Published in final edited form as:

J Mol Biol. 2013 September 23; 425(18): 3476–3487. doi:10.1016/j.jmb.2013.07.002.

A balanced ratio of proteins from gene *G* and frameshift-extended gene *GT* is required for phage lambda tail assembly

Jun Xu¹, Roger W. Hendrix, and Robert L. Duda[‡]

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

Abstract

In bacteriophage λ , the overlapping open reading frames *G* and *T* are expressed by a programmed translational frameshift similar to that of the *gag-pol* genes of many retroviruses, to produce the proteins gpG and gpGT. An analogous frameshift is widely conserved among other dsDNA tailed phages in their corresponding “*G*” and “*GT*” tail genes even in the absence of detectable sequence homology. The longer protein gpGT is known to be essential for tail assembly, but the requirement for the shorter gpG remained unclear because mutations in gene *G* affect both proteins. A plasmid system that can direct the efficient synthesis of tails was created and used to show that gpG and gpGT are both essential for correct tail assembly. Phage complementation assays under conditions where levels of plasmid-expressed gpG or gpGT could be altered independently revealed that the correct molar ratio of these two related proteins, normally determined by the efficiency of the frameshift, is also crucial for efficient assembly of functional tails. Finally, the physical connection between the G and T domains of gpGT, a consequence of the frameshift mechanism of protein expression, appears to be important for efficient tail assembly.

Keywords

assembly chaperones; protein complex assembly; protein-protein interactions; length regulation; bacteriophage assembly

Introduction

In bacteriophage λ , two open reading frames, *G* and *T*, located between the genes encoding the major tail protein and the tail length tape measure protein, produce two proteins that are related by a programmed translational frameshift (Figure 1) ¹. Translation of the *G* open reading frame (orf) terminates near the beginning of the *T* orf to produce gpG (gpG indicates the protein gene product of gene *G*). The second protein gpGT is made when about 3-4% of the ribosomes reading the *G* orf shift back one base on the mRNA at a slippery sequence at the *G-T* junction and continue in the -1 frame to the end of gene *T*. A frameshift analogous to the λ *G/GT* prototype appears to be conserved among most of the long-tailed dsDNA phages ², often in the absence of detectable sequence similarities. The conservation of the

© 2013 Elsevier Ltd. All rights reserved.

[‡]To whom correspondence should be addressed. duda@pitt.edu.

¹Present address: Mason Business Center, The Procter & Gamble Company, 8700 S Mason-Montgomery Road, Mason, OH 45040, USA

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.

frameshift among phages from a broad spectrum of hosts argues for a conservation of function and tail assembly mechanism. However, little is known about how gpG and gpGT function in the λ tail assembly pathway.

Genes *Z*, *U*, *V*, *G*, *T*, *H*, *M*, *L*, *K*, *I*, and *J* are all required for λ tail assembly³. Early studies showed that gpV was the major tail shaft protein⁴ and that the tail fiber (and host range determinant) was composed of gpJ^{5; 6; 7}. Many of the protein products of these essential genes were identified in SDS polyacrylamide gels⁸. The order of interaction of the protein gene products and the overall assembly pathway (Figure 1 (b)) were determined in large part from the results of *in vitro* complementation experiments between extracts of mutant infected cells⁹, and additional experiments by Katsura and many others as reviewed by Katsura¹⁰. From these studies it was clear that a slow sedimenting (15S) precursor produced by genes *J*, *I*, *L*, and *K* could be converted by the actions of genes *H*, *G*, and *M* to a faster sedimenting (25S) initiator of tail shaft polymerization, but the exact composition and nature of the 15S and 25S complexes remained unclear and no images of these complexes were ever obtained by electron microscopy. The exact roles of many of the proteins have remained unclear to this day, although it was later shown that gpH functions as a tape measure protein that determines the overall length of tails^{11; 12}. The function of gene “*T*” in the production of gpGT by frameshifting was not discovered until much later¹; however we have updated the pathway in Figure 1B to show the likely entry point of gpGT into the tail assembly scheme.

The protein gpG is expressed to a high level during phage infections, similar to the level seen for the major tail protein gpV^{1; 8}. This is surprising because neither gpG nor gpGT is present in the mature tail¹, especially since phages typically express structural proteins in approximately the same ratio as they are found in the mature structure, with the notable exceptions of scaffolding or chaperone proteins. Earlier studies using an amber mutant in the *T* region showed that gpGT is required for formation of the initiator and no visible tail structure was seen in the absence of gpGT^{1; 9; 13} (Note that because of the overlapping relationship of the two “genes”¹, amber mutants in the *G* and *TORF*’s fail to complement each other genetically. As a result, both types of mutants were initially considered to map to the same gene (*G*). The *T* amber mutant used here is called *Gan901* in the earlier literature.) Because any mutation in the *G* region will affect not only gpG but also the essential gpGT, it is not known whether gpG is required for tail assembly. One extreme possibility is that gpG is not essential, but it is expressed only to provide a basis for producing the appropriate amount of gpGT through an inefficient frameshift. This seems unlikely, since the expression level of most phage structural genes can be determined at the level of translational initiation efficiency, and the variation among genes can be as much as 800 fold^{8; 14; 15}. Furthermore, the expression of gpG itself is likely controlled at the level of translational initiation, which means gpGT’s expression level is also indirectly controlled at translation initiation. An alternative view is that gpG is essential. Since the frameshift is highly conserved among different phages, even though often no primary sequence similarity can be detected, this would argue that the outcome from the frameshift expression mechanism—producing two proteins sharing the same N-terminus in a fixed ratio—is important.

A plasmid based tail assembly system was established in which the 11 kb tail gene region was cloned under the control of a T7 promoter. The resulting construct can produce active tails. Through a series of complementation experiments in which we can decouple the relative amounts of the two proteins from the efficiency of the frameshift, we show that gpG is indeed essential and that the appropriate ratio of gpG to gpGT is important for successful tail assembly.

