Phage lysis: do we have the hole story yet?
Ry Young

In infections of Gram-negative bacteria, lysis is a three step process, with a choice of two effectors for each step. At a precise, allele-specific time, the inner membrane (IM) is fatally permeabilized by either a holin or a pinholin. This allows a muryalytic enzyme, either a canonical endolysin, escaping from the cytoplasm, or a SAR endolysin, activated in the periplasm, to degrade the peptidoglycan. Surprisingly, a third class of lysis protein, the spanin, is required for disruption of the outer membrane (OM). Key steps are regulated by membrane protein dynamics, both in terms of bilayer topology and subcellular distribution, by the energization of the membrane, and by holin-specific inhibitors called antiholins.

Addresses
Department of Biochemistry and Biophysics, Center for Phage Technology, 2128 TAMU, Texas A&M University, College Station, TX 77843-2128, United States

Corresponding author: Young, Ry (ryland@tamu.edu, ryland.young@gmail.com)

Introduction
In the infection cycle of double-stranded DNA phages, the lysis pathway involves at least two proteins, a small membrane protein called the holin and a muryalytic enzyme called the endolysin [1]. The holin accumulates harmlessly in the membrane until, at an allele-specific time, ‘triggering’, a term that stands for fatal permeabilization of the bilayer that allows the endolysin to escape from the cytoplasm and attack the peptidoglycan. Triggering occurs prematurily if the infected cells are subjected to conditions that cause a sudden decrease in the proton motive force (PMF); that is, energy poisons, membrane damage by external agents, or sudden anaerobiosis. Historically, the fatal membrane lesions formed by holin triggering have been called ‘holes’, as opposed to channels or pores.

The main goal here is to highlight progress in the past few years that, at least for phages of Gram-negative hosts, has advanced our understanding of this event, which is, numerically, the most common fate of earthly cells. At the outset, it must be noted that this overview devotes no space to consideration of lysis in Gram-positive bacteria, where most attention has been focused on the biotechnological utility of phage endolysins [2]. Also omitted is consideration of the ‘single gene lysis’ systems of small ssDNA and ssRNA phages, in which a single phage protein channels the host into autolysis [3].

The lambda lysis paradigm
Lysis genes are often found in ‘cassettes’ (Figure 1a), and in the case of the best-studied lysis systems, phage lambda and the lambdoid phage 21, four genes producing five protein products are clustered immediately downstream of the single late gene promoter. For both, the S gene encodes the lysis effector, the lambda holin S105 and the phage 21 pinholin, S2168, that triggers for lethal hole formation; the protein names reflect the length in AA residues [4,5] (Figure 1b). In both cases, an upstream translational start is used for production of an antiholin (S107; S2171). The negative-dominant character of each antiholin form is due to an N-terminal positively charged residue that confers a PMF-sensitive block on a topological change of the first transmembrane domain (TMD1) required for lytic function (Figure 1c). The last genes in both cassettes encode the spanin subunits (see below) [6].

Lambda lysis has been conveniently characterized by thermal induction of a lambda prophage, after which, under strictly defined conditions, lysis of nearly all the cells of the culture occurs abruptly at 50 min, or some 42 min after the turn-on of late gene expression [7]. Other lysis systems have been studied in this convenient system by substituting genes for heterologous holins, endolysins, and spanins or even substituting whole lysis cassettes [5,8–10].

