database (PF05100: Phage_tail_L). The genes encoding these proteins all lie in the same genomic position as the *L* gene (Figure 1A). An alignment of these homologues with gpL and other diverse phage-encoded gpL homologues highlights their common features (Figure 1B). Each of the purified gpL homologues listed above yielded protein solutions with a red colour, implying that this is a conserved property of these proteins. The gpL homologue from phage N15 (gpL_{N15}) was the most soluble, and was purified with the greatest yield; thus, preparations of this protein had the most intense reddish colour (Figure 2A). Further *in vitro* characterization was performed on this protein due to its optimal behaviour.

To assess the possibility that gpL might possess multiple domains and to determine which part of the protein was responsible for imparting the red colour, we subjected purified gpL_{N15} to partial proteolysis by trypsin and monitored the progress of the reaction by SDS-PAGE. A stable fragment with an apparent molecular weight of approximately 20 kDa (Figure 2B) was visualized on the gel after 120 minutes of digestion, and this band persisted even after 24 hours of digestion (data not shown). MALDI-TOF mass spectrometry was used to determine the molecular weight of the persisting fragment. The identified weight of 18,983 Da corresponds to residues 1-173 of gpL_{N15}, which has a predicted weight of 18,982 Da, assuming post-translational removal of the initial Met residue (Figure 2C). We detected no proteolytically stable fragments encompassing any portion of the C-terminal region of gpL_{N15}. Therefore, we concluded that the C-terminal region of gpL is at least partially

unstructured under these conditions and thus was rapidly proteolyzed into fragments too small to be detected by SDS-PAGE.

To further define the domain architecture of gpL_{N15}, 6-His tagged versions of the domains defined by the proteolysis experiment, N-gpL_{N15} (residues 1-173) and C-gpL_{N15} (residues 174-232), were overexpressed and purified. Purified preparations of C-gpL_{N15} were reddish in colour, while purified N-gpL_{N15} was colourless. Far-UV circular dichroism (CD) spectra acquired for preparations of full-length gpL_{N15} and N-gpL_{N15} exhibited an ellipticity minimum at 216nm (Figure 2D), which is characteristic of β -strand containing proteins. By contrast, the spectrum for C-gpL_{N15} displayed an ellipticity minimum at 203nm (Figure 2D), suggesting the presence of some random coil structure. We conclude that N-gpL_{N15} folds into a stable domain containing β -strands while C-gpL_{N15}, which may be partially unfolded, is responsible for the red colour of the gpL homologues.

gpL_{N15} coordinates a [4Fe-4S]²⁺ iron-sulphur cluster

Since reddish colour in a protein sample can result from the presence of bound iron, we used spectroscopic methods to investigate the metal binding properties of gpL_{N15}. The UV/visible absorbance spectrum of gpL_{N15} revealed two broad peaks between 310–370 and 400–440 nm in addition to the large peak at 280 nm (Figure 3A). Consistent with its reddish color, a similar spectrum was seen for C-gpL_{N15} while N-gpL_{N15} displayed a typical protein spectrum with a peak only at 280 nm. Since spectra similar to those seen for gpL_{N15} and C-gpL_{N15} are often observed for proteins that coordinate iron-sulphur (Fe-S) clusters^{33–35}, we determined the concentration of iron and other metals in gpL_{N15} preparations using inductively coupled plasma-atomic emission spectroscopy (ICP-AES). Although gpL_{N15} was reddish in colour and exhibited unusual peaks in its absorbance spectrum (Figure 3A), purified preparations of this protein contained only 0.07 ± 0.03 mol Fe/mol gpL_{N15}. Since Fe-S clusters are sensitive to oxygen³⁶, we surmised that these low values might be a result of oxygen exposure. Therefore, we purified gpL_{N15} aerobically and incubated it with Fe(NH₄)₂(SO₄)₂ and Na₂S under anaerobic conditions to reconstitute the putative Fe-S cluster. The reconstituted protein was then purified by Ni-NTA chromatography under anaerobic conditions to remove unbound iron and sulphur. These gpL_{N15} preparations were a deeper reddish-brown colour than the non-reconstituted protein and gave an enhanced absorbance in the visible region of the UV spectrum (Figure 3B). ICP-AES analysis on the reconstituted protein sample revealed the presence of 3.7 ± 0.5 mol Fe/mol gpL_{N15}, suggesting the presence of a [4Fe-4S] cluster.

We used electron paramagnetic resonance (EPR) to investigate the type of Fe-S cluster coordinated by gpL_{N15} and determine its oxidation state. No EPR signals were detected from the anaerobically reconstituted gpL_{N15}, suggesting that there were no unpaired electrons in the Fe-S cluster. To induce unpaired electrons ($s = 1/2$), gpL_{N15} was reduced with 10 mM dithionite. However, even after this treatment, only a weak EPR signal was detected. By contrast, oxidizing gpL_{N15} by exposing the protein to air for six hours resulted in a detectable EPR signal with a g-value³⁷ of 2.014 (Figure 3C). The EPR spectrum and g-value of oxidized gpL_{N15} was consistent with that of proteins coordinating a [3Fe-4S]¹⁺ cluster, similar to the well-studied Fe-S cluster protein aconitase, which coordinates a

[3Fe-4S] cluster in its inactive state and a [4Fe-4S] cluster when catalytically active³⁸. The [3Fe-4S]¹⁺ form of aconitase displays a g-value of 2.01³⁹. The lack of EPR signals from reconstituted gpL_{N15} was likely due to the lack of unpaired electrons in its [4Fe-4S]²⁺ cluster, while the strong EPR signals detected in oxidized reconstituted gpL_{N15} were likely a result of oxidative damage to the Fe-S cluster resulting in a [3Fe-4S]¹⁺ cluster. Together, the UV-visible absorption scans, ICP-AES analysis, and EPR data, indicate that gpL_{N15} most likely coordinates a [4Fe-4S]²⁺ cluster.

The oxygen sensitivity of the gpL_{N15} Fe-S cluster was assessed under both aerobic and anaerobic conditions. After storage of the reconstituted gpL_{N15} in air for 24 hrs, the protein lost its colour and characteristic peaks in a UV/visible absorbance scan (Figure 4A). By comparison, gpL_{N15} maintained its red colour and there was only a slight change in its absorbance properties (Figure 4A) after 24 hours of anaerobic storage. A timecourse of the loss of absorbance at 415 nm (Figure 4B) revealed that most of the colour was lost within the first 105 minutes, emphasizing the instability of the cluster under aerobic conditions.

