An ORFan No More: The Bacteriophage T4 *39.2* Gene Product, NwgI, Modulates GroEL Chaperone Function

Debbie Ang and Costa Georgopoulos[1]
Department of Biochemistry, University of Utah, Salt Lake City, Utah 84112-5650

ABSTRACT Bacteriophages are the most abundant biological entities in our biosphere, characterized by their hyperplasticity, mosaic composition, and the many unknown functions (ORFans) encoded by their immense genetic repertoire. These genes are potentially maintained by the bacteriophage to allow efficient propagation on hosts encountered in nature. To test this hypothesis, we devised a selection to identify bacteriophage-encoded gene(s) that modulate the host *Escherichia coli* GroEL/GroES chaperone machine, which is essential for the folding of certain host and bacteriophage proteins. As a result, we identified the bacteriophage RB69 gene *39.2*, of previously unknown function and showed that homologs of *39.2* in bacteriophages T4, RB43, and RB49 similarly modulate GroEL/GroES. Production of wild-type bacteriophage T4 Gp39.2, a 58-amino-acid protein, (a) enables diverse bacteriophages to plaque on the otherwise nonpermissive *groES* or *groEL* mutant hosts in an allele-specific manner, (b) suppresses the temperature-sensitive phenotype of both *groES* and *groEL* mutants, (c) suppresses the defective UV-induced PolV function (UmuCD) of the *groEL*44 mutant, and (d) is lethal to the host when overproduced. Finally, as proof of principle that Gp39.2 is essential for bacteriophage growth on certain bacterial hosts, we constructed a T4 *39.2* deletion strain and showed that, unlike the isogenic wild-type parent, it is incapable of propagating on certain *groEL* mutant hosts. We propose a model of how Gp39.2 modulates GroES/GroEL function.

DURING the 1940s, Alfred Hershey's pioneering experiments led to the isolation and characterization of bacteriophage T2 rapid lysis mutants (Hershey 1946; Hershey and Rotman 1948). Seymour Benzer exploited the corresponding bacteriophage T4 *rII* class of mutants (Lederberg and Lederberg 1953) in his classic experiments on genetic fine structure through DNA recombination, thus redefining our concept of the gene, the cistron, and the various distinct classes of mutations (Benzer 1955, 1959). His work clearly showed that T4 possesses two genes, *rIIA* and *rIIB*, which are essential for growth on bacteriophage λ lysogens, but not on nonlysogens. Shortly thereafter, Dick Epstein, on the basis of his Ph.D. thesis work with UV-inactivation curves of T4 on *E. coli* B and K-12 (λ), hypothesized that, in an analogous manner, T4 may possess gene(s) essential for growth on *Escherichia coli* B but not K-12 (λ) (Epstein 1958a,b). The experiment testing this hypothesis led to the classic discovery of the amber (nonsense) mutants (Epstein *et al.* 1964; Stahl 1995; Epstein *et al.* 2012), also independently discovered by Allan Campbell with bacteriophage λ (Campbell 1961). Although Dick Epstein's original idea that bacteriophage T4 possesses genes that specifically enable it to grow on *E. coli* B but not *E. coli* K-12 (λ) is thus far unsubstantiated, recent genetic analyses coupled with rapid DNA sequencing techniques have indeed shown that even closely related bacteriophages carry their own sets of unique genes that most likely favor their growth on certain bacterial hosts in nature (see below).

Recently there has been renewed interest in the vast number and variety of bacteriophages found in the environment (reviewed in Wommack and Colwell 2000; Hendrix 2002; Rohwer and Edwards 2002; Casjens 2008). Studies of related bacteriophages show that while genes encoding morphological components such as capsid proteins and tail fibers tend to be conserved, the rest of the genome displays remarkable diversity. One of the most interesting findings is the existence of novel open reading frames (referred to as ORFans; Fischer and Eisenberg 1999) whose functions are

doi: 10.1534/genetics.111.135640
Manuscript received October 9, 2011; accepted for publication November 7, 2011
Available freely online through the author-supported open access option.
This article is dedicated to the memory of our dear friend and colleague Richard (Dick) Epstein (March 14, 1929–March 6, 2011), the discoverer of the bacteriophage T4 amber mutations and an inspiration to generations of geneticists.
[1]Corresponding author: Department of Biochemistry, University of Utah, 15 N. Medical Dr. East, Rm. 4100, Salt Lake City, UT 84112-5650. E-mail: costa@biochem.utah.edu

unknown and that have no significant homology to other sequences currently found in databases (Nolan *et al.* 2006; Comeau *et al.* 2007, 2008, 2010; Zuber *et al.* 2007; Hatfull 2008; Denou *et al.* 2009; Petrov *et al.* 2010). Whereas many of these ORFan genes are nonessential under standard laboratory conditions, their occasional presence in related bacteriophages suggests that they are important for survival in the wild. The difficulty lies in determining the functions of these seemingly nonessential genes, whose effects may be subtle or observable only under specific conditions.

Because each bacteriophage relies heavily on its bacterial host(s) for propagation, studying this relationship has offered many insights into fundamental bacteriophage and host functions. For example, the GroEL/GroES chaperone machine of *E. coli* was originally discovered through host mutations that block the propagation of certain bacteriophages (reviewed in Friedman *et al.* 1984; Ang *et al.* 2000; Georgopoulos 2006). Subsequent work showed that the *groEL* and *groES* genes form an operon and are coordinately expressed under normal housekeeping conditions. Their expression is increased under conditions of stress such as heat shock (reviewed in Zeilstra-Ryalls *et al.* 1991; Georgopoulos 2006). The importance of the GroES/GroEL chaperone machine is exemplified by the fact that it is almost universally conserved and highly homologous to the essential Hsp10/Hsp60 proteins residing in chloroplasts and our own mitochondria (Horwich and Fenton 2009; Hartl *et al.* 2011). Furthermore, it is the only chaperone machine assisting the correct folding of certain essential proteins in *E. coli* (Houry *et al.* 1999; Kerner *et al.* 2005), thus is itself absolutely essential for life (Fayet *et al.* 1989).

Electron microscopy and X-ray crystallography showed that GroEL's native structure is a tetradecamer composed of two back-to-back heptameric rings (reviewed in Sigler *et al.* 1998). The much smaller GroES cochaperone is a heptamer that binds to one or both ends of GroEL through its mobile loops (Landry *et al.* 1993; Sigler *et al.* 1998). GroEL captures unfolded substrates via the hydrophobic surface of the apical domains of one of its rings (referred to as the *cis*-ring). Subsequent ATP binding at the *cis*-ring of GroEL leads to massive *en bloc* rearrangements of the GroEL subunits, leading to GroES binding and release of the substrate into the GroEL–GroES cavity, whose interior is now lined with mostly hydrophilic residues (Xu *et al.* 1997; Sigler *et al.* 1998). Some of the apical domain hydrophobic groups now exposed on the GroEL surface are bound by the seven mobile loops of the GroES cochaperone. This GroEL–cochaperone interaction is the result of a delicate balance between disorder and order of the cochaperone mobile loops (Landry *et al.* 1993, 1996; Richardson and Georgopoulos 1999; Richardson *et al.* 1999; Shewmaker *et al.* 2001, 2004). When bound to GroEL, the disordered cochaperone mobile loops become structurally ordered, although physical interaction with GroEL occurs only through a universally conserved tripeptide in the mobile loop (Landry *et al.* 1993, 1996; Xu *et al.* 1997). Earlier work showed that all of our *groES* mutations alter residues in the mobile loop, resulting in either weakened or strengthened interactions with GroEL (Landry *et al.* 1993, 1996; Zeilstra-Ryalls *et al.* 1996; Richardson *et al.* 1999; Richardson 2000; Shewmaker *et al.* 2004). The as-yet unfolded released substrate in the cavity has approximately 10 sec, the average time required for ATP hydrolysis, to reach a folding-competent conformation. ATP hydrolysis results in binding of ATP/substrate to the *trans*-ring, thus releasing GroES and the folded substrate from the *cis*-ring (Sigler *et al.* 1998; Horwich and Fenton 2009; Hartl *et al.* 2011).

Interestingly, whereas the temperate bacteriophage λ requires both the *groES* and *groEL* gene products for the correct assembly of its prohead, bacteriophage T4 does not require GroES because it encodes its own cochaperone called Gp31, which is distantly related to GroES (<10% sequence identity) (Nivinskas and Black 1988; Keppel *et al.* 1990; Koonin and Van Der Vies 1995). Many of the so-called T4-like bacteriophages encode Gp31 homologs, some of which are capable of substituting for the *E. coli* GroES protein in host function (Ang *et al.* 2000; Ang *et al.* 2001; Keppel *et al.* 2002). Despite the lack of sequence similarity, both GroES and Gp31 assemble into very similar structures (Hunt *et al.* 1996; Hunt *et al.* 1997). However, the cavity found in the GroEL/Gp31 complex is larger than that formed with GroES (Hunt *et al.* 1996, 1997; Landry *et al.* 1996; Clare *et al.* 2009). The observed differences may have evolved to accommodate the major bacteriophage coat protein Gp23, which, at approximately 56,000 Da, is at the upper limit for the cavity of the GroEL–GroES complex. As opposed to other GroEL-bound substrates, Gp23 seems to occupy a unique position under the GroEL–Gp31 dome (Clare *et al.* 2009). Furthermore, Snyder and Tarkowski (2005) have presented evidence suggesting that the nascent N-terminal region of Gp23 targets it directly to GroEL, thus potentially speeding up the rate of folding during bacteriophage T4 intracellular growth.