A unified model for phage lysis in Gram-negative hosts
In the past few years there have been multiple studies using fluorescence, transmission, and cryoelectron microscopy that have profoundly affected our view of the lysis pathway [11,12,13]. The cartoon in Figure 2 attempts to summarize in outline a coherent model of the two different molecular strategies by which the temporal scheduling of lysis is effected. In the left and right panels are depicted canonical holin–endolysin lysis, as exemplified by lambda, and the pinholin–SAR endolysin pathway, exemplified by phage 21, respectively. When late gene expression commences, the holin S105 of phage lambda and the pinholin S2168 begin to accumulate in the membrane, each as homodimers or as heterodimers with
Lysis cassettes and holin topology. (a) The lysis cassettes of phages λ and lambdoid phage 21 are shown. In both lambdoid phages, the lysis genes are proximal to the late promoter, pR'. Lengths of primary translation products are shown above each gene; upper and lower numbers for S and S' respectively. Gene color codes: turquoise = canonical holin, red = pinholin; black = canonical endolysin (R transglycosylase); dark blue = SAR endolysin; brown = i-spanin; green = o-spanin. Same color in two genes indicates sequence similarity. (b) Dual start motifs for λ holin and 21 pinholin genes. Star codons for each product indicated by star; red = antiholin, green = holin (S105) or pinholin (S'168). Red and green rectangles indicate Shine–Dalgarno sequences. Inverted arrows indicate RNA stem-loops that control choice of start codons. (c) Topological dynamics of holins and pinholins. In the energized membrane (PMF = ~180 mV), the TMD1s of the antiholin forms (S107 and S'171) are inhibited from entering (S107) or exiting (S'171) the bilayer, whereas TMD1 of the pinholin (S'168) exists spontaneously during pinhole pathway. Panels (a) and (c) adapted from Wang et al. [49]. Panel (b) adapted from Young [50].

their cognate antiholins. These proteins continue to accumulate harmlessly, freely mobile in the cytoplasmic membrane, while virion assembly is proceeding in the cytoplasm. In both cases, when the holins and pinholins reach an allele-specific critical concentration, suddenly nucleation and rapid formation of large two-dimensional aggregates, or ‘rafts’, occurs, based on studies with holin–GFP and pinholin–GFP fusions (Supp. Figure 1). Rafts are very large and number only a few per cell for the canonical lambda holin. In contrast, the rafts in the pinholin situation are smaller and more numerous. Raft formation is tightly coupled with triggering and hole formation.

S-holes
It has long been known that the S105 lesions were large and non-specific, as attested by to the ability of S105 (and other holins [9,10,14–16]) to support lysis with heterologous endolysins and, more dramatically, with hybrid endolysin-β-galactosidase fusions of ~0.5 MDa native mass [17]. However, it was not until recently that cryo-EM studies and tomography revealed the ‘S-holes’ to be micron-scale interruptions in the IM, averaging >340 nm in diameter and ranging to ~1 μm, but small in number, 1–3 per cell (Supp. Figure 2; upper panel) [11]. These massive lesions allow the endolysin R to escape from the cytoplasm and attack the peptidoglycan. The model is supported by the lysis morphology, as observed in phase-contrast videomicroscopy (Supp. Figure 2; lower panels) [18**]. Induced lysogens are seen to elongate until suddenly undergoing a ‘blow-out’ at a localized region; the cell violently empties of refractile material, including the progeny virions, while at least momentarily maintaining rod shape. It is thought that the localized disruption reflects the huge holin lesion and the resultant massive degradation of the peptidoglycan opposite the hole (Figure 2, left).
**Pinholes and SAR endolysins**

In the pinholin/SAR endolysin pathway, where after triggering the rafts were found to be smaller and more numerous, cross-linking, cysteine-accessibility and modeling studies have indicated that the entire pinholin population of ~7000 molecules participate in the formation of ~10³ heptameric ‘pinholes’, estimated to have a ~2 nm diameter (Supp. Figure 3) [15**,19**]. Lesions of this size are clearly too small to allow transit of mature endolysins. Instead, pinholins require a SAR endolysin secreted by the host sec system [20]. SAR stands for signal-anchor-release; SAR domains are N-terminal TMDs (or signal anchors) that serve to engage the sec pathway and then anchor the exported endolysin to the bilayer in a membrane potential-dependent fashion (Figure 3). Premature degradation of the murein is avoided because the tethered enzyme is inactive. Depolarization of the membrane by the pinholin causes quantitative release of the tethered SAR endolysin molecules, which then refold into an enzymatically active form, leading to generalized murein degradation. In consequence, the pinholin–endolysin system has a distinct morphological signature observable by phase-contrast videomicroscopy: after pinholin triggering, the cell begins to shorten and round up before suddenly exploding (Supp. Figure 2) [18**]. Presumably this reflects the uniform dispersal of the SAR endolysin within the periplasm, such that murein degradation is not localized to the region opposite a massive endolysin lesion but is instead distributed throughout the cytoplasm after release and activation (Figure 2, right).