Conserved cysteine residues in C-gpL_{N15} are responsible for coordinating the Fe-S cluster

An alignment of diverse gpL homologues (Figure 1B) revealed four completely conserved Cys residues in the C-terminal region of these proteins. Since Cys residues commonly coordinate Fe-S clusters⁴⁰, we surmised that these four Cys residues might coordinate the Fe-S cluster in gpL and its homologues. The location of these Cys residues in the C-terminal region of the gpL homologues is consistent with our observation that purified C-gpL_{N15} protein solution was red in colour while N-gpL_{N15} was not. To investigate the contribution of the conserved Cys residues to the red colouration of gpL_{N15}, each conserved cysteine was individually substituted with Ala and the mutant proteins were overexpressed and purified aerobically. Each protein possessed the same level of solubility as wild-type gpL_{N15}. The three mutant proteins that were investigated by CD spectroscopy (C184A, C221A, C228A) displayed CD spectra that were indistinguishable from the wild-type spectrum (Figure 5A). Strikingly, purified preparations of each of the cysteine mutants were colourless. In addition, the UV/visible absorbance scans of these proteins showed none of the unusual peaks in the visible region that were seen for the wild-type protein (Figure 5B). These data imply that each of the conserved Cys residues in gpL and its homologues plays a key role in coordinating the Fe-S cluster.

To confirm that the gpL_{N15} Fe-S cluster is coordinated by four Cys residues, we measured the number of reactive Cys residues in purified gpL_{N15} by mixing it with Ellman's reagent (5,5'-Dithio-bis(2-nitrobenzoic acid) or DTNB). DTNB reacts in a 1-to-1 ratio with reduced thiol groups to produce a yellow colour that absorbs maximally at 412 nm. Measurement of the absorbance at 412 nm after reaction with DTNB provides an accurate quantitation of the number of free Cys residues within a protein^{41,42}. Anaerobically reconstituted gpL_{N15} was found to contain 2.9 ± 0.5 moles of free thiol groups per mole of protein. Since there are a total of seven Cys residues in gpL_{N15}, this result supports our conclusion that four Cys residues are involved in coordinating the Fe-S cluster and thus are blocked from reacting with DTNB.

Conserved Cys residues in the C-terminal domain of gpL are necessary for tail assembly

λ gpL has four Cys residues, located at positions 173, 182, 205 and 212, that are highly conserved amongst related phages. The experiments described above establish that these four conserved cysteines play an important role in the formation of a 4Fe-4S cluster in gpL_{N15}. λ gpL also has four additional Cys residues, located at positions 12, 39, 140, and 208, that are not conserved amongst its homologues. To investigate whether the Cys residues implicated in Fe-S cluster binding are required for the biological activity of λ gpL, an *in vivo* complementation assay was used. Each of the eight Cys residues in gpL was individually substituted with Ser and the ability of each mutant to complement a λ phage mutant bearing an amber nonsense mutation in the *L* gene (*Lam*) was assessed. As seen in Figure 6A, an *Lam* phage lysate produced a titre of 10^9 plaque forming units (pfu) per ml when plated on cells expressing wild-type gpL from a plasmid. Three of the four non-conserved Cys residues, when substituted with Ser, complemented the *Lam* phage with the same efficiency as wild type lambda, while Cys140Ser showed some deficiency in complementation. However, when substituted with Met, an amino acid found at this position in some gpL homologues, the mutant Cys140Met complemented as well as wild type. Taken together, these results imply that none of the non-conserved cysteines play an essential role in producing functional lambda tails. In contrast to the non-conserved Cys residues, a reduction in titre of greater than 10^6 -fold was observed when the *Lam* lysate was plated on cells expressing any of the four highly conserved Cys to Ser substitutions (Figure 6A). Mutation of these positions to Ala also abrogated the biological function of gpL (data not shown). These results imply that each of the four conserved cysteines play an essential role in producing functional tails.

To determine if the Cys to Ser substitutions that failed to complement the *Lam* phage affect tail tube assembly, these substitutions were introduced into a plasmid (pETail) that expresses all of the λ tail genes under the control of a T7 promoter²⁵. Protein expression was induced from the plasmid in the presence of ³⁵S-methionine. The resulting tail preparations were loaded on a 10–30% glycerol velocity sedimentation gradient. Following sedimentation, the gradients were fractionated and the fractions where wild type λ tails characteristically run were analyzed by SDS-PAGE. The number of tails produced by each mutant was quantified by visualization of the major tail subunit, gpV (Figure 6B). The number of tails produced by each of the mutants of the non-conserved cysteines was approximately the same high level produced by the wild type control. By contrast, no tails were detected for the conserved cysteine mutants, Cys173Ser, Cys182Ser and Cys205Ser, and the Cys212Ser mutant supported production of only 4.2 ± 1.0 % of tails compared to wild type. The individual substitution of each of these four positions with Ala also resulted in a defect in tail assembly, with no tails produced (data not shown). These results imply that the Fe-S coordinating activity of gpL is required for formation of an initiator complex that can promote tail tube formation. The requirement of gpL for formation of the tail tube initiator has been previously demonstrated³¹.

Some distantly related gpL homologues appear to coordinate an Fe-S cluster while others do not

To search for additional phage structural proteins related to gpL that might coordinate FeS clusters, we performed iterative PSI-BLAST searches using multiple query proteins (see Materials and Methods for details). This process led to the identification of a family of proteins encoded in the tail gene regions of some *Pseudomonas*, *Burkholderia* and alpha-proteobacteria phages that all possess four completely conserved Cys positions in their C-terminal regions (Supplementary Table 1). The N-terminal regions of these proteins are members of the Pfam family DUF2163 (PF09931) and the C-terminal regions correspond to a family named Phage_BR0599 (PF09356). Although gpL and its family members display no detectable pairwise sequence identity with this other family of proteins, employment of the profile-profile comparison program, HHpred shows that a search with gpL reveals high confidence matches (> 95% probability) to both of the above-mentioned families containing the *Pseudomonas* proteins, indicating that these proteins are likely homologous to the N- and C-terminal domains of gpL (Supplementary Table 2). The putative Fe-S cluster-coordinating domains of the *Pseudomonas* phage-related proteins are distinguished from the λ -related family by the insertion of additional protein sequence between the two pairs of Cys residues putatively involved in Fe-S cluster coordination (Figure 6). Nevertheless, distinct sequence similarities are seen at the C-termini of all these proteins.

