Phages will out: strategies of host cell lysis

Ry Young, Ing-Nang Wang and William D. Roof

Most phages accomplish host lysis using a muralytic enzyme, or endolysin, and a holin, which permeabilizes the membrane at a programmed time and thus controls the length of the vegetative cycle. By contrast, lytic single-stranded RNA and DNA phages accomplish lysis by producing a single lysis protein without muralytic activity.

R. Young, I-N. Wang and W.D. Roof are in the Dept of Biochemistry and Biophysics, Texas A&M University, College Station, TX 77843-2128, USA.*
**tel: +1 409 845 2087,*
fax: +1 409 862 4178,
e-mail: ryland@tamu.edu

Bacteriophages must escape from the host cell and disperse their progeny to find new prey. The principal barrier to this is the continuous meshwork of peptidoglycan, a strong, stable structure that allows the bacterial envelope to withstand internal osmotic pressure. Filamentous phages, because of their unique morphology and morphogenesis, can extrude through the envelope without fatal consequences for the host. However, all other phages must either degrade or otherwise compromise the peptidoglycan to cause lysis.

Strategies for host lysis

There are at least two different strategies for phage lysis[1]. All double stranded (ds) DNA phages produce a soluble, muralytic enzyme known as an endolysin. To degrade the cell wall, endolysins require a second lysis factor, a small membrane protein designated a holin. This requirement stems from the fact that most endolysins lack a secretory signal sequence, and thus the holin, which somehow permeabilizes the membrane, is required for the endolysin to gain access to the murein. The holin thus controls the timing of lysis (Fig. 1). In many cases, a holin inhibitor is also produced, allowing fine-tuning of the lytic schedule. Finally, in many phages with Gram-negative hosts, two overlapping genes encode the auxiliary lysis proteins Rz and Rz1. The function of these proteins is uncertain, but they are thought to interact with some feature of the outer membrane or with its links to the cell wall. We think this complex multigenic strategy, holin–endolysin lysis, applies generally to dsDNA phages because it confers the capacity for evolutionary optimization of the timing of lysis and allows flexibility with respect to physiological and environmental conditions (Box 1). By contrast, simple lytic phages, such as the single-stranded (ss) DNA phage ϕX174 and the ssRNA phages, have a single lysis gene. Three entirely different types of single lysis genes are known, for which the prototypes are ϕX174 *E*, MS2 *L* and Qβ A_2. No muralytic enzyme activity has been detected in the lysates of these phages, and sequence analysis suggests that these lysis genes do not encode such an enzyme. In these cases, the lysis activity appears to have evolved late, after the replicative and morphogenesis functions, because it is either a secondary activity of a morphogenesis protein or is encoded in a short alternate reading frame embedded in an essential morphogenetic cistron. A number of significantly different models have been proposed for the mode of action of these single-gene systems[2,3]. Recent evidence indicates that, in the case of ϕX174 at least, the single lysis protein inhibits a specific step in murein biosynthesis. Here, advances since the last major reviews of these general types of lysis[1,2,4] will be summarized.

How does holin–endolysin lysis work?

Phage λ is convenient for studying lysis because an entire culture can be established in a synchronous vegetative cycle by lysogenic induction, allowing very precise and reproducible assessment of the timing of the lytic event[1]. The holin and endolysin genes, *S* and *R*, respectively are located at the beginning of the late transcriptional unit of λ (Fig. 2a), which is turned on approximately 8 min into the vegetative cycle and is constitutively expressed thereafter. The endolysin activity accumulates in the cytosol but has no access to its substrate until, at a genetically predetermined time, the holin forms a lethal membrane lesion that terminates respiration and allows the *R*-encoded endolysin to attack the murein (Fig. 1). The simplest notion is that the lesion is some sort of 'hole', which allows the endolysin to cross the cytoplasmic membrane. It should be noted, however, that the permeabilizing action of the holin on the membrane might also serve to activate a muralytic enzyme already externalized in an inactivated state, as proposed for the LytA amidase from *Streptococcus*[5,6]. In any case, destruction of the cell wall and lysis follow immediately and, given the synchrony achieved in lysogenic inductions, the result is an abrupt culture-wide loss of turbidity. In a sense, macroscopic lysis can be thought of as a reporter event for holin function, because endolysin activity is harmless inside the cytosol and is irrelevant to either the timing or extent of the lethal function of the holin. By contrast, deleting the holin gene permits continued respiration, macromolecular

0966-842X/00/$ - see front matter PII: S0966-842X(00)01705-4

synthesis and an increase in cell mass, with endolysin and progeny virions accumulating to high levels in the cytosol. A common characteristic of holin-based systems is their dependence on the energized membrane; addition of an energy poison triggers holin action essentially instantaneously[1]. This 'premature lysis' phenomenon indicates that the timing mechanism for holins depends on the energized membrane. The two most interesting questions about this process are: what is the nature of the 'hole' that permits transit of fully folded endolysins across the membrane, and how is the timing of lysis programmed into the structure of the holin gene?