**Regulation of lysis**

The operational cartoon in Figure 2 shows the common strategy but different tactics by which holins and pinholins control the timing of lysis. Underlying this simplified operational scheme is a remarkable pattern of membrane protein dynamics at the level of bilayer topology and subcellular distribution.

**Diverse gymnastics of SAR endolysins**

The molecular basis of SAR endolysin regulation has been studied for three different phages, with crystal structures determined for the inactive and active forms of two of these (Figure 3) [20–23]. All three systems have
SAR endolysins. (a) Structures of inactive, membrane-tethered (left) and active, released R21 (right). SAR domain is modeled as TMD in structure of inactive form. In this activation event, the catalytic triad Glu35-Asp44-Thr50 is assembled when SAR domain folds up against the body of the protein. Adapted from Sun et al. [22]. (b) General regulation of SAR endolysins. These enzymes are catalytically inactive when tethered to the membrane. They release from the membrane spontaneously at a slow rate but quantitatively if the PMF collapses. Release of the SAR domain causes refolding to the active state. (c) SAR domains (yellow highlights) cause different mechanisms of activation. SAR endolysins are homologs of the canonical lysozyme T4 E, with the ExxD/Cx5T catalytic triad (highlighted in blue). SAR domains have a high proportion of Gly/Ala/Ser residues, at the expense of Leu, compared to normal TMDs [20]. SAR domains that include a Cys residue (red) cause disulfide bond isomerization after release, resulting in liberation of a catalytic Cys (Lyz of phage P1) [21] or uncaging of the catalytic Glu (Lyz of phage ERA103) [22]. R21 activation is non-covalent (panel A). The activation of R21[550W], the enzyme that causes release of Shiga toxin from E. coli O157:H7 cells [51], is under investigation (Q. Sun, G. Kuty, T. Pang, J. Sacchettini and R. Young, unpublished).

Different regulatory schemes, including dramatic conformational rearrangements and two different pathways of disulfide bond isomerization, all consequent to the escape of the SAR domain from the bilayer. SAR domains are always N-terminal TMDs with a relatively hydrophilic character (reduced content of Leu residues compensated by enriched Gly/Ala/Ser). Moreover, SAR domains exit the bilayer spontaneously at a specific low rate but are instantly released when the PMF is collapsed.

Topological changes in the pinhole pathway

The topological behavior of the S21 pinholin is even stranger, as revealed by recent genetic, biochemical, and high resolution real-time fluorescence microscopy studies [13**,19**,24–27]. The pinholin accumulates as a dimer in which both TMD1, which is a SAR domain, and TMD2 of each monomer are stabilized in the membrane by intramolecular and intermolecular helix–helix interactions (Supp. Figure 2). It is not until both TMD1s spontaneously exit the membrane and then interact as aqueous-phase helices that the pinholin dimer is activated for the raft formation pathway. Thus TMD1 is an intrinsic intramolecular inhibitor of lysis while in the bilayer but a positive effector once relocalized to the periplasm, whereas TMD2 is the actual pinhole-former. The pinholin product, S21’71, has an extra positively charged residue at the N-terminus that inhibits escape of TMD1 from the bilayer (Figure 1). Thus, S21’71, if dimerized with the pinholin S21’68 form, creates an inactive heterodimer and slows the ‘lysis clock’.

The lambda antiholin

Opposite topological dynamics underlie the inhibitory role of S107, the lambda antiholin. In this case, the extra positive charge at the N-terminus of the antiholin prevents TMD1 from entering the membrane, giving rise to inactive S105:S107 heterodimers, again slowing the lysis clock (Figure 1) [28–32]. In both cases, the depolarization that occurs at the instant of triggering removes a barrier for the topological change of TMD1 (entry for S107, exit for S21’71) and thus converts the inactive holin–antiholin heterodimers into fully active molecules. Since in normal conditions, the holin–antiholin ratio is ~2:1, and the antiholin preferentially heterodimerizes with the holin, the number of active holin or pinholin dimers would triple at the instant of triggering.