Our PSI-BLAST searches also revealed a family of shorter phage-encoded gpL homologues that possess an N-terminal domain similar to gpL, but possess no C-terminal region that could bind an Fe-S cluster. Most of these proteins are found within Pfam family DUF1833 (PF08875), and are components of phages infecting various bacterial genera including *Pseudomonas*, *Burkholderia*, *Xanthomonas* and *Salmonella* (Supplementary Tables 1). We also detected sequence similarity between the N-terminal domain of gpL and other phage proteins that are not members of Pfam families. The largest group of these proteins include the Enterobacteria phage T5 pb3 protein and closely related homologues. Only approximately 150 residues at N-terminus of these large (>900 residue) putative tail proteins are related to gpL. Although the sequence similarities among these diverse gpL homologues are low, comparisons with HHpred show high probability matches of these sequences to HMMs corresponding to the four Pfam families containing putative gpL homologues (Supplementary Table 2). A common evolutionary origin for these gpL homologues is supported by their common genomic position with most (33/57) being position two ORFs downstream of the gene encoding the Tape Measure Protein, and the rest being within 5 ORFs of this gene (Supplementary Table 1).

It is notable that our PSI-BLAST searches detected putative homologues of gpL with or without Fe-S coordinating domains encoded in 75% of genomes of siphophages infecting Gram-negative bacteria, yet we found no putative gpL homologues in genomes of siphophages infecting Gram-positive bacteria. We conclude that possession of gpL-related proteins is a unique feature of Gram-negative infecting siphophages.

Discussion

The data presented here provide the first example of a viral structural protein that coordinates an iron-sulphur cluster (Fe-S cluster). As such, the role of this Fe-S cluster in the morphogenesis of siphophages likely represents a new category of Fe-S cluster function. Previously, Fe-S clusters have been shown to be involved in a number of diverse functions including electron transfer^{45,46}, substrate binding and catalysis^{36,38,47}, sensing and regulating protein expression levels^{48,49}, and maintaining the structural integrity of proteins^{50,51}. At this point, we can only speculate about the role of the gpL Fe-S cluster. This Fe-S cluster likely does not perform an enzymatic function because our attempts to reduce it to a $[4\text{Fe-4S}]^{1+}$ form were unsuccessful, implying that the cluster has a low redox potential. Thus, we hypothesize that the gpL Fe-S cluster serves to stabilize gpL in its assembled form within the phage particle and/or plays a role in the conformational changes required for gpL to carry out its function in assembly or DNA injection. Supporting a role for the Fe-S cluster in stabilizing gpL_{N15}, the C-terminal Fe-S cluster-binding region of aerobically purified gpL_{N15}, which displays low Fe-S cluster occupancy, appears to be unstructured as assessed by CD and proteolysis studies. We expect that this region becomes more structured when gpL_{N15} is bound to the Fe-S cluster. A comparable structural role for Fe-S clusters is exemplified by the DNA-repair enzymes, endonuclease III and MutY. The Fe-S clusters in these proteins control the folding of flexible loops that are important for enzyme function⁵⁰⁻⁵³. Although the Fe-S cluster appears to be quite unstable in air in uncomplexed gpL, this cluster may become much more stable when gpL is incorporated into a phage tail. This stabilization may occur through intermolecular and intramolecular interactions of gpL and/or by burial of the Fe-S clusters in the interior of the tail tip.

Although the role of the Fe-S cluster coordinated by gpL is unknown, it is likely crucial for function. Substitutions of any of the four highly conserved C-terminal domain Cys residues of gpL with Ser or Ala caused a complete loss of biological activity. Substitutions of the homologous residues in gpL_{N15} led to a marked loss of Fe-S cluster coordination as was detected by a loss of reddish colouration in the purified mutant proteins. The essential role of Fe-S cluster coordination is also supported by the complete conservation of the four Cys containing positions in alignments of the C-terminal domains of diverse gpL homologues from both the λ and *Pseudomonas* phage (DUF2163) families. Our finding that four different purified gpL homologues display reddish colouration supports the conclusion that binding of Fe-S clusters is a universal and essential property of this group of proteins. However, our identification of likely gpL homologues that possess no C-terminal region that could bind an Fe-S cluster indicates that Fe-S cluster formation is not an absolutely required feature of the tails of siphophages infecting Gram-negative bacteria. We also conclude that the function fulfilled by the N-terminal domain of gpL is separable from that of the C-terminal domain.

While this study presents the first known example of a phage morphogenetic protein coordinating an Fe-S cluster, there is growing evidence that metals, including iron, may play a role in the function of a variety of phage tail proteins. For example, the membrane-piercing proteins from myophages P2 and phi92 are stabilized by iron ions coordinated by six histidine residues positioned at their tips⁶². Similarly, His residues within the long tail fibre

of myophage T4 coordinate seven iron ions and a large conformational transition of the baseplate of the *Lactococcus* siphophage p2 is controlled by binding of Ca^{2+} ions¹³. These examples, coupled with our discovery of the Fe-S cluster coordinated by gpL, emphasize that determination of the roles of metals in the functioning of phage tails will be an important area of future investigation.

Materials and Methods

Plasmid construction, protein purification, and protein quantification

The *L* genes from phages λ , N15, HK022, and ϕ 80 were PCR amplified from phage lysates and the F-pyocin gene PA0623 was amplified from a *Pseudomonas aeruginosa* PA01 bacterial culture. The genes were cloned into a pET15b expression vector (*Novagen*), producing N-terminally 6-His tagged proteins. Gene L_{λ} was subsequently sub-cloned into the pAD100 vector⁶³ without the addition of an affinity tag and gene L_{N15} was sub-cloned into a pET21d vector (*Novagen*) producing a C-terminal 6-His tagged protein. Regions of DNA encoding residues 1-173 (N- L_{N15}) and 174-251 (C- L_{N15}) were cloned into a pET15b and pET21d expression vectors respectively, using the In-Fusion method (*Clontech Laboratories Inc.*). Quickchange mutagenesis (*Stratagene*) was used to create all mutations.