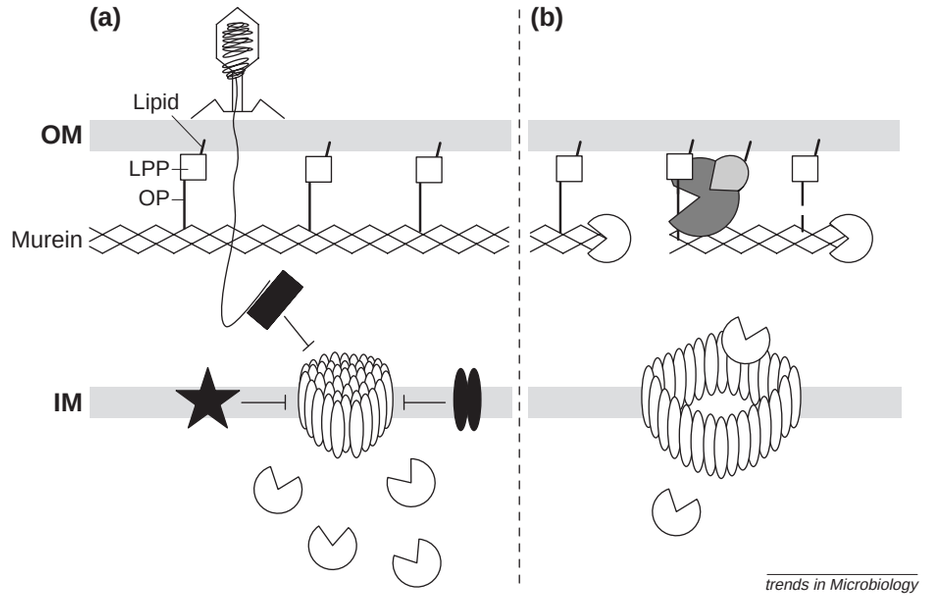

Fig. 1. The fate of the Gram-negative envelope during holin–endolysin lysis. The envelope comprises inner (IM) and outer (OM) membranes and murein, which is linked to the OM via oligopeptide (OP) links to the lipoprotein (LPP) tethered to the OM by lipid moieties. **(a)** Pre-hole configuration. Holins (clear ovals; e.g. $S^{\lambda}105$) accumulate in membrane aggregates and endolysin (notched clear circles; e.g. λ R) accumulates in the cytosol. Triggering of holin action can be inhibited by inhibitors, which can be membrane proteins orthologous (filled ovals; e.g. $S^{\lambda}107$) or non-orthologous (star; e.g. P2 LysA) to the holin, or a periplasmic protein, as proposed for the T4 rI protein[26] (rectangle). The binding of the periplasmic inhibitor to DNA is hypothetical and proposed purely to explain the observation that the signal for T4 lysis inhibition requires the injection process. **(b)** Hole configuration. An oligomeric membrane lesion is formed by the holin, allowing escape of endolysin, which attacks the murein. Also shown is a model for a complex of the λ Rz (dark grey) and Rz1 (light grey) attacking the oligopeptide links to the OM, with Rz1 tethered to the OM by its lipid moiety.

Holins, endolysins and mosaic lysis cassettes

Holins constitute one of biology's most diverse classes of functional homologs. Five years ago, a survey of phage and prophage sequences identified 19 holin or putative holin genes, separated into 12 unrelated orthologous groups[4], and there are now even more (Table 1). Most holin sequences are short and, overall, hydrophobic in nature, with a highly hydrophilic carboxy-terminal domain. Holins can be grouped into two general classes based on primary structure analysis (Fig. 2b). Class I holins, such as S^{λ}, are usually 95 residues or longer and have three potential transmembrane (TM) domains. Class II holins are usually smaller, at approximately 65–95 residues, and the distribution of charged and hydrophobic residues indicates two TM domains. Phage endolysins are also diverse, with four different kinds of muralytic activities directed against the three different covalent linkages that maintain the integrity of the cell wall: (1) glycosylase and (2) transglycosylase activities targeting the glycosidic linkages; and (3) amidase and (4) endopeptidase activities targeting the oligopeptide crosslinks[1,7,8]. Some confusion has arisen because the endolysins of phages T4, T7 and λ have all been called 'lysozyme'; in fact, these three proteins have different muralytic activities (glycosylase, amidase and transglycosylase, respectively). Recently, the endolysin of the *Staphylococcus aureus* phage 187 was shown to be a single 70-kDa polypeptide with domains specifying two different muralytic activities[7,8]. The *in vivo* complementation data available suggest that any holin will work with any endolysin, even when holin and endolysin pairs are chosen from phages of Gram-positive and Gram-negative hosts[9]. Within lambdoid coliphages, the lysis genes are grouped into a 'cassette', which is clearly a recombinational mosaic; each lambdoid phage can have as its holin an ortholog of the class I holin S^{λ}, or the class II holin S^{21}, and an ortholog of the λ transglycosylase or the 21 glycosylase as its endolysin (Fig. 2).

Box 1. Phage lysis from first principles

Two concepts are fundamental to our understanding of the molecular events in phage lysis. First, the most important regulatory decision in a phage infective cycle is when to effect host lysis. Lysis effected too soon would sacrifice the most productive period of intracellular virion assembly. However, extending the period of intracellular virion accumulation in a single infected cell must be balanced against the opportunity to infect new hosts. These considerations mean that there is evolutionary pressure towards an optimum lysis time[a,b]. Moreover, once the precise time to lyse the host has been decided, the major task in effecting lysis is then to subvert the continuous integrity of the peptidoglycan as rapidly as possible; that is, lysis should be saltatory. A lysis mechanism which would be associated with a slow deterioration of the host's capacity for macromolecular synthesis would significantly handicap a phage in comparison to a competitor with more decisive lytic ability, that is, 'If it were done when 'tis done, then 'twere well it were done quickly.' (Shakespeare)

References

a Abedon, S.T. (1989) Selection for bacteriophage latent period length by bacterial density: a theoretical examination. *Microbiol. Ecol.* 18, 79–88
b Wang, I-N. *et al.* (1996) The evolution of phage lysis timing. *Evol. Ecol.* 10, 545–558