To better understand the relationship between T4 and its host *E. coli*, we have used a direct genetic selection to identify additional bacteriophage genes that regulate GroE function. By using various *E. coli groEL* and *groES* mutants to screen a bacteriophage λ DNA library of the T4-like bacteriophage RB69, we identified the product of gene *39.2* as a specific modulator of GroE function. By testing its effect on a variety of host and bacteriophage mutants *in vivo* we conclude that the ability of the Gp39.2 protein to suppress GroE defects is limited to situations in which either the GroEL or its cochaperone mutant exhibits a weakened affinity for the other partner. Finally, we show that deletion of the *39.2* gene from the T4 genome inhibits its growth on certain *E. coli* hosts, potentially justifying its retention in the genome.

Materials and Methods

Strains and plasmids

See Table 1 for a list of strains, bacteriophages, and plasmids used in this study. Bacteriophages RB69, RB49, and RB43 were kind gifts of Dr. Henry Krisch (Toulouse). Bacteriophage T4 K10 and plasmid pBSPLO (−) were kind gifts of Dr. Ken Kreuzer (Duke University). Cells were grown in L-broth (10 g

Table 1 Strains, bacteriophages and plasmids

Strains	Genotype or description	Reference or source
B178	K-12 W3110 *galE sup*°	Our collection
C600	K-12 *thr-1 leuB*6 *lacY*1 *supE*44 *rfbD*1 *thi-1 tonA*21 F-	Our collection
DH10B	K-12 F– *mcrA* Δ(*mrr-hsdRMS-mcr*BC) Φ80*lacZ*ΔM15 Δ*lacX*74 *recA*1 *endA*1 *araD*139 Δ(*ara leu*)7697 *galU galK rpsL nupG*; used for cloning purposes	Invitrogen
CG217	B178 *groES*30 (A31V)	Landry *et al.* (1993)
CG3009	B178 *groEL*515 (A383T)	Zeilstra-Ryalls *et al.* (1993)
CG3010	B178 *groEL*44 (E191G)	Zeilstra-Ryalls *et al.* (1993)
CG3012	B178 *groEL*140 (S201F)	Zeilstra-Ryalls *et al.* (1993)
CG3013	B178 *groES*619 (G24D)	Zeilstra-Ryalls *et al.* (1993)
CG3015	B178 *groEL*673 (G173D)	Zeilstra-Ryalls *et al.* (1993)
CG3017	B178 *groEL*173 (Y199C)	Revel *et al.* (1980); Richardson (2000)
CG3388	B178 *groEL*3388 (E191G V190I)	Klein and Georgopoulos (2001)
DA1415	B178 *groES*42 (G23D)	Landry *et al.* (1993)
Bacteriophages		
λZAP II	Accepts 0–10 kb inserts, which can be excised *in vivo* via the pBluescript phagemid	Stratagene (Agilent)
λ*b*2cI	Clear-plaque former	Our collection
T4Do	Wild-type T4	Our collection
T4 K10	*38amB262 51amS29 denAnd28 denB- rIIB ΔrIIPT8*	Selick *et al.* (1988)
RB69	T4-like bacteriophage	J. Karam
RB49	T4-like bacteriophage	H. Krisch
RB43	T4-like bacteriophage	H. Krisch
K3	T4-like bacteriophage	H. Krisch
OX2	T4-like bacteriophage	H. Krisch
T2	T4-like bacteriophage	Our collection
T5	T4-like bacteriophage	Our collection
T6	T4-like bacteriophage	Our collection
T4 Δ*39.2*	T4Do carrying deletion of *39.2* gene	This work
Plasmids		
pAA	pBSPLO (−) carrying regions flanking T4 *Δ39.2*	This work; Selick *et al.* (1988)
pAC	pMPMK4–*39.2*^{T4} Kan^{R}, high copy	This work; Mayer (1995)
pDA1844	pAC–*39.2* ^{T4} (C13A)	This work
pDA1846	pAC–*39.2* ^{T4} (C31A C34A)	This work
pAQ	pMPMK4–*39.2*^{RB49} Kan^{R}	This work; Mayer (1995)

tryptone, 5 g yeast extract, 5 g NaCl per liter, pH 7) or on L-plates (L-broth with 10 g agar/liter) with antibiotics where appropriate (final concentrations 100 μg/ml ampicillin, 50 μg/ml kanamycin, 10 μg/ml chloramphenicol).

Isolation of λZAP II clones with the ability to propagate on groE mutants

A λZAP II library of RB69 DNA (prepared by Stratagene, a gift from Jim Karam, Tulane University) was used to infect fresh overnight cultures of *E. coli* B178 *groEL*44, *groEL*515, and *groES*42 at a multiplicity of infection (m.o.i.) of 0.01. Following incubation at room temperature for adsorption, the infected cells were mixed with 3 ml molten soft agar (L-broth with 0.6% agar), poured onto fresh L-plates, and incubated at 37° overnight. The next day, plaques were purified according to standard procedures, followed by preparation of stock lysates. Growth on the corresponding mutant host was confirmed prior to further characterization.

Identification of the RB69 39.2 open reading frame

To identify the region(s) of the RB69 genome represented by the λZAP II clones, DNA was isolated from each bacteriophage candidate by freeze–thaw, *i.e.*, 95° for 10 min, −20° for 20 min, 95° for 10 min, and then used as a template in a standard PCR reaction according to the manufacturer's protocol (Pfx; GibcoBRL/Invitrogen; primers forward: *5′ ccc agt cac gac gtt gta aaa cg* and reverse: *5′ agc gga taa caa ttt cac aca gg*). Gel-purified PCR products (Qiagen) were sequenced from the ends using forward and reverse primers, followed by BLASTn comparison (Altschul and Lipman 1990) to identify the genomic region represented by each clone. Subcloning of individual open reading frames into the λZAP II vector allowed us to identify the gene of interest.

Cloning and mutagenesis of T4 39.2

The 430-bp region encompassing the T4 *39.2* open reading frame was initially PCR amplified from genomic DNA using primers derived from the published sequences of Huang (1986) and Sanson and Uzan (1992): *5′ gca gga aga gaa aga gca aga aca aat ag* and *5′ gaa cca att ata tct aga gat ttc aat att ag*. Following sequence verification, the minimal gene was then PCR amplified from genomic DNA and cloned into vector pMPMK4 (Mayer 1995) using restriction sites *Eco*RI and *Pst*I, thus placing the gene under the regulation of the P_{BAD} arabinose-inducible promoter (Guzman *et al.* 1995; Mayer 1995). The primers used to amplify the minimal T4

39.2 gene were: *5′ gtt gag gaa ttc acc atg ccg ctt tat gat* and *5′ ggt ggt ctg cag tca tta ccc ttt aag caa gtc gta*. The resulting plasmid was named pAC.

Site-directed mutagenesis of the C13 and the C31, C34 residue codons in T4 *39.2* was accomplished using the QuikChange method (Stratagene) and pAC as template. Primers *5′ caa tcc aaa gac gct gca aaa gaa tac g* and *5′ cgt att ctt ttg cag cgt ctt tgg att g* were used for the C13A mutation. Primers *5′ ctg aaa gag ata ctg atg tag ctc ctg atg ctc atc ggc tgg ctg ttc* and *5′ gaa cag cca gcc gat gag cat cag gag cta cat cag tat ctc ttt cag* were used for the C31A C34A mutations. The resulting plasmids were named pDA1844 and pDA1846, respectively.

Homologs of 39.2 in other T4-like bacteriophages

To test if other T4-like bacteriophages encode *39.2*, DNA from T4, RB49, RB43, T2, T6, OX2 and K3 was PCR-amplified using primers for the minimal T4 *39.2* gene: *5′ gct tac ttt aag gaa tac ata tgc cgc ttt atg at* and *5′ ggt ggt gga tcc tta ccc ttt aag caa gtc gta*. Products were processed and sequenced as before. The RB49 and RB43 *39.2* homologs were PCR-amplified using primer pairs *5′ gtt gag gaa ttc acc atg cca tta tac gat tat* and *5′ ggt ggt ctg cag gac tca tga cac gcg cac*, and *5′ gct tac tct cac gaa tac ata tgc caa cat ata cc* and *5′gat cgt aga cgt tct aga tat ttt aga cat agg g*, respectively. The RB49 *39.2* gene was cloned into pMPMK4 at the *Eco*RI and *Pst*I sites, yielding plasmid pAQ.

Construction of a T4 39.2 deletion mutant

A *39.2* null mutation in bacteriophage T4 was constructed using the method of Selick *et al.* (1988). Essentially, the regions flanking *39.2* in T4 were amplified in two separate reactions according to the manufacturer's protocol (Pfx; Invitrogen), using primer pairs #1, *5′ ctg ata acg gat cct atg cac* and #2, *5′ tat att tca taa tta cga gct ccc ggg cat atg aat tat tcc tta aag* and #3, *5′ ctt taa gga ata att cat atg ccc ggg agc tcg taa tta tga aat ata,* and #4, *5′ gtt tcg cta gca gaa gag cca at*. The two gel-purified products, which share homology in the region spanning the deletion of *39.2*, were fused in a single standard PCR reaction (Pfx; Invitrogen) using primers #1 and #4. The product was digested with *Bam*HI and *Nhe*I, and cloned into pBSPLO (−), digested with *Bam*HI and *Xba*I, resulting in plasmid pAA. Strain C600 (*supE*44), transformed with pAA, was infected with bacteriophage T4 K10, which carries amber mutations in two essential genes, *38* and *51* (Selick *et al.* 1988). The resulting lysate was plated on lawns of *E. coli* B178 (*sup*°) to select integrants, *i.e.*, bacteriophage that have integrated the plasmid by homologous recombination at the site of *39.2*. [pBSPLO (−) encodes *supF*, which suppresses the amber mutations in T4 K10 in a *sup*° host.] Candidates were purified on B178 lawns to isolate single plaques. To recover plasmid-free bacteriophage, single plaques were resuspended in L-broth and used to infect C600 bacteria. The resulting lysates were plated on C600 lawns. Plaques were screened on B178 and C600 lawns to find those that had lost the plasmid (growth only on C600). Candidates were further screened by PCR for loss of the *39.2* allele. Four out of 17 sequenced candidates had lost the *39.2* gene. One of these T4 K10 Δ*39.2* bacteriophages was extensively backcrossed to wild-type T4Do, continuously selecting for loss of the amber mutations and screening by PCR for the simultaneous loss of the *39.2* gene to arrive at the isogenic T4 Δ*39.2* bacteriophage used in this work.