Results

Biologically active λ tails made from a plasmid

λ tail assembly has been studied extensively using phage infection or induction of lysogens carrying amber or deletion mutations in the tail region of the prophage¹⁰. Constructing targeted point mutations is difficult in that context, so we cloned ~11 kb of DNA containing the essential λ tail genes (including *Z*, *U*, *V*, *G*, *T*, *H*, *M*, *L*, *K*, *I*, and *J*, but not tail fiber genes *tfa* and *stf*) into a T7 promoter-based vector, creating pETail4. Figure 2 (a) shows an SDS-PAGE gel of the radio-labeled proteins expressed from pETail4. Note the presence of gpH* which is the processed form of gpH, an indication of completed tail assembly¹⁶. The proteins expressed from pETail4 assemble into biologically active tails as determined by their ability to join free heads to form infective phages in complementation assays¹⁷. The expression system is highly efficient, as judged by the high yield from *in vitro* complementation with a head donor extract. Extracts from cells carrying pETail4 gave more than 40-fold higher complementation activity (data not shown) than extracts from an induced lysogen defective in head assembly.

The tails made using pETail4 were purified using a modification of a protocol developed for phage HK97 prohead purification (see Materials & Methods). Figure 2 (b) shows the proteins from the purified tails separated by SDS-PAGE and visualized by Coomassie blue stain. Only five proteins were seen on the gel, namely gpU, gpM, gpV, gpH* (cleaved product of gpH) and gpJ. The products of the remaining tail genes *I*, *K*, *L*, and *Z* have essential roles in the early steps of tail assembly, but were not found in tails here. An electron micrograph of negatively stained purified tails is shown in Figure 2 (c) demonstrating that their morphology is normal and uniform. This plasmid-based tail production system (using a modified host, see Materials and Methods) became the basis for all further studies, since it makes it easy to build plasmids containing different subsets of the tail genes and to make mutations at will.

G is required for tail assembly

Neither gpG nor gpGT is incorporated into mature λ tails or phage¹. An amber mutation in the *T* region blocks tail assembly, indicating that gpGT is essential⁹. However the special relationship between gpG and gpGT makes it less straightforward to determine whether gpG is required because any mutation in the *G* region will also affect gpGT. To ask whether gpG is essential, we created a mutation (gpGT-mut) that fuses the *G* and *T* orf's and therefore allows gpGT to be expressed without a frameshift (Figure 3(a)). If the resulting plasmid construct can produce active tails, it would argue that gpG is not required because gpGT alone can support tail assembly. On the other hand, if the resulting construct does not produce active tails, it would be consistent with the view that gpG is required for the assembly but does not prove it because expression of gpGT without frameshift will result in a 30 fold increase in the amount of gpGT produced. The non-complementation result might be due to an abnormally large amount of gpGT blocking tail assembly.

The *GT* fusion mutation was introduced into pETail4 to create plasmid pETail5. Radio-labeling of the proteins produced from pETail5 shows the overproduction of gpGT (Figure 3(c)). Notably absent on the pETail5 gel was the gpH* band that was clearly observed for pETail4 (the cleavage product of gpH generated during tail assembly¹⁸) which indicates that tail assembly has not reached completion. A complementation assay carried out by mixing free heads with an extract prepared from cells expressing pETail5 failed to produce tail activity above background level, which indicates that the *GT* fusion mutation blocks tail assembly (data not shown).

gpGT alone is not enough to complement λ with an amber mutation in the G region

We next used complementation experiments to more specifically test whether gpG is essential. In this experiment, we tested whether cells that express gpGT only from a plasmid can support the growth of λ phage with an amber mutation in the G region, which produces neither gpG nor gpGT. If the cells expressing gpGT alone do not complement phages expressing neither gpG nor gpGT but do complement phages expressing gpG but not gpGT, this would argue that gpG is indeed required for tail assembly.

In each plate in the complementation experiments, increasing concentrations of different amber mutant and wild -type phages were spotted on lawns of cells containing a plasmid that expresses the proteins to be tested. If the proteins expressed from the plasmid can complement the defective amber phage, clear spots will be formed. The lower the phage concentration required to form a clear spot, the better the complementation.

Plasmids pT7-5-GT, pT7-5-GT-mut and pET21a-T, were tested for their ability to complement phages λ *Gam*, which produces neither gpG or gpGT, λ *Tam*, which produces gpG only, and λ *c857*, a wild-type control. The plasmids and phages used are described in Table 1 and the results are shown in Figure 4.

The plasmid pT7-5-GT-mut which makes gpGT alone does not complement phage λ *Gam447cIts1* which produces neither gpG nor gpGT (Figure 4(a) 2). However, λ *Gam* can be complemented by pT7-5-GT, which expresses both gpG and gpGT (Figure 4(a) 1). An important control shows that plasmid pT7-5-GT-mut is able to complement the λ *Tam* phage, which produces gpG but not gpGT (Figure 4(a) 2). Taken together, the results show that complementation only works when both gpG and gpGT are present, either provided from plasmids or from the phages themselves, which argues that gpG is required for productive tail assembly. We also found that a plasmid that expresses the *T*orf alone (pET21a-T) cannot complement the λ *Tam* phage that makes gpG, but no gpGT (Figure 4(b) 3), showing that the *T* domain of gpGT cannot efficiently replace the missing gpGT, even in the presence of gpG. We also note that plasmid pT7-5-GT, which expresses both gpG and gpGT only very weakly complements the λ *Tam* phage, which produces only gpG (Figure 4(b) 1). The 100 fold reduced complementation in this situation suggests that the ratio of G to GT may be important, perhaps disrupted in this case due to excess gpG produced by the λ *Tam* phage that cannot be overcome by supplying both gpG and gpGT (in their normal ratio) in *trans*.