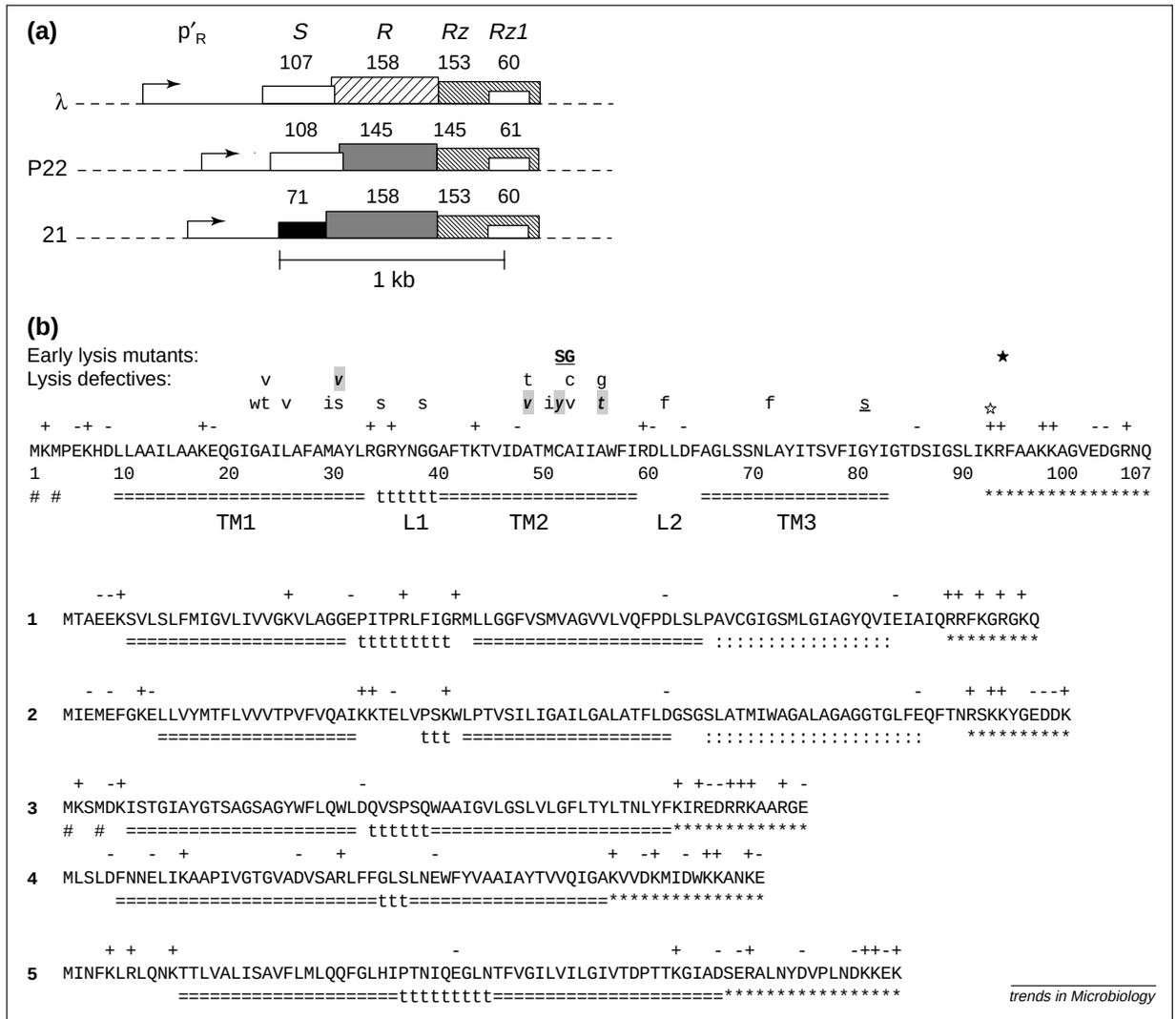

Fig. 2. Holin genes and functions. **(a)** Lambdoid lysis cassettes consisting of the holin–antiholin gene (*S*), the endolysin gene (*R*) and the *Rz/Rz1* gene, immediately downstream of the sole late promoter, p'_R. Orthologous relationships are indicated by the shading and hatching. The size of each gene in codons is shown above each reading frame. (The gene names for *S* and *R* orthologs in P22 are *13* and *19,* respectively.) **(b)** Holin sequences, domains and mutations. The sequence of $S^{\lambda}107$ is shown with predicted side-chain charges, transmembrane (TM) domains, connector loop (L) domains, predicted turn (ttttttt) domains, dual starts (#) and hydrophilic carboxy-terminal domains (*****) indicated. Above the S^{λ} sequence, missense mutations leading to a block or delay in lysis (lower case) or accelerated lysis (upper case, underlined and bold) are shown. Only charge-neutral mutations are shown. Shading indicates 'anti-dominant' (formerly 'early-dominant') alleles. Clear star and filled star indicate positions of nonsense mutations leading to lysis-negative or early-lysis phenotypes, respectively. Below the S^{λ} sequence, examples of five non-orthologous holin sequences annotated as above, are shown. Class I holins: (1) Y from P2 (*Escherichia coli*); (2) gp*14* from ϕ29 (*Bacillus subtilis*). Class II holins: (3) S^{21} (*E. coli*); (4) gp*17.5* from T7 (*E. coli*); (5) holin from DT1 (*Streptococcus thermophilus*).

Holin genetics and biochemistry

S^{λ} has been studied extensively by genetic analysis and more recently by physical and biochemical approaches. Recent gene fusion and protease-accessibility studies have demonstrated conclusively that the carboxy-terminal hydrophilic domain of S^{λ} is cytosolic[10]. Until recently, the topology in the membrane had been uncertain. Circular dichroism analysis of purified S^{λ} in the detergent octylglucoside shows ~40 residues in an α-helical conformation, suggesting two TM domains[11]. However, in a landmark paper, Graschopf and Bläsi[12] have shown that fusing a secretory signal sequence to S^{λ} confers a signal-peptidase-dependent lytic function, suggesting that the amino terminus of S^{λ} must be outside the membrane for lysis. Moreover, chemical modification studies of S^{λ} in membranes have shown that there are three TM cores approximately in the center of the hydrophobic regions discernible in the primary structure[13] (Fig. 2b). These data strongly suggest that the functional form of S^{λ} has three TM domains, with 'N-out, C-in' topology (Fig. 3a). By contrast, the simple primary structure of the prototype class II holin, S^{21}, dictates that it almost certainly has two TM domains, with 'N-in, C-in' membrane topology (Fig. 3a).

To make a membrane pore large enough for pre-folded endolysins, which range from 15 kDa to >70 kDa, holins must be part of a multiprotein

complex. Crosslinking studies suggest that in the membrane S^{λ} is a homo-oligomer of at least six subunits[14]. Purified S^{λ} diluted out of a chaotropic solution can cause permeabilization of liposomes loaded with fluorescent dye. This *in vitro* hole-forming activity is abolished by the same mutations that block hole formation *in vivo*[11]. These data suggest that the hole formed by S^{λ} is a homo-oligomeric complex.

S^{λ} has been subjected to extensive genetic analysis. The phenotypes of nonsense and frameshift mutations have revealed that the cytoplasmic carboxy-terminal domain is dispensable for hole formation and serves a regulatory role dependent on its positively charged residues[10]. Many missense mutations with altered lysis phenotypes have been mapped to the TM1 loop 1 (L1)–TM2 region (Fig. 2b)[15]. As would be expected for an oligomerizing protein, some lysis-defective S^{λ} alleles are dominant. However, in some cases, this dominance takes on an unexpected form – lysis-defective alleles, co-expressed at equal levels with wild-type S^{λ}, cause lysis at the same time or earlier than two wild-type alleles under the same conditions. This bizarre phenotypic trait, which we call anti-dominance (formerly early-dominance[15]), might derive from the necessity of propagating a particular conformational change throughout an oligomeric aggregate of holin molecules. An anti-dominant mutant protein might not only be less capable of assuming the necessary conformation but also less resistant to assuming that conformation as it is assembled into an oligomer nucleated by wild-type protein.

Of all S mutations, the most instructive are two mutations in the same position within TM2: A52V and A52G (Fig. 2b)[16]. These mutations cause plaque-formation defects for opposite reasons. A52V confers a complete lysis defect, resulting in futile intracellular hyper-accumulation of virions, whereas A52G causes catastrophically early lysis, at about 20 min, before the first virion is assembled. If nothing else, this is formal proof that holins have two essential functions: to cause lysis by permeabilizing the membrane for the endolysin and, just as important, to delay lysis until the program of virion assembly has generated an appropriate number of progeny (Box 1).

Table 1. Holins by class and host

Host[a]	Class[b]	No. different orthologous groups[c]	Total no. holin sequences
Gram-negative	I	7	12
	II	6	19
	?	1	12
Gram-positive	I	5	26
	II	9	30
	?	3	5
Mixed	?	1	9

[a]Host of phage, prophage or cryptic prophage in which holin gene has been identified. Mixed refers to a family of possible holin sequences found adjacent to muralytic enzyme sequences in several bacteria and archaea[53,54].
[b]Class I and II distinction based on whether primary structure has three or two potential transmembrane (TM) domains, respectively. A question mark indicates that the number of TM domains is unclear from inspection of sequence, and thus assignment to class I or II is not possible.
[c]Number of holin sequence families.