It must be emphasized that the dynamics of these antiholin systems, however elegant, are not the key to lysis timing in either the holin or pinholin pathways. Precisely timed triggering still occurs in mutants in which the antiholin forms are not produced, albeit a few minutes earlier. RNA secondary structures (Figure 1) control the ratio of holins (or pinholins) to antiholins, and mutations that alter these structures such that the antiholin form predominates are not only non-lytic alleles but also have dominant character (which is how they were discovered) [4,5,7,33,34]. There is some evidence that host factors may regulate the efficiency of expression of the lambda holin gene or the holin–antiholin ratio by affecting the formation of these structures (K. Nam, ‘Translational regulation of the S gene of bacteriophage lambda’, Ph.D. thesis, Texas A&M University, 1991) [33].

Triggering and the critical concentration

A lynchpin of the current model is that the holins and pinholins exhibit critical concentration behavior for raft formation, based on the precedent of the well-studied bacteriorhodopsin system (Figure 2; Supp. Figure 4) [35,36]. After induction, mobile bacteriorhodopsin monomers accumulate in the membrane until, at a critical
concentration, beginning the assembly of the ‘purple membrane’ array. Missense mutations at intermolecular interfaces can raise or lower the critical concentration significantly. Presumably, similar effects explain the phenotypic variance seen in genetic studies of lambda S, which, although the wt allele triggers at 50 min, can give rise to missense mutants triggering as early as 11 min or as late as 90 min [7,37] (K. To, R. White, J. Dewey, C. Dankenbring, R. Young, unpublished).

Also implicit in this scheme is the notion that the holin or pinholin rafts spontaneously convert to holes (or pinholes). Possible clues to this behavior come from studies on the pinhole pathway [13**,27]. Genetics and chemical probing studies indicated that in the pinhole, TMD2 of the pinholin has a more hydrophilic surface facing the lumen, whereas in the pre-triggering dimer, the same hydrophilic faces are sequestered against each other and the still-integrated TMD1. It is proposed that raft formation involves intimate helical packing, resulting in a lipid-depleted two-dimensional aggregate that leaks ions along the interfaces of the hydrophilic surfaces (Supp. Figure 3). Local reduction in the membrane potential could stimulate reorientation of these surfaces to allow hydration, more ion leakiness, and thus a concerted shift to the pinhole arrangement. A few local pinholes could then lead to more depolarization and, ultimately, total collapse of the PMF.

Although short on specifics, this scheme accounts for the universal holin susceptibility to premature triggering by anything that reduces membrane energization. S105 has two TMDs, 1 and 3, with relatively hydrophilic faces; experiments are in progress to test if these faces line the lumen of the S-hole (K. To, unpublished).

Spanins

One might expect that once the murein has been degraded, lysis and dispersal of the viral progeny would be automatic. It was perhaps not too surprising to discover that this was not true for phages of the mycolata, which encode, in addition to the holin and endolysin, an enzyme, LysB, to degrade the murein esters of the waxy outer membrane-equivalent [38–40]. However, it was definitely unexpected to find that, even after destruction of the peptidoglycan, lysis of Gram-negative hosts requires yet another functional class of protein, a spanin, for disruption of the outer membrane (OM) [18**]. Moreover, two completely different types of spanins, two-component spanins and unimolecular spanins, have been identified [6].

