Recent Progress in Bacteriophage Research

Edited by
Graham F. Hatfull

Printed Edition of the Special Issue Published in *Viruses*

MDPI

Graham F. Hatfull (Ed.)

Recent Progress in Bacteriophage Research

MDPI

This book is a reprint of the special issue that appeared in the online open access journal *Viruses* (ISSN 1999-4915) in 2012 (available at: http://www.mdpi.com/journal/viruses/special_issues/bacteriophage).

Table of Contents

Preface

Nearly one hundred years after the discovery of bacteriophages, these remarkable viruses continue to fascinate; providing new biological insights of general importance, and generating new tools for molecular biology. Although phage investigations waned after about sixty years of study, they have been embraced with renewed enthusiasm, fueled by appreciation of their broad roles in ecology and geochemistry, genomics, and the search for a response to antibiotic resistant bacterial pathogens. The conclusion that phage particles represent the majority of all biological entities in the biosphere, that the population is dynamic and old, and is highly diverse genetically, suggests a deep richness that warrants investigation from a multitude of different perspectives.

In this volume, some of these new topics in bacteriophage research are discussed. The variety of contributions reflects the broad areas currently under study. Several chapters discuss the nature of bacteriophage populations in their natural environments, and explore the details of interactions between the phages and their hosts, including bacterial biofilms. With the realization that complex microbial biofilms are common throughout the biosphere, it is of interest how phages engage with these, and how phages could be exploited to develop more effective antimicrobial therapies.

The dynamic nature of bacterial and phage populations suggests that nearly four billion years of microbial sparring has driven the development of numerous strategies for phage resistance, and coevolution of phages to access new hosts. One of the most exciting recent discoveries is the Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR) elements present in many bacterial strains that provide immunity to phage infection (as discussed by Fineran and colleagues in this volume). These are not only of mechanistic interest, but also of potential use as general tools in molecular biology.

Most bacteriophages have well-defined host ranges that vary from a single strain to different (but closely related) genera. Host range necessarily imposes a barrier to genetic exchange, and it is thus useful to consider groups of phages as they pertain to specific hosts. Several chapters in this volume follow this approach, and groups of phages that infect *Staphylococci*, the dairy bacteria, and some coliphages; the latter focusing on newly discovered *Leviviruses*.

Several prototype phages have been the subjects of intense investigations for many years, including λ and its relatives, and T4 and its relatives. Although well studied, these continue to yield new and exciting insights, and contributions in this volume discuss how λ interacts with its receptor, and regulatory aspects of DNA replication and gene expression.

The therapeutic application of bacteriophages remains a current topic of interest, and there are several specific types on infections that would seem well suited to such an approach. Wound infections and burns provide some examples where topical applications of phages could control bacterial infections, especially where antibiotic resistant strains are involved. The development of phages targeted to control *Pseudomonas* infections is presented in this volume.

Bacteriophage studies clearly represent an active area of exploration, with recent progress at both the molecular and population levels. With a vast and genetically diverse population to explore, we can anticipate continued progress and the emergence of many new exciting stories in the years to come.

Graham F. Hatfull
Guest Editor

Review

Understanding Bacteriophage Specificity in Natural Microbial Communities

Britt Koskella * and Sean Meaden

BioSciences, University of Exeter, Cornwall Campus, Tremough, TR10 9EZ, UK;
E-Mail: sm341@exeter.ac.uk

* Author to whom correspondence should be addressed; E-Mail: B.L.Koskella@Exeter.ac.uk;
Tel.: +44-(0)-1326-371-878.

Abstract: Studying the coevolutionary dynamics between bacteria and the bacteriophage viruses that infect them is critical to understanding both microbial diversity and ecosystem functioning. Phages can play a key role in shaping bacterial population dynamics and can significantly alter both intra- and inter-specific competition among bacterial hosts. Predicting how phages might influence community stability and apparent competition, however, requires an understanding of how bacteria-phage interaction networks evolve as a function of host diversity and community dynamics. Here, we first review the progress that has been made in understanding phage specificity, including the use of experimental evolution, we then introduce a new dataset on natural bacteriophages collected from the phyllosphere of horse chestnut trees, and finally we highlight that bacterial sensitivity to phage is rarely a binary trait and that this variation should be taken into account and reported. We emphasize that there is currently insufficient evidence to make broad generalizations about phage host range in natural populations, the limits of phage adaptation to novel hosts, or the implications of phage specificity in shaping microbial communities. However, the combination of experimental and genomic approaches with the study of natural communities will allow new insight to the evolution and impact of phage specificity within complex bacterial communities.

1. Introduction

Whether found in the soil [1,2], the leaf [3,4], the ocean [5,6], or the human body [7,8], microbial communities are proving to be more dynamic and diverse than could have been predicted. This incredible diversity is seen both within environments and among environments [7], but how it is generated and maintained is unclear. Early ideas about microbial diversity posited that "Everything is everywhere, but the environment selects [9]." However, this view of the microbial world remains hotly debated [10]. Much of the work testing this tenet has focused on the abiotic environment, such as salinity [11] or soil types [12], but the biotic environment is likely to be just as, if not more, important in shaping selection on microbial populations. Bacteria in any given environment face strong selection pressures from other microbes, predators, viruses, and in the case of bacteria living within another organism, the host immune response. In contrast to the abiotic environment, biotic "environments" have the potential to evolve in response to any changes in the microbial community, making them highly dynamic and capable of driving divergence among populations [13,14]. Bacteriophages (phages) represent perhaps the most ubiquitous of these biotic drivers [15–17]. To understand the role of phage-mediated selection in generating diversity, however, we need good insight into how specific phages are to their bacterial hosts.

For phages to alter the composition of a microbial community, there must exist a degree of specificity such that some hosts are more resistant to local phages than others or are better able to respond to phage-mediated selection. There is clear evidence that not all bacteria are infected by all phages, and indeed that most phages can only infect a subset of bacterial species (Table 1; [18]), but our understanding of phage host range is far from complete. Can phages easily adapt to infect new bacterial types as they become common? Can the same phage lineage shift from one bacterial species to another? These questions are far from new [19], but the development of recent techniques and the power of comparative genomics are moving us towards more satisfying answers. Experimental evolution (Box 1) provides one powerful approach to address these knowledge gaps, as the bacteria-phage interaction can be observed in the absence of other abiotic or biotic selection pressures; as such it has offered key advances in our understanding of the evolution of bacterial resistance to phages and reciprocal adaptations of phages to overcome such resistance. However, there are many reasons that the outcome of coevolution in a test tube might not be predictive of coevolution in nature, given the added biotic and abiotic complexity of most microbial ecosystems. For example, although experimental evolution studies have almost exclusively focused on phage adaptation within a population of one, or at most a few, bacterial species, most bacterial communities are highly rich, and therefore most bacterial species to which phage are adapting are rare. A comparison of culture-independent sequencing studies of microbial communities from the leaf surface, soil, atmosphere and the human body shows that the most dominant species in each given community represents a mere 2%–5% of sequences [20]. Given the heterogeneity and diversity of these microbial communities, it is unclear how a phage with narrow host range could evolve and be maintained.

For virulent bacteriophages, *i.e.*, those that reproduce within and then lyse their host cells, success depends on the chance event of encountering a susceptible host cell in the environment, and is most certainly reduced as a function of community diversity, dispersal, and exposure to the harsh conditions outside of a cell. It makes intuitive sense, therefore, that those phages with a larger host range should be at an inherent advantage. The data gathered so far, however, do not clearly support this intuition. First, many phages seem to be specific to a single bacterial species, and are often specific to only a few strains within that species [21–23] (Table 1). Second, there is building evidence that phages are "locally adapted" to their bacterial hosts [3,24], indicating a degree of specialization to common bacterial strains or species in a given population. Third, although phages do tend to increase their host range during the initial stages of coevolution, there is evidence that this expansion is short-lived [25]. The underlying mechanisms of phage infectivity and bacterial resistance are of course key to the evolution of phage host range [22,26], and have been the focus of extensive review elsewhere [21,27]. The data make it clear that phage infectivity is a complex function of adsorption [28], structural change of both host and phage [29], transport of nucleic acid into the bacterial cell, and avoidance of degradation once inside the cell [30], and is thus a result of both phage and bacterial phenotype. In addition, host susceptibility/resistance to phage can be determined by plasmids hosted by the bacterial cell [31,32]. This form of phage-plasmid interaction could lead to broad phage host range due to horizontal transfer of the plasmid among bacterial species within a community. A more thorough understanding of these interacting mechanisms will allow us to better predict the potential for host range expansion/contraction and therefore the effect of phages on microbial communities under both natural and therapeutic settings.

The more general question of why parasites specialize is of course not specific to bacteria-phage interactions, and we can apply much of the current coevolutionary theory to understanding the evolution of phage specificity. Many phages act as obligate parasites, as they are both unable to reproduce outside of their host cells and require cell lysis to transmit, thus killing their hosts. However, we acknowledge that, although virulent phages are obligate killers, other phages integrate into the host genome and their fitness relies on host reproduction. In these cases, the acquisition of a prophage can confer beneficial phenotypic change to the bacterial hosts, and therefore this latter relationship acts more synergistically than antagonistically. In either case, the question of host range for phages that are in the lytic cycle, and being transmitted among cells can be broken into two parts: first, the specificity of host resistance against infecting parasites; and second, the specificity of parasites on different hosts. It is often difficult to tease these two processes apart, but a recent review suggests that the failure to infect "nonhost" species (*i.e.*, those not considered to be hosts for the pathogen in question) may be the result of pathogen evolution leading to specialization on its own source host species and not the result of host evolution for resistance [33]. By reviewing studies across many host-parasite systems, the authors find a general trend towards decreasing parasite infection success on hosts of increasing genetic distance from the focal host. It remains to be determined whether this pattern is ubiquitous for bacteria-phage interactions, especially given the broad host ranges of some phages (Table 1).

The most supported evolutionary argument for why parasites specialize on given hosts, despite the clear advantage of a broad host range, is that there exists a trade-off between fitness and the breadth of parasite infectivity or host resistance [34]: In other words, the idea that "a jack of all trades is a master of none." This can be explained either by antagonistic pleiotropy, a situation where an adaptation that is advantageous in one host is deleterious in another, or else by selection for a less efficient but more general mechanism of infection. Support for this trade-off has been found for phage $\phi 2$, where individuals with broader host range within populations of its bacterial host, *Pseudomonas fluorescens*, were shown to pay a cost for this increased breadth relative to phages with narrow host ranges [35]. Specific evidence of antagonistic pleiotropy has also been found; during experimental host range expansion of phage $\phi 6$, spontaneous mutants able to infect novel hosts were found to be less infective to their native hosts in seven out of nine cases [36]. These trade-offs are also likely to be common in host populations. Indeed, recent results from experimental evolution (see Box 1) of *Prochlorococcus* hosts and their associated phages demonstrate that resistance to one phage genotype often came with the added cost of increased susceptibility to another phage genotype [37]. Similarly, experimental evolution of *P. syringae* in either single phage or multiple phage environments shows that bacteria evolved with multiple phages paid a higher cost of resistance than those evolving with single phages [38].

Box 1. Experimental evolution of phage specificity.

Experimental evolution of bacteriophage specificity has offered some key insights into the underlying process, evolutionary consequences, and fitness costs of host range expansion. The power of this method is that it allows replicate lines, started with genetically identical phages, to be passaged on homogeneous or heterogeneous hosts populations under a range of conditions (such as density and resources) for many thousands of generations (Figure 1). During this time, both the host bacterium and the phage can be frozen in time and resurrected at the end of the experiment, at which point [39] the fitness of evolved phages can be compared directly to both the ancestral types and phages experiencing a different selection regime.

This approach has been used to demonstrate a number of key features of phage specificity, and has gone some way in explaining both when and how phage host range is likely to expand. First, in terms of range expansion within a host population (*i.e.*, the evolution of "generalist" phages capable of infecting more genotypes of a given bacterial species), there is evidence that phage $\phi 2$ is more likely to increase its host range during experimental coevolution with its bacterial host, *P. fluorescens*, than when the bacterial population is held constant [35]. Similar results were found in coevolving populations of phage SBW25$\phi 2$ and *P. fluorescens* [40]. Furthermore, the emergence of evolved "generalist" phenotypes of both bacteria and phages during experimental coevolution has been demonstrated both in a marine cyanobacteria and cyanophage system [5] and in a *P. fluorescens* and phage SBW25$\phi 2$ system [41]. Second, in terms of host range expansion to novel hosts, experimental evolution has provided evidence that phage $\phi 6$ populations are more likely to evolve expanded host range when there is strong competition for hosts (*i.e.*, when the focal host is rare in a population) [42].

It has also been shown that during the early stages of such a host shift, the likelihood of successful adaptation to a novel host is increased when contact with the native host is maintained, as this prevents extinction of the phage [43].

Figure 1. Illustrative example of experimental evolution of phage host range, where: (**A**) independent lines of genetically identical phage populations are propagated under different treatment regimens (e.g., different bacterial host species); and then (**B**) tested for infectivity on focal and alternate hosts. (**C**) Outcomes of these experiments might be a directional change towards increased host range over time (a), an initially increasing but then stable host range, perhaps indicative of coevolutionary response by the host population (b), or a decrease in host range associated with antagonistic pleiotropy during specialization on the focal host (c).
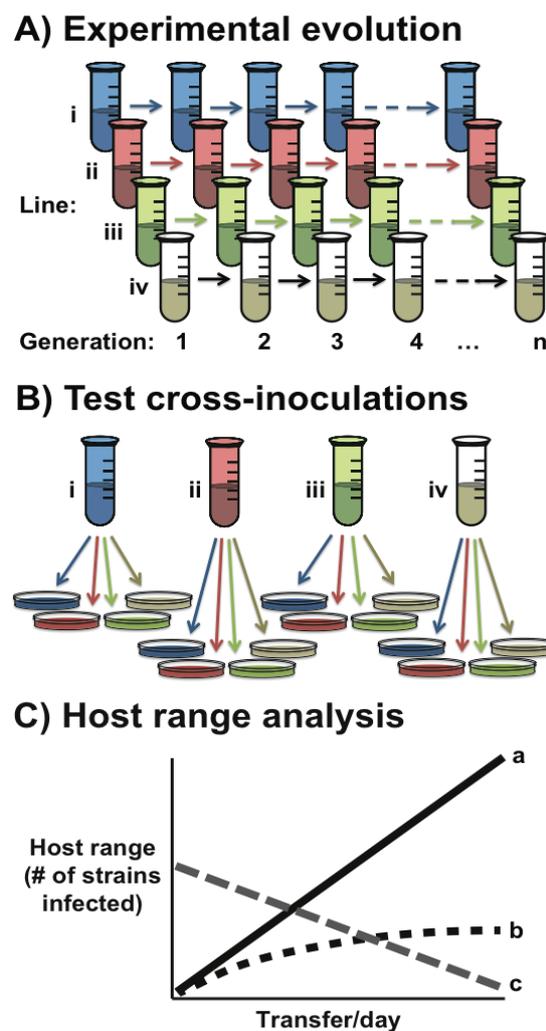
An experimental coevolution approach can also be taken to identify the mutations underlying gains or losses of host types. For example, host range expansion of phage ϕ6 to a novel host was

found to be associated with a single nucleotide change [44]. Furthermore, phages experimentally coevolved with *P. fluorescens* hosts evolved increased host range over time and the phage genotypes with the broadest host ranges were found to have the most nonsynonymous amino acid changes, especially in the phage tail fiber gene [45]. Finally, experimental evolution of phage λ on populations of *Escherichia coli* that had lost the receptor used for phage attachment were found to evolve the ability to infect the bacterial host via a novel receptor following the spread of key precursor mutations, suggesting that phage host shifting can occur via entirely new innovations [46]. Clearly, the power of experimental evolution in understanding phage host range has not been fully exploited and moving forward this approach will offer further insight to the evolution of phage specificity in complex bacterial communities, fitness trade-offs between broad and narrow host ranges, and the potential limits of host shifting among phages.

1.1. The Structure of Bacteria-Phage Interaction Networks

Studies of phages in natural populations have confirmed that, like many other parasites, phages are well adapted to their local host populations [3,24,47]. However, whether individual phages are specialized on certain genotypes/species, the frequencies of which differ across space, or whether coevolution is driving among-population divergence remains to be determined. Importantly, phage specificity can only be meaningfully evaluated within a culturable reference set of host and/or parasite genotypes; and a different reference panel of host or pathogen genotypes may reveal different levels of specificity. Choosing an appropriate reference set is often challenging, as the interaction networks and species ranges of most phages is not known, but it is of central importance to choose a panel that is biologically meaningful and/or informative to the predictions being tested.

Those studies that have looked at host range of individual phages from the environment demonstrate large variation in specificity, both within and across bacterial species. In fact, some phages that appear to be "generalist," in the sense that they can infect bacteria spanning genera, fail to infect a subset of strains or species within genera (Table 1). This apparent inconsistency, which is likely the result of both specific phage adaptations and the subsequent evolution of bacterial resistance in some lineages, including via transfer of plasmids, makes it difficult to decipher specific rules regarding phage host range. For example, populations of *Candidatus* isolated from two separate sludge bioreactors were found to differ primarily in genomic regions encoding phage defense mechanisms, despite global dispersal of the strains among the two sites [48]. It is only when many of these studies are compared and datasets are compiled that clear patterns emerge. A recent meta-analysis of the statistical structure of bacteria-phage interactions shows that the infection networks are non-random and are typically nested [18]. This means that the highly resistant bacterial strains/species are only infected by broad host range phages, whereas the highly susceptible bacteria are infected by phages with narrow to broad host ranges. The observed lack of modularity may be suggestive of a true continuum of phage host range. Alternatively it might reflect the fact that most studies included in the analysis examine interactions either within a single bacterial species (*i.e.*, across multiple strains/genotypes) or between phages from one environment on bacteria from entirely different

communities, habitats, and even continents. To determine the statistical structure of bacteria and phage communities in nature requires analysis of phage host range on representative bacterial hosts from the same local habitat.

1.2. The Evolutionary Implications of Phage Host Range

Given that phages are relatively specific, are capable of rapid adaptation, and are often obligate killers of their host cells, they can impose strong selection on bacterial populations and shape microbial communities. The "Kill the Winner" hypothesis posits that phages adapt to and preferentially infect the lineages of bacteria with the highest frequencies in the population, for example, those with higher metabolic fitness. Evidence for this hypothesis has been collected in a number of ways, including experimental coevolution of phages and bacteria in a test tube [49], monitoring of population change over time [50,51], and using metagenomic approaches [16]. Killing the winner is a form of negative frequency-dependent selection, as bacterial fitness is an indirect function of its frequency in the community. Specifically, bacterial species that are rare and free from phage attack will increase in frequency until the point at which an infective phage is introduced, either via mutation or migration, and spreads through the population. At this point, the common bacterial species will be at a relative disadvantage and may decrease in frequency.

Clearly, this type of dynamic is only possible if specificity underlying infection exists, but it also requires a time lag during which previously rare bacterial species can increase in frequency and remain free from phage attack. The length of this lag, *i.e.*, the time it takes for a phage mutant to arise by mutation or immigration, will dictate how common bacterial species can become before being targeted by coevolving phages. There is also likely to be a lag in the time it takes for a rare phage genotype to increase in frequency as its respective host becomes common, especially in populations where bacterial densities, and therefore rate of encounter, are low. Extending this theory to bacteria-phage interactions, we might predict that oscillatory dynamics should be more pervasive in relatively closed microbial communities, such as the human gut, than in highly connected communities, such as the ocean. This can be extended to predict that microbial communities with higher gene flow of bacteria and/or phages should show greater species evenness than closed communities, where frequencies are fluctuating over time. The data on phage infection of common bacterial species remains scarce, but the evidence we do have is in line with the above prediction. Estimates from marine communities suggest that cyanobacterial cell lysis by cyanophages ranges from a mere 0.005% to 3.2% per day, depending on the season [52]. Similarly, for bacterial isolates collected from the surface of tree leaves, a habitat that is open to constant immigration, only 3% of bacteria were found to be susceptible to local phages. This is in stark contrast to the interior of those same leaves, a more protected and closed microenvironment, where 45% of bacterial isolates were found to be susceptible to local phages [3]. It is important to note, however, that a number of differences exist among these habitats beyond the potential for immigration [53]. A comparable result was found in natural soil samples, where 33 to 40% of bacteria could be lysed by phages from the same sample [24]. Given the paucity of studies

that have explored bacteria-phage dynamics in nature, the ubiquity of phage-mediated negative frequency-dependent selection remains unclear. Future studies exploring the natural prevalence of phage infection, coevolutionary dynamics of bacteria and phages over time, and the evolution of phage host range in either natural or experimental communities are still needed to predict how phages influence bacterial communities.

1.3. The Applied Implications of Phage Host Range

In addition to their potential role in shaping bacterial community composition, phages are key players in shaping the evolution of bacterial genomes [54]. As lytic bacteriophages reproduce within the host and reassemble, bacterial chromosomal DNA can be inadvertently packaged into the viral capsid along with the viral DNA. This mistake will lead to *generalized* transduction (in contrast to *specialized* transduction by prophages) and can move chromosomal DNA from one bacterial host to another. During this movement among hosts, phages can transfer genes encoding toxins or virulence factors, and thus critically alter the bacterial phenotype. For example, phage-mediated transfer of pathogenicity islands between *Listeria monocytogenes* and *Staphylococcus aureus* has been demonstrated in raw milk [55] and phage-mediated transfer of antibiotic resistance has been demonstrated among species of *Enterococcus* [56]. It is increasingly clear that transduction can occur across distantly related bacterial species, and even the seemingly highly conserved 16S rRNA gene has been found within the genome of a broad host range transducing phage [57]. Thus understanding phage specificity among bacterial strains and species is key to predicting potential movement of genes across bacterial species and habitats, and thus the potential emergence of novel pathogens.

An understanding of phage specificity is also central to predicting the success and consequences of phage therapy, *i.e.*, the use of phage or cocktails of phages to control the growth and/or virulence of pathogenic bacteria. The utility of phage therapy is often called into question because of the apparent specificity of phages [58]. However, this specificity is also a clear advantage of phage therapy over more general treatments, such as antibiotics, since the non-target bacterial populations should remain relatively undisturbed. The first steps in testing the benefits of a potentially therapeutic phage are to test (a) whether the phage is too specific to be effective against the standing strain variation of a pathogen in a host population and (b) the likelihood that the phage will affect other non-pathogenic bacteria, either immediately due to a large host range or over short evolutionary timescales as the phage evolves. For example, recent work from silage of dairy farms found a great deal of strain-to-strain variation in susceptibility of the food-borne pathogen, *L. monocytogenes*, to phages collected from silage. They tested the host range of 114 listeriaphages and found that 12% of these phages had narrow host ranges and could infect fewer than half of the strains tested, representing the nine major serotypes of *L. monocytogenes*. However, another 29% of the phages were capable of infecting nearly all of the strains tested, suggesting that these phages would be good candidates for therapeutic control of the pathogen [59]. Furthermore, given the ease of full genome sequencing, it is now possible to scan the phage genome for virulence factors and known toxin-encoding genes to ensure the phage will not act to increase the harm caused by a given pathogen.

2. Results and Discussion

2.1. Specificity within a Natural Phyllosphere Environment

The microbial community within eukaryotic hosts is a relatively closed system that holds the potential for long periods of uninterrupted bacteria-phage coevolution, especially if the hosts are long-lived. Recent work examining bacteria and phages from the horse chestnut phyllosphere (*i.e.*, the above-ground, aerial habitat of the plant) has demonstrated strong local adaptation of phages to bacteria collected from the same leaf [3]. Given that the culturable bacterial communities found within these leaves differed among the trees sampled, it is unclear whether this result demonstrates adaptation of multiple phages to common bacterial strains (which differ among populations) or whether it suggests phage adaptation to infect the common bacterial species within a given community. In other words the result could indicate species sorting according to infection success or it could suggest coevolution of bacteria and phages within each population. One way to tease these two possibilities apart would be to examine specific phage clones from each population and measure their host ranges both within the bacterial community from which they were isolated and from other communities.

Table 1. Examples of phage specificity from natural populations. Habitat refers to the environment from which the samples were selected, and host to the bacterial species used to first visualize the phage. For each study, the number of phages tested is reported and the host range of these phages is described depending on whether the phages were able to infect bacteria from multiple species and/or multiple genera. Finally, we report whether there was variability in phage infectivity on different strains/genotypes within a single bacterial species. In all cases, "n/a" is reported when the phages were not tested in a way that allowed for a given comparison. Two cases show both within-species specificity and an ability to infect multiple species, and these are highlighted in bold to emphasize the difficulty in describing a given phage as "generalist" versus "specialist." Note that this table is for illustrative purposes and is not exhaustive. For a formal meta-analysis of bacteria-phage infection networks, see recent review by Flores and coauthors [60].

Habitat	Host	# Phages tested	Host range Multi-species	Multi-genus	Within-species specificity	Reference
Rhizosphere	Pseudomonas	5	4	0	n/a	Campbell *et al*. 1995 [61]
Sewage	Multiple hosts	11	n/a	11	n/a	Jensen *et al*. 1998 [62]
Industrial	Leuconostoc	6	0	0	Yes	Barrangou *et al*. 2002 [63]
Marine	Vibrio	13	**10**	n/a	**Yes**	Comeau *et al*. 2005 [64]
Soil	Burkholderia	6	**6**	n/a	**Yes**	Seed and Dennis 2005 [65]
Effluent	Salmonella	66	n/a	0	Yes	McLaughlin *et al*. 2006 [66]
Marine	Cellulophaga	46	0	0	Yes	Holmfeldt *et al*. 2007 [67]

As a first step towards this, we randomly selected 144 culturable bacterial isolates from across the eight trees sampled in the original experiment [3] and inoculated each isolate with a dilution series of the sympatric phage community (*i.e.*, filtered leaf homogenate, note that no enrichment procedure was used). For the 14 bacteria that were susceptible to their local phages, a single phage plaque was isolated and re-inoculated into an overnight culture of the bacteria it was able to infect. These co-cultures were then filtered, creating a high titer phage inoculum made up of a single phage clone. All 14 of these phages were cross-inoculated onto each of the 144 original bacterial host isolates, as well as 17 previously characterized *P. syringae* isolates representing nine different pathovars, to determine host range (Figure 2). In addition, we measured the susceptibility of each bacterium in the reference panel to ten phages that were previously collected from sewage and enriched on *P. syringae* pathovar tomato. It is important to note that the host range examined here represents the phage's *plaquing* host range (*i.e.,* the range of hosts a given phage can successfully infect and lyse in soft agar [21]) and may be an underrepresentation of its *productive* (*i.e.*, phage-producing) host range. We sequenced ≈800 bp of the 16S ribosomal RNA region from all hosts that were found to be susceptible to phage. Prior to sequencing, bacterial isolates were grown in KB broth overnight. These overnight cultures were then diluted 1:5 in PCR grade water and used as PCR template (5 μL) in reactions with universal 16S primers 515f (5-GTGCCAGCMGCCGCGGTAA-3) and 1492r (5′-GGTTACCTTGTTACGACTT-3′) [68]. Diluted PCR products (50 ng/μL) were sequenced in both directions, and then aligned and compared against the NCBI database using Geneious (v2.5) software. Individual isolates were assigned to a given genera and species when possible according to the top BLAST hits associated with the sequence (with an e-score of 0.0). Geneious was used first to align the sequences (using MUSCLE, multiple sequence comparison by log-expectation [69]) and then a consensus neighbor-joining tree was assembled with pairwise distances calculated using the "Jukes-Cantor" formula.

We examined the resulting network of bacteria-phage interactions using a recently described method [70]. Relative to the null model, the network shows very little evidence of nestedness, in that the phages with broad host range do not tend to infect the more resistant hosts (the network is only 1% more nested than expected by chance). Instead, there is evidence that the network is highly modular; the network shows 82% of interactions occur between isolates from the same module relative to a null model of 66%. This is suggestive that the pattern of local adaptation observed previously [3] is indicative of many phages each coevolving with a subset of the bacterial host community. However, given that the phages and bacteria were collected from across eight separate microbial "populations," represented by eight different tree hosts [3], a larger analysis will be required to confirm that the modularity represents bacteria-phage coevolution at small spatial scales or whether it indeed indicates that these natural microbial communities are harboring phages with largely non-overlapping host ranges. The phages examined clearly fall across the continuum of "generalist" to "specialist" (note that none were restricted to a single bacterial isolate), but most are restricted to infecting fewer than one third of the bacterial isolates. Of the 13 phages isolated from the phyllosphere, 5 are capable of infecting both *Pseudomonas* and *Erwinia* species and most are capable

of infecting multiple pathovars within *P. syringae* (Figure 2). Interestingly, the phages isolated from sewage had a relatively wider host range across the *Pseudomonas* isolates, infecting a mean of 10.9 (SD = 3.21) hosts, than did the phages from the phyllosphere, infecting a mean of 7.57 (SD = 3.18) hosts. This is not surprising, as the method of searching for phage in one environment using a host bacterium from another is likely to bias the resulting isolates towards more "generalist" phages. Finally, we found two phages collected from the horse chestnut leaf that are capable of infecting a previously characterized strain of *P. syringae* pathovar aesculi (Pae), the causal agent of bleeding canker disease in horse chestnut trees. These phages were not, however, capable of infecting the other three strains of Pae we tested, suggesting that there is variation in susceptibility of Pae strains to phages despite the relatively low genetic diversity typically found among isolates of this rapidly emerging pathogen [71].

2.2. The Importance of Dose in Measuring Specificity

In addition to taking into consideration the appropriate reference panel when measuring phage specificity, it is also necessary to take into account that bacterial sensitivity is likely to depend critically on both phage titer and test conditions. Most studies examining the specificity of parasites within their local communities [18,72–74] treat infectivity as a binary trait and examine bacterial sensitivity at a single phage titer and without taking into account environmental effects on the outcome of the interaction. This is often a necessary step given the large sample sizes of many studies. However, when environmental heterogeneity is taken into account it becomes clear that infection specificity is not simply a binary trait but rather that it can depend on local resources [75], temperature [76], experimental approach [21], and dose [67]. For example, examination of *Cellulophaga baltica* strains and their associated phages isolated from coastal waters showed differences of up to 6 orders of magnitude in bacterial sensitivity to the same titer of phage [67]. This variation can be critical to predicting the impact of phages on bacterial communities, as fine-scale differences among bacterial genotypes or species in sensitivity to the same phage would mean a fitness advantage of one over the other. Therefore, studies that treat infection success as binary would fail to predict this apparent competition. Similarly, if broad host range phages are able to infect more hosts simply because they reach higher prevalence within the environment (given the greater number of available hosts), this specificity can be considered context-dependent.

Figure 2. Neighbor-joining tree showing the phylogenetic relationships among bacteria used in this study and their susceptibility to bacteriophages from the phyllosphere (black) or from sewage (blue). Trees are based on 16S rRNA gene sequences (~800 bp). Bacterial isolates in red have been classified previously to the pathovar level. Phages 4a and 102b had identical host ranges, despite being isolated from separate leaves, and their profiles have thus been collapsed into one.

As an example, we measured whether the bacterial isolates used in our host range tests showed variation in sensitivity to the phages from the phyllosphere. To do this, we inoculated multiple hosts with a dilution series of the same phage inocula (and thus necessarily the same phage titer). In this way, we were able to compare the plaque forming units (PFUs) from each inoculum on the lawn of one bacterial host versus another. If the relationship between phage titer and phage infection success for a given inoculum were the same regardless of the host being tested, we would expect to find no significant difference in PFUs across hosts. On the other hand, if different measures of PFUs are found for the same phage inoculum across different hosts, this would suggest that resistance is a quantitative trait, and therefore that host range may be dose-dependent. Overall, we found a significant interaction between phage inoculum and bacterial strain on the PFU per μL observed across 5 replicate dilution series (general linear model with log PFU as a response variable and phage and bacteria as explanatory variables; $F_{5, 156} = 3.25$, $p < 0.01$). Furthermore, when we included only the bacteria-phage combinations for which the test was fully reciprocal (*i.e.*, each of the two phages were infective to the same three hosts), we again found this interaction effect (Figure 3; $F_{1, 29} = 8.03$, $p < 0.01$). We later tested whether the density of bacterial cells in the soft agar overlay could explain this result and were able to rule out this possibility, as PFU was not correlated with bacterial density within the ranges used. This suggests that direct competition between these three bacteria in the presence of either of these phages is likely to be biased towards the least susceptible, even though all three are within the host range of the phage, and reinforces the idea that infection is not a binary trait. However, given that all bacteria were susceptible overall, these results suggest that dose will only affect the binary outcome of host resistance when phage titer is very low. Moving forward, researchers should consider whether bacterial resistance to phage should be considered a quantitative or qualitative trait with regard to their specific question of interest. For example, if experimental evolution lineages are being compared to a control treatment of the same bacterial strain, resistance might be treated as a binary trait to aid in statistical comparison. In this case, the only caution would be in interpreting a negative result such as lack of a cost of resistance, as the approach may have lumped different degrees of resistance into a single phenotype. On the other hand, resistance should be considered as a quantitative trait, varying for example across multiple environmental conditions or dose, if the goal is to make predictions regarding how phages will alter the competitive hierarchy of bacterial species in a community.

Figure 3. Results of the reciprocal cross-inoculation where the same phage inoculum was spotted in a dilution series onto lawns of each of three different *Erwinia sp.* bacterial isolates. The number of plaque forming units (PFUs) per microliter of inocula was measured for each cross, and the means across five replicates are shown on the Y-axis. Variation in PFU within a given inocula (*i.e.*, among the blue bars or among the yellow bars) represents variation in phage success across bacterial hosts. Note that the important comparison is within-phage variation, as between-phage variation reflects absolute differences in phage titer.
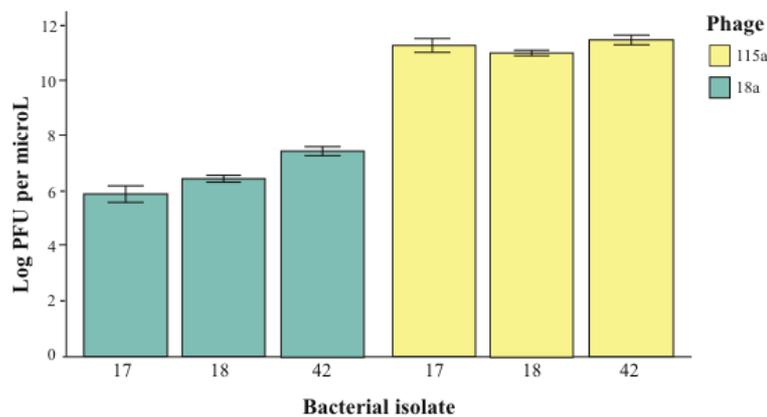
3. Conclusions

We set out to highlight the importance of understanding phage host range as a key factor in predicting (i) how phages shape microbial communities; (ii) when genes encoding virulence and toxins might be transferred among bacterial species; and (iii) the potential success of phages as therapeutic agents against bacterial pathogens. Recent advances in experimental evolution techniques and comparative genomics have given us important new insight, including that phage host range can be altered by single mutations [44], but is often the result of complex epistatic interactions among mutations [45], and that selection can lead to increased host range [42], but that this increase often carries a significant fitness cost [36,38,77]. Studies of the host ranges of natural phages have demonstrated a high variability of specificity, ranging from phages with extremely narrow ranges of hosts within a single species to those that can infect bacteria across genera. A key new insight comes from comparison across studies and systems, and suggests that bacteria-phage networks tend to be statistically nested [18,70]. However, our understanding of these networks remains limited, as very few whole community analyses have been completed and, importantly, estimates of phage host range are only as good as the reference panel against which they've been tested. Furthermore, phage infection is rarely if ever a binary trait, and therefore studies that measure infectivity at a single dose, and/or under a single environmental treatment, are likely to miss key aspects of phage specificity. Moving forward, data from natural populations in which the community of both bacteria and phages can be properly represented and tested in a way that captures variation in infection success beyond a

simple "yes/no", will help elucidate the complex networks of interactions between phages and their bacterial host.

Acknowledgments

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Gómez, P.; Buckling, A. Bacteria-phage antagonistic coevolution in soil. *Science* **2011**, *332*, 106–109.
2. Griffiths, R.I.; Thomson, B.C.; James, P.; Bell, T.; Bailey, M.; Whiteley, A.S. The bacterial biogeography of British soils. *Environ. Microb.* **2011**, *13*, 1642–1654.
3. Koskella, B.; Thompson, J.N.; Preston, G.M.; Buckling, A. Local biotic environment shapes the spatial scale of bacteriophage adaptation to bacteria. *Am. Nat.* **2011**, *177*, 440–451.
4. Lindow, S.E.; Brandl, M.T. Microbiology of the Phyllosphere. *Appl. Environ. Microbiol.* **2003**, *69*, 1875–1883.
5. Marston, M.F.; Pierciey, F.J.; Shepard, A.; Gearin, G.; Qi, J.; Yandava, C.; Schuster, S.C.; Henn, M.R.; Martiny, J.B.H. Rapid diversification of coevolving marine Synechococcus and a virus. *PNAS* **2012**, *109*, 4544–4549
6. Pommier, T.; Douzery, E.J.P.; Mouillot, D. Environment drives high phylogenetic turnover among oceanic bacterial communities. *Biol. Lett.* **2012**, *8*, 562–566.
7. Smillie, C.S.; Smith, M.B.; Friedman, J.; Cordero, O.X.; David, L.A.; Alm, E.J. Ecology drives a global network of gene exchange connecting the human microbiome. *Nature* **2011**, *480*, 241–244.
8. Hooper, L.V.; Littman, D.R.; Macpherson, A.J. Interactions Between the Microbiota and the Immune System. *Science* **2012**, *336*, 1268–1273.
9. De Wit, R.; Bouvier, T. "Everything is everywhere, but, the environment selects"; what did Baas Becking and Beijerinck really say? *Environ. Microb.* **2006**, *8*, 755–758.
10. O'Malley, M.A. "Everything is everywhere: but the environment selects": Ubiquitous distribution and ecological determinism in microbial biogeography. *Stud. Hist. Philos. Biol. Biomed. Sci.* **2008**, *39*, 314–325.
11. Lin, W.; Wang, Y.; Li, B.; Pan, Y. A biogeographic distribution of magnetotactic bacteria influenced by salinity. *ISME J.* **2011**, *6*, 475–479.

12. Chu, H.; Fierer, N.; Lauber, C.L.; Caporaso, J.G.; Knight, R.; Grogan, P. Soil bacterial diversity in the Arctic is not fundamentally different from that found in other biomes. *Environ. Microb.* **2010**, *12*, 2998–3006.

13. Buckling, A.; Rainey, P.B. The role of parasites in sympatric and allopatric host diversification. *Nature* **2002**, *420*, 496–499.

14. Childs, L.M.; Held, N.L.; Young, M.J.; Whitaker, R.J.; Weitz, J.S. Multiscale model of CRISPR-induced coevolutionary dynamics: Diversification at the interface of Larmarck and Darwin. *Evolution* **2012**, *66*, 2015–2029.

15. Suttle, C.A. Viruses in the sea. *Nature* **2005**, *437*, 356–361.

16. Rodriguez-Valera, F.; Martin-Cuadrado, A.-B.; Rodriguez-Brito, B.; Pasic, L.; Thingstad, T.F.; Rohwer, F.; Mira, A. Explaining microbial population genomics through phage predation. *Nat. Rev. Microbiol.* **2009**, *7*, 828–836.

17. Clokie, M.R.J.; Millard, A.D.; Letarov, A.V.; Heaphy, S. Phages in nature. *Bacteriophage* **2011**, *1*, 31–45.

18. Flores, C.O.; Meyer, J.R.; Valverde, S.; Farr, L.; Weitz, J.S. Statistical structure of host–phage interactions. *PNAS* **2011**, *108*, 288–297

19. Frisch, A.W.; Levine, P. Specificity of the Multiplication of Bacteriophage. *J. Immunol.* **1936**, *30*, 89–108.

20. Fierer, N.; Lennon, J.T. The generation and maintenance of diversity in microbial communities. *Am. J. Bot.* **2011**, *98*, 439–448.

21. Hyman, P.; Abedon, S.T. Bacteriophage host range and bacterial resistance. *Adv. Appl. Microbiol.* **2010**, *70*, 217–248.

22. Duplessis, M.; Moineau, S. Identification of a genetic determinant responsible for host specificity in Streptococcus thermophilus bacteriophages. *Mol. Microbiol.* **2001**, *41*, 325–336.

23. Miklič, A.; Rogelj, I. Characterization of lactococcal bacteriophages isolated from Slovenian dairies. *Int. J. Food Sci. Technol.* **2003**, *38*, 305‑311.

24. Vos, M.; Birkett, P.J.; Birch, E.; Griffiths, R.I.; Buckling, A. Local adaptation of bacteriophages to their bacterial hosts in soil. *Science* **2009**, *325*, 833.

25. Hall, A.R.; Scanlan, P.D.; Morgan, A.D.; Buckling, A. Host–parasite coevolutionary arms races give way to fluctuating selection. *Ecol. Lett.* **2011**, *14*, 635–642.

26. Riede, I.; Degen, M.; Henning, U. The receptor specificity of bacteriophages can be determined by a tail fiber modifying protein. *EMBO J.* **1985**, *4*, 2343.

27. Rakhuba, D.; Kolomiets, E.; Szwajcer Dey, E.; Novik, G. Bacteriophage receptors, mechanisms of phage adsorption and penetration into host cell. *Pol. J. Microbiol.* **2010**, *59*, 145–155.

28. Chatterjee, S.; Rothenberg, E. Interaction of Bacteriophage l with Its *E. coli* Receptor, LamB. *Viruses* **2012**, *4*, 3162–3178.

29. Mahony, J.; van Sinderen, D., Structural aspects of the interaction of dairy phages with their Host bacteria. *Viruses* **2012**, *4*, 1410–1424.

30. Richter, C.; Chang, J.T.; Fineran, P.C. Function and Regulation of Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/CRISPR Associated (Cas) Systems. *Viruses* **2012**, *4*, 2291–2311.

31. De Vos, W.M.; Underwood, H.M.; Lyndon Davies, F. Plasmid encoded bacteriophage resistance in Streptococcus cremoris SK11. *FEMS Microbiol. Lett.* **1984**, *23*, 175–178.

32. Deng, Y.-M.; Harvey, M.L.; Liu, C.-Q.; Dunn, N.W. A novel plasmid-encoded phage abortive infection system from Lactococcus lactis biovar. diacetylactis. *FEMS Microbiol. Lett.* **1997**, *146*, 149–154.

33. Antonovics, J.; Boots, M.; Ebert, D.; Koskella, B.; Poss, M.; Sadd, B.M. The origins of specificity by means of natural selection: Evolved and nonhost resistance in host-pathogen interactions. *Evolution* **2012**, *6*, 1–9.

34. Straub, C.S.; Ives, A.R.; Gratton, C. Evidence for a trade-off between host-range breadth and host-use efficiency in aphid parasitoids. *Am. Nat.* **2011**, *177*, 389–395.

35. Poullain, V.; Gandon, S.; Brockhurst, M.A.; Buckling, A.; Hochberg, M.E. The evolution of specificity in evolving and coevolving antagonistic interactions between a bacteria and its phage. *Evolution* **2008**, *62*, 1–11.

36. Duffy, S.; Turner, P.E.; Burch, C.L. Pleiotropic costs of niche expansion in the RNA bacteriophage $\Phi 6$. *Genetics* **2006**, *172*, 751‑757.

37. Avrani, S.; Schwartz, D.A.; Lindell, D. Virus-host swinging party in the oceans: Incorporating biological complexity into paradigms of antagonistic coexistence. *Mob. Gen. Elem.* **2012**, *2*, 88–95.

38. Koskella, B.; Lin, D.M.; Buckling, A.; Thompson, J.N. The costs of evolving resistance in heterogeneous parasite environments. *Proc. R. Soc. B: Biol. Sci.* **2012**, *279*, 1896–1903.

39. Gaba, S.; Ebert, D. Time-shift experiments as a tool to study antagonistic coevolution. *TREE* **2009**, *24*, 226–232.

40. Hall, A.R.; Scanlan, P.D.; Buckling, A. Bacteria—Phage Coevolution and the Emergence of Generalist Pathogens. *Am. Nat.* **2011**, *177*, 44–53.

41. Buckling, A.; Rainey, P.B. Antagonistic coevolution between a bacterium and a bacteriophage. *Proc. R. Soc. B: Biol. Sci.* **2002**, *269*, 931–936.

42. Bono, L.M.; Gensel, C.L.; Pfennig, D.W.; Burch, C.L. Competition and the origins of novelty: Experimental evolution of niche-width expansion in a virus. *Biol. Let.* **2013**, *9*, doi:10.1098/rsbl.2012.0616.

43. Dennehy, J.J.; Friedenberg, N.A.; Holt, R.D.; Turner, P.E. Viral ecology and the maintenance of novel host use. *Am. Nat.* **2006**, *167*, 429–439.

44. Duffy, S.; Burch, C.L.; Turner, P.E. Evolution of host specificity drives reproductive isolation among RNA viruses. *Evolution* **2007**, *61*, 2614–2622.

45. Scanlan, P.D.; Hall, A.R.; Lopez-Pascua, L.D.C.; Buckling, A. Genetic basis of infectivity evolution in a bacteriophage. *Molec. Ecol.* **2011**, *20*, 981–989.

46. Meyer, J.R.; Dobias, D.T.; Weitz, J.S.; Barrick, J.E.; Quick, R.T.; Lenski, R.E. Repeatability and contingency in the evolution of a key innovation in phage lambda. *Science* **2012**, *335*, 428–432.

18

47. Held, N., Whitaker, R.J. Viral biogeography revealed by signatures in *Sulfolobus islandicus* genomes. *Environ. Microb.* **2009**, *11*, 457–466.

48. Kunin, V.; He, S.; Warnecke, F.; Peterson, S.B.; Garcia Martin, H.; Haynes, M.; Ivanova, N.; Blackall, L.L.; Breitbart, M.; Rohwer, F.; *et al.* A bacterial metapopulation adapts locally to phage predation despite global dispersal. *Gen. Res.* **2008**, *18*, 293–297.

49. Fuhrman, J.A.; Schwalbach, M. Viral Influence on Aquatic Bacterial Communities. *Biol. Bull.* **2003**, *204*, 192–195.

50. Rodriguez-Brito, B.; Li, L.; Wegley, L.; Furlan, M.; Angly, F.; Breitbart, M.; Buchanan, J.; Desnues, C.; Dinsdale, E.; Edwards, R.; *et al.* Viral and microbial community dynamics in four aquatic environments. *ISME J.* **2010**, *4*, 739–751.

51. Shapiro, O.H.; Kushmaro, A.; Brenner, A. Bacteriophage predation regulates microbial abundance and diversity in a full-scale bioreactor treating industrial wastewater. *ISME J.* **2010**, *4*, 327–36.

52. Waterbury, J.B.; Valois, F.W. Resistance to co-occurring phages enables marine Synechococcus communities to coexist with cyanophages abundant in seawater. *Appl. Environ. Microb.* **1993**, *59*, 3393.

53. Yu, X.; Lund, S.P.; Scott, R.A.; Greenwald, J.W.; Records, A.H.; Nettleton, D.; Lindow, S.E.; Gross, D.C.; Beattie, G.A. Transcriptional responses of Pseudomonas syringae to growth in epiphytic versus apoplastic leaf sites. *PNAS.* **2013**, *110*, 425–434

54. Fineran, P.C.; Petty, N.K.; Salmond, G.P.C. Transduction: Host DNA Transfer by Bacteriophages. In *The Encyclopedia of Microbiology*; Schaechter, M., Ed.; Elsevier, 2009.

55. Chen, J.; Novick, R.P. Phage-mediated intergeneric transfer of toxin genes. *Science* **2009**, *323*, 139–141.

56. Mazaheri Nezhad Fard, R.; Barton, M.; Heuzenroeder, M. Bacteriophage‐mediated transduction of antibiotic resistance in enterococci. *Lett. Appl. Microb.* **2011**, *52*, 559–564.

57. Beumer, A.; Robinson, J.B. A Broad-Host-Range, Generalized Transducing Phage (SN-T) Acquires 16S rRNA Genes from Different Genera of Bacteria. *Appl. Environ. Microb.* **2005**, *71*, 8301–8304.

58. Loc-Carrillo, C.; Abedon, S.T. Pros and cons of phage therapy. *Bacteriophage* **2011**, *1*, 111–114.

59. Vongkamjan, K.; Switt, A.M.; den Bakker, H.C.; Fortes, E.D.; Wiedmann, M. Silage Collected from Dairy Farms Harbors an Abundance of Listeriaphages with Considerable Host Range and Genome Size Diversity. *Appl. Environ. Microb.* **2012**, *78*, 8666–8675.

60. Flores, C.O.; Meyer, J.R.; Valverde, S.; Farr, L.; Weitz, J.S. Statistical structure of host–phage interactions. *PNAS.* **2011**, *108*, 288–297.

61. Campbell, J.I.A.; Albrechtsen, M.; Sørensen, J. Large Pseudomonas phages isolated from barley rhizosphere. *FEMS Microb. Ecol.* **1995**, *18*, 63–74.

62. Jensen, E.C.; Schrader, H.S.; Rieland, B.; Thompson, T.L.; Lee, K.W.; Nickerson, K.W.; Kokjohn, T.A. Prevalence of Broad-Host-Range Lytic Bacteriophages of Sphaerotilus natans, Escherichia coli, andPseudomonas aeruginosa. *Appl. Environ. Microb.* **1998**, *64*, 575–580.

63. Barrangou, R.; Yoon, S.-S.; Breidt, J.F.; Fleming, H.P.; Klaenhammer, T.R. Characterization of Six Leuconostoc fallax Bacteriophages Isolated from an Industrial Sauerkraut Fermentation. *Appl. Environ. Microb.* **2002**, *68*, 5452–5458.

64. Comeau, A.M.; Buenaventura, E.; Suttle, C.A. A Persistent, Productive, and Seasonally Dynamic Vibriophage Population within Pacific Oysters (Crassostrea gigas). *Appl. Environ. Microb.* **2005**, *71*, 5324–5331.

65. Seed, K.D.; Dennis, J.J. Isolation and characterization of bacteriophages of the Burkholderia cepacia complex. *FEMS Microb. Lett.* **2005**, *251*, 273–280.

66. McLaughlin, M.R.; Balaa, M.F.; Sims, J.; King, R. Isolation of Salmonella Bacteriophages from Swine Effluent Lagoons. Journal article number J-10632 of the Mississippi Agricultural and Forestry Experiment Station. *J. Environ. Qual.* **2006**, *35*, 522–528.

67. Holmfeldt, K.; Middelboe, M.; Nybroe, O.; Riemann, L. Large Variabilities in Host Strain Susceptibility and Phage Host Range Govern Interactions between Lytic Marine Phages and Their Flavobacterium Hosts. *Appl. Environ. Microb.* **2007**, *73*, 6730–6739.

68. Turner, S.; Pryer, K.M.; Miao, V.P.W.; Palmer, J.D. Investigating Deep Phylogenetic Relationships among Cyanobacteria and Plastids by Small Subunit rRNA Sequence Analysis1. *J. Euk. Microb.* **1999**, *46*, 327–338.

69. Edgar, R.C. MUSCLE: multiple sequence alignment with high accuracy and high throughput, *Nucl. Acid. Res.* **2004**, *32*, 1792–1797.

70. Weitz, J.S.; Poisot, T.; Meyer, J.R.; Flores, C.O.; Valverde, S.; Sullivan, M.B.; Hochberg, M.E. Phage—bacteria infection networks. *Trends Microb.* **2012**, *21*, 82–91.

71. Green, S.; Laue, B.; Fossdal, C.G.; A'Hara, S.W.; Cottrell, J.E. Infection of horse chestnut (*Aesculus hippocastanum*) by *Pseudomonas syringae* pv. *aesculi* and its detection by quantitative real-time PCR. *Plant Path.* **2009**, *58*, 731–744.

72. Anderson, T.K.; Sukhdeo, M.V.K. Host Centrality in Food Web Networks Determines Parasite Diversity. *PLoS One* **2011**, *6*, e26798.

73. Poulin, R.; Moulillot, D. Parasite specialization from a phylogenetic perspective: A new index of host specificity. *Parasitol.* **2003**, *126*, 473–480.

74. Luijckx, P.; Ben-Ami, F.; Mouton, L.; Du Pasquier, L.; Ebert, D. Cloning of the unculturable parasite Pasteuria ramosa and its Daphnia host reveals extreme genotype–genotype interactions. *Ecol. Lett.* **2011**, *14*, 125–131.

75. Poisot, T.; Lepennetier, G.; Martinez, E.; Ramsayer, J.; Hochberg, M.E. Resource availability affects the structure of a natural bacteria–bacteriophage community. *Biology Letters* **2011**, *7*, 201–204.

76. Seeley, N.D.; Primrose, S.B. The Effect of Temperature on the Ecology of Aquatic Bacteriophages. *J. Gen. Virol.* **1980**, *46*, 87–95.

77. Lennon, J.; Khatana, S.; Marston, M.; Martiny, J. Is there a cost of virus resistance in marine cyanobacteria? *ISME J.* **2007**, *1*, 300.

Article

Spatial Vulnerability: Bacterial Arrangements, Microcolonies, and Biofilms as Responses to Low Rather than High Phage Densities

Stephen T. Abedon

Department of Microbiology, The Ohio State University, 1680 University Dr., Mansfield, OH 44906, USA; E-Mail: abedon.1@osu.edu; Tel.: +1-419-755-4343; Fax: +1-419-755-4327

Abstract: The ability of bacteria to survive and propagate can be dramatically reduced upon exposure to lytic bacteriophages. Study of this impact, from a bacterium's perspective, tends to focus on phage-bacterial interactions that are governed by mass action, such as can be observed within continuous flow or similarly planktonic ecosystems. Alternatively, bacterial molecular properties can be examined, such as specific phage-resistance adaptations. In this study I address instead how limitations on bacterial movement, resulting in the formation of cellular arrangements, microcolonies, or biofilms, could *increase* the vulnerability of bacteria to phages. Principally: (1) Physically associated clonal groupings of bacteria can represent larger targets for phage adsorption than individual bacteria; and (2), due to a combination of proximity and similar phage susceptibility, individual bacteria should be especially vulnerable to phages infecting within the same clonal, bacterial grouping. Consistent with particle transport theory—the physics of movement within fluids—these considerations are suggestive that formation into arrangements, microcolonies, or biofilms could be either less profitable to bacteria when phage predation pressure is high or require more effective phage-resistance mechanisms than seen among bacteria not living within clonal clusters. I consider these ideas of bacterial 'spatial vulnerability' in part within a phage therapy context.

Keywords: adsorption; bacteriophage; biofilms; cellular arrangements; ecology; microcolonies; particle transport; phages; phage therapy

1. Introduction

Environments can be distinguished in terms of the degree of spatial structure that they exhibit, where spatial structure is a description of the extent to which diffusion, motility, and environmental mixing are constrained. Important spatially structured bacterial habitats include soils; sediments; surface tissues of plants, animals, and fungi; and bacterial biofilms in general. The latter are found both as suspended aggregates and on most submerged surfaces. It is an oft-repeated assertion that the majority of bacteria, or at least a large fraction, may be found within biofilms rather than as planktonic organisms [1,2]. Naturally occurring bacteria thus exist to a great extent as spatially structured populations or communities. Furthermore, and pertinent to fields as diverse as medicine [3] and civil engineering [4], pathogenic or nuisance bacteria found in the biofilm state can be resistant to both antibiotics and disinfectants—a resistance that can be more a function of the phenotypic plasticity of bacteria, *i.e.*, varying metabolic states, rather than due to either genetically acquired resistance or diffusion barriers to chemical penetration into biofilms. Development of alternative methods of biofilm removal therefore is desirable [5].