Regulation of holin function

What makes S^{λ} trigger 'hole-formation' at 50 min, rather than earlier or later and, in view of the ability of energy poisons to trigger premature lysis, what is the role of the energized membrane? Part of the answer lies in the fact that, despite spanning only 107 codons, S^{λ} actually encodes two proteins with opposing functions: the holin, or $S^{\lambda}105$, and a specific inhibitor, $S^{\lambda}107$, by virtue of independent translational initiations at codons 3 and 1, respectively (Fig. 3b)[15,17]. The timing of lysis depends on the proportion of the two proteins, which is 2:1 in favor of $S^{\lambda}105$ under standard conditions in broth culture, and the distinct inhibitor function of $S^{\lambda}107$ depends on a positively charged residue at position 2. Moreover, Graschopf and Bläsi have shown that fusing a signal sequence to the amino terminus of $S^{\lambda}107$ abolishes its inhibitor capacity, converting it into a lytic holin[12]. These considerations have led to a compelling model for the triggering event in holin function, its energy dependence and the inhibitor mechanism. In this model, S^{λ} is initially inserted as a 'helical hairpin' comprising TM2 and TM3. Hole formation would require S^{λ} to assume its final three TM domain topology by penetration of the amino terminus through the bilayer (Fig. 3a), an event known to be inhibited by membrane potential in other proteins with analogous topology[18,19]. The $S^{\lambda}107$ inhibitor is therefore defective in spontaneous triggering because its extra amino-terminal positive charge blocks traversal of the bilayer. Moreover, it would be susceptible to triggering by the addition of exogenous cyanide or carbonylcyanide-m-chlorophenylhydrazone (CCCP), which would depolarize the membrane and thus reduce the energetic barrier.

Although this is an attractive hypothesis, it does not address what happens after the creation of the three TM domain topology. It does imply that the presence of $S^{\lambda}107$ in its two TM domain state can retard hole formation by excess $S^{\lambda}105$. Artificially altering the ratio in favor of the inhibitor or holin retards or accelerates the onset of lysis, respectively[20], suggesting that the ability of $S^{\lambda}107$ to retard lysis is a population event that can integrate the physical state of many S molecules. Normally, there are ~$1–5 \times 10^3$ S proteins in the membrane at the time of spontaneous triggering. However, it is not known whether S is distributed uniformly in the membrane or collected in clusters, and attempts to visualize the membrane lesions by ultrastructural techniques have not been rewarding[21]. In any case, the dual start motif does not account for all of the timing properties of S^{λ}, because alleles producing only $S^{\lambda}105$ still have a sharply defined lysis time, albeit somewhat earlier than wild type. Thus, the dual start motif represents fine-tuning

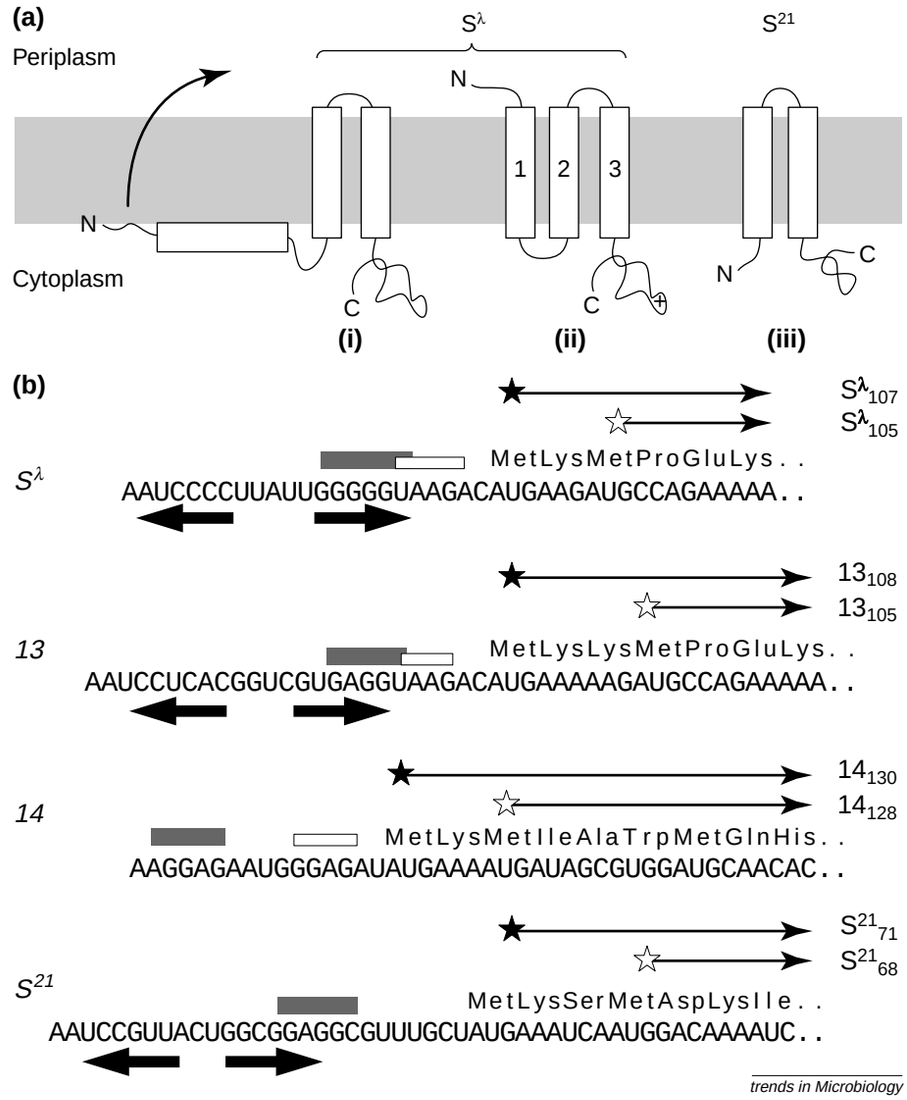

Fig. 3. Holin topologies. **(a)** Putative transmembrane (TM) topology of S^{λ} and S^{21}. **(i)** shows a model for the triggering step which would be opposed by membrane potential, especially for $S^{\lambda}107$, with its additional positive charge at the amino terminus. **(b)** Dual starts of class I and class II holin genes. Filled and empty stars show starts of the long (inhibitor) and short (effector) form of each holin. Shaded and empty boxes show the Shine–Dalgarno sites for the long and short forms, respectively. The inverted pairs of arrows show RNA stem loops controlling dual starts. The *13* gene is from the *Salmonella typhimurium* P22 phage, and *14* is from the ϕ29 *Bacillus subtilis* phage.

of holin function. It would seem likely that the holin–holin inhibitor ratio is actively regulated under physiological conditions where a more extended vegetative cycle would be favored, but no evidence for such regulation has been reported.

Recently, preliminary mutational analysis of the prototype class II holin gene, *S*^{21}, has shown that it also has a functional dual start motif leading to the synthesis of a holin, $S^{21}68$, and a holin inhibitor, $S^{21}71$ (Fig. 3b)[22]. However, the amino terminus of S^{21} is almost certainly cytosolic, given its simple primary structure, and thus penetration of the bilayer is not an issue. The apparent conservation of the dual start motif between both class I and class II could be a coincidence, reflecting the enormous pressure for optimization of these small genes. How much the differences in the dual start regulation between the class I and class II prototypes might also extend to other features of holin function is a question that awaits systematic genetic, physiological and biochemical analysis of S^{21}.