Genes like Russian dolls

The last genes of the lambda lysis cassette are also the strangest: Rs, which encodes an integral IM protein, and Rz1, which is wholly embedded in the +1 reading frame of the Rs gene and encodes an OM lipoprotein (Figure 1). Recently phase contrast videomicroscopy has revealed that both genes are required for lysis [18**]. In the absence of Rs or Rz1 function, the lysis process terminates in spherical cells lacking the cell wall and bounded by the outer membrane (OM) (Supp. Figure 2). At the molecular level, the Rz and Rz1 proteins were shown to be the subunits of a complex that spans the entire periplasm, connecting the IM and OM, and thus designated as a spanin (Figure 4) [41]. Rz is an 153 AA integral membrane protein, or i-spanin, with an N-terminal domain and a periplasmic domain shown to be both highly helical and subject to coiled-coil oligomerization [6,42]. Mature Rz1 is a 40 AA OM lipoprotein, or o-spanin, with a proline content of 25%. Co-precipitation experiments have shown that the Rz-Rz1 spanin complex assembles in the envelope during the morphogenesis period, via C-terminal interactions [41]. To add to the litany of unique features, both Rz and Rz1 have recently been shown to accumulate as covalent homodimers, linked by three intermolecular disulfide bonds (two in the Rz dimer, one in the Rz1 dimer) [43]. Thus the spanin complex is actually an Rz2:Rz12 tetramer. Oddly, complementation studies have shown that spanin function requires that at least one of the two subunit homodimers is covalently linked, although it does not matter which one.

Spanin function

The interesting questions are: (1) how does the Rz2:Rz12 spanin complex function to disrupt the OM? (2) How is this regulated, in view of the finding that the complexes accumulate throughout the morphogenesis period. Figure 4 shows the current model, partly based on biochemical studies with the soluble periplasmic domains of the spanin subunits and partly conceived as shavings from Ockham’s Razor [18**]. The regulation is thought to be inherent in the fact that since the complexes span the periplasm, they are trapped within the lacunae of the cross-linked peptidoglycan. After holin triggering and cell wall degradation by the endolysin, the spanin complexes are free to diffuse laterally and form aggregates, providing the free energy for a conformational change that disrupts the OM. The nature of the disruption is unknown but the simplest notion is that the oligomerized complexes bring the surfaces of the OM and IM together and cause membrane fusion.

The membrane fusion model arose from the kernel ideas that there has to be some reason for the spanins to connect the IM and OM and that there is no evidence of enzymatic activity or catalytic domains. The idea was indirectly supported by the discovery that many phages, including the iconic coliphage T1 (Figure 1), an entirely distinct type of spanin, the unimolecular or u-spanin (Figure 4), in which features of the i-spanin and o-spanin subunits are combined and which complements an RsRz1 double defect [6]. The prototypic u-spanin, T1 gp11, is an OM lipoprotein, so that its N-terminal lipoylated Cys is exported by the Lol system to the inner leaflet of the
6 Growth and development: prokaryotes

Figure 4

Spanin function and gene architecture. (a) Holin-endolysin function in the absence of spanin function results in spherical cells bounded by the OM. Cryo-EM image of E. coli cell in which SR genes have been expressed, after triggering. Note inset: shows efflux of cytosolic material from S-hole. Scale bar is 500 nm. From Dewey et al. [11]. (b) Model for spanin structure and function. Two component spanin complexes consist of an integral IM subunit, the i-spanin (Rz) and an OM lipoprotein, the lipoprotein o-spanin (Rz1), that interact via a C-terminal/C-terminal interaction, spanning the periplasm through the meshwork of the peptidoglycan (PG). Alternatively, a unimolecular spanin, u-spanin (T1 gpT1), which is an OM lipoprotein but also has a C-terminal TMD, spans the periplasm as a covalent chain. The Rz/Rz1 complex is shown as a heterodimer for simplicity; in vivo, each subunit is actually a disulfide-linked homodimer. The current model is that the spanin complex undergoes oligomerization when freed from the meshwork of the PG, providing free energy for a collapsing conformational change that brings the two membranes together (only the collapsing change is shown). In the two-component system, the oligomerization is due to the high coiled-coil propensity of the highly α-helical periplasmic domain of the i-spanin. For the u-spanin, the collapsing change probably involves beta sheet formation. The resultant adjacency of the bilayers and the fusogenic properties of the spanin periplasmic domains are thought to lead to membrane fusion, thus disrupting the OM. (c) Spanin gene architecture. Two component spanin genes (red = i-spanin; black = o-spanin) are highly diverse and are found in three architectures, each with multiple, dissimilar sequences: embedded (o-spanin gene entirely within i-spanin gene) (e.g., T7; also see Figure 1a); overlapping (o-spanin gene extends beyond end of i-spanin gene; e.g., P2); and separated (genes completely separated; e.g., T4). U-spanin (red-black hatch; T1) genes are also diverse. Panels b and c adapted from Summer et al. [5].