Bacteriophages, the viruses of bacteria, are a possible alternative to antibiotics, or disinfectants, as antibacterial agents. Such phage therapy or phage-mediated bacterial biocontrol [6,7] has shown promise against bacterial biofilms [8,9]. Rather than a relatively new aspect of phage study, however, the exploration of phage infection of spatially structured bacterial populations goes back to the beginning of the phage era. The first generally recognized bacteriophage study [10], that of Twort [11,12], considered in particular the phage impact on bacteria growing as colonies. Though subsequent studies of phage interaction with macroscopic bacterial colonies have been relatively rare, observation of phage-induced lysis of microscopic colonies has been a routine facet of phage biology, with a growing literature considering explicitly the dynamics of bacterial lysis within the context of phage plaque formation [9,13–17]. Given the ubiquity of biofilms within natural environments, phage interaction with spatially structured bacterial populations should be somewhat relevant to our appreciation of phage environmental microbiology in general [13,18]. Similarly, improved understanding of such interactions may possess applied significance, such as helping to inform phage choice as anti-biofilm agents [19,20] or phage modification to improve anti-biofilm properties [21,22].

Here I explore the costs to bacteria of 'group living' that can result from exposure to phages. This I term a 'spatial vulnerability' because bacteria that are physically attached together—as arrangements, microcolonies, "macrocolonies" [23], or otherwise within biofilms—display less mobility relative to each other than do equivalent bacteria found as physically isolated cells. The result, if physically associated bacteria are clonally related [1], can be a greater negative impact resulting from phage exposure than if the same bacteria instead existed as free cells. I argue that benefits associated with group living therefore are accessible to bacteria only to the extent that their vulnerability to phages nevertheless is small. Murray and Jackson [24] provide comparable though more general arguments based on particle transport theory, that is, the physics of un-self-propelled movement within fluids as

applied to aquatic viruses. In general, being large is possible only if the pressure of viral predation is sufficiently low.

Mechanisms that can reduce bacterial exposure to phages include existence within environments into which phage penetration is difficult or when bacteria exist at sufficiently low densities that they are unable to support phage amplification to "inundative" densities [25]. The latter can be described as an avoidance of 'kill the winner' mechanisms [26,27] or, equivalently, bacterial existence within numerical refuges [28,29]. A general implication is that biofilms may tend to persist particularly within environments in which the densities of phages targeting those bacteria are relatively low or, alternatively, that biofilm-forming bacteria *must* possess substantial phage-resistance mechanisms [30–32] in order to maintain their populations within environments where phage predation pressures are relatively high.

The common theme is that living as physically associated and therefore spatially structured clonal groups, in and of itself, should not be expected to serve bacteria as a phage-resistance mechanism. Rather, I argue from first principles that group living can result in greater bacterial vulnerability to phages than may be experienced by bacteria that instead are physically separated from their clonal relations. With this perspective in mind, the utility of phages as an anti-bacterial as well as a specifically anti-biofilm strategy may be appreciated as an explicit ecological reversal of exactly those circumstances in which biofilms otherwise may flourish: Application of sufficient densities of phages, where phages otherwise are lacking, such that uncontrolled bacterial proliferation can be reversed.

2. Results and Discussion

In this study the primary question being asked is what might be the ecological costs to bacteria, in light of phage predation, that are associated with bacterial growth as arrangements or microcolonies. To answer this question, I generally employ ecological models, arguments, and scenarios—that is, considerations of how bacteria may interact with their environments—and this is rather than primarily enlisting evolutionary approaches or perspectives. In addition, as the study represents a relatively novel exploration the ideas presented, I limit discussion to less complex scenarios, avoiding addressing for instance consideration of stochasticity, the growth in size of bacterial arrangements, or simulations of phage-bacterial ecological dynamics.

2.1. Phage Adsorption to Free Bacteria

The interaction between phages and those bacteria that exist as individual, planktonic cells—here, collectively, "free" bacteria—is fairly straightforward. Beginning with phage attachment to a susceptible bacterium, phage-genome uptake occurs, initiating the infection proper. At some point mature virions must be released from the infected bacterium, beginning an extracellular search for new bacteria to infect [33]. This search is driven by a combination of phage diffusion, fluid flow including environmental mixing, and bacterial as well as bulk environmental movement [18]. Upon sufficient mixing, all bacteria within an environment are then equally likely to encounter a particular phage that

has just been released from a specific bacterium. That is, spatial structure can be said to largely *not* exist given a combination of bacteria that are "free" and substantial environmental mixing. In this section I consider the basics of phage adsorption, focusing particularly on issues of encounter rates between phages and bacteria rather than mechanisms of phage attachment or subsequent phage initiation of infection. For visualization of the spatial scale of environments in which these interactions take place, see Abedon [34].

2.1.1. Phage Movement towards Bacterial Targets

The extracellular search at its most basic consists of a process of virion diffusion. Such diffusion, due to the comparatively small size of phages, occurs at a rate that is substantially greater than the diffusion of free-floating bacteria. As a result, phage extracellular movement towards an idealized bacterium can be described [24] (p. 104) as "simple diffusion to a single sphere". The likelihood of phage-bacterium collision, even in an environment lacking in spatial structure, thus is a function of phage diffusion much more so than the diffusion of target bacteria. Given the substantially larger size of bacteria relative to phages, the likelihood of phage-bacterium encounter is governed by bacterial target size much more so than phage diameter [35]. Free phages thus can be considered to rapidly diffuse among relatively large and stationary bacterial targets (Figure 1).

Figure 1. Illustration of phage and bacterial contributions to phage adsorption rates. Generally phages are relatively small and bacteria somewhat larger. Since diffusion rates are inversely proportional to particle size, whereas target size is proportional to particle size, the result is that phage diffusion (larger arrows pointing right) is a more important contributor to phage adsorption than is bacterial diffusion (smaller arrows point left) while bacterial target size is more important than phage target size to the likelihood of phage-bacterial encounter. An approximate doubling of total bacterial size (lower right) consequently affects target size but has little relevant impact on combined diffusion rates. Note that arrow lengths reflect an assumption that phages are one-tenth the diameter of the coccus and one-twentieth the diameter of the diplococcus.
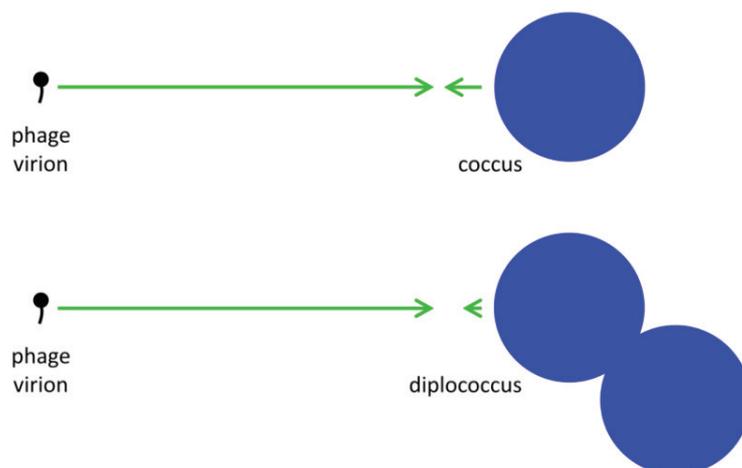
Rates of virus diffusion along with the size of target bacteria are not the only factors affecting rates of encounter between these entities, as so too do their environmental densities. For instance, and as considered in some detail by Murray and Jackson [24], the specific rate of virus adsorption is a function of both the size as well as number of adsorbable targets. Bacteria thus represent smaller targets relative to protozoa and as a consequence individual viruses are less likely to encounter individual bacteria in comparison to individual protozoa. Protozoa, however, tend to display lower population densities than do bacteria, resulting in lower rates of virus encounter with protozoa despite the latter's larger individual sizes. These lower rates of encounter can have the effect of keeping prey densities below 'winner' concentrations (Section 1) despite larger target sizes for individual prey organisms.

Ioannou *et al.* [36] provide both theory and experimental evidence that, in light of predation, the costs to prey of increasing in size—and therefore becoming more absolutely visible to predators—may be offset by prey decrease in abundance. In particular, larval prey that are present at lower densities will be relatively less visible due to, on average, their existing at greater distances from their three-spined stickleback predator (a type of fish), just as bacteria existing at lower densities too, on average, are present at greater distances from individual phages relative to bacteria found in higher density bacterial populations. Consistent with these considerations, in subsequent sections I suggest that bacterial existence as arrangements or as microcolonies can result in greater bacterial vulnerability to phages due to resulting increases in overall target size. I suggest in addition that such vulnerabilities may be avoided at least in part when bacterial populations are relatively rare, that is, should bacteria enter into what have been described, for free bacteria, as numerical refuges [28,29].

2.1.2. Basic Adsorption Calculations

The higher the density at which phage virions are present within environments then the more likely that a given bacterium will encounter a phage, potentially resulting in bacterial conversion from phage uninfected to phage infected. Thus,

$$N_t = N_0 e^{-kPt} \tag{1}$$

where N_t is the density of bacteria that are not phage infected at the end of some interval (t), N_0 is the density of uninfected bacteria at the beginning of that interval, P is the density of free phages (that is, phages that are both unadsorbed and no longer associated with their parental infection), and kPt represents an actual multiplicity of phage *infection* [9,19,25,37], that is, MOI$_{actual}$ as defined by Kasman *et al.* [38]. The phage adsorption rate constant, k, is the probability that a single phage within a specified volume will encounter and then adsorb a single bacterium. This value is based in part on the rate of phage diffusion along with the size of bacterial targets (Section 2.1.1). Note in this equation that phage densities are presumed to remain constant over the course of the interval, t, a situation that may be readily approximated particularly when bacterial densities are low [39].

The more bacteria that are present within a given environment then the more likely that a specific phage will encounter some bacterium, such that,

$$P_t = P_0 e^{-kNt} \qquad (2)$$

where P_0 is the initial phage density and P_t is that density after time, t. This equation in particular describes the loss of free phages as a function of bacterial adsorption. Substantial declines in phage titers will occur due to bacterial adsorption, however, only if bacterial densities are relatively high or t is relatively large. Consequently, and as is true also with Equation (1), for this study I employ the simplifying assumption that phage densities do not vary over time. Operationally, this means that I am placing greater emphasis on consideration of bacterial vulnerability to phages than I am on the dynamics of phage generation and loss.

2.2. Phage Interaction with Bacterial Arrangements

We can increase the complexity of phage-bacterial interactions by considering bacteria that are found predominantly as arrangements rather than as otherwise "free" cells (note that generally, in using the term "arrangement", I am implying microcolony as well, that is, clonally related bacteria that by some means are found attached to one another). For example, bacteria can be arranged as doublets of cells (such as diplococci), strings of bacteria (streptococci or streptobacilli), or other, more complicated forms (staphylococci, tetrads, or sarcinae), and even, as indicated above, as microcolonies as well as biofilms. These arrangements are formed in the course of bacterial division and they differ in terms of the number of divisions that take place prior to cell separation as well as in terms of the planes of those divisions. Forming into arrangements presumably provides bacteria with selective advantages, as considered in Section 2.3.1. This is just as the specific shapes that different bacterial strains and species display, such as coccus *versus* bacillus *versus* spirillum, can be viewed as presumptively adaptive [40,41] or biofilm phenotypes can be seen as improvements in some manner upon the planktonic state [23].

Existing as arrangements, or as microcolonies, may be costly in the face of phage-mediated predation. We can consider this proposed elevation in costs as a consequence of increases in the overall target size of arrangements relative to individual bacteria, which is relevant especially in combination with increased potential for phage propagation within arrangements. At an extreme, arrangement target size could increase directly as a multiple of the number of bacteria found within an arrangement (*i.e.*, ten bacteria as a single target could be ten-times as likely to become phage adsorbed as a single bacterium). Again at an extreme, once an arrangement has become phage infected, then complete loss of all bacteria found in that arrangement could occur. In this section, I consider limitations on these extremes. I nevertheless retain the general conclusion that group living could increase bacterial vulnerability to phages.

2.2.1. Increased Target Size

The likelihood of a bacterium encountering a phage, as indicated in Equation (1), is kPt, where k is a function in part of the bacterium's target size [24,35]. If bacteria form into arrangements, then the likelihood that a specific bacterium encounters a phage may be lower due to partial shading of bacteria

by other bacteria [42] or, alternatively, because of shading that results from bacterial association with surfaces. To reflect these issues, I will use the term $\ddot{k}$ to describe reductions in phage adsorption rates to bacteria that stem from shading, such that $\ddot{k} < k$. Note that the umlaut's intention is to imply a description of properties associated with bacterial arrangements, with the double dots literally suggestive of a diplococcus.

The rate of phage adsorption to a bacterial arrangement can be described as $n\ddot{k}P$, where n is the number of bacteria making up an individual arrangement. That is, the *target size* of an arrangement increases by a factor of n relative to free bacteria while at the same time decreases by a factor of $\ddot{k}/k$. The increase due to n, however, likely is greater than the decrease described by $\ddot{k}/k$, at least so long as arrangements are not sequestered within phage-excluding volumes such as (perhaps) defects in the glass walls of chemostats [29]. Larger arrangements, in other words, almost inevitably will tend to serve as larger targets for phage encounter than will either individual bacteria or smaller arrangements.

Figure 2. Shading of bacteria by bacteria. Shown is a progression starting with two "free" coccus-shaped bacteria (left) which is followed by a diplococcus displaying some degree of attachment (middle) that in turn is followed by a diplococcus displaying maximal attachment as well as minimized surface-to-volume ratio (right), *i.e.*, existing as a combined-volume sphere of $2^{1/3}$-fold increased radius over an individual cell (see calculation, below). The left-hand lack of arrangement shows no shading whereas the right-hand arrangement shows an approximation of maximal shading for a combined spherical shape. The middle arrangement displays some intermediate degree of shading and therefore some intermediate overall target size between maximal and minimal (holding cell volumes constant). Note that the volume of a sphere, V_1, is equal to $\left(\frac{4}{3}\right)\pi r_1^3$. Twice its volume ($V_2$) therefore is $\left(\frac{8}{3}\right)\pi r_1^3$, which as a sphere is equal to $\left(\frac{4}{3}\right)\pi r_2^3$. For $\left(\frac{4}{3}\right)\pi r_2^3 = \left(\frac{8}{3}\right)\pi r_1^3$, then $r_2 = 2^{1/3}r_1$. With such shading, then, diameter increases by only $2^{1/3} = 1.26$ fold.

Generally it is the diameter or 'breadth' of bacterial targets that is crucial to determining viral contact rates [24,35]. For example, the target size of paired, spherical bacteria (diplococci) will range between ~1.26 ($= 2^{1/3} = n\ddot{k}$), which is the increased diameter of a two-fold larger volume, and approximately two (n) times larger than the target size of individual cocci. These values in other words range from where shading is substantial (1.26 times) to where shading instead is minimal (~2 times; for illustration, see Figure 2). Given diversity in arrangement shape it is clear that using arrangement diameter as a proxy for target size is a simplification, though one which I retain both for the sake of

mathematical convenience and because assuming that targets are spherical may be the most reasonable of default assumptions. Clearly though, and as indicated in the above calculation (Figure 2), surface area (as equivalent to the "~2 times" calculation) provides a more intuitive perspective on target size and particularly so given non-spherical as well as relatively immobile targets. The larger and more important point, however, is that in terms of target size, arrangements should be inherently more vulnerable to phage encounter than individual bacteria.

2.2.2. Increased Multiplicity of Adsorption

To visualize the impact of forming into arrangements, with shading affecting phage adsorption rates, compare Equation (1) with

$$A_t = A_0 e^{-n\ddot{k}Pt} \qquad (3)$$

Here A stands for arrangement and A_t is the number of arrangements that have *not* been phage adsorbed over an interval, t, given a constant phage density, P. So long as $n\ddot{k} > k$ holds, then $N_t/N_0 > A_t/A_0$. That is, fewer arrangements will remain fractionally unadsorbed (A_t/A_0) than would individual, free bacteria (N_t/N_0), holding bacterial size and adsorption susceptibility otherwise constant. Here $n\ddot{k}Pt$ is equivalent to MOA$_{actual}$ for arrangements. Note though that it is my preference to instead use the term multiplicity of *adsorption*, *i.e.*, MOA, rather than multiplicity of *infection* because while an arrangement can be wholly adsorbed by a phage, subsequent infection of the whole arrangement is a more complicated process *versus* the infection of individual phage-adsorbed bacteria.

A complementary perspective on the above assertion—that is, that fewer arrangements will remain unadsorbed by phages relative to free bacteria, $N_t/N_0 > A_t/A_0$—is that MOA for arrangements can be up to n-fold higher than that for individual cells. A quantity that I will call MOA$_{input}$ (M) can, after Kasman *et al.* [38], be set equal to the density of phages divided by the density of phage targets. The density of arrangements (A_0), as phage targets, is expected to be n-fold lower than that of free bacteria, *i.e.*, $A_0 = N_0/n$, assuming a constancy in both cell size and total species biomass [36,43]. Holding phage numbers constant, then M for arrangements (M_A) is expected to be n-fold higher than M for free bacteria (M_N), since $M_A = P/(N_0/n)$ whereas $M_N = P/N_0$. The fraction of targets expected to remain unabsorbed, in turn, is readily calculated as e^{-M}, which is the frequency of the zero category—bacteria (N_t/N_0) or arrangements (A_t/A_0) experiencing no phage adsorption—given a Poisson distribution of phages adsorbing to targets. The larger M then the smaller the fraction of cells or arrangements remaining unadsorbed, and therefore $N_t/N_0 > A_t/A_0$ if $M_A > M_N$. More precisely, we can consider instead MOA$_{actual}$, which are $M_A = n\ddot{k}Pt$ *versus* $M_N = kPt$. With M defined in this manner, then the fraction of phage targets expected to remain unadsorbed is equal to $e^{-n\ddot{k}Pt}$ ($= A_t/A_0$) and e^{-kPt} ($= N_t/N_0$), respectively, which are restatements of Equations (3) and (1), respectively. Note that $e^{-n\ddot{k}Pt} < e^{-kPt}$ if as expected $n\ddot{k}Pt > kPt$, implying that $A_t/A_0 < N_t/N_0$.

These considerations come with the caveat that increases in the likelihood of arrangement adsorption that occur as a function of n, that is, as n contributes to arrangement diameter and therefore to target size, may be slowed to the extent that adsorption rates to the individual cells making up an

arrangement, $\ddot{k}$, also may decline as n increases. It may be harder, that is, for environmental phages on average to encounter *individual* bacteria that are found within larger arrangements (such as a large microcolony) *versus* individual bacteria that are found in smaller arrangements (such as a diplococcus). Arrangement vulnerability to adsorption, given this tendency, therefore might increase *less rapidly* as a function of the number of bacteria that they contain.

2.2.3. Phage Propagation within Arrangements

The expression $n\ddot{k}Pt$ should adequately describe the likelihood of arrangement encounter with a phage. Further, $n\ddot{k}Pt$ defines the actual multiplicity, a.k.a., multiplicity of adsorption, of a bacterial arrangement with the value $\ddot{k}$ specifying inefficiencies in this phage adsorption process in comparison to free bacteria. In addition, phage adsorption to arrangements is not identical to phage adsorption to free bacteria because only a fraction of the number of bacteria making up an arrangement become initially phage adsorbed (*i.e.*, $1/n$) rather than all of the bacteria making up a free bacterium (1/1). Furthermore, with bacterial arrangements an initial ("primary") phage adsorption could give rise to a variety of subsequent outcomes including infection of only the adsorbed bacterium, subsequent infection of a fraction of the bacteria found within the arrangement, or indeed subsequent infection of all of the bacteria making up an arrangement. The latter result, of course, is the more costly of outcomes to the affected bacteria, just as it is in terms of prey aggregation more generally [36]. The result is that an additional assumption must be made to argue that the vulnerability of bacterial arrangements to phages can be greater than that seen for individual bacteria. This assumption, in particular, is that the efficiency of phage propagation among bacteria found within arrangements must be greater than that which can be sustained among free bacteria.

When phage densities are higher, then the likelihood of phage adsorption of a given bacterium also should be higher, as described by Equation (1). Importantly, then, the density of phages immediately surrounding a lysing bacterium, that is, as made up predominantly of those phages released from that bacterium, will be the highest phage densities that can be readily attained within a given environment. The phages released in this burst will then diffuse outward, declining in density as they do. The result should be a higher rate of phage adsorption to any susceptible bacterium found within the immediate vicinity of a lysing bacterium, but lower rates at increasing distances (assuming that phage intrinsic adsorption ability does not substantially increase over the course of environmental diffusion).

While free bacteria can randomly find themselves in the vicinity of a lysing bacterium, bacteria that are found in arrangements can be spatially constrained to that vicinity. The result is a higher likelihood of phage infection of other bacteria found within the same phage-infected arrangement than to other environmental bacteria (Figure 3). This argument is similar to an observation made by Babic *et al.* [44] that transfer of conjugative transposons among bacteria found in arrangements (chains) too can be quite efficient and this is for similar reasons, *i.e.*, a constraining of bacterial location in arrangements to within the vicinity of agents infecting the same arrangement (p. 1): "Since many bacterial species grow naturally in chains, this intrachain transfer is likely a common mechanism for accelerating the spread of conjugative elements within microbial communities."

Figure 3. Illustration of the tendency of phages to display biases towards acquisition of locally available bacteria. Here shown to the right is phage acquisition of a bacterium (blue) that is found as part of the same arrangement as a lysing bacterium (red with dashed border). The green arrows represent outwardly diffusing phage progeny released upon bacterial lysis while the shorter, gray arrows illustrate the tendency of those phages that are released immediately adjacent to an uninfected bacterium to encounter that bacterium. Contrasting this second bacterium looming large in the vicinity of an adjacent phage burst, even at a high plankton bacterial density of 10^8 per mL, each free bacterium (left) occupies a total environmental volume of 10^4 μm^3 (1 cm = 10^4 μm, meaning that 1 mL = 1 cm^3 = 10^{12} μm^3, where 10^{12} μm^3/10^8 bacteria = 10^4 μm^3/bacterium). This density in turn implies an average distance between bacteria of about $10^{4/3}$ (*i.e.*, the cube root of 10^4 μm^3), or more than 10 μm, which one may compare with a typical bacterium diameter of about 1 μm. Thus, bacteria in arrangements can be not-unreasonably described as having local densities that should encourage phage adsorption with higher likelihood than that seen among planktonic, individual bacteria.
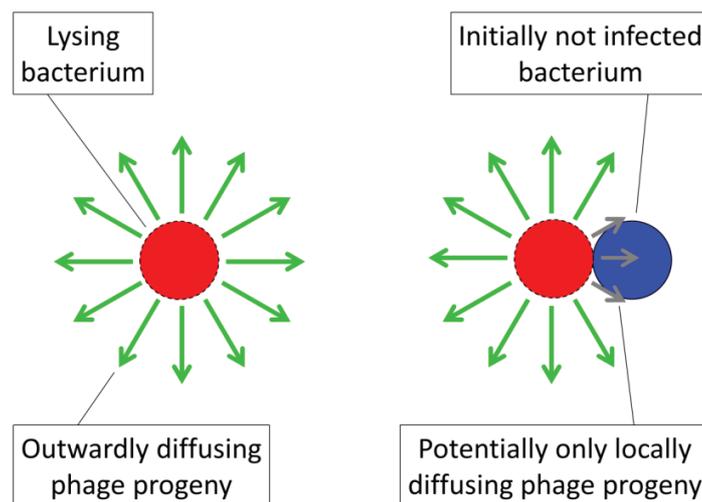
Analogously, greater bacterial densities found within bacterial arrangement can be viewed as possessing a higher "mass" relative to that associated with planktonic bacteria, where by "mass"—potentially confusingly—I am referring to immediately local *densities* of bacteria. This "mass" of bacteria within a bacterial arrangement would be more likely to exceed a "critical" level, such that phage propagation can be sustained, than may be achieved by an equivalent number of more locally dilute, free bacteria. This perspective is just as one can consider for nuclear fission, *i.e.*, radioactive decay, and associated chain reactions [45], which occurs more rapidly given higher "fuel" densities. Indeed, the idea of a "critical mass" can be directly equated with phage proliferation thresholds, that is, those bacterial densities at which rates of phage population growth are perfectly balanced by rates of phage loss [19,25]. Each is a description of target densities (atoms or bacteria) that can be sufficiently high that collision with targets (by neutrons or phages), in combination with

subsequent proliferation (via fission or infection), balances any losses that can occur due either to interactions with nontargets or movement away from the focus "mass".

A target bacterium that is found in the immediate vicinity of a phage, in other words, displays a much higher local density from the perspective of that phage than may be the case for bacteria that are randomly dispersed throughout an environment. A local concentration of bacterial "mass" thus can result in a high propensity for bacterial adsorption by phages that have been generated within the same "mass". See Abedon [34], by contrast, for illustration of the relatively low propensity for phages to randomly encounter free-floating bacteria found within fluid environments; see too Hagens and Loessner [46] as well as Goodridge [47]. As default assumptions, therefore, we can view bacterial arrangements as both larger targets for phage adsorption than individual bacteria and as locally higher bacterial densities, densities that may be better able to support local phage propagation and population growth than more diffuse populations of free bacteria. Consistent with the analogy with nuclear fission, which can be controlled by the insertion of neutron-absorbing substances, we can question the efficiency of phage acquisition and then infection of seemingly adjacent bacteria, e.g., perhaps as may be separated by extracellular matrix.

2.2.4. Inefficiencies in Phage Propagation

Notwithstanding proposed tendencies for phages to more readily acquire bacteria that are within their immediate vicinity, it is as noted possible for inefficiencies to exist in the sequential phage infection of bacteria co-occupying the same arrangements. To incorporate such inefficiencies into models of arrangement vulnerability to phages, I employ the term, $\ddot{n}$. This represents the number of bacteria within an arrangement that will be lost, on average, as a consequence of phage adsorption of a single bacterium within that arrangement. This number, $\ddot{n}$, can range up to the total number of bacteria making up an arrangement (n). Another way of viewing $\ddot{n}$, however, is that $\ddot{n} > 1$ implies a phage reproductive number within a bacterial arrangement that is greater than 0 such that some degree of phage propagation within an arrangement occurs along with consequent bacterial death. If insufficient phage release from infected bacteria and/or insufficient subsequent bacterial infection occurs within an arrangement, or subsequent infections don't happen fast enough, then complete eradication of an arrangement by an infecting phage may not happen, such that $\ddot{n} < n$.

These ideas can be expressed as,

$$N_t = N_0 - \left(1 - e^{-n\ddot{k}Pt}\right)\ddot{n}A_0 = N_0 - \left(1 - e^{-n\ddot{k}Pt}\right)\ddot{n}N_0/n \tag{4}$$

where $(1 - e^{-n\ddot{k}Pt})$ is the fraction of arrangements that become individually phage adsorbed over some interval, t, and $\ddot{n}$ is the number of bacteria per arrangement that are lost to this adsorption (assuming, for simplicity, that $\ddot{n}$ is independent of the actual multiplicity of phage adsorption to a given arrangement). Were $\ddot{n} = 1$, then though arrangements are more likely to be adsorbed than free cells, nevertheless no more bacteria would be lost per arrangement adsorption. Indeed, to the extent that the initial bacterial infection is less likely, that is, given $\ddot{k} < k$ along with $\ddot{n} = 1$, then overall existence as an arrangement could result in *less* vulnerability to phages rather than more. Such a

situation would occur, for example, were phage infections abortive or perhaps could result instead were infections considerably reduced in burst size or extended in latent period such that phage propagation through an arrangement were substantially impaired. Alternatively, the equality $\ddot{n} = n$ would imply complete arrangement loss following each phage adsorption of an arrangement. The parameter, $\ddot{n}$, is thus a description of arrangement vulnerability to phages post-adsorption, ranging from minimal ($\ddot{n} = 1$, or even $\ddot{n} = 0$) to maximal ($\ddot{n} = n$). A visual summary of the models represented by Equations (1) and (4) is presented in Figure 4.

Figure 4. The model. Parameters include P (density of phages in environment), k (phage adsorption constant), $\ddot{k}$ (phage adsorption constant considering reductions due to shading of bacteria by bacteria found within bacterial arrangements), n (number of bacteria found per arrangement), N (bacterial density of overall environment), L (phage latent period, which is the duration of a phage infection), and $\ddot{n}$ (number of bacteria per arrangement lost subsequent to phage infection of one cell in the arrangement). Likelihood of phage adsorption of bacterial arrangements is $n\ddot{k}$ and density of arrangements within environments is equal to $N/n = N_0/n$ (or indeed $n_0\ddot{k}$ and N_0/n_0, respectively, to reflect that n changes as a function of time in the figure). The inequality $t \geq 2L$ indicates how phage acquisition of bacteria within a bacterial arrangement, according to this model, involves at least two sequential rounds of phage infection. The absence of cells in the lower right is intentional as too is the reduction in cell number to n_t in the lower left. Both of these reductions in cell number, going from middle to bottom, indicate phage-induced bacterial lysis.
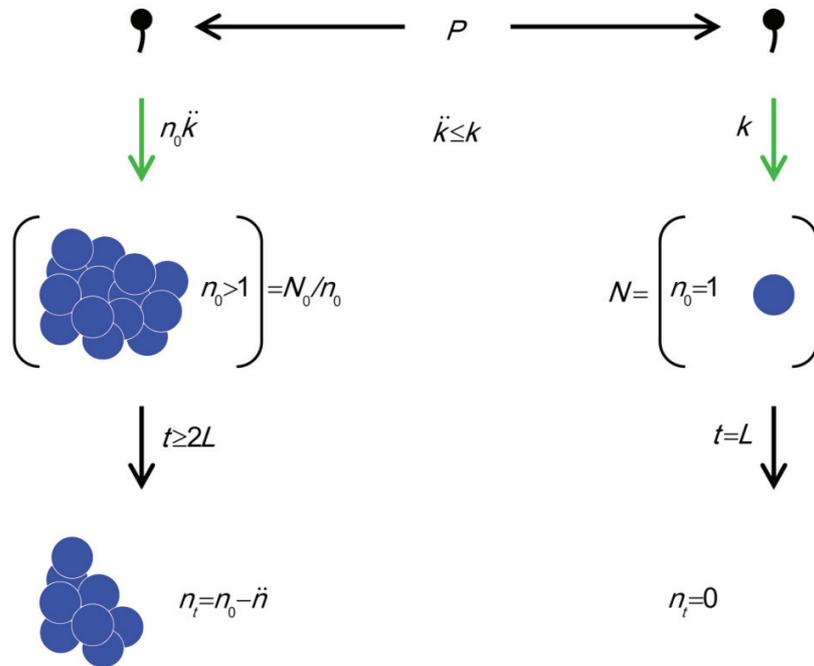
2.2.5. An Important Special Case

Equation (4) is less applicable given higher levels of phage adsorption, as can occur over longer periods. This is because, as noted, Equation (4) fails to take into account the impact, on the overall fraction of bacteria that are lost, of multiple phage adsorptions of individual arrangements. That is, for instance, we might have an expectation of greater bacterial loss with greater levels of arrangement adsorption by environmental phages under conditions where otherwise $\ddot{n} < n$, but Equation (4) does not reflect the possibility that if one adsorbing phage fails to clear an arrangement then perhaps more than one adsorbing phage will, with greater likelihood, succeed in doing so. One way to address this concern is simply by setting $\ddot{n} = n$, that is, an assumption of complete arrangement loss per primary phage adsorption.

Alternatively, one can limit one's considerations to circumstances in which phage multiplicities of adsorption to arrangements are relatively low (<<1). In such cases the equation can be simplified as

$$N_t = N_0 - \left(n\ddot{k}Pt\right)\ddot{n}A_0 = N_0 - \ddot{k}Pt\ddot{n}N_0 \tag{5}$$

which compares, for a free bacterium, with,

$$N_t = N_0 - kPtN_0 \tag{6}$$

Note that N_t (as defined by Equation (5)) is smaller than N_t (as defined by Equation (6)) when $\ddot{n} > k/\ddot{k}$. That is, when this inequality holds then bacteria found within bacterial arrangements are more vulnerable to phages than are phage-susceptible bacteria that are "free". In words: Bacteria found within arrangements are more susceptible to phage attack if losses due to existence within a phage-adsorbed arrangement, $\ddot{n}$, are greater than increases in *individual* bacterial vulnerability to "primary" adsorptions that come from *not* existing within an arrangement (that is, $k/\ddot{k}$ where $k > \ddot{k}$).

If a bacterium is half as likely to be subject to primary phage adsorption when found within an arrangement ($\ddot{k}/k = 0.5$ such that $k/\ddot{k} = 2$) but each primary adsorption results in the loss of ten bacteria ($\ddot{n} = 10$), then the resulting $10 > 2$ would imply that arrangements are more vulnerable to phages than are free bacteria. Indeed, in this example bacteria found within arrangements would be five times more vulnerable. Note further that if arrangements can reduce $\ddot{n}$ to 1, then arrangements will be expected to display lower vulnerability to phages than free bacteria so long as individual bacteria found within arrangements are less vulnerable to primary phage adsorptions than free bacteria (*i.e.*, again, such that $k > \ddot{k}$). This is an observation that could very well explain the utility of abortive infection systems to bacteria, *i.e.*, phage-resistance mechanisms in which both adsorbed bacteria and adsorbing phages die [30]. Alternatively, for $\ddot{n} = 0$, which is the case given successful bacterial display of, for example, anti-phage restriction-modification systems, then biologically the inequality, $\ddot{n} > k/\ddot{k}$, no longer holds, *i.e.*, $0 > k/\ddot{k}$. Logically, though, in this case arrangements should be no more or less vulnerable to phages than free bacteria since, in fact, neither would be vulnerable. See Table 1, under "Higher" phage densities for summary. Note, though, that calculations relevant to the "Lower" phage densities portion of the Table, *i.e.*, Equation (7), are not discussed until Sections 2.3.3 and 2.3.4.

Table 1. Summary of predictions as a function of phage densities in environments and phage potential to acquire bacteria sequentially within bacterial arrangements (recall in interpreting the table that the inequality, $N_t / N_0 > A_t / A_0$, implies greater success over time in the face of phage-mediated predation for free bacteria *versus* bacteria found within arrangements while $N_t / N_0 < A_t / A_0$ implies the opposite). Generally, $k/\ddot{k} \geq 1$. Calculations relevant to lower phage densities, as found in the bottom portion of the table, are not discussed until Section 2.3.3 and especially section 2.3.4.

Environmental Phage Density (*P*)	Phage Propagation Ability Through Arrangements ($\ddot{n}$)			
	Higher	**Lower**		
Higher (bacterial losses dominate dynamics)	For $\ddot{n} > k/\ddot{k}$, $$\frac{N_t}{N_0} > \frac{A_t}{A_0}$$ [lesser or no impediments to phage propagation within arrangements]	For $\ddot{n} < k/\ddot{k}$, $$\frac{N_t}{N_0} < \frac{A_t}{A_0}$$ [impediments less than absolute]	For $\ddot{n} = 1$, $$\frac{N_t}{N_0} \leq \frac{A_t}{A_0}$$ [e.g., abortive infections]	For $\ddot{n} = 0$, $$\frac{N_t}{N_0} = \frac{A_t}{A_0}$$ [e.g., phage restriction]
Lower (bacterial gains dominate dynamics)	For $\mu_A - \mu_N > P(\ddot{k}\ddot{n} - k)$, $$\frac{N_t}{N_0} < \frac{A_t}{A_0}$$ [which, as $P \to 0$, is more likely]	$$\frac{N_t}{N_0} < \frac{A_t}{A_0}$$	[assuming phage-independent advantages to arrangement formation, *i.e.*, $\mu_A - \mu_N > 0$, and that $\mu_A - \mu_N > P(\ddot{k}\ddot{n} - k)$ holds, which is likely given both $P \to 0$ and $\ddot{n} \to 0$]	

2.3. Utility of Group Living in Light of Phages

Above I argue that the likelihood of adsorption by any given phage can be lower for individual bacteria found within arrangements ($\ddot{k}$) than for free bacteria (k) but nonetheless that arrangements can be more vulnerable to exploitation by phages. If we assume that arrangements nevertheless provide bacteria with selective benefits, then it should be possible to consider how large these benefits must be to offset costs stemming from this presumed increased susceptibility of arrangements to phage infection. Before moving on to that issue, however, I first address two underlying considerations, (1) how existence within clonal associations in fact might benefit bacteria and (2) experimental evidence that phages can propagate through bacterial arrangements and/or microcolonies.

2.3.1. Selective Benefits of Living in Arrangements

Fitness advantages accrued by bacteria from living within groups generally can result from increased short-term growth rates, greater long-term rates of reproduction, increased population resistance to extinction, or greater competitive ability in terms of more effectively sustaining population densities. The latter might be accomplished by gaining better access to nutrients.

Adherence to surfaces in conjunction with subsequent growth as a microcolony, for instance, can retain a bacterial population within the vicinity of key resources, such as flowing water [48].

Continued association of cells following division similarly may allow for more effective penetration into resource-supplying substrates, *i.e.*, burrowing, particularly as seen among filamentous microorganisms growing in spatially structured environments such as soils. Aggregated cells also will tend to have larger collective activity domains [49]. This may be beneficial to bacteria by allowing better retention of extracellular regulatory molecules, particularly such as towards quorum sensing [1,50], or for the concentrating of other beneficial extracellular factors such as exoenzymes or exotoxins. Reduced levels of sharing with unrelated organisms of any molecules generated by these extracellular factors may be possible given bacteria association within arrangements and this may be so simply given greater densities of those cells within a given microvolume.

Growing in a single location can allow for more effective interspecific interactions [49] including crossfeeding [51] as well as development of closely spatially integrated microbial consortia [52,53]. Such associations might be enhanced by cells that grow to higher densities within a specific location, including as cell arrangements or instead as cells that simply fail to disperse and thereby form into microcolonies (or, instead, simply don't disperse very far). More generally, the development of favorable physiochemical gradients might be more readily achieved when cells are living within groups consisting at least in part of related individuals, *i.e.*, as consistent with the above-noted idea of cellular aggregates possessing larger collective activity domains *versus* bacteria that are not found in association with related bacteria. The result may be a higher potential to contribute to synergistic interspecific interactions. Indeed, "Any transfer between two such organisms will be immensely inefficient unless the concentrations of cells and substrates are high… Such interactions can take place between groups of bacteria where the size of the group means an enlarged activity domain and the retention of reactants in the vicinity of the group." [49] (pp. 56 and 58).

Bacteria can display movement when found within disfavorable environments, such as can be associated with planktonic or actively dispersing cells, and then cease such movement, such as through adherence to a surface, when found within a favorable environment. This is a behavior that, minimally, is equivalent to the concept of orthokinesis, that is, where the speed of an organism is modified by external cues or, indeed, thigmokinesis, where speed is modified by degree of physical contact with something else. Included among favorable environments and external cues can be other microbial species from which a given bacterium can derive benefits. Association with other microorganisms on a surface, rather than being stable, can as a consequence of thigmokinesis instead be associated with the invasion of later-arriving organisms. The result can be an interspecific competition for space or resources that can put a premium on competitive ability, which in turn may benefit from more effective arrangement or microcolony formation [50,54,55]. Full or partial replacement of one microbial species with another on surfaces more generally is an example of ecological succession, as has been documented particularly well in terms of dental biofilms [49,56].

Cellular clones growing in a single location have the potential to interact with each other cooperatively towards mutual benefit and do so in ways that are less available to free bacteria. This

can include elevation of colonies into more nutrient-rich microenvironments [48], certain degrees of cellular differentiation and/or physiological specialization, fruiting body formation, and even seemingly intentional cell cannibalism under starvation conditions [1,57]. Perhaps consistent with this idea of bacteria being able to mutually cooperate when living within single-species groups, cells growing as biofilms are known to display an increased resistance to toxic chemicals found in their environments, which clinically can include both antibiotics and disinfectants [3,4]. Increased resistance may also be seen given association other microbial species [58]. Extracellular polymer production itself is often described as a potentially cooperative activity among related, physically associated bacteria [55].

Growth as arrangements may give rise to a greater resistance to engulfment by protozoa, and biofilms otherwise may be more resistant to protozoa-mediated predation than planktonic bacteria [59]. The resulting adaptations, such as "resisting ingestion, by becoming too large or too long… making themselves inaccessible, by growing in aggregates or biofilms", [41] (p. 3), however, may conflict with avoidance of phages by these same bacteria. Biofilm formation can also serve as a means of immune system resistance during infections by bacterial pathogens and this is due at least in part to interference with the action of phagocytes, though not necessarily solely via a direct blocking of bacterial engulfment [60,61,62]. More generally, utility that comes with being larger, whether individually per cell or instead as a consequence of cell-to-cell associations, is exploited by most non-bacterial cellular organisms including animals, plants, fungi, and protists.

2.3.2. Susceptibility of Bacterial Arrangements and Microcolonies to Phage Exploitation

To what degree are bacteria living in arrangements in fact phage susceptible? This question must be addressed particularly since assertions have been made that biofilm formation serves bacteria as an inherently phage-resistant state, as I review elsewhere [9]. Costerton *et al.* [63] in particular noted that (p. 440), "the gellike state of the predominantly polysaccharide biofilm matrix limits the access of antibacterial agents, such as… bacteriophage… Therefore, biofilm bacteria are substantially protected from… bacteriophage…" Elsewhere [8,9], however, I review the substantial potential, under experimental conditions, for phages to in fact considerably impact biofilms. Phages may not be well equipped to drive biofilm bacteria completely to extinction in the course of propagating on those bacteria, and phages that specialize in targeting biofilm bacteria may not always be highly prevalent in environments. Nevertheless, there is no evidence that I know of that would appear to indicate that biofilms are *not* relatively susceptible, or in many cases even highly susceptible to exploitation by phages.

As with the phage potential to clear bacterial biofilms, evidence also exists that phages can propagate through bacterial arrangements. Barron *et al.* [64], for example, state that "the phage released from a single coccus may infect other cocci in that chain." The evidence supporting that claim is an observation by Friend and Slade [65]. They found that one-step growth curves on a group A streptococcal strain in fact were two-step in practice, implying an initial arrangement adsorption followed by post-burst adsorption within the same arrangement; see also earlier work by Kjem, 1958

and 1964, as cited by Friend and Slade, as well as work by Fischetti *et al.* [66]. In the Friend and Slade study, two-step curves were then reduced to one-step curves through the separation of streptococci into individual, that is, "free" bacteria using sonication.

The tendency for phage plaques to be clear, particularly in their centers, also can be viewed as an indication of the tendency for phages to prevent phage-sensitive bacteria from propagating in the immediate vicinity of phage bursts. Indeed, plaque formation explicitly occurs within a context of phage penetration into and subsequent clearance of the bacterial microcolonies that make up bacterial lawns [67]. Consistently, Doolittle *et al.* [68] observed phage propagation within single-species bacterial biofilms, describing the dynamics as plaque-like. Conversely, plaque cloudiness can signify limitations on the ability of phages to clear bacterial microcolonies [9,13,69]. It should *not* be controversial therefore that phages can propagate to at least some extent into bacterial arrangements or microcolonies. Indeed, in phage therapy it is often assumed that phages can be quite adept at propagating through bacterial biofilms, a process that I have described elsewhere as an 'active penetration' [19]. These claims all come with the caveat, however, that such active penetration is not necessarily going to be the case for every combination of phage, bacterial arrangement, and circumstance.

2.3.3. Phage-Mediated Costs of Existing as Arrangements

For the sake of mathematical convenience, I consider especially the phage impact on instantaneous bacterial population growth rates rather than other aspects of bacterial fitness in addressing the extent to which phages might affect bacterial arrangements. Focusing on instantaneous growth rates, in particular, greatly simplifies the mathematics while at the same time avoids consideration of the difficult issue of exactly how bacterial arrangements propagate over longer time periods. The conclusions I reach, however, should be qualitatively applicable to other situations.

These considerations of phage impact on the instantaneous population growth of bacteria existing within arrangements can be easily formulated as a differential equation in which changes in bacterial density (N) are considered in terms of rates of cell division (μ) *versus* declines due to phage adsorption,

$$dN/dt = \mu_A N - n\ddot{k}P\ddot{n}N/n = N(\mu_A - \ddot{k}P\ddot{n}) \tag{7}$$

which compares with

$$dN/dt = \mu_N N - kPN = N(\mu_N - kP) \tag{8}$$

for free bacteria. Note that μ_A and μ_N are growth rates of bacteria associated with arrangements and free bacteria, respectively.

The expression $n\ddot{k}P\ddot{n}N/n$ in Equation (7), or simply kPN in Equation (8), describes those bacteria that have been lost from the unadsorbed bacterial pool, N, as a consequence of phage adsorption either of themselves (Equation (7) and Equation (8)) or of the arrangement in which those bacteria are located (Equation (7)). As noted above (*i.e.*, $A_0 = N_0/n$ in Section 2.2.2), the expression N/n as found in Equation (7) is a description of the density of arrangements consisting of n bacteria that are found in

the environment in question. Also as above, the parameter $\ddot{n}$ describes the number of bacteria that will be lost to phage infection given phage adsorption of an arrangement. Lastly, $n\ddot{k}P$ is a description of the per-arrangement rate of bacterial adsorption by phages given an environmental phage density of P.

Per bacterium, Equation (7) thus describes adsorptions that occur at a rate of $\ddot{k}P$ and each of those adsorptions results in a loss of $\ddot{n}$ bacteria. This compares with the rate of loss of individual bacteria as described by kP in Equation (8). The expression, μN with either subscript, by contrast, describes in both equations the gains in bacterial density that occur as a consequence of bacterial replication. Thus, if $\ddot{k}P\ddot{n} > \mu_A$, then the bacterial population will experience a net decline in number whereas $\ddot{k}P\ddot{n} < \mu_A$ indicates net gains and $\ddot{k}P\ddot{n} = \mu_A$ defines a steady state. For Equation (8) the equivalent expressions instead are respectively $kP > \mu_N$, $kP < \mu_N$, and $kP = \mu_N$, where P in the latter can be described as an inundation threshold or even phage minimum inhibitory concentration, that is, of bacteria [25]. In the absence of phages, the bacterial population will simply grow at rates specified by μ_A or μ_N.

In Equation (7) the dynamics of microcolony formation are not considered and nor are various other complications such as multiple adsorption by "environmental" phages of individual bacterial arrangements, where by "environmental" I am distinguishing those phages defining P (=environmental) from phages that instead are explicitly propagating through bacterial arrangements. What Equation (7) nevertheless indicates is that the phage impact on bacterial arrangements varies as a function of phage density (P) in combination with the susceptibility of individual bacteria making up an arrangement to phages ($\ddot{k}\ddot{n}$). Specifically, the larger $\ddot{k}P\ddot{n}$ then the more bacteria found within arrangements that are lost to phage infection. For example, twice as many bacteria will be lost per unit time for bacterial arrangements ($\ddot{k}P\ddot{n}$ from Equation (7)) *versus* free bacteria (kP from Equation (8)) if $\ddot{k}\ddot{n}/k = 2$.

Alternatively, a total of $\ddot{k}\ddot{n}/k$-fold more arrangement-associated bacteria will be lost *versus* free bacteria for any given phage density, P. Forming into an arrangement, and thereby incurring costs of additional vulnerability to phages, therefore should be worthwhile to bacteria *only* to the extent that μ_A, or some other measure of bacterial fitness, increases as a consequence of group living to a larger extent than bacterial fitness decreases as a result of incurring a greater spatial vulnerability to phages, *i.e.*, as described by $\ddot{k}\ddot{n}/k$.

2.3.4. Importance of Reduced Vulnerability to Phages

The extreme situations with regard to Equation (7) are as follows: (i) If no phages are present in an environment ($P = 0$) then there will be no phage-associated cost to arrangement, microcolony, or biofilm formation (in which case no advantage is required from group living to offset costs of phage adsorption) or (ii) if phages are present at effectively infinite densities ($P = \infty$), or simply sufficiently high densities, then potentially no amount of phage-independent benefit to group living could offset increases in phage-associated vulnerabilities. The latter situation is just as dire for free bacteria, however, *i.e.*, for which $\ddot{n} = 1$ but nonetheless where it is possible for $\ddot{k}P\ddot{n} \approx kP \approx \infty$ (Equation (7) *versus* Equation (8)). At a minimum, however, $\mu_A = \ddot{k}P\ddot{n}$ is necessary for bacterial fitness, as measured here in terms of increases in bacterial growth rates, to offset costs due to phage adsorption,

and this compares with $\mu_N = kP$ for free bacteria; see Abedon and Thomas-Abedon [19] along with references cited, or Abedon [25], for derivation of the latter.

For circumstances in which $\ddot{k}P\ddot{n} > kP$, then the growth rate or other measure of the fitness of bacterial arrangements must be greater than that of free bacteria by that amount, e.g., $\mu_A - \mu_N > \ddot{k}P\ddot{n} - kP = P(\ddot{k}\ddot{n} - k)$, to offset increased phage-associated costs that are borne by bacterial arrangements. This fitness improvement, however, need not be substantial unless phage densities (P) are also substantial. Thus, as $\ddot{n}$ or $\ddot{k}$ increase so too does the potential for phages to block the evolution of bacterial arrangements, but at the same time such increases do not serve as absolute blocks on this evolution. The alternative perspective is that given sufficiently high phage densities—but not too high, as indicated in the previous paragraph—then evolution could tend to favor reductions in $\ddot{n}$ even if bacteria otherwise experience benefits from forming into arrangements, that is, reduced formation of arrangements could serve as a bacterial anti-phage strategy. In simple terms, a coccus might encounter a phage approximately half as often as a diplococcus.

A generalization on these considerations is that bacteria that have formed into arrangements will have to avoid, on a per-bacterium basis, equivalent increases in vulnerability to phages in order to partake of whatever net advantages may be associated with forming into arrangements. This reduced vulnerability, furthermore, can result either from existing in environments in which phage densities are low (where phage-independent bacterial fitness "gains" can dominate competitive dynamics) or, alternatively, from greater resistance by individual bacteria to phage attack when phage densities are high, that is, when phage-mediated losses dominate competitive dynamics [30,31]. This idea of predator-independent aspects to fitness dominating prey evolution when predator densities are low (here bacteria and phages, respectively) but predator-resistance dominating when prey densities are high is a standard conclusion from community ecology. It is also seen in phage-bacterial chemostat studies as phage-sensitive bacteria, if they possess a growth-rate advantage to phage-resistant bacteria, out-compete phage-resistant bacteria when phage densities are lower but those same phage-sensitive bacteria are at a competitive disadvantage if phages instead are more numerous [70,71]. See Table 1 for a general summary of fitness expectations for arrangements *versus* free bacteria given higher *versus* lower arrangement vulnerability to phages.

2.3.5. Reduced Bacterial Densities as Phage-Resistance Strategy

Perhaps the simplest approach to reducing arrangement vulnerability to phages would be for those bacteria to exist at lower population densities, population densities, that is, which are insufficient to support phage population growth to levels at which those phages can substantially impact arrangement fitness. In other words, bacteria that are *less* able to attain "winner" densities within a given environment—as in "kill the winner" (Section 1)—may as a consequence be better able to exploit biofilm niches. This idea is similar to the conclusions of Ioannou *et al.* [36], as described in Section 2.1.1, who suggest that one means by which prey species can offset the costs of possessing greater individual sizes is by displaying lower population densities such that, as a prey type, they effectively are a greater distance, on average, from their predator and therefore less likely to be

desirable to the predator. A major difference between the scenario presented by Ioannou *et al.* and that of kill the winner, however, is seen in terms of the degree of specialization of the predator species. For phages, specialization can be extreme such that bacteria phage-susceptibility types can "hide" by failing to support phage replication to inundative densities, that is, by not being "winners". For the visual predators considered by Ioannou *et al.*, by contrast, it is not that prey avoid predation by impacting predator densities but instead that they avoid being consumed by being less visible to predators that already exist at some more or less fixed density.

Biofilms often can occupy only a small fraction of total environmental volumes, such as in aquatic environments. In such circumstances, biofilm bacteria as a consequence may be inherently unable to achieve winner-level densities across an environment. This low potential, furthermore, may be particularly the case to the extent that biofilms consist of mixed populations of bacteria rather than monocultures of specific phage susceptibility types, thereby implying even lower densities of individual bacterial phage susceptibility types than available surfaces might maximally hold. Alternatively, biofilms found within highly structured environments, such as undisturbed soils [18], may over time tend to be mostly sequestered within micro-localities away from phages to which they are susceptible. Physically associated clonal groups of bacteria, such as can make up biofilms, in other words, may not be so much inherently resistant to phage attack as either relatively unexposed to specific phages or inherently less able to support the population growth of those phages to inundative densities [9].

An additional possibility may hold: Those bacteria which as a matter of luck are resistant to all phages present within environments might be able to attain and then sustain "winner" densities, either as free bacteria or as arrangements. This latter idea is ecologically similar to the domination of animal-rich ecosystems by plants that are resistant to forage by those herbivores which happen to be present within an environment [72]. That is, bottom-up control on bacterial density (nutrient availability) rather than top-down control (predator prevalence) would be expected to operate within ecosystems where predators are lacking whereas a combination of bottom-up and top-down control may hold when differences exist among prey in terms of their resistance to predation [70]. We might therefore predict a surfeit of bacteria existing as arrangements under three distinguishable circumstances: (i) where the abundance of specific phage susceptibility types of bacteria generally is low (potentially top-down control, 'kill the winner', and high bacterial diversity); (ii) where phage-resistant organisms dominate (combination of bottom-up and top-down control, and potentially lower bacterial diversity); or (iii) where phages simply are absent (bottom-up control with the diversity of bacteria therefore determined by factors other than phage-mediated predation).

3. Experimental Section

See Results and Discussion.

4. Conclusions

Group living bestows benefits—else why live in groups?—but also engenders costs. One cost comes from an increased vulnerability to exploitation that group living creates, such as a greater potential for infection of individual bacteria by bacteriophages. Here I have provided a simple model for quantifying those costs, one that points to the idea that bacterial arrangements along with microcolonies and biofilms may persist particularly under circumstances where high rates of lytic phage infection are unlikely. This conclusion is broadly consistent with Murray and Jackson's [24] suggestion, pp. 113 and 114, that "…for any given viral concentration, a large particle is more likely to have a virus reach its surface… Any given virus, however, is far more likely to reach a small, presumably bacterial, particle than a larger one in the ocean because of the greater abundance of small particles…" That is, individual prey size and vulnerability are not the only variables controlling the susceptibility of prey *populations* to predators, with prey population density also playing a key ecological role.

This perspective contrasts with notions that group living among bacteria might *directly* serve as a means of avoiding phage-mediated predation. Observation of the existence of specific bacterial types at high environmental densities and particularly as arrangements, microcolonies, or biofilms, however, would suggest the existence of effective mechanisms by which their vulnerability to phages has been reduced, just as pathogen-resistance mechanisms have, of course, evolved in multicellular organisms such as animals [32]. One reasonable scenario explaining such a situation is that diversity may exist within bacterial communities in terms of phage susceptibility even when that diversity is not superficially obvious. In this case, densities of individual bacteria types wouldn't be as high as they would appear—perhaps as a consequence of "kill the winner" (loss of bacterial populations existing at higher environmental densities)—because densities of specific phage susceptibility types would not be appreciably high. That situation represents the default assumption that one might make with regard to large, relatively open systems, that is, stabilizing frequency-dependent selection acting on rare phage susceptibility types that results in substantial bacterial diversity particularly in terms of phage resistance [73].