Not all holin genes use dual start motifs to generate a holin inhibitor. In the phages P1 and P2, which are dissimilar to lambdoid phages and to one another, there are separate genes encoding holin inhibitors[23,24]. Nothing is known about the mode of inhibition in either case, except that the P2 holin inhibitor LysA appears to be a membrane protein (Fig. 1). The most sophisticated example of holin control is also one of the oldest phenomena in molecular genetics: lysis inhibition (LIN) in T4 infections[25]. In T4, where *t* and *e* are the holin and endolysin genes, respectively, infected cells can enter the LIN state, in which lysis is blocked and virions continue to accumulate in the infected cell for hours[26]. The signal to induce LIN is a secondary infection by T4 during the infective cycle. Somehow, the attempted injection of the superinfecting phage DNA, or at least some step beyond simple adsorption, is detected and relayed to the holin by products of the well known T4 *r* genes (Fig. 1)[26]. Over more than 30 years, a number of *r* loci have been found that give rise to mutations abolishing the capacity for LIN. Among these, five loci have been firmly established: *rI*, *rIIA*, *rIIB*, *rIII* and *rV*. Recently, sequencing revealed that *rV* is allelic to *t*, which demonstrates that the ultimate recipient of the *r*-mediated signal is the *t*-encoded holin[27].

Rz/*Rz1*: nested genes in phages of Gram-negative hosts

The third DNA sequence in the lysis cassette contains two genes, *Rz* and *Rz1*, with the latter occupying the last third of *Rz* in a +1 reading frame (Fig. 2a)[28]. Recently, complementation and suppression analysis have established that both of these nested genes are required for lysis if the outer membrane is stabilized by millimolar concentrations of divalent cations[29]. In the absence of *Rz* or *Rz1*, and in the presence of divalent cations, infected cells form meta-stable spherical

shapes that gradually lose refractility, indicating that S and R are both functioning. *Rz1* encodes a 60-residue prolipoprotein, processing of which by signal peptidase II yields a mature, Cys-lipoylated 40-residue protein located in the outer membrane[28]. The function of Rz and Rz1 is unknown, but could account for an endopeptidase activity in λ lysates originally ascribed to R (Ref. 30). An attractive idea is that Rz and Rz1 form a complex and cause the cleavage of the oligopeptide links between the murein and the outer membrane protein Lpp. According to this model, the spherical cell phenotype results from failure to cleave these links to the compromised residual murein under conditions where the outer membrane is stabilized by cation.

```
Qin     CVPKQKNMPYALTQRSIPQILPLPSEAKQPKPPKECSPTCSEILQQKLSFMLKLLTNATSQE
P2      CTSAPPAPTPVIVPNACPKVSLCPMPGSDPQTNGDLSADIRQLENALARCASQVKMIKHCQDENDAQTRQPAQ
PS119   CSSKENALCHPQPKPPAPPAWAMMPPSNSLQLLDETFSVSGTESSATKQH
DLP12   CTSKQSVSQ  CVKPPRPPAWIMQPPPDWQTPLDGIISPSERG
21      CTSKQSVSQ  CVKPRLPPAWIMQPPPDWQTPLNGIISPSERG
λ       CTSKQSVSQ  CVKPPPPPAWIMQPPPDWQTPLDGIISPSERG
        1         10        20        30        40
H19     CGSTPPVQVP CVKPPAPPATIMQPAPNWQTPLHGI SPSESG
ES18    CSSTPPVPQCQQVNPPPPPAWIMHPAPDWQTPLNGIISPSKSE
P22     CKSPPPVQSQRPEPAAWAMEKAQDLQQMLDSIITVSEVESTG
PS3     CKSPPPVQSQR    PEPAAWSMEKAQDLQQMLNSIITVSEVESTG
PS34    CKSPPPVQSQR    PEPAAWAMEKAQDLQQMLNSIITVSEVESTR
T7      CASESKL        PEPPMVSVDSSLMVEPNLTTEMLNVFSQ
                                                  trends in Microbiology
```

Fig. 4. Sequences of predicted mature Rz1 lipoproteins[29]. The phage or cryptic prophage sources are indicated (DLP12 and Qin are defective prophages of *Escherichia coli*). Residues conserved in at least seven Rz1 sequences are shaded. The numbering is for the mature sequence of λ Rz1.

Rz/Rz1 orthologs are found in many phages of Gram-negative bacteria, but not all. Rz1 sequences, as well as having a conserved Cys residue for signal peptidase II processing and lipoylation, also contain many Pro residues (Fig. 4). It is unexpected to find two overlapping polypeptide sequences responsible for the same phenotype. To our knowledge, the only other example also involves lysis genes: the putative holin embedded out of frame within the endolysin gene of the *Staphylococcus* phage 187 (Ref. 7).

Single-gene lysis: lysis without muralytic enzymes

The simplest, smallest, lytic bacteriophages are the ssDNA phages, represented by ϕX174 (and the related *Microviridae* α3, G4 and S13), with ten genes occupying a 5.4-kb genome, and the ssRNA phages, which have just three or four genes in <4.3 kb (Ref. 1). In each of these cases, a single gene is required for lysis: *E* in ϕX174, *L* in MS2/GA classes of RNA phages and A_2 in the Qβ/SP classes (Fig. 5). The *E* and *L* genes encode small membrane proteins (Fig. 6), and are embedded within a major morphogenesis gene in an alternate reading frame[31,32]. By contrast, the Qβ/SP phages have no distinct lysis gene. Rather, the 41-kDa A_2 maturation protein, which binds the host sex pilus, has a second role as a lysis protein[33,34] (Fig. 5). Expression of plasmid-borne *E*, *L* or A_2 causes quantitative host lysis, although not with the saltatory profile of holin–endolysin lysis[34–37]. Despite these interesting distinctions, nothing has been published relevant to the mode of RNA phage lysis since the last major review on phage lysis[1]. However, significant recent literature has arisen about ϕX174 E, which we will briefly consider.