Protein expression was induced in BL21 (DE3) tail cells with 1 mM isopropyl-beta-D-thiogalactopyranoside (IPTG) followed by 3 hours of incubation with shaking at 37 °C. Cells were harvested by centrifugation, flash frozen in liquid nitrogen, and stored at -70 °C. Bacterial cell pellets were thawed on ice and resuspended in binding buffer (50 mM Tris pH 7.5, 500 mM NaCl, 5 mM Imidazole, 2 mM DTT) supplemented with EDTA-free complete protease inhibitor cocktail (*Roche*). Egg white lysozyme was added to a final concentration of 1 mg/mL and the mixture was incubated on ice for 30 minutes with occasional mixing. The cells were lysed by sonication and incubated with 10 $\mu\text{g}/\text{ml}$ of DNase (*Sigma-Aldrich*) at room temperature for 15 minutes with gentle rocking. The cell debris was collected at 15000 rpm, 4 °C for 30 minutes and the resulting supernatant was incubated with 1 mL Ni-NTA agarose (*Qiagen*) for 30 minutes at 4 °C with gentle rocking. The mixture was transferred to a column and the resin was washed with approximately 100 mL of 50 mM Tris pH 7.5, 500 mM NaCl, 30 mM Imidazole, 2 mM DTT. Protein was eluted with 50 mM Tris pH 7.5, 500 mM NaCl, 250 mM Imidazole, 2 mM DTT and dialyzed for 4 hours in 10 mM Tris pH 8.0, 100 mM NaCl, 2 mM DTT. The sample was transferred to fresh dialysis buffer and equilibrated an additional 16 hours at 4 °C. The purified protein was filtered through a 0.22 micron filter (*Millipore*) and stored at 4 °C. Protein concentration was determined using the Bradford Assay.

Domain mapping by partial trypsin digest

Trypsin was mixed with purified gp L_{N15} in a molar ratio of 1:500 protease:protein and was incubated at room temperature with occasional mixing. Aliquots were removed at varying time points and immediately mixed with PMSF to a final concentration of 0.2 mM to inhibit the proteolysis reaction. A portion of each aliquot was then mixed with 2X SDS sample buffer in a 1:1 ratio and boiled for five minutes. To separate and visualize the proteolyzed fragments, the samples were run on a SDS polyacrylamide gel. Samples for which stable

fragments were visualized on the gel were sent for MALDI-TOF mass spectrometry analysis at the Advanced Protein Technology Center at The Hospital for Sick Children (Toronto). The MALDI-TOF results were mapped onto the protein sequence to identify the degraded fragments.

Circular dichroism and absorbance spectroscopy

Circular dichroism (CD) experiments were performed at 25 °C in an Aviv Circular Dichroism Model 202 spectrometer. Experiments were collected at wavelengths ranging from 200 to 260 nm in 1 nm increments with a five second averaging time. Three independent scans were collected and averaged for analysis. Data were collected in a quartz cuvette with a 0.1 cm path length. Unless otherwise noted, far-UV scans were measured at a protein concentration of 25 µM. UV absorbance spectra on non-reconstituted proteins were collected at room temperature on a Varian Cary 300 Bio UV-visible spectrophotometer at wavelengths ranging from 200 to 700 nm in a 1 cm quartz cuvette. UV absorbance spectra acquired on the reconstituted protein and for the DTNB assay were collected from 200 to 1100 nm using an Agilent 8453 UV-visible spectrophotometer in a 1 cm quartz cuvette.

Anaerobic reconstitution of gpL_{N15}

Purified gpL_{N15} was dialyzed into 50 mM Tris-HCl pH 7.5, 250 mM NaCl (buffer A) and then concentrated to approximately 1 mM before being transferred into an anaerobic chamber, which was maintained in an 95% N₂ and 5% H₂ atmosphere with less than 2 ppm of O₂ at 25 °C. gpL_{N15} was left to equilibrate anaerobically for at least 1.5 hours followed by a 1.5 hour incubation with 10 mM DTT. Freshly prepared Fe(NH₄)₂(SO₄)₂ in buffer A was added to each protein sample to a final concentration of 10 mM and incubated for 2 minutes before the addition of 10 mM Na₂S (also prepared in buffer A). The reconstitution mixture was incubated for 1 hour before FeS precipitate was removed using centrifugation and excess Fe and S was removed by applying the entire mixture to a PD-10 size exclusion column. The reconstituted protein was further purified with Ni-NTA agarose, using the same protocol described for the aerobic protein purification with one exception; DTT was not present in any of the buffers during anaerobic purification. Excess salt and imidazole were removed from the eluted protein using a PD-10 column.

DTNB assay

Reconstituted protein samples were diluted in buffer A, mixed with DTNB assay solution (7.8 M guanidine hydrochloride, 1.3 mM EDTA, and 0.35 mM 5, 5' dithio-bis(2-nitrobenzoic acid) that was freshly prepared in 1M Tris pH 8.8), and incubated anaerobically for 10 minutes before measuring the UV absorbance at 412 nm. A standard curve was created using a set of 2-mercaptoethanol dilutions made in buffer A and the number of occupied thiol groups was calculated for each protein sample.

Inductively coupled plasma-atomic emissions spectroscopy

The concentration of iron in the reconstituted and non-reconstituted protein samples was determined by ICP-AES. Samples were prepared in buffer A and analyzed using an Optima

7300 DV; ICP-AES data was collected by the ANALEST service facility at the University of Toronto.

Electron paramagnetic resonance spectroscopy

Approximately 12 μ M of reconstituted protein was either reduced with 10 mM dithionite, exposed to air for 6 hours, or was left untreated. Samples were stored in microfuge tubes, frozen in liquid nitrogen, and were sent frozen to the National Biomedical Electron Paramagnetic Resonance Center at the Medical College of Wisconsin. Upon arrival, samples were thawed and transferred to a 4 mm quartz EPR tube and frozen once again in liquid nitrogen. EPR spectra were obtained using a Bruker E500 ELEXSYS spectrometer with an Oxford Instruments ESR-9 helium flow cryostat and a Bruker DM0101 cavity. The parameters used to collect EPR spectra were as follows: liquid helium temperature (10 K), microwave frequency 9.633 GHz, and modulation amplitude from 5–10 G.