We recognize that it is not strictly true that plasmid pT7-5-GT-mut, which produces gpGT as fusion protein, makes no gpG. As a result of translational frameshifting at the normal site, we expect that a small amount of a new protein will be made in which the C-terminal 5 amino acids of gpG are replaced with a different sequence (Figure 3(b)). The SDS gel and Western blot in Figure 3(d) confirm this expectation and we refer to the C-terminal altered gpG as gpG*. The effect of the GT-mut fusion mutation is therefore two fold. First by making gpG* rather than gpGT as the product of the frameshift, it changes the gpG (or gpG*) to gpGT ratio by roughly 900 fold (from 30:1 gpG:gpGT to 1:30 gpG*:gpGT). Second the mutation leads to the 5 amino acid swap at the C-terminus of gpG mentioned above. The fact that the GT-mut mutant does not make functional tails may be because gpG* is not functional, but we believe a more likely explanation is the gpG:gpGT ratio is inappropriate in the mutant. We show in the following section that the ratio must be approximately correct to allow efficient assembly. We have not seen any indication of biological activity for gpG*, either stimulating or inhibiting tail production, but would not expect to because in all experiments it is expressed in much lower amounts than gpGT. We do not believe that the production of gpG* in the mutant significantly affects any of our conclusions about tail assembly.

The gpG/gpGT ratio is important for tail assembly

The frameshift provides a mechanism to produce two proteins that share the same N - terminus. The ratio of the two proteins normally is determined by the efficiency of the frameshifting. We again used a spot complementation experiment to investigate whether the ratio between gpG and gpGT is important for tail assembly. In this experiment, we provided gpG and gpGT from an arabinose controlled plasmid expression system. The arabinose promoter has the property that the expression level can be modulated in a linear fashion by varying the amount of arabinose inducer added ¹⁹. The ratio of gpGT to gpG inside the cell can be represented by the formula below where the subscripts *plasmid* and *phage* indicate the source of the protein.

$$Ratio_{gpGT/gpG} = \frac{gpGT_{plasmid} + gpGT_{phage}}{gpG_{plasmid} + gpG_{phage}} \quad (\text{Formula 1})$$

Assuming the proteins produced from the input phages stay in similar levels ($gpGT_{phage}$ and gpG_{phage} stay the same) across different concentrations of inducer used, a higher amount of inducer will result in a higher concentration of the protein produced from the plasmid (higher level of $gpGT_{plasmid}$ and $gpG_{plasmid}$) and hence change the ratio of gpG vs. gpGT.

The plasmids used for the arabinose-modulated expression experiments are described in Table 1 and the experimental results are shown in Figure 5. Panel (a) shows the complementation results for λ *Gam* (amber in the *G* region which produces neither gpG nor gpGT) on plasmids pBAD33-GT (produces wild type gpG and gpGT), pBAD33-GT-mut (produces gpGT only) and pBAD33-T (produces the *T* part of gpGT). As the concentration of arabinose increases, pBAD33-GT begins to complement and complementation signal increases with increasing inducer (Figure 5 (a) 1a-1f), while plasmid pBAD33-GT-mut, which provides gpGT only, does not complement under any of the tested concentrations of arabinose (Figure 5(a), 2a-2f). This result is consistent with the previous result that gpG is required for tail assembly and supplying gpGT alone *in trans* does not complement a phage that produces neither gpG nor gpGT. pBAD33-T does not complement, showing as expected that the *T* region alone cannot substitute for gpG and gpGT.

Figure 5 (b) shows the complementation results with λ *Tam* (amber mutation in the *T* region which produces gpG but not gpGT) on the same set of plasmids as in Figure 5 (a). As the concentration of arabinose increases, pBAD33-GT begins to complement λ *Tam*, but not to the extent seen for λ *Gam* arguably because the excess gpG produced by the λ *Tam* phage cannot be overcome by supplying both gpG and gpGT (in their normal ratio) *in trans*. For pBAD33-GT-mut, which provides gpGT only, as the concentration of arabinose increases, λ *Tam* begins to complement (Figure 5(b) 2a-2d) and then as the arabinose concentration further increases the complementation level decreases (Figure 5 (b) 2e-2f). Both of these results show that either too much gpGT or too little gpGT supplied from the plasmid does not support efficient assembly and only at an intermediate concentration of gpGT, corresponding to an intermediate gpGT/ gpG ratio, does the complementation work well. This result supports the idea that the ratio of gpG to gpGT is important for assembly.

When the T domain of gpGT ("gpT") is expressed alone from the plasmid and supplied to the phage that makes gpG alone (Figure 5(b) 3a-3f), a weak complementation signal can be seen, but it is exquisitely dependent on the level of expression from the plasmid. This shows that gpT can have some limited functional activity and, because this activity was not evident when gpT was supplied to the phage that makes neither gpG nor gpGT (Figure 5(a) 3a-3f), we conclude that both gpG and gpT must be present in order for the activity of gpT to be

realized. In parallel to what we saw when gpGT was supplied to the phage producing only gpG (Figure 5(b) 2a-f), the gpT-dependent tail assembly activity decreases at the highest level of gpT expression, arguing that gpT is most effective at a particular gpT/ gpG ratio. Importantly, the amount of complementation activity at the optimal gpT/ gpG ratio (Figure 5(b) 3e) is only about 1% of the complementation activity supported by the optimal gpGT/ gpG ratio (Figure 5(b) 2d). This result argues that the physical connection between the G and T domains of gpGT is important for achieving efficient assembly of functional tails.

Figure 5 panel (c) shows the complementation results from plating our wild-type phage λ c857 on the same set of plasmids. Expression of the normal ratio of gpG and gpGT from plasmid pBAD33-GT at any level has no effect on the growth of wild-type phage (Figure 5(c) 1a-f), but at the highest concentration of arabinose induction of pBAD33-GT-mut, plasmid expression clearly inhibits the growth of wild type phage (Figure 5(c) 2f). This again is consistent with the explanation that the excess gpGT provided from the plasmid disturbs the ratio of gpG and gpGT. The effect of disturbance of pBAD33-GT-mut on wild type phage is not as severe as on λ Tam, even though in theory the ratio is disturbed more in wild type phage under higher concentration of arabinose because wild type phage produces gpGT too (Formula 1). A possible explanation for this observation is given in the Discussion. We also noted a weak inhibitory effect of the highest level of expression of “gpT” on the growth of wild -type phage. This is shown by the slightly poorer growth of λ c857 on the cells producing maximal “gpT” (Figure 5(c) 3f). This supports the conclusion that “gpT” can have biological activity, even when it is not properly attached to gpG, if it is expressed at sufficiently high levels.