Selection of rifamycin-resistant mutants

To determine the frequency of rifamycin-resistant mutants in UV-irradiated cultures, we essentially followed the protocol of Donnelly and Walker (1989). Overnight cultures were diluted 1:100 into LB and incubated at 37° with aeration. At OD_{595nm} 0.2, arabinose was added to the appropriate cultures to a final concentration of 0.02%. Incubation at 37° was continued for all cultures until they reached approximately OD_{595nm} 0.8. Cells were centrifuged and resuspended in an equal volume of 10 mM $MgSO_4$. For each time point, a 1-ml aliquot of cells was placed on the sterile surface of a plastic petri dish lid and covered with aluminum foil. The lid was positioned under a prewarmed UV lamp, the foil was removed, and counting of time began. Cells were transferred to 1.5-ml microfuge tubes on ice and kept covered with foil whenever possible to prevent photoreactivation. Dilutions were spread on L-plates to determine rates of survival. In addition, 0.25 ml of each time point was diluted into 5 ml L-broth to allow outgrowth of rifamycin-resistant mutants. The tubes were covered with aluminum foil and incubated at 37° overnight with aeration. Dilutions were spread on L-plates with or without rifamycin (100 μg/ml) and incubated overnight at 37°. Colonies were counted the next day to determine the frequency of rifamycin-resistant mutants.

Results

Selection of RB69 clones that allow bacteriophage growth on E. coli groE mutant strains

E. coli groES and *groEL* mutants were first identified by their resistance to either bacteriophage λ or T4 infection (reviewed in Friedman *et al.* 1984; Zeilstra-Ryalls *et al.* 1991; Ang *et al.* 2000; Georgopoulos 2006). Later work showed that, whereas both GroEL and GroES are required for λ head morphogenesis, only GroEL is required for the folding of bacteriophage T4's major capsid protein Gp23. GroES is not required because T4 encodes its own distant cochaperone ortholog named Gp31 (Nivinskas and Black 1988; Keppel *et al.* 1990; Van Der Vies *et al.* 1994). Our interest in homologs of T4 Gp31 led us to the λZAP II library carrying 4- to 10-kbp fragments of genomic DNA isolated from the T4-like bacteriophage RB69, which we used to infect lawns of *E. coli groEL*44, *groEL*515, and *groES*42 mutants. From our previous studies, we knew that T4 Gp31 expressed from a plasmid allows bacteriophage λ to propagate on our *groES* mutants, as well as the *groEL*515 mutant (Keppel *et al.* 2002). Thus, we expected to identify the RB69 homolog of the T4 *31* cochaperone gene by infecting the *groES*42 and *groEL*515 mutants with the library and isolating plaque formers.

We had shown earlier that bacteriophage T4 growth is blocked on the *groEL*44 mutant because its Gp31 product does not interact with the mutant GroEL44 protein (Shewmaker *et al.* 2004). As such, we used *E. coli groEL*44 as a negative control here, under the assumption that the putative RB69 Gp31 homolog likewise would not interact with GroEL44.

Consistent with our prediction, on *E. coli groES*42 lawns we isolated λZAP II plaque formers from the RB69 DNA library at the expected frequency of 3×10^{-3} per plaque former on wild-type *E. coli* (since the RB69 genome is only 167,560 bp). As expected, the λZAP II vector alone yielded no plaques on *E. coli groES*42 ($<10^{-8}$).

Our first surprise was that no plaques were found on *E. coli groEL*515. We had anticipated the putative λZAP II-*31*^{RB69} recombinant bacteriophage to grow on this host since bacteriophage λ propagates on *E. coli groEL*515 cells expressing Gp31^{T4} from a plasmid (Keppel *et al.* 2002). Our surprise continued when we unexpectedly found a high frequency of plaque formers (3×10^{-3}) on *E. coli groEL*44 lawns, similar to the frequency found on *E. coli groES*42. Subsequent plaque purification and retesting showed that these two groups of λZAP II recombinant bacteriophage behave similarly; *i.e.*, all members of the two groups grow on both *E. coli groES*42 and *E. coli groEL*44, but none grow on *E. coli groEL*515. In addition, these λZAP II recombinant plaque formers grow with an efficiency of 1.0 on our *E. coli groES*30 and *groES*619 mutants, but do not grow on the *E. coli groEL*140, *groEL*673, or *groEL*173 mutants ($<10^{-8}$) (Table 2). Taken together, these observations suggested that, instead of isolating the putative *31*^{RB69} cochaperone gene, we had identified a novel RB69 gene that modulates the host GroEL/GroES chaperone in an allele-specific manner.

Identification of the gene responsible for bacteriophage growth on E. coli groE strains

We proceeded to identify the putative RB69 gene(s) responsible for the seemingly allele-specific plating pattern of our λZAP II recombinant candidates. To determine the minimal region(s) of the RB69 genome represented by these clones, candidate inserts were sequenced from their ends (see *Materials and Methods*). Comparison with the highly homologous T4 genome indicated that all candidates carried DNA from the region defined by the genes *goF* (aka *comC*-α), a putative antitermination factor (Takahashi and Yoshikawa 1979; Stitt *et al.* 1980), and gene *39*, a component of T4 DNA topoisomerase (Huang 1986), located at approximately kbp 6 on the physical RB69 map (Figure 1). While most open reading frames in this region have homologs in T4, interestingly, and consistent with the genome plasticity of the bacteriophage T4 family, there is a unique gene, *orf*005c, that has been substituted for the T4 *39.1* gene. A search of the T4 genomic DNA sequence showed that no homolog of *orf*005c exists, while a similar search of the RB69 genome indicated that it likewise has no homolog of *39.1*.

Next, we reconstructed a λZAP II recombinant bacteriophage encompassing the common overlap among our λZAP II candidates, confirming that this ~3100-bp RB69 DNA fragment is indeed responsible for allowing bacteriophage growth on the *groEL*44 and *groES*42 mutants. Subcloning specific RB69 open reading frames into the λZAP II vector allowed us to identify the 59-amino-acid-coding *39.2* gene as solely responsible for the suppression phenotype (data not shown).

Table 2 Putative RB69 gene allows λ growth only on those *E. coli groE* mutants that result in a low-affinity GroEL/GroES interaction

Bacterial host	λZAP II vector	λZAP II-RB69 recombinant[a]
B178	+[b]	+
*groES*42 (G23D)[c]	−[d]	+
*groES*619 (G24D)	−	+
*groES*30 (A31V)	−	+
*groEL*44 (E191G)	−	+
*groEL*140 (S201F)	−	−
*groEL*515 (A383T)	−	−
*groEL*673 (G173D)	−	−
*groEL*173 (Y199C)	−	−

[a] The λZAP II-RB69 recombinant was selected for growth on the *E. coli groEL*44 mutant host. It is representative of all candidates isolated on either *E. coli groEL*44 or *groES*42.
[b] Growth of bacteriophage at an efficiency of 1.0.
[c] The amino acid substitution resulting from each *groE* mutation is indicated in the parentheses following the allele designation.
[d] No growth of bacteriophage ($<10^{-8}$).

Bacteriophage homologs of 39.2

Because T4 is the prototype for this family of bacteriophages, we PCR amplified and sequenced the equivalent region in T4. Near the end of the *39.2* gene, we noted several discrepancies with the published sequence of Huang (1986), which predicts a protein of 45 amino acid residues. Correction of these differences results in a predicted Gp39.2^{T4} protein more similar in sequence and length (58 amino acid residues) to that of Gp39.2^{RB69}, with a homology of 91% and identity of 82% between the two (BLASTp). This is accurately reflected in the T4T sequence found at the GGC T4-like genome website (http://phage.ggc.edu/).

Using gene *39.2*^{T4}-specific PCR primers, we amplified and sequenced the corresponding *39.2* homologs of the closely related bacteriophages T2, T6, K3, and OX2. As seen, these protein sequences are highly conserved with an overall pairwise identity above 80% (Figure 2). A striking feature in the alignment is the conservation of four cysteine residues (highlighted in red), typical of a zinc-binding motif. Subsequently, the DNA sequences of the more distantly related pseudo-T-even bacteriophages RB43 and RB49 were completed and made available on the GGC T4-like genome website. Although both bacteriophages RB43 and RB49 possess Gp39.2-coding genes, the auxiliary genes in their immediate neighborhood are by and large different from those shown in Figure 1 for T4 and RB69. Overall, the predicted Gp39.2^{RB43} and Gp39.2^{RB49} sequences are 30–40% identical to that of Gp39.2^{T4}, as well as to each other (Figure 2). Interestingly, Gp39.2^{RB49} has only three of the highly conserved cysteine residues. However, the putative zinc-binding motif is retained since a histidine residue replaces the second cysteine (shown in red and indicated by an arrow at the bottom of Figure 2).