In vivo complementation assays

Plasmids carrying either wild-type or cysteine mutant gene *L* were freshly transformed into BL21(DE3)- tail cells. A single colony was picked from each mutant and grown in LB-Amp at 37 °C to an OD₆₀₀ of 0.8. Cells were suspended in 0.7% molten agar supplemented with IPTG to a final concentration of 0.1 mM, then poured onto LB/Agar ampicillin plates containing 10 mM MgSO₄. Serial dilutions of a λ Lam phage lysate were spotted in 4 μ L aliquots onto the bacterial lawn and incubated overnight at 37 °C. Zones of clearance were scored the following day to determine the complementing ability of the expressed protein.

Preparation and analysis of ³⁵S-labeled tail extracts

BL21(DE3)- cells transformed by pETail plasmids (REF Gibbs & Theriault) carrying wild type or mutant gpL genes were grown at 37 °C in M9-MC medium (1.33% Na₂HPO₄•7H₂O, 0.3% KH₂PO₄, 0.05% NaCl, 0.1% NH₄Cl, 0.4% (w/v) glucose, 1 mM MgSO₄, 0.1 mM CaCl₂ and 1% Difco MC medium) to a cell density of 3×10⁸ cells/mL. The temperature was shifted to 30 °C and 1 mM IPTG was added to induce protein expression for 30 minutes, followed by the addition of 200 mg/mL rifampicin to shut down host gene expression. After 20 minutes, ³⁵S-Met (5–10 mCi/mL) was added and the culture was grown for 1 hour. The cells were harvested by centrifugation for 10 minutes at 8,000 r.p.m and the cell pellet was resuspended in 1/10 culture volume of cold lysis buffer (50 mM Tris-HCl, pH 8.0, 5 mM EDTA, 0.1% Triton X-100) and the mixture was warmed to 24 °C for five minutes. Subsequently, 7.5 mM MgSO₄ and 20 mg/mL DNase I was added and the mixture was incubated an additional five minutes. The cell debris was removed by centrifugation at 8,000 r.p.m. for 15 minutes at 4 °C. The resulting cleared tail lysate was loaded onto 10–30% glycerol gradient in TKG buffer (20 mM Tris-HCl pH 7.5, 100 mM potassium glutamate) and centrifuged in an SW41 rotor at 40,000 r.p.m. for 3.5 hours at 4°. After centrifugation fractions were collected and analyzed by SDS-PAGE, followed by autoradiography.

Identification of gpL homologues

A PSI-BLAST³² search against the non-redundant database using the NCBI web server was carried out for 4 iterations with λ gpL as the initial query. Several diverse prophage proteins

that were hit in this initial search were used as queries in subsequent PSI-BLAST searches. The identity of these prophage proteins as gpL homologues was confirmed by determining that their genomic position within the prophage was as expected (i.e. the genes encoding these proteins lay in tail-encoding regions and were generally 2 ORFs away from the gene encoding the TMP). A search with ZP_01219717.1 (hypothetical protein P3TCK_17682 from *Photobacterium profundum* 3TCK) detected putative gpL homologues in the DUF2163 family. A search with ZP_01899409.1 (hypothetical protein PE36_00310 from *Moritella* sp. PE36) detected the gpL homologue of phage PY54 and YP_002002053.1 (putative phage associated protein from *Neisseria gonorrhoeae* NCCP11945). A search with YP_002002053.1 detected the gpL homologues in phage T5 and related phages, and the homologues in Pfam family DUF2163/BR0599. Searches with HHpred (<http://toolkit.tuebingen.mpg.de/hhpred>) were used to confirm the similarities of proteins to gpL. Our set of phage genomes considered here includes only those in the RefSeq complete genome set found at NCBI. Our database included genomes present as of August, 2012.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank Paul Sadowski for critical reading of the manuscript. We also thank Deborah Zamble for advice and for aid in anaerobic protein purification. This work was supported by Operating Grants from the Canadian Institutes for Health Research to K.L.M. and A.M.E. (FRN 62796) and A.R.D. (FRN 77680) and the National Institutes of Health to R.W.H. (RO1-GM47795).

References

1. Ackermann HW. 5500 Phages examined in the electron microscope. *Arch Virol.* 2007; 152:227–43. [PubMed: 17051420]
2. Katsura, I. Tail Assembly and Injection. *Lambda II*; Cold Spring Harbor: 1983.
3. Leiman PG, Chipman PR, Kostyuchenko VA, Mesyanzhinov VV, Rossmann MG. Three-dimensional rearrangement of proteins in the tail of bacteriophage T4 on infection of its host. *Cell.* 2004; 118:419–29. [PubMed: 15315755]
4. Plisson C, White HE, Auzat I, Zafarani A, Sao-Jose C, Lhuillier S, Tavares P, Orlova EV. Structure of bacteriophage SPP1 tail reveals trigger for DNA ejection. *EMBO J.* 2007; 26:3720–8. [PubMed: 17611601]
5. Kanamaru S, Leiman PG, Kostyuchenko VA, Chipman PR, Mesyanzhinov VV, Arisaka F, Rossmann MG. Structure of the cell-puncturing device of bacteriophage T4. *Nature.* 2002; 415:553–7. [PubMed: 11823865]
6. Kostyuchenko VA, Leiman PG, Chipman PR, Kanamaru S, van Raaij MJ, Arisaka F, Mesyanzhinov VV, Rossmann MG. Three-dimensional structure of bacteriophage T4 baseplate. *Nat Struct Biol.* 2003; 10:688–93. [PubMed: 12923574]
7. Leiman PG, Shneider MM, Kostyuchenko VA, Chipman PR, Mesyanzhinov VV, Rossmann MG. Structure and location of gene product 8 in the bacteriophage T4 baseplate. *J Mol Biol.* 2003; 328:821–33. [PubMed: 12729757]
8. Leiman PG, Shneider MM, Mesyanzhinov VV, Rossmann MG. Evolution of bacteriophage tails: Structure of T4 gene product 10. *J Mol Biol.* 2006; 358:912–21. [PubMed: 16554069]
9. Leiman PG, Kostyuchenko VA, Shneider MM, Kurochkina LP, Mesyanzhinov VV, Rossmann MG. Structure of bacteriophage T4 gene product 11, the interface between the baseplate and short tail fibers. *J Mol Biol.* 2000; 301:975–85. [PubMed: 10966799]