Discussion

G is essential for tail assembly

The conservation of the frameshift strategy at the region corresponding to the λ *G/GT* region among a broad range of tailed phages² argues that the frameshift plays a central role in tail assembly and makes it very interesting to know what function gpG and gpGT provide in the tail assembly pathway. The protein gpG is expressed at about the same high level as major tail protein gpV, but it is not incorporated into mature tails^{1; 8}. It is known from earlier experiments that gpGT is essential, as amber mutations in the *T* region block tail assembly⁹. However, evidence supporting an essential role for gpG was lacking because any mutation in the *G* region will also affect gpGT. To address this question, we created a plasmid-based tail assembly system that can produce active tails and found that introducing a mutation that allows gpGT to be expressed without frameshifting abolishes tail assembly. This result is consistent with the view that gpG is essential. However, the mutation also resulted in a 30-fold increase of gpGT relative to the other tail proteins. That is to say, the relative ratios of the tail proteins have been altered due to the mutation. Consequently, the inability of the gpGT-overexpressing plasmid pETail5 to produce active tails could be solely the result of a non-functional ratio of tail proteins. To further investigate the role of gpG in tail production, we used a plasmid-by-phage complementation system in which plasmids providing gpG and/ or gpGT *in trans* were tested for their ability to complement amber mutations in the *G* or *T* regions of an infecting phage. From the different combinations of complementation tests (Figure 4), we can draw the following conclusions: 1) gpG and gpGT produced from the plasmid are functional, as they can complement amber mutations in the phage; 2) complementation only works when both gpG and gpGT are present, either provided from a phage or from a plasmid. These results argue strongly that gpG is essential for tail assembly. They also confirm the essential nature of gpGT.

The gpG/gpGT ratio is important

The effect of the λ tail gene frameshift is twofold. On the one hand it creates two proteins, gpG and gpGT, with gpGT containing almost the entire sequence of gpG. On the other hand the molar ratio of the two proteins is determined by the efficiency of frameshifting. The arabinose-controlled plasmid by phage complementation system provides us a tool to regulate the relative amounts of gpG and gpGT independently of the frameshift by regulating the expression of gpG and gpGT from plasmids. When we modulated the expression of gpGT by inducing under a wide range of arabinose concentrations, we found that there is a gpGT expression level at which complementation works best, and this argues that there is a gpGT/ gpG ratio that is optimal for tail assembly.

A key assumption in the experiments with the arabinose mediated expression system is that a higher concentration of arabinose will result in higher expression from the plasmid in each individual cell. It has been shown for the plasmids we used here that a higher concentration of arabinose results in a higher level of protein expression in cell cultures and that the level of protein expression can be modulated linearly over a range of arabinose concentrations¹⁹. This result can be explained simply if the level of the expressed protein in each individual cell is modulated by the level of the arabinose inducer. However it is also consistent with a model in which individual cells are either fully repressed or fully induced, with the graded response to inducer reflecting the fraction of the population in the fully induced state. It has been argued that the latter scenario is likely to be the case²⁰. However, our results appear to be incompatible with the latter scenario and consistent with the view that protein expression is modulated in a graded manner in each individual cell. In the complementation experiment in which gpGT is supplied to the phage that cannot make gpGT, the level of complementation at first increases with increasing amounts of arabinose inducer but then decreases sharply at the highest levels of arabinose. This decrease at high arabinose is simply explained if gpGT is increasing in each cell to a point where the gpGT/ gpG ratio is too high to allow tail assembly. It is inconsistent with the model in which the fraction of fully induced cells in the population is increasing with increasing arabinose, unless we assume that high concentrations of arabinose itself are strongly inhibitory to tail assembly. We can rule out the latter possibility because in experiments in which the normal gpGT/ gpG ratio is maintained (e.g., (Figure 5(c) 1a-f), maximal levels of plaque formation are seen at all arabinose concentrations, including the highest.

We have also investigated the significance of the gpGT/ gpG ratio using a variation of the complementation conditions used in Figure 4: the basal levels of protein expression levels of gpG and/ or gpGT from T7 promoter-based plasmids (uninduced in host BL21(DE3) Δ tail) were elevated by removing the pLysS plasmid that was present in the experiment shown in Figure 4 (Studier, 1990). These experiments recapitulated the part of the arabinose results showing that complementation efficiency declines substantially when too much gpGT is supplied (data not shown). Taken together, the complementation results provide strong evidence that efficient assembly of tails requires not only the presence of both gpG and gpGT but their presence in the appropriate molar ratio as well.

gpG and gpGT act as chaperones for gpH and recruit the major tail protein for assembly

In a separate publication²¹, we show that gpG and the “G” part of gpGT both bind to almost the entire length of gpH and thus appear to act as chaperones for the tail tape measure protein. We also show that gpH only functions efficiently when complexed with gpG and gpGT and we believe that gpG and gpGT keep the tape measure protein soluble and competent for assembly. In addition, we show that the “T” part of gpGT binds to the lambda major tail protein, gpV, and also induces gpV to polymerize into tubular forms that resemble tails. These biochemical results place constraints on how we can understand the roles of gpG

and gpGT in tail assembly. We suggest that tape measure protein gpH, as a coiled -coil of a few molecules, is initially coated with a mixture of gpG and gpGT and then the gpG and gpGT are replaced by major tail subunit gpV, which forms a tube around the coiled -coil of gpH. We believe that this transition most likely includes a crucial role for the activity of the “T” domain of gpGT in binding gpV and in converting it to a polymerization -competent form. Possibly it is successful polymerization of gpV around gpH that displaces the gpG and gpGT chaperones. Testing this or any other model for how gpV becomes part of the tail and how gpG and gpGT depart will require, we believe, microscopic or other structural evidence about the nature of the complex of the gpG and gpGT chaperones with the tape measure protein. In particular it will be important to determine how molecules of gpG and gpGT are distributed in the complex, as that is likely to illuminate the question of why the gpG:gpGT ratio is important for correct function. None of this information is currently available.