OM [44], but it also has a C-terminal TMD embedded in the IM. Besides the different topology, the periplasmic domain of gpI1 is predicted to be mostly beta-sheet and has no Cys residues. Thus the u-spanin is clearly a completely unrelated solution to the same operational task. The parallel notion here is that the gpI1 protomers, once liberated from the meshwork of the peptidoglycan by endolysin action, undergo oligomerization that provides the free energy for a collapsing conformational change and membrane fusion.

Definitive evidence addressing the membrane fusion model should soon be available. In any case, whatever the mode by which these spanin proteins disrupt the OM, the genetic diversity involved turned out to be quite profound [6]. There are not only many unrelated families of embedded spanin genes arranged like Rs/Rs1 (Figures 1 and 4), but also multiple families in which the o-spanin gene extends beyond the end of the i-spanin gene (overlapped) and also where the two genes are entirely distinct (separated). Add to this mix the existence of multiple u-spanin genes like T1 I1 and one is left with an evolutionary puzzle that should prove intriguing to theorists.

Conclusions and perspective
Considerable progress has been made in fleshing out the operational aspects of the lysis pathway in Gram-negative hosts. In overview, the pathway has three steps, each with two fundamentally different options: timed triggering by the holin or pinholin, cell wall degradation by the endolysin or SAR endolysin, and then OM disruption by the i-spanin/o-spanin complex or the u-spanin. There are thus eight possible combinations for these three steps, but only six are functionally compatible, considering that pinholins require SAR endolysins. Remarkably, all six have been found in nature or shown to function in the laboratory. One combination, comprising a pinholin, SAR endolysin, and i-spanin/o-spanin ensemble, has been shown to be responsible for the release of Shiga-toxin in EHEC infections [45], so there is even some clinical
relevance to this story. More information is needed about the structure of the S-holes, specifically, and in the holes formed by canonical holins in general. Beyond this, technological innovation, possibly using planar lipid bilayer systems, will be required to get at the specifics of the current model, especially the interactions of the holin proteins with the energized membrane. Also needed are studies directly addressing the proposed membrane-fusion model for spanin function. If the general idea is validated, the spanin systems should be regarded as simple, genetically flexible platforms for studying the membrane fusion process that underlies so many key biological processes. Finally, space limitations have precluded discussion of new developments in the oldest experimental system in molecular genetics, T4 lysis-inhibition, or LIN. For interested readers it is worth noting that, besides its historical status, LIN is the only case where it is known that environmental information is used for real-time, antiholin-mediated control of holin triggering [46–48].

Acknowledgements
This work was supported by National Institutes of Health grant NIGMS27099 to the author, as well as the Sadie Hatfield Professorship in Agriculture at Texas A&M University. The author is grateful to all of the students, post-doctoral scientists, and faculty collaborators for their many consequential insights into this fundamental biological phenomenon over more than 35 years.

Appendix A. Supplementary data
Supplementary material related to this article can be found, in the online version, at http://dx.doi.org/10.1016/j.mib.2013.08.008.

References and recommended reading
Papers of particular interest, published within the period of review, have been highlighted as:

** of outstanding interest

19. First systematic videomicroscopy of lysis morphology and evidence for spanin essentiality.
21. Discovery of SAR sequences, a signal conferring dynamic membrane topology.
8 Growth and development: prokaryotes


47. Tran TA, Struck DK, Young R: The T4 RI antiholin has an N-terminal signal anchor release domain that targets it for degradation by DegP. J Bacteriol 2007, 189:7618-7625.