10. Kostyuchenko VA, Navruzbekov GA, Kurochkina LP, Strelkov SV, Mesyanzhinov VV, Rossmann MG. The structure of bacteriophage T4 gene product 9: the trigger for tail contraction. *Structure*. 1999; 7:1213–22. [PubMed: 10545330]
11. Thomassen E, Gielen G, Schutz M, Schoehn G, Abrahams JP, Miller S, van Raaij MJ. The structure of the receptor-binding domain of the bacteriophage T4 short tail fibre reveals a knitted trimeric metal-binding fold. *J Mol Biol*. 2003; 331:361–73. [PubMed: 12888344]
12. Leiman PG, Arisaka F, van Raaij MJ, Kostyuchenko VA, Aksyuk AA, Kanamaru S, Rossmann MG. Morphogenesis of the T4 tail and tail fibers. *Virology*. 2010; 7:355. [PubMed: 21129200]
13. Sciara G, Bebeacua C, Bron P, Tremblay D, Ortiz-Lombardia M, Lichiere J, van Heel M, Campanacci V, Moineau S, Cambillau C. Structure of lactococcal phage p2 baseplate and its mechanism of activation. *Proc Natl Acad Sci U S A*. 2010; 107:6852–7. [PubMed: 20351260]
14. Veessler D, Spinelli S, Mahony J, Lichiere J, Blangy S, Bricogne G, Legrand P, Ortiz-Lombardia M, Campanacci V, van Sinderen D, Cambillau C. Structure of the phage TP901-1 1.8 MDa baseplate suggests an alternative host adhesion mechanism. *Proc Natl Acad Sci U S A*. 2012; 109:8954–8. [PubMed: 22611190]
15. Veessler D, Robin G, Lichiere J, Auzat I, Tavares P, Bron P, Campanacci V, Cambillau C. Crystal structure of bacteriophage SPP1 distal tail protein (gp19.1): a baseplate hub paradigm in gram-positive infecting phages. *J Biol Chem*. 2010; 285:36666–73. [PubMed: 20843802]
16. Goulet A, Lai-Kee-Him J, Veessler D, Auzat I, Robin G, Shepherd DA, Ashcroft AE, Richard E, Lichiere J, Tavares P, Cambillau C, Bron P. The opening of the SPP1 bacteriophage tail, a prevalent mechanism in Gram-positive-infecting siphophages. *J Biol Chem*. 2011; 286:25397–405. [PubMed: 21622577]
17. Lossi NS, Dajani R, Freemont P, Filloux A. Structure-function analysis of HsiF, a gp25-like component of the type VI secretion system, in *Pseudomonas aeruginosa*. *Microbiology*. 2011; 157:3292–305. [PubMed: 21873404]
18. Casjens SR, Hendrix RW. Locations and amounts of major structural proteins in bacteriophage lambda. *J Mol Biol*. 1974; 88:535–45. [PubMed: 4476800]
19. Buchwald M, Siminovitch L. Production of serum-blocking material by mutants of the left arm of the lambda chromosome. *Virology*. 1969; 38:1–7. [PubMed: 4891222]
20. Dove WF. Action of the lambda chromosome. I. Control of functions late in bacteriophage development. *J Mol Biol*. 1966; 19:187–201. [PubMed: 5967283]
21. Mount DW, Harris AW, Fuerst CR, Siminovitch L. Mutations in bacteriophage lambda affecting particle morphogenesis. *Virology*. 1968; 35:134–49. [PubMed: 5652675]
22. Wang J, Hofnung M, Charbit A. The C-terminal portion of the tail fiber protein of bacteriophage lambda is responsible for binding to LamB, its receptor at the surface of *Escherichia coli* K-12. *J Bacteriol*. 2000; 182:508–12. [PubMed: 10629200]
23. Katsura I. Morphogenesis of bacteriophage lambda tail. Polymorphism in the assembly of the major tail protein. *J Mol Biol*. 1976; 107:307–26. [PubMed: 1003470]
24. Murialdo H, Siminovitch L. The morphogenesis of bacteriophage lambda. IV. Identification of gene products and control of the expression of the morphogenetic information. *Virology*. 1972; 48:785–823. [PubMed: 4555611]
25. Xu, J. PhD Thesis. University of Pittsburgh; 2001. A conserved frameshift strategy in dsDNA long tailed bacteriophages.
26. Katsura I. Determination of bacteriophage lambda tail length by a protein ruler. *Nature*. 1987; 327:73–5. [PubMed: 2952887]
27. Katsura I, Hendrix RW. Length determination in bacteriophage lambda tails. *Cell*. 1984; 39:691–8. [PubMed: 6096021]
28. Roessner CA, Ihler GM. Proteinase sensitivity of bacteriophage lambda tail proteins gpJ and pH in complexes with the lambda receptor. *J Bacteriol*. 1984; 157:165–70. [PubMed: 6228546]
29. Scandella D, Arber W. An *Escherichia coli* mutant which inhibits the injection of phage lambda DNA. *Virology*. 1974; 58:504–13. [PubMed: 4132241]
30. Scandella D, Arber W. Phage lambda DNA injection into *Escherichia coli* pel- mutants is restored by mutations in phage genes V or H. *Virology*. 1976; 69:206–15. [PubMed: 1108413]