The physical linkage between the G and T domains of gpGT is important for tail assembly

A consequence of the frameshift mechanism of expression is that the T protein domain is always found joined to the G domain as a part of gpGT and never as a free standing protein. Given the strong conservation of the frameshift mechanism among tailed phages, this suggests that the physical connection between the G and T domains of gpGT may be important for the function of the T domain. Our complementation experiments provide support for this view. When the T domain is artificially expressed as a separate protein (gpT), it has only a very weak ability to complement a phage that cannot make gpGT. As with the complementation of the same phage by a plasmid producing gpGT, the complementation by gpT shows a maximum at a particular gpT/ gpG ratio. However, the maximum level of complementation provided by gpT is reduced at least 100-fold relative to the parallel complementation by gpGT. Since the only difference between these two situations is ostensibly the presence or absence of the physical connection between the G and T domains of gpGT, this argues that that physical connection is an essential part of the function of the T domain and therefore an essential part of tail assembly.

Materials and Methods

Plasmid construction

pETail4: 16KB BamHI-EcoRI fragment of λ DNA containing the entire 11 tail genes was cloned into pBluescript(sk-) (Stratagene) BamHI-EcoRI sites to create pBS- λ -tail I. The large fragment of pBS- λ -tail-I BswI-HindII digestion was filled in and circularized to create pBS- λ -tail 3. pBS- λ -tail3 was cut with BamHI, filled in and then cut with NsiI and ligated to 2.6 kb BsaAI-NsiI fragment of pBS- λ -tail I to create pBSTail 5. The SacI XhoI fragment of pBSTail 5 was inserted between the SacI-XhoI sites of pET21+ to create pETail I. BstBI-Xho I digestion of PCR fragment produced by using primers J1 (gaggagtttcgaaag), J2 (ggctcgagacgaacctgtgaac) and pETail I as template was inserted back into the BstBI-XhoI sites of pETail I. The resulting construct removes gene *lom*. pETail 4 contains all of genes *Z*, *U*, *V*, *G*, *T*, *H*, *M*, *L*, *K*, *I*, and *J*, including 236 bp before the start of gene *Z* and 30 bp after the end of gene *J*.

pETail5: PstI-MscI fragment of pBRUM-1 (Uma Chandran, PhD thesis, 1990, University of Pittsburgh) containing part of G-T was cloned into PstI and EcoRV sites of pBluescript(sk-) to create pBS-GT.

The PCR product using GT1 (gccctgcagcgcattgagcatctc) and GT2 (gaaggcctttcccgagaacagg) as primer and pBRUM-1 as template was cut with PstI-StuI and inserted back into pBS-GT PstI and XmnI-partial digest vector. The resulting construct (pBS-GT-mut) added one more G right after the slippery sequence which allowed gpGT to

be synthesized without a frameshift. pETail5 was created by replacing the ApaI -Nsi I fragment of pETail4 with the ApaI-NsiI fragment from pBS-GT-mut. The resulting plasmid contains all the tail genes and *GT*-fusion mutation.

pT7-5-GT: MscI fragment of pBRUM-1 (Uma Chandran, PhD thesis, 1990, University of Pittsburgh) was cloned into pT7-5 SmaI site to create pT7-5-VT. The 1 kb XbaI HincII fragment containing G/ GT of pT7-5-VT was inserted into XbaI-SmaI sites of pT7-5. The resulting plasmid will express gpG and gpGT.

pT7-5-GT-mut: The ApaI-NsiI fragment of pBS-GT-mut was inserted into pT7-5-VT ApaI-NsiI site to create pT7-5-VT-mut. The 1 kb XbaI HincII fragment of pT7-5-VT-mut was inserted into XbaI-SmaI sites of pT7-5 to create pT7-5-GT-mut.

pBAD33-GT: SacI HindIII fragment of pT7-5-GT was inserted into SacI HindIII sites of pBAD33.

pBAD33-GT-mut: SacI HindIII fragment of pT7-5-GT-mut was inserted into SacI HindIII sites of pBAD33.

pET21a-T: pET21a was digested with EcoRI, blunted with mung bean nuclease and then cut with Hind III. The resulting vector was ligated to the pBS-GT XmnI/ HindIII fragment to create pET21a-T. The resulting plasmid contains only the T part of GT, fused to the T7 tag from the vector.

pBAD33-T: pET21a was digested with EcoRI, blunted with mung bean nuclease, and then cut with Hind III. The resulting vector was ligated to the pBS-GT XmnI/ HindIII fragment to create pET21a-T. The resulting plasmid contains only the T part of GT and fused to the T7 tag from the vector. The XbaI HindIII fragment of pET21a-T was inserted into XbaI HindIII sites of pBAD33 to create pBAD33-T.

Modification of the expression host

The host used to express T7 promoter based plasmids, BL21(DE3), has a λ derivative prophage, DE3, bearing the T7 RNA polymerase gene integrated into the chromosome. To avoid interference from the λ tail genes in the DE3 prophage, either by *trans* expression or by recombination with the plasmid, we knocked out the tail genes using an integrating suicide vector gene replacement strategy²². DNA flanking both sides of the region to be deleted was cloned into the PKO5 vector²². PKO5 has a temperature sensitive replication origin from pSC101, a chloramphenicol resistance gene (CAM) and a *SacB* gene which is a counter-selective marker in the presence of sucrose. The plasmid was transformed into BL21(DE3) and the transformed cells were plated on a LB/ CAM plate and incubated at 42°C. Because of the *ts* replication origin, cells with chloramphenicol resistance must have the plasmid integrated into the chromosome by homologous recombination. The colonies that grew were plated onto a plate with 5% sucrose at 30°C to select for cells that lost the *SacB* gene. The loss of the plasmid is due to a second recombination event which either leaves the original chromosome intact or a deletion. The cells with the deletion were identified by colony PCR and designated BL21(DE3 Δ tail).