Alternatively, phages to which high-density bacterial populations are sensitive simply may not have reached those bacterial populations. That situation represents the default assumption that one might make with regard to small, relatively closed systems. The utility of phage therapy as an antibacterial strategy is that in many instances the infection to be treated can be described in the latter terms: small, relatively closed systems, such as localized or even systemic bacterial infections. Efficient bacterial eradication, including of bacterial biofilms, often can be achieved, therefore, simply by "opening" these systems sufficiently to phages, and the phage therapy literature, such as in terms of successful clinical treatment of infections in humans [74,75], appears to be consistent with that scenario.

Acknowledgments

Conflicts of Interest

The author declares no conflict of interest.

References

1. Stoodley, P.; Sauer, K.; Davies, D.G.; Costerton, J.W. Biofilms as complex differentiated communities. *Ann. Rev. Microbiol.* **2002**, *56*, 187–209.

2. Kjelleberg, S.; Givskov, M. *The Biofilm Mode of Life: Mechanisms and Adaptations*; Horizon Biosciences: Norfolk, UK, 2007.

3. Ramage, G.; Culshaw, S.; Jones, B.; Williams, C. Are we any closer to beating the biofilm: Novel methods of biofilm control. *Curr. Opin. Infect. Dis.* **2010**, *23*, 560–566.

4. Gino, E.; Starosvetsky, J.; Kurzbaum, E.; Armon, R. Combined chemical-biological treatment for prevention/rehabilitation of clogged wells by an iron-oxidizing bacterium. *Environ. Sci. Technol.* **2010**, *44*, 3123–3129.

5. Cos, P.; Tote, K.; Horemans, T.; Maes, L. Biofilms: An extra hurdle for effective antimicrobial therapy. *Curr. Pharm. Des.* **2010**, *16*, 2279–2295.

6. Abedon, S.T. Kinetics of phage-mediated biocontrol of bacteria. *Foodborne Pathog. Dis.* **2009**, *6*, 807–815.

7. Loc-Carrillo, C.; Abedon, S.T. Pros and cons of phage therapy. *Bacteriophage* **2011**, *1*, 111–114.

8. Abedon, S.T. Bacteriophages and biofilms. In *Biofilms: Formation, Development and Properties*; Bailey, W.C., Ed.; Nova Science Publishers: Hauppauge, NY, USA, 2010; Chapter 1, pp. 1–58.

9. Abedon, S.T. *Bacteriophages and Biofilms: Ecology, Phage Therapy, Plaques*; Nova Science Publishers: Hauppauge, NY, USA, 2011.

10. Abedon, S.T.; Thomas-Abedon, C.; Thomas, A.; Mazure, H. Bacteriophage prehistory: Is or is not Hankin, 1896, a phage reference? *Bacteriophage* **2011**, *1*, 174–178.

11. Twort, F.W. An investigation on the nature of ultra-microscopic viruses. *Lancet* **1915**, *ii*, 1241–1243.

12. Twort, F.W. An investigation on the nature of ultra-microscopic viruses. *Bacteriophage* **2011**, *1*, 127–129.

13. Abedon, S.T.; Yin, J. Impact of spatial structure on phage population growth. In *Bacteriophage Ecology*; Abedon, S.T., Ed.; Cambridge University Press: Cambridge, UK, 2008; Volume 15, Chapter 4, pp. 94–113.

14. Krone, S.M.; Abedon, S.T. Modeling phage plaque growth. In *Bacteriophage Ecology*; Abedon, S.T., Ed.; Cambridge University Press: Cambridge, UK, 2008; Volume 15, Chapter 16, pp. 415–438.

15. Abedon, S.T.; Yin, J. Bacteriophage plaques: Theory and analysis. *Meth. Mol. Biol.* **2009**, *501*, 161–174.

16. Gallet, R.; Shao, Y.; Wang, I.N. High adsorption rate is detrimental to bacteriophage fitness in a biofilm-like environment. *BMC Evol. Biol.* **2009**, *9*, 241.

17. Gallet, R.; Kannoly, S.; Wang, I.N. Effects of bacteriophage traits on plaque formation. *BMC Microbiol.* **2011**, *11*, 181.

18. Abedon, S.T. Communication among phages, bacteria, and soil environments. In *Biocommunication of Soil Microorganisms*; Witzany, G., Ed.; Springer: New York, NY, USA, 2011; Volume 23, Chapter 2, pp. 37–65.

19. Abedon, S.T.; Thomas-Abedon, C. Phage therapy pharmacology. *Curr. Pharm. Biotechnol.* **2010**, *11*, 28–47.

20. Gill, J.J.; Hyman, P. Phage choice, isolation and preparation for phage therapy. *Curr. Pharm. Biotechnol.* **2010**, *11*, 2–14.

21. Lu, T.K.; Collins, J.J. Dispersing biofilms with engineered enzymatic bacteriophage. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 11197–11202.

22. Goodridge, L.D. Designing phage therapeutics. *Curr. Pharm. Biotechnol.* **2010**, *11*, 15–27.

23. Mondes, R.; O'Toole, G.A. The developmental model of microbial biofilms: Ten years of a paradigm up for review. *Trends Microbiol.* **2009**, *17*, 73–87.

24. Murray, A.G.; Jackson, G.A. Viral dynamics: A model of the effects of size, shape, motion, and abundance of single-celled planktonic organisms and other particles. *Mar. Ecol. Prog. Ser.* **1992**, *89*, 103–116.

25. Abedon, S. Phage therapy pharmacology: Calculating phage dosing. *Adv. Appl. Microbiol.* **2011**, *77*, 1–40.

26. Thingstad, T.F. Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems. *Limnol. Oceanogr.* **2000**, *45*, 1320–1328.

27. Thingstad, T.F.; Bratbak, G.; Heldal, M. Aquatic Phage Ecology. In *Bacteriophage Ecology*; Abedon, S.T., Ed.; Cambridge University Press: Cambridge, UK, 2008; Chapter 10, pp. 251–280.

28. Chao, L.; Levin, B.R.; Stewart, F.M. A complex community in a simple habitat: An experimental study with bacteria and phage. *Ecology* **1977**, *58*, 369–378.

29. Schrag, S.; Mittler, J.E. Host-parasite persistence: The role of spatial refuges in stabilizing bacteria-phage interactions. *Am. Nat.* **1996**, *148*, 348–377.

30. Hyman, P.; Abedon, S.T. Bacteriophage host range and bacterial resistance. *Adv. Appl. Microbiol.* **2010**, *70*, 217–248.

31. Labrie, S.J.; Samson, J.E.; Moineau, S. Bacteriophage resistance mechanisms. *Nat. Rev. Microbiol.* **2010**, *8*, 317–327.

32. Abedon, S.T. Bacterial 'immunity' against bacteriophages. *Bacteriophage* **2012**, *2*, 50–54.

33. Abedon, S.T. Lysis of lysis inhibited bacteriophage T4-infected cells. *J. Bacteriol.* **1992**, *174*, 8073–8080.

34. Abedon, S.T. Envisaging bacteria as phage targets. *Bacteriophage* **2011**, *1*, 228–230.

35. Stent, G.S. *Molecular Biology of Bacterial Viruses*; WH Freeman and Co.: San Francisco, CA, USA, 1963.

36. Ioannou, C.C.; Bartumeus, F.; Krause, J.; Ruxton, G.D. Unified effects of aggregation reveal larger prey groups take longer to find. *Proc. Biol. Sci* **2011**, *278*, 2985–2990.

37. Abedon, S.T. Phage Population Growth: Constraints, Games, Adaptation. In *Bacteriophage Ecology*; Abedon, S.T., Ed.; Cambridge University Press: Cambridge, UK, 2008; Volume 15, Chapter 3, pp. 64–93.

38. Kasman, L.M.; Kasman, A.; Westwater, C.; Dolan, J.; Schmidt, M.G.; Norris, J.S. Overcoming the phage replication threshold: A mathematical model with implications for phage therapy. *J. Virol.* **2002**, *76*, 5557–5564.

39. Abedon, S.T. Bacteriophage T4 resistance to lysis-inhibition collapse. *Genet. Res.* **1999**, *74*, 1–11.

40. Young, K.D. The selective value of bacterial shape. *Microbiol. Mol. Biol. Rev.* **2006**, *70*, 660–703.

41. Young, K.D. Bacterial morphology: Why have different shapes? *Curr. Opin. Microbiol.* **2007**, *10*, 596–600.

42. Azeredo, J.; Sutherland, I.W. The use of phages for the removal of infectious biofilms. *Curr. Pharm. Biotechnol.* **2008**, *9*, 261–266.

43. Turner, G.F.; Pitcher, T.J. Attack abatement: A model for group protection by combined avoidance and dilution. *Am. Nat.* **1986**, *128*, 228–240.

44. Babic, A.; Berkmen, M.B.; Lee, C.A.; Grossman, A.D. Efficient gene transfer in bacterial cell chains. *MBio* **2011**, *2*, e00027-11.

45. Abedon, S.T. Phages, ecology, evolution. In *Bacteriophage Ecology*; Abedon, S.T., Ed.; Cambridge University Press: Cambridge, UK, 2008; Volume 15, Chapter 1, pp. 1–28.

46. Hagens, S.; Loessner, M.J. Bacteriophage for biocontrol of foodborne pathogens: Calculations and considerations. *Curr. Pharm. Biotechnol.* **2010**, *11*, 58–68.

47. Goodridge, L.D. Phages, bacteria, and food. In *Bacteriophage Ecology*; Abedon, S.T., Ed.; Cambridge University Press: Cambridge, UK, 2008; Chapter 12, pp. 302–331.

48. Xavier, J.B.; Foster, K.R. Cooperation and conflict in microbial biofilms. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 876–881.

49. Wimpenny, J. Ecological determinants of biofilm formation. *Biofouling* **1996**, *10*, 43–63.

50. Nadell, C.D.; Xavier, J.B.; Levin, S.A.; Foster, K.R. The evolution of quorum sensing in bacterial biofilms. *PLoS Biol.* **2008**, *6*, e14.

51. Crombach, A.; Hogeweg, P. Evolution of resource cycling in ecosystems and individuals. *BMC Evol. Biol.* **2009**, *9*, 122.

52. Hoffmeister, M.; Martin, W. Interspecific evolution: Microbial symbiosis, endosymbiosis and gene transfer. *Environ. Microbiol.* **2003**, *5*, 641–649.

53. Searcy, D.G. Metabolic integration during the evolutionary origin of mitochondria. *Cell Res.* **2003**, *13*, 229–238.

54. Kreft, J.U. Biofilms promote altruism. *Microbiology* **2004**, *150*, 2751–2760.

55. Nadell, C.D.; Bassler, B.L. A fitness trade-off between local competition and dispersal in *Vibrio cholerae* biofilms. *Proc. Natl. Acad. Sci USA* **2011**, *108*, 14181–14185.

56. Takahashi, N.; Nyvad, B. The role of bacteria in the caries process: Ecological perspectives. *J. Dent. Res.* **2011**, *90*, 294–303.

57. Kolter, R. Biofilms in lab and nature: A molecular geneticist's voyage to microbial ecology. *Int. Microbiol.* **2010**, *13*, 1–7.

58. Whiteley, M.; Ott, J.R.; Weaver, E.A.; McLean, R.J.C. Effects of community composition and growth rate on aquifer biofilm bacteria and their susceptibility to betadine disinfection. *Environ. Microbiol.* **2001**, *3*, 43–52.

59. Matz, C. Biofilms as refuge against predation. In *The Biofilm Mode of Life: Mechanisms and Adaptations*; Kjelleberg, S., Givskov, M., Eds.; Horizon Bioscience: Norfolk, UK, 2007; Chapter 11, pp. 195–213.

60. Thurlow, L.R.; Hanke, M.L.; Fritz, T.; Angle, A.; Aldrich, A.; Williams, S.H.; Engebretsen, I.L.; Bayles, K.W.; Horswill, A.R.; Kielian, T. *Staphylococcus aureus* biofilms prevent macrophage phagocytosis and attenuate inflammation *in vivo*. *J. Immunol.* **2011**, *186*, 6585–6596.

61. Arciola, C.R. Host defense against implant infection: The ambivalent role of phagocytosis. *Int. J. Artif. Organs* **2010**, *33*, 565–567.

62. Jensen, P.O.; Givskov, M.; Bjarnsholt, T.; Moser, C. The immune system *vs. Pseudomonas aeruginosa* biofilms. *FEMS Immunol. Med. Microbiol.* **2010**, *59*, 292–305.

63. Costerton, J.W.; Cheng, J.-J.; Geesey, G.G.; Ladd, T.I.; Nickel, J.C.; Dasgupta, M.; Marrie, T.J. Bacterial biofilms in nature and disease. *Ann. Rev. Microbiol.* **1987**, *41*, 435–464.

64. Barron, B.A.; Fischetti, V.A.; Zabriskie, J.B. Studies of the bacteriophage kinetics of multicellular systems: A statistical model for the estimation of burst size per cell in streptococci. *J. Appl. Bacteriol.* **1970**, *33*, 436–442.

65. Friend, P.L.; Slade, A.D. Characteristics of group A streptococcal bacteriophages. *J. Bacteriol.* **1966**, *92*, 148–154.

66. Fischetti, V.A.; Barron, B.; Zabriskie, J.B. Studies on streptococcal bacteriophages. I. burst size and intracellular growth of group A and group C streptococcal bacteriophages. *J. Exp. Med.* **1968**, *127*, 475–488.

67. Kaplan, D.A.; Naumovski, L.; Rothschild, B.; Collier, R.J. Appendix: A model of plaque formation. *Gene* **1981**, *13*, 221–225.

68. Doolittle, M.M.; Cooney, J.J.; Caldwell, D.E. Tracing the interaction of bacteriophage with bacterial biofilms using fluorescent and chromogenic probes. *J. Indust. Microbiol.* **1996**, *16*, 331–341.

69. Abedon, S.T. Bacteriophage Intraspecific Cooperation and Defection. In *Contemporary Trends in Bacteriophage Research*; Adams, H.T., Ed.; Nova Science Publishers: Hauppauge, NY, USA, 2009; Chapter 7, pp. 191–215.

70. Bohannan, B.J.M.; Lenski, R.E. Effect of prey heterogeneity on the response of a food chain to resource enrichment. *Am. Nat.* **1999**, *153*, 73–82.

71. Bohannan, B.J.M.; Lenski, R.E. Linking genetic change to community evolution: Insights from studies of bacteria and bacteriophage. *Ecol. Lett.* **2000**, *3*, 362–377.

72. Terborgh, J.; Lopez, L.; Nunez, V.; Rao, M.; Shahabuddin, G.; Orihuela, G.; Riveros, M.; Ascanio, R.; Adler, G.H.; Lambert, T.D.; *et al*. Ecological meltdown in predator-free forest fragments. *Science* **2001**, *294*, 1923–1926.

73. Abedon, S.T. Phage evolution and ecology. *Adv. Appl. Microbiol.* **2009**, *67*, 1–45.

74. Kutter, E.; De Vos, D.; Gvasalia, G.; Alavidze, Z.; Gogokhia, L.; Kuhl, S.; Abedon, S.T. Phage therapy in clinical practice: Treatment of human infections. *Curr. Pharm. Biotechnol.* **2010**, *11*, 69–86.

75. Abedon, S.T.; Kuhl, S.J.; Blasdel, B.G.; Kutter, E.M. Phage treatment of human infections. *Bacteriophage* **2011**, *1*, 66–85.

Article

Bacteriophages with the Ability to Degrade Uropathogenic *Escherichia Coli* Biofilms

Andrew Chibeu [1,*], **Erika J. Lingohr** [1], **Luke Masson** [2,3], **Amee Manges** [4], **Josée Harel** [5], **Hans-W. Ackermann** [6], **Andrew M. Kropinski** [1,7,*] **and Patrick Boerlin** [8]

[1] Laboratory for Foodborne Zoonoses, Public Health Agency of Canada, Guelph, ON, N1G 3W4, Canada; E-Mail: erika.lingohr@phac-aspc.gc.ca

[2] Biotechnology Research Institute, National Research Council of Canada, 6100 Royalmount Avenue, Montréal, QC H4P 2R2, Canada; E-Mail: luke.masson@cnrc-nrc.gc.ca

[3] Département de microbiologie et immunologie, Université de Montréal, 2900, boul. Édouard-Montpetit, Montréal, QC H3T 1J4, Canada

[4] Department of Epidemiology, Biostatistics and Occupational Health, McGill University, 1020 avenue des Pins Ouest, Montréal, QC H3A 1A2, Canada; E-Mail: amee.manges@mcgill.ca

[5] Groupe de Recherche sur les Maladies Infectieuses du Porc (GREMIP) and Centre de Recherche en infectiologie porcine (CRIP), Université de Montréal, Faculté de médecine vétérinaire, Saint-Hyacinthe, QC J2S 7C6, Canada; E-Mail: josee.harel@umontreal.ca

[6] Felix d'Herelle Reference Center for Bacterial Viruses, Department of Microbiology, Immunology and Infectionlogy, Faculty of Medicine, Laval University, QC G1K 4C6, Canada; E-Mail: ackermann@mcb.ulaval.ca

[7] Department of Molecular and Cellular Biology, University of Guelph, ON N1G 2W1, Canada

[8] Department of Pathobiology, Ontario Veterinary College, University of Guelph, ON N1G 2W1, Canada; E-Mail: pboerlin@uoguelph.ca

* Authors to whom correspondence should be addressed;
E-Mails: andrew.chibeu@phac-aspc.gc.ca (A.C.); kropinsk@queensu.ca (A.M.K.);
Tel.: +1-519-826-2616; Fax: +1-519-822-2280.

Abstract: *Escherichia coli*-associated urinary tract infections (UTIs) are among the most common bacterial infections in humans. UTIs are usually managed with antibiotic therapy, but over the years, antibiotic-resistant strains of uropathogenic *E. coli* (UPEC) have

emerged. The formation of biofilms further complicates the treatment of these infections by making them resistant to killing by the host immune system as well as by antibiotics. This has encouraged research into therapy using bacteriophages (phages) as a supplement or substitute for antibiotics. In this study we characterized 253 UPEC in terms of their biofilm-forming capabilities, serotype, and antimicrobial resistance. Three phages were then isolated (vB_EcoP_ACG-C91, vB_EcoM_ACG-C40 and vB_EcoS_ACG-M12) which were able to lyse 80.5% of a subset (42) of the UPEC strains able to form biofilms. Correlation was established between phage sensitivity and specific serotypes of the UPEC strains. The phages' genome sequences were determined and resulted in classification of vB_EcoP_ACG-C91 as a *SP6likevirus*, vB_EcoM_ACG-C40 as a *T4likevirus* and vB_EcoS_ACG-M12 as *T1likevirus*. We assessed the ability of the three phages to eradicate the established biofilm of one of the UPEC strains used in the study. All phages significantly reduced the biofilm within 2–12 h of incubation.

1. Introduction

Urinary tract infections (UTIs) are among the most common bacterial infections in humans. They account for more than seven million visits to physicians' offices per year in the United States of America [1]. Approximately 20% of women develop UTIs sometime during their lifetime. Above the age of 50, men and women have a similar incidence of UTIs [1–5]. UTIs have tremendous economic impact on health care systems in both direct and indirect costs associated with treatment [6,7]

Uropathogenic strains of *Escherichia coli* (UPECs) account for about 75%–85% of UTIs [4,8]. There has been an evolution toward antibiotic resistance in UPECs, with decreasing susceptibility to first-line agents such as ampicillin, nitrofurantoin, sulphamethoxazole/trimethoprim (SXT) and fluoroquinolones [4,9–13]. An alternative to antimicrobial treatment would be of great relevance to treat multiresistant UTIs [5], as well as to avoid the continuous selection of resistant pathogens and to safeguard the efficacy of antimicrobial agents for cases of emergency and life-threatening infections.

In hospital-acquired infections as well as in geriatrics patients, most UTI are associated with indwelling catheters which act as foci for biofilm formation [14]. Biofilms play a significant role in the ability of bacteria to withstand killing by host immune responses and antibiotics. There is a need to develop new therapeutic strategies to eradicate biofilm infections [15].

The development of biofilms by bacteria does not protect cells from bacteriophage killing. These viruses can penetrate the extracellular matrix that binds macromolecules and prevents their diffusion into the biofilm [16–18] and kill cells [19–21]. Furthermore, certain phages have evolved to deal with capsules by possessing virion-associated polysaccharide depolymerases [22–24] leading us to believe that they are ideal agents to help reduce biofilm-associated infections as well as to kill planktonic cells.

In this paper, we describe a subset of a collection of 253 *E. coli* isolates from UTIs specifically selected on their ability to form biofilms and the activity of three environmentally isolated bacteriophages on these strains. To assess their efficacy and safety as therapeutic agents against UPEC, the phages were characterized in detail with regards to their spectrum of activity, genetic structure, as well as activity on biofilms.

2. Results and Discussion

2.1. Biofilm Forming Capabilities, Serotype and Antimicrobial Resistance

Each of the 253 UTI *E. coli* isolates was screened for static biofilm formation in 96-well untreated polystyrene microtiter plates following growth in artificial urine medium as described in the Experimental section. Isolates were considered positive for biofilm formation if the crystal violet-stained biofilm had an OD_{600} equal to or greater than 3-fold the value obtained in the well containing bacteria-free medium. The UTI *E. coli* strain CTF073 was used as a positive control and 12 wells of this strain were included with each plate of samples of UPEC assayed for biofilm formation. The plotted results were the average of three independent experiments, each with six replicates per isolate. Only 42 out of the 253 *E. coli* isolates tested in this study produced biofilms in microtiter plate wells with the levels of biofilm formation varying among isolates (Figure 1). This is in contrast to a previous study [25] where a majority of the UTI strains tested produced biofilms *in vitro*. A possible reason for the different outcomes in the two studies is the fact that UPEC do not form biofilm well on polystyrene plates in urine or artificial urine medium as compared to M9 minimal medium used by [25]. Among the 42 biofilm-forming isolates, Can67 showed the highest level of biofilm formation (OD_{600} above 1.0) whereas MSHS94 and Can72 resulted in weak biofilms formation with an OD_{600} below 0.2.

The biofilm-forming UTI *E. coli* isolates totalized 22 different somatic O antigens and 14 different H antigens (Table 1). We then investigated their susceptibility profiles against 15 antimicrobials. The investigated isolates were frequently resistant with 32 (78%) isolates displaying resistance to antimicrobial agents, while 16 (39%) isolates displayed multidrug resistance phenotypes. These results are in agreement with previous studies that have shown a high frequency of antimicrobial resistance among UPEC [4,9–13]. However, inspection of the different levels of biofilm formation on microtiter plates (Figure 1) and comparing this to the results on antimicrobial susceptibility (Table 1) showed that there was no apparent direct correlation between antimicrobial resistance and the amount of biofilm formed by each isolate.

Figure 1. Urinary tract infection (UTI) *E. coli* isolates positive for biofilm formation on 96-well microtiter plates. Isolates were considered as biofilm formers if the OD_{600} for the crystal violet-stained biofilm was equal to or greater than 3-fold the OD_{600} for a bacteria free medium. Data points represent an average of three independent experiments, each with 6 replicate wells for each isolate tested. Isolate CTF073 is a positive control for biofilm formation. Error bars indicate standard error of means.
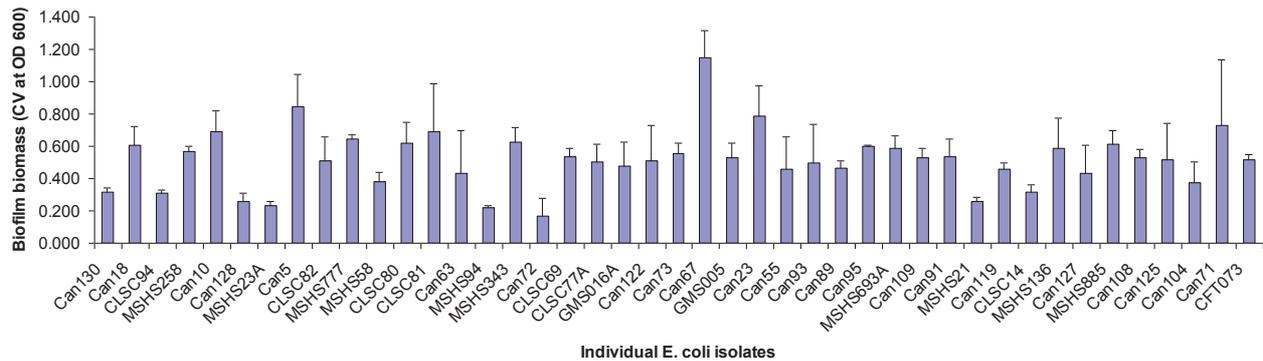
Table 1. List of UTI *E. coli* forming biofilms, their serotypes, respective antimicrobials resistances and sensitivity to the three isolated phages ACG-C91, ACG-C40 and ACG-M12. AMC = Amoxicillin/Clavulanic acid; AMP = Ampicillin; FOX = Cefoxitin; TIO = Ceftiofur; CRO = Ceftriaxone; CHL = Chloramphinicol; CIP = Ciprofloxacin; KAN = Kanamycin; NAL = Nalidixic acid; STR = Streptomycin; SOX = Sulfisoxazole, TCY = Tetracycline; SXT = Sulphamethoxazole/Trimethoprim. + symbolises sensitivity to phage; - resistance to phage; nd, not determined.

Strain	Serotype	Antimicrobials to which isolates are resistant	ACG-C40	ACG-C91	ACG-M12
Can130	O1:NM		+	+	-
Can18	O2:H4		-	+	-
CLSC94	O2:H7		+	+	-
MSHS258	O2:H18		+	+	-
Can10	O8:H10		-	-	-
Can128	O8:NM	SOX, TCY	+	-	-
MSHS23A	O6:H1	AMP, STR, SOX,	+	-	+
Can5	O6:H1	NAL	+	-	+
CLSC82	O6:H1	AMP	+	-	+
MSHS777	O6:H1	STR	+	-	+
MSHS58	O6:H25	KAN	-	-	-
CLSC80	O11:H18	STR, SOX	+	-	-
CLSC81	O14:H4		+	-	+
Can63	O14:H31	TCY	-	-	+
MSHS94	O18ac:H7	AMP, STR, SOX, TCY	+	+	+

Table 1. *Cont.*

Strain	Serotype	Antimicrobials to which isolates are resistant	ACG-C40	ACG-C91	ACG-M12
MSHS343	O18ac:H7		+	+	+
Can72	O18ac:NM	AMP	+	-	-
CLSC69	O21:H14		-	-	-
CLSC77A	O22:H1		-	-	-
GMS016A	O25:H1		+	+	+
Can122	O25:H4	AMP, TIO, CRO, CIP, NAL, SOX, TCY, SXT	-	-	-
Can73	O25:H4	AMP, CIP, NAL	-	-	-
Can67	O25:H4	AMC, AMP, CIP, NAL	-	-	-
GMS005A	O35:H10		+	-	-
Can23	O68:H18		+	-	-
Can55	O75:H7	AMC, AMP, FOX, TIO, CRO	+	+	-
Can93	O75:H7		-	+	-
Can89	O78:H5		+	-	-
Can95	O106:H18	CHL, TCY	-	-	-
MSHS693A	O117:H5		+	-	-
Can109	O134:H31	NAL	+	+	-
Can91	O135:H6		+	+	+
MSHS21	O135:H11		+	-	-
Can119	O153:H18	AMP, TIO, CRO, NAL, STR, SOX, SXT	+	+	-
CLSC14	O153:NM	AMP, STR, SOX, TCY, SXT	+	-	-
MSHS136	O166:H15	AMC, AMP	-	-	-
Can127	OR:H4	AMP,CHL, TCY	-	+	+
MSHS885	OR:H4	AMP, CHL, STR, SOX, TCY	+	+	-
Can108	OR:H4	SOX, TCY, STX	+	+	-
Can125	OR:H4	AMP, CIP, NAL, SOX, SXT	+	-	+
Can104	OR:H7		-	+	+
Can71	OR:H40		+	-	+
CFT073	nd	nd	+	+	+

It is important to note that the ability of biofilms to withstand antibiotic killing is a function of their mode of growth. Antimicrobial agents have been shown to penetrate biofilms at different rates depending on the particular agent and the biofilm [26]. The antimicrobial susceptibility tests carried out in this study using the broth microdilution method did not test the susceptibility of the bacteria to the antimicrobial agent after biofilm formation. The antimicrobials were added to cell suspensions at time of media inoculation and not after the biofilms had formed. The results of [27] demonstrated that older biofilms of *E. coli* resisted ampicillin treatment to a greater extent than their younger counterparts. However, in our study, a majority of the biofilm forming *E. coli* UTI isolates (78%) were resistant to at least one antimicrobial agent. This compares to 53.1% of the non-biofilm forming *E. coli*

UTI isolates that were resistant to at least one out of the 15 antimicrobials tested (results not shown). This is consistent with what is known about antimicrobial resistance among biofilm forming bacterial strains [28].

2.2. Phage Morphology

UPEC strains Can 91, Can 40 and MSHS1210 were used as propagating strains for the isolation of phages from serially diluted CsCl purified phage concentrate. All three isolates were from midstream urine specimens with Can 40 and Can 91 being part of the CANWARD study whereas MSHS1210 was from the McGill University student health service. The single unique phages isolated from UPEC strains Can 91, Can 40 and MSHS1210 were vB_EcoP_ACG-C91, vB_EcoM_ACG-C40 and vB_EcoS_ACG-M12, respectively.

Transmission electron microscopy revealed that phage ACG-C40 has an elongated head, a neck, and a contractile tail with tail fibers (Figure 2a). Its head is 110×82 nm, while the extended tail is 114×8 nm. Phage ACG-M12 (Figure 2b) has an isometric head of about 157 nm in diameter between opposite apices and a relatively flexible tail of 172×7 nm, which terminate in 1–2 fibers of 12 nm in length. Phage ACG-C91 (Figure 2c) has an isometric head of 65–68 nm and a short tail of 12×8 nm, which carries fibers of 13 nm in length with a terminal swelling. Based on their morphology, ACG-C40 is classified as a member of the family *Myoviridae*, ACG-M12 is part of the *Siphoviridae* and ACG-C91 belongs to the family *Podoviridae*.

Figure 2. Negative staining of phages ACG-C91 (**a**); ACG-M12 (**b**) and ACG-C40 (**c**) with 2% uranyl acetate or 2% phosphotungstate. Final magnification is × 297,000. Bars indicate 100 nm.
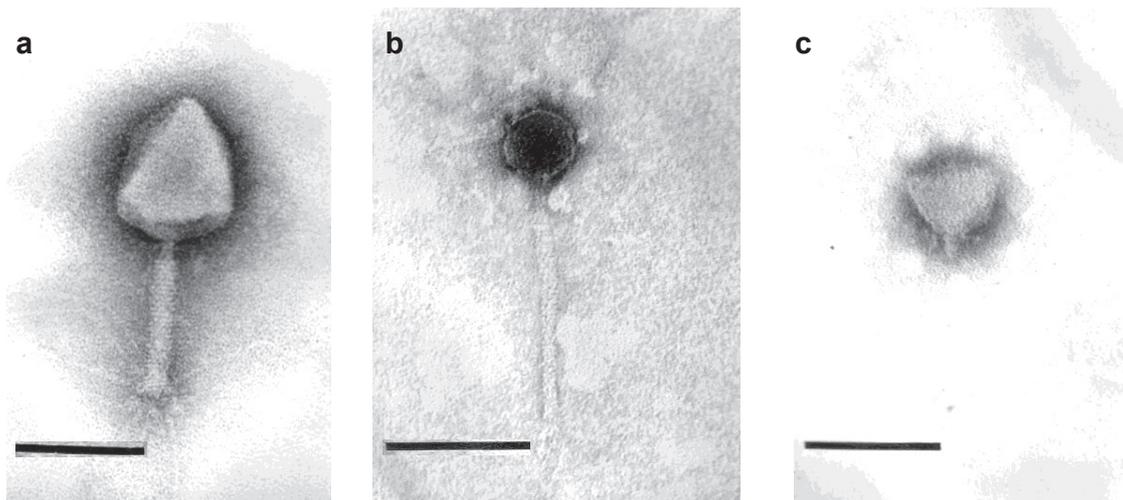
2.3. Lytic Spectra of Phages against Biofilm Forming UTI Isolates and Relation to Isolates' Serotypes

The lytic activity of ACG-C91, ACG-C40 and ACG-M12 tested against the biofilm forming *E. coli* UTI isolates. In the spot tests, 33 (80.5%) of the biofilm forming *E. coli* UTI isolates were infected by at least one of the three phages (Table 1).

Out of the three phages isolated in this study, phage ACG-C40 had the broadest host range, with the ability to lyse 28 (66.67%) of the biofilm forming UTI *E. coli* isolates tested (Table 1). It is not unusual for phages in the family *Myoviridae* to have such a large host range. A previous study by [29], showed that bacteriophage T4 lysed 41.1% of the 69 clinical isolates of *E. coli* tested. Other T4-like phages such as AR1 have been observed to have similar broad host ranges [30]. The ability of phage ACG-C40 to cause lysis in a wide range of *E. coli* isolates with different serotypes makes it an important candidate for phage therapy applications. The broad host range of UTI *E. coli* isolates is an indicator of how well adapted phage ACG-C40 is to infecting UPEC, which in itself is a desirable property for a phage to be used in phage therapy against such strains. The criteria of choosing phages which are well adapted for infection of targeted hosts for therapeutic purposes have been applied in similar previous studies [31–33].

The correlation between host serotype and phage sensitivity is seen in the fact that among the biofilm forming strains, all the O25:H4 strains were resistant to the three tested phages. The O6:H1 strains were sensitive to phages ACG-C40 and ACG-M12. Further in depth studies to identify the receptors of the tested phages and the receptors spatial orientation on the surface of the host strains under tested conditions would give insight as to why the phages are resistant or susceptible to all strains of a given serotype.

2.4. Salient Bacteriophage Genome Features

Genome sequencing and bioinformatics analysis of the three isolated phages revealed that all three are non-temperate and that none carries any known bacterial virulence genes for humans or animals. None of the three phages possessed a demonstrable extracellular polysaccharide or exopolysaccharide (EPS) depolymerase gene. The importance of phage-associated depolymerases in biofilm eradication has been recognized in previous studies [22–24,34]. Depolymerase-producing phages were found to be more effective in eradicating mature biofilms than non-depolymerase producing phages [34]. Phages producing depolymerases have thus been used in concert with antimicrobials to facilitate deeper penetration of antimicrobials by degrading the EPS [23,34]. However, phages that do not produce EPS depolymerases have also found use in biofilm degradation. Such phages include naturally occurring phages such as T4 [16,17,35] and phages such as T7 engineered to express recombinant dispersin B (DsbB) [36].

2.4.1. Phage ACG-C40

The sequence of the phage genome consisted of 167,396 bp (G+C content 35%) which is close to the genome size of other "T4-like viruses" [37]. It was predicted to encode 282 ORFs and 10 tRNAs. The latter clustered between 68, 043 and 68, 989 bp on the phage's genome.

The DNA of this virus was resistant to digestion by all restriction endonucleases tested. An analysis of the genome of this phage (Additional file: Figure A1; Table A1) reveals that it does, like coliphage T4, encode, a glucosyl transferase (orf 060). Bacteriophage T4 contains glucosylated 5-hydroxymethylcytosine (hmdCyt) instead of cytosine which makes its DNA resistant to most restriction enzymes except for EcoRV and NdeI [38]. Based on the presence of the glucosyl transferase gene and the resistance to enzymes, we conclude that bacteriophage ACG-C40 also contains hypermodified bases, probably glucosylated hmdCyt.

The genome of this phage revealed that the phage potentially encodes two versions of the Hoc protein; a shorter and a longer version which may result from translational frameshifting generating Hoc protein with C termini of different lengths (supplementary file: Table S1).

2.4.2. Phage ACG-C91

Phage ACG-C91 has a 43,731 bp genome with G+C content of 45% and predicted 55 ORFs (supplementary file: Table S2). The phage DNA was sensitive to restriction enzymes BglI, XbaI, EcoRV, EcoRI, NdeI and SalI.

2.4.3. Phage ACG-M12

Phage ACG-M12 has a 46, 054 bp genome with a G+C content of 44% and 77 ORFs (supplementary file: Table S3). The phage DNA was sensitive to all restriction enzymes tested except NaeI and SmaI.

2.5. Comparative Genomics

Pairwise comparisons using CoreGenes 3.0 [39] at default stringency setting ("75") revealed that at the protein level, ACG-C40, ACG-C91 and ACG_M12 were 85.9% ,78.8% and 77.3% similar to bacteriophages T4, SP6 and RTP, respectively. Bacteriophage ACG-C40 genome encodes 239 proteins homologous to T4 proteins, ACG-C91 genome has 41 encoded proteins homologous to those of SP6 and ACG-M12 genome encodes 58 proteins with homology to RTP proteins (supplementary files: Table A1; Table A2; Table A3).

A Mauve alignment [40,41] of phage ACG-C91 DNA against the genome of phage SP6 shows regions of sequence similarity (coloured red) with two regions where there are no DNA sequence similarities (coloured white). The regions of non-sequence similarity correspond to the regions encoding phage ACG-C91 internal virion protein proteins (orf 38), putative tail protein (orf 41) and the endosialidase (orf 53). The genome region encoding the endosialidase is lacking altogether in

phage SP6 (supplimentary file: Table A1; supplementary file: Figure A2a). This protein is found in K1-specific phages, such as K1F and K1-5, that encode virion-associated endosialidases hydrolyzing the K1 polysialic acid structure of the K1 capsule-producing *E.coli* strains. Mauve alignment of ACG-M12 against the genome of phage RTP was also performed and it revealed regions of sequence similarity (red) interspersed with four regions where no sequence similarity exists (white). The regions of genome homology disparity include phage ACG-M12 genome regions encoding: hypothetical proteins orf2, orf3 and orf4 (1025 bp–1588 bp), the major tail protein orf26 (12435 bp–13091 bp), putative tail fibre orf47 (29255–32074) and conserved hypothetical proteins orf62 and orf63 (supplementary file: Figure A2b).

2.5. Bacteriophage Eradication of Established Biofilms

UTI *E. coli* isolate Can 91 was selected to determine the effectiveness of the isolated phages to degrade established biofilms because of the isolate's ability to form strong biofilms in microtiter wells (OD_{600} 0.538) after 48 hrs and its sensitivity to all the three phages isolated. Phages ACG-C40, ACG-C91 and ACG-M12 treatment of preformed UTI *E. coli* Can 91 biofilms yielded reductions (based on OD_{600} measurement of CV-stained cells) in biofilm mass compared to untreated controls (Figure 3). All of the three phages displayed, at low concentrations (10^5 PFU/mL), a steady reduction of biofilm biomass over time with the largest percentage reduction of biofilm biomass being realized after 8 h incubation. The same pattern was also observed at higher phage concentrations (10^7 PFU/mL and 10^9 PFU/mL). At all phage concentrations, it was evident that the biofilm started re-establishing itself after 24 h (Figure 3).

These results indicate that under the conditions tested, the biofilm reduction by the phages is not dose-dependent, a fact that could be taken advantage of if the phages are to be applied *in vivo* to remove established biofilms. It would mean that low titers of the phages were as effective as using a higher titer in eradicating established biofilms. Re-establishment of biofilms after 24 h exposure to the phages *in vitro* may be attributed to development of bacterial resistance against the phages. This could be overcome by using a cocktail of phages. It may also be a good strategy to augment phage treatment with chemical antimicrobials which will be effective in prevention of the re-establishment of biofilms by the phage resistant cells.

Further studies need to be carried out to assess the efficacy of phage ACG-C40, ACG-C91 and ACG-M12 in the prevention of biofilm formation for comparison with this study where we have assayed their efficacy in the eradication of established biofilms.

Figure 3. Phage disruption of established *E. coli* strain Can 91 biofilm. Biofilms grown in polystyrene microtiter plate wells for 48 h, were initially inoculated with 10^5, 10^7 and 10^9 pfu of phages (**a**) ACG-C40 and (**b**) ACG-C91 (**c**) ACG-M12. After 2, 4, 8 and 24 h of phage treatment at 37 °C, average biofilm biomass in corresponding microtiter plate wells, were scored relatively to untreated control samples (100%) and represented on the Y-axes. Three independent experiments were performed, each starting from a different overnight culture and each with six repeats for each parameter combination. Error bars indicate standard error of means.
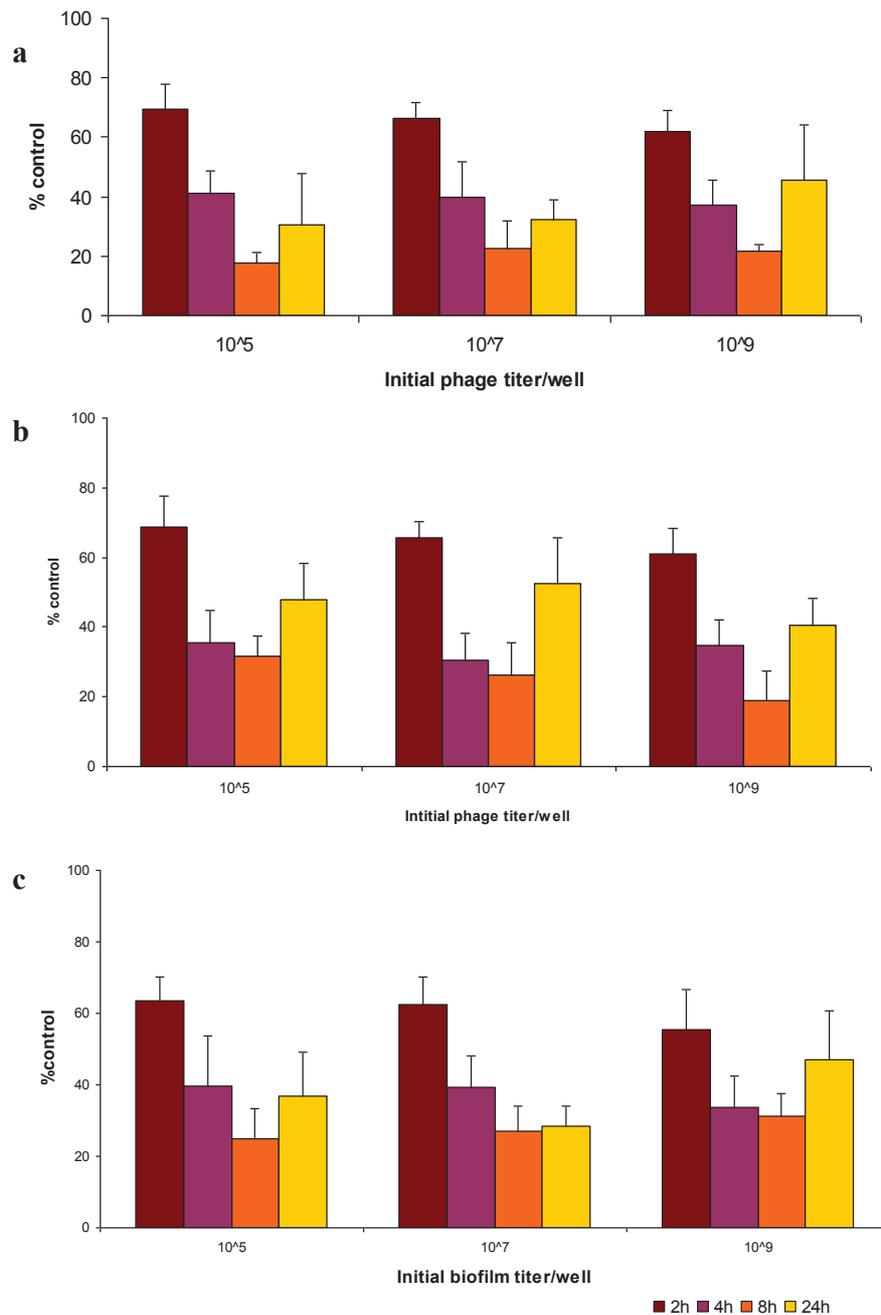
3. Experimental Section

3.1. Sampling and E. coli *Strain Isolation*

A group of 123 strains of UTI *E. coli* were isolated from women with community-acquired UTI either from a student health service (McGill University, Montréal, QC, Canada) or at a community health center also in Montréal. To ensure no bias in the selection of isolates due to recurrent UTI or treatment failures, only one isolate per patient was used.

An additional 130 isolates were part of the ongoing Canadian national surveillance study (CANWARD) for testing urinary culture pathogens for antimicrobial susceptibilities. The study involved out-patients attending hospital clinics and emergency rooms, and in patients on medical and surgical wards, and in intensive care units [42]. Each study site was asked to submit only clinically significant isolates, as defined by local site criteria, from out- and in-patients with urinary infections. Fifty (2009) to 100 (2007, 2008) consecutive urinary tract isolates per year per medical center site (one isolate per patient) were collected. Primary isolate identification was performed by the submitting medical center site and confirmed by the coordinating laboratory, as required, based on morphological characteristics and spot tests [43]. If an isolate identification made by the coordinating laboratory did not match that provided by the submitting site, the isolate was removed from the study. All isolates were stored at −70 °C in skim milk.

At the Department of Epidemiology, Biostatistics and Occupational Health of McGill University, the isolates from the McGill University student health service and the community health center, also in Montréal, as well as those from the CANWARD study were recovered on MacConkey and CLED agar (Uricult dipslides: Orion Diagnostica, Oy, Finland). A single representative colony was picked from the MacConkey side of the Uricult slide and grown in LB broth and archived in 15% glycerol at −80 °C. When needed, further testing, based upon indole production and the presence of lysine or ornithine decarboxylase, was also performed to confirm the bacterial identification.

3.2. Culture Conditions

For all experiments, bacteria were grown in LB medium or synthetic urine medium. The latter medium was based on that devised by [44], and its composition and method of sterilization have been described previously [45]. Briefly the medium was composed of $CaCl_2$, 0.65 g/L; $MgCl_2$, 0.65 g/L; NaCl, 4.6 g/L; Na_2SO_4, 2.3 g/L; Na_3-citrate, 0.65 g/L; Na_2-oxalate, 0.02 g/L; KH_2PO_4, 2.8 g/L; KCl, 1.6 g/L; NH_4Cl, 1.0 g/L; urea, 25 g/L; creatinine, 1.1 g/L; Luria-Bertani (LB) broth (Difco, Sparks, MD, USA), 10 g/L. The pH of the medium is adjusted to pH 5.8 prior to sterilization by filtration through a 0.2-µm filter.

3.3. E. coli *Serotyping*

Identification of somatic (O) and flagellar (H) antigens was performed by standard agglutination methods [46]. The isolates were serotyped at the Public Health Agency of Canada, Laboratory for Foodborne Zoonoses, Guelph, ON, Canada.

3.4. Susceptibility Testing

The broth microdilution method was used (Sensititre System; Trek Diagnostics, Cleveland, OH, USA), following the protocols of the Canadian Integrated Program for Antimicrobial Resistance Surveillance [47,48]. Each *E. coli* isolate was tested for the following antimicrobial agents (breakpoints are indicated in parentheses): amikacin ($\geq$64 µg/mL), amoxicillin-clavulanic acid ($\geq$32 and $\geq$16 µg/mL, respectively), ampicillin ($\geq$32 µg/mL), cefoxitin ($\geq$32 µg/mL), ceftiofur ($\geq$8 µg/mL), ceftriaxone ($\geq$2 µg/mL) (11), chloramphenicol ($\geq$32 µg/mL), ciprofloxacin ($\geq$4 µg/mL), gentamicin ($\geq$16 µg/mL), kanamycin ($\geq$64 µg/mL), naladixic acid ($\geq$32 µg/mL), streptomycin ($\geq$64 µg/mL), sulfisoxazole ($\geq$512 µg/mL), tetracycline ($\geq$16 µg/mL), and trimethoprim-sulfamethoxazole ($\geq$4 and $\geq$76 µg/mL, respectively).

3.5. Bacteriophage Isolation

Three liters of preliminary treated sewage sample from the Guelph water treatment plant were centrifuged at 3,000 RPM for 20 min to eliminate solid debris. The clarified supernatant was concentrated to 400 mL using the tangential flow system as described in [49]. NaCl (20.5 g) was then added to the concentrate and stirred until it dissolved before 40 g of polyethylene glycol was added to precipitate the phage present in the concentrate. This was left overnight while being stirred gently with a magnetic stirrer at 4 °C. Precipitated phage was recovered by centrifugation at 3,000 RPM for 10 min and resuspended in 3 mL SM buffer [50]. This was subjected to two rounds of CsCl gradient purification steps and the resulting viral band isolated.

E. coli isolates Can 91, Can 40 and MSHS1210 were used as propagating strains for the isolation of single plaques from serially diluted CsCl purified phage concentrate. All three isolates were from midstream urine specimens with Can 40 and Can 91 being part of the CANWARD study whereas MSHS1210 was from the McGill University student health service.

LB agar (Difco) plugs from standard agar overlay method [51] were resuspended in 500 µL SM buffer, diluted and used for further single plaque isolation on the respective *E. coli* propagation strains. This was repeated three times for each strain to ensure unique phages for each UPEC isolate. The single unique phages isolated from *E. coli* isolates Can 91, Can 40 and MSHS1210 were named vB_EcoP_ACG-C91, vB_EcoM_ACG-C40 and vB_EcoS_ACG-M12, respectively, following the naming convention of [52].

3.6. Transmission Electron Microscopy

Phages were pelleted at 25,000 × g for 1 h, using a Beckman high-speed centrifuge and a JA-18.1 fixed-angle rotor (Beckman, Palo Alto, CA, USA). Phage pellets were washed twice under the same conditions in neutral 0.1 M ammonium acetate and resuspended. The phages were then deposited on copper grids with carbon-coated Formvar films and stained with 2% uranyl acetate (pH 4) or 2% phosphotungstate (pH 7.2). They were examined in a Philips EM electron microscope. Magnification was calibrated using T4 phage tails as size standards.

3.7. Bacteriophage DNA Isolation, Restriction Analysis and Sequencing

To separate phage from bacterial debris, crude phage lysates were centrifuged at 10,000 × g for 15 min at 4 °C. Bacterial nucleic acids in the supernatants were digested with pancreatic DNase 1 and RNase A, each to a final concentration of 10 μg/mL (Sigma-Aldrich Canada Ltd., Oakville, ON, USA). Phage particles were then precipitated in the presence of 10% w/v (final concentration) polyethylene glycol (PEG-8000; Sigma-Aldrich) at 4 °C overnight. The precipitated phage particles were recovered by centrifugation, resuspended in TM buffer (10 mM Tris-HCl, pH 7.8, 1 mM $MgSO_4$). The DNA was extracted from a portion of the precipitated viral particles using the SDS/proteinase K method modified from [50], followed by extraction with phenol:chloroform:isoamyl alcohol (25:24:1, vol/vol), ethanol precipitation and resolution in 10 mM Tris-HCl (pH 7.5). The DNA was characterized spectrophotometrically.

Phage DNA was digested with restriction enzymes BamHI, BglI, EcoRV, HindIII, NaeI, XbaI, NdeI, PstI, and SmaI (New England Biolabs, Hertfordshire, UK), according to the manufacturer's instructions. Phage Lambda DNA (Fermentas Inc., Hanover, MD, USA) digested by the same enzymes was used as a positive control for the restriction digest analysis.

The DNA was subjected to pyrosequencing (454 technology) at the McGill University and Génome Québec Innovation Centre (Montréal, QC, Canada) to between 195 and 930 fold coverage.

3.8. Bioinformatic Analysis

The sequences were rearranged to resemble their GenBank homologs and annotated using MyRAST [53]. The generated gbk file was incorporated into Kodon (Applied Maths, Austin, TX, USA) and proof-read. Proteins were generated using gbk2faa program [54] and screened for homologs using BLAST algorithm [55]. In addition they were searched for conserved motifs in the Pfam database [56] and transmembrane helix predictions using TMHMM [57] and Phobius [58]. The sequence was also screened for tRNAs using tRNAscan-SE program [59]. The final sequences were converted to Sequin format [60] using gbk2sqn [61] before deposition into GenBank. The accession numbers are ACG-C91 JN986844, ACG-M12 JN986845 and ACG-C40 JN986846.

3.9. Biofilm Assay

Each of the 253 *E. coli* isolates was screened for biofilm formation using the 96-well plate assay as described in [62], with minor modifications. Briefly, *E. coli* isolates were used to inoculate 5 mL of synthetic urine medium and grown for 16 h at 37 °C. A 1:100 dilution of each of the cultures was made in synthetic urine media and 200 µL of each diluted culture was added to six wells in an untreated 96-well polystyrene flat-bottomed, 96-well microtiter plates (Costar; Corning Inc., Corning, NY, USA). The plate was covered and incubated 37 °C for 48 h without agitation. Negative control wells that contained 200 µL of sterile synthetic urine media only were included. Following incubation, planktonic bacteria were removed by pipetting out the culture and washing the wells twice with PBS to remove loosely bound cells, using a multichannel pipettor. A 1% (wt/vol; 210 µL) crystal violet (CV) solution was added to the wells for 10 min. The CV stain was then decanted and the unbound stain removed by washing the plate in a water tray. The plates were left to dry following which 200 µL of 95% ethanol added in each well to solubilize the bound CV from the stained *E. coli* biofilms. The absorbance of CV at 600 nm was measured in a Wallac-Victor^2 1420 Multilabel Counter (Perkin-Elmer, Boston, MA, USA).

3.10. Bacteriophage Host Range on Biofilm Forming UPEC Isolates

The host range of phages ACG-C91, ACG-C40 and ACG-M12 against biofilm forming *E. coli* UTI isolates was determined by standard spot tests [63]. Briefly, this involved mixing 200 µL overnight culture of each isolate with 3 mL molten top LB agar (0.6% agar) and pouring on 96 mm diameter plates containing bottom agar. This was left to solidify then 10 µL of phage to be tested was spotted in the middle of the plate and left to dry before incubation inverted overnight at 37 °C.

3.11. Bacteriophage Activity against Established Biofilms

E. coli isolate CLSC 94 was used to test the effect of the two environmentally isolated phages ACG-C40 and ACG-C91 on established biofilms. The method of [64] with minor modifications was used. Briefly, stationery-phase culture of *E. coli* CLSC94 was diluted 1:100 and 200 µL of the inoculum added to 12 wells of an untreated 96-well polystyrene microtiter plate. This was incubated for 24 h at 37 °C without agitation. Negative control wells that contained 200 µL of sterile synthetic urine media only were included. Following incubation period, planktonic bacteria were removed by pipetting out the culture, and washing the wells twice with PBS to remove loosely bound cells. The wells were then treated with 200 µL of ACG-C40 and ACG-C91 phage lysate at titers of 10^5, 10^7 and 10^9 PFU/mL added to each of the 12 wells. Phages were also added to the negative control wells which contained only UPEC cells grown in sterile synthetic urine media. After incubation periods of 2 h, 4 h, 8 h and 24 h, planktonic bacteria were removed and wells were washed twice with PBS to remove loosely bound cells. The remaining biofilm was stained as described above.

4. Conclusions

We were able to isolate phages that are effective against UPECstrains and are able to degrade biofilms formed by the UPEC strains. The phages' DNA sequences were determined. They were screened and found to be devoid of undesirable laterally transferable virulence and antimicrobial resistance determinants on the basis of homologies with known virulence and resistant genes available in GenBank.

The study resulted in detailed characterization of a large set of UPEC with regards to biofilm forming capabilities and antimicrobial resistance. The candidate phages with adequate lytic spectrum for therapeutic purposes (including efficacy for disruption of existing biofilms) were characterized in detail. The phages have a promising potential for phage therapy of UTIs caused by biofilm forming UPEC.