31. Katsura I, Kuhl PW. Morphogenesis of the tail of bacteriophage lambda. III. Morphogenetic pathway. *J Mol Biol.* 1975; 91:257–73. [PubMed: 1237629]
32. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 1997; 25:3389–402. [PubMed: 9254694]
33. Khoroshilova N, Popescu C, Munck E, Beinert H, Kiley PJ. Iron-sulfur cluster disassembly in the FNR protein of *Escherichia coli* by O₂: [4Fe-4S] to [2Fe-2S] conversion with loss of biological activity. *Proc Natl Acad Sci U S A.* 1997; 94:6087–92. [PubMed: 9177174]
34. Smith AD, Jameson GN, Dos Santos PC, Agar JN, Naik S, Krebs C, Frazzon J, Dean DR, Huynh BH, Johnson MK. NifS-mediated assembly of [4Fe-4S] clusters in the N- and C-terminal domains of the NifU scaffold protein. *Biochemistry.* 2005; 44:12955–69. [PubMed: 16185064]
35. Jakimowicz P, Cheesman MR, Bishai WR, Chater KF, Thomson AJ, Buttner MJ. Evidence that the *Streptomyces* developmental protein WhiD, a member of the WhiB family, binds a [4Fe-4S] cluster. *J Biol Chem.* 2005; 280:8309–15. [PubMed: 15615709]
36. Beinert H. Iron-sulfur proteins: ancient structures, still full of surprises. *J Biol Inorg Chem.* 2000; 5:2–15. [PubMed: 10766431]
37. Eaton GR, Eaton SS. Electron Paramagnetic Resonance Spectroscopy. *ChemInform.* 2004; 35:37–48.
38. Beinert H, Kennedy MC. Aconitase, a two-faced protein: enzyme and iron-regulatory factor. *The FASEB Journal.* 1993:1442–1449.
39. Emptage MH, Dreyers JL, Kennedy MC, Beinert H. Optical and EPR characterization of different species of active and inactive aconitase. *J Biol Chem.* 1983; 258:11106–11. [PubMed: 6309830]
40. Andreini C, Bertini I, Cavallaro G, Najmanovich RJ, Thornton JM. Structural analysis of metal sites in proteins: non-heme iron sites as a case study. *J Mol Biol.* 2009; 388:356–80. [PubMed: 19265704]
41. Riddles PW, Blakeley RL, Zerner B. Reassessment of Ellman's reagent. *Methods Enzymol.* 1983; 91:49–60. [PubMed: 6855597]
42. Ellman GL. Tissue sulfhydryl groups. *Arch Biochem Biophys.* 1959; 82:70–7. [PubMed: 13650640]
43. Golinelli MP, Akin LA, Crouse BR, Johnson MK, Meyer J. Cysteine ligand swapping on a deletable loop of the [2Fe-2S] ferredoxin from *Clostridium pasteurianum*. *Biochemistry.* 1996; 35:8995–9002. [PubMed: 8688437]
44. Golinelli MP, Chatelet C, Duin EC, Johnson MK, Meyer J. Extensive ligand rearrangements around the [2Fe-2S] cluster of *Clostridium pasteurianum* ferredoxin. *Biochemistry.* 1998; 37:10429–37. [PubMed: 9671512]
45. He WZ, Malkin R. Reconstitution of iron-sulfur center B of Photosystem I damaged by mercuric chloride. *Photosynthesis Research.* 1994; 41:381–388. [PubMed: 24310152]
46. Cammack R. Variations, Iron-Sulfur Clusters in Enzymes: Themes and Variations. *Advances in Inorganic Chemistry: Iron-Sulfur Proteins.* 1992; 38:281–322.
47. Beinert H, Kennedy MC, Stout CD. Aconitase as Ironminis signSulfur Protein, Enzyme, and Iron-Regulatory Protein. *Chem Rev.* 1996; 96:2335–2374. [PubMed: 11848830]
48. Dimple B, Ding H, Jorgensen M. *Escherichia coli* SoxR protein: sensor/transducer of oxidative stress and nitric oxide. *Methods Enzymol.* 2002; 348:355–64. [PubMed: 11885291]
49. Johnson DC, Dean DR, Smith AD, Johnson MK. Structure, function, and formation of biological iron-sulfur clusters. *Annu Rev Biochem.* 2005; 74:247–81. [PubMed: 15952888]
50. Fromme JC, Verdine GL. Structure of a trapped endonuclease III-DNA covalent intermediate. *EMBO J.* 2003; 22:3461–71. [PubMed: 12840008]
51. Fu W, O'Handley S, Cunningham RP, Johnson MK. The role of the iron-sulfur cluster in *Escherichia coli* endonuclease III. A resonance Raman study. *J Biol Chem.* 1992; 267:16135–7. [PubMed: 1644800]
52. Thayer MM, Ahern H, Xing D, Cunningham RP, Tainer JA. Novel DNA binding motifs in the DNA repair enzyme endonuclease III crystal structure. *EMBO J.* 1995; 14:4108–20. [PubMed: 7664751]

53. Guan Y, Manuel RC, Arvai AS, Parikh SS, Mol CD, Miller JH, Lloyd S, Tainer JA. MutY catalytic core, mutant and bound adenine structures define specificity for DNA repair enzyme superfamily. *Nat Struct Biol.* 1998; 5:1058–64. [PubMed: 9846876]
54. Pell LG, Liu A, Edmonds L, Donaldson LW, Howell PL, Davidson AR. The X-ray crystal structure of the phage lambda tail terminator protein reveals the biologically relevant hexameric ring structure and demonstrates a conserved mechanism of tail termination among diverse long-tailed phages. *J Mol Biol.* 2009; 389:938–51. [PubMed: 19426744]
55. Yang F, Forrer P, Dauter Z, Conway JF, Cheng N, Cerritelli ME, Steven AC, Pluckthun A, Wlodawer A. Novel fold and capsid-binding properties of the lambda-phage display platform protein gpD. *Nat Struct Biol.* 2000; 7:230–7. [PubMed: 10700283]
56. Pell LG, Kanelis V, Donaldson LW, Howell PL, Davidson AR. The phage lambda major tail protein structure reveals a common evolution for long-tailed phages and the type VI bacterial secretion system. *Proc Natl Acad Sci U S A.* 2009; 106:4160–5. [PubMed: 19251647]
57. Edmonds L, Liu A, Kwan JJ, Avanesky A, Caracoglia M, Yang I, Maxwell KL, Rubenstein J, Davidson AR, Donaldson LW. The NMR structure of the gpU tail-terminator protein from bacteriophage lambda: identification of sites contributing to Mg(II)-mediated oligomerization and biological function. *J Mol Biol.* 2007; 365:175–86. [PubMed: 17056065]
58. Maxwell KL, Yee AA, Booth V, Arrowsmith CH, Gold M, Davidson AR. The solution structure of bacteriophage lambda protein W, a small morphogenetic protein possessing a novel fold. *J Mol Biol.* 2001; 308:9–14. [PubMed: 11302702]
59. Michel-Briand Y, Baysse C. The pyocins of *Pseudomonas aeruginosa*. *Biochimie.* 2002; 84:499–510. [PubMed: 12423794]
60. Kondou Y, Kitazawa D, Takeda S, Tsuchiya Y, Yamashita E, Mizuguchi M, Kawano K, Tsukihara T. Structure of the central hub of bacteriophage Mu baseplate determined by X-ray crystallography of gp44. *J Mol Biol.* 2005; 352:976–85. [PubMed: 16125724]
61. Leiman PG, Shneider MM. Contractile tail machines of bacteriophages. *Adv Exp Med Biol.* 2012; 726:93–114. [PubMed: 22297511]
62. Browning C, Shneider MM, Bowman VD, Schwarzer D, Leiman PG. Phage pierces the host cell membrane with the iron-loaded spike. *Structure.* 2012; 20:326–39. [PubMed: 22325780]
63. Davidson AR, Sauer RT. Folded proteins occur frequently in libraries of random amino acid sequences. *Proc Natl Acad Sci U S A.* 1994; 91:2146–50. [PubMed: 8134363]