Protein expression

T7 promoter controlled plasmids were expressed in BL21(DE3) or BL21(DE3)pLysS cells^{23; 24} as described previously²⁵, except BL21(DE3 Δ tail) was used. Overnight cells grown in LB/ Amp (50 μ g/ ml Amp) (BL21(DE3)) or LB/ Amp/ Cam (50 μ g/ ml Amp, 25 μ g/ ml Cam) (BL21(DE3)pLysS) were inoculated with 1:2000 dilution. The cells were grown at 37°C with vigorous shaking to a density of approximately 4×10^8 cells/ ml and

induced with IPTG at a final concentration of 0.4 mM for pT7-5 based vectors and 1mM for pET21 based vectors (Novagen). The cells were grown for another 3.5 hours, chilled and harvested by centrifugation. Radio-labeling was performed as previously described²⁶. Briefly, BL21(DE3) containing appropriate plasmid was grown at 37°C in RG-glucose medium until the culture was at $\sim 4 \times 10^8$ cells/ ml. IPTG was added at a final concentration of 0.4 mM to induce for 15 minutes after which rifampicin was added to a final concentration of 0.2 mg/ ml, to shut down expression of host genes for 15 min. ³⁵S-methionine was added to 10 μ Ci/ ml and the culture was incubated for another 15 min. The cells were then harvested by centrifugation for further analysis.

Tail purification

Plasmid pETail4 or pT7-5-Tail4 containing the complete set of tail genes was expressed in BL21(DE3) pLysS cells as described above. Induced cells were harvested and lysed in lysis buffer (50 mM Tris-HCl, pH 8.0, 5 mM EDTA). The lysate was clarified by centrifugation, and tails precipitated by adding potassium glutamate to 0.7 M and PEG 8000 to 5.0% (w/ v) followed by sedimentation to collect the precipitate. The pellet was resuspended in TKG buffer (20 mM Tris-HCl, pH 7.5, 100 mM potassium glutamate), loaded onto a 10~30% glycerol gradient and centrifuged in the SW28 rotor (Beckman) at 27 krpm for 5 hrs. The tail band was extracted with a 20G needle and the tails were pelleted by ultracentrifugation in a Ti45 rotor (Beckman) at 40,000 rpm for 4.5 hrs. The tail pellet was resuspended in a small volume, typically 1 ml (from 1 liter culture), of TKG buffer.

SDS polyacrylamide gel electrophoresis

SDS Gel electrophoresis was performed as described in²⁷ in a 1 mm mini-gel apparatus. Protein samples were mixed with proper amount of 4 $\times$ SDS sample buffer and boiled for 3 minutes. The samples were then loaded onto a gel and run at 150 V constant voltage until the blue dye ran out of the gel.

Spot EOP complementation tests

Spot tests compare the relative EOP (efficiency of plating) of amber mutant phage (lacking gpGT, or both gpG & gpGT) grown on a hosts synthesizing various amounts and ratios of gpG and gpGT (from a plasmid). Spot tests are variants of phage T4 techniques^{28; 29} and have also been used with HK97³⁰. Overnight culture of cells carrying the appropriate plasmid (150 μ l) were plated in 2.5 ml soft agar overlays onto tryptone agar plates with appropriate antibiotics (and for some experiments arabinose was added to different levels to induce varying levels of gene expression: $2 \times 10^{-6}\%$, $2 \times 10^{-5}\%$, $2 \times 10^{-4}\%$, $2 \times 10^{-3}\%$, or $2 \times 10^{-2}\%$) (w/ v) arabinose was added). The soft agar on the plates was allowed to set for a few minutes before proceeding. Phage stocks (wild -type or mutant) were serially diluted to obtain a set of stocks with 10^8 , 10^7 , 10^6 , 10^5 , 10^4 , and 10^3 , plaque forming units per ml and a 4 μ l aliquot of each dilution was spotted onto the lawn. The plate was incubated at 37°C overnight and complementation scored by recording the number of clear spots formed by growth of phage within the spot or by recording an image of the plate using a scanner or camera. The number of spots appear to be a true indicator of complementation, as distinguished from marker rescue due to recombination between the phage and plasmid genes. This is internally controlled for in the panels of Figure 5 where in each vertical set of data the opportunity for recombination always the same (both plasmid and phage are present), but spots (indicating complementation) appear only at particular levels of the inducer, arabinose.

Acknowledgments

This work was supported by NIH grant R01 GM47795 to RWH. We thank Jon Beckwith for providing us the pBAD vector series.