Acknowledgments

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Foxman, B. Epidemiology of urinary tract infections: Incidence, morbidity, and economic costs. *Dis.-a-Mon.* **2003**, *49*, 53–70.
2. Sanford, J.P. Urinary tract symptoms and infections. *Ann. Rev. Med.* **1975**, *25*, 485.
3. Nicolle, L.E.; Ronald, A.R. Recurrent urinary tract infection in adult women: Diagnosis and treatment. *Infect. Dis. Clin. N. Am.* **1987**, *1*, 793–806.
4. Nicolle, L.E. Epidemiology of urinary tract infections. *Clin. Microbiol. Newslett.* **2002**, *24*, 135–140.
5. Czaja, C.A.; Hooton, T.M. Update on acute uncomplicated urinary tract infection in women. *Postgrad. Med.* **2006**, *119*, 39–45.
6. Foxman, B.; Barlow, R.; D'Arcy, H.; Gillespie, B.; Sobel, J.D. Urinary tract infection: Self-reported incidence and associated costs. *Ann. Epidemiol.* **2000**, *10*, 509–515.
7. Russo, T.A.; Johnson, J.R. Medical and economic impact of extraintestinal infections due to *Escherichia coli*: Focus on an increasingly important endemic problem. *Microbes Infect.* **2003**, *5*, 449–456.
8. Bacheller, C.D.; Bernstein, J.M. Urinary tract infections. *Med. Clin. N. Am.* **1997**, *3*, 719–730.
9. Gupta, K.; Hooton, T.M.; Stamm, W.E. Increasing antimicrobial resistance and the management of uncomplicated community-acquired urinary tract infections. *Ann. Intern. Med.* **2001**, *135*, 41–50.
10. Gupta, K. Addressing antibiotic resistance. *Am. J. Med.* **2002**, *113*, 29S–34S.

11. Karlowsky, J.A.; Kelly, L.J.; Thornsberry, C.; Jones, M.E.; Sahm, D.F. Trends in antimicrobial resistance among urinary tract infection isolates of *Escherichia coli* from female outpatients in the United States. *Antimicrob. Agents Chemother*. **2002**, *46*, 2540–2545.

12. Karlowsky, J.A.; Thornsberry, C.; Jones, M.E.; Sahm, D.F. Susceptibility of antimicrobial-resistant urinary *Escherichia coli* isolates to fluoroquinolones and nitrofurantoin. *Clin. Infect. Dis.* **2003**, *36*, 183–187.

13. Zhanel, G.G.; Hisanaga, T.L.; Laing, N.M.; DeCorby, M.R.; Nichol, K.A.; Palatnik, L.P.; Johnson, J.; Noreddin, A.; Harding, G.K.; Nicolle, L.E.; *et al*. NAUTICA Group. Antibiotic resistance in outpatient urinary isolates: Final results from the North American Urinary Tract Infection Collaborative Alliance (NAUTICA). *Int. J. Antimicrob. Agents* **2005**, *26*, 380–388.

14. Morris, N.S.; Stickler, D.J.; McLean, R.J. The development of bacterial biofilms on indwelling urethral catheters. *World J. Urol*. **1999**, *17*, 345–350.

15. Stewart, P.S.; Costerton, J.W. Antibiotic resistance of bacteria in biofilms. *Lancet* **2001**, *358*, 135–138.

16. Doolittle, M.M.; Cooney, J.J.; Caldwell, D.E. Lytic infection of *Escherichia coli* biofilms by bacteriophage T4. *Can. J. Microbiol*. **1995**, *42*, 12–18.

17. Doolittle, M.M.; Cooney, J.J.; Caldwell, D.E. Tracing the interaction of bacteriophage with bacterial biofilms using fluorescent and chromogenic probes. *J. Indust. Microbiol*. **1996**, *16*, 331–341.

18. Lacroix-Gueu, P.; Briandet, R.; Leveque-Fort, S.; Bellon-Fontaine, M.N.; Fontaine-Aupart, M.P. *In situ* measurements of viral particles diffusion inside mucoid biofilms. *C. R. Biol*. **2005**, *328*, 1065–1072.

19. Wood, H.L.; Holden, S.R.; Bayston, R. Susceptibility of *Staphylococcus epidermidis* biofilm in CSF shunts to bacteriophage attack. *Eur. J. Pediatr. Surg*. **2001**, *11*, S56–S57.

20. Sillankorva, S.; Oliveira, R.; Vieira, M.J.; Sutherland, I.W.; Azeredo, J. Bacteriophage φS1 infection of *Pseudomonas fluorescens* planktonic cells *versus* biofilms. *Biofouling* **2004**, *20*, 133–138.

21. Curtin, J.J.; Donlan, R.M. Using bacteriophages to reduce formation of catheter-associated biofilms by *Staphylococcus epidermidis*. *Antimicrob. Agents Chemother*. **2006**, *50*, 1268–1275.

22. Hughes, K.A.; Sutherland, I.W.; Jones, M.V. Biofilm susceptibility to bacteriophage attack: The role of phage-borne polysaccharide depolymerise. *Microbiology* **1998**, *144*, 3039–3047.

23. Hughes, K.A.; Sutherland, I.W.; Clark, J.; Jones, M.V. Bacteriophage and associated polysaccharide depolymerases-novel tools for study of bacterial biofilms. *J. Appl. Microbiol*. **1998**, *85*, 583–590.

24. Hanlon, G.W.; Denyer, S.P.; Ollif, C.J.; Ibrahim, L.J. Reduction in exopolysaccharide viscosity as an aid to bacteriophage penetration through *Pseudomonas aeruginosa* biofilms. *Appl. Environ. Microbiol*. **2001**, *67*, 2746–2753.

25. Rijavec, M.; Müller-Premru, M.; Zakotnik, B.; Zgur-Bertok, D. Virulence factors and biofilm production among *Escherichia coli* strains causing bacteraemia of urinary tract origin. *J. Med. Microbiol*. **2008**, *57*, 1329–1334.

26. Hoyle, B.D.; Alcantara, J.; Costerton, J.W. *Pseudomonas aeruginosa* biofilm as a diffusion barrier to piperacillin. *Antimicrob. Agents Chemother.* **1992**, *36*, 2054–2056.

27. Ito, A.; Taniuchi, A.; May, T.; Kawata, K.; Okabe, S. Increased antibiotic resistance of *Escherichia coli* in mature biofilms. *Appl. Environ. Microbiol.* **2009**, *75*, 4093–4100.

28. Gilbert, P.; Maira-Litran, T.; McBain, A.J.; Rickard, A.H.; Whyte, F.W. The physiology and collective recalcitrance of microbial biofilm communities. *Adv. Microbial. Physiol.* **2002**, *46*, 203–256.

29. Sillankorva, S.; Oliveira, D.; Moura, A.; Henriques, M.; Faustino, A.; Nicolau, A.; Azeredo, J. Efficacy of a broad host range lytic bacteriophage against *E. coli* adhered to urothelium. *Curr. Microbiol.* **2011**, *62*, 1128–1132.

30. Goodridge, L.; Gallaccio, A.; Griffiths, M.W. Morphological, host range, and genetic characterization of two coliphages. *Appl. Environ. Microbiol.* **2003**, *69*, 5364–5371.

31. Chibani-Chennoufi, S.; Sidoti, J.; Bruttin, A.; Dillmann, M.L.; Kutter, E.; Qadri, F.; Sarker, S.A.; Brüssow, H. Isolation of *Escherichia coli* bacteriophages from the stool of pediatric diarrhea patients in Bangladesh. *J. Bacteriol.* **2004**, *186*, 8287–8294.

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

33. Nishikawa, H.; Yasuda, M.; Uchiyama, J.; Rashel, M.; Maeda, Y.; Takemura, I.; Sugihara, S.; Ujihara, T.; Shimizu, Y.; Shuin, T.; *et al*. T-even-related bacteriophages as candidates for treatment of *Escherichia coli* urinary tract infections. *Arch. Virol.* **2008**, *153*, 507–515.

34. Verma, V.; Harjai, K.; Chibber, S. Structural changes induced by a lytic bacteriophage make ciprofloxacin effective against older biofilm of *Klebsiella pneumoniae*. *Biofouling* **2010**, *26*, 729–737.

35. Corbin, B.D.; McLean, R.J.; Aron, G.M. Bacteriophage T4 multiplication in a glucose-limited *Escherichia coli* biofilm. *Can. J. Microbiol.* **2001**, *47*, 680–684.

36. Lu, T.K.; Collins, J.J. Dispersing biofilms with engineered enzymatic bacteriophage. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 11197–11202.

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

38. Carlson, K.; Raleigh, E.A.; Hattman, S. Restriction and Modification. In *Bacteriophage T4*; Karam, J.D., Ed.; ASM Press: Washington, DC, USA, 1994; pp. 369–381.

39. Mahadevan, P.; King, J.F.; Seto, D. CGUG: *In silico* proteome and genome parsing tool for the determination of "core" and unique genes in the analysis of genomes up to ca. 1.9 Mb. *BMC Res. Notes* **2009**, *2*, 168.

40. Darling, A.C.; Mau, B.; Blattner, F.R.; Perna, N.T. Mauve: Multiple alignment of conserved genomic sequence with rearrangements. *Genome Res.* **2004**, *14*, 1394–1403.

41. Kropinski, A.M.; Borodovsky, M.; Carver, T.J.; Cerdeño-Tárraga, A.M.; Darling, A.; Lomsadze, A.; Mahadevan, P.; Stothard, P.; Seto, D.; Van Domselaar G.; *et al*. *In silico* identification of genes in bacteriophage DNA. *Methods Mol. Biol.* **2009**, *502*, 57–89.

42. Canadian Antimicrobial Resistance Alliance. Available online: http://www.can-r.com/ (accessed on 26 March 2012).

43. Zhanel, G.G.; Adam, H.J.; Low, D.E.; Blondeau, J.; Decorby, M.; Karlowsky, J.A.; Weshnoweski, B.; Vashisht, R.; Wierzbowski, A.; Hoban, D.J. Canadian antimicrobial resistance alliance (CARA). Antimicrobial susceptibility of 15,644 pathogens from Canadian hospitals: results of the CANWARD 2007–2009 study. *Diagn. Microbiol. Infect. Dis.* **2011**, *69*, 291–306.

44. Griffith, D.P.; Musher, D.M.; Itin, C. Urease: The primary cause of infection-induced urinary stones. *Investig. Urol.* **1976**, *13*, 346–350.

45. Stickler, D.J.; Morris, N.S.; Winters, C. Simple physical model to study the formation and physiology of biofilms on urethral catheters. *Methods Enzymol.* **1999**, *310*, 494–501.

46. Ewing, E.H. Genus *Escherichia*. In *Identification of Enterobacteriaceae*, 4th ed.; Edwards, P.R., Ewing, W.H., Eds.; Elsevier Science: New York, NY, USA, 1986; pp. 96–134.

47. Government of Canada. Canadian Integrated Program for Antimicrobial Resistance Surveillance (CIPARS) 2005. Public Health Agency of Canada: Guelph, Ontario, Canada, 2007.

48. Cockerill, F.R.; Wikler, M.A.; Bush, K. Performance standards for antimicrobial susceptibility testing; twentieth informational supplement M100-S20; Clinical and Laboratory Standards Institute: Wayne, PA, USA, 2010.

49. Casas, V.; Rohwer, F. Phage metagenomics. *Methods Enzymol.* **2007**, *421*, 259–268.

50. Sambrook, J.; Fritsch, E.F.; Maniatis, T. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor: New York, NY, USA, 1989; Volume 2.

51. Adams, M.H. Bacteriophages. Interscience Publishers: New York, NY, USA, 1959.

52. Kropinski, A.M.; Prangishvili, D.; Lavigne, R. Position paper: The creation of a rational scheme for the nomenclature of viruses of Bacteria and Archaea. *Environ. Microbiol.* **2009**, *11*, 2775–2777.

53. myRAST- The SEED Server. Available online: http://blog.theseed.org/servers/presentations/t1/drast-overview.html (accessed on 9 February 2012).

54. GBKFAA. Available online: http://lfz.corefacility.ca/gbk2faa/ (accessed on 9 February 2012).

55. Basic Local Alignment Search Tool (BLAST). Available online: http://blast.ncbi.nlm.nih.gov/ (accessed on 9 February 2012).

56. Finn, R.D.; Mistry, J.; Tate, J.; Coggill, P.; Heger, A.; Pollington, J.E.; Gavin, O.L.; Gunasekaran P.; Ceric, G.; Forslund, K.; *et al*. The Pfam protein families database. *Nucleic Acids Res.* **2010**, *38*, D211–D222.

57. Sonnhammer, E.L.; Von, H.G.; Krogh, A. A hidden Markov model for predicting transmembrane helices in protein sequences. In Proceedings of the International Conference on Intelligent Systems for Molecular Biology, Montreal, QC, Canada, 28 June–1 July 1998; pp. 175–182.

58. Kall, L.; Krogh, A.; Sonnhammer, E.L. A combined transmembrane topology and signal peptide prediction method. *J. Mol. Biol.* **2004**, 338, 1027–1036.

59. Lowe, T.M.; Eddy, S.R. tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* **1997**, *25*, 955–964.

64

60. Benson, D.A.; Karsch-Mizrachi, I.; Lipman, D.J.; Ostell, J.; Sayers, E.W. GenBank. *Nucleic Acids Res.* **2011**, *39*, D32–D37.

61. Genbank to Sequin File Converter. Available online: http://lfz.corefacility.ca/gbk2sqn/ (accessed on 9 February 2012).

62. O'Toole, G.A.; Pratt, L.A.; Watnick, P.I.; Newman, D.K.; Weaver, V.B.; Kolter, R. Genetic approaches to study of biofilms. *Methods Enzymol.* **1999**, *310*, 91–109.

63. Kutter, E. Phage Host Range and Efficiency of Plating. In *Bacteriophages: Methods and Protocols*; Clokie, M.R.J., Kropinski, A.M., Eds.; Humana Press: New York, NY, USA, 2009; Volume 1, pp. 141–151.

64. Soni, K.A.; Nannapaneni, R. Removal of *Listeria monocytogenes* biofilms with bacteriophage P100. *J. Food Prot*. **2010**, *73*, 1519–1524.

Communication

Lysogenic Conversion and Phage Resistance Development in Phage Exposed *Escherichia coli* Biofilms

Pieter Moons [†], David Faster and Abram Aertsen *

Laboratory of Food Microbiology, Department of Microbial and Molecular Systems (M²S), Faculty of Bioscience Engineering, Katholieke Universiteit Leuven, Kasteelpark Arenberg 22, Leuven 3001, Belgium; E-Mails: pieter.moons@ua.ac.be (P.M.); david.faster@biw.kuleuven.be (D.F.)

[†] Present address: Laboratory of Medical Microbiology, Vaccine & Infectious Disease Institute, Universiteit Antwerpen, Universiteitsplein 1, S6.21, Antwerpen 2610, Belgium.

* Author to whom correspondence should be addressed; E-Mail: abram.aertsen@biw.kuleuven.be; Tel.: +32-16-321752; Fax: +32-16-321960.

Abstract: In this study, three-day old mature biofilms of *Escherichia coli* were exposed once to either a temperate Shiga-toxin encoding phage (H-19B) or an obligatory lytic phage (T7), after which further dynamics in the biofilm were monitored. As such, it was found that a single dose of H-19B could rapidly lead to a near complete lysogenization of the biofilm, with a subsequent continuous release of infectious H-19B particles. On the other hand, a single dose of T7 rapidly led to resistance development in the biofilm population. Together, our data indicates a profound impact of phages on the dynamics within structured bacterial populations.

Keywords: *Escherichia coli*; biofilm; lysogenic conversion; resistance development; Shiga toxin; phage

1. Introduction

Many bacterial populations are organized as biofilms, which consist of cells attached to a surface embedded in a matrix of variable composition [1]. These biofilms structurally vary from flat layers of

cells to complex structured communities, consisting of tower or mushroom shaped micro-colonies interspersed with water channels that allow access of nutrients and removal of metabolites [2–4]. Bacterial biofilms present a medical hazard, since they confer increased resistance against antimicrobials and the host immune system [5], and are linked to persistent infections [6–8].

Apart from being a predominant species among the facultative anaerobic bacteria in the gastrointestinal tract [9], *E. coli* is considered a major zoonotic food-borne pathogen [10] and is one of the main causes of nosocomial infections [11], such as those associated with biofilm formation on urinary catheters [12]. Based on their disease-associated virulence factors, the species is divided into pathotypes, which are known to cause diarrhoeal disease, extra-intestinal and urinary tract infections, sepsis and meningitis [13,14]. Virulence characteristics include a variety of adhesion and colonization factors, the formation of attaching and effacing lesions, the production of toxins and the presence of antibiotic resistance genes [13,15,16]. Among diarrheagenic *E. coli*, those producing the potent Shiga (or Vero) toxins (Stxs) are the most virulent and can cause the potentially fatal haemolytic uremic syndrome [17,18]. Genes encoding Stxs are typically harbored by temperate phages (Stx-phages) integrated in the genome of such pathogenic *E. coli* strains, which makes them subject to lateral gene transfer [19,20].

The worldwide emergence of multi-drug resistant pathogens and empty antibiotic development pipelines reduce medical treatment options and necessitate the research into alternative therapies. One such option is phage therapy, which makes use of lytic phages as natural bacterial enemies whose narrow host range minimizes their impact on the normal flora [21–24]. In fact, phages are well suited to affect biofilms, since infection leads to local enrichment of viral particles. Moreover, many bacteriophages possess enzymes capable of bacterial lysis and less common biofilm matrix degradation [22,25,26], aiding in the accessibility of biofilm cells toward viral particles [27,28] and leading to biofilm dispersal [29–31].

In this study, we examined the impact of a temperate or lytic phage on mature *E. coli* biofilms, with particular interest in the lateral transfer of virulence determinants and phage resistance development, respectively.

2. Results and Discussion

2.1. Lysogenic Conversion of a Mature E. coli Biofilm with an Stx-Encoding Phage

In order to examine to which extent phage encoded virulence factors could be captured and spread within an existing *E. coli* biofilm, we decided to make acquisition of the naturally *stx1*-encoding temperate H-19B phage [32] readily detectable through selective plating by equipping it with an antibiotic resistance marker. Since only a small fragment of the H-19B genome sequence is currently known [32], a random transposon mutagenesis procedure was followed to tag this (pro)phage. More specifically, *E. coli* MG1655 was first lysogenized with H-19B, after which a random Tn10-transposon library of ca. 10,000 clones was constructed in this lysogen using the λNK1324 hop protocol described by Kleckner *et al.* [33]. Assuming some of the Tn10 insertions to be located within

the H-19B prophage, mitomycin C was subsequently used to induce the prophage in the obtained pool of transposon mutants. To isolate H-19B::Tn10 mutants within the corresponding phage lysate, it was first plaqued on *E. coli* MG1655. Since the Tn10 transposon codes for the *cat* gene and confers chloramphenicol resistance, lysogens arising in the middle of turbid plaques were scored for the presence or absence of chloramphenicol resistance. As such, nine chloramphenicol resistant lysogens were obtained, which were further confirmed to simultaneously have acquired the *stx1* operon of H-19B. The main advantage of this procedure is that it automatically disregards phage mutants compromised in lytic or lysogenic development. Moreover, in contrast to phage recombineering protocols [34], this protocol can be applied without prior knowledge of the phage's genome sequence and resembles the method used previously by Acheson *et al.* [35] to look for phage encoded exported proteins.

From one of the obtained H-19B::Tn10 lysogens, the Tn10 insertion site was determined and found to map within a gene bearing homology to the *nleG* virulence genes. NleG proteins are effectors of the type 3 secretion system that are thought to mimic eukaryotic E3 ubiquitin ligases [36], and are generally found in the late region of phage genomes [37]. This insertion underscores the presence of additional virulence genes to be present in H-19B, supports the observation that neither lytic nor lysogenic behaviour is affected in H-19B::Tn10 and due to the uptake of an additional piece of DNA (*i.e,.* the transposon) adds evidence to the plasticity of Stx genomes [38]. Subsequently, a mature three-day old *E. coli* biofilm was only once exposed to a small number (*i.e.*, 150 viral particles spread over one hour) of the corresponding H-19B::Tn10 derivative in order to mimic an accidental exposure. Subsequently, the spread of this virulence conferring phage genome throughout the biofilm population was tracked. More specifically, the effluent of the biofilm was examined for the presence of (i) cells lysogenized with H-19B::Tn10, and (ii) free H-19B::Tn10 phage (Figure 1). Interestingly, from this analysis it became clear that lysogenic conversion of the biofilm proceeded very rapidly, with the emergence of *ca.* 10^6 CFU/mL of H-19B::Tn10 lysogens on a total effluent cell count of circa 10^9 CFU/mL (*i.e.*, 0.06% conversion) after 24 h, and an above 50% conversion reached after five days. Moreover, at the end of the experiment the ratio between total and lysogenized cells within the actual biofilm itself corresponded to that observed in the effluent, demonstrating that, at least at the end of the experiment, lysogenic conversion had occurred in the entire biofilm Furthermore, during the first two days after phage exposure, the concentration of free H-19B::Tn10 phage in the effluent quickly rose from the applied 50 PFU/mL to 10^5–10^6 PFU/mL, after which it remained stable at this level. It can be anticipated that part of the biofilm is being lysed, leading to a local enrichment in H-19B::Tn10 concentration. Nevertheless, it remains unclear exactly to what extent the emergence of H-19B::Tn10 lysogens is the result of *de novo* lysogenic conversion of pre-existing wild-type cells or of clonal enrichment of the first converted cells.

Figure 1. Numbers of viable cells (CFU/mL) and phages (PFU/mL) in the effluent of a mature *E. coli* biofilm exposed for 1 hour to 50 PFU/mL (total of 150 PFU per biofilm) of H-19B::Tn10 phage at day 0, after sampling the effluent for initial total cell count. Dark blue bars represent total cell concentration, while light blue bars represent the concentration of H-19B::Tn10 lysogens. White markers represent free H-19B::Tn10 phage. Dark and light grey bars represent total cell counts and lysogen cell counts, respectively, in the attached biofilm obtained after dismantling the flow cell setup. The data represent the average and standard deviations of three biological replicates.
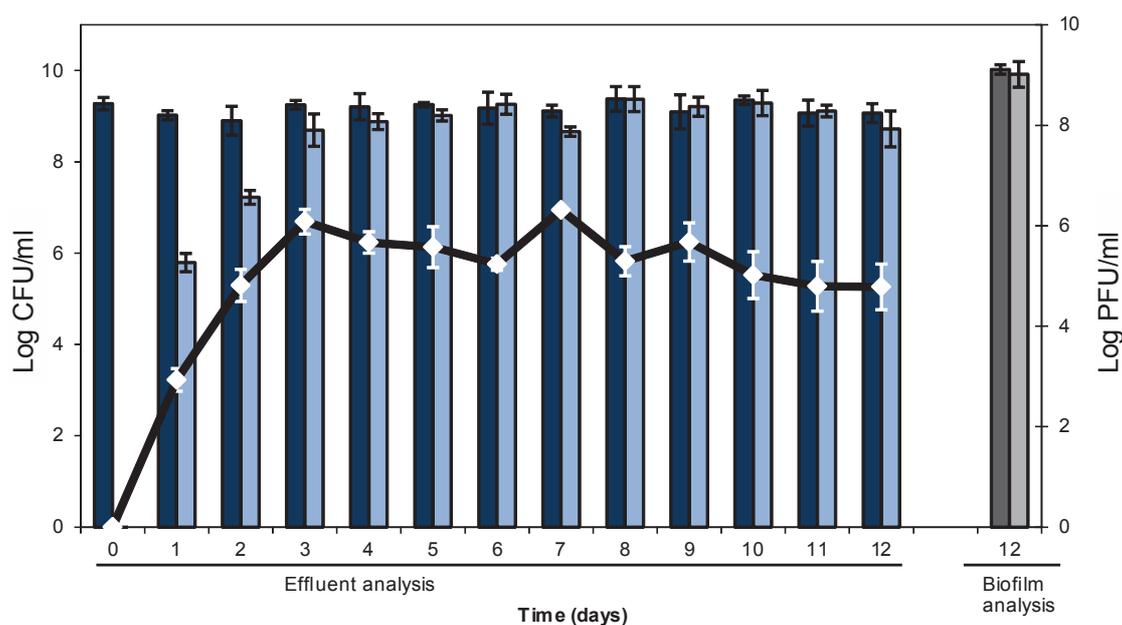
2.2. Resistance Development of E. coli *biofilms Against a Lytic Phage*

In a second approach, biofilm dynamics were examined upon single exposure of a 3-day old mature *E. coli* biofilm to *ca.* 10^8 particles per milliliter of the obligatory lytic T7 phage. The high phage titer selected in this approach was considered to best reflect the clinical application of phage in combatting biofilm-related infections. After three additional days of incubation, cells within the biofilm were harvested and enumerated. A comparison between the number of viable cells within the T7-exposed or -unexposed biofilm revealed an almost 100-fold reduction (*p*-value in two tailed homoscedastic T-test <0.05) in the phage exposed biofilm (Figure 2). However, in between the normal colonies observed during plating, colonies with a mucoid appearance were observed in biofilms treated with T7 phages (Figure 3A), but not in untreated biofilms. This phenotype was earlier shown to be correlated with T7 resistance due to a mutation resulting in the production of excess capsular polysaccharide preventing T7 adhesion [39] due to physical blocking of the phage binding site [40], and actual T7 resistance of these colonies could be confirmed as in Figure 3B. After examining *ca*. 400 colonies of both types, phage resistant clones with a non-mucoid phenotype or phage sensitive clones with a mucoid phenotype were not observed. In turn, this enabled us to accurately determine the number of

resistant cells in the biofilm. However, a large variation in the number of resistant cells was found between biological replicates (Table 1), ranging from circa 0.05% to over 28%. As in Luria-Delbruck fluctuation experiments, this variation likely stemmed from clonal enrichment of stochastically pre-existing T7-resistant mutants that spontaneously arise within a population. Since 3 days after initial phage exposure T7 particles still remained present in the effluent and the biofilm, on-going selection for T7-resistance development seems warranted upon longer incubation of the biofilm.

Figure 2. Number of viable cells (CFU) within a mature 3-day old *E. coli* biofilm exposed once, for one hour, to either water (control; dark bars) or *ca.* 10^8 PFU/mL of T7 phage (light bars) followed by an additional three days of biofilm development. The data represent the average and standard deviations of three biological replicates.
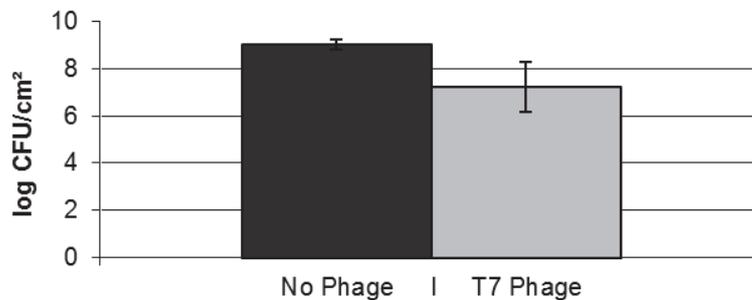
Figure 3. (**A**) mucoid (bottom) and non mucoid (top) colony of *E. coli*. (**B**) Phage resistance of a bacterial colony was demonstrated by streaking a line of *E. coli* cells through a perpendicularly streaked line of T7 solution. T7 phage resistant cells (bottom) were capable of growth past this line, while T7 sensitive cells (top) were not.
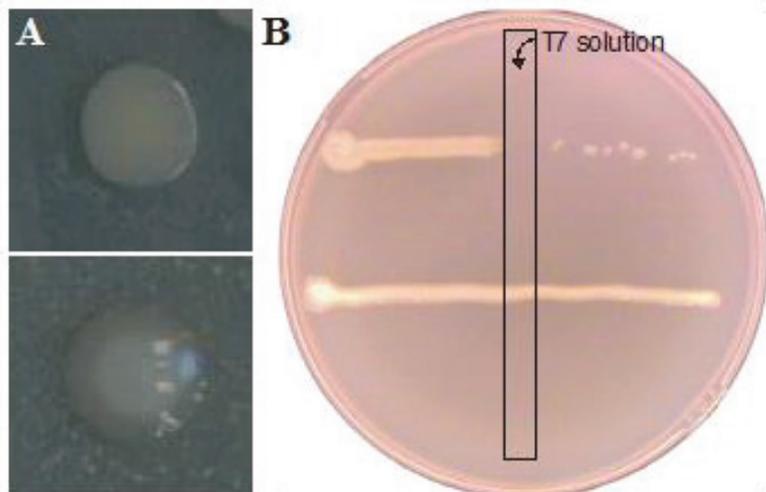
Table 1. Viable cell counts within three mature three-day old *E. coli* biofilms each exposed once for one hour to *ca.* 10^8 PFU/mL of T7 phage, followed by three additional days of growth before plating.

Cell counts (log CFU/cm²)	Biofilm 1	Biofilm 2	Biofilm 3
Total cell count	7.31	5.82	7.91
T7 resistant cells	6.76	2.49	7.05
T7 sensitive cells	7.17	5.82	7.85

3. Experimental Section

3.1. Strains, Standard Culture Conditions and Chemicals

E. coli MG1655 [41], MG1655 *lacZ::Tc* [42], and MG1655 H-19B (*i.e.*, lysogenized with *stx1* encoding phage H-19B; [43]) were used in this study, and were cultured at 37 °C under shaking conditions in Lysogeny Broth (LB). In order to determine the number of bacteria (as colony forming units or CFU) in a sample, a dilution series was prepared and plated on LB agar (1.5% agar), after which the concentration was expressed as CFU/mL.

To obtain a H-19B phage lysate from lysogens, cells from 1 mL bacterial culture were pelleted by centrifugation ($6000 \times$ g, 15 min), after which 50 µL of $CHCl_3$ was added to sterilize the supernatant and to release progeny phage from infected but non-lysed cells. In case high phage titers were required, the cultures were grown in the presence of 2 µg/mL mitomycin C which induces the release of temperate phages. To obtain a T7 phage [44] lysate, T7 was propagated on MG1655, after which 1 mL of cells were pelleted by centrifugation ($6000 \times$ g, 15 min) and the resulting supernatant was sterilized by addition of 50 µL of $CHCl_3$. In order to determine the number of phages (as plaque forming units or PFU) in a sample of lysate or effluent, a dilution series was prepared and plated on *E. coli* MG1655 grown in LB soft agar (0.7% agar), after which the concentration was expressed as PFU/mL.

Where necessary, growth media were supplemented with tetracycline (20 µg/mL) and/or chloramphenicol (30 µg/mL) (Applichem, Darmstadt, Germany).

3.2. Analysis of H-19B::Tn10

After its construction as described in the text, the Tn10 insertions site in H-19B::Tn10 was mapped by subcloning of the Tn10 encoded chloramphenicol resistance marker in pUC18. Subsequently, genomic DNA flanking the Tn10 transposon was sequenced using a primer (5'-AAGCACCGCCGGACATC-3') reading outwards of the transposon. In addition, the isolated H-19B::Tn10 phage was also confirmed by PCR (directly on plaques or crude lysate) to still carry its stx1 locus, by using primers flanking this region (5'-CAGTGGATCCTGGCACGGAAACATGGGT-3' and 5'-TCAGTCTAGATTACGTCTTTGCAGT CGAGAAGTC-3').

3.3. Setup for Biofilm Formation

Biofilms were grown at 30 °C in three-channel flow chambers (Biocentrum DTU: Technical University of Denmark, Soltofts Plads 221 DK-2800 Kgs, Lyngby) with individual channel dimensions of $1 \times 4 \times 40$ mm^3 [45] that were covered with a microscope glass coverslip (st1; Knittel Gläser, Braunschweig, Germany). The setup makes use of a 16-channel peristaltic pump (Watson Marlow 205S, Zellik, Belgium) that feeds each channel with a flow of 3 mL/h (flow rate of 0.2 mm/s) of AB-trace medium [2 g/L $(NH_4)_2SO_4$; 6 g/L $Na_2HPO_4.2H_2O$; 3 g/L KH_2PO_4; 3 g/L NaCl; 9.5 mg/L $MgCl_2$; 11.1mg/L $CaCl_2$ and 0.1 mL/L of the following trace metal mixture (200 mg/L $CaSO_4.2H_2O$; 200 mg/L $FeSO_4.7H_2O$; 20 mg/L $MnSO_4.H_2O$; 20 mg/L $CuSO_4.5H_2O$; 20 mg/L $ZnSO_4.7H_2O$; 10 mg/L $CoSO_4.7H_2O$; 10 mg/L $NaMoO_4.H_2O$; 5 mg/L H_3BO_3)] supplemented with 0.3 mM glucose and 1 μg/mL thiamine dichloride. Bubble traps were placed in each channel before the flow cell to remove air bubbles. Before use, the flow system was sterilized by flushing with a solution of 0.5% sodium hypochlorite for 4 h, and rinsed with approximately 0.2 L of sterile water before the medium was pumped through.

Bacterial cultures for inoculation were prepared by diluting an overnight LB broth culture 1/100 in fresh LB medium and regrowing it for 4 h at 30 °C under shaking conditions. To inoculate the flow cells the medium flow was stopped, flow chambers were turned with the glass coverslip down and 250 μL of the diluted cell suspension was carefully injected through the silicon tubes into each flow channel with a small syringe. After 1 h, to allow adsorption of the cells to the coverslip surface, the flow channels were turned upright and the flow was resumed. For experiments with H-19B::Tn10 and T7, *E. coli* MG1655 *lac::Tc* and *E. coli* MG1655 were used, respectively. When macroscopically visible mature *E. coli* biofilms were formed after 3 days, they became exposed for 1 h to either sterile water, 50 PFU/mL of H-19B::Tn10 or *ca.* 10^8 PFU/mL of T7 via a separate input channel.

3.4. Analysis of Biofilms

Biofilms were analyzed by determining the number of viable cells (CFU) or phage particles (PFU) in either the effluent or in the biofilm itself. Effluent was collected using a small connector inserted in the tubing behind the flowcell, while the actual biofilm was harvested by carefully disconnecting the flowcells and vigorously pipetting up and down the channels with 250 μL potassium phosphate buffer (10 mM, pH 7.0).

After vortexing, dilution series of collected samples were plated on LB agar (in case *E. coli* MG1655 was used) or LB agar supplemented with tetracycline (in case *E. coli* MG1655 *lac::Tc* was used) to enumerate total cell counts, and on LB agar supplemented with chloramphenicol to enumerate H-19B::Tn10 lysogens. Phages were enumerated by first sterilizing the collected samples with $CHCl_3$, and subsequently plating them on *E. coli* MG1655.

72

4. Conclusions

Biofilms, already recalcitrant to therapy, can become an even larger burden through the acquisition of novel virulence determinants such as those conveyed by temperate phages. In Stx phages, the toxin genes are associated with the late gene region, and become highly expressed when the (pro)phage enters the lytic cycle which can be induced at initiation of the bacterial SOS response [19,20,46]. Of intestine isolated *E. coli* strains, 10%–30% were shown to be able to contribute to Shiga toxin production upon infection through transfer of the phage [19,47,48], while also actual *in vivo* transfer was demonstrated [35,49,50]. Certain antibiotics, through induction of the SOS response [51], even lead to the release of temperate phages [52] and thus contribute to the spread of phages (and genes) to the normal flora and elevated production of toxins.

In this study, we report the integration of a chloramphenicol resistance marker in the temperate H-19B Stx-phage without compromising either the lysogenic or the lytic cycle. This phage was shown to rapidly and massively establish itself as a prophage within a mature *E. coli* biofilm, despite the very low initial dose. During and after this establishment, the biofilm started to produce a high number of free H-19B virions, thus contributing to the further dissemination of the phage and its virulence determinants. Moreover, since the production and release of H-19B phage typically coincides with the production and release of its Shiga-toxin [19,20,53], such biofilms likely become a source of continuous toxin production. These findings underscore the ability of temperate phages to rapidly and stably establish themselves within a susceptible but matrix-embedded population, and demonstrate that biofilms can serve as a source of new phage particles and their possible toxins.

Upon challenging a biofilm with lytic phages, an initial reduction in cell numbers is often observed [54,55], but complete eradication by a single phage is never achieved and the establishment of equilibrium between virus and host with stable numbers of both organisms was reported earlier [56]. In addition, resistance is thought to quickly arise within the biofilm, but is only scarcely supported by quantitative data. Here, we report the application of a single pulse of lytic T7 phage to an existing *E. coli* biofilm and demonstrate a reduction of the biofilm even several days after phage application. Nevertheless, this lysis was far from complete and, although highly variable, up to 30% of phage resistant cells could be isolated. This development of resistance was earlier described as a mutual, escalating arms race in which phage arise that are capable of infecting resistant strains, from which in turn bacteria evolve that are resistant to those phages [57,58]. These results confirm the therapeutic potential of lytic phages mainly lies in the use of either well-characterized phage cocktails or by combining them with antibiotics. Synergistic phage-antibiotic combinations on *E. coli* biofilms were observed earlier [59,60] and show great promise for adequate treatment.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Flemming, H.C.; Wingender, J. The biofilm matrix. *Nat. Rev. Microbiol.* **2010**, *8*, 623–633.
2. Hall-Stoodley, L.; Costerton, J.W.; Stoodley, P. Bacterial biofilms: From the natural environment to infectious diseases. *Nat. Rev. Microbiol.* **2004**, *2*, 95–108.
3. Moons, P.; Michiels, C.W.; Aertsen, A. Bacterial interactions in biofilms. *Crit. Rev. Microbiol.* **2009**, *35*, 157–168.
4. Høiby, N.; Ciofu, O.; Johansen, H.K.; Song, Z.J.; Moser, C.; Jensen, P.; Molin, S.; Givskov, M.; Tolker-Nielsen, T.; Bjarnsholt, T. The clinical impact of bacterial biofilms. *Int. J. Oral Sci.* **2011**, *3*, 55–65.
5. Høiby, N.; Bjarnsholt, T.; Givskov, M.; Molin, S.; Ciofu, O. Antibiotic resistance of bacterial biofilms. *Int. J. Antimicrob. Agents* **2010**, *35*, 322–332.
6. Costerton, J.W.; Stewart, P.S.; Greenberg, E.P. Bacterial biofilms: a common cause of persistent infections. *Science* **1999**, *284*, 1318–1322.
7. Francolini, I.; Donelli, G. Prevention and control of biofilm-based medical-device-related infections. *FEMS Immunol. Med. Microbiol.* **2010**, *59*, 227–238.
8. Hall-Stoodley, L.; Stoodley, P.; Kathju, S.; Høiby, N.; Moser, C.; Costerton, J.W.; Moter, A.; Bjarnsholt, T. Towards diagnostic guidelines for biofilm-associated infections. *FEMS Immunol. Med. Microbiol.* **2012**, *65*, 127–145.
9. Beloin, C.; Roux, A.; Ghigo, J.M. *Escherichia coli* biofilms. *Curr. Top. Microbiol. Immunol.* **2008**, *322*, 249–289.
10. Newell, D.G.; Koopmans, M.; Verhoef, L.; Duizer, E.; Aidara-Kane, A.; Sprong, H.; Opsteegh, M.; Langelaar, M.; Threfall, J.; Scheutz, F.; *et al.* Food-borne diseases—The challenges of 20 years ago still persist while new ones continue to emerge. *Int. J. Food Microbiol.* **2010**, *139*, S3–S15.
11. Cantón, R.; Akóva, M.; Carmeli, Y.; Giske, C.G.; Glupczynski, Y.; Gniadkowski, M.; Livermore, D.M.; Miriagou, V.; Naas, T.; Rossolini, G.M.; *et al.* Rapid evolution and spread of carbapenemases among Enterobacteriaceae in Europe. *Clin. Microbiol. Infect.* **2012**, *18*, 413–431.
12. Jacobsen, S.M.; Stickler, D.J.; Mobley, H.L.; Shirtliff, M.E. Complicated catheter-associated urinary tract infections due to *Escherichia coli* and *Proteus mirabilis*. *Clin. Microbiol. Rev.* **2008**, *21*, 26–59.
13. Kaper, J.B.; Nataro, J.P.; Mobley, H.L. Pathogenic *Escherichia coli*. *Nat. Rev. Microbiol.* **2004**, *2*, 123–140.
14. Chaudhuri, R.R.; Henderson, I.R. The evolution of the *Escherichia coli* phylogeny. *Infect. Genet. Evol.* **2012**, *12*, 214–226.

15. Johnson, T.J.; Nolan, L.K. Pathogenomics of the virulence plasmids of *Escherichia coli*. *Microbiol. Mol. Biol. Rev.* **2009**, *73*, 750–774.

16. Farfan, M.J.; Torres, A.G. Molecular mechanisms that mediate colonization of Shiga toxin-producing *Escherichia coli* strains. *Infect. Immun.* **2012**, *80*, 903–913.

17. Werber, D.; Krause, G.; Frank, C.; Fruth, A.; Flieger, A.; Mielke, M.; Schaade, L.; Stark, K. Outbreaks of virulent diarrheagenic *Escherichia coli*—are we in control? *BMC Med.* **2012**, *10*, 11.

18. Melton-Celsa, A.; Mohawk, K.; Teel, L.; O'Brien, A. Pathogenesis of Shiga-toxin producing *Escherichia coli*. *Curr. Top. Microbiol. Immunol.* **2012**, *357*, 67–103.

19. Schmidt, H. Shiga-toxin-converting bacteriophages. *Res. Microbiol.* **2001**, *152*, 687–695.

20. Herold, S.; Karch, H.; Schmidt, H. Shiga toxin-encoding bacteriophages—genomes in motion. *Int. J. Med. Microbiol.* **2004**, *294*, 115–121.

21. Minot, S.; Sinha, R.; Chen, J.; Li, H.; Keilbaugh, S.A.; Wu, G.D.; Lewis, J.D.; Bushman, F.D. The human gut virome: Inter-individual variation and dynamic response to diet. *Genome Res.* **2011**, *21*, 1616–1625.

22. Donlan, R.M. Preventing biofilms of clinically relevant organisms using bacteriophage. *Trends Microbiol.* **2009**, *17*, 66–72.

23. Loc-Carrillo, C.; Abedon, S.T. Pros and cons of phage therapy. *Bacteriophage* **2011**, *1*, 111–114.

24. Ryan, E.M.; Gorman, S.P.; Donnelly, R.F.; Gilmore, B.F. Recent advances in bacteriophage therapy: how delivery routes, formulation, concentration and timing influence the success of phage therapy. *J. Pharm. Pharmacol.* **2011**, *63*, 1253–1264.

25. Azeredo, J.; Sutherland, I.W. The use of phages for the removal of infectious biofilms. *Curr. Pharm. Biotechnol.* **2008**, *9*, 261–266.

26. Rodríguez-Rubio, L.; Martínez, B.; Donovan, D.M.; Rodríguez, A.; García, P. Bacteriophage virion-associated peptidoglycan hydrolases: potential new enzybiotics. *Crit. Rev. Microbiol.* **2012**, doi:10.3109/1040841X.2012.723675.

27. Hughes, K.A.; Sutherland, I.W.; Jones, M.V. Biofilm susceptibility to bacteriophage attack: The role of phage-borne polysaccharide depolymerase. *Microbiology* **1998**, *144*, 3039–3047.

28. Domenech, M.; García, E.; Moscoso, M. *In vitro* destruction of *Streptococcus pneumoniae* biofilms with bacterial and phage peptidoglycan hydrolases. *Antimicrob. Agents Chemother.* **2011**, *55*, 4144–4148.

29. Rice, S.A.; Tan, C.H.; Mikkelsen, P.J.; Kung, V.; Woo, J.; Tay, M.; Hauser, A.; McDougald, D.; Webb, J.S.; Kjelleberg, S. The biofilm life cycle and virulence of *Pseudomonas aeruginosa* are dependent on a filamentous prophage. *ISME J.* **2009**, *3*, 271–282.

30. Meng, X.; Shi, Y.; Ji, W.; Zhang, J.; Wang, H.; Lu, C.; Sun, J.; Yan, Y. Application of a bacteriophage lysin to disrupt biofilms formed by the animal pathogen *Streptococcus suis*. *Appl. Environ. Microbiol.* **2011**, *77*, 8272–8279.

31. Siringan, P.; Connerton, P.L.; Payne, R.J.; Connerton, I.F. Bacteriophage-Mediated dispersal of *Campylobacter jejuni* biofilms. *Appl. Environ. Microbiol.* **2011**, *77*, 3320–3632.

32. Neely, M.N.; Friedman, D.I. Functional and genetic analysis of regulatory regions of coliphage H-19B: Location of shiga-like toxin and lysis genes suggest a role for phage functions in toxin release. *Mol. Microbiol.* **1998**, *28*, 1255–1267.

33. Kleckner, N.; Bender, J.; Gottesman, S. Uses of transposons with emphasis on Tn10. *Meth. Enzymol.* **1991**, *204*, 139–180.

34. Serra-Moreno, R.; Acosta, S.; Hernalsteens, J.P.; Jofre, J.; Muniesa, M. Use of the lambda Red recombinase system to produce recombinant prophages carrying antibiotic resistance genes. *BMC Mol. Biol.* **2006**, *7*, 31.

35. Acheson, D.W.; Reidl, J.; Zhang, X.; Keusch, G.T.; Mekalanos, J.J.; Waldor, M.K. *In vivo* transduction with shiga toxin 1-encoding phage. *Infect. Immun.* **1998**, *66*, 4496–4498.

36. Wu, B.; Skarina, T.; Yee, A.; Jobin, M.C.; Dileo, R.; Semesi, A.; Fares, C.; Lemak, A.; Coombes, B.K.; Arrowsmith, C.H.; *et al*. NleG Type 3 effectors from enterohaemorrhagic *Escherichia coli* are U-Box E3 ubiquitin ligases. *PLoS Pathog.* **2010**, *6*, e1000960.

37. Ogura, Y.; Ooka, T.; Iguchi, A.; Toh, H.; Asadulghani, M.; Oshima, K.; Kodama, T.; Abe, H.; Nakayama, K.; Kurokawa, K.; *et al*. Comparative genomics reveal the mechanism of the parallel evolution of O157 and non-O157 enterohemorrhagic *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 17939–17944.

38. Smith, D.L.; Rooks, D.J.; Fogg, P.C.; Darby, A.C.; Thomson, N.R.; McCarthy, A.J.; Allison, H.E. Comparative genomics of Shiga toxin encoding bacteriophages. *BMC Genom.* **2012**, *13*, 311.

39. Radke, K.L.; Siegel, E.C. Mutation preventing capsular polysaccharide synthesis in *Escherichia coli* K-12 and its effect on bacteriophage resistance. *J. Bacteriol.* **1971**, *106*, 432–437.

40. Scholl, D.; Adhya, S.; Merril, C. Escherichia coli K1's capsule is a barrier to bacteriophage T7. *Appl. Environ. Microbiol.* **2005**, *71*, 4872–4874.

41. Guyer, M.S.; Reed, R.R.; Steitz, J.A.; Low, K.B. Identification of a sex-factor-affinity site in *E. coli* as gamma delta. *Cold Spring Harb. Symp. Quant. Biol.* **1981**, *45*, 135–140.

42. Moons, P.; van Houdt, R.; Aertsen, A.; Vanoirbeek, K.; Engelborghs, Y.; Michiels, C.W. Role of quorum sensing and antimicrobial component production by *Serratia. plymuthica* in formation of biofilms, including mixed biofilms with *Escherichia coli*. *Appl. Environ. Microbiol.* **2006**, *72*, 7294–7300.

43. Aertsen, A.; Faster, D.; Michiels, C.W. Induction of Shiga toxin-converting prophage in *Escherichia coli* by high hydrostatic pressure. *Appl. Environ. Microbiol.* **2005**, *71*, 1155–1162.

44. Dunn, J.J.; Studier, F.W. Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. *J. Mol. Biol.* **1983**, *166*, 477–535.

45. Christensen, B.B.; Sternberg, C.; Andersen, J.B.; Palmer, R.J.; Nielsen, A.T.; Givskov, M.; Molin, S. Molecular tools for study of biofilm physiology. *Meth. Enzymol.* **1999**, *310*, 20–42.

46. Fogg, P.C.; Saunders, J.R.; McCarthy, A.J.; Allison, H.E. Cumulative effect of prophage burden on Shiga toxin production in *Escherichia coli*. *Microbiology* **2012**, *158*, 488–497.

76

47. James, C.E.; Stanley, K.N.; Allison, H.E.; Flint, H.J.; Stewart, C.S.; Sharp, R.J.; Saunders, J.R.; McCarthy, A.J. Lytic and lysogenic infection of diverse *Escherichia coli* and *Shigella.* strains with a verocytotoxigenic bacteriophage. *Appl. Environ. Microbiol.* **2001**, *67*, 4335–4337.

48. Gamage, S.D.; Strasser, J.E.; Chalk, C.L.; Weiss, A.A. Nonpathogenic *Escherichia coli* can contribute to the production of Shiga toxin. *Infect. Immun.* **2003**, *71*, 3107–3115.

49. Tóth, I.; Schmidt, H.; Dow, M.; Malik, A.; Oswald, E.; Nagy, B. Transduction of porcine enteropathogenic *Escherichia coli* with a derivative of a shiga toxin 2-encoding bacteriophage in a porcine ligated ileal loop system. *Appl. Environ. Microbiol.* **2003**, *69*, 7242–7247.

50. Cornick, N.A.; Helgerson, A.F.; Mai, V.; Ritchie, J.M.; Acheson, D.W. *In vivo* transduction of an Stx-encoding phage in ruminants. *Appl. Environ. Microbiol.* **2006**, *72*, 5086–5088.

51. Hastings, P.J.; Rosenberg, S.M.; Slack, A. Antibiotic-induced lateral transfer of antibiotic resistance. *Trends Microbiol.* **2004**, *12*, 401–404.

52. Zhang, X.; McDaniel, A.D.; Wolf, L.E.; Keusch, G.T.; Waldor, M.K.; Acheson, D.W. Quinolone antibiotics induce Shiga toxin-encoding bacteriophages, toxin production, and death in mice. *J. Infect. Dis.* **2000**, *181*, 664–670.

53. Aertsen, A.; van Houdt, R.; Michiels, C.W. Construction and use of an *stx1* transcriptional fusion to *gfp*. *FEMS Microbiol. Lett.* **2005**, *245*, 73–77.

54. Carson, L.; Gorman, S.P.; Gilmore, B.F. The use of lytic bacteriophages in the prevention and eradication of biofilms of *Proteus mirabilis* and *Escherichia coli*. *FEMS Immunol. Med. Microbiol.* **2010**, *59*, 447–455.

55. Chibeu, A.; Lingohr, E.J.; Masson, L.; Manges, A.; Harel, J.; Ackermann, H.W.; Kropinski, A.M.; Boerlin, P. Bacteriophages with the ability to degrade uropathogenic *Escherichia coli* biofilms. *Viruses* **2012**, *4*, 471–487.

56. Corbin, B.D.; McLean, R.J.; Aron, G.M. Bacteriophage T4 multiplication in a glucose-limited Escherichia coli biofilm. *Can. J. Microbiol.* **2001**, *47*, 680–684.

57. Buckling, A.; Rainey, P.B. Antagonistic coevolution between a bacterium and a bacteriophage. *Proc. Biol. Sci.* **2002**, *269*, 931–936.

58. Kashiwagi, A.; Yomo, T. Ongoing phenotypic and genomic changes in experimental coevolution of RNA bacteriophage Qβ and *Escherichia coli*. *PLoS Genet.* **2011**, *7*, e1002188.

59. Lu, T.K.; Collins, J.J. Engineered bacteriophage targeting gene networks as adjuvants for antibiotic therapy. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 4629–4634.

60. Ryan, E.M.; Alkawareek, M.Y.; Donnelly, R.F.; Gilmore, B.F. Synergistic phage-antibiotic combinations for the control of *Escherichia coli* biofilms *in vitro*. *FEMS Immunol. Med. Microbiol.* **2012**, *65*, 395–398.

Review

Photodynamic Inactivation of Mammalian Viruses and Bacteriophages

Liliana Costa [1], Maria Amparo F. Faustino [2], Maria Graça P. M. S. Neves [2], Ângela Cunha [1] and Adelaide Almeida [1,*]

[1] Department of Biology and CESAM, University of Aveiro, Aveiro 3810-193, Portugal;
E-Mails: lcosta@ua.pt (L.C.); acunha@ua.pt (A.C.)

[2] Department of Chemistry and QOPNA, University of Aveiro, Aveiro 3810-193, Portugal;
E-Mails: faustino@ua.pt (M.A.F.F.); gneves@ua.pt (M.G.P.M.S.N.)

* Author to whom correspondence should be addressed; E-Mail: aalmeida@ua.pt;
Tel.: +351-234-370-350; Fax: +351-234-372-587.

Abstract: Photodynamic inactivation (PDI) has been used to inactivate microorganisms through the use of photosensitizers. The inactivation of mammalian viruses and bacteriophages by photosensitization has been applied with success since the first decades of the last century. Due to the fact that mammalian viruses are known to pose a threat to public health and that bacteriophages are frequently used as models of mammalian viruses, it is important to know and understand the mechanisms and photodynamic procedures involved in their photoinactivation. The aim of this review is to (i) summarize the main approaches developed until now for the photodynamic inactivation of bacteriophages and mammalian viruses and, (ii) discuss and compare the present state of the art of mammalian viruses PDI with phage photoinactivation, with special focus on the most relevant mechanisms, molecular targets and factors affecting the viral inactivation process.

78

Nomenclature

AlPcS$_4$	Aluminum phthalocyanine tetrasulfonate
AZT	Azidothymidine
BVDV	Bovine viral diarrhea virus
DMTU	Dimethylthiourea
EMCV	Encephalomyocarditis virus
HAV	Hepatitis A virus
HBV	Hepatitis B virus
HCV	Hepatitis C virus
HIV	Human immunodeficiency virus
HPV	Human papillomatosis virus
HSV	Herpes simplex virus
LED	Light emitting diode
MB	Methylene blue
NM	Not mentioned
NQ	Not quantified
Pc$_4$	Silicon phthalocyanine
PDI	Photodynamic inactivation
PS	Photosensitizer
ROS	Reactive oxygen species
SFV	Semliki Forest virus
SHV	Suid herpes virus
SOD	Superoxide dismutase
SSB	Singlet strand breaks
Tri-Py$^+$-Me-PF	5-(pentafluorophenyl)-10,15,20-tris(1-methylpyridinium-4-yl)porphyrin tri-iodide
VSV	Vesicular stomatitis virus
VZV	Varicella zoster virus
$^1O_2$	Singlet oxygen
$^3O_2$	Molecular oxygen
^1PS	Ground state photosensitizer
$^3PS*$	Triplet excited state photosensitizer

1. Introduction

Humans are exposed to pathogenic viruses through various routes and the development of viral-induced diseases is a common occurrence.

Although the transmission of viral diseases has been reduced by the development of good water supplies and hygienic-based procedures for a whole range of human activities [1], pathogenic viruses are still the causative agents of many diseases in humans and other species. The most usual human

diseases caused by viruses include the common cold (coronaviruses), influenza (influenza viruses), chickenpox (varicella zoster virus), cold sores (herpes simplex virus), gastroenteritis and diarrhoea (caliciviruses, rotaviruses and adenoviruses) [2,3]. Pathogenic viruses are also implicated in serious diseases, such as Ebola (Ebola virus), AIDS (immunodeficiency viruses), avian influenza and sudden acute respiratory syndrome (SARS) (SARS-coronavirus), and they are also an established cause of cancer (papillomavirus, hepatitis B and C viruses, Epstein–Barr virus, Kaposi's sarcoma-associated herpes virus, human T-lymphotropic virus, and Merkel cell polyomavirus) [4].