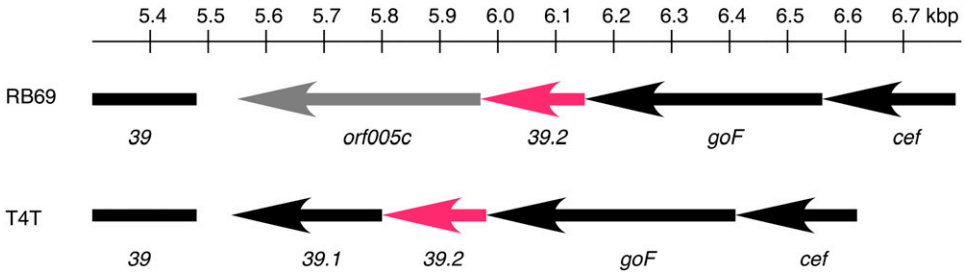

Figure 1 Comparison of the RB69 and T4 *g39-cef* regions. Arrows indicate open reading frames and their direction of transcription. *goF* is also known as *comC-α* (Takahashi and Yoshikawa 1979; Sanson and Uzan 1992). The RB69 *orf005c* gene is not a homolog of T4T *39.1*. Only the beginning of gene *39* is shown. The scale above refers to the physical RB69 genome. The figure was drawn using information available on the GGC T4-like genome website at http://phage.ggc.edu/.

Despite their relatively limited identity with Gp39.2^{T4}, both Gp39.2^{RB43} and Gp39.2^{RB49} also suppress various GroE mutant phenotypes (see below and data not shown).

Construction of wild-type and mutant 39.2 plasmids

The minimal *39.2* genes of bacteriophages RB69, T4, and RB49 were PCR amplified and cloned under the tight regulation of the arabinose-inducible promoter P_{BAD} (Guzman *et al.* 1995; Mayer 1995). Preliminary random PCR mutagenesis experiments with the minimal *39.2*^{T4} gene suggested that a mutation in one or more of the highly conserved cysteine residues (shown in Figure 2) abolishes Gp39.2 biological function. However, because these mutants had multiple changes, we constructed two *39.2* mutant clones by site-directed mutagenesis of plasmid pAC, carrying the minimal *39.2*^{T4} gene. Changing the C13 codon to an alanine codon resulted in plasmid pDA1844. Similarly, changing the C31 and C34 codons to alanine codons resulted in a double mutant in plasmid pDA1846. Upon induction by arabinose, both mutant plasmids express levels of Gp39.2 comparable to the wild-type clone (data not shown), although they completely lack activity in the biological assays shown below.

Suppression of bacteriophage resistance by Gp39.2

E. coli wild-type, *groEL*44, *groEL*515, and *groES*42 cells were transformed with the plasmids expressing wild-type and mutant Gp39.2^{T4}, as well as wild-type Gp39.2^{RB69} and Gp39.2^{RB49} to test the effect of Gp39.2 on *groE* mutant phenotypes. We first tested bacteriophage growth by spotting serial dilutions of a representative set of bacteriophages, namely λ, T5, RB43, RB49, RB69, and T4, whose growth is blocked on the *groEL*44 mutant. As a control, we used bacteriophage T4 *31* (L35I), whose mutant Gp31 (L35I) protein has a higher affinity for GroEL44, enabling the bacteriophage to grow normally on *groEL*44 bacteria (Keppel *et al.* 1990; Van Der Vies *et al.* 1994; Richardson and Georgopoulos 1999; Richardson *et al.* 1999). Clearly, only wild-type Gp39.2^{T4} enables these bacteriophages to propagate on *groEL*44 bacteria (Figure 3A), although both the wild-type and C13A mutant Gp39.2 proteins are expressed at comparable levels. Analogous results were observed in certain other mutant *groE* strains and with plasmids expressing wild-type Gp39.2^{RB69}, Gp39.2^{RB43}, and Gp39.2^{RB49} (data not shown). As expected from our original selection with the λZAP II-RB69 recombinant clones and the results shown in Table 2, expression of wild-type Gp39.2^{T4} did not alter the bacteriophage plating properties of *groEL*515 mutant bacteria.

Suppression of temperature sensitivity by Gp39.2

The *groEL*44 strain does not grow at 42°. Thus, we asked whether expression of wild-type Gp39.2^{T4} and Gp39.2^{RB49} can allow its growth at the nonpermissive temperature. As

```
OX2       MPLYDYKCQSKDCAKEYEKIKKISERDTDVCPDCHRIAIRLVSAPKHVNGGFYDLLKG- 58
T6        MPLYDYKCQSKDCAKEYEKIKKISERDTDVCPDCHRIAIRLVSAPKHVNGGFYDLLKG- 58
T4        MPLYDYKCQSKDCAKEYEKIKKISERDTDVCPDCHRLAVRLVSAPKHVNGGFYDLLKG- 58
K3        MPMYDYKCQSEDCGHEYEKIKKISERENDVCPKCHRLSTRRPSAPKHVNGGFYDLLKG- 58
T2        MPMYDYKCQSEDCGHEYEKIKKISERENDVCPKCHRLSTRRPSAPKHVNGGFYDLLKG- 58
RB69      MPLYDYKCQSEDCGHEYEKIKRISERENDVCPECHRLSTRRVSAPKHVNGGFYDLLKKG 59
          **:*******:**.:******:****:.****.***:: *  ***************

T4        MPLYDYKCQSKDCAKEYEKIKKISERDTDVCPDCHRLAVRLVSAPKHVNGGFYDLLKG 58
RB43      MPTYTYKCT--KCEKEFSKMSSIATRDAPKDCECGGEAHRTVDKPQMVKGGFPYV--- 53
          ** * ***   .* **:.*:..*: **:    :*   * * *. *: *:***  :

T4        MPLYDYKCQSKDCAKEYEKIKKISERDTDVCPDCHRLAVRLVSAPKHVNGGFYDLLKG-- 58
RB49      MPLYDYSCEEHG---QFQKIRKISEREKAECPECGKECKQGLTAPRGINGGYYDKNVRVS 57
          ******.*:.:.   :::**:*****:.  **:* : . : ::**: :***:**

RB43      MPTYTYKCTKCEKEFSKMSSIATRDAPKDC-ECGGEAHRTVDKPQMVKGGFPYV------ 53
RB49      MPLYDYSCEEHG-QFQKIRKISEREK-AECPECGKECKQGLTAPRGINGGY-YDKNVRVS 57
          ** * *.* :   :*.*: .*: *:   :* *** *.:: :  *: ::**: *
```

Figure 2 ClustalW comparisons of predicted Gp39.2 protein sequences of T4-like bacteriophages. Conserved putative zinc finger residues are indicated in red. A recent search of 39.2 sequences in the GenBank Database suggests that Gp39.2^{RB49} is so far unique in that it has a histidine rather than cysteine at the position indicated by the arrow. Asterisks indicate identity, colons indicate conservation among groups with strongly similar properties, and periods indicate conservation among groups with weakly similar properties (Larkin *et al.* 2007).

seen in Figure 3B, expression of wild-type Gp39.2^{T4} and Gp39.2^{RB49}, but not Gp39.2^{T4} (C13A) or Gp39.2^{T4} (C31A C34A), partially restores growth of *groEL*44 mutant bacteria at 42°. Similarly, wild-type Gp39.2^{RB69} and Gp39.2^{RB43} also partially suppress *groEL*44 growth at 42° (data not shown). It may be that Gp39.2 enables the mutant GroEL44 to fold, at least partially, all of its essential bacterial substrates at 42°.

Overexpression of Gp39.2 is toxic to E. coli growth

We observed that expression of Gp39.2 can be toxic to *E. coli* growth (Figure 4). By varying the amount of arabinose inducer present in the medium, we manipulated the intracellular Gp39.2 levels to show that toxicity depends on the *groE* genetic background. Specifically, we found that *groEL*515 bacteria are more sensitive to Gp39.2 levels than wild-type bacteria, which in turn are more sensitive than *groEL*44 bacteria. However, at very high levels, Gp39.2 is toxic even to *groEL*44 bacteria (data not shown). Thus, it appears that Gp39.2-mediated toxicity is directly or indirectly related to the function of the GroEL/GroES chaperone machine.

Gp39.2 expression enables folding of the UmuC protein

As stated earlier, the GroEL/GroES chaperone machine is absolutely essential for *E. coli* viability because it is the only one that folds certain essential *E. coli* proteins (Houry *et al.* 1999; Kerner *et al.* 2005). Employing the observations of Donnelly and Walker (1989, 1992), we endeavored to demonstrate directly that Gp39.2 can play a role in the folding of a specific *E. coli* protein, UmuC. The UmuC and UmuD' (a cleaved product of UmuD) proteins make up DNA polymerase V (UmuD'$_2$C; PolV), which is required for SOS-induced mutagenesis in *E. coli*. Specifically, Tang *et al.* (1999) showed that PolV is highly "error prone," capable of bypassing abasic lesions on UV-damaged DNA. Donnelly and Walker (1989; 1992) showed that the GroEL/GroES chaperone machine is required for PolV function and that GroEL interacts directly with UmuC. Petit *et al.* (1994) further demonstrated in a purified protein system that GroEL/GroES enables UmuC to fold properly, leading to assembly of an active PolV enzyme.

Using Donnelly and Walker's procedure (1989), we showed that indeed *E. coli groEL*44 mutant bacteria are deficient in UV-induced mutagenesis by monitoring the appearance of *E. coli* rifamycin-resistant (Rif^R) mutants following UV irradiation (data not shown). We then transformed the *groEL*44 bacteria with plasmid pAC encoding the wild-type *39.2*^{T4} gene, grew the resulting transformants in the presence or absence of arabinose, and followed the appearance of Rif^R mutants as a function of UV dose. As shown in Figure 5, expression of Gp39.2^{T4} substantially increases the appearance of Rif^R mutants following UV mutagenesis. We conclude that Gp39.2 directly or indirectly enables the mutant GroEL44 chaperone to help UmuC reach its native active state.