Phage tail tips are critical for assembly, host interaction and DNA injection.

Phage λ gpL is a conserved non-contractile tail tip protein.

gpL coordinates an iron-sulphur cluster.

Iron-sulfur cluster binding is a conserved feature of the gpL family.

The iron-sulphur cluster is likely required for biological activity

Figure 1. Homologues of gpL

A) The genomic region surrounding the *L* gene from phages λ , N15, HK022, and phi80 and the F-type pyocin from *Pseudomonas aeruginosa* PAO1. Genes encoding proteins with conserved functions are labeled by protein name and coloured similarly. The abbreviations used denote Tail Terminator Protein (TrP), Tail Tube Protein (TTP), Tape Measure Protein (TMP), and Tail Fibre (TF). The other genes are labeled according to the λ proteins or homologues that they encode. B) A sequence alignment of gpL homologues from Enterobacteria phages N15, T1, and ES18, and from the F-pyocin of *P. aeruginosa* PAO1 is shown. The pairwise sequence identity of these proteins to λ gpL ranges from 23 to 41%, and the average pairwise identity of all these sequences is 35%. The conserved Cys residues are boxed. The first 9 residues of the phage T1 homologue are not shown. The alignment was performed with MAFFT.

Figure 2. Mapping of gpL domains

A) The protein solution of the phage N15 gpL protein possesses a distinct reddish colour. B) gpL_{N15} was proteolyzed with trypsin and samples were taken at the time points indicated beneath the lanes in the gel. The digested protein was visualized on a 12% SDS polyacrylamide gel stained with Coomassie Blue. C) The gpL_{N15} sequence is displayed and the N- and C-terminal domains are highlighted in blue and yellow, respectively. The trypsin cut site at Arg173 is coloured pink and the highly conserved cysteine residues in the putative C-terminal domain are coloured red. D) Far-UV CD spectra for full-length gpL_{N15} (—), C-gpL_{N15} (●●●●), and N-gpL_{N15} (----).

Figure 3. Absorbance and EPR spectroscopic analysis of gpLN15

A) Absorbance spectra of full-length gpLN15 (—), C-gpLN15 (●●●●), and N-gpLN15 (----) purified under aerobic conditions. B) Absorbance spectrum of full-length gpLN15 purified aerobically (----) compared to the spectrum obtained on the same sample that was reconstituted with $\text{Fe}(\text{NH}_4)_2(\text{SO}_4)_2$ under anaerobic conditions (—). C) The EPR spectrum of reconstituted and oxidized gpLN15. A g-value of 2.014 was obtained.

Figure 4. Stability of the Fe-S cluster of gpL_{N15}

A) An absorbance spectrum of anaerobically reconstituted gpL_{N15} (—) was compared to samples stored aerobically (●●●●●) and anaerobically (----) for 24 hours. B) Anaerobically reconstituted gpL_{N15} was incubated aerobically and an absorbance spectrum was measured every 15 minutes up to 120 minutes and then every 30 minutes up to 240 minutes. A plot of absorbance versus time monitored at 415 nm is shown in the inset.

Figure 5. Spectroscopic and functional characterization of mutant proteins bearing substitutions of conserved cysteine residues in the C-terminal domain of gpL

A) Far UV CD spectra and B) absorbance spectra of wild-type gpL_{N15} and the indicated mutants with Ala substitutions of the four conserved Cys residues in the C-terminal domain of gpL_{N15}.

Figure 6.

A) Effect of individual amino acid substitutions on the biological activity of *gpL* *in vivo*. Lawns of amber-non-suppressing *E. coli* cells were spotted with dilution series of λ L amber (λ Lam 756 cI₈₅₇) and of λ wild type (λ cI₈₅₇). The leftmost spot in each series has 10⁵ phages, and the phage concentration decreases by ten-fold for each step to the right. The mutant gpL being expressed in the cells is shown to the left of the corresponding panels. Cys173, Cys182, Cys205 and Cys212 are the conserved cysteines of gpL; Cys12, Cys39, Cys140 and Cys208 are the nonconserved cysteines. B) Tail production by gpL amber and mutant alleles in pETail, which contains all 11 genes required for complete tails. The figure shows the relative amounts of tails produced in four representative cases; the notation under each panel indicates the gene *L* mutation in that experiment. In each case, expression of the tail proteins from the plasmid was induced and the proteins were labeled with ³⁵S-methionine, and the resulting lysates were fractionated on a 10–30% velocity sedimentation glycerol gradient. Twelve fractions were collected from each gradient and analyzed by SDS PAGE and autoradiography, with fraction 1 from the bottom of the gradient and fraction 12 from the top. Lane 13 is the unfractionated lysate. Complete tails run at roughly fraction 8 in these gradients, and we take the amount of major tail protein, gpV, in this fraction as an indication of the amount of tails made.

Figure 7. Alignment of C-terminal domains of gpL homologues from *Pseudomonas* and *Burkholderia* phages

The C-terminal domains of putative gpL homologues from *Pseudomonas* phages B3, D3112, and YuA, and *Burkholderia* phage BcepNazgul are aligned to the the C-terminal gpL domains of some of the λ gpL homologues shown in Fig. 1B. The conserved Cys residues are boxed.