References

1. Levin ME, Hendrix RW, Casjens SR. A programmed translational frameshift is required for the synthesis of a bacteriophage lambda tail assembly protein. *J Mol Biol.* 1993; 234:124–39. [PubMed: 8230192]
2. Xu J, Hendrix RW, Duda RL. Conserved translational frameshift in dsDNA bacteriophage tail assembly genes. *Mol Cell.* 2004; 16:11–21. [PubMed: 15469818]
3. Parkinson JS. Genetics of the left arm of the chromosome of bacteriophage lambda. *Genetics.* 1968; 59:311–25. [PubMed: 5705826]
4. Buchwald M, Steed-Glaister P, Siminovitch L. The morphogenesis of bacteriophage lambda. I. Purification and characterization of lambda heads and lambda tails. *Virology.* 1970; 42:375–89. [PubMed: 4099070]
5. 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]
6. Dove WF. Action of the lambda chromosome. I. Control of functions late in bacteriophage development. *J Mol Biol.* 1966; 19:187–201. [PubMed: 5967283]
7. Mount DW, Harris AW, Fuerst CR, Siminovitch L. Mutations in bacteriophage lambda affecting particle morphogenesis. *Virology.* 1968; 35:134–49. [PubMed: 5652675]
8. 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]
9. Katsura I, Kuhl PW. Morphogenesis of the tail of bacteriophage lambda. III. Morphogenetic pathway. *J Mol Biol.* 1975; 91:257–73. [PubMed: 1237629]
10. Katsura I. Tail assembly and injection. In: Hendrix, R.; Roberts, J.; Stahl, F.; Weisberg, RA., editors. *Lambda II.* Cold Spring Harbor Laboratory; Cold Spring Harbor, NY: 1983. p. 331–346.
11. Katsura I. Determination of bacteriophage lambda tail length by a protein ruler. *Nature.* 1987; 327:73–5. [PubMed: 2952887]
12. Katsura I, Hendrix RW. Length determination in bacteriophage lambda tails. *Cell.* 1984; 39:691–8. [PubMed: 6096021]
13. 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]
14. Ray PN, Pearson ML. Evidence for post-transcriptional control of the morphogenetic genes of bacteriophage lambda. *J Mol Biol.* 1974; 85:163–75. [PubMed: 4835731]
15. Sampson LL, Hendrix RW, Huang WM, Casjens SR. Translation initiation controls the relative rates of expression of the bacteriophage lambda late genes. *Proc Natl Acad Sci U S A.* 1988; 85:5439–43. [PubMed: 2969591]
16. Hendrix RW, Casjens SR. Protein cleavage in bacteriophage lambda tail assembly. *Virology.* 1974; 61:156–9. [PubMed: 4415352]
17. Weigle J. Assembly of phage lambda in vitro. *Proc Natl Acad Sci U S A.* 1966; 55:1462–6. [PubMed: 5227665]
18. Tsui LC, Hendrix RW. Proteolytic processing of phage lambda tail protein gpH: timing of the cleavage. *Virology.* 1983; 125:257–64. [PubMed: 6220513]
19. Guzman LM, Belin D, Carson MJ, Beckwith J. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol.* 1995; 177:4121–30. [PubMed: 7608087]
20. Siegle DA, Hu JC. Gene expression from plasmids containing the araBAD promoter at subsaturating inducer concentrations represents mixed populations. *Proc Natl Acad Sci U S A.* 1997; 94:8168–72. [PubMed: 9223333]

21. Xu J, Hendrix RW, Duda RL. Chaperone-protein interactions that mediate assembly of the bacteriophage lambda tail to the correct length. *Journal of Molecular Biology*. 2013 accepted, under revision.
22. Link AJ, Phillips D, Church GM. Methods for generating precise deletions and insertions in the genome of wild -type *Escherichia coli*: application to open reading frame characterization. *J Bacteriol*. 1997; 179:6228–37. [PubMed: 9335267]
23. Studier FW. Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system. *J Mol Biol*. 1991; 219:37–44. [PubMed: 2023259]
24. Studier FW, Rosenberg AH, Dunn JJ, Dubendorff JW. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol*. 1990; 185:60–89. [PubMed: 2199796]
25. Duda RL, Hempel J, Michel H, Shabanowitz J, Hunt D, Hendrix RW. Structural transitions during bacteriophage HK97 head assembly. *J Mol Biol*. 1995; 247:618–35. [PubMed: 7723019]
26. Duda RL, Martincic K, Hendrix RW. Genetic basis of bacteriophage HK97 prohead assembly. *J Mol Biol*. 1995; 247:636–47. [PubMed: 7723020]
27. Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*. 1970; 227:680–5. [PubMed: 5432063]
28. Benzer S. Fine Structure of a Genetic Region in Bacteriophage. *Proc Natl Acad Sci U S A*. 1955; 41:344–54. [PubMed: 16589677]
29. Doermann AH, Bochner L. The identification of complex genotypes in bacteriophage T4. I. *Methods. Genetics*. 1970; 66:417–28. [PubMed: 4933468]
30. Ross PD, Cheng N, Conway JF, Firek BA, Hendrix RW, Duda RL, Steven AC. Crosslinking renders bacteriophage HK97 capsid maturation irreversible and effects an essential stabilization. *Embo J*. 2005
31. Katsura I, Kuhl PW. Morphogenesis of the tail of bacteriophage lambda. II. In vitro formation and properties of phage particles with extra long tails. *Virology*. 1975; 63:238–51. [PubMed: 1089336]

Highlights

- Protein gpG & frameshift-extended gpGT are both required for phage λ tail assembly
- When the ratio of gpG to gpGT is changed from ~30:1 to ~1:30, assembly is blocked
- Graded complementation tests show that the gpG/gpGT ratio is a critical factor

Figure 2. Plasmid pETail4 produces λ tails

(a) Proteins expressed from plasmid pETail4 were radio-labeled with ^{35}S -Methionine and separated by 12.5 % SDS-PAGE and autoradiographed. The identification of bands is based on previous studies ⁸. The identity of gpJ, gpH, gpH*, gpGT, gpV, gpG, gpU and gpM was previously established. The identity of gpL, gpI, gpK and gpZ is not firmly established. gpH* is produced by assembly-dependant cleavage of gpH, and its presence is an indication of assembly ¹⁶. (b) Tails produced from plasmid pETail4 were purified using a protocol similar to one developed for purifying HK97 procapsids ²⁵. Proteins from the purified tails were separated by 12.5% SDS-PAGE and visualized using Coomassie brilliant blue R250 stain. Only 5 proteins were seen on the gel, gpJ, gpH*, gpV, gpU and gpM. (c) An electron micrograph of purified tails. The tails are uniform in length and have the same number of disks as the tails on a normal phage, indicating correct assembly.

Figure 3. Construction and effects of the GT-mut fusion mutation that produces gpGT without a frameshift