Proof of principle: Gp39.2 is essential for T4 survival on certain hosts

To determine if the seemingly nonessential *39.2* gene of T4 is in fact essential for bacteriophage growth on certain hosts,

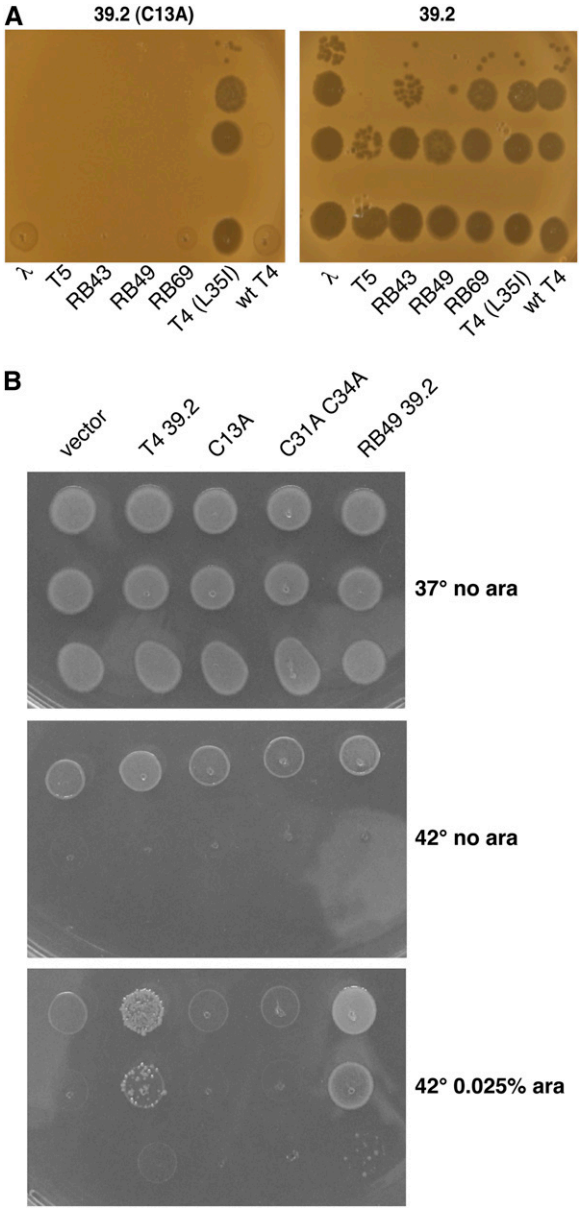

Figure 3 Wild-type but not mutant Gp39.2 suppresses various GroE mutant phenotypes. (A) Suppression of bacteriophage resistance in *groE* mutants. Dilutions (10-fold) of GroE-dependent bacteriophages were spotted on lawns of *E. coli groEL*44 bacteria transformed with either pDA1844 (*39.2* C13A) or plasmid pAC (*39.2*). The medium contained 0.01% arabinose and the plates were incubated overnight at 37°. (B) Suppression of the GroE temperature-sensitive phenotype by *39.2* expression. Dilutions (100-fold) of *E. coli groEL*44 transformed with pMPMK4 vector, pAC, pDA1844, pDA1846, or pAQ plasmids, respectively, were spotted on L-plates with or without 0.025% arabinose as indicated and incubated at 37° or 42° overnight. The nature of the plasmid-encoded protein is indicated along the top of the figures.

we constructed an isogenic T4 strain with an in-frame deletion of *39.2*, using the insertion/substitution system of Selick *et al.* (1988). The T4 Δ*39.2* mutant and its isogenic wild-type parent were tested on our standard wild-type *E. coli* K-12 host at various temperatures, and observed to plate equally well (Figure 6, A and B). When tested on our *groE* mutant hosts, the

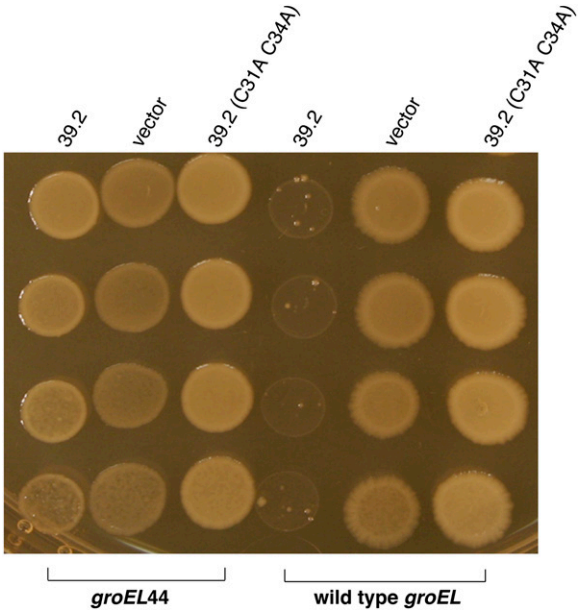

Figure 4 Expression of Gp39.2 is toxic for wild-type *E. coli*. The B178 *groEL*44 strain and its isogenic wild-type parent were transformed with plasmids pAC, pMPMK4, and pDA1846. Cultures of these transformants were serially diluted 1:20 and spotted on an L-plate containing 50 μg/ml kanamycin and 0.05% arabinose. The plate was incubated at 37° overnight. The nature of the plasmid-encoded protein is indicated along the top of the figure.

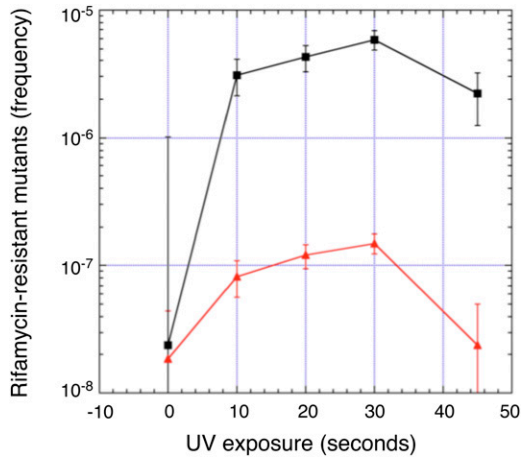

Figure 5 Expression of T4 Gp39.2 restores UV-induced mutagenesis in *E. coli groEL44*. *E. coli groEL*44 cells were transformed with plasmid pAC (wild-type Gp39.2) and grown in the presence (black) or absence (red) of 0.02% arabinose. The experiment was performed twice. Frequency of mutagenesis to Rif^R was determined as described in *Materials and Methods*.

two bacteriophages exhibited a differential growth pattern on *E. coli groEL*44. At the semi-permissive temperature of 30°, wild-type T4 forms plaques, albeit small, while T4 Δ*39.2* forms no plaques (Figure 6A). We found four additional strains and conditions under which the wild-type T4 parental strain grows much better than T4 Δ*39.2*. The differential plating behavior of the two bacteriophages on *E. coli groEL*3388 at 39° is shown in Figure 6B. *E. coli groEL*3388, as well as the other three strains, carries the original *groEL*44 (E191G) mutation and an additional intragenic mutation that suppresses the temperature-sensitive phenotype of *groEL*44 bacteria, allowing them to form colonies at 43° (Klein and Georgopoulos 2001). The intragenic mutation in *groEL*3388 results in the V190I substitution. The single *groEL* (V190I) mutation was also independently isolated in our laboratory as an extragenic suppressor of the *groES*619 temperature-sensitive phenotype at 42° (Zeilstra-Ryalls *et al.* 1994). Clearly, *groEL*3388 mutant bacteria completely block growth of the T4 Δ*39.2* mutant. Although we have not yet analyzed the biochemical properties of the GroEL3388 protein, if T4 or its close relatives encounter a bacterial host in nature possessing a GroEL protein similar to GroEL3388, certainly those bacteriophages expressing Gp39.2 will be able to propagate on it.

Discussion

Because their short growth cycles (~40 min) terminate upon host lysis, bacteriophages such as λ, T5, and T4 must complete their development in a timely fashion. In the case of T4, the most abundant protein produced by the bacteriophage is its major capsid subunit, Gp23, whose correct folding depends entirely on the host GroEL chaperone and the T4-encoded Gp31 cochaperone. Since both the *groEL* and *groES* genes are essential for *E. coli* viability (Fayet *et al.* 1989), any mutations that block bacteriophage growth cannot completely abolish the activities of these two genes. Through extensive genetic, biochemical, and biophysical analyses, we showed that most of our *groEL*, *groES*, and gene *31* mutants interfere with the normal GroEL/cochaperone cycle (Landry *et al.* 1993; Zeilstra-Ryalls *et al.* 1994; Landry *et al.* 1996; Richardson and Georgopoulos 1999; Richardson *et al.* 1999; Richardson 2000; Klein and Georgopoulos 2001; Shewmaker *et al.* 2004). Specifically, the *groEL*44, *groES*30, *groES*42, and *groES*619 mutations result in proteins with reduced affinity for the corresponding wild-type cochaperone or GroEL partner. In contrast, the *groEL*515, *groEL*673, *groEL*173, and *groEL* (V190I) mutations result in mutant proteins that tend to prolong their interaction with the cochaperone partner. On the basis of these results and those reported here, we propose the following model to explain the suppression of *groES* and *groEL* mutations by Gp39.2. Figure 7 shows the backbone ribbon structure of a single GroEL subunit in either the "closed" (which binds substrate) or "open" (which binds GroEL's cochaperones) conformation (Braig *et al.* 1994; Xu *et al.* 1997; Sigler *et al.* 1998). The large difference between the two GroEL conformations is highlighted by the R322 and K178 residues, which form a salt-bridge in the closed conformation (likely stabilizing the structure), but which are found far apart in the open GroEL conformation. We suggest that, either directly or indirectly, Gp39.2 shifts the equilibrium in favor of the GroEL open conformation. Gp39.2 can accomplish this by either destabilizing the closed or stabilizing the open GroEL conformation. This model is consistent with our bacteriophage plating results and Gp39.2 suppression of both host temperature sensitivity and the UmuC defect in our *groE* mutant strains. It is also consistent