The enhanced implication of viruses in severe infectious diseases and the increasing knowledge about the complex mechanisms of viral pathogenesis have greatly contributed to the rapid development of antiviral drugs. Consequently, the use of antivirals has largely increased in the last years and resistance to antiviral drugs is now well documented for several pathogenic viruses [5–10]. Moreover, as viruses are genetically flexible, they may mutate quickly and mutations come as no surprises, leading to the development of resistance to conventional antiviral drugs. Consequently, the emergence of antiviral drug can become a great problem, such the resistance observed for bacteria relative to antibiotics. So, alternative methods unlikely to cause resistance are required. Photodynamic inactivation (PDI) of viruses represents a promising and inexpensive potential alternative to meet that need.

The sensitivity of viruses to photodynamic procedures was reported in the 1930s [11,12] but only within the last 30 years, with the development of new active molecules, namely photosensitizers (PS), and an increment of light technologies (lasers, LED, portability, *etc.*), have photodynamic techniques for the inactivation of viruses received growing attention [13]. Most of the clinical applications of PDI for treatment of infections have so far been directed to viral lesions [14]. Clinical PDI was first applied to the treatment of herpes infection in the early 1970s [15], particularly for herpes genitalis. Since then, a great variety of viruses has been effectively inactivated by photodynamic treatment using *in vitro* conditions [16] but, considering the clinical use of viral PDI, the procedures are limited to the treatment of papillomatosis, caused by human papillomatosis virus (HPV), like laryngeal papillomatosis [17] and epidermodysplasia verruciformis [18] and, in a small scale, to the treatment of viral complications in AIDS patients [19,20]. However, considerable progress has been made in the viral photodynamic disinfection of blood products. The major threat of viral contamination in blood and blood products comes from the immunodeficiency viruses (HIV) [21], hepatitis viruses [21–23], cytomegalovirus [23], human parvovirus B19 [24] and human T-cell lymphotropic virus type I and type II [23]. HIV has been inactivated *in vitro* following a photodynamic procedure [25–39]. The photoinactivation of hepatitis viruses in blood products has also been successfully tested against the hepatitis C virus (HCV) [37,40–42], hepatitis B virus (HBV) [43] and hepatitis A virus (HAV) [44]. Inactivation of cytomegalovirus [45], human parvovirus B19 [46] and human T-cell lymphotropic virus [47] in blood products was also efficiently achieved after photodynamic treatment.

The availability of a simple and quantitative assay to follow the viral photoinactivation process is important. Traditional viral quantification techniques, such as *in vitro* viral cultures, are time-consuming and labor-intensive processes. Molecular quantitative methods such as nucleic acid amplification procedures, including real time PCR, are rapid and sensitive but detect only viral nucleic acid and do

not determine infectivity. When the virucidal properties of different photosensitizing compounds are initially evaluated, bacteriophages can be useful as surrogates of mammalian viruses. The reasons for their use are: (i) the detection methods are much simpler, faster and cheaper than those of mammalian viruses, avoiding the advanced facilities and equipment needed for propagating human pathogens; (ii) they are non-pathogenic to humans; (iii) they can be grown to higher titers than most mammalian viruses and, therefore, enhancing the sensitivity of the assay; (iv) the results of bacteriophages assays are available within several hours post-inoculation, instead of the days or weeks required by mammalian viruses infectivity-based assays; (v) they are at least as resistant as the mammalian viruses to environmental factors and to water treatment [48].

It has been shown that enveloped viruses are significantly more sensitive to photodynamic destruction than non-enveloped viruses [49,50]. As most of the bacteriophages are non-enveloped, they are more difficult to suffer photoinactivation than the enveloped viruses. In general, this property makes them good indicators to evaluate the efficiency of viral PDI. A PDI protocol that is effective to inactivate a non-enveloped phage will most likely be effective against enveloped mammalian viruses.

Several bacteriophages were used in photoinactivation studies as surrogates for mammalian viruses, e.g., MS2 [44], M13 [51,52], PM2 [53], Qβ [54–56], PRD1 [57], λ [58,59], ϕ6 [60], R17 [60], *Serratia* phage *kappa* [61], T5 [62], T3 [63], T7 [57,64] and T4-like [65–68], and the results show that they are effectively photoinactivated.

2. Antimicrobial PDI

PDI is a simple and controllable method for the inactivation of microorganisms based on the production of reactive oxygen species (ROS) (free radicals and singlet oxygen). This technology requires the combined action of oxygen, light and a photosensitizer (PS), which absorbs and uses the energy from light to produce those ROS [69]. Therefore, the photodynamic effects depend on multiple variables including: the structural features of the PS, the concentrations of PS and molecular oxygen, and the properties of the light used (e.g., wavelength, type, dose and fluence rate) [66,67,69–72]. Changes in any of these parameters will affect the rate of microbial photoinactivation [66,67,73,74].

The majority of PS used in PDI is derived from tetrapyrrolic macrocycles known as porphyrins. These chromophores and their analogs, such as chlorins and bacteriochlorins, are involved in very important biological functions, such as respiration (heme group) and photosynthesis (chlorophyll and bacteriochlorophyll (Figure 1). Based on these macrocycles, the scientific community was able to develop a number of synthetic analogs, such as *meso*-tetraarylporphyrins, phthalocyanines, texaphyrins, porphycenes and saphyrins, which proved to have very promising features for being used as PS (Figure 2) [16]. Also, non-tetrapyrrolic derivatives, such as the naturally occurring hypericin, or synthetic dyes like toluidine blue O, rose bengal, eosin, methylene blue (MB) and fullerenes, were considered in many PDI studies (Figure 3) [71].

In order to be efficient, photosensitizing agents used for viral PDI must bind specifically to vital viral components, such as lipid envelope (when present), the protein coat or to the nucleic acids [55].

Figure 1. Structure of some tetrapyrrolic macrocycles with natural occurrence.

chlorophyll a

Heme group

bacteriochlorophyll a
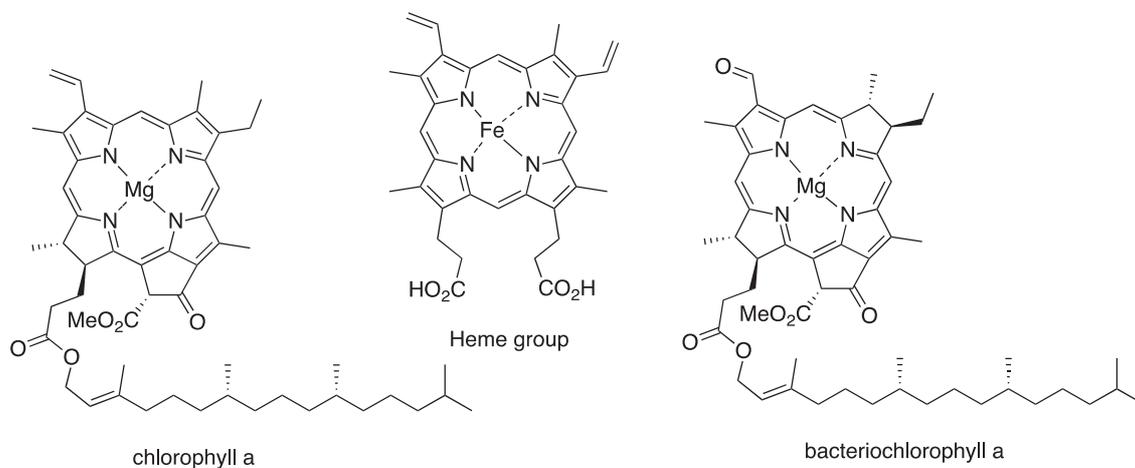
Figure 2. Skeletons of some synthetic pyrrolic macrocycles used as photosensitizers.

meso-tetraarylporphyrin

phthalocyanine

porphycene

texaphyrin
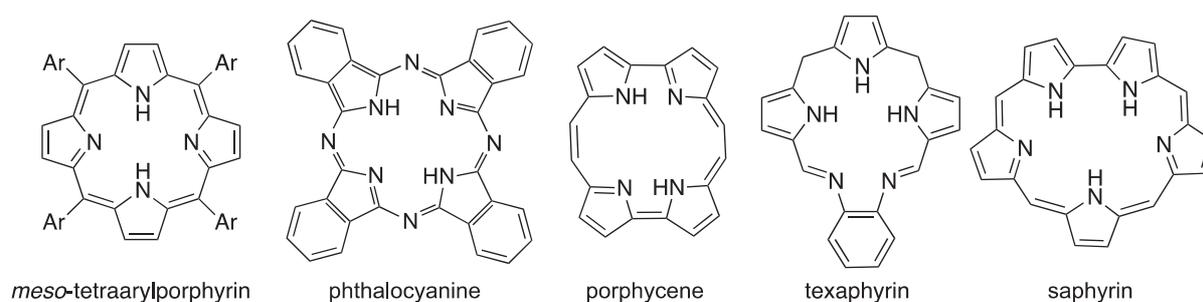
saphyrin

Figure 3. Structure of some non-tetrapyrrolic photosensitizers.

hypericin

methylene blue

rose bengal

psoralen

merocyanine 540

fullerene
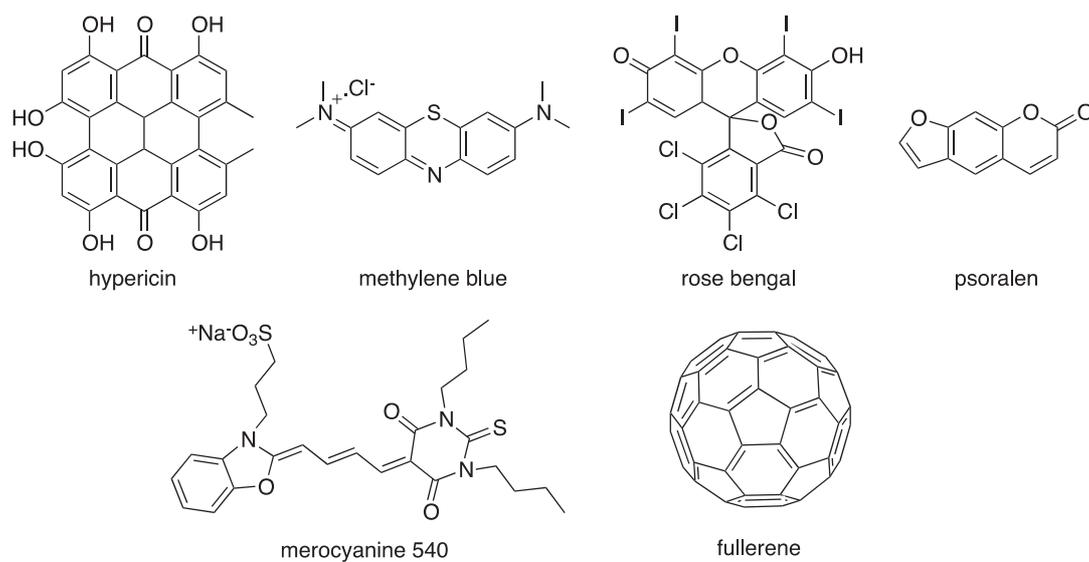
The efficiency of mammalian viruses and bacteriophages PDI has been described for porphyrin derivatives, chlorin derivatives, chlorophyll derivatives, phthalocyanine derivatives, hypericin, methylene blue, rose bengal, merocyanine 540, proflavine, and fullerene derivatives (Table 1).

Table 1. Some PS used for mammalian viruses and bacteriophages PDI.

Photosensitizer	Microorganism	PDI	Reference
Mammalian viruses			
Hematoporphyrin derivative	HSV-1	7 log	[75]
	HSV-1	<0.8 log	[36]
Uroporphyrin	Adenovirus	7 log	[76]
Natural metalloporphyrin derivatives	HIV-1	<0.8 log	[36]
Chlorophyll derivatives	VSV	~6 log	[77]
7-despropionate-7-hydroxypropylmesopyropheophorbide a	BVDV	~5 log	[78]
	EMCV	~0.2 log	
Benzoporphyrin derivative monoacid ring A	HIV-1	>4 log	[33]
Glycoconjugated *meso*-tetraarylporphyrin derivatives	HSV-1	6 log	[79]
	HSV-2	6 log	
Metallo tetrasulfonated *meso*-tetraarylporphyrin derivatives	HIV-1	≤2 log	[36]
Tetrasulfonated *meso*-tetraarylporphyrin derivatives	HIV-1	≤2 log	[36]
	HAV	~4 log	[44]
meso-Tetrakis(1-methylpyridinium-4-yl)porphyrin	HAV	~4 log	[44]
meso-Tetrakis(1-butylpyridinium-4-yl)porphyrin	HAV	>3.8 log	[44]
meso-Tetrakis(1-octylpyridinium-4-yl)porphyrin	HAV	>3.9 log	[44]
Cationic β-vinyl substituted *meso*-tetraphenylporphyrin derivatives	HSV-1	<3 log	[80]
Aluminum dibenzodisulfophthalocyanine	HIV-1	3.7 log	[49]
Aluminum phthalocyanine tetrasulfonate	HIV-1	>5 log	[49]
	VSV	4.2 log	[82]
	Adenovirus	4 log	[76]
Silicon phthalocyanine derivative	VSV	4 log	[82]
Cationic phthalocyanines	HIV-1	>5 log	[49]
	HSV-1	≥5 log	[83]
Hypericin	HIV-1	NQ	[30]
	VSV	4–5 log	
	Influenza virus	NQ	
	Sendai virus	NQ	

Table 1. *Cont.*

Photosensitizer	Microorganism	PDI	Reference
Mammalian viruses			
Methylene blue	VSV	4.7 log	[81]
	HSV-1	5 log	[84]
	SHV-1	2.5 log	[84]
	HCV	<2 log	[41]
	HIV-1	<2 log	[41]
	Adenovirus	7 log	[76]
	Dengue virus	5–6.4 log	[74]
	Enterovirus 71	~8 log	[85]
	Vaccinia virus	5 log	[86]
Phenothiazine derivatives	VSV	>4.4 log	[60]
Rose bengal	Vaccinia virus	5 log	[86]
	HIV-1	NQ	[30]
	VSV	4–5 log	
	Influenza virus	NQ	
	Sendai virus	NQ	
	Adenovirus	7 log	[76]
Buckminsterfullerene	SFV	7 log	[50]
	VSV	7 log	
Merocyanine 540	HSV-1	5–6 log	[45]
Bacteriophages			
Glycoconjugated *meso*-tetraarylporphyrins	T7 phage	<3 log	[64]
	T7 phage	<3.5 log	[87]
Tetrasulfonated *meso*-tetraarylporphyrin derivatives	MS2 phage	>3.8 log	[44]
meso-Tetrakis(1-methylpyridinium-4-yl)porphyrin	λ phage	<7 log	[58]
	MS2 phage	>4.1 log	[44]
	T4 phage	7 log	[66,67]
	T7 phage	<4 log	[88]
5-(pentafluorophenyl)-10,15,20-tris(1-methylpyridinium-4-yl)porphyrin	T4 phage	7 log	[66–68]
5-(4-methoxicarbonylphenyl)-10,15,20-tris(1-methylpyridinium-4-yl)porphyrin	T4 phage	7 log	[66]
5-(4-carboxyphenyl)-10,15,20-tris(1-methylpyridinium-4-yl)porphyrin	T4 phage	3.9 log	[66]
5,10-bis(4-carboxyphenyl)-15,20-bis(1-methylpyridinium-4-yl)porphyrin	T4 phage	1.4 log	[66]
5,15-bis(4-carboxyphenyl)-10,20-bis(1-methylpyridinium-4-yl)porphyrin	T4 phage	1.2 log	[66]
5,10,15-tris(1-methylpyridinium-4-yl)-20-phenylporphyrin	T7 phage	1.7 log	[88]

Table 1. *Cont.*

Photosensitizer	Microorganism	PDI	Reference
Methylene blue	*Serratia* phage *kappa*	>4 log	[61]
	M13 phage	2.2 log	[52,81]
	f2 phage	5 log	[56]
	Qβ phage	7–8 log	[56]
	Qβ phage	7–8 log	[89]
Phenothiazine derivatives	R17 phage	4–7 log	[60]
	φ6	4–6.5 log	
Rose bengal	PRD1 phage	~3.5 log*	[57]
	T7 phage	~4.5 log*	
Riboflavin	λ phage	<4 log	[59]
Proflavine	*Serratia* phage *kappa*	4 log	[61]
	T3 phage	7–11 log	[63]
Polyhydroxylated fullerene	MS2 phage	~4 log	[90]
	PRD1 phage	~2.5 log*	[57]
	T7 phage	~3.5 log*	
	MS2 phage	~5 log*	

*log(N/N0).

Besides this, viral PDI has also been described for phthalocyanine derivatives [81], methylene blue [53,62,91,92], toluidine blue O [53,62,93], neutral red [93], proflavine [93], azure B [53] and merocyanine 540 [45,47,94].

3. Mechanisms of Photodynamic Inactivation

The mechanisms of PDI are based on the ability of the PS to absorb energy from light and transfer that energy to molecular oxygen. In the dark, the electronic configuration of a PS exists in the so-called ground state. The absorption, by the PS, of a photon at an appropriate wavelength initially leads to the production of an unstable, electronically-excited state of the PS molecule (the lifetime of this state ranges from 10^{-9} to 10^{-6} s) [95]. The excited PS molecule can then decay to the ground state by emission of light (radiative pathway - fluorescence) or by intersystem crossing, affording the excited triplet state which has a longer lifetime (10^{-3} to 10 s) [95]. At this point, the PS can reach the ground state either by spin inversion followed by phosphorescence emission, or by a non-radiative process. Due to the longer lifetime of the PS triplet state, this excited state can also react in one of two ways (Figure 2): by initiating photochemical reactions that can directly generate reactive oxygen species (ROS) (type I pathway), or indirectly by energy transfer to molecular oxygen (type II pathway), leading to the formation of singlet oxygen (Figure 4). These events afford toxic species which are responsible for the irreparable oxidative damages induced to important biological targets [1,69,95,96].

Figure 4. Schematic representation of the photosensitization process (adapted from [97]).
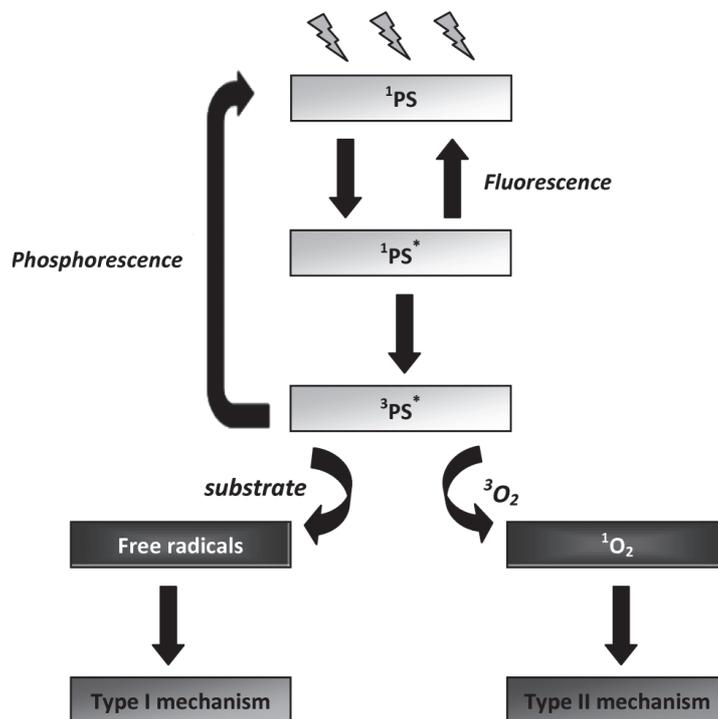
3.1. Type I and Type II Mechanisms

Type I mechanism involves hydrogen-atom abstraction or electron-transfer between the excited PS and a substrate, yielding free radicals [Equations (1) and (2)]. These radicals can react with oxygen to form active oxygen species, such as the superoxide radical anion [Equation (3)]. Superoxide is not particularly reactive in biological systems but, when protonated, can lead to the production of hydrogen peroxide and oxygen [Equations (4) and (5)] or highly reactive hydroxyl radicals [Equations (6)–(8)] [98]. Type II photooxidation is considerably less complex mechanistically than type I and in general there are far fewer products [99]. In this pathway, the excited triplet state PS (^{3}PS*) can transfer the excess energy to molecular oxygen (^{3}O$_2$) and relax to its ground state (^{1}PS) creating an excited singlet molecular oxygen (^{1}O$_2$) [Equation (9)] [69]. ^{1}O$_2$ is highly electrophilic and can interact with numerous enzymes, leading to the inhibition of protein synthesis and molecular alteration of DNA strands, which alters the transcription of the genetic material during its replication (mutagenic effect) and, in this way, leading to microbial death [Equation (10)] [98,100]. Like nucleic acids and proteins, unsaturated lipids are also prominent targets of ^{1}O$_2$ and free radical attack. Lipid peroxidation-ensuing reactions can alter surrounding proteins, nucleic acids and other molecules, in addition to the lipids themselves [98]. Therefore, it is likely that damage of different kinds caused to the viral envelope is important in the process of microbial inactivation [13].

$$\text{SubstrateH}_2 + \text{PS} \rightarrow \text{PSH}^{\bullet} + \text{SubstrateH}^{\bullet} \tag{1}$$

$$\text{PS* + Substrate} \rightarrow \text{PS}^{\bullet -} + \text{Substrate}^{\bullet +} \text{ or PS* + Substrate} \rightarrow \text{PS}^{\bullet +} + \text{Substrate}^{\bullet -} \tag{2}$$

$$PS^{\bullet-} + {}^3O_2 \rightarrow PS + O_2^{\bullet-} \tag{3}$$

$$O_2^{\bullet-} + H^+ \rightleftharpoons HOO^{\bullet} \tag{4}$$

$$2\ HOO^{\bullet} \rightarrow H_2O_2 + O_2 \tag{5}$$

$$H_2O_2 + Fe^{2+} \rightarrow HO^{\bullet} + OH^- + Fe^{3+} \tag{6}$$

$$BiomoleculeH + HO^{\bullet} \rightarrow Biomolecule^{\bullet} + H_2O \tag{7}$$

$$Biomolecule^{\bullet} + {}^3O_2 \rightarrow Biomolecule\text{-}OO^{\bullet} \rightarrow products \tag{8}$$

$$PS^* + {}^3O_2 \rightarrow PS + {}^1O_2 \tag{9}$$

$$Biomolecules + {}^1O_2 \rightarrow oxidative\ products \tag{10}$$
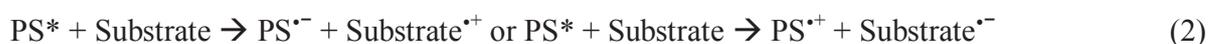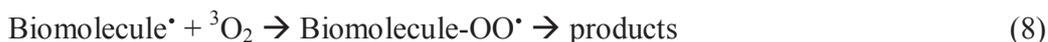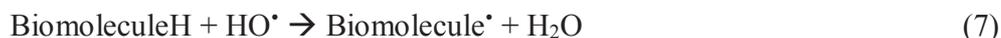
Both type I and type II mechanisms can occur simultaneously or exclusively, and the ratio between these processes depends on the PS used and on the concentrations of substrate and oxygen [95]. The competition between organic substrates and molecular oxygen for the ${}^3PS^*$ determines whether the reaction pathway is type I or type II and the predominant mechanism can be changed during the course of the PDI process [101].

3.2. Evaluation of the Specific Involvement of Type I and Type II Mechanisms

An important goal in the investigation of viral PDI is to identify the type of mechanism involved (type I or type II) in the presence of a selected PS [102]. The simple detection of a reactive species does not necessarily explain the mechanism by which a specific PS induces the toxic effect. It is generally easier to draw a negative conclusion, *i.e.*, if singlet oxygen is absent, it cannot be the reactive species responsible for the photodynamic effect [103]. The simplest approach for determining whether singlet oxygen (type II mechanism) or free radicals (type I mechanism) is involved in the photodynamic process is to study the inhibitory effects of various scavengers, *i.e.*, compounds that can intercept these ROS at high rates and in a putatively selective manner [99,101,104].

3.2.1. Type I Mechanism Scavengers

A first line of defence against ROS is, of course, the protection against their formation. However, the interception of the damaging species once formed, to prevent it from further deleterious reactions, is also a deactivation strategy of defence. In general, free radical scavengers neutralize the radical species by donating one of their own electrons. The quenching agents themselves are not particularly toxic before and after the electron donation [105].

Three different types of quenching are possible, which include the transfer of the radical character with the formation of a reactive scavenger-derived radical; trapping of free radicals with the formation of a stable or inert free radical trap; and molecules which mimic quenching enzyme activities. In general, scavenger molecules either prevent free radicals from being formed or remove them before they can damage vital molecular components [105].

Several free radical scavengers have been used to evaluate the specific involvement of type I mechanism during mammalian viruses and bacteriophages PDI with different PS (Table 2).

Table 2. Free radical scavengers used in mammalian viruses and bacteriophages PDI.

PS	Scavenger	Microorganism	Scavenger protection	Reference
Mammalian viruses				
Aluminum phthalocyanine tetrasulfonate	Reduced glutathione Mannitol Glycerol SOD	VSV	Little/no effect Little/no effect Little/no effect Little/no effect	[106]
Polyhydroxylated fullerene	Glutathione (2.0 mM)	SFV VSV	no effect no effect	[50]
	Hydroquinone (2.0 mM)	SFV VSV	no effect no effect	[50]
Merocyanine 540	Glutathione (10 and 30 mmol L^{-1}) Cysteamine (10 and 30 mmol L^{-1}) SOD (1.5 to 29 U mL^{-1})	HSV-1	30-50% 60-70% no effect	[45]
Methylene blue	Mannitol (100 mM) Glycerol (10 mM) SOD (300 U mL^{-1}) Catalase (30 U mL^{-1})	HSV-1	24% 24% 24% 24%	[84]
Bacteriophages				
5,10,15-(4-β-D-glucosylphenyl)-20-phenylporphyrin	DMTU (0.1–5.0 mM)	T7 phage	44%	[64]
5,10.15,20-Tetrakis(4-β-D-glucosylphenyl) porphyrin	DMTU (0.1–5.0 mM)	T7 phage	79%	[64]
5,10,15-(4-β-D-galactosylphenyl)-20-(pentafluorophenyl)-porphyrin	DMTU (0.1–5.0 mM)	T7 phage	89%	[87]
5-(pentafluorophenyl)-10,15,20-tris(1-methylpyridinium-4-yl)porphyrin	D-mannitol (100 mM)	T4 phage Qβ	20% no effect	[107]
	L-cysteine (100 mM)	T4 phage	9%	[107]
5,10,15,20-tetrakis(1-methylpyridinium-4-yl)porphyrin	D-mannitol (100 mM)	T4 phage	no effect	[107]
Proflavine	L-cysteine (0.025 M)	T3 phage	75–80%	[63]
Polyhydroxylated fullerene	SOD	MS2 phage	no effect	[90]

3.2.1.1. Free Radicals in PDI of Mammalian Viruses

Free radical species had, in general, little or no effect on the photoinactivation of the studied mammalian viruses (Table 2). In fact, it can be observed that the rate of inactivation of

HSV [45,84,106], influenza virus [108], Semliki Forest virus (SFV) and VSV [50] in the presence of different PS and scavengers like glutathione, D-mannitol, glycerol, superoxide dismutase (SOD), catalase and hydroquinone was not significantly affected. Although this data suggest that free radicals are not major players in the viral inactivation process, the participation of type I reaction pathways cannot be ruled out, as was shown by the considerable level of protection afforded by glutathione and cysteamine when merocyanine 540 was used as PS for inactivation of HSV-1 [45].

3.2.1.2. Free Radicals in PDI of Bacteriophages

The photoinactivation rate of some bacteriophages can be reduced in the presence of free radical scavengers, suggesting a contribution of radical species in the inactivation process (Table 2). In particular, it was reported that the inhibition of T7 phage photoinactivation in the presence of glycoconjugated *meso*-tetraarylporphyrins varied according to the structure of the PS and the concentration of dimethylthiourea (DMTU) [64,87]. In fact, T7 phage PDI by *meso*-tetrakis(4-β-D-glucosylphenyl)porphyrin [64] and 5,10,15-(4-β-D-galactosylphenyl)-20-(pentafluorophenyl)porphyrin [87] seemed to be mainly mediated by free radical species, as revealed by the protection effect of free radical scavenger DMTU, contrary to T7 phage photosensitization by 5,10,15-(4-β-D-glucosylphenyl)-20-phenylporphyrin, which revealed a significantly smaller contribution from type I mechanism. The highest inhibition was reached at about 1.0 mM of DMTU; further increase in scavenger concentration did not decrease the slope of photoinduced inactivation of phages. However, in spite of inhibiting the efficacy of the PS, DMTU did not completely inhibit T7 phage PDI [64,87]. Similar results were reported for T3 phage in the presence of L-cysteine as the scavenger and proflavine as the PS. However, the photoinactivation rate of MS2 by a polydroxylated fullerene was not affected by the presence of SOD, suggesting a negligible contribution of radical species, such as the superoxide radical anion [90]. T4-like phage PDI was also little or not affected by the presence of free radical scavengers L-cysteine and D-mannitol in the presence of porphyrin derivatives, leading to the conclusion that free radical species are not major participants in phage PDI [107].

3.2.2. Type II Mechanism Quenchers

In general, the action of chemical singlet oxygen quenchers involves the reaction of singlet oxygen with the quenching agent, producing an oxidized product. Another possibility is the deactivation of singlet oxygen to ground state ($^3O_2$) by physical quenching, achieved by either energy or charge transfer, without consumption of oxygen or product formation [101,109]. Residues of histidine, tryptophan and tyrosine in proteins are considered to be major natural quenchers of singlet oxygen [110].

Several singlet oxygen quenchers have been used to evaluate the specific involvement of type II mechanism during viral PDI with different PS (Table 3).

Table 3. Singlet oxygen quenchers used on mammalian viruses and bacteriophage PDI.

PS	Quencher	Microorganism	Quencher protection	Reference
Mammalian viruses				
Aluminum phthalocyanine tetrasulfonate	Sodium azide	VSV	significant effect	[106]
	Tryptophan	VSV	Significant effect	
Rose bengal	β-carotene	Influenza virus	Significant effect	[108]
	Sodium azide			
Hypericin	Sodium azide	HIV	Significant effect	[111]
Methylene blue	Imidazole (5.0 and 10 mM)	HSV-1	55%–75%	[84]
Bacteriophages				
5,10,15-(4-β-D-galactosylphenyl)-20-(pentafluorophenyl)porphyrin	Sodium azide (0.1–5.0 mM)	T7 phage	38%	[87]
5-(pentafluorophenyl)-10,15,20-tris(1-methylpyridinium-4-yl)porphyrin	Sodium azide (100 mM)	T4 phage	80%	[107]
		Qβ	39%	
	L-histidine (50 mM)	T4 phage	74%	
meso-tetrakis(1-methylpyridinium-4-yl)porphyrin	Sodium azide (100 mM)	T4 phage	90%	[107]
	L-histidine (100 mM)	T4 phage	78%	
5,10,15,20-Tetrakis(4-β-D-glucosylphenyl)porphyrin	1,3-diphenylisobenzofuran (0.1–5.0 mM)	T7 phage	42%	[64]
5,10,15-(4-β-D-glucosylphenyl)-20-phenylporphyrin	1,3-diphenylisobenzofuran (0.1-5.0 mM)	T7 phage	74%	[64]
Polyhydroxylated fullerene	β-carotene	T7 phage	69%	[57]
		PRD1 phage	56%	
	β-carotene (26 μM)	MS2 phage	50%–60%	[90]
Rose bengal	Sodium azide (3.5–35 mM)	M13 phage	31%	[52]

3.2.2.1. Singlet Oxygen in PDI of Mammalian Viruses

Singlet oxygen seems to be the most important mediator of virucidal activity (Table 3) on mammalian viruses. The rate of viral photoinactivation is significantly inhibited by oxygen removal or by addition of singlet oxygen quenchers, such as β-carotene, imidazole, L-histidine or sodium azide [45,84,106–108]. Hypericin may induce photochemical alterations on HIV major capsid protein p24, which are inhibited by sodium azide, suggesting that the damage results from singlet oxygen [111]. When merocyanine 540 [45], phthalocyanine derivatives [106] or rose bengal [108] were used as PS, the results suggest that $^1O_2$ is the main cytotoxic species involved in VSV photoinactivation, while type I reactants such as hydroxyl radicals are less important.

3.2.2.2. Singlet Oxygen in PDI of Bacteriophages

Considering the PDI of bacteriophages in the presence of singlet oxygen quenchers, the results (Table 3) suggest that, in most of the studied cases, singlet oxygen is an important mediator of the

toxic effect induced by PDI. However, the participation of free radicals cannot be ruled out. For instance, the inactivation of M13 bacteriophage by MB was inhibited from 1.72 log to 0.54 log by sodium azide in a quencher dose-dependent mode, up to a concentration of 3.5 mM. However, photoinactivation occurred even in the presence of sodium azide, suggesting that both type I and type II mechanisms may be involved in the M13 photoinactivation process. In the presence of quencher concentrations ranging from 3.5 to 35 mM, a sodium azide protective effect was not observed, as evidenced by increasing rates of M13 phage photoinactivation, reaching a plateau thereafter [52]. Also, the effect of singlet oxygen quenchers and of hydrogen peroxide indicated singlet oxygen as the main factor responsible for the loss of biological activity of bacteriophage M13 by rose bengal [51].

The efficiency of 5,10,15-(4-β-D-galactosylphenyl)-20-(pentafluorophenyl)porphyrin to photoinactivate T7 phage decreased in 38% in the presence of sodium azide [87]. This result, and the ones obtained in the presence of DMTU (Table 2), proved that for this PS, both mechanisms play a role in T7 phage photoinactivation, with type I being the predominant one. Similar results were obtained by Gábor *et al.* [64] in the presence of glycoconjugated *meso*-tetraarylporphyrin derivatives as PS and using 1,3-diphenylisobenzofuran as the singlet oxygen quencher. When T7 phage was phototreated with 5,10,15,20-tetrakis(4-β-D-glucosylphenyl)porphyrin, the rate of inactivation decreased 42% in the presence of 1,3-diphenylisobenzofuran. When 5,10,15-(4-β-D-glucosylphenyl)-20-phenylporphyrin was used, the rate of protection substantially increased (74%). It can then be concluded that the type of PDI mechanism depends on the PS structure, with the symmetric derivative exerting its toxic effect mainly via the generation of free radicals, whether the asymmetric derivative proceeds mainly by singlet production [64]. However, in the study of Egyeki *et al.* [87] using the same asymmetric 5,10,15-(4-β-D-galactosylphenyl)-20-(pentafluorophenyl)porphyrin as PS, and the same phage, the toxic effect occurred mainly via free radical generation. Besides this, the contribution of type I and type II processes was PS concentration-dependent and the sum of the photoinactivation rate measured in the presence of scavengers was smaller than the one measured without the scavengers. This result may imply a synergism between singlet oxygen and hydroxyl radical-mediated damages or it can also be supposed that the efficiency of neither scavenger is 100% [64,87].

A recent study showed that irradiation of polyhydroxylated fullerene suspensions (40 μM) in the presence of β-carotene reduced the photoinactivation rate of PRD1 and T7 phages, demonstrating singlet oxygen involvement [57]. Also, when the T4-like phage was irradiated in the presence of porphyrin derivatives and singlet oxygen quenchers sodium azide and L-histidine, the rate of phage inactivation was considerably reduced, suggesting that singlet oxygen may be an important mediator of the virucidal activity of these PS [107]. However, from the data obtained, other inactivation mechanisms cannot be excluded [57,107].

Although some data about the importance of the type I and II mechanisms in PDI of bacteriophages are discrepant, in general, it seems that the type II pathway is more important than the type I mechanism in phage PDI. On the other hand, there are only a few studies focusing on the simultaneous effect of singlet oxygen and free radicals scavengers under the same protocol of viral PDI [64,84,87,90,106,107].

4. Molecular Targets of Antiviral PDI

The short-lived ROS generated by photodynamic mechanisms are responsible for the damage induced to critical molecular targets [112]. Different viral targets, such as the envelope lipids and proteins, capsid and core proteins and the nucleic acid can be attacked by singlet oxygen and/or other ROS (hydrogen peroxide, superoxide and hydroxyl radicals) to achieve the loss of infectivity [84]. For a better understanding of the photoinactivation process, the knowledge of how the molecular targets are affected by PDI assumes a great importance [113]. For this reason, a detailed photophysical and photochemical study of the interactions between the toxic species generated by the PS and key biomolecules such as lipids, proteins and nucleic acids is essential for the knowledge and prediction of photosensitization process efficiency [114]. However, the studies performed show that the primary target of PDI depends on the chemical structure of the PS, the targeted virus and the mechanism of photoinactivation [64].

4.1. Nucleic Acids

Depending upon the viruses, the nucleic acid can be either DNA or RNA (single or double stranded). The size of the nucleic acid also varies depending on the viruses. Several studies have shown that both DNA and RNA mammalian viruses and phages are efficiently inactivated by PDI. There is now considerable information that PS like MB can bind to and penetrate viral membranes, whereupon they intercalate with nucleic acids. Upon activation by light, the generated ROS can cause the destruction of the nucleic acids, particularly at guanine residues, preventing viral replication [115]. However, there is a difference in target selectivity depending on the mechanism involved: sugar moieties are usually attacked by radicals (generated via type I process) and guanine residues are the targets of singlet oxygen (generated via type II process) [97].

4.1.1. DNA Damage

From the four DNA bases, guanine is the most susceptible component to suffer a type I photosensitization reaction, due to the fact that it exhibits the lowest oxidation potential among DNA bases and it is the only base that can be oxidized by singlet oxygen (type II process) [116].

The treatment of viruses with MB and other heterocyclic dyes resulted in the damage of viral DNA [53,65,75,76] either by base modification or base loss, single strand breaks (SSB), or cross-links of DNA with proteins [34,75,81,88,117]. It is known that cationic porphyrins can bind to nucleic acids via intercalation into base pairs or self-stacking, inducing lesions upon photoinactivation due to the easy oxidation of guanine residues [118–120].

The binding of cationic porphyrins to DNA is presumably due to the electrostatic interaction between the positively-charged substituents in the porphyrin macrocycle and the negatively charged phosphate oxygen atoms of DNA [120]. However, porphyrin binding to DNA is not a prerequisite for an efficient photosensitization, since free porphyrins are more effective in virus inactivation than the

DNA-bound species [88]. This observation, which is in conflict with the generally accepted idea that the porphyrin molecule must be in close vicinity with the site of photosensitized damage, may be explained by the lower quantum yield of singlet oxygen by the bound porphyrin when compared with the free one [88].

4.1.1.1. Damages in the DNA of Mammalian Viruses

Viral DNA is thought to be a critical target structure for PDI by MB and light [93]. DNA isolated from adenovirus treated with 1.3 μM MB exhibited a smear in Southern blot analysis, indicative of random DNA fragmentation [76]. MB plus light treatment of HSV-1 gives rise to DNA damage and blocks DNA replication [121].

4.1.1.2. Damage in the DNA of Bacteriophages

An internal component of T4 phage has been suggested as an important target because MB needs to cross the outer barrier made by its protein capsids in order to produce a significant effect [65]. In fact, some of the irradiated phages could still inject functional genetic material but have lost their ability to form plaques, suggesting that their DNA was damaged. Protein synthesis was also severely impaired [65]. Treatment of M13 phage with MB and aluminum phthalocyanine tetrasulfonate (AlPcS$_4$) caused strand breaks and piperidine-labile bonds in DNA, which is correlated with the loss of infectivity. This is in agreement with the proposal that lesions of the viral genome might be responsible for the lethality induced by sensitization [81]. DNA strand cleavage was found to be MB concentration and light dose dependent. Viral inactivation and DNA damage were found to be oxygen-dependent processes. However, DNA damage was not correlated with the loss of PM2 phage infectivity, as observed in transfection studies which measured the infectivity of the extracted viral DNA, indicating that DNA from MB-treated phage was just as capable of generating progeny virus as the untreated controls [53]. The observed DNA damage is not correlated with loss of phage infectivity and may not be the prime target of viral PDI, because 100% of closed circular DNA was recovered from the MB phototreated PM2 phage [53]. Concerning the effects of PDI on isolated viral DNA, treatment of M13mp2 DNA with increasing concentrations of MB, in the presence of light, yielded increasing amounts of 8-oxo-7,8-dihydro-2'-deoxyguanosine (8-oxodguo), a prevalent adduct produced by singlet oxygen and perhaps by oxygen free radicals. At 100 μM MB, 1 residue of 8-oxodguo was produced for every 40 residues of deoxyguanosine in DNA. Thus, treatment of M13mp2 DNA with MB plus light resulted in putative alterations at deoxyguanosine residues that impede the progression of DNA synthesis *in vitro* [116].

4.1.2. RNA Damage

RNA has been suggested to be a key factor in viral PDI with many PS, but direct evidence of a correlation between RNA damage and infectivity loss has not been reported yet, as is the case of VSV

when treated with phthalocyanine derivatives [81]. In RNA, as for DNA [71], guanine is suggested as the major target for oxidation by photosensitizing agents and light.

4.1.2.1. Damage in the RNA of Mammalian Viruses

VSV genome was damaged by 30 μg mL^{-1} of a chlorophyll derivative and red light illumination which caused a decrease of as much as 85% in RNA polymerase activity, which can be due to damage in the viral RNA polymerase complex, and 98% inhibition of viral RNA synthesis in 6 hours [77]. According to Moor *et al.* [82], the RNA and/or the RNA polymerase complex of VSV might be a major target for its photoinactivation by AlPcS$_4$ and MB. MB and phthalocyanine derivatives inactivated VSV and inhibited fusion of the virus envelope with Vero cells. The degree of inhibition was small compared to the extent of virus inactivation, suggesting that non-membrane targets, like the viral RNA, might be involved in VSV photoinactivation. However, there is no report of a correlation between RNA damage and loss of infectivity [81]. Photoinactivation of HIV-1 by MB and light lead to destruction of its RNA [34].

4.1.2.2. Damage in the RNA of Bacteriophages

Following MB plus light exposure, the Qβ RNA genome exhibited sufficient lethal lesions to account for phage inactivation [122]. However, the protein component of the phage also exerted some effect in viral PDI [122]. In a comparison of RNA photoinactivation using MB and rose bengal as the PS, Schneider *et al.* [54] suggested a causal relationship between 8-oxodguo formation in RNA and R17 and Qβ bacteriophage inactivation. However, no direct relationship between photodynamically induced RNA damage and viral inactivation was described [54]. 8-oxodguo formation or oxidative damage of Qβ RNA alone does not directly account for the lethal event of the virus. Directly treating extracted phage RNA with MB and light caused a loss of activity in the infectious RNA assay but there was a much greater loss of activity if the phage RNA was treated with MB and light in the phage *per se*. The results demonstrated that Qβ RNA infectious activity is significantly more affected by photoinactivation in its protein-associated virion state as compared with its purified isolated polymer state [92,122]. Inactivation of purified RNA by MB and light, in the absence of proteins, most likely occurs due to oxidative damage to the RNA at the site at which MB is bound and might involve oxidized bases such as 8-oxoguanine or strand breaks [122].

In spite of the reduced number of reports focusing on the damage induced by PDI in the nucleic acids of mammalian viruses and bacteriophages, it can be concluded that both DNA and RNA are potential targets of viral PDI. However, there are no studies specifically focusing on the damages induced to DNA and RNA of both mammalian viruses and bacteriophages under the same PDI protocol.

4.2. Outer Structures

Enveloped viruses are inactivated more rapidly than non-enveloped viruses because the destruction of the envelope structure is generally accompanied by loss of virus infectivity [13,40,94,123,124]. The

damages caused by photodynamic reactions on unsaturated lipids present in their envelopes and/or on major envelope proteins, which act as PS binding-sites, modify their structure and avoid cell infection and virus replication [50,84]. However, some studies showed that non-enveloped viruses can also be efficiently inactivated by the toxic action of PS [55,56,58,62,64–67,73,81,87,88,94,122].

The higher susceptibility to PDI of enveloped viruses, relatively to non-enveloped viruses, indicates that the viral envelope may be a more important target than nucleic acids for photosensitization. It also indicates that the unsaturated lipids present in the envelope, as well as the major envelope proteins, are important PDI targets. However, as far as it is known, no studies focus on the degradation of viral envelope lipids after PDI or even on other viral internal lipids. There are, however, many studies about the effects of PDI on viral envelope proteins as well as on other core proteins.

The statement that enveloped viruses are more easily inactivated than non-enveloped ones is only based in indirect studies which compare the inactivation results of enveloped and non-enveloped viruses. The enveloped viruses used in PDI protocols [30,36,45,77,81–83] were only assayed for their protein alterations and no additional experimental work was done concerning their lipids. However, the results of PDI obtained by Lytle *et al.* [94] with the enveloped φ6 phage, although indirectly, are in good accordance with what is reported in the literature about the major contribution from lipids for the viral photoinactivation process.

Relative to proteins degradation by PDI, the results of different studies showed that the main damage is the formation of protein cross-links, followed by other types of damage, which include loss of proteins, alterations in protein molecular conformation, mass and charge, and alterations in protein band intensity (Table 4).

When proteins are irradiated with UV or visible light in the presence of a PS, photooxidation of sensitive amino acid residues such as cysteine, L-histidine, tyrosine, methionine and tryptophan, and covalent cross-linking of peptide chains can be observed, leading to the formation of molecular aggregates [125,126], disrupting their normal folding conformation, thus forcing them into other conformations that affect their normal functioning [127]. In fact, the formation of cross-linked/aggregated material appears to be a major consequence of photosensitized-mediated protein oxidation [128], and it has been demonstrated that the formation of protein cross-links is not a primary photodynamic event, but a secondary reaction between the photooxidation products of sensitive amino acid residues and other groups in the protein [126].

The PS *per se* can induce alterations in the folding of some enzymes, leading to the exposure of some amino acid residues normally shielded in the protein, and to the shielding of others usually exposed in the molecule. These protein modifications lead to changes in properties such as solubility, proteolytic susceptibility, absorbance, and fluorescence emission of several of their amino acids. These alterations are mainly mediated by hydrogen peroxide and hydroxyl radical generation, although singlet oxygen mediated reactions could also occur [129]. The amino acids located in the surface of the protein are photooxidized at a much faster rate than the residues buried in the interior of the molecule. If a protein is completely unfolded, susceptible amino acids may also be attacked and photodegraded [103,130].

Table 4. Degradation of viral outer structures after mammalian viruses and bacteriophages PDI.

Virus	Type of damage	PS	Reference
Enveloped-mammalian viruses			
HSV-1	Viral envelope (reduced ability to adhere to and penetrate host cells)	Merocyanine 540	[45]
	Viral envelope (prevention of viral adsorption and host penetration)	Phthalocyanine derivatives	[131]
	Glycoprotein D; loss of proteins; dimerization; protein cross-links; alterations in protein molecular mass and charge	Phthalocyanine derivatives	[83]
HSV-2	Viral envelope (prevention of viral adsorption and host penetration)	Phthalocyanine derivatives	[131]
HSV	Protein cross-links	Phthalocyanine derivatives	[132]
VZV	Viral envelope (prevention of viral adsorption and host penetration)	Phthalocyanine derivatives	[131]
HIV	Major capsid protein p24	Hypericin	[111]
HIV-1	Loss of infectivity; loss of fusion function; membrane proteins cross-links	Hypericin	[30]
	Loss of infectivity; loss of fusion function; membrane proteins cross-links	Rose bengal	[30]
	p24 and gp120 proteins; protein cross-links	MB	[34]
	Inhibition of cell fusion activity of Env proteins	Natural and sulfonated tetraarylporphyrins	[36]
VSV	Loss of infectivity; loss of fusion function; cross-linking of G and M proteins	Hypericin	[30]
	Loss of infectivity; loss of fusion function; cross-linking of G and M proteins	Rose bengal	[30]
	Inhibition of fusion of the envelope to Vero cells; G protein	MB	[81]
	Inhibition of fusion of the envelope to Vero cells; G protein	Aluminum phthalocyanine tetrasulfonate	[81]
	G and M proteins; protein cross-links	Phthalocyanine derivatives	[82]
	G, M, L and N proteins; protein cross-links	Chlorophyll derivatives	[77]
Influenza virus	Loss of infectivity; loss of fusion function; cross-linking of G and M proteins	Hypericin	[30]
	Loss of infectivity; loss of fusion function; cross-linking of G and M proteins	Rose bengal	[30]
	Loss of infectivity; HA fusion protein; protein cross-links	Rose bengal	[108]

Table 4. *Cont*.

Virus	Type of damage	PS	Reference
Enveloped-mammalian viruses			
Sendai virus	Loss of infectivity; loss of fusion function; cross-linking of G and M proteins	Hypericin	[30]
	Loss of infectivity; loss of fusion function; cross-linking of G and M proteins	Rose bengal	[30]
Vaccinia virus	Histidine residues in virus proteins	Rose bengal	[86]
Human cytomegalovirus	Viral envelope (reduced ability to adhere to and penetrate host cells)	Merocyanine 540	[45]
Sindbis virus	Viral envelope (reduced ability to adhere to and penetrate host cells)	Merocyanine 540	[47]
	Viral capsid protein	Hypericin	[133]
Friend erythroleukemia virus	Viral envelope (reduced ability to adhere to and penetrate host cells)	Merocyanine 540	[134]
Non-enveloped mammalian viruses			
Adenovirus	Not damaged	Phthalocyanine derivatives	[131]
Enterovirus 71	Appearance/disappearance of protein bands; increase of the protein band intensity	Methylene blue	[85]
T7 phage	Protein capsid; loosening of the protein-DNA interaction	Glycoconjugated *meso*-tetraarylporphyrins	[64]
	Capsid and core proteins; loosening of protein-DNA interaction	Glycoconjugated *meso*-tetraarylporphyrins	[87]
	Capsid proteins; protein cross-links	*meso*-Tetrakis(1-methylpyridinium-4-yl)porphyrin	[88]
	Capsid proteins; protein cross-links	Polyhydroxylated fullerene	[57]
M13 phage	Coat protein	Methylene blue Aluminum phthalocyanine tetrasulfunate	[81]
PRD1 phage	Capsid proteins; protein cross-links; phospholipids (less affected)	Polyhydroxylated fullerene	[57]
Qβ phage	Coat and maturation (A) proteins; formation of protein carbonyls; RNA-protein cross-links	Methylene blue	[89]
	RNA-protein cross-links	Methylene blue	[92]
MS2 phage	A protein	Polyhydroxylated fullerene	[57]

4.2.1. Damage on Mammalian Viral Outer Structures

It has been shown that enveloped viruses can be inactivated due to protein damage [30,82,83,131]. However, while the same treatment is reported to be ineffective against some non-enveloped viruses [83,131], the results from Wong *et al.* [85] showed that even a non-enveloped virus can be efficiently inactivated due to the damage induced by PDI to its viral proteins (Table 4).

The proteins in the viral envelope of HSV-1 were considered to be major targets of merocyanine 540 photosensitization [45]. Some phthalocyanine derivatives have been shown to induce cross-links in HSV protein that might be responsible for the observed loss of infectivity [132]. Protein analysis by SDS-PAGE, after treatment with phthalocyanine derivatives, revealed irreversible changes in the HSV-1 envelope proteins, which were reflected by the loss of many proteins, the appearance of cross-linked material on the top of the gel and by alterations in the molecular mass and molecular charge of the proteins. These alterations contribute, in all likelihood to HSV-1 inactivation [83].

In VSV treated with 3.75–30 μL mL^{-1} of chlorophyll derivatives and light, the M protein band was not detected, which was accompanied by a decrease in the intensity of the G protein band [77]. Large complexes of proteins were also detected on the top of the gel, indicating that viral PDI cross-linked the proteins [77]. Using a fusion assay and protein analysis, it was shown that MB and AlPcS$_4$ caused a decrease in the intensity of the G-protein (which is known to play a crucial role in binding VSV to the host cell) band and a slight decrease in the intensity of M protein (matrix protein) band and protein cross-links. However, the observed damage in viral proteins could not account for VSV PDI [82]. VSV was inactivated by MB and phthalocyanine derivatives, which inhibited the fusion of the virus envelope to Vero cells. However, the degree of this inhibition was small compared to the extent of virus inactivation (43% inhibition *vs.* 4.7 log or 99.998% inactivation, for MB) [81]. Abe and Wagner [81] also found few changes in the relative abundance of VSV G protein after MB and AlPcS4 phototreatment, and they also observed additional protein bands on SDS-PAGE analysis [81]. It was found, by Western blot analysis, that HIV-1 p24 and gp120 proteins were altered in size, possibly due to cross-linking, after MB phototreatment [34]. However, using the same PS, AlpcS$_4$ and MB, no changes in protein patterns after SDS-PAGE of the viral proteins were observed, under conditions that caused complete VSV inactivation [135].

The results from Vzorov *et al.* [36] indicated that the porphyrins inhibited the cell fusion activity of HIV Env proteins (a biological function that is important for viral entry as well as induction of viral cytopathic effects) when expressed from recombinant vectors. These results showed that the viral Env protein is an important target of these compounds [36].

PDI of influenza virus by rose bengal altered the HA fusion protein and led to protein cross-links [108].

Photoinactivation of vaccinia virus with rose bengal significantly altered the concentration and oxidized histidine in vaccinia virus protein, suggesting that inactivation was attributed to alterations in viral proteins, as opposed to nucleic acids [86].

Treatment of of influenza and Sindbis viruses by hypericin [30], lead to an extensive cross-linking of the envelope proteins, which may have impaired the capacity of the viruses to adhere to and penetrate the host cells.

The protein profile of the non-enveloped enterovirus 71 was considerably altered after a low dose PDI and a MB concentration ≥ 0.5 μM, as revealed by a smearing and the disappearance of several protein bands [85]. However, enterovirus 71 PDI was also due to damages in the viral genome [85].