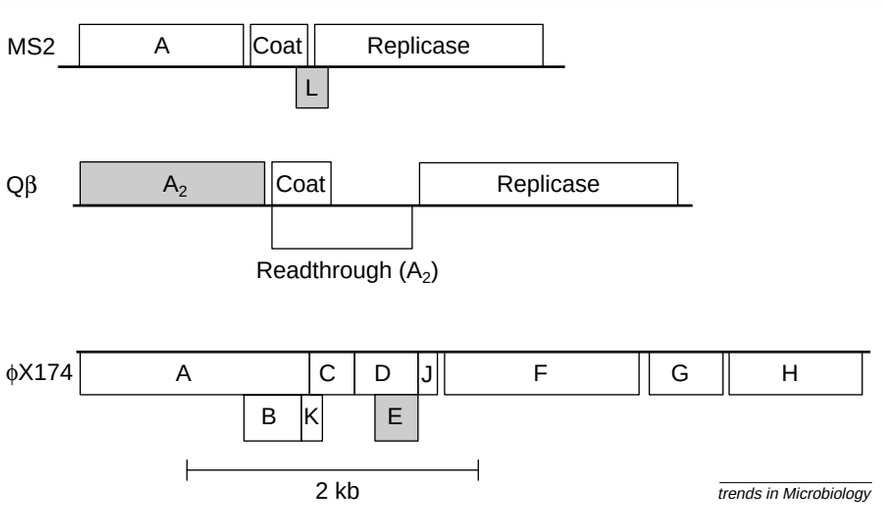

Fig. 5. Small lytic phages with single-gene lysis systems. Lysis genes are shaded. ϕX174 is shown linearized at the start of the *A* gene.

Multiple models for E lysis

E is a 91-residue integral membrane protein that, when expressed, is sufficient for lysis if the host is actively growing[2]. E comprises two domains, a hydrophobic amino-terminal domain and a positively charged carboxy-terminal domain (Fig. 6). The amino-terminal domain of the protein is conserved in related phages, suggesting that this is the 'business' end of the molecule. This notion is supported by the fact that the entire carboxy-terminal domain can be replaced by heterologous sequences, including the bulk of *lacZ*, without loss of lytic function[38,39]. A host gene, *slyD*, is absolutely required for lysis by wild-type E. *slyD* encodes a 196-residue protein, the first 150 residues of which constitute a peptidyl-prolyl *cis-trans* isomerase (PPIase) orthologous to human FKBP12 (Refs 40,41). In addition to the PPIase activity, SlyD has a carboxy-terminal domain rich in His and Cys residues, which confers a metal[42] and

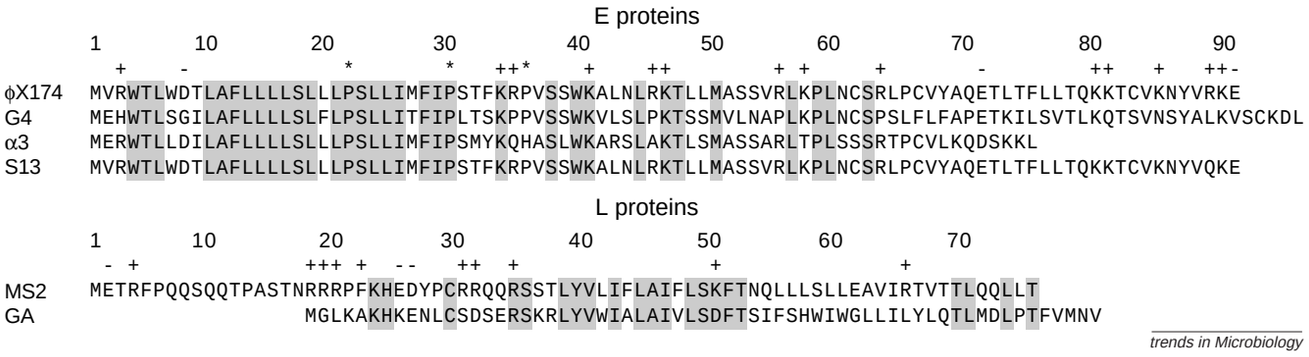

Fig. 6. E and L lysis proteins. Conserved residues are shaded. Proline residues conserved in the hydrophobic region of E are indicated by asterisks. Positively (+) and negatively (–) charged resides are also indicated.

nucleotide binding[43] capacity (and also causes it to contaminate essentially all preparations of purified His-tagged proteins[44]). Given the fact that *slyD* encodes a PPIase, it is probable that it is not required by E in its lytic function, but rather is necessary for the folding of E or its delivery to the membrane. Less than 500 molecules of E are present at the time of lysis[38].

Lysis by E requires continued host cell division[45]. Ultrastructural studies and light microscopy of cells undergoing E-mediated lysis show large bulges emerging from the midpoint or developing septum[40,46]. These bulged intermediates can be stabilized by 0.1 M $MgSO_4$, which presumably acts as an osmoticant and an outer-membrane stabilizing agent[47]. These forms are indistinguishable from those observed during penicillin-induced lysis of growing cells. These observations suggest that E might act by interfering with host peptidoglycan synthesis, to the extent that the attempt to lay down new murein at the developing septum is compromised. According to this view, the result is a greatly weakened region of the cell wall at the midpoint or septal area, which in turn leads to the formation of the bulged intermediates and eventual bursting of the cell. However, other models for E lysis have also been proposed. Several reports suggest that autolytic processes are induced by E action[48,49]. The difficulty with this model is that there is no molecular definition of the so-called 'autolytic' state, and so one mystery is simply replaced by another. Witte, Lubitz and colleagues have proposed an intricate model in which E acts by forming an oligomeric 'transmembrane tunnel' which spans the entire envelope, based on electron micrographs that show localized lesions in the envelope, varying in diameter from 50 to 200 nm, located at the midpoint or poles of cells. The phase-contrast images show that the lysed cell becomes less refractile but otherwise maintains normal cell shape. The tunnel model was supported by the finding that only limited amounts of a periplasmic enzyme activity were found in the cleared lysate, compared to the quantitative recovery of cytosolic activities. Moreover, these authors reported that Pro21 in the hydrophobic domain (Fig. 6) was required for the lytic function of an E–streptavidin chimera, and proposed that SlyD was involved in catalysing a peptidyl-prolyl isomerization that was required for everting the E sequence in the envelope[50]. However, fluorescence microscopy of cells undergoing lysis mediated by a chimera formed between the amino-terminal hydrophobic domain of E and green fluorescence protein (GFP) has shown that the lysis protein is uniformly distributed in the membrane, rather than clustered around lytic lesions (W.D. Roof *et al.*, submitted). Thus, there has been no consensus on the basic mode of E-mediated lysis, much less on its molecular mechanism.

A clear resolution of this controversy could now be at hand, from genetic data. In the original search for the target of *E*, by selecting for survivors of E-mediated lysis, only *slyD* mutations were obtained. This now seems understandable, in view of the fact that *slyD* is a non-essential gene that presumably catalyses an essential step in the folding or membrane insertion of E. A more stringent mutant selection resulted in E-resistant mutations which mapped to *mraY* (Ref. 51). *mraY* encodes translocase I, which catalyses the formation of the first lipid-linked murein precursor[52]. Moreover, induction of a plasmid clone of *E* results in blockage of murein synthesis, as measured by incorporation of labeled diaminopimelate (T. Bernhardt, unpublished). Thus, E appears to act as a 'protein antibiotic', inhibiting cell wall synthesis, thereby promoting a septal catastrophe as the growing cell

Questions for future research

- What are the structures of class I and class II holins, and their 'holes' in the membrane, before and after triggering?
- What is the cause of spontaneous triggering and how is it programmed in time?
- How do the different types of holin inhibitors work? (e.g. $S^{\lambda}107$, $S^{21}71$, P2 LysA, P1 LydB and the T4 r system)
- Do Rz and Rz1 form a complex? Do they attack the links between the outer membrane and murein? How did they evolve as interacting partners encoded by different reading frames of the same gene?
- Does E inhibit murein synthesis, form an oligomeric 'transmembrane tunnel' or cause lysis in a fundamentally different way?
- Do the two different RNA phage single-gene lysis systems work in the same way as E?

attempts septation. By analogy, the finding that E acts as an inhibitor of peptidoglycan synthesis suggests that the other single-gene lysis systems could also do so. If so, it seems possible that the MS2 L protein, which, like E, is a membrane protein, could target MraY, because MraY is the only integral membrane protein that functions in the cytosolic steps of murein precursor biosynthesis. By contrast, the Qβ A_2 protein, presumably soluble when not incorporated into the virion, might target a different step in the pathway.