(a). An extra G was inserted right after the slippery sequence at an XmnI site (see Figure 1). The extra G allows gpGT to be expressed without a frameshift. (see (b)) The slippery sequence is underlined and the extra G inserted is boxed. Before the mutation, about 3.5% of the ribosomes change open reading frame at the slippery sequence to synthesize gpGT and the amino acid sequence at the junction point of gpGT is VSAG→KVFD (VSAG is in the G reading frame and KVFD in the T reading frame). After the insertion of a single nucleotide G right after the slippery sequence, the ribosomes can read directly into the T reading frame without a frameshift as shown in (b). The gpGT synthesized this way has exactly the same

amino acid sequence as the ones synthesized by a frameshift mechanism. (b) A -1 frameshift can create a new protein gpG* in the *GTmut* fusion mutant. As in (a), a fraction of ribosomes change open reading frame at the slippery sequence to synthesize gpG* in which the C-terminal 5 amino acids of gpG (CSTVS) are replaced with a different sequence (GVRR). (c). Comparison of proteins made by pETail4 and pETail5. An autoradiograph of a SDS-PAGE gel shows that in pETail5, gpGT is overexpressed and the gpG band disappears as expected. (d). Proteins made by plasmids with the wild-type and *GTmut* frameshift sites. Left panel. Radio-labeled proteins produced by pT7-5-GT which expresses wild type gpG and gpGT, and pT7-5-GT-mut, which contains the *GT* fusion mutation. Cells containing the appropriate plasmid were radio-labeled as described in Materials and Methods. The proteins were separated on a 12.5% SDS polyacrylamide gel and visualized by autoradiography. **Right panel.** Western blot of extracts expressing pT7-5-GT and pT7-5-GT-mut. The proteins were expressed, separated by SDS-PAGE, transferred to PVDF membrane and probed with anti-gpG antibody.

Figure 4. Protein gpG is required for tail assembly

Cells containing plasmids pT7-5-GT-mut (1), pT7-5-GT (2), pET21a-T (3), and empty vector control plasmid pT7-5 (4) were spotted with three types of phages (see Table 1), λ Gam (amber mutation in the G region), λ Tam (amber mutation in the T region) and λ -wt, (our wild type λ). The first row indicates the plasmid being tested, the next row indicates the proteins produced from the plasmid and the last column shows the type of phage being spotted and the defect in each phage. Each plate was spotted from left to right with 0.004 ml of 10-fold serial dilutions of phages, where the left-most spot has about 10^6 phages.

Panel 2(a) shows that plasmid pT7-5-GT-mut which contains the *GT* fusion mutation and produces gpGT only, clearly does not complement λ Gam which produces neither gpG nor gpGT, but panel 2(b) shows it does complement λ Tam which produces gpG but not gpGT. This suggests gpG is required for phage growth. Plasmid pT7-5-GT contains the wild type G/ GT region and produces both gpG and gpGT. Plasmid pT7-5-GT supports the growth of λ Gam very well (panel 1(a)) and λ Tam to a modest level (panel 1(a)). These results show that both gpG and gpGT have to be present to support phage growth. The fact the pET21a-T, which expresses the *T* frame fused to a T7 tag in the vector, does not complement any of the amber phages (panels 3(a) and 3(b)) indicates that the *G* part of gpGT is important for assembly too.

Figure 5. A balanced ratio of gpG and gpGT is important for tail assembly

An arabinose promoter-based system was used to study the effect of the ratio between gpG and gpGT. The genes were cloned into the pBAD33 vector (Guzman et al., 1995) which has the arabinose promoter. Three plasmids, pBAD33-GT, pBAD33-GT-mut and pBAD33-T, were tested. pBAD33-GT contains the wild type *G/GT* region and expresses both gpG and gpGT. pBAD33-GT-mut contains the *GT* fusion mutation which expresses gpGT only. pBAD33-T contains the T7-tag fused *T* frame which allows the *T* frame to be expressed independent of gpG. Each column spanning through all the three panels represent the same plasmid being tested. The amount of protein expressed from the plasmid is modulated by changing the concentration of arabinose in the plate. Six different arabinose concentration

were tested. (0 , $2 \times 10^{-6}\%$, $2 \times 10^{-5}\%$, $2 \times 10^{-4}\%$, $2 \times 10^{-3}\%$, $2 \times 10^{-2}\%$). Three types of phage were tested, λTam , λGam and $\lambda-wt$ (see Table 1).

(a). Phage λGam , which produces neither gpG nor gpGT, was tested on three different plasmids. In agreement with previous results, λGam can only be complemented with plasmid pBAD33-GT, which provides both gpG and gpGT. Neither gpGT alone nor gpT complements. The higher the concentration of arabinose, the better pBAD33-GT complements λGam . **(b). Phage λTam** which produces gpG but not gpGT, was tested on three different plasmids. Similar to the result in A, increasing the concentration of arabinose allows pBAD33-GT to complement better, but to a lower level than with λGam in A. pBAD33-GT-mut complements best at an arabinose concentration of $2 \times 10^{-4}\%$, but decreases when the concentration is raised. **(c). Phage $\lambda-wt$** was tested as a control. Note that the maximum arabinose concentration (2(f)), pBAD33-GT-mut interferes with the ability of $\lambda-wt$ to grow. A similar, but weak inhibitory effect on $\lambda-wt$ growth was seen for “gpT”, but only at the highest level of induction (3(c)).

Table 1
Plasmids and phages used in the spot complementation experiments

Plasmid	Vector	Proteins Expressed	Description
pT7-5-GT	pT7-5	gpG, gpGT	wild type <i>G/GT</i> region
pT7-5-GT-mut	pT7-5	gpGT, (gpG [*])	<i>GT</i> fusion mutation (Fig. 3)
pET21a-T	pET21a (Novagen)	gpT	T7 tag in pET21a fused to the T orf
pBAD33-GT	pBAD33 (Guzman et al., 1995)	gpG, gpGT	Wild type <i>G/GT</i> region
pBAD33-GT-mut	pBAD33	gpGT, (gpG [*])	<i>GT</i> fusion mutation (Fig. 3)
pBAD33-T	pBAD33	gpT	T7-tagged-gpT in pET21a-T transferred to pBAD33
Phage	Short name	Mutations	Phenotypes
<i>λGam447cIts1</i>	<i>λGam</i>	<i>G</i> , <i>cI</i>	no gpG, no gpGT, <i>cI</i> temperature sensitive mutation
<i>λTan901cIts1</i> [*]	<i>λTam</i>	" <i>T</i> ", <i>cI</i>	no gpGT, <i>cI</i> temperature sensitive mutation
<i>λcI857</i>	<i>λ</i> -wt	<i>cI857</i>	wild type, <i>cI</i> temperature sensitive mutation

^{*} *Tan901* appears earlier as *Gan901* (Kuhl and Katsura, 1975); we renamed it here for ease of discussion because it maps to gene *T* (Levin et al., 1993).