4.2.2. Damages on Bacteriophage Outer Structures

In spite of the limited available data for enveloped bacteriophages, substantially higher photoinactivation rates compared with other non-enveloped phages were described [94]. The photoinactivation by merocyanine 540 of four bacteriophages, two non-enveloped phages without lipids (phi X174 and T7), a non-enveloped phage with lipids (PRD1), and an enveloped phage with an external lipoprotein envelope (phi 6) was studied by Lytle *et al.* [94]. The survival curves of the different viruses clearly demonstrated different levels of sensitivity to photoinactivation by this PS, with phi 6 being the most sensitive, followed by T7 (21-fold less sensitive). While both PRD1 and phi 6 have lipid components, only phi 6 was photoinactivated by the PS. Thus, the internal lipid components of PRD1 were not sufficient to allow photoinactivation by merocyanine 540. A higher inactivation rate with a fullerene derivative was also observed by Hotze *et al.* [57] for a phage without lipids (T7 phage) than for PRD1 phage. The dissimilarities in phage composition resulted from differential resistance to singlet oxygen by the outer structures, since PRD1 has a double capsid with an internal lipid membrane, whereas T7 has a single proteinaceous capsid lacking lipids, and both phages contain double stranded DNA with similar GC content (48% for T7 and 51% for PRD1) [57]. Phage proteins were significantly affected by photosensitization (30–92%) when compared to the relatively smaller effect on nucleic acids in both PRD1 and T7, and lipids in PRD1 phage ($\leq 13\%$), as assessed by FTIR spectra analysis [57]. The higher T7 phage inactivation is consistent with greater damage to its proteinaceous capsid. Besides this, SDS-PAGE analysis further evidenced that oxidative cross-linking of capsid proteins induced by exogenous singlet oxygen is the likely cause of phage inactivation [57]. The high propensity for MS2 phage inactivation by this PS (compared to PRD1 and T7 phages) possibly arises from damage to its A protein, which is necessary for infecting its host *Escherichia coli* since it contains highly reactive amino acids such as methionine, cysteine, histidine, and tyrosine and not to damages to the nucleic acid [57]. Glycosylated substituted porphyrins led to structural changes at the protein capsid and/or loosening of the protein-DNA interaction, which can be responsible for T7 phage inactivation [64]. Besides of the alteration of the DNA structure, the phototreatment pointed to significant alterations in the protein structure and/or in the DNA-protein interaction, which may be the cause of photodynamic inactivation [87,88]. The alterations in the DNA secondary structure might also be the result of photochemical damage in phage capsid proteins and consequent disruption of the phage particle. Photomodification of core proteins can also lead to phage inactivation, even if the primary structure of the DNA part is preserved, since these proteins play an important role in the early events of infection and DNA penetration [87]. The damage of T7

nucleoprotein is a complex process and clearly both phage DNA and protein capsid are affected by photoreactions [88]. Irradiation of Qβ bacteriophage in the presence of increasing concentrations of MB resulted in exponentially increasing amounts of viral RNA-protein cross-linkage products, and this is probably the most important event in viral inactivation [92]. The RNA genome of Qβ bacteriophage contained sufficiently lethal lesions following MB plus light exposure to account for the resulting phage inactivation. Nevertheless, the data also indicate that the protein component of the phage somehow contributes to the inactivation of the phage [122]. The protein component of Qβ phage is involved in the process of photoinactivation because the formation of protein carbonyls and RNA-protein cross-links were efficiently formed by MB plus light exposure [89]. The close correlation of cross-link formation with phage inactivation and the expectation that even one such cross-link in a phage genome would be lethal makes the RNA-protein cross-link lesion a strong candidate for the primary inactivating lesion of Qβ phage exposed to MB and light [122].

Little alteration of M13 phage proteins on SDS-PAGE after MB and AlPcS$_4$ photoinactivation was observed by Abe and Wagner [81]. The results of Zupán *et al.* [136], suggested that the tetracationic porphyrin *meso*-tetrakis(1-methylpyridinium-4-yl)porphyrin did not interact with capsid proteins and did not disturb protein-DNA interaction, even if it has a strong stabilization effect on the intraphage DNA.

5. Resistance to PDI and Recovery of Viability

The development of increasing numbers of antiviral agents over the past decades, in the same way as with antibiotics, has provided the clinician with therapeutic options previously unavailable. With the increasing utilization of antiviral drugs, however, has come an enhanced appreciation of the development of antiviral resistance [1,7,137,138,139,140]. Drug resistance is costly to the health service, to the patient who fails to gain maximum therapeutic benefit, and for the community in which resistant viruses may be spread [9].

There is now an urgent need for the development of novel, convenient and inexpensive measures for combating antimicrobial-untreatable infections and limiting the development of additional antimicrobial resistant microorganisms. Photodynamic technology may provide one approach to meet this need, both in terms of therapy and in terms of sterilization, by a mechanism that is markedly different from that typical of most antimicrobials [1,141,142].

As mentioned before, photosensitization involves the generation of singlet oxygen and free radical species, which cause molecular damage. Whether microorganisms could develop resistance to these active oxygen species is still questionable [143] and, consequently, the development of microbial resistance to photosensitization is still under debate. Until now, the development of microbial resistance to PDI is not known and is thought very improbable to be developed. In general, the development of resistance to PDI by microbial strains should be considered as an unlikely event since this process is typically multi-target, with ROS causing damage to many microbial components, which is at a variance with the mechanism of action of most antimicrobial drugs [139,144,145]. In contrast to most common antimicrobials, the number of molecular alterations required to ensure survival would

be too great and the microorganism would require multi-site mutations to become highly resistant, an event with significantly lower probability than single-site mutations, which is often sufficient for conferring resistance to small-molecule inhibitors [42,146]. This particular property of antimicrobial PDI is important regarding the repeated treatment of chronic and/or recurrent infections [139].

Antimicrobial PDI, when compared to standard treatments which may require application for several weeks to achieve an effective killing of the microorganism, shortly after initiation of light exposure, exhibits serious and irreversible damage of microorganisms [66,68]. This damage does not allow the creation or operation of any kind of anti-drug or mutagenic mechanism. Antimicrobial PDI is therefore very effective and, up until now, no photosensitization-resistant mutants have been found [68].

5.1. Resistance of Mammalian Viruses and Recovery of Viability after Photosensitization

Data from North *et al.* [33] show that HIV azidothymidine (AZT)-resistant strains were as susceptible as the AZT-sensitive ones to photosensitization with a benzoporphyrin derivative. This finding comes as no surprise since the mechanisms of action of AZT (inhibition of reverse transcription) and light-activated benzoporphyrin derivative are different. Thus, mutations in the virus that occur at the reverse transcriptase level will not affect photodynamic destruction [33]. Studies focusing on the possible development of viral resistance are extremely scarce and little is known about the recovery of viral viability after consecutive photodynamic treatments.

5.2. Bacteriophage Resistance and Viability Recovery after Photosensitization

Concerning bacteriophages, there is only one study focusing on the possible development of viral resistance after photosensitization [68]. After 10 consecutive cycles of photodynamic treatment, a T4-like phage, in the presence of the tricationic porphyrin 5-(pentafluorophenyl)-10,15,20-tris(1-methylpyridinium-4-yl)porphyrin (Tri-Py$^+$-Me-PF) at 5.0 μM under white light irradiation, exhibited no changes in the rate of photoinactivation during the course of the experiments, meaning that no resistance was observed. If phage resistance would occur, important reductions on phage photoinactivation efficiency would be detected between experiments. Besides that, T4-like phage did not recover its viability after exposure to Tri-Py$^+$-Me-PF during 120 min of irradiation [68]. In a preliminary study by Perdrau and Todd [12], all attempts at reactivating the inactivated *Staphylococcus* phage by MB were unsuccessful.

6. Factors Affecting Viral PDI

6.1. Effect of the Number of Charges, Symmetry, Size of Meso Substituent Groups and Photosensitizer Concentration

It has been shown that the location and binding site of the PS, which is highly dependent on the structure and intramolecular charge distribution, is an important factor in microbial PDI [143,147].

In terms of molecular structure, molecular charge is important in determining antimicrobial activity. Positively charged PS are generally more efficient and can act at lower concentrations than neutral and anionic PS molecules [144]. The positive charges on the PS molecule appear to promote a tight electrostatic interaction between the positively charged PS and the negatively charged sites at the viral capsids and envelopes, orientating the PS toward sites which are critical for the stability and metabolism of a particular microorganism [44,147,148]. This kind of association increases the efficiency of the photoinactivation process.

Cationic PS photodamage can be induced in nucleic acid or viral outer structures by PS binding or by PS localized in its vicinity [136]. For instance, it is more likely that positively charged PS will be effective in causing nucleic acid damage than will neutral or anionic congeners, which mainly act against the outer side of the microorganism [149].

The symmetry and the size of the chain of *meso* substituent groups also affect the photodynamic effect. PS with opposite charged groups are more symmetrical than PS with adjacent charged groups. The adjacent positive charges in the PS macrocycle should result in a molecular distortion due to electrostatic repulsion [150]. The toxicity of a PS can be modulated by the introduction of selected substituents on the macrocycle periphery. In this way, the physicochemical properties of a synthetic PS can be manipulated in order to enhance its interactions with the structural features of the viruses, such as viral capsids, and to minimize the interactions with plasma membranes or mammalian cell membranes [44].

The amphiphilic nature of a PS is another important feature affecting PDI efficiency and can be modulated by the introduction of adequate functionalities in the macrocycle periphery, such as different numbers of positive charges, an asymmetrical charge distribution, or introduction of aromatic hydrocarbon side chains [16,151].

PS concentration is also an important parameter that must be taken into account since viral PDI was shown to be strongly influenced by PS concentration. Increasing the PS concentration reduces the time needed to achieve complete viral inactivation, thus increasing the efficiency of a particular PDI protocol [66].

6.1.1. Mammalian Viruses PDI

Complete inactivation of VSV (4.2 log) can be obtained by treating it with 1.0 µM of the anionic phthalocyanine derivative $AlPcS_4$ and 5 min illumination with red light. For the neutral phthalocyanine derivative (Pc_4), complete inactivation (4 log) was achieved using a much lower amount of PS (4.5 nM) in combination with 10 min illumination [82]. The inactivation of VSV in PBS showed a linear relationship with illumination time [82]. Inactivation of the fusion activity of VSV, influenza and Sendai viruses was reached with nanomolar concentrations of hypericin and rose bengal and was absolutely dependent upon light and increased with increasing time of illumination [30]. HAV in PBS or plasma was completely inactivated within 10 min (>3.7 log) by the cationic symmetric porphyrin *meso*-tetrakis(1-methylpyridinium-4-yl)porphyrin. In contrast, inactivation of HAV to 3.6 log with the anionic symmetric porphyrin *meso*-tetrakis(4-sulfonatephenyl)porphyrin required 90 min [44].

The rate and extent of inactivation appeared to vary with the nature of the *meso* substituent groups [44]. HIV and VSV lost infectivity upon illumination with hypericin and rose bengal in a concentration-dependent manner [30].

6.1.2. Bacteriophage PDI

MS2 phage inactivation has been observed with neutral porphyrin derivatives. However, this required higher irradiation periods (30 min) than for the cationic ones (1 min) [44]. Neutral glycosylated substituted porphyrins can also significantly photoinactivate the T7 phage [64,87]. The T4-like phage PDI was achieved by exposing the phage in the presence of six cationic porphyrins at different concentrations (0.5, 1.0 and 5.0 µM) to white light for 270 min. The results showed that phage photoinactivation varied according with the PS concentration, with higher concentrations being the most efficient ones [66]. The T4-like phage PDI also varied with the number of porphyrin charges, with tri- and tetracationic porphyrin derivatives being more effective in viral inactivation that the dicationic ones, which inactivated the phage below the limit of detection. Tetra- and tricationic porphyrin derivatives (*meso*-tetrakis(1-methylpyridinium-4-yl)porphyrin and 5-(pentafluorophenyl)-10,15,20-tris(1-methylpyridinium-4-yl)porphyrin, respectively) lead to complete T4-like phage inactivation (~7 log) after 270 min of irradiation with 40 W m^{-2} [66]. This tetracationic porphyrin showed similar results in another study (7 log of reduction) for lambda phage inactivation, when irradiated with light of 658 nm [58]. Increasing porphyrin concentration at a fixed light dose leads to increased viral inactivation [58]. A concentration-dependent effect was also detected with a porphyrin derivative [87], but over 2.0 µM of PS the process was saturated. A further increase in porphyrin concentration did not lead to a higher inactivation rate of T7 phage. Aggregation and/or photobleaching of PS are likely explanations [87]. Cationic *meso*-tetrakis(1-alkylpyridinium-4-yl)porphyrin derivatives with different alkyl substituent groups were tested for MS2 phage inactivation but, with the exception of 5,10,15,20-tetrakis(4-sulfonatophenyl)porphyrin, showed toxicity even in the absence of light [44].

In a study conducted by Gábor *et al.* [64], the porphyrin derivative with symmetrical glycosylated groups was found to be twice as effective as the asymmetrical one on the inactivation process of T7 phage. According to Costa and colleagues [66], the rate of T4-like phage inactivation was also dependent on the lipophilic character of the *meso*-substituent groups. The presence of a lipophilic aryl group in one of the *meso* positions of the porphyrin core appears to have an important role in phage inactivation, affecting the rate and efficiency of T4-like phage [66]. Casteel *et al.* [44] have also observed differences in the photoinactivation rate of MS2 phage when they used PS with different alkyl substituent groups and concluded that the rate and extent of inactivation appeared to vary with the nature of the *meso* substituent groups.

6.2. Effect of Different Light Sources and Fluence Rate on Antimicrobial PDT

PDT requires a source of light to activate the PS by exposing it to visible or near-visible light at a specific wavelength [152]. The light source for PDT must also exhibit suitable spectral characteristics

coinciding preferentially with the maximum absorption wavelength range of the PS, applied in order to generate enough ROS to produce an efficient toxic effect [153].

In parallel with the advances in chemistry (related with the discovery and synthesis of new and more efficient PS) there has also been much activity in developing new light sources, better suited for the photosensitization process. Briefly, these include user-friendly lasers frequently based on solid state laser diodes, as well as inexpensive light emitting diodes (LED) and filtered broad-band lamps [154].

PS activation has been achieved via a variety of light sources, such as arc plasma discharge lamps, metal halogen lamps, slide projector illumination assemblies, and a variety of lasers. For treatment of larger areas, non-coherent light sources, such as tungsten filament, quartz halogen, xenon arc, metal halide, and phosphor-coated sodium lamps, are in use. Recently, non-laser light sources, such as LED, have also been applied in PDT. These light sources are much less expensive and small, lightweight and highly flexible, its lifetime can reach up to one hundred thousands hours, and can be manufactured to wavelengths that activate commercially available PS [152,155–159].

At first glance, the available literature on fluence rate effects for PDT seems contradictory. Some studies indicate less damage at low fluence rate, others indicate more killing at lower, compared to higher, fluence rates for the same total fluence and some indicate no influence of fluence rate at all [152,157,158]. A reduction in the fluence rate lowers the rate of oxygen consumption, thereby extending the radius over which singlet oxygen may be formed and consequently increasing the phototoxic effect [159]. Qin *et al.* [160] showed that an increase in the fluence rate increases microbial damage, although, it seems to have an upper limit of photons to observe this effect. Since each PS molecule can only absorb one photon at a time, when the number of light photons bypasses the number of PS molecules, the PS will no longer be able to absorb the photons "in excess" and the rate of PDI will not increase. In fact, if the number of photons is higher than this limit, the antimicrobial effect will decrease because the dye in suspension will not absorb all the excess light [160]. Schindl *et al.* [161] referred that the biological effect of light depends on the fluence, irrespective of the time over which this dose is delivered. Maclean *et al.* [162] also indicate that the inactivating light may be applied at high irradiance over a short time or at lower irradiance over a longer time. A numerical model, assuming that the rate of photodynamic damage occurring at time *t* is proportional to the fluence rate at that time and the local concentrations of PS and oxygen can be established. However, according to this model, relatively low fluence rates can be nearly as effective as high fluence rate sources if applied over the same period of time [163].

There is also a direct correlation between the phototoxic effect and the PS concentration and light fluence. With a lowering of the PS concentration, more light has to be applied to achieve identical effects, and *vice versa*. Lower doses of PS require higher activating light fluences, and higher fluence requires a longer duration of light application [96].

6.2.1. Effect of Light on Mammalian Viruses PDI

The effects of dengue virus inactivation were increased with the increase of MB concentration, the enhancement of power density of the light source and the extension of illumination time, as well as the

decrease of illumination distance. This enabled the narrow bandwidth light system to kill or inactivate the enveloped virus at much greater distance in much shorter time [74]. VSV in the presence of MB was rapidly inactivated by red (provided by LED incident light at 272 W cm^{-2}) or green-yellow light (provided by low-pressure sodium lamps at a fluence rate of 165 W cm^{-2}) but slower by white light (provided by a bank of fluorescent tubes at a fluence rate of 42 W cm^{-2}) [46], showing that higher power densities produce a high rate of viral inactivation than low fluence rates. Wagner *et al.* [164] also showed that red light of 9 W m^{-2}, given at a total dose of 1.8×10^4 and 3.2×10^4 J m^{-2}, inactivated MB-treated VSV by 6 and $\geq$7 log, respectively. VSV inactivation was linearly dependent on the fluence rate of red light illumination [165].

6.2.2. Effect of Light on Bacteriophage PDI

In terms of what is known about phage PDI, only one study focusing on the effect of different light sources and power densities [67] exists. In this study, cationic porphyrin derivatives (*meso*-tetrakis(1-methylpyridinium-4-yl)porphyrin and 5-(pentafluorophenyl)-10,15,20-tris(1-methylpyridinium-4-yl)porphyrin), when irradiated with different sources of light (fluorescent PAR lamps, sun light and halogen lamp) with fluence rates ranging from 40 W m^{-2} to 1690 W m^{-2}, efficiently photoinactivated non-enveloped phages. All light sources tested lead to reductions of about 7 log for the somatic T4-like phage. However, the rate and the extent of inactivation were dependent on the light source, namely when low fluence rates were used (40 W m^{-2}) and on the energy dose, being considerably more effective when light was delivered at a lower fluence rate. However, depending on the light source used, different irradiation periods were required to inactivate T4-like phage to the limits of detection. The results also showed that the efficacy of T4-like phage inactivation, using the same fluence rate, was dependent on the light source used, in particular when the light is delivered at a low fluence rate. M13 phage was phototreated with 5.0 μM MB and was inactivated in an irradiation dose-dependent manner [52]. Kastury and Platz [58] showed that increasing the concentration of a PS at a fixed light dose leads to increased viral inactivation as does an increase in the total light exposure at a fixed PS concentration. The inactivation rate of T1 bacteriophage increased with increasing fluence rate, indicating that the distance of the sample from the light source is a variable which must be controlled [73]. At higher PS concentrations, the inactivation rate reaches a maximum and then decreases, because the filtering effect of the dye decreases the effective fluence rate [73]. In a simple model purposed by Lee *et al.* [56], the phage survival ratio can also be considered as a decreasing exponential fraction of the light fluence (assuming that the fluence is uniform throughout the system).

7. Conclusion

The efficiency of different types of PS in viral PDI has been proved for different types of mammalian viruses and bacteriophages, whether they are enveloped or non-enveloped, for either DNA or RNA viruses. Even though enveloped viruses are more easily inactivated than non-enveloped ones, several studies confirm that non-enveloped mammalian viruses and phages can be efficiently

inactivated by PDI. The type of viral nucleic acid has not been described as an important factor affecting viral photoinactivation but, as far as it is known, no studies specifically focus on the photoinactivation behaviour of DNA and RNA viruses. However, RNA phage MS2 was highly susceptible to photoinactivation when compared with DNA phages under the same conditions of photosensitization.

The type of mechanisms involved in the process of viral photosensitization was already elucidated and singlet oxygen and free radical species were identified as important contributors for an effective viral PDI. However, the contribution of singlet oxygen seems to be more pronounced in mammalian viruses and bacteriophage PDI. There are, however, few studies simultaneously comparing the contribution of both types of mechanisms (type I and type II) involved in viral PDI. The primary targets for the photoinactivation of viruses, whether treating mammalian viruses or phages, are the outer structures. Although there are several studies about the specific effects of PDI on viral proteins, for different types of mammalian viruses and phages, there are no studies concerning the specific effects of PDI on viral lipids. However, it has been clearly shown that enveloped viruses are more easily inactivated than their non-enveloped counterparts, which imply that the lipids present on viral envelopes are important targets of viral PDI.

PS are effective in inactivating the phages to the limits of detection in a way that they do not recover viability, avoiding the development of viral resistance. Nothing is known yet for the particular case of mammalian viruses but, as the viral targets are the same for mammalian viruses and phages, it is also expected that no resistance will be developed in the case of mammalian viruses. Besides that, antiviral PDI is equally effective whether the mammalian virus is sensitive or resistant to conventional antiviral agents. Taking into account all these advantages, PDI for viral inactivation can be regarded as a promising alternative therapy to conventional antiviral treatments, namely for the disinfection of blood and blood products, preventing viral contamination and for the treatment of wound and burn infections. Viral PDI has a fast mode of action and has also the additional benefits of being more economical and an environmental friendly technology, which might be successfully used also in the environmental field for wastewater, drinking water and fish-farming water disinfection.

Different PS concentrations and different light sources and fluence rates were tested, showing that they are important PDI parameters that must inevitably be taken into account when a viral photosensitization protocol has to be elaborated. The inactivation of mammalian viruses and phages can be attained at micromolar-level PS concentrations and different light sources are equally effective, depending on the final dose at which the viruses are exposed to. Besides that, PS can also be modulated by the addition of different *meso* substituent groups and positive charges in order to facilitate their interactions with the viruses, making them more efficient for mammalian viruses and phage PDI.

The similarity of the results obtained for mammalian viruses and bacteriophages show that they exhibit a similar behaviour when submitted to viral photoinactivation techniques: (i) the PS used for viral PDI were equally effective in the photoinactivation of mammalian viruses and bacteriophages; (ii) the mechanism of mammalian viruses and bacteriophage photosensitization involves the

production of singlet oxygen (type II mechanism) with a slight contribution of free radical species (type I mechanism); (iii) singlet oxygen and free radicals were shown to affect viral nucleic acids and also the proteins and lipids present in the mammalian viruses and bacteriophage outer surfaces, with the latter being considerably more affected by PDI; and (iv) the rate and extent of mammalian viruses and phage PDI is also affected by the same factors, like the PS concentration and number of positive charges, the nature and position of *meso* substituent groups, the fluence rate and energy dose. Consequently, it is important to persist in the development of more PDI phage studies to clarify some aspects of viral PDI, such as influence of viral nucleic acid type (DNA or RNA) in the photoinactivation efficiency and the possibility of viral resistance development and viability recovery after photosensitization. It will also be important to study the synergistic effect between viral PDI and antiviral classical methodologies using bacteriophages as models of mammalian virus' photoinactivation.

Acknowledgments

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Jori, G.; Brown, S.B. Photosensitized inactivation of microorganisms. *Photochem. Photobio. Sci.* **2004**, *3*, 403–405.
2. Van Der Poel, W.H.; Vinjé, J.; van Der Heide, R.; Herrera, M.I.; Vivo, A.; Koopmans, M.P. Norwalk-like calicivirus genes in farm animals. *Emerg. Infect. Dis.* **2000**, *6*, 36–41.
3. Blerkom, L.V.L. Role of viruses in human evolution. *Yearbk. Phys. Anthropol.* **2009**, *46*, 14–46.
4. Pulitzer, M.P.; Amin, B.D.; Busam, K.J. Merkel cell carcinoma: Review. *Adv. Anat. Pathol.* **2009**, *16*, 135–44.
5. Sullivan, V.; Biron, K.K.; Talarico, C.; Stanat, S.C.; Davis, M.; Pozzi, L.M.; Coen, D.M. A point mutation in the human cytomegalovirus DNA polymerase gene confers resistance to ganciclovir and phosphonylmethoxyalkyl derivatives. *Antimicrob. Agents Chemother.* **1993**, *37*, 19–25.
6. Smee, D.F.; Barnett, B.B.; Sidwell, R.W.; Reist, E.J.; Holy, A. Antiviral activities of nucleosides and nucleotides against wild-type and drug-resistant strains of murine cytomegalovirus. *Antivir. Res.* **1995**, *26*, 1–9.
7. Kimberlin, D.W.; Whitley, R.J. Antiviral resistance: Mechanisms, clinical significance, and future implications. *J. Antimicrob. Chemother.* **1996**, *37*, 403–421.

8. Jabs, D.A.; Enger, C.; Forman, M.; Dunn, J.P. for The cytomegalovirus retinitis and viral resistance study group. Incidence of foscarnet resistance and cidofovir resistance in patients treated for cytomegalovirus retinitis. *Antimicrob. Agents Chemother.* **1998**, *42*, 2240–2244.

9. Pillay, D.; Zambon, M. Antiviral drug resistance. *Br. Med. J.* **1998**, *317*, 660–662.

10. Smee, D.F.; Sidwell, R.W.; Kefauver, D.; Bray, M.; Huggins, J.W. Characterization of wild-type and cidofovir-resistant strains of camelpox, cowpox, monkeypox, and vaccinia viruses. *Antimicrob. Agents Chemother.* **2002**, *46*, 1329–1335

11. Schultz, E.W.; Krueger, A.P. Inactivation of *Staphylococcus* bacteriophage by methylene blue. *Proc. Soc. Exp. Biol. Med.* **1928**, *26*, 100–101.

12. Perdrau, J.R.; Todd, C. The photodynamic action of methylene blue on certain viruses. *Proc. R. Soc. Lond. B Biol. Sci.* **1933**, *112*, 288–298.

13. Käsermann, F.; Kempf, C. Buckminsterfullerene and photodynamic inactivation of viruses. *Rev. Med. Virol.* **1998**, *8*, 143–151.

14. Hamblin, M.R.; Hasan, T. Photodynamic therapy: A new antimicrobial approach to infectious disease? *Photochem. Photobiol. Sci.* **2004**, *5*, 436–450.

15. Felber, T.D.; Smith, E.B.; Knox, J.M.; Wallis, C.; Melnick, J.L. Photodynamic inactivation of herpes simplex: Report of a clinical trial. *J. Am. Med. Assoc.* **1973**, *92*, 223–289.

16. Almeida, A.; Cunha, A.; Faustino, M.A.F.; Tomé, A.C.; Neves, M.G.P.M.S. Porphyrins as Antimicrobial Photosensitizing Agents. In *Photodynamic Inactivation of Microbial Pathogens: Medical and Environmental Applications*; Royal Society of Chemistry: Cambridge, UK, 2011; pp. 83–160.

17. Mullooly, V.M.; Abramson, A.L.; Shikowitz, M.J. Dihemato-porphyrin ether-induced photosensitivity in laryngeal papilloma patients. *Lasers Surg. Med.* **1990**, *10*, 349–356.

18. Karrer, S.; Szeimies, R.M.; Abels, C.; Wlotzke, U.; Stolz, W.; Landthaler, M. Epidermodysplasia verruciformis treated using topical 5-aminolaevulinic acid photodynamic therapy. *Br. J. Dermatol.* **1999**, *140*, 935–938.

19. Lavie, G.; Mazur, Y.; Lavie, D.; Meruelo, D. The chemical and biological properties of hypericin—A compound with a broad spectrum of biological activities. *Med. Res. Rev.* **1995**, *15*, 111–119.

20. Smetana, Z.; Malik, Z.; Orenstein, A.; Mendelson, E.; Ben-Hur, E. Treatment of viral infections with 5-aminolevulinic acid and light. *Lasers Surg. Med.* **1997**, *21*, 351–358.

21. Sloand, E.M.; Pitt, E.; Klein, H.G. Safety of the blood supply. *J. Am. Med. Assoc.* **1995**, *274*, 1368–1373.

22. Mannucci, P.M. Outbreak of hepatitis A among Italian patients with haemophilia. *Lancet* **1992**, *339*, 819.

23. Klein, H.G. Oxygen carriers and transfusion medicine. *Artif. Cell. Blood Substit. Biotechnol.* **1994**, *22*, 123–135.

24. Azzi, A.; Fanci, R.; Ciappi, S.; Zakrzewska, K.; Bosi, A. Human parvovirus B19 infection in bone marrow transplantation patients. *Am. J. Hematol.* **1993**, *44*, 207–209.

108

25. Asanaka, M.; Kurimura, T.; Toya, H.; Ogaki, J.; Kato Y. Anti-HIV activity of protoporphyrin. *AIDS* **1989**, *3*, 403–404.

26. Dixon, D.W.; Marzilli, L.G.; Schinazi R.F. Porphyrins as agents against the human immunodeficiency virus. *Ann. N. Y. Acad. Sci.* **1990**, *616*, 511–513.

27. Lambrecht, B.; Mohr, H.; Knuver-Hopf, J.; Schmitt, H. Photoinactivation of viruses in human fresh plasma by phenothiazine dyes in combination with visible light. *Vox Sang.* **1991**, *60*, 207–213.

28. Levere, R.D.; Gong, Y.F.; Kappas, A.; Bucher, D.J.; Wormser, G.; Abraham, N.G. Heme inhibits human immunodeficiency virus 1 replication in cell cultures and enhances the antiviral effect of zidovudine. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 1756–1759.

29. Matthews, J.L.; Sogandares-Bernal, F.; Judy, M.; Gulliya, K.; Newman, J.; Chanh, T.; Marengo-Rowe, A.J. Inactivation of viruses with photoactive compounds. *Blood Cell.* **1992**, *18*, 75–88.

30. Lenard, J.; Rabson, A.; Vanderoef, R. Photodynamic inactivation of infectivity of humam immunodeficiency virus and other enveloped viruses using hypericin and rose bengal: Inhibition of fusion and syncytia formation. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 158–162.

31. Neurath, A.R.; Strick, N.; Jiang, S. Rapid prescreening for antiviral agents against HIV-1 based on their inhibitory activity in site-directed immunoassays—Approaches applicable to epidemic HIV-1 strains. *Antivir. Chem. Chemother.* **1993**, *4*, 207–214.

32. Debnath, A.K.; Jiang, S.; Strick, N.; Lin, K.; Haberfield, P.; Neurath, A.R. 3-Dimensional structure-activity analysis of a series of porphyrin derivatives withanti-HIV-1 activity targeted to the v3 loop of the gp120 envelope glycoprotein of the human-immunodeficiency-virus type 1. *J. Med. Chem.* **1994**, *37*, 1099–1108.

33. North, J.; Coombs, R.; Levy, J. Photodynamic inactivation of free and cell-associated HIV-1 using the photosensitizer, benzoporphyrin derivative. *J. Acquir. Immune Defic. Syndr.* **1994**, *7*, 891–898.

34. Bachmann, B.K.-H.J.B.; Lambrecht, B.; Mohr, H. Target structures for HIV-1 inactivation by methylene blue and light. *J. Med. Virol.* **1995**, *47*, 172–178.

35. Song, R.; Witvrouw, M.; Schols, D.; Robert, A.; Balzarini, J.; de Clercq, E.; Bernadou, J.; Meunier, B. 1997. Anti-HIV activities of anionic metalloporphyrins and related compounds. *Antivir. Chem. Chemother.* **1997**, *8*, 85–97.

36. Vzorov, A.N.; Dixon, D.W.; Trommel, J.S.; Marzilli, L.G.; Compans, R.W. Inactivation of human immunodeficiency virus type 1 by porphyrins. *Antimicrob. Agents Chemother.* **2002**, *46*, 3917–3925.

37. Vanyur, R.; Heberger, K.; Jakus, J. Prediction of anti-HIV-1 activity of a series of tetrapyrrole molecules. *J. Chem. Inform. Comput. Sci.* **2003**, *43*, 1829–1836.

38. Dairou, J.; Vever-Bizet, C.; Brault, D. Interaction of sulfonated anionic porphyrins with HIV glycoprotein gp120: photodamages revealed by inhibition of antibody binding to V3 and C5 domains. *Antivir. Res.* **2004**, *61*, 37–47.

39. Marchesan, S.; Da Ros, T.; Spalluto, G.; Balzarini, J.; Prato, M. Anti-HIV properties of cationic fullerene derivatives. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3615–3618.

40. North, J.; Freeman, S.; Overbaugh, J.; Levy, J.; Lansman, R. Photodynamic inactivation of retrovirus by benzoporphyrin derivative: A feline leukemia virus model. *Transfusion* **1992**, 32, 121–128.

41. Müller-Breitkreutz, K.; Mohr, H. Hepatitis C and human immunodeficiency virus RNA degradation by methylene blue/light treatment of human plasma. *J. Med. Virol.* **1998**, *56*, 239–245.

42. Cheng, Y.; Tsou, L.K.; Cai, J.; Aya, T.; Dutschman, G.E.; Gullen, E.A.; Grill, S.P.; Chen, A.P.-C.; Lindenbach, B.D.; Hamilton, A.D.; *et al*. A novel class of meso-tetrakis-porphyrin derivatives exhibits potent activities against hepatitis C virus genotype 1b replicons *in vitro*. *Antimicrob. Agents Chemother.* **2010**, *54*, 197–206.

43. Lin L.; Hu, J. Inhibition of hepadnavirus reverse transcriptase RNA interaction by porphyrin compounds. *J. Virol.* **2008**, *82*, 2305–2312.

44. Casteel, B.M.J.; Jayaraj, K.; Avram, G.; Bail, L.M.; Sobsey, M.D. Photoinactivation of hepatitis A virus by synthetic porphyrins. *Photochem. Photobiol.* **2004**, *80*, 294–300.

45. O'Brien, J.M.; Gaffney, D.K.; Wang, T.P.; Sieber, F. Merocyanine 540 sensitized photoinactivation of enveloped viruses in blood products: Site and mechanism of phototoxicity. *Blood* **1992**, *80*, 277–285.

46. Mohr, H.; Bachmann, B.; Klein-Struckmeier, A.; Lambrecht, B. Virus inactivation of blood products by phenothiazine dyes and light. *Photochem. Photobiol.* **1997**, *65*, 441–445.

47. Sieber, F.; O'Brien, J.M.; Krueger, G.J.; Schober, S.L.; Burns, W.H.; Sharkis, S.J.; Sensenbrenner, L.L. Antiviral activity of merocyanine 540. *Photochem. Photobiol.* **1987**, *46*, 707–711.

48. Leclerc, H.; Edberg, S.; Pierzo, V.; Delattre, J.M. Bacteriophages as indicators of enteric viruses and public health risk in groundwaters. A review. *J. Appl. Microbiol.* **2000**, *88*, 5–21.

49. Rywkin, S.; Ben-Hur, E.; Malik, Z.; Prince, A.M.; Li, Y.S.; Kenney, M.E.; Oleinick, N.L.; Horowitz, B. New phthalocynanines for photodynamic virus inactivation in red blood cell concentrates. *Photochem. Photobiol.* **1994**, *60*, 165–170.

50. Käsermann, F.; Kempf, C. Photodynamic inactivation of enveloped viruses by buckminsterfullerene. *Antivir. Res.* **1997**, *34*, 65–70.

51. DiMascio, P.; Wefers, H.; Do-Thi, H-P.; Lafleur, M.V.M.; Sies, H. Singlet molecular oxygen causes loss of biological activity in plasmid and bacteriophage DNA and induces single strand breaks. *Biochim. Biophys. Acta* **1989**, *1007*, 151–157.

52. Abe, H.; Ikebuchi,K.; Wagner, S.J.; Kuwabara, M.; Kamo, N.; Sekiguchi, S. Potential involvement of both type I and type II mechanisms in M13 virus inactivation by methylene blue photosensitization. *Photochem. Photobiol.* **1997**, *66*, 204–208.

53. Specht, K.G. The role of DNA damage in PM2 viral inactivation by methylene blue photosensitization. *Photochem. Photobiol.* **1994**, *59*, 506–514.

54. Schneider, J.E.; Philips, J.R.; Pye, Q.; Maidt, M.L.; Price, S.; Floyd, R.A. Methylene blue and rose bengal photoinactivation of RNA bacteriophages: Comparative studies of 8-oxoguanine formation in isolated RNA. *Arch. Biochem. Biophys.* **1993**, *301*, 91–97.

55. Jockush, S.; Lee, D.; Turro, N.J.; Leonard, E.F. Photoinduced inactivation of viruses: Adsorption of methylene blue, thionine and thiopyronine on Qβ bacteriophage. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 7446–7451.

56. Lee, D.; Foux, M.; Leonard, E.F. The effects of methylene blue and oxygen concentration on the photoinactivation of Qβ bacteriophage. *Photochem. Photobiol.* **1997**, *65*, 161–165.

57. Hotze, E.M.; Badireddy, A.R.; Chellam, S.; Wiesner, M.R. Mechanisms of bacteriophage inactivation via singlet oxygen generation in UV illuminated fullerol suspensions. *Environ. Sci. Technol.* **2009**, *43*, 6639–6645.

58. Kasturi, C.; Platz, M.S. Inactivation of lambda phage with 658 nm light using a DNA binding porphyrin sensitizer. *Photochem. Photobiol.* **1992**, *56*, 427–429.

59. Martin, C.B.; Wilfong, E.; Ruane, P.; Goodrich, R.; Platz, M. An action spectrum of the riboflavin-photosensitized inactivation of lambda phage. *Photochem. Photobiol.* **2005**, *81*, 474–480.

60. Wagner, S.J.; Skripchenkol, A.; Robinenel, D.; Foley, J.W.; Cincotta, L. Factors affecting virus photoinactivation by a series of phenothiazine dyes. *Photochem. Photobiol.* **1998**, *67*, 343–349.

61. Brendel, M. Different photodynamic action of proflavine and methylene blue on bacteriophage. I. Host cell reactivation of *Serratia*phage *kappa*. *Mol. Gen. Genet.* **1970**, *108*, 303–311.

62. Yamamoto, N. Photodynamic inactivation of bacteriophage and its inhibition. *J. Bacteriol.* **1957**, *6*, 510–521.

63. Witmer, H.; Fraser, D. Photodynamic action of proflavine on coliphage T3 II. Protection by L-cysteine. *J. Virol.* **1971**, *7*, 319–322.

64. Gábor, F.; Szolnoki, J.; Tóth, K.; Fekete, A.; Maillard, P.; Csík, G. Photoinduced inactivation of T7 phage sensitized by symmetrically and asymmetrically substituted tetraphenyl porphyrin: comparison of efficiency and mechanism of action. *Photochem. Photobiol.* **2001**, *73*, 304–311.

65. Kadish, L.L.; Fisher, D.B.; Pardee, A.B. Photodynamic inactivation of free and vegetative bacteriophage T4. *Biochim. Biophys. Acta* **1967**, *138*, 57–65.

66. Costa, L.; Alves, E.; Carvalho, C.M.B.; Tomé, J.P.C.; Faustino, M.A.F.; Neves, M.G.P.M.S.; Tomé, A.C.; Cavaleiro, J.A.S.; Cunha, A.; Almeida, A. Sewage bacteriophage photoinactivation by cationic porphyrins: a study of charge effect. *Photochem. Photobiol. Sci.* **2008**, *7*, 415–422.

67. Costa, L.; Carvalho, C.M.B.; Faustino, M.A.F.; Neves, M.G.P.M.S.; Tomé, J.P.C.; Tomé, A.C.; Cavaleiro, J.A.S.; Cunha, Â.; Almeida, A. Sewage bacteriophage inactivation by cationic porphyrins: influence of light parameters. *Photochem. Photobiol. Sci.* **2010**, *9*, 1126–1133.

68. Costa, L.; Tomé, J.P.C.; Neves, M.G.P.M.S.; Tomé, A.C.; Cavaleiro, J.A.S.; Faustino, M.A.F.; Cunha, Â.; Gomes, N.C.M.; Almeida, A. Evaluation of resistance development and viability recovery by a non-enveloped virus after repeated cycles of aPDT. *Antivir. Res.* **2011**, *91*, 278–282.

69. DeRosa, M.C.; Crutchley, R.J. Photosensitized singlet oxygen and its applications. *Coord. Chem. Rev.* **2002**, *233–234*, 351–371.

70. Capella, M.A.M.; Capella, L.S.A light in multidrug resistance: Photodynamic treatment of multidrug-resistant tumors. *J. Biomed. Sci.* **2003**, *10*, 361–366.

71. Castano, A.P.; Demidova, T.N.; Hamblin, M.R. Mechanisms in photodynamic therapy: Part one-photosensitizers, photochemistry and cellular localization. *Photodiagn. Photodyn.* **2004**, *1*, 279–293.

72. Prates, R.A.; da Silva, E.G.; Yomada, A.M.; Jr.; Suzuki, L.C.; Paula, C.R.; Ribeiro, M.S. Light parameters influence cell viability in antifungal photodynamic therapy in a fluence and rate fluence dependent manner. *Laser Phys.* **2009**, *19*, 1038–1044.

73. Welsh, J.N.; Adams, M.H. Photodynamic inactivation of bacteriophage. *J. Bacteriol.* **1954**, *1*, 122–127.

74. Huang, Q.; Fu, W-L.; Chen, B.; Huang, J-F.; Zhang, X.; Xue, Q. Inactivation of dengue virus by methylene blue/narrow bandwidth light system. *J. Photochem. Photobiol. B Biol.* **2004**, *77*, 39–43.

75. Schnipper, L.E.; Lewin, A.A.; Swartz, M.; Crumpacker, C.S. Mechanisms of photodynamic inactivation of herpes simplex viruses; comparison between methylene blue, light plus electricity, and hematoporphyrin plus light. *J. Clin. Investig.* **1980**, *65*, 432–438.

76. Schagen, F.H.E.; Moor, A.C.E.; Cheong, S.C.; Cramer, S.J.; van Ormondt, H.; van der Eb, A.J.; Dubbelman T.M.A.R.; Hoeben, R.C. Photodynamic treatment of adenoviral vectors with visible light: An easy and convenient method for viral inactivation. *Gene Ther.* **1999**, *6*, 873–881.

77. Lim, D.-S.; Ko, S.-H.; Kim, S.-J.; Park, Y.-J.; Park, J.-H.; Lee, W.-Y. Photoinactivation of vesicular stomatitis virus by a photodynamic agent, chlorophyll derivatives from silkworm excreta. *J. Photochem. Photobiol. B Biol.* **2002**, *67*, 149–156.

78. Sagristá, M.L.; Postigo, F.; De Madariaga, M.A.; Pinto, R.M.; Caballero, S.; Bosch, A.; Vallés, M.A.; Mora, M. Photodynamic inactivation of viruses by immobilized chlorine-containing liposomes. *J. Porphyrin Phthalocyanines* **2009**, *13*, 578–588.

79. Tomé, J.P.C.; Neves, M.G.P.M.S.; Tomé, A.C.; Cavaleiro, J.A.S.; Mendonça, A.F.; Pegado, I.N.; Duarte, R.; Valdeira, M.L. Synthesis of glycoporphyrin derivatives and their antiviral activity against herpes simplex virus types 1 and 2. *Bioorg. Med. Chem.* **2005**, *13*, 3878–3888.

80. Silva, E.M.P.; Giuntini, F.; Faustino, M.A.F.; Tomé, J.P.C.; Neves, M.G.P.M.S.; Tomé, A.C.; Silva, A.M.S.; Santana-Marques, M.G.; Ferrer-Correia, A.J.; Cavaleiro, J.A.S.; *et al.* Synthesis of cationic β-vinyl substituted *meso*-tetraphenylporphyrins and their *in vitro* activity against herpes simplex virus type 1. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3333–3337.

81. Abe, H.; Wagner, S.J. Analysis of viral DNA, protein and envelope damage after methylene blue, phthalocyanine derivative or merocyanine 540 photosensitization. *Photochem. Photobiol.* **1995**, *61*, 402–409.

112

82. Moor, A.C.E.; Wagenaars-van Gompel, A.E.; Brand, A.; Dubbelman, T.M.A.R.; Van Steveninck, J. Primary targets for photoinactivation of vesicular stomatitis virus by AlPcS$_4$ or Pc$_4$ and red light. *Photochem. Photobiol.* **1997**, *65*, 465–470.

83. Smetana, Z., Ben-Hur, E.; Mendelson, E.; Salzberg, S.; Wagner, P.; Malik, Z. Herpes simplex virus proteins are damaged following photodynamic inactivation with phthalocyanines. *J. Photochem. Photobiol. B Biol.* **1998**, *44*, 77–83.

84. Müller-Breitkreutz, K.; Mohr, H.; Briviba, K.; Sies, H. Inactivation of viruses by chemically and photochemically generated singlet molecular oxygen. *J. Photochem. Photobiol. B Biol.* **1995**, *30*, 63–70.

85. Wong, T.-W.; Huang, H.-J.; Wang, Y.-F.; Lee, Y.-P.; Huang, C.-C.; Yu, C.-K. Methylene blue-mediated photodynamic inactivation as a novel disinfectant of enterovirus 71. *J. Antimicrob. Chemother.* **2010**, *65*, 2176–2182.

86. Turner, G.S.; Kaplan, C. Photoinactivation of vaccinia virus with rose bengal. *J. Gen. Virol.* **1968**, *3*, 433–443.

87. Egyeki, M.; Turóczy, G.; Majer, Zs.; Tóth, K.; Fekete, A.; Maillard, Ph.; Csík, G. Photosensitized inactivation of T7 phage as surrogate of non-enveloped DNA viruses: Efficiency and mechanism of action. *Biochim. Biophys. Acta* **2003**, *1624*, 115–124.

88. Zupán, K.; Egyeki, M.; Tóth, K.; Fekete, A.; Herényi, L.; Módos, K.; Csík, G. Comparison of the efficiency and the specificity of DNA-bound and free cationic porphyrin in photodynamic virus inactivation. *J. Photochem. Photobiol. B Biol.* **2008**, *90*, 105–112.

89. Schneider, J.E.; Jr.; Tabatabale, T.; Maidt, L.; Smith, R.H.; Nguyen, X.; Pye, Q.; Floyd, R.A. Potential mechanisms of photodynamic inactivation of virus by methylene blue I. RNA-protein crosslinks and other oxidative lesions in Qβ bacteriophage. *Photochem. Photobiol.* **1998**, *67*, 350–357.

90. Badireddy, A.R.; Hotze, E.M.; Chellam, S.; Alvarez, P.J.J.; Wiesner, M.R. Inactivation of bacteriophages via photosensitization of fullerol nanoparticles. *Environ. Sci. Tech.* **2007**, *41*, 6627–6632.

91. Marotti, J.; Aranha, A.C.C.; Eduardo, C.D.P.; Ribeiro, M.S. Photodynamic therapy can be effective as a treatment for herpes simplex labialis. *Photomed. Laser Surg.* **2009**, *27*, 357–363.

92. Floyd, R.A.; Schneider, J.E.; Dittmer, D.P. Methylene blue photoinactivation of RNA viruses. *Antivir. Res.* **2004**, *61*, 141–151.

93. Wallis, C.; Melnick, J.L. Photodynamic inactivation of animal viruses: A review. *Photochem. Photobiol.* **1965**, *4*, 159–170.

94. Lytle, C.D.; Budacz, A.P.; Keville, E.; Miller, S.A.; Prodouz, K.N. Differential inactivation of surrogate viruses with merocyanine 540. *Photochem. Photobiol.* **1991**, *54*, 489–493.

95. Via, L.D.; Magno, S.M. Photochemotherapy in the treatment of cancer. *Curr. Med. Chem.* **2001**, *8*, 1405–1418.

96. Schmidt-Erfurth, U.; Hasan, T. Mechanisms of action of photodynamic therapy with verteporfin for the treatment of age-related macular degeneration. *Surv. Ophthalmol.* **2000**, *45*, 195–214.

97. Wainwright, M. Photodynamic antimicrobial chemotherapy (PACT). *J. Antimicrob. Chemother.* **1998**, *42*, 13–28.

98. Bonnett, R. *Chemical Aspects of Photodynamic Therapy*; Gordon and Breach Science Publishers: Amsterdam, The Netherlands, 2000.

99. Girotti, A.W. Photosensitized oxidation of membrane lipids: Reaction pathways, cytotoxic effects, and cytoprotective mechanisms. *J. Photochem. Photobiol. B Biol.* **2001**, *63*, 103–113.

100. Calin, M.A.; Parasca, S.V. Light sources for photodynamic inactivation of bacteria. *Laser Med. Sci.* **2009**, *24*, 453–460.

101. Min, D.B.; Boff, J.M. Chemistry and reaction of singlet oxygen in foods. *Compr. Rev. Food Sci. Food Saf.* **2002**, *1*, 58–72.

102. Maisch, T.; Bosl, C.; Szeimies, R.M.; Lehn, N.; Abels, C. Photodynamic effects of novel XF porphyrin derivativeson prokaryotic and eukaryotic cells. *Antimicrob. Agents Ch.* **2005**, *49*, 1542–1552.

103. Ochsner, M. Photophysical and photobiological processes in the photodynamic therapy of tumours. *J. Photochem. Photobiol. B Biol.* **1997**, *39*, 1–18.

104. Wondrak, G.T.; Jacobson, M.K.; Jacobson, E.L. Identification of quenchers of photoexcited states as novel agents for skin photoprotection. *J. Pharmacol. Exp. Therapeut.* **2005**, *312*, 482–491.

105. Sies, H. Oxidative stress: Oxidants and antioxidants. *Exp. Physiol.* **1997**, *82*, 291–295.

106. Rywkin, S.; Lenny, L.; Goldstein, J.; Geacintov, N.E.; Margolis-Nunno, H.; Horowitz, B. Importance of type I and type II mechanisms in the photodynamic inactivation of viruses in blood with aluminum phthalocyanine derivatives. *Photochem. Photobiol.* **1992**, *56*, 463–469.

107. Costa, L.; Tomé, J.P.C.; Faustino, M.A.F.; Neves, M.G.P.S.; Tomé, A.C.; Cavaleiro, J.A.S.; Cunha, A.; Almeida, A. Involvement of type I and type II mechanisms on the photoinactivation of non-enveloped DNA and RNA bacteriophages. *Environ. Sci. Technol.* **2012**, submitted for publication.

108. Lenard, J.; Vanderoef, R. Photoinactivation of influenza virus fusion and infectivity by rose bengal. *Photochem. Photobiol.* **1993**, *58*, 527–531.

109. Bisby, R.H.; Morgan, C.G.; Hamblett, I.; Gorman, A.A. 1999. Quenching of singlet oxygen by trolox c, ascorbate, and amino acids: effects on pH and temperature. *J. Phys. Chem. A* **1999**, *103*, 7454–7459.

110. Baker, A.; Kanofsky, J.R. Quenching of singlet oxygen bybiomolecules from Ll210 leukemia cells. *Photochem. Photobiol.* **1992**, *55*, 523–528.

111. Degar, S.; Prince, A.M.; Pascual, D.; Lavie, G.; Levin, B.; Mazur, Y.; Lavie, D.; Ehrlich, L.S.; Carter, C.; Meruelo, D. Inactivation of the human immunodeficiency virus by hypericin: Evidence for photochemical alterations of p24 and a block in uncoating. *AIDS Res. Hum. Retrovir.* **1992**, *8*, 1929–1936.

112. Wainwright, M. Local treatment of viral disease using photodynamic therapy. *Int. J. Antimicrob. Agents* **2003**, *21*, 510–520.

113. Garcia, G.; Sarrazy, V.; Sol, V.; Morvan, C.L.; Granet, R.; Alves, S.; Krausz, P. DNA photocleavage by porphyrin–polyamine conjugates. *Bioorg. Med. Chem.* **2009**, *17*, 767–776.

114. Miranda, M.A. Photosensitization by drugs. *Pure Appl. Chem.* **2001**, *73*, 481–486.

115. Wainwright, M. The use of methylene blue derivatives in blood product disinfection. *Int. J. Antimicrob. Agents* **2000**, *16*, 381–394.

116. McBride, T.J.; Schneider, J.E.; Floyd, R.E.; Loeb, L.A. Mutations induced by methylene blue plus light in single stranded M13mp2. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 6866–6870.

117. OhUigin, C.; McConnell, D.J.; Kelly, J.M.; van der Putten, W.J.M. Methylene blue photosensitised strand cleavage of DNA: Effects of dye binding and oxygen. *Nucleic Acids Res.* **1987**, *15*, 7411–7427.

118. Mettath. S.; Munson, B.R.; Pandey, R.K. DNA interaction and photocleavage properties of porphyrins containing cationic substituents at the peripheral position. *Bioconjugate Chem.* **1999**, *10*, 94–102.

119. Kubát, P.; Lang, K.; Anzenbacher, P.; Jr.; Jursíkova, K.; Král, V.; Ehrenberg, B. Interaction of novel cationic *meso*-tetraphenylporphyrins in the ground and excited states with DNA and nucleotides. *J. Chem. Soc. Perkin Trans. 1* **2000**, *1*, 933–941.

120. Caminos, D.A.; Durantini, E.N. Interaction and photodynamic activity of cationic porphyrin derivativesbearing different patterns of charge distribution with GMP and DNA. *J. Photochem. Photobiol. A: Chem.* **2008**, *198*, 274–281.

121. Müller-Breitkreutz, K.; Mohr, H. Infection cycle of herpes viruses after photodynamic treatment with methylene blue and light. *Transfusions Medizin* **1997**, *34*, 37–42.

122. Schneider, J.E.; Jr.; Pye, Q.; Floyd, R.A. Qβ bacteriophage photoinactivated by methylene blue plus light involves inactivation of its genomic RNA. *Photochem. Photobiol.* **1999**, *70*, 902–909.

123. Smetana, Z.; Mendelson, E.; Manor, J.; Van Lier, J.E.; Ben-Hur, E.; Salzberg, S.; Malik, Z. Photodynamic inactivation of herpes simplex viruses with phthalocyanine derivatives. *J. Photochem. Photobiol. B: Biol.* **1994**, *22*, 37–43.

124. Ben-Hur, E.; Horowitz, B. Virus inactivation in blood. *AIDS* **1996**, *11*, 1183–1190.

125. Girotti, A.W. Photodynamic action of protoporphyrin IX on human erythrocytes: Cross-linking of membrane proteins. *Biochem. Biophys. Res. Comm.* **1976**, *72*, 1367–1374.

126. Verweij, H.; van Steveninck, J. Model studies on photodynamic cross-linking. *Photochem. Photobiol.* **1982**, *35*, 265–267.

127. Macdonald, I.J.; Dougherty, T.J. Basic principles of photodynamic therapy. *J. Porphyrin Phthalocyanines* **2001**, *5*, 105–129.

128. Davies, M.J. Singlet oxygen-mediated damage to proteins and its consequences. *Biochem. Biophys. Res. Commun.* **2003**, *305*, 761–770.

129. Afonso, S.G; Enriquez, S.R.; Batlle, C.A.M. The photodynamic and non photodynamic actions of porphyrins. *Braz. J. Med. Biol. Res.* **1999**, *32*, 255–266.

130. Jori, G.; Galiazzo, G.; Tamburro, A.M.; Scoffone, E. Dye-sensitized photooxidation as a tool for determining the degree of exposure of amino acid residues in proteins. *J. Biol. Chem.* **1970**, *245*, 3375–3383.

131. Malik, Z.; Ladan, H.; Nitzan, Y.; Smetana, Z. Antimicrobial and antiviral activity of porphyrin photosensitization. *Proc. SPIE* **1993**, *2078*, 305–312.

132. Malik, Z.; Smetana, Z.; Mendelson, E., Wagner, P.; Salzberg, S.; Ben-Hur, E. Alteration in herpes simplex virus proteins following photodynamic treatment with phthalocyanines. *Photochem. Photobiol.* **1996**, *63*, 59S.

133. Yip, L.; Hudson, J.B.; Gruszecka-Kowalik, E.; Zalkow, L.H.; Neil Towers, G.H.N. Antiviral activity of a derivative of the photosensitive compound hypericin. *Phytomedicine* **1996**, *2*, 185–190.

134. Sieber, F.; Krueger, G.J.; O'Brien, J.M.; Schober, S.L.; Sensenbrenner, L.L.; Sharkis, S.J. Inactivation of Friend erythroleukemia virus and Friend virus-transformed cells by merocyanine 540-mediated photosensitization. *Blood* **1989**, *73*, 345–350.

135. Melki, R.; Gaudin, Y.; Blondel, D. Interaction between tubulin and the viral matrix protein of vesicular stomatitis virus: possible implications in the viral cytopathic effect. *Virology* **1994**, *202*, 339–347.

136. Zupán, K.; Herényi, L.; Tóth, K.; Majer, Z.; Csík, G. Binding of cationic porphyrin to isolated and encapsidated viral DNA analyzed by comprehensive spectroscopic methods. *Biochemistry* **2004**, *43*, 9151–9159.

137. Yoshikawa, T.T. Antimicrobial resistance and aging: beginning of the end of the antibiotic era? *J. Am. Geriatr. Soc.* **2002**, *50*, S226–S229.

138. Malik, Z.; Gozhansky, S.; Nitzan, Y. Effects of photoactivated HPD on bacteria and antibiotic resistance. *Microbios Lett.* **1982**, *21*, 103–112.

139. Maisch, T.; Szeimies, R-M.; Jori, G.; Abels, C. Antibacterial photodynamic therapy in dermatology. *Photochem. Photobiol. Sci.* **2004**, *3*, 907–917.

140. Pillay, D. Emergence and control of resistance to antiviral drugs in resistance in herpes viruses, hepatitis B virus, and HIV. *Commun. Dis. Public Health* **1998**, *1*, 5–13.