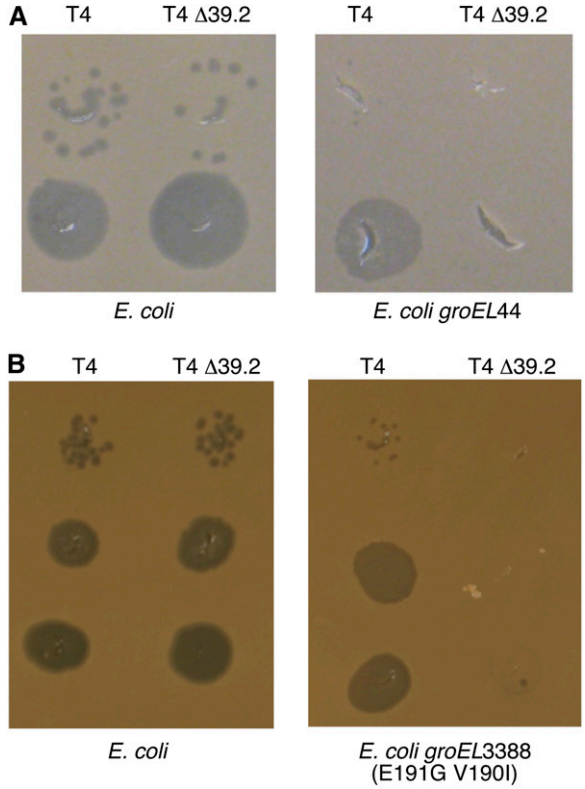

Figure 6 T4 Δ*39.2* does not propagate on some *E. coli* strains. Serial dilutions (10-fold) of wild-type T4 and T4 Δ*39.2* were spotted on lawns of *E. coli* B178 strains, as indicated above and below, respectively. (A) Plates were incubated overnight at 30°. (B) Plates were incubated overnight at 39°.

with the observation that expression of Gp39.2 is best tolerated in the *groEL*44 host, followed by the wild-type host, and is most toxic in the *groEL*515 host. We postulate that toxicity is observed in the *groEL*515 mutant because GroEL515 is held longer in the open state, rarely releasing the bound cochaperone heptamer, thus blocking the normal cycle of chaperone binding and timely release of substrate. Because Gp39.2 specifically suppresses weak GroEL–cochaperone interactions, we propose to rename its gene *nwgI* (normalizes weak GroE interactions).

Our observations also lead us to postulate that Gp39.2 acts through GroEL and not the cochaperone. The GroES protein of *E. coli* exhibits very low identity (<10%) with $Gp31^{T4}$, $Gp31^{RB43}$, and $Gp31^{RB49}$ (aka CocO; Ang *et al.* 2001). In addition, $Gp31^{T4}$, $Gp31^{RB43}$, and $Gp31^{RB49}$ are only 30–40% identical to one other. The fact that Gp39.2 expression suppresses all GroEL44 defects (which are due to GroEL44's weak interaction with its cochaperones; Richardson *et al.*, 1999; Richardson 2000; Shewmaker *et al.*, 2004) regardless of the cochaperone suggests that Gp39.2 engages directly or indirectly with GroEL to shift its equilibrium toward the open conformation, thus favoring interaction with any cochaperone. The isolation of *E. coli* mutants that tolerate high levels of Gp39.2 will perhaps help locate the putative GroEL–Gp39.2 interface.

Our failure to obtain plaque formers from the λZAP II-RB69 library on our "high affinity" *groEL*515, *groEL*673, and *groEL*173 mutants suggests that either the RB69 genome does not encode such a function or the function does not suppress these particular GroE mutant defects. However, on the basis of our model of the biological role of Gp39.2, it would not be surprising to discover bacteriophages that do indeed encode a function opposite that of Gp39.2, *i.e.*, a polypeptide that shifts the GroEL equilibrium toward the closed state. This putative function would be beneficial in hosts encoding a GroEL with a higher affinity for the bacteriophage-encoded Gp31, which would again slow down the chaperone cycle and timely maturation of Gp23, thus reducing the yield of bacteriophage progeny. We predict that such bacterial hosts should display a differential plating pattern between isogenic wild-type T4 and T4 Δ*39.2* bacteriophages. Thus far, we have not detected any such hosts among the various Enterobacteriaceae collections tested.

What is the origin of the bacteriophage 39.2 gene?

A search of the extant nucleotide sequence database shows that the Gp39.2 family of bacteriophage proteins is encoded only by those T4-like bacteriophages that can propagate on Enterobacteriaceae. Gp39.2 also belongs to the larger superfamily of so-called FmdB or CxxC_CxxC_SSSS proteins that contain a putative zinc finger domain found in a wide range of bacteria, but not Enterobacteriaceae. The superfamily is characterized by two CxxC motifs separated by approximately 17 amino acid residues. Consistent with the prediction that members of this superfamily likely bind zinc, we have shown that a monomer of $Gp39.2^{T4}$ binds a molecule of zinc (unpublished data of Steve Alam, University of Utah). This suggests that zinc binding is necessary for the structure/function of Gp39.2 and helps explain why the $Gp39.2^{T4}$ (C13A or C31A C34A) mutant proteins are inactive in *in vivo* biological experiments.

Curiously, no function has been assigned yet to FmdB, the founding member of this superfamily. The original *fmdB* gene, encoding a 112-amino-acid polypeptide, was named thus because it immediately follows the *fmdA* (formamidase) gene of *Methylophilus methylotrophus* (Wyborn *et al.* 1996) and is simply referred to as "putative FmdB regulatory protein" in the literature. Most nonbacteriophage FmdB family members carry the putative zinc-binding region at their amino-terminal end, and a 20- to 60-amino-acid extension at their carboxyl ends, making them longer than Gp39.2 family members. There is a handful of recently identified and sequenced *Myoviridae* bacteriophages, very distant members of the T4-like superfamily, that all encode a *bona fide fmdB* gene of 81 amino acids in length. The original member of this small group is the Vi01 bacteriophage of *Salmonella typhimurium* (Pickard *et al.* 2010), whose putative FmdB product is 96% identical to the *S. enterica* bacteriophage Det7 FmdB-like protein (Walter *et al.* 2008; S. Casjens, personal communication), and 94% identical to both the *E. coli* O157:H7 bacteriophage CBA120 (Kutter *et al.* 2011) and the *Shigella boydii* SboM-AG3 bacteriophage FmdB-like proteins (Anany *et al.* 2011).

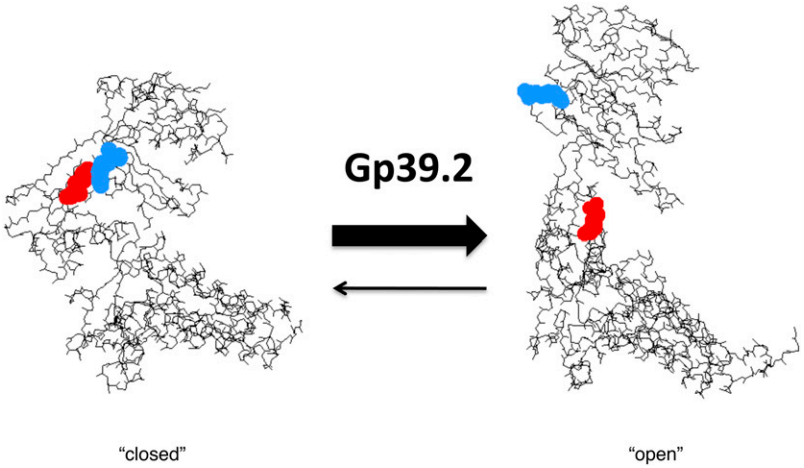

Figure 7 Model: Gp39.2 favors GroEL/cochaperone interaction by shifting the equilibrium of GroEL toward the "open" state. To explain our genetic results, we propose that the Gp39.2 family of bacteriophage-encoded proteins directly or indirectly shifts the equilibrium of GroEL toward the open conformation, thus favoring binding of GroEL's cochaperones. Only one of the seven GroEL subunits of the *cis*-GroEL ring is shown here, based on the corresponding X-ray structure of Braig *et al.* (1994), Xu *et al.* (1997), and Sigler *et al.* (1998). The basic R322 and acidic E178 residues are highlighted in blue and red, respectively, to emphasize the massive *en bloc* rearrangements in the two structures (see text for details).

Recently, Arbiol *et al.* (2010) provided evidence that at least two members of the greater T4-like family exchange genes by a novel mechanism of modular gene shuffling. Specifically, using bacteriophage RB43 and its very close relative coliphage phi1, Arbiol *et al.* (2010) identified small regulatory cassettes, termed PeSLs, which contain an early type T4 promoter sequence and a transcription terminator. Modular shuffling of genes is likely mediated by recombination between PeSLs, and this genetic exchange likely proceeds through small circular DNA intermediates formed by the recombination event. Surprisingly, such small circular DNAs are detected not only in extracts of RB49/phi1-infected hosts but also in purified encapsidated virions. These novel and exciting results offer an efficient mechanism for readily exchanging genes without destroying the parental DNA molecules in the process. For example, members of the Vi01 family of bacteriophages may have acquired a host *fmdB* gene(s). Subsequent co-infection of a common enterobacterial host with an ancestral T4-like bacteriophage may have resulted in exchange of the *fmdB* gene between the two bacteriophages and eventual evolution to the present family of *39.2* genes.