Conclusion

Rapid progress has been made in identifying a large number of holin genes and their regulators. The membrane topology and the domain structure of the major classes of holins are emerging. Genetic and biochemical systems have been established to address the two compelling questions of holin function: what is the 'hole', and how is holin action scheduled so precisely. Moreover, genetic data suggest that the lysis proteins of simple phages can function without muralytic activity by inhibiting cell wall synthesis, like the cell wall antibiotics of fungal origin.

Acknowledgements

We thank all the members of the Young laboratory, past and present, for their help and encouragement and Sharyll Pressley for her reliable clerical support. This work was supported by PHS grant GM27099 and funds from the Robert A. Welch Foundation and the Texas Agricultural Experiment Station.

References

1 Young, R. (1992) Bacteriophage lysis: mechanism and regulation. *Microbiol. Rev.* 56, 430–481
2 Witte, A. *et al.* (1990) φX174 protein E-mediated lysis of *Escherichia coli*. *Biochimie* 72, 191–200
3 Walderich, B. *et al.* (1988) Induction of the autolytic system of *Escherichia coli* by specific insertion of bacteriophage MS2 lysis protein into the bacterial cell envelope. *J. Bacteriol.* 170, 5027–5033
4 Young, R. and Bläsi, U. (1995) Holins: form and function in bacteriophage lysis. *FEMS Microbiol. Rev.* 17, 191–205
5 Diaz, E. *et al.* (1996) The two-step lysis system of pneumococcal bacteriophage EJ-1 is functional in Gram-negative bacteria: triggering of the major pneumococcal autolysin in *Escherichia coli*. *Mol. Microbiol.* 19, 667–681
6 Diaz, E. *et al.* (1989) Subcellular localization of the major pneumococcal autolysin: a peculiar mechanism of secretion in *Escherichia coli*. *J. Biol. Chem.* 264, 1238–1244
7 Loessner, M.J. *et al.* (1999) Evidence for a holin-like protein gene fully embedded out of frame in the endolysin gene of *Staphylococcus aureus* bacteriophage 187. *J. Bacteriol.* 181, 4452–4460
8 Navarre, W.W. *et al.* (1999) Multiple enzymatic activities of the murein hydrolase from staphylococcal phage φ11. Identification of a D-alanyl-glycine endopeptidase activity. *J. Biol. Chem.* 274, 15847–15856
9 Steiner, M. *et al.* (1993) The missing link in phage lysis of Gram-positive bacteria: gene *14* of *Bacillus subtilis* phage φ29 encodes the functional homolog of lambda S protein. *J. Bacteriol.* 175, 1038–1042
10 Bläsi, U. *et al.* (1999) The C-terminal sequence of the lambda holin constitutes a cytoplasmic regulatory domain. *J. Bacteriol.* 181, 2922–2929
11 Smith, D.L. *et al.* (1998) Purification and biochemical characterization of the lambda holin. *J. Bacteriol.* 180, 2531–2540
12 Graschopf, A. and Bläsi, U. (1999) Molecular function of the dual-start motif in the λ S holin. *Mol. Microbiol.* 33, 569–582
13 Gründling, A. *et al.* (1999) Biochemical and genetic evidence for three transmembrane domains in the class I holin, λ S. *J. Biol. Chem.* 275, 769–776
14 Zagotta, M.T. and Wilson, D.B. (1990) Oligomerization of the bacteriophage lambda *S* protein in the inner membrane of *Escherichia coli*. *J. Bacteriol.* 172, 912–921
15 Raab, R. *et al.* (1988) Dominance in lambda *S* mutations and evidence for translational control. *J. Mol. Biol.* 199, 95–105
16 Johnson-Boaz, R. *et al.* (1994) A dominant mutation in the bacteriophage lambda *S* gene causes premature lysis and an absolute defective plating phenotype. *Mol. Microbiol.* 13, 495–504
17 Bläsi, U. and Young, R. (1996) Two beginnings for a single purpose: the dual-start holins in the regulation of phage lysis. *Mol. Microbiol.* 21, 675–682
18 Cao, G. and Dalbey, R.E. (1994) Translocation of N-terminal tails across the plasma membrane. *EMBO J.* 13, 4662–4669
19 Dalbey, R.E. *et al.* (1995) Directionality in protein translocation across membranes: the N-tail phenomenon. *Trends Cell Biol.* 5, 380–383
20 Chang, C-Y. *et al.* (1995) *S* gene expression and the timing of lysis by bacteriophage λ. *J. Bacteriol.* 177, 3283–3294
21 Reader, R.W. and Siminovitch, L. (1971) Lysis defective mutants of bacteriophage lambda: on the role of the *S* function in lysis. *Virology* 43, 623–637
22 Barenboim, M. *et al.* (1999) Characterization of the dual start motif of a class II holin gene. *Mol. Microbiol.* 32, 715–727
23 Ziermann, R. *et al.* (1994) Functions involved in bacteriophage P2-induced host cell lysis and identification of a new tail gene. *J. Bacteriol.* 176, 4974–4984
24 Schmidt, C. *et al.* (1996) Three functions of bacteriophage P1 involved in cell lysis. *J. Bacteriol.* 178, 1099–1104
25 Hershey, A.D. (1946) Mutation of bacteriophage with respect to type of plaque. *Genetics* 31, 620–640
26 Paddison, P. *et al.* (1998) The roles of the bacteriophage T4 *r* genes in lysis inhibition and fine-structure genetics: a new perspective. *Genetics* 148, 1539–1550
27 Dressman, H.K. and Drake, J.W. (1999) Lysis and lysis inhibition in bacteriophage T4: *rV* mutations reside in the holin *t* gene. *J. Bacteriol.* 181, 4391–4396
28 Kedzierska, S. *et al.* (1996) The *Rz1* gene product of bacteriophage lambda is a lipoprotein localized in the outer membrane of *Escherichia coli*. *Gene* 168, 1–8
29 Zhang, N. and Young, R. (1999) Complementation and characterization of the nested *Rz* and *Rz1* reading frames. *Mol. Gen. Genet.* 262, 659–667
30 Taylor, A. (1971) Endopeptidase activity of phage λ endolysin. *Nat. New Biol.* 234, 144–145
31 Sanger, F. *et al.* (1977) Nucleotide sequence of bacteriophage φX174 DNA. *Nature* 265, 687
32 Atkins, J.F. *et al.* (1979) Binding of mammalian ribosomes to MS2 phage RNA reveals an overlapping gene encoding a lysis function. *Cell* 18, 247–256
33 Model, P. *et al.* (1979) Characterization of Op3, a lysis-defective mutant of bacteriophage f2. *Cell* 18, 235–244
34 Winter, R.B. and Gold, L. (1983) Overproduction of bacteriophage Qβ maturation (A_2) protein leads to cell lysis. *Cell* 33, 877–885
35 Young, K.D. and Young, R. (1982) Lytic action of cloned φX174 gene *E*. *J. Virol.* 44, 993–1002
36 Henrich, B. *et al.* (1982) Lysis of *Escherichia coli* by induction of cloned φX174 genes. *Mol. Gen. Genet.* 185, 493–497
37 Remault, E. *et al.* (1982) Functional expression of individual plasmid-coded RNA bacteriophage MS2 genes. *EMBO J.* 1, 205–209