141. Reddi, E.; Ceccon, M.; Valduga, G.; Jori, G.; Bommer, J.C.; Elisei, F.; Latterini, L.; Mazzucato, U. Photophysical properties and antibacterial activity of *meso*-substituted cationic porphyrins. *Photochem. Photobiol.* **2002**, *75*, 462–470.

142. Jori, G.; Coppellotti, O. Inactivation of pathogenic microorganisms by photodynamic techniques: Mechanistic aspects and perspective applications. *Anti-Infect. Agents Med. Chem.* **2007**, *6*, 119–131.

143. Minnock, A.; Vernon, D.I.; Schofield, J.; Griffiths, J.; Parish, J.H.; Brown, S.B. Mechanism of uptake of a cationic water-soluble pyridinium zinc phthalocyanine across the outer membrane of *Escherichia coli*. *Antimicrob. Agents Chemother.* **2000**, *44*, 522–527.

144. Demidova, T.; Hamblin, M. Effects of cell-photosensitizer binding and cell density on microbial photoinactivation. *Antimicrob. Agents Chemother.* **2005**, *6*, 2329–2335.

145. Jori, G.; Fabris, C.; Soncin, M.; Ferro, S.; Coppellotti, O.; Dei, D.; Fantetti, L.; Chiti, G.; Roncucci, G. Photodynamic therapy in the treatment of microbial infections: Basic principles and perspective applications. *Lasers Surg. Med.* **2006**, *38*, 468–481.

146. Wainwright, M. Photoantimicrobials—So what's stopping us? *Photodiagn. Photodyn.* **2009**, *6*, 167–169.

147. Merchat, M.; Bertolini, G.; Giacomini, P.; Villanueva, A.; Jori, G. *Meso*-substituted cationic porphyrins as efficient photosensitizers of Gram-positive and Gram-negative bacteria. *J. Photochem. Photobiol.* **1996**, *32*, 153–157.

148. Dowd, S.E., Pillai, S.D.; Wang, S.; Corapcioglu, M.Y. Delineating the specific influence of virus isoelectric point and size on virus adsorption and transport through sandy soils. *Appl. Environ. Microbiol.* **1998**, *64*, 405–410.

149. Wainwright, W.; Photoantimicrobials—A PACT against resistance and infection. *Drugs Future* **2004**, *29*, 85–93.

150. Kessel, D.; Raymund, L.; Vicente, M.G.H. Localization and photodynamic efficacy of two cationic porphyrins varying in charge distribution. *Photochem. Photobiol.* **2003**, *78*, 431–435.

151. Banfi, S.; Caruso, E.; Buccafurni, L.; Battini, V.; Zazzaron, S.; Barbieri, P.; Orlandi, V. Antibacterial activity of tetraaryl-porphyrin photosensitizers: An *in vivo* study on Gram negative and Gram positive bacteria. *J. Photochem. Photobiol. B Biol.* **2006**, *85*, 28–38.

152. Konopka, K.; Goslinski, T. Photodynamic therapy in dentistry. *Crit. Rev. Oral Biol. Med.* **2007**, *8*, 694–707.

153. Robertson, C.A.; Evans, D.H.; Abrahamse, H. Photodynamic therapy (PDT): A short review on cellular mechanisms and cancer research applications for PDT. *J. Photochem. Photobiol. B Biol.* **2009**, *96*, 1–8.

154. Brancaleon, L.; Moseley, H. Laser and non-laser light sources for photodynamic therapy. *Laser Med. Sci.* **2002**, *17*, 173–186.

155. Veenhuizen, R.B.; Stewart, F.A. The importance of fluence rate in photodynamic therapy: Is there a parallel with ionizing radiation dose-rate effects? *Radiother. Oncol.* **1995**, *37*, 131–135.

156. Allison, R.R.; Mota, H.C.; Sibata, C.H. Clinical PD/PDT in North America: an historical review. *Photodiagn. Photodyn.* **2004**, *1*, 263–277.

157. Juzeniene, A.; Juzena, P.; Ma, L-W.; Iani, V.; Moan, J. Effectiveness of different light sources for 5-aminolevulinic acid photodynamic therapy. *Laser Med. Sci.* **2004**, *19*, 139–149.

158. Kübler, A.C. Photodynamic therapy. *Med. Laser Appl.* **2005**, *20*, 37–45.

159. Lukšiene, Z. New approach to inactivation of harmful and pathogenic microorganisms by photosensitization. *Food Tech. Biotechnol.* **2005**, *43*, 411–418.

160. Qin, Y.; Luan, X.; Bi, L.; He, G.; Bai, X.; Zhou, C.; Zhang, Z. Toluidine blue-mediated photoinactivation of periodontal pathogens from supragingival plaques. *Laser Med. Sci.* **2008**, *23*, 49–54.

161. Schindl, A.; Rosado-Sholosser,B.; Trautinger, F. Reciprocity regulation in photobiology: An overview (in German). *Hautarzt* **2001**, *52*, 779–785.

162. Maclean, M.; MacGregor, S.J.; Anderson, J.G.; Woolsey, G.A. The role of oxygen in the visible-light inactivation of *Staphylococcus aureus*. *J. Photochem. Photobiol. B Biol.* **2008**, *92*, 180–184.

163. Langmack, K.; Mehta, R.; Twyman, P.; Norris, P. Topical photodynamic therapy at low fluence rates—theory and practice. *J. Photochem. Photobiol. B Biol.* **2001**, *60*, 37–43.

164. Wagner, S.J.; Storry, J.R.; Mallory, D.A.; Stromberg, R.R.; Benade, L.E.; Friedman, L.I. Red cell alterations associated with virucidal methylene blue phototreatment. *Transfusion* **1993**, *33*, 30–36.

165. Wagner, S.J. Virus inactivation in blood components by photoactive phenothiazine dyes. *Transfus. Med. Rev.* **2002**, *16*, 61–66.

Review

Function and Regulation of Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/CRISPR Associated (Cas) Systems

Corinna Richter, James T. Chang and Peter C. Fineran *

Department of Microbiology and Immunology, University of Otago, PO Box 56, Dunedin 9054, New Zealand; E-Mails: ricco896@student.otago.ac.nz (C.R.); james.chang@otago.ac.nz (J.T.C.)

* Author to whom correspondence should be addressed; E-Mail: peter.fineran@otago.ac.nz; Tel.: +64-3-479-7735; Fax: +64-3-479-8540.

Abstract: Phages are the most abundant biological entities on earth and pose a constant challenge to their bacterial hosts. Thus, bacteria have evolved numerous 'innate' mechanisms of defense against phage, such as abortive infection or restriction/modification systems. In contrast, the clustered regularly interspaced short palindromic repeats (CRISPR) systems provide acquired, yet heritable, sequence-specific 'adaptive' immunity against phage and other horizontally-acquired elements, such as plasmids. Resistance is acquired following viral infection or plasmid uptake when a short sequence of the foreign genome is added to the CRISPR array. CRISPRs are then transcribed and processed, generally by CRISPR associated (Cas) proteins, into short interfering RNAs (crRNAs), which form part of a ribonucleoprotein complex. This complex guides the crRNA to the complementary invading nucleic acid and targets this for degradation. Recently, there have been rapid advances in our understanding of CRISPR/Cas systems. In this review, we will present the current model(s) of the molecular events involved in both the acquisition of immunity and interference stages and will also address recent progress in our knowledge of the regulation of CRISPR/Cas systems.

---

1. Introduction

Bacteria and Archaea are frequently exposed to stresses such as infection from bacteriophages (phage) and other genetic elements. These events can result in horizontal gene transfer (HGT), which is mediated by transduction, transformation or by conjugation of mobile elements such as plasmids [1,2]. HGT can be beneficial for adaptation and survival by rapidly promoting the transfer of genes encoding antibiotic resistance, virulence factors, or the ability to degrade toxic compounds [1]. However, in the absence of any selective advantage, foreign genetic elements can have a fitness cost. This is most evident when prokaryotes are infected by virulent phages. The sheer scale of viral abundance ($\sim 10^{31}$), diversity and rates of infection ($\sim 10^{25}$/s) [3,4] has resulted in the evolution of prokaryotic defenses to maintain the balance in this arms race [5].

Bacteria possess multiple methods to regulate genetic flux and resist phage infection. These mechanisms include the mutation or masking of cell surface receptors, restriction-modification, abortive infection and the **c**lustered **r**egularly **i**nterspaced **s**hort **p**alindromic **r**epeats (CRISPR) systems [5,6]. CRISPR systems are a widespread mechanism that equips bacteria with a sequence-specific heritable 'adaptive immune system' that has a genetic memory of past genetic incursions (for recent reviews see [7–9]).

CRISPRs use small non-coding RNAs for defense and function in conjunction with **C**RISPR associated (Cas) proteins. The mechanism of CRISPR/Cas interference involves three phases (Figure 1). Firstly, resistance is acquired via the integration of short sequences from foreign genetic elements (termed spacers) into repetitive genetic elements known as CRISPR arrays. Secondly, CRISPR arrays are then transcribed and processed into small RNAs (crRNAs) by Cas proteins. In the third and final step, targeting of the invading phage or plasmid is mediated by a Cas protein complex that contains crRNAs. During this stage, the crRNA-Cas protein complex then interferes, in a sequence-specific manner, with the foreign nucleic acids.

Despite sharing mechanistic similarities, there is significant diversity amongst CRISPR/Cas systems. There are three major CRISPR/Cas system types (I-III), which are characterized by a signature protein [10]. The main types are further divided into subtypes (e.g., I-A to I-F) based mainly on the presence of a subtype-specific set of Cas proteins and for some types the repeat sequence of the associated CRISPR array is also taken into account. The last five years have seen rapid advances in our mechanistic/molecular understanding of CRISPR/Cas systems. In this review, we provide an historical context, highlight the current models for the adaptation, processing and interference stages and finally, discuss the regulation of the CRISPR/Cas systems.

2. Discovery of CRISPR/Cas Systems

The first report of interspaced palindromic repeat sequences was made by Ishino *et al.* who detected five 29 bp repeats with 32 bp spacers near the *iap* gene of *Escherichia coli* [11]. Further reports were made for *Mycobacterium tuberculosis* [12], *Haloferax* spp. [13] and *Archaeoglobus fulgidus* [14,15]; however, they were not detected in eukaryotic or virus sequences [15]. Mojica and co-workers performed a comparative *in silico* study of those repetitive elements to determine structure and sequence similarity, as well as their phylogenetic distribution [16] showing that CRISPRs display a high degree of homology between phylogenetically distant species and a wide distribution in bacteria and archaea. To date CRISPR/Cas systems have been found in almost 50% of bacterial and 85% of archaeal genome sequences available [17].

Figure 1. Overview of clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR associated (Cas) adaptive immunity. (**A**) Adaptation. The CRISPR arrays are composed of short repeats and intervening sequences derived from foreign invaders. Upon infection with a foreign element (e.g., phages or plasmids), part of the genome is typically incorporated into the leader end of the CRISPR array and the repeat is duplicated. The CRISPR arrays are located adjacent to a cluster of *cas* genes; (**B**) crRNA generation. The CRISPRs are transcribed into pre-crRNAs that are then processed into mature crRNAs; (**C**) Interference. The crRNA, in a complex with Cas proteins, binds and degrades the target nucleic acid of the invading element.
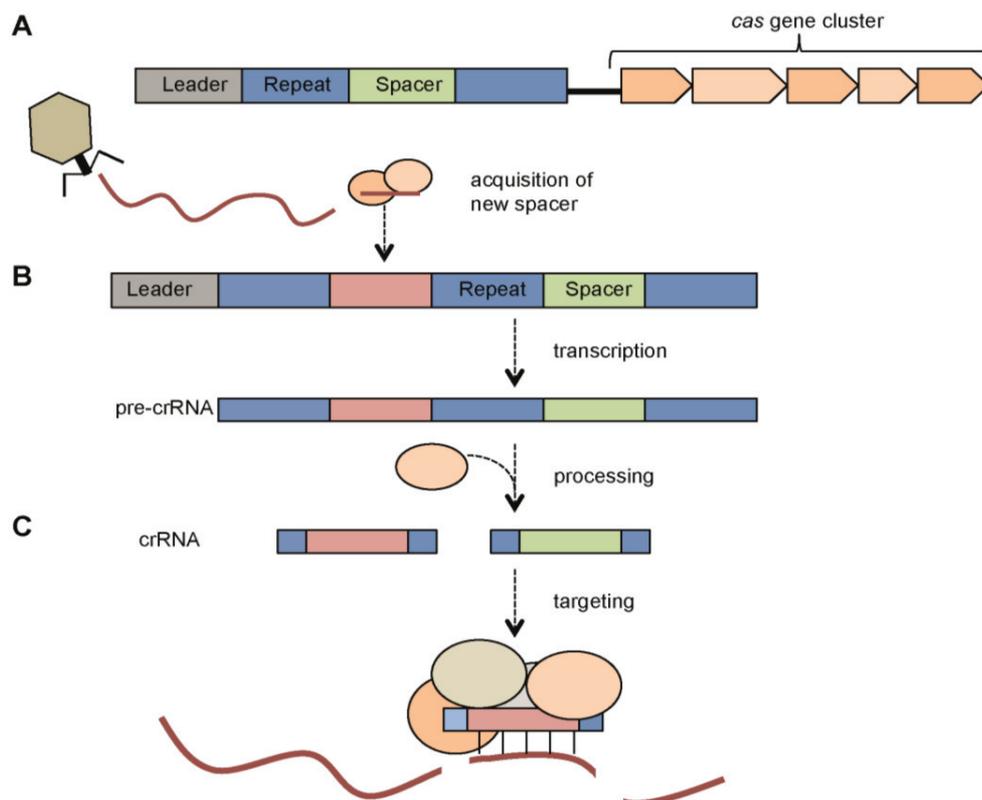
Terminology for these repeated elements was inconsistent and included interspaced short sequence repeats (SSR) [18], spacer interspersed direct repeats (SPIDR) [15] or short regularly spaced repeats (SRSRs) [16]. For simplicity, the name CRISPR (clustered regularly interspaced short palindromic repeats) was accepted since it reflects the most important characteristics [19].

An important step towards the understanding of the role of CRISPRs was the identification of four genes closely located to the CRISPR arrays [19], which were termed **C**RISPR-**as**sociated (*cas*) genes 1-4. Cas1-4 were only present in genomes containing CRISPR arrays and their predicted functions included helicase and nuclease activities, which led to the hypothesis that CRISPR/Cas systems might be involved in DNA repair [20].

A critical discovery came in 2005 with three independent studies showing that spacers matched sequences of extrachromosomal origin, including phages, prophages and plasmids [21–23]. Furthermore, there was a positive correlation between the presence of spacers matching a particular phage and phage resistance [22,23]. In combination with detailed bioinformatics analyses of the Cas proteins, it was hypothesized that CRISPR/Cas systems provide an RNAi-like mechanism of resistance against invading genetic elements [21–25].

The first direct evidence that CRISPR/Cas could protect bacteria from phages or plasmids was provided by two key studies. Firstly, Barrangou *et al*. obtained phage resistant mutants following phage challenge of *Streptococcus thermophilus*. The phage resistant strains had incorporated new spacers matching to the viral genome while removal or addition of spacers matching to the phages resulted in phage sensitivity or resistance, respectively [26]. Subsequently, Marraffini and Sontheimer showed that CRISPR/Cas systems can also prevent both conjugation and transformation of plasmids in *Staphylococcus epidermidis* [27]. Recently, it was shown that the *Streptococcus pneumoniae* CRISPR/Cas can prevent natural transformation. Non-capsulated *S. pneumoniae* with CRISPRs programmed to target capsule genes were unable to be transformed into capsulated strains during infection in mice [28].

3. Components of CRISPR/Cas Systems

CRISPR/Cas systems comprise the following critical elements: a CRISPR array, an upstream leader sequence and the *cas* genes (Figure 1A) that are discussed below [19,29].

3.1. CRISPR Arrays: Repeats, Spacers and Leader Sequence

In the CRISPR array, repeats alternate with spacer sequences [11,30–32]. Within arrays, the repeats are typically identical in terms of length and sequence, but small differences do occur [16,29]. In particular, the repeat at the end of an array is often truncated or deviates more from the consensus sequence [17,33]. Repeats are usually between 23 and 47 bp in length and those of the same subtype share a consensus sequence in type I-C, I-E, I-F, and II [10,34,35]. Many repeat sequences show palindromes or short inverted repeats [16] and a predicted stable secondary structure in the form of a stem-loop [34]. However, there are repeats, which are not palindromic and are predicted to be

unstructured. This difference in the ability to form secondary structures also has implications for the mechanism of pre-crRNA processing, which will be discussed later [10,34].

In contrast, the spacer sequences are mostly unique within a genome. Many spacers have been matched to sequences originating from extra-chromosomal sources such as phage or plasmids and other transferable elements [21–23,36]. Indeed, it is the spacers that confer the sequence-specific immunity against those extra-chromosomal agents [26,33,37]. The sequences in the foreign genome from which spacers are derived are termed protospacers. It is likely that the majority of spacers are phage- or plasmid-derived, but the limited depth of sequence data on these abundant and diverse elements leads to an underestimation. Interestingly, some CRISPR/Cas systems also possess spacers that match sequences elsewhere within their own genome. Their function, if any, is not resolved, but it has been proposed that they have been accidentally incorporated [38]. Spacer length can vary slightly throughout an array (typically by 1–2 nt) and spacer lengths up to 72 bp have been reported, but usually the size is similar to that of the repeats in the same array [17].

The third component of the CRISPR array is the leader sequence, which is located upstream of the first repeat [15]. This AT-rich sequence is about 200–500 bp long and includes the promoter necessary for transcription of the array (see below) [39–42]. The leader region is also important for the acquisition of new spacers [43].

A single genome can harbor more than one CRISPR array. Those can vary considerably in size with the largest identified to date, in *Haliangium ochraceum* DSM 14365, containing 587 repeats [35]. The length of CRISPR arrays appears to correlate with the degree of activity in spacer acquisition, with longer arrays being more active than short arrays and those with degenerate repeat sequences [33].

3.2. Cas Proteins

In close proximity to the CRISPR array are genes that encode the **C**RISPR **as**sociated (Cas) proteins (Figure 1A) [19,20,24,25,44]. Cas proteins provide the enzymatic machinery required for the acquisition of new spacers from, and targeting, invading elements.

CRISPR/Cas systems are currently classified into type I, II and III, based on the phylogeny and presence of particular Cas proteins [10]. There is further division within each type into subtypes (e.g., type I is composed of type I-A to I-F). The Cas proteins are important for the differentiation between both the major CRISPR/Cas types and the subtypes. Two main groups of Cas proteins can be distinguished. The first group, which includes the core Cas proteins, is found across multiple types or subtypes. Cas1 and Cas2 (Cas2 is sometimes fused to another Cas protein) are found in every Cas operon across the three main types, and Cas1 is considered the universal marker of CRISPR/Cas systems [10,24,25,44]. Cas3, Cas9, and Cas10 are each specific for one major type, serving as signature proteins for type I, II, and III, respectively [10].

The second major group consists of proteins only found within the gene clusters of one particular subtype [10,25]. For example, the Cse and Csy proteins are subtype-specific for type I-E and type I-F systems, respectively. The subtype specific proteins of some systems form a complex involved in targeting and interference, which in the type I-E system is referred to as Cascade (CRISPR-associated

complex for antiviral defense) [45–49]. These complexes will be discussed in detail below. The systems can be complex since some bacteria contain multiple CRISPR/Cas subtypes, each of which can have multiple CRISPR arrays that function with the appropriate Cas cluster [18]. In summary, the *cas* cluster for each CRISPR/Cas subtype contains genes encoding Cas1 and Cas2, the signature protein defining the major type and a set of subtype specific proteins. Interestingly, the total number of proteins differs between the types of system, even though the functional principles are similar in the cases studied thus far. It is possible that unexpected functions may reside within Cas proteins that are yet to be characterized.

4. CRISPR/Cas Mechanism

As highlighted in Figure 1, the mechanism of CRISPR/Cas defense involves three stages. Resistance must first be acquired by integrating spacers into the CRISPR arrays. These arrays are then transcribed and processed into short crRNAs. Finally, a crRNA-Cas ribonucleoprotein complex targets the invading nucleic acid for degradation. In the following sections we outline the processes of spacer acquisition, crRNA biogenesis and interference.

4.1. Adaptation via Spacer Acquisition

The least characterised stage of CRISPR/Cas defense is adaptation, whereby new spacers are acquired from phages and plasmids and, with duplication of the repeat, are added to the leader proximal end of the CRISPR array (Figure 2). A number of metagenomic studies of phage-bacterial dynamics in environmental [50,51] and human [52] niches have provided clear evidence that CRISPR/Cas systems actively acquire new spacers from mobile genetic elements over short time-scales. The first demonstration of spacer incorporation from phages and plasmids in the laboratory was in the type II-A system of *S. thermophilus* [26,53]. Very recently, studies have observed spacer acquisition in the *E. coli* type I-E [43,54,55], *P. aeruginosa* type I-F [56], *Streptococcus agalactiae* type II-A [57] and *Sulfolobus solfataricus* type I and III-B [58] systems. The most well characterized system is the *E. coli* type I-E, so this will be the focus of the following discussion.

4.1.1. Cas Proteins Required for Adaptation

In the *E. coli* type I-E system, Cas1 and Cas2 are required for spacer incorporation from plasmids and phages [43,55]. These recent results are consistent with the dispensable role of Cas1 and Cas2 in crRNA biogenesis / targeting in type I-E [45], type II-A [59] and type III-A [60] systems. The type I-E data is of broad relevance because Cas1 and Cas2 are conserved across all CRISPR/Cas types [10,24]. It is interesting that an insertion mutation of *csn2* inhibited spacer acquisition in the *S. thermophilus* type II-A system, suggesting that, in addition to Cas1 and Cas2, other Cas proteins might play a role [26].

Figure 2. Model of CRISPR/Cas adaptation (based on type I-E systems). (**A**) Naïve acquisition. Cas1 and Cas2 are required for acquisition of new spacers. The first repeat at the leader end of the CRISPR array is duplicated and incorporates a new spacer sequence from a protospacer. The final nucleotide of the repeat is not duplicated but is provided by the PAM nucleotide immediately adjacent to the protospacer; (**B**) Priming acquisition. Expression of the CRISPR array and generation of the crRNAs against the foreign DNA results in binding/targeting, which is hypothesized to aid in the generation of precursors for integration (possibly single- or double-stranded). Productive interference is not required for priming. Priming acquisition requires Cascade-crRNA, Cas3 and Cas1 and Cas2 and results in new spacers derived from the same strand as the initial spacer.
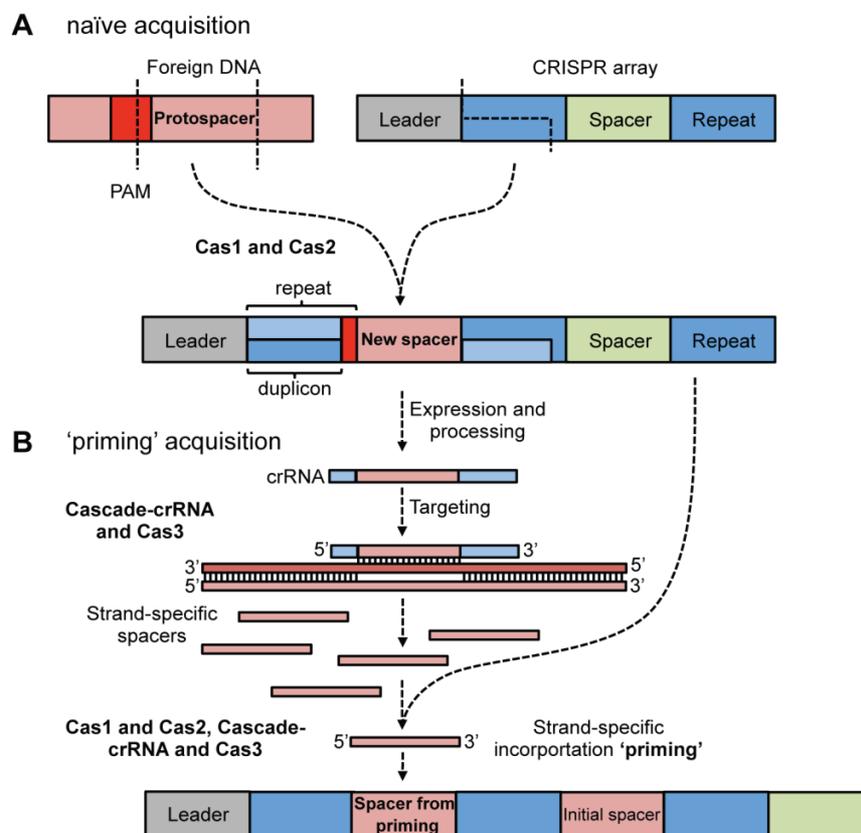
Exactly how Cas1 and Cas2 function is unclear, but biochemical and structural studies offer some clues [61–63]. Cas1 is a metal-dependent endonuclease and the *Pseudomonas* Cas1 cleaves dsDNA, generating ~80 bp fragments [61]. The 80 bp dsDNA fragments generated from Cas1 are larger than expected for integration (~32–33 nt), indicating that other factors are involved. Cas1 dimers contain a stirrup-like structure and a positively charged surface that together might be involved in binding dsDNA [9]. Cas2 from *S. solfataricus* cleaves ssRNA at U-rich regions *in vitro* [64], but ssRNA or ssDNA binding or cleavage was not detected for *Desulfovibrio vulgaris* Cas2 [65]. Recently, the *Bacillus halodurans* Cas2 dimer was shown to contain Mg^{2+}-dependent endonuclease activity against

dsDNA and generated 120 bp products [66]. Despite these studies the exact mechanistic roles of Cas1 and Cas2 in acquisition are unknown.

There is increasing evidence that some Cas1 and Cas2 proteins may function together directly (e.g., via protein-protein interactions). For example, the type I-A system from *Thermoproteus tenax* contains a Cas1-Cas2 fusion protein, which interacts with Csa4 and Csa1, in what the authors name Cascis (CRISPR associated complex for the integration of spacers) [67]. It is noteworthy that the type I-F Cas3 is a Cas2-Cas3 hybrid, containing an N-terminal Cas2-like domain fused to Cas3 [10,24].

4.1.2. Protospacer Selection and Incorporation into CRISPR Arrays

The minimal CRISPR requirements for spacer acquisition in the type I-E system are a single 'repeat' and 60 bp of upstream sequence [43]. As the CRISPR promoter is not contained within this 60 bp, this suggests that CRISPR transcription is not required for incorporation, but indicates this region might be recognized by Cas1 and/or Cas2 to enable the directionality of integration at the leader end of the CRISPR array.

In type I-E arrays with multiple repeats, the leader proximal repeat is the one that is duplicated (Figure 2A) [43]. In addition, the protospacer added to the CRISPR array contains the last nt of the PAM motif, which becomes the final nt of the 5' repeat (Figure 2A) [54,55,68]. Therefore, only the first 28 nt of the repeat are duplicated and thus, this unit has been called the 'duplicon' [68] and might provide fidelity to the direction of incorporation [69]. This duplicon mechanism cannot apply to all CRISPR/Cas types since the last nucleotide of some PAMs (e.g., CCN in *S. solfataricus*) is not conserved, so could not contribute to the conserved repeats [58]. Interestingly, internal spacer acquisition was recently observed for one CRISPR array in *S. solfataricus* [58], which challenges the dogma that spacer acquisition is always leader proximal [69].

One critical question is how do the CRISPR/Cas systems acquire foreign DNA and not DNA from their own chromosome? In *S. thermophilus*, examination of the protospacers in the phage genomes enabled the identification of a short 3' protospacer adjacent motif (PAM) [37]. PAM sequences have since been identified via bioinformatics for other CRISPR/Cas systems [70]. In *E. coli*, spacer selection was shown to require the presence of a PAM [43], but the short length of PAMs cannot provide the level of discrimination to avoid sampling chromosomal DNA. A self / non-self mechanism does occur, since the incorporation of spacers from the chromosome is rare [43]. It is possible that DNA modification, such as by a restriction-modification-like mechanism, might play a role.

4.1.3. Priming Acquisition

In the type I-E system there is now evidence that acquisition is divided into two stages; (1) naïve- and (2) priming-based acquisition [54,55]. As discussed above naïve acquisition is the initial spacer incorporation from a new foreign element. During priming, the initial spacer 'primes' the array for the addition of multiple invader-derived spacers from the same DNA strand as the initial spacer (Figure 2B) [54,55]. Cas1, Cas2, Cascade, Cas3 and a crRNA that matches the invading element are

required for priming [55]. The nuclease activity of Cas3 was proposed to be required for the feedback loop, generating precursors for incorporation following the initial targeting event [54]. However, spacers that match the foreign DNA but do not support targeting, due to single nucleotide mutations, still promote priming [55]. Priming is proposed to allow adaptation to phage and plasmids that mutate and escape the initial spacers [55] and multiple spacers would increase resistance and decrease the frequency of escape.

In summary, details of CRISPR adaptation are beginning to come to light, but there are still many questions regarding the exact mechanisms of protospacer recognition, spacer generation, repeat duplication, integration of the new spacer and the roles of the Cas proteins.

4.2. Expression, crRNA Generation and Interference

Following the acquisition of spacers, the CRISPR array is expressed as a long pre-crRNA that is processed into crRNAs. The mature crRNAs form part of a ribonucleoprotein complex, which targets and degrades the foreign genetic material. As will be outlined below, these processes are similar, yet differ between type I, II, and III systems (Figure 3).

4.2.1. Type I Systems

In type I CRISPR/Cas systems, regions within each repeat form a stem-loop, which is bound by the processing endoribonuclease. The endoribonucleases in type I systems belong to the group of Cas6 proteins. Characterized members of this group include Cas6b in type I-B [71], Cas5d in type I-C [49], Cas6e (CasE, Cse3) in type I-E [45,72,73] and Cas6f (Csy4) in type I-F [42,74]. The pre-crRNA is cleaved in a sequence-specific manner at the downstream base of the stem-loop, producing the mature crRNA consisting of a short 5' repeat handle followed by the spacer sequence and the stem-loop of the next repeat (Figure 3B) [49,71,74,75]. The crRNA and the endoribonuclease stay associated, and might serve as nucleation point [48,49,76] for the formation of Cascade (Figure 3C). Cascade is a ribonucleoprotein complex formed by the subtype-specific Cas proteins and the crRNA, initially identified in type I-E, but later also found in type I-C and I-F [9,45,46,48,49].

Figure 3. crRNA generation and target interference in type I, II and III CRISPR/Cas systems. (**A**) Transcription of the CRISPR array into a pre-crRNA; (**B**) Processing of the pre-crRNA into mature short crRNAs. In type I, RNA cleavage is performed by Cas6-homologues, which bind the repeat stem-loop and stay associated for Cascade formation. In type II, tracrRNA is required for binding and processing of the pre-crRNA by Cas9 and RNaseIII. In type III, Cas6 binds to non-structured repeats and processes the pre-crRNA into crRNA and then dissociates; (**C**) Target binding and cleavage. Type I Cascade binds the DNA target before recruiting Cas3 for degradation. In type II, Cas9 stays associated with the tracrRNA:crRNA complex after processing and subsequently binds and cleaves target DNA. The type III-B CMR-complex binds spacer sequence and targets RNA. It is hypothesized that a type III-A Csm complex forms and this system targets DNA.
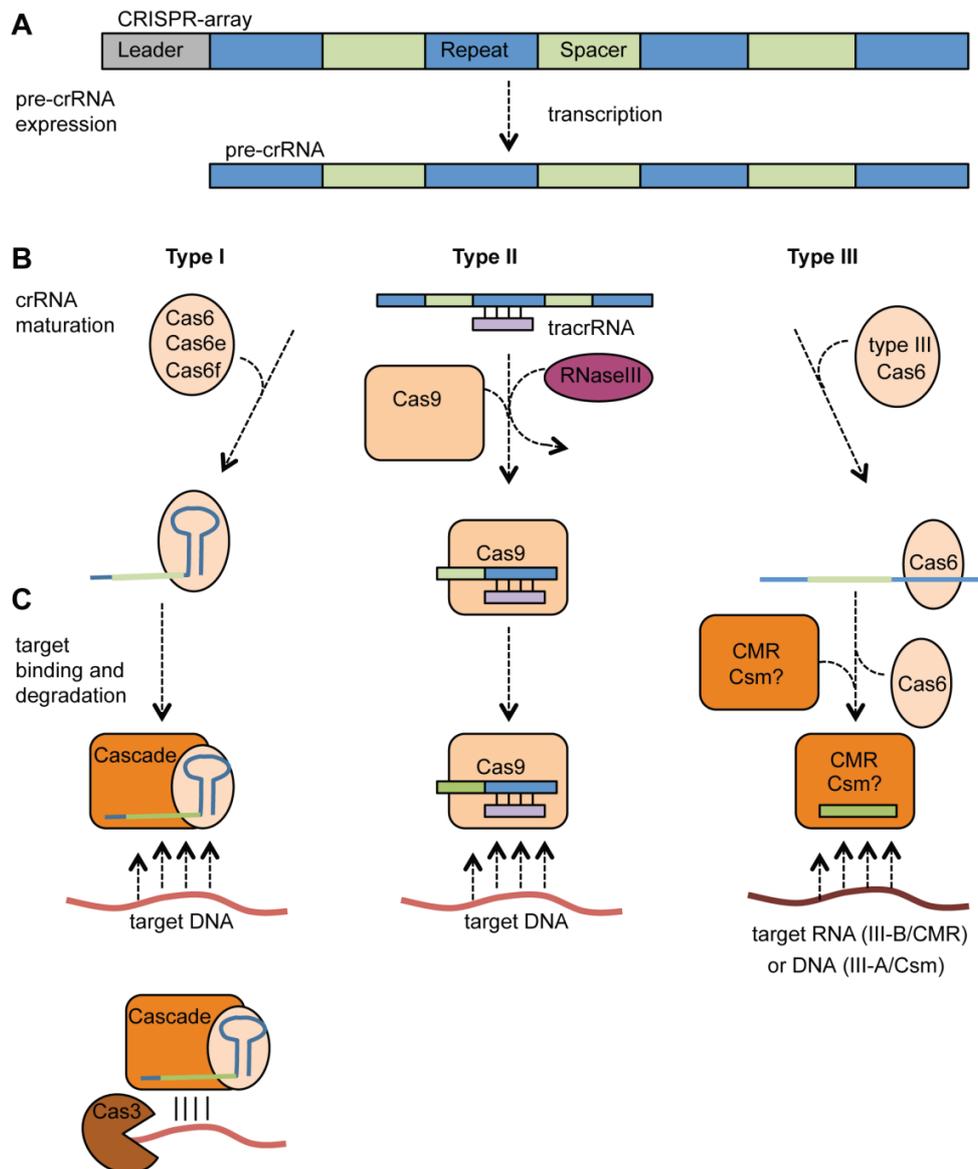
The type I-E Cascade has a shape referred to as seahorse-like and has a hexameric backbone formed by Cas7 (CasC, Cse4). The stem-loop of the crRNA is bound to the endoribonuclease Cas6e, which forms the "head" of the seahorse. Cse1 (CasA) forms the "tail" and binds the 3' end of the crRNA at the other end. Cse1 is bridged to Cas6e by a Cse2 (CasB) dimer, and linked by Cas5 (CasD) to the sixth Cas7 protein of the backbone [46,47]. The I-F and I-C complexes share common features with the I-E Cascade in terms of their shape, but also show variation, which is mainly due to different proteins involved. Different numbers of subtype-specific proteins constitute Cascade, three in the type I-C, four in I-F systems, while five distinct proteins are present in the type I-E Cascade [10]. Additional variation occurs at the level of the stoichiometry of individual protein subunits in different subtype Cascade complexes. When three or four proteins are used, the endoribonuclease, with the crRNA bound via the stem-loop, is located at one end of the complex. A homo-multimer of six proteins (Csy3 and Csd2 for I-F, and I-C, respectively) forms the neck or backbone of the complex, which holds the spacer and ensures it is extended for optimal presentation to invading nucleic acids. The end of the complex is formed by a hetero-dimer [48,49]. A type I-A complex from *S. solfataricus* (termed aCASCADE by the authors) has also been partially characterized and contained the backbone protein, Cas7 (Csa2), Cas5a, Csa5, Cas6 and crRNAs [77]. Although not directly shown, based on the data from other type I systems, the type I-A Cascade is likely to consist of the Cas6 endoribonuclease bound to the crRNA, a backbone of six Cas7 proteins and a Cas5a/Csa5 heterodimer.

During the final step of interference, the type I Cascade containing a specific crRNA recognizes and binds the appropriate target DNA (Figure 3C). In type I-E systems, a short loop within CasA recognizes potential targets via the PAM sequence and it is proposed to bind them before the target DNA is incorporated into Cascade [78]. In addition to the PAM requirement in the target, a non-contiguous 7 nt sequence (nucleotides in position 1–5 and 7–8) within the first 8 nt of the spacer, relative to the 5′ end of the spacer within the crRNA, are essential for initial target recognition and hence are termed the 'seed sequence' [48,79]. Cascade with bound crRNA specifically binds to negatively supercoiled target DNA [80]. The negative supercoiled topology serves as energy source for formation of an R-loop. Furthermore, Cascade induces DNA bending before Cas3 is recruited for subsequent DNA degradation (Figure 3C) [80]. Cas3 is the signature protein of type I systems and possesses ATP-dependent helicase activity as well as the ability to cleave ssDNA via an HD nuclease domain [81–83]. Even though the exact mechanism is not yet fully understood, it is hypothesized that once recruited, Cas3 further unwinds and cleaves the target DNA in a 5' to 3' direction, while Cascade dissociates from the target and is recycled [80,83].

4.2.2. Type II Systems

A very different mechanism of CRISPR expression and crRNA maturation was discovered in the type II-A system of *Streptococcus pyogenes* [84]. An abundant short *trans*-encoded transcript is essential for processing of pre-crRNA into mature crRNA. Therefore, these RNAs are termed *trans*-activating crRNAs (tracrRNA). These tracrRNAs contain a 25 bp stretch with almost perfect complementarity to the type II CRISPR repeats within the pre-crRNA (Figure 3B). The type II repeats

do not form stem-loops and the authors suggest this deficiency is overcome by pairing with the tracrRNA [84].

In addition to the tracrRNA, Cas9 (Csn1) and host RNase III are needed for the processing of pre-crRNA into mature crRNA. The requirement of RNase III is the first discovery of a non-Cas protein being an essential part of CRISPR/Cas defense machinery (Figure 3B) [84]. Interestingly, the *S. thermophilus* type II-A system provides phage resistance in *E. coli* when recombinantly expressed [59], which suggests diverse RNase III enzymes are compatible for crRNA generation. Cas9 is not only involved in crRNA processing but, along with the tracrRNA:crRNA hybrid, recognizes and degrades the target (Figure 3C) [85]. In complex with the tracrRNA:crRNA structure, Cas9 binds the target dsDNA and creates a double strand break by cleaving the complementary and non-complementary strand with its HNH- and RuvC-like domains, respectively [85]. *In vivo*, blunt cleavage of phage and plasmid dsDNA was observed at a position 3 bp from the PAM within the protospacer [53]. The PAM is important for the affinity towards the target and may have implications on the unwinding of the duplex, as cleavage of targets with mutated PAMs is only impaired in dsDNA but not ssDNA. Indeed, phages with PAM mutations also have the ability to avoid targeting by type II systems [37]. Furthermore, similar to the seed sequence in type I systems, complementarity between crRNA and target over a 13 bp stretch proximal to the PAM is required for interference [85] and hence, phages with mutations in this region of the protospacer region can evade interference [37].

4.2.3. Type III Systems

In type III CRISPR/Cas systems, expression and interference are predominantly similar to those described for type I systems. However, the significant difference is that the endoribonuclease Cas6 does not bind to a repeat stem-loop but to the first bases of the repeat and that further crRNA maturation steps occur (Figure 3B) [86,87]. Indeed, not all repeats in those systems contain a strong palindrome and those associated with type III-A are predicted to be mostly non-structured [34]. In *Pyrococcus furiosus*, the binding site of Cas6 is located within nt 2–8 of the repeat and cleaves at a distal site between bases 22 and 23 resulting in crRNAs with an 8 nt repeat handle on the 5' end, the spacer sequence and 22 nt of the following repeat at the 3' end [86–88]. Interestingly, this RNA is not the final mature crRNA product. Cas6, unlike the type I endoribonucleases, only transfers the crRNA to the targeting complex termed CMR for type III-B where they are further processed (Figure 3C) [10,89,90]. In the type III-A system of *S. epidermidis*, a similar mechanism for processing by Cas6 was observed and the *csm2*, *csm3* and *csm5* genes were required for further 3' processing and crRNA maturation [60].

The *P. furiosus* CMR complex includes the full subset of Cmr proteins (Cmr1–6) and can contain two different species of crRNA [90], both with the 8 nt 5' end derived from the repeat and either 31 nt or 37 nt of the spacer [91]. The type III-B CMR complex from *S. solfataricus* contains an additional protein, Cmr7 and a stoichiometry of one protein of each of Cmr1–6 and 6 Cmr7 proteins [92]. Interestingly, the *P. furiosus* CMR complex has been shown to target RNA [89,90,92]. Basepairing between the last 14 nt of the crRNA and the RNA target seem to be essential for cleavage [90]. Target RNA cleavage occurs 14 nt upstream from the 3' end of the two mature crRNA species, generating two

cleavage sites [90]. In contrast to type III-B, DNA is the target of type III-A systems [27], but there is not yet data about complex formation for the Csm proteins.

The ability of CRISPR/Cas systems to target DNA raised the question of how they avoid targeting their own CRISPR arrays, which have perfect complementarity to the crRNAs they produce. For the type III-A system, this is avoided by requiring spacer:protospacer complementarity and an absence of base-pairing to the 5' handle in the crRNA [93]. This ensures that only non-self targets are licensed for degradation, but whether the same principle applies to other types remains unknown.

5. Regulation of CRISPR/Cas Systems

As CRISPR/Cas systems function as a defense mechanism their expression might be expected to respond to invasion by extrachromosomal elements. This was recently proven in a shotgun proteomics approach in *S. thermophilus*, which showed an increase in Cas protein expression following phage infection [94]. Likewise, CRISPR and Cas expression were increased by phage infection in *Thermus thermophilus* [39]. Furthermore, CRISPR/Cas systems can be modulated in response to UV and play a role in sensitivity to DNA damage, hinting towards other possible roles besides the neutralization of invading elements [63,67]. However, to date little is known about how CRISPR/Cas systems are regulated in response to those external stimuli or during periods in which the system is not required [63,67]. Under certain conditions it could be viewed as favorable to down-regulate CRISPR/Cas activity when beneficial elements are available. However, most evidence demonstrates up-regulation of CRISPR/Cas upon exposure to phage. Since CRISPR/Cas systems are not 100% effective, some beneficial elements will be acquired and when they provide a selective advantage they are likely to be maintained. An overview of the regulatory inputs is shown in Figure 4.

Figure 4. Regulation of CRISPR/Cas activity. Network model summarizing the regulation of CRISPR/Cas activity. Triangular and flat arrowheads indicate positive and negative effects on CRISPR/Cas activity, respectively for *E. coli* (red), *Salmonella* (cyan), *M. xanthus* (green), *Sulfolobus* (black), *T. thermophilus* (purple) and *S. thermophilus* (blue). For details see text.
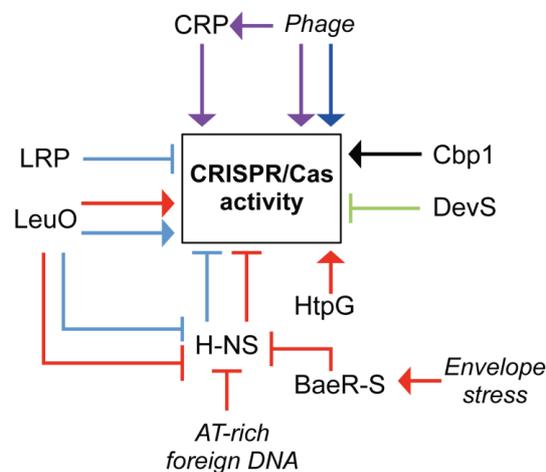
Most information about CRISPR/Cas regulation is available for the type I-E systems in *E. coli* and *Salmonella enterica* serovar Typhi backgrounds. The Histone-like nucleoid structuring protein (H-NS) is a global regulator involved in compacting the bacterial chromosome, via high affinity binding to AT-rich, curved DNA (reviewed by Dorman *et al.* [95]). Many promoters are located in close proximity to curved DNA. Binding of H-NS to these promoters prevents RNA polymerase (RNAP) binding, which results in gene-silencing [95]. In *E. coli*, H-NS negatively regulates *cas* gene expression, due to H-NS binding sites near the *cas* operon promoters [41,96]. It is hypothesized that H-NS will bind to invading nucleic acids when they enter the cell, due to the higher AT-content [97,98]. Sequestration of H-NS is predicted to free the *cas* and/or LeuO promoter (see below) for RNAP recognition and thus activate expression, allowing effective CRISPR/Cas-mediated defense [41,96,99].

LRP (lysine-responsive regulatory protein) is a negative regulator for *cas* expression in *S. typhi* [100]. LRP functions in a similar fashion to H-NS by binding to the *cas* promoter and competitively excluding binding of RNAP. This appears to be independent of H-NS. However, LRP is able to bind simultaneously alongside H-NS, indicating that the two proteins interact to generate a nucleosome structure for *cas* gene repression [100]. Interestingly, LRP does not regulate the *cas* operon in *E. coli*, suggesting that the type I-E CRISPR/Cas systems in these two closely-related organisms function differently at the regulatory level [41,100].

LeuO, a LysR-type transcriptional regulator that also responds to amino acid starvation, affects *cas* gene expression via competition with H-NS in *S.* Typhi and *E. coli* [96,100,101]. LeuO binds to a region flanking the *cas* promoter and the H-NS binding site. The binding of LeuO to the DNA competes with H-NS binding, and relieves the inhibition of *cas* expression by enabling promoter recognition by RNAP [96,100,101]. During amino acid starvation expression of LeuO is increased in response to accumulation of the small molecules guanosine 3'-diphosphate 5'-triphosphate and guanosine 3',5'-bis(diphosphate), collectively called (p)ppGpp [102]. Interestingly, accumulation of (p)ppGpp does not occur during phage lambda infection in *E. coli* [103]. In theory, it is possible that infection with other phages triggers amino acid starvation, leading to LeuO-dependent activation of the CRISPR/Cas system.

The BaeR-S two-component regulatory system is activated in response to envelope stress (reviewed by MacRitchie *et al.* [104]). Phage infection potentially causes envelope stress by accumulation of viral proteins in the membrane. Upon detection of membrane stress, the histidine sensor kinase BaeS is activated by phosphorylation and in turn activates the BaeR protein. Activated BaeR gains the ability to bind DNA, modulating gene expression [104]. In *E. coli*, the BaeR binding site is located at the H-NS binding site near the *cse1* (*casA*) promoter. Therefore, bound BaeR serves as an antagonist for H-NS binding, predicted to result in the release of the *cas* promoters for RNA polymerase recognition and *cas* expression [105,106].

The high-temperature protein G (HtpG) chaperone affects the CRISPR/Cas system in *E. coli* by stabilizing Cas3. If HtpG is absent, the CRISPR/Cas system is no longer able to efficiently prevent phage infection [107]. However, HtpG is not present in all bacteria containing CRISPR/Cas systems. Indeed, most archaeal strains that contain Cas3 homologues do not carry an HtpG homolog, suggesting

that it is not universally required. Therefore, other bacteria may require other chaperones in order for CRISPR/Cas interference to occur [108].

Other systems that affect CRISPR/Cas regulation were identified in *Myxococcus xanthus*, *T. thermophilus* and Sulfolobales. In *Myxococcus*, the *dev* operon is co-transcribed with *cas* genes, and this operon is negatively auto-regulated by DevS (a Cas5 protein) [109]. *T. thermophilus* contains 12 CRISPRs, and 3 different *cas* operons, type I-E, type III-A and type III-B [10,39]. During phage infection, the *cas* genes for the type I-E and type III-A systems were up-regulated via CRP (cAMP receptor protein) [39,110]. When CRP is bound to cAMP, it recognizes promoters and activates gene expression. Interestingly, inactivation of *crp* did not abolish the *cas* activation, suggesting both *crp*-dependent and -independent pathways. In addition, some *cas* genes and CRISPR arrays were regulated by phage infection in a CRP-independent manner [39]. In *Sulfolobus islandicus* and *S. solfataricus*, the expression of pre-crRNA is regulated by Cbp1 (CRISPR DNA repeat binding protein), which directly binds the DNA repeats and might influence the activity of promoters and terminators present within some spacers or repeats [111,112]. In summary, CRISPR/Cas systems can be regulated at the level of transcription of the *cas* genes and CRISPR arrays and post-translational level on the Cas proteins (Figure 4). The picture of how these systems are regulated is far from complete and diverse and complex regulatory strategies exist.

6. Conclusions

CRISPR/Cas systems are widespread and versatile prokaryotic defense mechanisms providing adaptive and heritable immunity to extrachromosomal elements. In recent years research performed on these systems has led to a deeper understanding of the underlying mechanistic principles. Despite this rapid increase in knowledge, many questions remain to be answered. For example, challenges in the future include elucidating why some prokaryotes carry multiple CRISPR/Cas types and how their activity is coordinated. Furthermore, the adaptation step is still poorly understood, especially with respect to the specific roles of the Cas proteins involved and how foreign DNA is selected for integration. Major progress has been made regarding structure, function and interactions of some Cas proteins and their complexes. However, for many subtype-specific proteins, functional and structural information is still lacking and their precise role in the CRISPR/Cas mechanism remains unknown. Finally, the question, how CRISPR/Cas activity is regulated and to what extend it plays a role in areas not related to defense, e.g., DNA repair, needs to be addressed. Future research on CRISPR/Cas systems will also open up vast opportunities to utilize them, such as genome engineering for the creation of more robust industrial starter strains or as a gene targeting or gene silencing mechanism similar to that used in eukaryotic cells.

Acknowledgments

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Frost, L.S.; Leplae, R.; Summers, A.O.; Toussaint, A. Mobile genetic elements: The agents of open source evolution. *Nat. Rev. Microbiol.* **2005**, *3*, 722–732.

2. Koonin, E.V.; Wolf, Y.I. Genomics of bacteria and archaea: The emerging dynamic view of the prokaryotic world. *Nucleic Acids Res.* **2008**, *36*, 6688–6719.

3. Fuhrman, J.A. Marine viruses and their biogeochemical and ecological effects. *Nature* **1999**, *399*, 541–548.

4. Hendrix, R.W. Bacteriophage genomics. *Curr. Opin. Microbiol.* **2003**, *6*, 506–511.

5. Labrie, S.J.; Samson, J.E.; Moineau, S. Bacteriophage resistance mechanisms. *Nat. Rev. Microbiol.* **2010**, *8*, 317–327.

6. Petty, N.K.; Evans, T.J.; Fineran, P.C.; Salmond, G.P. Biotechnological exploitation of bacteriophage research. *Trends Biotechnol.* **2007**, *25*, 7–15.

7. Marraffini, L.A.; Sontheimer, E.J. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat. Rev. Genet.* **2010**, *11*, 181–190.

8. Terns, M.P.; Terns, R.M. CRISPR-based adaptive immune systems. *Curr. Opin. Microbiol.* **2011**, *14*, 321–327.

9. Wiedenheft, B.; Sternberg, S.H.; Doudna, J.A. RNA-guided genetic silencing systems in bacteria and archaea. *Nature* **2012**, *482*, 331–338.

10. Makarova, K.S.; Haft, D.H.; Barrangou, R.; Brouns, S.J.; Charpentier, E.; Horvath, P.; Moineau, S.; Mojica, F.J.; Wolf, Y.I.; Yakunin, A.F.; *et al*. Evolution and classification of the CRISPR-Cas systems. *Nat. Rev. Microbiol.* **2011**, *9*, 467–477.

11. Ishino, Y.; Shinagawa, H.; Makino, K.; Amemura, M.; Nakata, A. Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J. Bacteriol.* **1987**, *169*, 5429–5433.

12. Groenen, P.M.; Bunschoten, A.E.; van Soolingen, D.; van Embden, J.D. Nature of DNA polymorphism in the direct repeat cluster of *Mycobacterium tuberculosis*; application for strain differentiation by a novel typing method. *Mol. Microbiol.* **1993**, *10*, 1057–1065.

13. Mojica, F.J.; Ferrer, C.; Juez, G.; Rodríguez-Valera, F. Long stretches of short tandem repeats are present in the largest replicons of the Archaea *Haloferax mediterranei* and *Haloferax volcanii* and could be involved in replicon partitioning. *Mol. Microbiol.* **1995**, *17*, 85–93.

134

14. Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, R.J.; Gwinn, M.; Hickey, E.K.; Peterson, J.D.; *et al*. The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*. *Nature* **1997**, *390*, 364–370.

15. Jansen, R.; van Embden, J.D.; Gaastra, W.; Schouls, L.M. Identification of a novel family of sequence repeats among prokaryotes. *OMICS* **2002**, *6*, 23–33.

16. Mojica, F.J.; Díez-Villaseñor, C.; Soria, E.; Juez, G. Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Mol. Microbiol.* **2000**, *36*, 244–246.

17. Grissa, I.; Vergnaud, G.; Pourcel, C. The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinformatics* **2007**, *8*, 172.

18. van Belkum, A.; Scherer, S.; van Alphen, L.; Verbrugh, H. Short-sequence DNA repeats in prokaryotic genomes. *Microbiol. Mol. Biol. Rev.* **1998**, *62*, 275–293.

19. Jansen, R.; Embden, J.D.; Gaastra, W.; Schouls, L.M. Identification of genes that are associated with DNA repeats in prokaryotes. *Mol. Microbiol.* **2002**, *43*, 1565–1575.

20. Makarova, K.S.; Aravind, L.; Grishin, N.V.; Rogozin, I.B.; Koonin, E.V. A DNA repair system specific for thermophilic Archaea and bacteria predicted by genomic context analysis. *Nucleic Acids Res.* **2002**, *30*, 482–496.

21. Mojica, F.J.; Díez-Villaseñor, C.; García-Martínez, J.; Soria, E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.* **2005**, *60*, 174–182.

22. Bolotin, A.; Quinquis, B.; Sorokin, A.; Ehrlich, S.D. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* **2005**, *151*, 2551–2561.

23. Pourcel, C.; Salvignol, G.; Vergnaud, G. CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* **2005**, *151*, 653–663.

24. Makarova, K.S.; Grishin, N.V.; Shabalina, S.A.; Wolf, Y.I.; Koonin, E.V. A putative RNA-interference-based immune system in prokaryotes: Computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol. Direct* **2006**, *1*, doi:10.1186/1745-6150-1-7.

25. Haft, D.H.; Selengut, J.; Mongodin, E.F.; Nelson, K.E. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Comput. Biol.* **2005**, *1*, e60.

26. Barrangou, R.; Fremaux, C.; Deveau, H.; Richards, M.; Boyaval, P.; Moineau, S.; Romero, D.A.; Horvath, P. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **2007**, *315*, 1709–1712.

27. Marraffini, L.A.; Sontheimer, E.J. CRISPR interference limits horizontal gene transfer in *Staphylococci* by targeting DNA. *Science* **2008**, *322*, 1843–1845.

28. Bikard, D.; Hatoum-Aslan, A.; Mucida, D.; Marraffini, L.A. CRISPR Interference Can Prevent Natural Transformation and Virulence Acquisition during *In Vivo* Bacterial Infection. *Cell Host Microbe* **2012**, *12*, 177–186.