The recent work of Yonesaki and colleagues exemplifies another potential role for apparently nonessential genes encoded by T4-like bacteriophages. By constructing an amber mutation in the T4 gene *61.5* (renamed *dmd*; encodes a 60-amino-acid protein), Kai *et al.* (1996) showed that it is essential for growth only on certain *E. coli* hosts. Specifically, on these restrictive hosts in the absence of a functional *61.5* (*dmd*) gene, T4 late mRNAs are completely destabilized by RNase LS, thus blocking T4 late protein expression and growth. RNase LS is encoded by *rnlA*, a toxin gene whose product is inhibited by the product of the downstream antitoxin gene *rnlB* (Koga *et al.* 2011). The *rnlA–rnlB* toxin–antitoxin pair is located on the CP4-57 defective prophage in *E. coli* K-12 (Blattner *et al.* 1997). Thus, it appears that T4-like bacteriophages have acquired and/or evolved the *dmd* gene to propagate on bacterial hosts possessing *rnlA*-like genes (apparently abundant in nature, judging from a BLASTp search for the *rnlA* gene).

Closing the circle

In the early 1970s in Geneva, Switzerland, one of us (C.G.) discussed with Dick Epstein whether one can isolate T4 mutants that propagate on *E. coli* K-12 strains but not on *E. coli* B, regardless of the host's ability to suppress nonsense mutations. Encouraged by Epstein, C.G. screened and isolated two nonamber T4 mutants meeting these criteria and showed that the mutations map in genes *8* and *53* (Georgopoulos *et al.* 1977). These two genes encode structural components of the wedges that make up the T4 baseplate (Yap *et al.* 2010). The simplest interpretation is that the mutations alter Gp8 or Gp53 in such a way that T4 cannot interact properly with the host receptor(s) necessary for proper T4 attachment and/or DNA injection. It is interesting to note that at least one of the original T4 "amber" mutants, HL626, originally thought to be in gene *60* (Yegian *et al.* 1971), is analogous to the mutants of Georgopoulos *et al.* (1977). These results exemplify the power of simple microbial genetic selections and screens to detect many classes of mutant phenotypes, provided they are set up appropriately. However, as seen with the *39.2* (*nwgI*) story described here, even what appears to be an obvious selection can lead to unexpected yet interesting results.

Acknowledgments

The authors are grateful to Dr. Jim Karam for the generous gift of the λZAP II-RB69 library and thank Dick Epstein, Dominique Belin, France Keppel, Dana Carroll, and Sherwood Casjens for useful discussions. This work was initiated at the University of Geneva, Switzerland, supported by FN-3100AO-114063. Currently, National Institutes of Health R01 grant GM88207 supports this work.

Literature Cited

Altschul, S. F., and D. J. Lipman, 1990 Protein database searches for multiple alignments. Proc. Natl. Acad. Sci. USA 87: 5509–5513.

Anany, H., E. J. Lingohr, A. Villegas, H. W. Ackermann, Y. M. She *et al.*, 2011 A *Shigella boydii* bacteriophage which resembles Salmonella phage ViI. Virol. J. 8: 242.

Ang, D., F. Keppel, G. Klein, A. Richardson, and C. Georgopoulos, 2000 Genetic analysis of bacteriophage-encoded cochaperonins. Annu. Rev. Genet. 34: 439–456.

Ang, D., A. Richardson, M. P. Mayer, F. Keppel, H. Krisch *et al.*, 2001 Pseudo-T-even bacteriophage RB49 encodes CocO, a cochaperonin for GroEL, which can substitute for *Escherichia coli*'s GroES and bacteriophage T4's Gp31. J. Biol. Chem. 276: 8720–8726.

Arbiol, C., A. M. Comeau, M. Kutateladze, R. Adamia, and H. M. Krisch, 2010 Mobile regulatory cassettes mediate modular shuffling in T4-type phage genomes. Genome Biol. Evol. 2: 140–152.

Benzer, S., 1955 Fine structure of a genetic region in bacteriophage. Proc. Natl. Acad. Sci. USA 41: 344–354.

Benzer, S., 1959 On the topology of the genetic fine structure. Proc. Natl. Acad. Sci. USA 45: 1607–1620.

Blattner, F. R., G. Plunkett 3rd. C. A. Bloch, N. T. Perna, V. Burland *et al.*, 1997 The complete genome sequence of Escherichia coli K-12. Science 277: 1453–1462.

Braig, K., Z. Otwinowski, R. Hegde, D. C. Boisvert, A. Joachimiak *et al.*, 1994 The crystal structure of the bacterial chaperonin GroEL at 2.8 A. Nature 371: 578–586.

Campbell, A., 1961 Sensitive mutants of bacteriophage lambda. Virology 14: 22–32.

Casjens, S. R., 2008 Diversity among the tailed-bacteriophages that infect the Enterobacteriaceae. Res. Microbiol. 159: 340–348.

Clare, D. K., P. J. Bakkes, H. Van Heerikhuizen, S. M. Van Der Vies, and H. R. Saibil, 2009 Chaperonin complex with a newly folded protein encapsulated in the folding chamber. Nature 457: 107–110.

Comeau, A. M., C. Bertrand, A. Letarov, F. Tetart, and H. M. Krisch, 2007 Modular architecture of the T4 phage superfamily: a conserved core genome and a plastic periphery. Virology 362: 384–396.

Comeau, A. M., G. F. Hatfull, H. M. Krisch, D. Lindell, N. H. Mann *et al.*, 2008 Exploring the prokaryotic virosphere. Res. Microbiol. 159: 306–313.

Comeau, A. M., C. Arbiol, and H. M. Krisch, 2010 Gene network visualization and quantitative synteny analysis of more than 300 marine T4-like phage scaffolds from the GOS metagenome. Mol. Biol. Evol. 27: 1935–1944.

Denou, E., A. Bruttin, C. Barretto, C. Ngom-Bru, H. Brussow *et al.*, 2009 T4 phages against *Escherichia coli* diarrhea: potential and problems. Virology 388: 21–30.

Donnelly, C. E., and G. C. Walker, 1989 groE mutants of *Escherichia coli* are defective in umuDC-dependent UV mutagenesis. J. Bacteriol. 171: 6117–6125.

Donnelly, C. E., and G. C. Walker, 1992 Coexpression of UmuD' with UmuC suppresses the UV mutagenesis deficiency of groE mutants. J. Bacteriol. 174: 3133–3139.

Epstein, R. H., 1958a A study of multiplicity-reactivation in the bacteriophage T4, pp. 101. Ph.D. thesis, University of Rochester, Rochester, NY.

Epstein, R. H., 1958b A study of multiplicity-reactivation in the bacteriophage T4. I. Genetic and functional analysis of T4D–K12 (lambda) complexes. Virology 6: 382–404.

Epstein, R. H., A. Bolle, C. M. Steinberg, E. Kellenberger, E. Boy De La Tour *et al.*, 1964 Physiological studies of conditional lethal mutants of bacteriophage T4D. Cold Spring Harb. Symp. Quant. Biol. 28: 375–394.

Epstein, R. H., A. Bolle, and C. M. Steinberg, 2012 Amber mutants of bacteriophage T4D: their isolation and genetic characterization. Genetics 190: 831–840.

Fayet, O., T. Ziegelhoffer, and C. Georgopoulos, 1989 The groES and groEL heat shock gene products of *Escherichia coli* are essential for bacterial growth at all temperatures. J. Bacteriol. 171: 1379–1385.

Fischer, D., and D. Eisenberg, 1999 Finding families for genomic ORFans. Bioinformatics 15: 759–762.

Friedman, D. I., E. R. Olson, C. Georgopoulos, K. Tilly, I. Herskowitz *et al.*, 1984 Interactions of bacteriophage and host macromolecules in the growth of bacteriophage lambda. Microbiol. Rev. 48: 299–325.

Georgopoulos, C., 2006 Toothpicks, serendipity and the emergence of the *Escherichia coli* DnaK (Hsp70) and GroEL (Hsp60) chaperone machines. Genetics 174: 1699–1707.

Georgopoulos, C., M. Georgiou, G. Selzer, and H. Eisen, 1977 Bacteriophage T4 mutants which propagate on *E. coli* K12 but not on *E. coli* B. Experientia 33: 1157–1158.

Guzman, L. M., D. Belin, M. J. Carson, and J. Beckwith, 1995 Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. J. Bacteriol. 177: 4121–4130.

Hartl, F. U., A. Bracher, and M. Hayer-Hartl, 2011 Molecular chaperones in protein folding and proteostasis. Nature 475: 324–332.

Hatfull, G. F., 2008 Bacteriophage genomics. Curr. Opin. Microbiol. 11: 447–453.

Hendrix, R. W., 2002 Bacteriophages: evolution of the majority. Theor. Popul. Biol. 61: 471–480.

Hershey, A. D., 1946 Mutation of bacteriophage with respect to type of plaque. Genetics 31: 620–640.

Hershey, A. D., and R. Rotman, 1948 Linkage among genes controlling inhibition of lysis in a bacterial virus. Proc. Natl. Acad. Sci. USA 34: 89–96.

Horwich, A. L., and W. A. Fenton, 2009 Chaperonin-mediated protein folding: using a central cavity to kinetically assist polypeptide chain folding. Q. Rev. Biophys. 42: 83–116.

Houry, W. A., D. Frishman, C. Eckerskorn, F. Lottspeich, and F. U. Hartl, 1999 Identification of in vivo substrates of the chaperonin GroEL. Nature 402: 147–154.

Huang, W. M., 1986 Nucleotide sequence of a type II DNA topoisomerase gene. Bacteriophage T4 gene 39. Nucleic Acids Res. 14: 7751–7765.

Hunt, J. F., A. J. Weaver, S. J. Landry, L. Gierasch, and J. Deisenhofer, 1996 The crystal structure of the GroES co-chaperonin at 2.8 A resolution. Nature 379: 37–45.

Hunt, J. F., S. M. Van Der Vies, L. Henry, and J. Deisenhofer, 1997 Structural adaptations in the specialized bacteriophage T4 co-chaperonin Gp31 expand the size of the Anfinsen cage. Cell 90: 361–371.