38 Maratea, D. *et al.* (1985) Deletion and fusion analysis of the ϕX174 lysis gene *E. Gene* 40, 39–46

39 Buckley, K.J. and Hayashi, M. (1986) Lytic activity localized to membrane-spanning region of ϕX174 E protein. *Mol. Gen. Genet.* 204, 120–125

40 Roof, W.D. *et al.* (1994) *slyD*, a host gene required for ϕX174 lysis, is related to the FK506-binding protein family of peptidyl-prolyl *cis-trans*-isomerases. *J. Biol. Chem.* 269, 2902–2910

41 Wülfing, C. *et al.* (1994) An *Escherichia coli* protein consisting of a domain homologous to FK506-binding proteins (FKBP) and a new metal binding motif. *J. Biol. Chem.* 269, 2895–2901

42 Hottenrott, S. *et al.* (1997) The *Escherichia coli* SlyD is a metal ion-regulated peptidyl-prolyl *cis-trans*-isomerase. *J. Biol. Chem.* 272, 15697–15701

43 Mitterauer, T. *et al.* (1999) Metal-dependent nucleotide binding to the *Escherichia coli* rotamase SlyD. *Biochem. J.* 342, 33–39

44 Scholz, C. *et al.* (1999) R73A and H144Q mutants of the yeast mitochondrial cyclophilin Cpr3 exhibit a low prolyl isomerase activity in both peptide and protein-folding assays. *FEBS Lett.* 443, 367–369

45 Witte, A. *et al.* (1998) Mutations in cell division proteins FtsZ and FtsA inhibit ϕX174-protein-E-mediated lysis of *Escherichia coli*. *Arch. Microbiol.* 170, 259–268

46 Bradley, D.E. *et al.* (1969) Structural changes in *Escherichia coli* infected with a ϕX174-type bacteriophage. *J. Gen. Virol.* 5, 113–121

47 Gschwender, H.H. and Hofschneider, P.H. (1969) Lysis inhibition of ϕX174, MS2, and Qβ-infected *Escherichia coli* bacteria by magnesium ions. *Biochim. Biophys. Acta* 190, 454–459

48 Lubitz, W. *et al.* (1984) Requirement for a functional host cell autolytic enzyme system for lysis of *Escherichia coli* by bacteriophage ϕX174. *J. Bacteriol.* 159, 385–387

49 Lubitz, W. *et al.* (1984) Lysis of *Escherichia coli* after infection with ϕX174 depends on the regulation of the cellular autolytic system. *J. Gen. Microbiol.* 130, 1079–1087

50 Witte, A. *et al.* (1997) Proline 21, a residue within the α helical domain of ϕX174 lysis protein E, is required for its function in *Escherichia coli*. *Mol. Microbiol.* 26, 337–346

51 Bernhardt, T.G. *et al.* Genetic evidence that the bacteriophage ϕX174 lysis protein inhibits cell wall synthesis. *Proc. Natl. Acad. Sci. U. S. A.* (in press)

52 Ikeda, M. *et al.* (1991) The *Escherichia coli mraY* gene encoding UDP-*N*-acetylmuramoyl-pentapeptide: undecaprenyl-phosphate phospho-*N*-acetylmuramoyl-pentapeptide transferase. *J. Bacteriol.* 173, 1021–1026

53 Brunskill, E.W. and Bayles, K.W. (1996) Identification of LytSR-regulated genes from *Staphylococcus aureus*. *J. Bacteriol.* 178, 5810–5812

54 Kawarabayasi, Y. *et al.* (1998) Complete sequence and gene organization of the genome of a hyper-thermophilic arachaebacterium, *Pyrococcus horikoshii* OT3. *DNA Res.* 5, 55–76

Horizontal gene transfer and the origin of species: lessons from bacteria

Fernando de la Cruz and Julian Davies

In bacteria, horizontal gene transfer (HGT) is widely recognized as the mechanism responsible for the widespread distribution of antibiotic resistance genes, gene clusters encoding biodegradative pathways and pathogenicity determinants. We propose that HGT is also responsible for speciation and sub-speciation in bacteria, and that HGT mechanisms exist in eukaryotes.

F. de la Cruz is in the Departamento de Biología Molecular (Unidad Asociada al C.I.B.), Universidad de Cantabria, Santander, Spain; J. Davies is in the Dept of Microbiology and Immunology, University of British Columbia, Vancouver, BC, Canada V6T 1Z3.*
**tel: +34 942 201942, fax: +34 942 201945, e-mail: delacruz@medi.unican.es*

The old topic of horizontal gene transfer (HGT) has become fashionable[1]. From the overwhelming surge of genome sequence information, more and more candidates for horizontally transferred genes are being identified: among prokaryotes[2]; from bacteria to eukaryotes[3]; from bacteria to archaea[4]; from animals to bacteria[5]; and so on. It is clear that genes have flowed through the biosphere, as in a global organism. HGT, once solely of interest for practical applications in classical genetics and biotechnology, has now become the substance of evolution. However, amid this new-found enthusiasm for an old subject, the existence of HGT in eukaryotes is viewed with skepticism, as there is a distinct lack of mechanistic studies that would help to explain such a broad-host-range process (i.e. one that operates not only between species and genera, but also between kingdoms). In this particular case it is apparent that, once again, we can learn from our vast experience with prokaryotes, especially as typical methods of carrying out HGT in prokaryotes, such as transfection or even bacterial conjugation[6,7], can be used for the genetic manipulation of eukaryotic cells. There are a number of well documented examples in which bacterial adaptation has been influenced by HGT. Studying these examples can provide an appreciation of the general mechanistic trends involved, as well as providing the foundation for considerations of a ubiquitous role of HGT in shaping modern eukaryotic species.

 PII: S0966-842X(00)01703-0