Kai, T., H. E. Selick, and T. Yonesaki, 1996 Destabilization of bacteriophage T4 mRNAs by a mutation of gene 61.5. Genetics 144: 7–14.

Keppel, F., B. Lipinska, D. Ang, and C. Georgopoulos, 1990 Mutational analysis of the phage T4 morphogenetic 31 gene, whose product interacts with the *Escherichia coli* GroEL protein. Gene 86: 19–25.

Keppel, F., M. Rychner, and C. Georgopoulos, 2002 Bacteriophage-encoded cochaperonins can substitute for *Escherichia coli*'s essential GroES protein. EMBO Rep. 3: 893–898.

Kerner, M. J., D. J. Naylor, Y. Ishihama, T. Maier, H. C. Chang *et al.*, 2005 Proteome-wide analysis of chaperonin-dependent protein folding in *Escherichia coli*. Cell 122: 209–220.

Klein, G., and C. Georgopoulos, 2001 Identification of important amino acid residues that modulate binding of *Escherichia coli* GroEL to its various cochaperones. Genetics 158: 507–517.

Koga, M., Y. Otsuka, S. Lemire, and T. Yonesaki, 2011 Escherichia coli rnlA and rnlB compose a novel toxin-antitoxin system. Genetics 187: 123–130.

Koonin, E. V., and S. M. Van Der Vies, 1995 Conserved sequence motifs in bacterial and bacteriophage chaperonins. Trends Biochem. Sci. 20: 14–15.

Kutter, E. M., K. Skutt-Kakaria, B. Blasdel, A. El-Shibiny, A. Castano *et al.*, 2011 Characterization of a ViI-like phage specific to *Escherichia coli* O157:H7. Virol. J. 8: 430.

Landry, S. J., J. Zeilstra-Ryalls, O. Fayet, C. Georgopoulos, and L. M. Gierasch, 1993 Characterization of a functionally important mobile domain of GroES. Nature 364: 255–258.

Landry, S. J., A. Taher, C. Georgopoulos, and S. M. Van Der Vies, 1996 Interplay of structure and disorder in cochaperonin mobile loops. Proc. Natl. Acad. Sci. USA 93: 11622–11627.

Larkin, M. A., G. Blackshields, N. P. Brown, R. Chenna, P. A. Mcgettigan *et al.*, 2007 Clustal W and Clustal X version 2.0. Bioinformatics 23: 2947–2948.

Lederberg, E. M., and J. Lederberg, 1953 Genetic studies of lysogenicity in *Escherichia coli*. Genetics 38: 51–64.

Mayer, M. P., 1995 A new set of useful cloning and expression vectors derived from pBlueScript. Gene 163: 41–46.

Nivinskas, R., and L. W. Black, 1988 Cloning, sequence, and expression of the temperature-dependent phage T4 capsid assembly gene 31. Gene 73: 251–257.

Nolan, J. M., V. Petrov, C. Bertrand, H. M. Krisch, and J. D. Karam, 2006 Genetic diversity among five T4-like bacteriophages. Virol. J. 3: 30.

Petit, M. A., W. Bedale, J. Osipiuk, C. Lu, M. Rajagopalan *et al.*, 1994 Sequential folding of UmuC by the Hsp70 and Hsp60 chaperone complexes of *Escherichia coli*. J. Biol. Chem. 269: 23824–23829.

Petrov, V. M., S. Ratnayaka, J. M. Nolan, E. S. Miller, and J. D. Karam, 2010 Genomes of the T4-related bacteriophages as windows on microbial genome evolution. Virol. J. 7: 292.

Pickard, D., A. L. Toribio, N. K. Petty, A. Van Tonder, L. Yu *et al.*, 2010 A conserved acetyl esterase domain targets diverse bacteriophages to the Vi capsular receptor of *Salmonella enterica* serovar Typhi. J. Bacteriol. 192: 5746–5754.

Revel, H. R., B. L. Stitt, I. Lielausis, and W. B. Wood, 1980 Role of the host cell in bacteriophage T4 development. I. Characterization of host mutants that block T4 head assembly. J. Virol. 33: 366–376.

Richardson, A., 2000 Function/structure studies of the bacteriophage T4-encoded co-chaperonin, Gp31, with emphasis on co-chaperonin mobile loop interactions with GroEL. Ph.D. Thesis, University of Geneva, Geneva.

Richardson, A., and C. Georgopoulos, 1999 Genetic analysis of the bacteriophage T4-encoded cochaperonin Gp31. Genetics 152: 1449–1457.

Richardson, A., S. M. Van Der Vies, F. Keppel, A. Taher, S. J. Landry *et al.*, 1999 Compensatory changes in GroEL/Gp31 affinity as a mechanism for allele-specific genetic interaction. J. Biol. Chem. 274: 52–58.

Rohwer, F., and R. Edwards, 2002 The phage proteomic tree: a genome-based taxonomy for phage. J. Bacteriol. 184: 4529–4535.

Sanson, B., and M. Uzan, 1992 Sequence and characterization of the bacteriophage T4 comC alpha gene product, a possible transcription antitermination factor. J. Bacteriol. 174: 6539–6547.

Selick, H. E., K. N. Kreuzer, and B. M. Alberts, 1988 The bacteriophage T4 insertion/substitution vector system: a method for introducing site-specific mutations into the virus chromosome. J. Biol. Chem. 263: 11336–11347.

Shewmaker, F., K. Maskos, C. Simmerling, and S. J. Landry, 2001 The disordered mobile loop of GroES folds into a defined beta-hairpin upon binding GroEL. J. Biol. Chem. 276: 31257–31264.

Shewmaker, F., M. J. Kerner, M. Hayer-Hartl, G. Klein, C. Georgopoulos *et al.*, 2004 A mobile loop order-disorder transition modulates the speed of chaperonin cycling. Protein Sci. 13: 2139–2148.

Sigler, P. B., Z. Xu, H. S. Rye, S. G. Burston, W. A. Fenton *et al.*, 1998 Structure and function in GroEL-mediated protein folding. Annu. Rev. Biochem. 67: 581–608.

Snyder, L., and H. J. Tarkowski, 2005 The N terminus of the head protein of T4 bacteriophage directs proteins to the GroEL chaperonin. J. Mol. Biol. 345: 375–386.

Stahl, F. W., 1995 The amber mutants of phage T4. Genetics 141: 439–442.

Stitt, B. L., H. R. Revel, I. Lielausis, and W. B. Wood, 1980 Role of the host cell in bacteriophage T4 development. II. Characterization of host mutants that have pleiotropic effects on T4 growth. J. Virol. 35: 775–789.

Takahashi, H., and H. Yoshikawa, 1979 Genetic study of a new early gene, comC-alpha, of bacteriophage T4. Virology 95: 215–217.

Tang, M., X. Shen, E. G. Frank, M. O'donnell, R. Woodgate *et al.*, 1999 UmuD'(2)C is an error-prone DNA polymerase, *Escherichia coli* pol V. Proc. Natl. Acad. Sci. USA 96: 8919–8924.

Van Der Vies, S. M., A. A. Gatenby, and C. Georgopoulos, 1994 Bacteriophage T4 encodes a co-chaperonin that can substitute for *Escherichia coli* GroES in protein folding. Nature 368: 654–656.

Walter, M., C. Fiedler, R. Grassl, M. Biebl, R. Rachel *et al.*, 2008 Structure of the receptor-binding protein of bacteriophage det7: a podoviral tail spike in a myovirus. J. Virol. 82: 2265–2273.

Wommack, K. E., and R. R. Colwell, 2000 Virioplankton: viruses in aquatic ecosystems. Microbiol. Mol. Biol. Rev. 64: 69–114.

Wyborn, N. R., J. Mills, S. G. Williams, and C. W. Jones, 1996 Molecular characterisation of formamidase from *Methylophilus methylotrophus*. Eur. J. Biochem. 240: 314–322.

Xu, Z., A. L. Horwich, and P. B. Sigler, 1997 The crystal structure of the asymmetric GroEL-GroES-(ADP)7 chaperonin complex. Nature 388: 741–750.

Yap, M. L., K. Mio, P. G. Leiman, S. Kanamaru, and F. Arisaka, 2010 The baseplate wedges of bacteriophage T4 spontaneously assemble into hubless baseplate-like structure in vitro. J. Mol. Biol. 395: 349–360.

Yegian, C. D., M. Mueller, G. Selzer, V. Russo, and F. W. Stahl, 1971 Properties of the DNA-delay mutants of bacteriophage T4. Virology 46: 900–919.

Zeilstra-Ryalls, J., O. Fayet, and C. Georgopoulos, 1991 The universally conserved GroE (Hsp60) chaperonins. Annu. Rev. Microbiol. 45: 301–325.

Zeilstra-Ryalls, J., O. Fayet, L. Baird, and C. Georgopoulos, 1993 Sequence analysis and phenotypic characterization of groEL mutations that block lambda and T4 bacteriophage growth. J. Bacteriol. 175: 1134–1143.

Zeilstra-Ryalls, J., O. Fayet, and C. Georgopoulos, 1994 Two classes of extragenic suppressor mutations identify functionally distinct regions of the GroEL chaperone of *Escherichia coli*. J. Bacteriol. 176: 6558–6565.

Zeilstra-Ryalls, J., O. Fayet, and C. Georgopoulos, 1996 In vivo protein folding: suppressor analysis of mutations in the groES cochaperone gene of *Escherichia coli*. FASEB J. 10: 148–152.

Zuber, S., C. Ngom-Bru, C. Barretto, A. Bruttin, H. Brussow *et al.*, 2007 Genome analysis of phage JS98 defines a fourth major subgroup of T4-like phages in *Escherichia coli*. J. Bacteriol. 189: 8206–8214.

Communicating editor: M. Johnston