29. Lillestøl, R.K.; Redder, P.; Garrett, R.A.; Brugger, K. A putative viral defence mechanism in archaeal cells. *Archaea* **2006**, *2*, 59–72.

30. Nakata, A.; Amemura, M.; Makino, K. Unusual nucleotide arrangement with repeated sequences in the *Escherichia coli* K-12 chromosome. *J. Bacteriol.* **1989**, *171*, 3553–3556.

31. She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-Weiher, C.C.; Clausen, I.G.; Curtis, B.A.; De Moors, A.; *et al.* The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 7835–7840.

32. Van Embden, J.D.; van Gorkom, T.; Kremer, K.; Jansen, R.; van der Zeijst, B.A.; Schouls, L.M. Genetic variation and evolutionary origin of the direct repeat locus of *Mycobacterium tuberculosis* complex bacteria. *J. Bacteriol.* **2000**, *182*, 2393–2401.

33. Horvath, P.; Romero, D.A.; Côuté-Monvoisin, A.C.; Richards, M.; Deveau, H.; Moineau, S.; Boyaval, P.; Fremaux, C.; Barrangou, R. Diversity, activity, and evolution of CRISPR loci in *Streptococcus thermophilus*. *J. Bacteriol.* **2008**, *190*, 1401–1412.

34. Kunin, V.; Sorek, R.; Hugenholtz, P. Evolutionary conservation of sequence and secondary structures in CRISPR repeats. *Genome Biol.* **2007**, *8*, doi:10.1186/gb-2007-8-4-r61.

35. Grissa, I.; Vergnaud, G.; Pourcel, C. CRISPRFinder: A web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res.* **2007**, *35*, W52–W57.

36. Brodt, A.; Lurie-Weinberger, M.N.; Gophna, U. CRISPR loci reveal networks of gene exchange in archaea. *Biol. Direct* **2011**, *6*, doi:10.1186/1745-6150-6-65.

37. Deveau, H.; Barrangou, R.; Garneau, J.E.; Labonté, J.; Fremaux, C.; Boyaval, P.; Romero, D.A.; Horvath, P.; Moineau, S. Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J. Bacteriol.* **2008**, *190*, 1390–1400.

38. Stern, A.; Keren, L.; Wurtzel, O.; Amitai, G.; Sorek, R. Self-targeting by CRISPR: Gene regulation or autoimmunity? *Trends Genet.* **2010**, *26*, 335–340.

39. Agari, Y.; Sakamoto, K.; Tamakoshi, M.; Oshima, T.; Kuramitsu, S.; Shinkai, A. Transcription profile of *Thermus thermophilus* CRISPR systems after phage infection. *J. Mol. Biol.* **2009**, *395*, 270–281.

40. Pougach, K.; Semenova, E.; Bogdanova, E.; Datsenko, K.A.; Djordjevic, M.; Wanner, B.L.; Severinov, K. Transcription, processing and function of CRISPR cassettes in *Escherichia coli*. *Mol. Microbiol.* **2010**, *77*, 1367–1379.

41. Pul, Ü.; Wurm, R.; Arslan, Z.; Geissen, R.; Hofmann, N.; Wagner, R. Identification and characterization of *E. coli* CRISPR-cas promoters and their silencing by H-NS. *Mol. Microbiol.* **2010**, *75*, 1495–1512.

42. Przybilski, R.; Richter, C.; Gristwood, T.; Clulow, J.S.; Vercoe, R.B.; Fineran, P.C. Csy4 is responsible for CRISPR RNA processing in *Pectobacterium atrosepticum*. *RNA Biol.* **2011**, *8*, 517–528.

43. Yosef, I.; Goren, M.G.; Qimron, U. Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*. *Nucleic Acids Res.* **2012**, *40*, 5569–5576.

44. Godde, J.S.; Bickerton, A. The repetitive DNA elements called CRISPRs and their associated genes: Evidence of horizontal transfer among prokaryotes. *J. Mol. Evol.* **2006**, *62*, 718–729.

45. Brouns, S.J.; Jore, M.M.; Lundgren, M.; Westra, E.R.; Slijkhuis, R.J.; Snijders, A.P.; Dickman, M.J.; Makarova, K.S.; Koonin, E.V.; van der Oost, J. Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **2008**, *321*, 960–964.

46. Jore, M.M.; Lundgren, M.; van Duijn, E.; Bultema, J.B.; Westra, E.R.; Waghmare, S.P.; Wiedenheft, B.; Pul, Ü.; Wurm, R.; Wagner, R.; *et al*. Structural basis for CRISPR RNA-guided DNA recognition by Cascade. *Nat. Struct. Mol. Biol.* **2011**, *18*, 529–536.

47. Wiedenheft, B.; Lander, G.C.; Zhou, K.; Jore, M.M.; Brouns, S.J.; van der Oost, J.; Doudna, J.A.; Nogales, E. Structures of the RNA-guided surveillance complex from a bacterial immune system. *Nature* **2011**, *477*, 486–489.

48. Wiedenheft, B.; van Duijn, E.; Bultema, J.B.; Waghmare, S.P.; Zhou, K.; Barendregt, A.; Westphal, W.; Heck, A.J.; Boekema, E.J.; Dickman, M.J.; *et al*. RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 10092–10097.

49. Nam, K.H.; Haitjema, C.; Liu, X.; Ding, F.; Wang, H.; DeLisa, M.P.; Ke, A. Cas5d Protein Processes Pre-crRNA and Assembles into a Cascade-like Interference Complex in Subtype I-C/Dvulg CRISPR-Cas System. *Structure* **2012**, *20*, 1574–1584.

50. Andersson, A.F.; Banfield, J.F. Virus population dynamics and acquired virus resistance in natural microbial communities. *Science* **2008**, *320*, 1047–1050.

51. Tyson, G.W.; Banfield, J.F. Rapidly evolving CRISPRs implicated in acquired resistance of microorganisms to viruses. *Environ. Microbiol.* **2008**, *10*, 200–207.

52. Pride, D.T.; Salzman, J.; Haynes, M.; Rohwer, F.; Davis-Long, C.; White, R.A., 3rd.; Loomer, P.; Armitage, G.C.; Relman, D.A. Evidence of a robust resident bacteriophage population revealed through analysis of the human salivary virome. *ISME J.* **2012**, *6*, 915–926.

53. Garneau, J.E.; Dupuis, M.E.; Villion, M.; Romero, D.A.; Barrangou, R.; Boyaval, P.; Fremaux, C.; Horvath, P.; Magadan, A.H.; Moineau, S. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **2010**, *468*, 67–71.

54. Swarts, D.C.; Mosterd, C.; van Passel, M.W.; Brouns, S.J. CRISPR interference directs strand specific spacer acquisition. *PLoS One* **2012**, *7*, e35888.

55. Datsenko, K.A.; Pougach, K.; Tikhonov, A.; Wanner, B.L.; Severinov, K.; Semenova, E. Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. *Nat. Commun.* **2012**, *3*, doi:10.1038/ncomms1937.

56. Cady, K.C.; Bondy-Denomy, J.; Heussler, G.E.; Davidson, A.R.; O'Toole, G.A. The CRISPR/Cas Adaptive Immune System of *Pseudomonas aeruginosa* Mediates Resistance to Naturally Occurring and Engineered Phages. *J. Bacteriol.* **2012**, *194*, 5728–5738.

57. Lopez-Sanchez, M.J.; Sauvage, E.; Da Cunha, V.; Clermont, D.; Ratsima Hariniaina, E.; Gonzalez-Zorn, B.; Poyart, C.; Rosinski-Chupin, I.; Glaser, P. The highly dynamic CRISPR1 system of *Streptococcus agalactiae* controls the diversity of its mobilome. *Mol. Microbiol.* **2012**, *85*, 1057–1071.

58. Erdmann, S.; Garrett, R.A. Selective and hyperactive uptake of foreign DNA by adaptive immune systems of an archaeon via two distinct mechanisms. *Mol. Microbiol.* **2012**, *85* 1044–1056.

59. Sapranauskas, R.; Gasiunas, G.; Fremaux, C.; Barrangou, R.; Horvath, P.; Siksnys, V. The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res.* **2011**, *39*, 9275–9282.

60. Hatoum-Aslan, A.; Maniv, I.; Marraffini, L.A. Mature clustered, regularly interspaced, short palindromic repeats RNA (crRNA) length is measured by a ruler mechanism anchored at the precursor processing site. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 21218–21222.

61. Wiedenheft, B.; Zhou, K.; Jinek, M.; Coyle, S.M.; Ma, W.; Doudna, J.A. Structural basis for DNase activity of a conserved protein implicated in CRISPR-mediated genome defense. *Structure* **2009**, *17*, 904–912.

62. Han, D.; Lehmann, K.; Krauss, G. SSO1450—A CAS1 protein from *Sulfolobus solfataricus* P2 with high affinity for RNA and DNA. *FEBS Lett.* **2009**, *583*, 1928–1932.

63. Babu, M.; Beloglazova, N.; Flick, R.; Graham, C.; Skarina, T.; Nocek, B.; Gagarinova, A.; Pogoutse, O.; Brown, G.; Binkowski, A.; *et al.* A dual function of the CRISPR-Cas system in bacterial antivirus immunity and DNA repair. *Mol. Microbiol.* **2011**, *79*, 484–502.

64. Beloglazova, N.; Brown, G.; Zimmerman, M.D.; Proudfoot, M.; Makarova, K.S.; Kudritska, M.; Kochinyan, S.; Wang, S.; Chruszcz, M.; Minor, W.; *et al.* A novel family of sequence-specific endoribonucleases associated with the clustered regularly interspaced short palindromic repeats. *J. Biol. Chem.* **2008**, *283*, 20361–20371.

65. Samai, P.; Smith, P.; Shuman, S. Structure of a CRISPR-associated protein Cas2 from *Desulfovibrio vulgaris*. *Acta Crystallogr. F Struct. Biol. Cryst. Commun.* **2010**, *66*, 1552–1556.

66. Nam, K.H.; Ding, F.; Haitjema, C.; Huang, Q.; Delisa, M.P.; Ke, A. Double-stranded Endonuclease Activity in B. halodurans Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)-associated Cas2 Protein. *J. Biol. Chem.* **2012**, Epub ahead of print.

67. Plagens, A.; Tjaden, B.; Hagemann, A.; Randau, L.; Hensel, R. Characterization of the CRISPR/Cas subtype I-A system of the hyperthermophilic crenarchaeon *Thermoproteus tenax*. *J. Bacteriol.* **2012**, *194*, 2491–2500.

68. Goren, M.G.; Yosef, I.; Auster, O.; Qimron, U. Experimental Definition of a Clustered Regularly Interspaced Short Palindromic Duplicon in *Escherichia coli*. *J. Mol. Biol.* **2012**, *423*, 14–16.

69. Westra, E.R.; Brouns, S.J. The rise and fall of CRISPRs—Dynamics of spacer acquisition and loss. *Mol. Microbiol.* **2012**, *85*, 1021–1025.

70. Mojica, F.J.; Diez-Villasenor, C.; Garcia-Martinez, J.; Almendros, C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* **2009**, *155*, 733–740.

138

71. Richter, H.; Zoephel, J.; Schermuly, J.; Maticzka, D.; Backofen, R.; Randau, L. Characterization of CRISPR RNA processing in *Clostridium thermocellum* and *Methanococcus maripaludis*. *Nucleic Acids Res.* **2012**, Epub ahead of print.

72. Sashital, D.G.; Jinek, M.; Doudna, J.A. An RNA-induced conformational change required for CRISPR RNA cleavage by the endoribonuclease Cse3. *Nat. Struct. Mol. Biol.* **2011**, *18*, 680–687.

73. Gesner, E.M.; Schellenberg, M.J.; Garside, E.L.; George, M.M.; MacMillan, A.M. Recognition and maturation of effector RNAs in a CRISPR interference pathway. *Nat. Struct. Mol. Biol.* **2011**, *18*, 688–692.

74. Haurwitz, R.E.; Jinek, M.; Wiedenheft, B.; Zhou, K.; Doudna, J.A. Sequence- and structure-specific RNA processing by a CRISPR endonuclease. *Science* **2010**, *329*, 1355–1358.

75. Sternberg, S.H.; Haurwitz, R.E.; Doudna, J.A. Mechanism of substrate selection by a highly specific CRISPR endoribonuclease. *RNA* **2012**, *18*, 661–672.

76. Haurwitz, R.E.; Sternberg, S.H.; Doudna, J.A. Csy4 relies on an unusual catalytic dyad to position and cleave CRISPR RNA. *EMBO J.* **2012**, *31*, 2824–2832.

77. Lintner, N.G.; Kerou, M.; Brumfield, S.K.; Graham, S.; Liu, H.; Naismith, J.H.; Sdano, M.; Peng, N.; She, Q.; Copie, V.; *et al*. Structural and Functional Characterization of an Archaeal Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR)-associated Complex for Antiviral Defense (CASCADE). *J. Biol. Chem.* **2011**, *286*, 21643–21656.

78. Sashital, D.G.; Wiedenheft, B.; Doudna, J.A. Mechanism of foreign DNA selection in a bacterial adaptive immune system. *Mol. Cell.* **2012**, *46*, 606–615.

79. Semenova, E.; Jore, M.M.; Datsenko, K.A.; Semenova, A.; Westra, E.R.; Wanner, B.; van der Oost, J.; Brouns, S.J.; Severinov, K. Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 10098–10103.

80. Westra, E.R.; van Erp, P.B.; Kunne, T.; Wong, S.P.; Staals, R.H.; Seegers, C.L.; Bollen, S.; Jore, M.M.; Semenova, E.; Severinov, K.; *et al*. CRISPR Immunity Relies on the Consecutive Binding and Degradation of Negatively Supercoiled Invader DNA by Cascade and Cas3. *Mol. Cell* **2012**, *46*, 595–605.

81. Howard, J.A.; Delmas, S.; Ivančić-Baće, I.; Bolt, E.L. Helicase dissociation and annealing of RNA-DNA hybrids by *Escherichia coli* Cas3 protein. *Biochem. J.* **2011**, *439*, 85–95.

82. Sinkunas, T.; Gasiunas, G.; Fremaux, C.; Barrangou, R.; Horvath, P.; Siksnys, V. Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. *EMBO J.* **2011**, *30*, 1335–1342.

83. Beloglazova, N.; Petit, P.; Flick, R.; Brown, G.; Savchenko, A.; Yakunin, A.F. Structure and activity of the Cas3 HD nuclease MJ0384, an effector enzyme of the CRISPR interference. *EMBO J.* **2011**, *30*, 4616–4627.

84. Deltcheva, E.; Chylinski, K.; Sharma, C.M.; Gonzales, K.; Chao, Y.; Pirzada, Z.A.; Eckert, M.R.; Vogel, J.; Charpentier, E. CRISPR RNA maturation by *trans*-encoded small RNA and host factor RNase III. *Nature* **2011**, *471*, 602–607.

85. Jinek, M.; Chylinski, K.; Fonfara, I.; Hauer, M.; Doudna, J.A.; Charpentier, E. A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* **2012**, *337*, 816–821.

86. Carte, J.; Wang, R.; Li, H.; Terns, R.M.; Terns, M.P. Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. *Genes Dev.* **2008**, *22*, 3489–3496.

87. Carte, J.; Pfister, N.T.; Compton, M.M.; Terns, R.M.; Terns, M.P. Binding and cleavage of CRISPR RNA by Cas6. *RNA* **2010**, *16*, 2181–2188.

88. Wang, R.; Preamplume, G.; Terns, M.P.; Terns, R.M.; Li, H. Interaction of the Cas6 riboendonuclease with CRISPR RNAs: Recognition and cleavage. *Structure* **2011**, *19*, 257–264.

89. Hale, C.R.; Majumdar, S.; Elmore, J.; Pfister, N.; Compton, M.; Olson, S.; Resch, A.M.; Glover, C.V., 3rd.; Graveley, B.R.; Terns, R.M.; *et al*. Essential features and rational design of CRISPR RNAs that function with the Cas RAMP module complex to cleave RNAs. *Mol. Cell* **2012**, *45*, 292–302.

90. Hale, C.R.; Zhao, P.; Olson, S.; Duff, M.O.; Graveley, B.R.; Wells, L.; Terns, R.M.; Terns, M.P. RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell* **2009**, *139*, 945–956.

91. Hale, C.; Kleppe, K.; Terns, R.M.; Terns, M.P. Prokaryotic silencing (psi)RNAs in *Pyrococcus furiosus*. *RNA* **2008**, *14*, 2572–2579.

92. Zhang, J.; Rouillon, C.; Kerou, M.; Reeks, J.; Brugger, K.; Graham, S.; Reimann, J.; Cannone, G.; Liu, H.; Albers, S.V.; *et al*. Structure and mechanism of the CMR complex for CRISPR-mediated antiviral immunity. *Mol. Cell* **2012**, *45*, 303–313.

93. Marraffini, L.A.; Sontheimer, E.J. Self versus non-self discrimination during CRISPR RNA-directed immunity. *Nature* **2010**, *463*, 568–571.

94. Young, J.C.; Dill, B.D.; Pan, C.; Hettich, R.L.; Banfield, J.F.; Shah, M.; Fremaux, C.; Horvath, P.; Barrangou, R.; VerBerkmoes, N.C. Phage-induced expression of CRISPR-associated proteins is revealed by shotgun proteomics in *Streptococcus thermophilus*. *PLoS One* **2012**, *7*, e38077.

95. Dorman, C.J. H-NS, the genome sentinel. *Nat. Rev. Microbiol.* **2007**, *5*, 157–161.

96. Westra, E.R.; Pul, Ü.; Heidrich, N.; Jore, M.M.; Lundgren, M.; Stratmann, T.; Wurm, R.; Raine, A.; Mescher, M.; van Heereveld, L.; *et al*. H-NS-mediated repression of CRISPR-based immunity in *Escherichia coli* K12 can be relieved by the transcription activator LeuO. *Mol. Microbiol.* **2010**, *77*, 1380–1393.

97. Lucchini, S.; Rowley, G.; Goldberg, M.D.; Hurd, D.; Harrison, M.; Hinton, J.C. H-NS mediates the silencing of laterally acquired genes in bacteria. *PLoS Pathog.* **2006**, *2*, e81.

98. Navarre, W.W.; Porwollik, S.; Wang, Y.; McClelland, M.; Rosen, H.; Libby, S.J.; Fang, F.C. Selective silencing of foreign DNA with low GC content by the H-NS protein in *Salmonella*. *Science* **2006**, *313*, 236–238.

99. Mojica, F.J.; Díez-Villaseñor, C. The on-off switch of CRISPR immunity against phages in *Escherichia coli*. *Mol. Microbiol.* **2010**, *77*, 1341–1345.

100. Medina-Aparicio, L.; Rebollar-Flores, J.E.; Gallego-Hernández, A.L.; Vázquez, A.; Olvera, L.; Gutiérrez-Ríos, R.M.; Calva, E.; Hernández-Lucas, I. The CRISPR/Cas Immune System Is an Operon Regulated by LeuO, H-NS, and Leucine-Responsive Regulatory Protein in *Salmonella enterica* Serovar Typhi. *J. Bacteriol.* **2011**, *193*, 2396–2407.

101. Hernández-Lucas, I.; Gallego-Hernández, A.L.; Encarnación, S.; Fernández-Mora, M.; Martínez-Batallar, A.G.; Salgado, H.; Oropeza, R.; Calva, E. The LysR-type transcriptional regulator LeuO controls expression of several genes in *Salmonella enterica* serovar Typhi. *J. Bacteriol.* **2008**, *190*, 1658–1670.

102. Majumder, A.; Fang, M.; Tsai, K.J.; Ueguchi, C.; Mizuno, T.; Wu, H.Y. LeuO expression in response to starvation for branched-chain amino acids. *J. Biol. Chem.* **2001**, *276*, 19046–19051.

103. Drahos, D.J.; Hendrix, R.W. Effect of bacteriophage lambda infection on synthesis of *groE* protein and other *Escherichia coli* proteins. *J. Bacteriol.* **1982**, *149*, 1050–1063.

104. MacRitchie, D.M.; Buelow, D.R.; Price, N.L.; Raivio, T.L. Two-component signaling and gram negative envelope stress response systems. *Adv. Exp. Med. Biol.* **2008**, *631*, 80–110.

105. Perez-Rodriguez, R.; Haitjema, C.; Huang, Q.; Nam, K.H.; Bernardis, S.; Ke, A.; DeLisa, M.P. Envelope stress is a trigger of CRISPR RNA-mediated DNA silencing in *Escherichia coli*. *Mol. Microbiol.* **2011**, *79*, 584–599.

106. Baranova, N.; Nikaido, H. The *baeSR* two-component regulatory system activates transcription of the *yegMNOB* (*mdtABCD*) transporter gene cluster in *Escherichia coli* and increases its resistance to novobiocin and deoxycholate. *J. Bacteriol.* **2002**, *184*, 4168–4176.

107. Yosef, I.; Goren, M.G.; Kiro, R.; Edgar, R.; Qimron, U. High-temperature protein G is essential for activity of the *Escherichia coli* clustered regularly interspaced short palindromic repeats (CRISPR)/Cas system. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 20136–20141.

108. Chen, C.C.; Ghole, M.; Majumder, A.; Wang, Z.; Chandana, S.; Wu, H.Y. LeuO-mediated transcriptional derepression. *J. Biol. Chem.* **2003**, *278*, 38094–38103.

109. Viswanathan, P.; Murphy, K.; Julien, B.; Garza, A.G.; Kroos, L. Regulation of *dev*, an operon that includes genes essential for *Myxococcus xanthus* development and CRISPR-associated genes and repeats. *J. Bacteriol.* **2007**, *189*, 3738–3750.

110. Shinkai, A.; Kira, S.; Nakagawa, N.; Kashihara, A.; Kuramitsu, S.; Yokoyama, S. Transcription activation mediated by a cyclic AMP receptor protein from *Thermus thermophilus* HB8. *J. Bacteriol.* **2007**, *189*, 3891–3901.

111. Deng, L.; Kenchappa, C.S.; Peng, X.; She, Q.; Garrett, R.A. Modulation of CRISPR locus transcription by the repeat-binding protein Cbp1 in *Sulfolobus*. *Nucleic Acids Res.* **2012**, *40*, 2470–2480.

112. Peng, X.; Brugger, K.; Shen, B.; Chen, L.; She, Q.; Garrett, R.A. Genus-specific protein binding to the large clusters of DNA repeats (short regularly spaced repeats) present in *Sulfolobus* genomes. *J. Bacteriol.* **2003**, *185*, 2410–2417.

Review

The Staphylococci Phages Family: An Overview

Marie Deghorain * and Laurence van Melderen *

Laboratoire de Génétique et Physiologie Bactérienne, Faculté de Sciences, IBMM, Université Libre de Bruxelles (ULB), Gosselies B-6141, Belgium

*   Author to whom correspondence should be addressed; E-Mails: lvmelder@ulb.ac.be (L.V.M.); mdeghora@ulb.ac.be (M.D.); Tel.: +32-2-650-97-76 (M.D.); +32-2-650-97-78 (L.V.M.); Fax: +32-2-650-97-70.

Abstract: Due to their crucial role in pathogenesis and virulence, phages of *Staphylococcus aureus* have been extensively studied. Most of them encode and disseminate potent staphylococcal virulence factors. In addition, their movements contribute to the extraordinary versatility and adaptability of this prominent pathogen by improving genome plasticity. In addition to *S. aureus*, phages from coagulase-negative *Staphylococci* (CoNS) are gaining increasing interest. Some of these species, such as *S. epidermidis*, cause nosocomial infections and are therefore problematic for public health. This review provides an overview of the staphylococcal phages family extended to CoNS phages. At the morphological level, all these phages characterized so far belong to the *Caudovirales* order and are mainly temperate *Siphoviridae*. At the molecular level, comparative genomics revealed an extensive mosaicism, with genes organized into functional modules that are frequently exchanged between phages. Evolutionary relationships within this family, as well as with other families, have been highlighted. All these aspects are of crucial importance for our understanding of evolution and emergence of pathogens among bacterial species such as *Staphylococci*.

Keywords: bacteriophages; *Staphylococcus*; horizontal transfer; virulence

1. Introduction

The vast majority of bacteria contain prophages, either integrated into their chromosome or as extra-chromosomal elements, accounting for substantial genetic variability. Not only do phages shape bacterial genome architecture, they also constitute major vehicles for horizontal gene transfer [1,2]. In addition, they contribute to virulence by encoding numerous virulence or fitness factors and by their movements within genomes (see below, section 2.2.) [2–5]. These mobile elements are responsible for gene disruption, provide anchor region for genomic rearrangements, protect bacteria from lytic infections or, in contrast, provoke cell lysis through prophage induction [2]. Thus, phages play essential roles in bacterial evolution and adaptation.

Phages are widespread in *Staphylococcus aureus* and have been extensively studied [1,3,6]. They were firstly used for the typing of clinical *S. aureus* isolates [7,8]. *S. aureus* is a major human and animal pathogen that causes both nosocomial and community-acquired infections. It colonizes skin and mucous membranes, with the anterior nares being the primary niche in humans. While found in healthy carriers, *S. aureus* is also responsible for a wide range of diseases, from mild skin infections to severe life-threatening infections, such as sepsis or endocarditis [6]. The number of prophages in *S. aureus* genome is generally high. All *S. aureus* genome sequenced so far do contain at least one prophage, and many strains carry up to four [1]. These encode numerous staphylococcal toxins responsible for pathogenesis [1,2,6].

Staphylococci also comprise coagulase-negative species (coagulase-negative *Staphylococci*, CoNS), which are distinguishable from *S. aureus* by the lack of coagulase-encoding gene. In contrast to *S. aureus*, which is only found in part of the population, these species belong to the commensal flora of healthy humans. Some species are associated to specific niches, and others appear to be more 'generalist' and are generally found on the body surface [9–12]. CoNS include human opportunistic pathogens often associated with medical devices. *S. epidermidis* is referred to as a frequent cause of nosocomial infections [9–12]. In addition, 'true' pathogens that are not associated with medical devices may also be problematic for public health. As an example, *S. saprophyticus* is considered as a frequent pathogen responsible for uncomplicated urinary tract infections [11,13]. Pathogenesis of CoNS species relies on factors required for their commensal mode of life or fitness (e.g., factors involved in adhesion, in biofilm formation and in persistence) and not on toxins, as observed for *S. aureus* [10,11]. As a consequence, less attention has been paid to these phages.

During the past decade, sequencing of *Staphylococci* genomes and extensive comparative genomic analyses have significantly increased the number of staphylococcal phages identified. Up to now, more than 68 *Staphylococci* phages and prophages sequences, mainly from *S. aureus*, are found in the [14]. In addition, 268 *Staphylococci* genomes are available on the PATRIC server [15] and offer a remarkable source of novel prophage sequences for further studies (see below).

In this review, we provide an overview of *Staphylococci* phages with a focus on their contribution to pathogenesis. A special interest is placed on the classification methods, as well as on the evolutionary relationships connecting staphylococcal phages. Phage classification is often problematic, due to the modular organization of phage genomes. Relationships between *Staphylococci* phages and

phages from other species are also discussed in an evolutionary perspective. Finally, the potential use of staphylococcal phages for bio-technological and medical applications is briefly addressed.

2. The Phages of *S. aureus*

2.1. Global Features of S. aureus *Phages*

2.1.1. Morphological Families

As the vast majority of phages, the *S. aureus* phages known so far are double-stranded DNA phages belonging to the *Siphoviridae* family of the *Caudovirales* order (Table S1) (reviewed in [2,4,6]). In general, they are temperate phages detected as prophage inserted in the chromosome, some of them being lytic due to mutations in the lysogeny functions (e.g., phiIPLA35 and phiIPLA88; [16]; or SA11; [17]). According to the morphological classification previously proposed by Ackermann [18], staphylococcal *Siphoviridae* are composed of an icosahedral capsid and a non-contractile tail ended by a base-plate structure. Capsids may adopt elongated or isometric shapes, and tail length varies from short (130 nm) to long (400 nm). A small number of *S. aureus Podoviridae* and *Myoviridae* phages, also belonging to the *Caudovirales* order, were described (Table S1). *Podoviridae,* such as the recently identified SAP-2 phage [19], are composed of a small icosahedral capsid and a short, non-flexible, non-contractile tail. *Myoviridae* phages, such as the well-known phage K [20], are characterized by an icosahedral capsid and a long contractile tail.

2.1.2. Genomic Characteristics of *S. aureus* Phages

A comparative study of *S. aureus* phage genomes performed by Pelletier and co-workers [3] revealed several key genomic features, which are globally applicable to all staphylococcal phages described so far. The analysis encompassed 27 genomes from *S. aureus* phages and prophages belonging to the three morphological families described above.

Genome size extends from less than 20 kb up to more than 125 kb [3]. In contrast to phages from *Mycobacterium* [21] or *Pseudomonas aeruginosa* [22], genome sizes are not uniformly distributed, and three categories can be established and used to classify *Staphylococci* phages (class I: < 20 kb; class II: ≈ 40 kb, class III: >125 kb; see section 4.1) [3]. Interestingly, genome size categories correlate with the morphological classification, *Podoviridae* harboring the smallest genomes (class I), *Myoviridae* the largest ones (class III), and *Siphoviridae* showing intermediate sizes (class II).

Coding regions are tightly packed with very few and small intergenic regions and a high gene density (1.67 genes/kb in average) [3]. Their GC content is similar to that of the host. *S. aureus* phages provide an impressive, mainly unexplored source of genetic diversity. On the 2,170 predicted proteins from the 27 phages analyzed in the study of Pelletier and co-workers, a function could be assigned using BLAST to 35% of the ORFs. No match was detected for 44% of these ORFS in Bacteria and Phages Gene Bank databases [3].

Genomes of *S. aureus Siphoviridae* display the typical structure of the morphological family (Figure 1a) [2,4,23]. Five functional modules are arranged as follows: lysogeny, DNA metabolism, DNA packaging and capsid morphogenesis, tail morphogenesis and host cell lysis. The DNA metabolism module can be divided into replication and regulation functions. When present, virulence factors are generally encoded downstream of the lysis module [2,4]. In some cases, they are inserted between the lysogeny and DNA metabolism modules as reported for phiNM1 to four prophages found in the *S. aureus* Newman strain [24]. Genes are generally transcribed on the same strand, except for small clusters, such as genes involved in host genome integration [3,4].

While harboring a modular structure as well, organization of *Podoviridae* genomes is different (Figure 1b) [3,19,25]. One major distinction resides in a smaller number of ORFs, as indicated by a smaller size (20 to 32 ORFs) [3,19,25]. Functional modules encoding DNA packaging and capsid morphogenesis, tail morphogenesis and lysis were identified, in addition to genes of unknown function. In contrast to *Siphoviridae*, modules are not well defined and tail and lysis genes are overlapping. In addition, the lysogeny module is absent as expected for lytic phages.

Genome organization of staphylococcal *Myoviridae* (e.g., phages K, G1, Twort) [3,20,26] is similar to *E. coli* T4 phage, the paradigm for *Myoviridae* [27]. Genomes are organized into functional modules of conserved genes (replication, structural elements), interrupted by large plastic regions encoding mainly genes of unknown function (Figure 1c). Structural modules found in staphylococcal *Myoviridae* phages are more closely related (in terms of gene content and organization) to modules found in staphylococcal *Siphoviridae* phages than to modules of *Myoviridae* found in different bacterial species.

Figure 1. Modular organization of *Staphylococci* phages genomes (**a**) *Siphoviridae* genomes. Colored boxes represent the five functional modules found in *Siphoviridae* genomes. Red: lysogeny, yellow: DNA metabolism, green: DNA packaging and capsid morphogenesis, blue: tail morphogenesis, pink: cell host lysis. Virulence genes (purple) are generally found downstream the lysis module, or inserted between the lysogeny and DNA metabolism module. A closer view of the DNA packaging and capsid morphogenesis shows the structural genes pattern typical of the Sfi21- and Sfi11- like phages genera (see text for details); (**b**) *Podoviridae* genomes. Lysis module (pink) and tail morphogenesis modules (blue) are overlapping. DNA metabolism genes (*i.e.*, single-strand DNA binding protein and DNA polymerase; green) are located in a region of genes of unknown function (gray), upstream to the tail module. An encapsidation protein is encoded next to the DNA polymerase in the staphylococcal *Podoviridae* genomes described so far. (**c**) *Myoviridae* genomes. The phage Twort genome is represented as an example. A large region encodes genes of unknown function (gray). DNA metabolism genes (yellow) are distributed in two distinct modules (known as replication modules), as well as lysis genes (pink) that are found upstream and downstream to DNA packaging and capsid (green) and tail modules (blue). TerS: small subunit terminase, TerL: large subunit terminase, Port: portal protein,

Prot: protease, MHP: major capsid protein, H: capsid morphogenesis protein, mHP: minor capsid protein, SS-DNA binding: single strand DNA binding protein.

(a)

Siphoviridae
genome size: 40kb

(b)
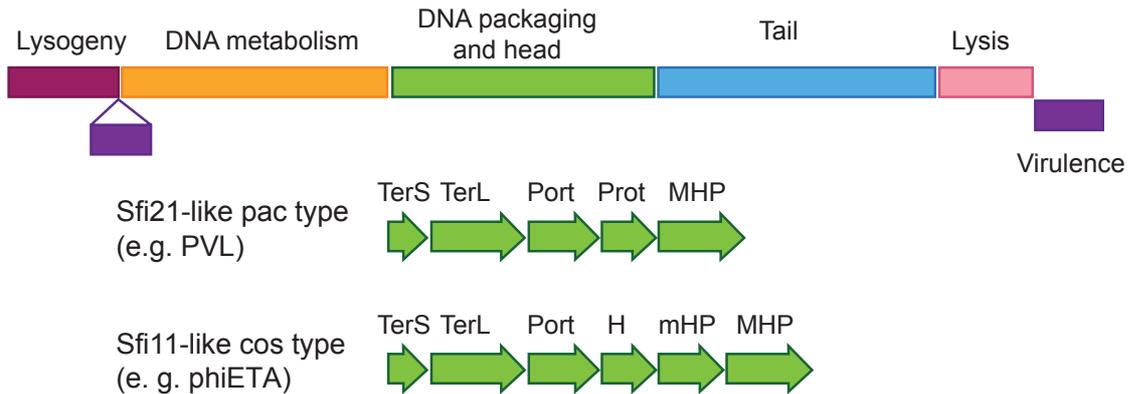
Podoviridae
genome size: <20kb

(c)
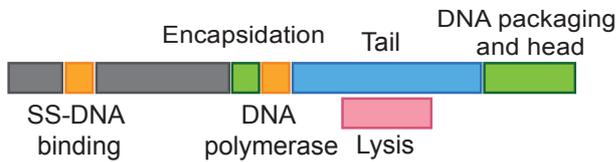
Myoviridae
genome size: >125kb
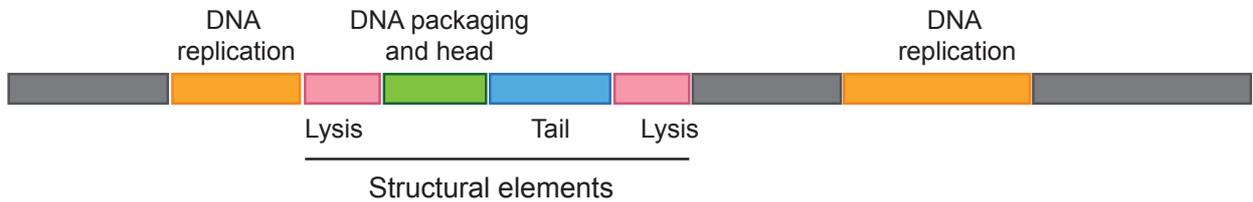
2.2. Role of Phages in S. aureus Pathogenesis

The *S. aureus* genome is mainly composed of a backbone of genes that are conserved among the different strains, both in terms of sequence and synteny [28,29]. These genes constitute the 'core genome' in opposition to the 'accessory genome', which is variable between strains and constituted by

integrated plasmids, transposons, genomic islands, pathogenicity islands (saPIs) and prophages. This 'accessory genome' may represent up to 25% of *S. aureus* genome and largely contributes to the high genetic and phenotypic plasticity of the pathogen. Indeed, one of the remarkable characteristics of *S. aureus* is represented by its versatility and ability to adapt to diverse and hostile environments [1,6, 28,30,31], and phages are playing an essential role in this phenomenon.

2.2.1. Phages-encoded Virulence Factors

Phages encode a large proportion of *S. aureus* virulence factors and provide the pathogen with a large variety of toxins, mainly allowing escaping host immune system (reviewed in [1,6,31]). Many factors have been described and characterized, such as the widespread immune modulator staphylokinase (*sak*) responsible for host tissue destruction, the chemotaxis inhibitory protein CHIP (*chp*), the staphylococcal inhibitor of complement SCIN (*scn*) and several superantigens (*sea*, *seg*, *sek*, *sek2*, *sep*, *seq*). These superantigens are enterotoxins causing food poisoning, toxic shock syndrome and necrotizing fasciitis. In addition, the bi-component cytotoxin Panton-Valentine leukocidin (PVL, encoded by *lukF-PV*, *lukS-PV*) and related leukocidins (*lukM*, *lukF*)) form pores into leukocytes and cause necrotic infections. Finally, the exfoliative toxin A (*eta*) is involved in severe skin infections.

In general, phages are carrying single virulence factor genes, although some exceptions have been reported. For examples, phiSa3 phages and relatives, such as phiN315, may encode up to five virulence factors, which form an immune escape complex (IEC) [5,31–33]. Virulence factor genes are not strictly associated to a specific phage and appear to be exchanged by horizontal gene transfer and recombination [2,6,31] (see section 2.2.3). As mentioned above, virulence genes are often located near the attachment site (*att*) of the prophage, *i.e.*, adjacent to the host chromosome. A possible origin for their acquisition by phages might be aberrant excision events from an ancestral bacterial chromosome [34]. They might also derive from mobile genetic elements, as suggested by the presence of transposase genes flanking toxin genes as in phiPV83 [35]. Thus, phages clearly impact virulence by positive lysogenic conversion, since they provide novel functions and activities to the host. Negative conversion also occurs as prophage insertion might inactivate genes [2,4,31]. In most of the cases, both phenomena occur simultaneously. As an example, phiSa3-related phages integrate in beta-hemolysin gene, rendering lysogens defective for beta-hemolyse but effective for IEC production. The integration site relies on the specificity of the phage integrase [36,37].

Expression of phage-encoded virulence genes is maximal upon entry in the lytic cycle, since latent promoters are activated and phage genome is replicated, leading to an increase in genome copy number [5,38,39], although some expression is detected during lysogeny [34].

2.2.2. Phage-mediated Mobilization of Virulence Factors: SaPIs Pathogenicity Islands

Phages are the primary vehicles for horizontal transfer between *S. aureus* strains. They spread chromosomally-encoded virulence determinants through generalized transduction. In addition, they are also responsible for the mobilization of SaPIs, which encodes major toxin genes, such as the toxic

shock syndrome toxin 1 and other superantigens (reviewed in [40]). SaPIs are widespread in *S. aureus* genomes. They are discrete chromosomal DNA segments that have been acquired by horizontal transfer. They are not mobile by themselves and rely on a helper phage for moving. SaPIs are replicated and mobilized either in response to SOS-induced excision of a helper prophage present in the same strain, either following the infection by a helper phage or by the joint entry of SaPI and helper phage [40]. The underlying molecular mechanism for induction is the specific interaction of a SaPI repressor and a de-repressor encoded by the helper phage. Different proteins of a particular helper phage may be involved in induction of different SaPIs. In particular, phi80alpha is able to mobilize at least five different SaPIs [40–42]. Hence, SaPIs mobilization represents a remarkable example of evolutionary adaptation involving pathogenicity islands and phages [42].

2.2.3. Phage Dynamics Contribute to *S. aureus* Evolution and Pathogenesis

Additionally to gene transfer, phages contribute to genetic alterations during infection, providing the species with broad genetic variations [5,38,43,44,45]. In addition to phage acquisition or excision (lysogenic conversion); duplication, ectopic integration and stable extra-chromosomal form of phages have been reported, increasing genetic diversity within bacterial populations [36,44]. Hence, generating heterogeneity within a population upon infection offers different virulence potentials and provides the pathogen with the ability to develop a flexible response to host defenses. It was shown that phage mobilization and atypical genomic integration are favored in pathogenic strains upon infection conditions, compared to colonizing strains in healthy carriers [5,36,44]. Factors causing phage induction *in situ* are environmental conditions that lead to bacterial DNA damage, such as antibiotics treatments or reactive oxygen species released by macrophages. Phage-mediated phenotypic diversification acts in concert with other mechanisms for genetic variations, such as recombination and mutations, and is under the influence of external factors, such as the presence of co-infecting species [43,45].

3. What about Phages in Non-*aureus Staphylococci*?

Comparison of CoNS genomes with *S. aureus* genomes revealed the inter-species conservation of both sequence and synteny of a large proportion of genes (core genes) with variable regions carrying species-specific genes [13,46–49]. Phage-encoded virulence factors responsible for *S. aureus* pathogenesis are absent in CoNS, in correlation with the difference in pathogenesis-mediated by CoNS. However, toxin and antibiotic resistance genes have been identified in several of these species. Although mobilization of virulence genes by phages has not been demonstrated, it is conceivable that phages might play a role in pathogenesis and evolution of CoNS, such as observed for *S. aureus*. Accordingly, phages were shown to impact genetic variability in *S. epidermidis* [50]. Although phage prevalence in clinical isolates might be underestimated [51], CoNS genomes described so far contain only few prophages or genomic islands, if any.

Most of the CoNS phages belong to the *Siphoviridae* family, while several virulent *Podoviridae* phages have been recently isolated directly from human anterior nares [52]. As for *S. aureus* phages, the first interest brought to phage infecting these species relied on their use for clinical isolate typing [53,54]. Several studies report the isolation and characterization of phages from *S. epidermidis* and *S. saprophyticus* [52,55–57], but only nine phages and prophages have been sequenced and studied, both at the molecular and physiological levels (Table S2). Among these, five phages and prophages are from *S. epidermidis* [51,58,59], one prophage from *S. carnosus* [48], two phages from *S. hominis* [37] and one phage from *S. capitis* [37]. The *S. epidermidis* vB_SepiS_phiIPLA5 (Table S2) is strictly lytic due to a defective lysogeny module [51,57], and the *S. carnosus* phiTM300 prophage (Table S2) appears to have lost its mobility, as it could not be induced upon mitomycin C treatment [48]. These CoNS phages show the general genomic features described for the *S. aureus* phages (see above, section 2.1.2; Table S2). Their genome size range falls into the class II proposed by the Pelletier group [3], with high gene density and a typical genomic organization into five functional modules (lysogeny, DNA metabolism, DNA packaging and capsid morphogenesis, tail morphogenesis and lysis). As for *S. aureus* phages, a function could be only assigned to a small proportion of the ORFs (from 29% to 53%, depending on the phage). Known virulence determinants were not found in these genomes. Comparative analysis of genome sequences revealed that *S. epidermidis* vB_SepiS_phiIPLA5, vB_SepiS_phiIPLA7, phiPH15 and phiCNPH82 are closely related, while phi909 showed a high similarity to *S. aureus* phages [48,51]. Interestingly, we also detected close relationships between *S. aureus* and CoNS phages (see below, section 4) [37].

Genomic islands related to SaPIs elements were identified in *S. haemolyticus* [46] and *S. saprophyticus* [13], although they usually lack superantigen-encoding genes. The first CoNS superantigen-bearing genomic island was recently described in *S. epidermidis* [59]. SePI-I encodes the staphylococcal enterotoxin C3 (SEC3) and enterotoxin-like toxin L (SEIL). Interestingly, the *seil* gene is homologous to those of *S. aureus*, indicating horizontal transfer events between *Staphylococci* species.

4. Classification and Evolution of the Staphylococcal Phage Family

4.1. Classification of Staphylococci Phages, a Long-Term Challenge

Early classifications proposed for *S. aureus* phages were based on their lytic properties, serotypes, morphology, on number and size of virion proteins and on genome size and organization as revealed through DNA hybridization or endonuclease restriction patterns ([18,60–63]; reviewed in [26]). More recently, progress in genomics and bio-informatics have allowed alternative classification methods.

Comparative genomic studies led to the subdivision of the morphological families into sub-families and genera. Following a classification proposed by Brussow and Desiere (although still not recognized by ICTV (International Comity on Taxonomy of Viruses)), *Siphoviridae* phages of low GC gram-positive bacteria, including *Staphylococci,* are categorized as Sfi21- or Sfi11-like phages by several authors [64] (Tables S1 and S2). This distinction is based on the capsid genes pattern as

reported for *Streptococcus thermophilus* phages [65]. Sfi21-like phages share characteristic features over the capsid region with *E. coli* HK97 phage and use a similar cos-site based strategy for DNA packaging (Figure 1a). In *S. aureus*, Sfi21-like genus is subdivided into three groups: the first being represented by phiPVL-phiPV83-phi13-phiSa3mw phages, the second by phiSLT-phiSa2mw-phi12 and the third by phiMu50A-phiN315 [4] (Table S1). While showing different capsid morphology (isometric or elongated), these phages encode typical capsid gene pattern of Sfi21-like phages (portal protein-protease-major capsid protein). Sfi11-like *pac*-type phages are related to the *B. subtilis* SPP1 phage and differ from Sfi21-like phages, notably by the lack of the protease-encoding gene (Figure 1a) [64, 65].

The two other families of the *Caudovirales* order were recently reassessed on the basis of protein similarities [66,67]. Among *Podoviridae*, *S. aureus* phages constitute a novel genus called 44AHJD-like. This genus belongs to the *Picovirinae* sub-family that also includes the phi29-like genus represented by the well-described *Bacillus* phi29 [67]. Among *Myoviridae*, the Twort phage is a representative of the Twort-like genus within the *Spounaviridae* sub-family [66]. This sub-family also includes the *Bacillus subtilis* SPO1 phage and relatives, as well as the *Lactobacillus plantarum* LP65 phage [66,68].

Other classification approaches based on specific marker genes found within *S. aureus Siphoviridae* genomes have been proposed [29,63,69,70]. They rely on PCR detection of genes representative of different phage types or categories. In addition to providing classification schemes, these methods are useful for *S. aureus* prophage detection, which is of great interest for epidemiological studies. The markers encompass genes coding for structural components, such as tail fibers, capsid proteins [63] or integrase genes [29,36]. In the latter case, classification correlates with distinct integration sites into the host chromosome. A good correlation between the type of integrase and virulence determinants has also been reported [36]. However, these methods do not provide information about mosaic structure, although detection of representatives of each functional module might provide some clues about mosaicism [70].

Recently, the group of Pelletier [3] proposed a classification taking into account the genome size, the gene organization, in addition to comparative nucleotide and protein sequence analysis. Using comparative genomic, Class II (*Siphoviridae*) was divided into three clades (A-C), and a new clade (D) was added later on by the Fischetti group after including *S. epidermidis* and additional *S. aureus* phage sequences [58]. Our group has refined this classification by extending the analysis to 85 phage and prophage genomes, among which 15 originated from CoNS [37]. The approach was based on the similarity of protein repertoires using tree-like and network-like methods [71,72]. Both methods established nine distinct clusters (data obtained with the network-like method is shown in Figure 2). Seven of these clusters (1–6, 9) are composed of *Siphoviridae* and constitute class II, according to the Pelletier classification. Clusters 7 and 8 constitute class III and I and are composed of *Myoviridae* and *Podoviridae*, respectively. They are unrelated to the other clusters. Within class II, our analysis generates seven related clusters, instead of the four clades previously proposed by the Pelletier group. Most importantly, one cluster (cluster 9) is composed exclusively of non-*S. aureus* phages and

constitutes an entirely new clade, as compared to the earlier classification. Four of the seven clusters are composed exclusively of *S. aureus* phages (clusters 2, 4, 5 and 6). Interestingly, the two last clusters (clusters 1 and 3) are composed of *S. aureus* and non-*S. aureus* phages, revealing close relationships between phages of different *Staphylococci* species.

Figure 2. Network representation of relationships between *Staphylococci* phages based on protein content (adapted from [37]). Circles represent the nine different clusters defined by Markov cluster algorithm (MCL). The color indicates the host species (magenta: *S. aureus*; purple: *S. aureus* and CoNS; blue: CoNS). The number of genomes is indicated into brackets. Cluster 8 corresponds to the class I (*Podoviridae*), clusters 1–6 and 9 to class II (*Siphoviridae*) and cluster 7 to class III (*Myoviridae*). Cluster 1 corresponds to clade A, cluster 4 to clade B, and clades C and D were split into two sub-clades (clusters 2 and 3, and 5 and 6, respectively). Cluster 9 constitutes a new clade. In this schematic representation, gray lines between distinct clusters indicate that at least 30% of homologous proteins are shared between at least two phage genomes. Following this analysis, two previously unclassified phages (2638A and 187) were included in cluster 6 and 1, respectively. PT1028 was not included in this analysis.
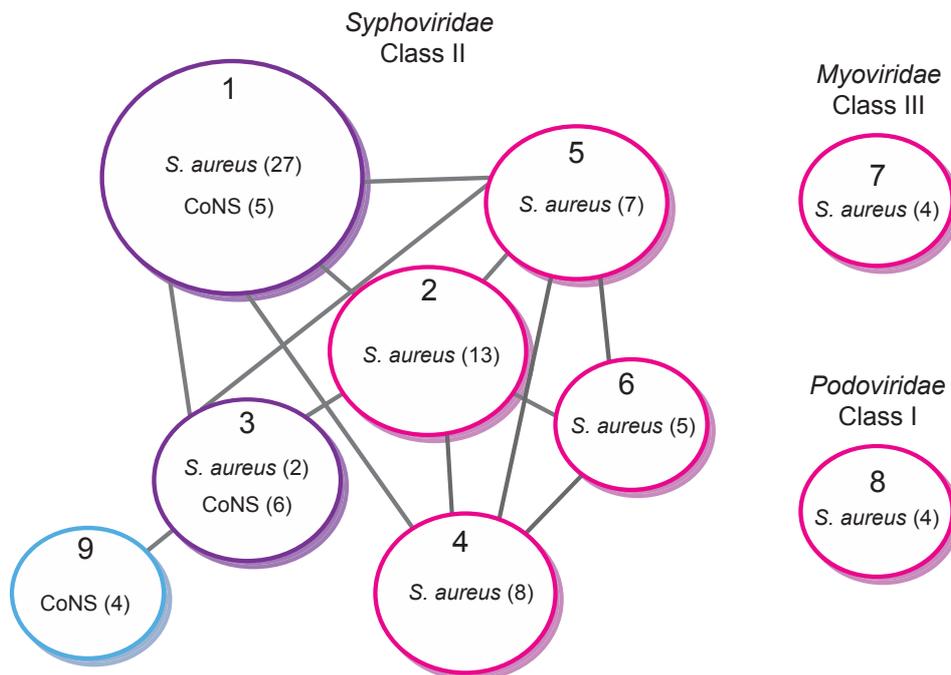
4.2. Modular Evolution of Staphylococci Phages

4.2.1. Extensive Genome Mosaicism in *Staphylococci* Phages

The mosaic gene organization is consistent with the theory of modular evolution based on module exchanges by horizontal transfer and recombination events [3,4,73]. This mosaicism can be viewed at either the nucleotide or amino acid level. At the nucleotide level, genome comparisons reveal exchange events that are likely to have occurred recently. Comparisons at the protein level identify homologous proteins that are shared by distantly related phages. They clearly derived from a common ancestor, but have diverged with time, and similarity at the nucleotide level is undetectable [73]. *S. aureus* phages are highly mosaic, indicating that gene exchange is common within this phage group [3,33,35,37,74–79]. Exchanges not only concern single genes, but also protein domains or a group of genes, such as functional modules [3,73]. Interestingly, *S. aureus* phages often share large, highly similar sequences with at least two other phages and with different phages along their genome. As an example, more than 80% of the ROSA genome is covered by large identical sequences from at least five other *S. aureus* phage genomes, with remarkable co-linearity. Exchange events are favored between phages of the same genome size range, which correspond to different morphological families (see above section 2.1.2) [79].

Different molecular mechanisms have been proposed to explain how these exchanges occur [23,73]. A first model invokes homologous recombination events between conserved sequences, which are generally found at gene borders, such as detected in lambdoid *E. coli* phages [80,81]. Recombination events require large homologous regions that are recognized either by host- or phage-encoded recombinases [80]. A second model proposes alternative recombination mechanisms, occurring randomly or between short sequences [82]. It is likely that these two mechanisms contribute to mosaicism generation. For instance, PVL-related phage genomes are composed of long regions that are shared between different phages, flanked by conserved sequences, called junctions, suitable for homologous recombination events. However, shuffling might occur at other regions, since short regions appear to be shared by phage genomes [35,74–76]. Even though these mechanisms certainly lead to a high percentage of non-viable phages, it is considered as a highly creative process that might provide countermeasures against anti-phage mechanisms evolved by bacteria, such as CRISPR sequences, toxin-antitoxin systems and restriction-modification [82].

4.2.2. Evolutionary Relationships within *Staphylococci* Phages

An interesting aspect of mosaicism is that comparison of a large number of genomes allows the establishment of phylogenetic relationships of specific regions instead of entire phages. Phylogenetic trees of representative genes revealed distinct evolutionary histories for different modules, thereby highlighting extensive mosaicism. This observation thus provides a tool to investigate evolution of *S. aureus* phages.

General clustering approaches, as used by our group to extend the classification of *Staphylococci* phages to CoNS species [37], are also useful to represent relationships among phage groups. In order to study the modular evolution of the CoNS StB12, StB27 and StB20 phages, families of homologous proteins were defined and the relative proportion of homologs in the nine clusters shown in Figure 2 was calculated. As previously shown [79], phages belonging to different morphological families do not share homologs. Our study confirms that morphogenesis module genes are in general less prone to horizontal swapping [3,73], although genes involved in host-phage interactions are notable exceptions. Clustering of *Staphylococcus Siphoviridae* phages relies mainly on structural features, indicating that within the overall similarities of virion structure, subtle variations allow differentiation. This is in accordance with a modular mechanism of evolution in which differentiation processes mainly rely on the exchange of a restricted number of genes or protein domains within structural modules [37]. Importantly, mosaicism encompasses genomes from both *Siphoviridae* phages from *S. aureus* and CoNS species, suggesting inter-species gene exchanges, which might be of crucial importance for *Staphylococci* pathogenesis.

4.3. Evolutionary Relationships between Staphylococci Phages and Other Species

Among the numerous approaches developed to outline evolutionary and functional relationships between phages [58,66,67,71,79,83–87], an original method developed by Lima-Mendez et al., [72,87] allowed them to study phage modular evolution on a large scale. The analysis was based on the clustering of phage proteins families in evolutionary conserved modules (ECMs) and the establishment of their distribution among phages. ECMs are defined as groups of protein that have a similar phylogenetic profile, meaning that they co-occur in genomes. In virulent phages, they are generally larger than functional modules, while in temperate phages, they tend to correspond to functional modules [87]. The clustering method allowed classification of phages into groups distinguished by different combination of ECMs. On the other hand, ECMs specific to particular phage groups can be identified. As an example, some modules were strictly associated with *Staphylococci* phages [72,87]. These ECMs encode virulence determinants, capsid and tail morphogenesis genes, or genes of unknown functions. The latter case is of special interest, since these ORFs of unknown functions might potentially be involved in phage/host interaction and may therefore include novel virulence factors or other proteins of medical relevance [72]. Other ECMs associated with *Staphylococci* phages were shared with *Streptococci* phages (modules involved in capsid and long non-contractile tail morphogenesis, phage integration/excision, replication or regulation functions), indicating an evolutionary link between these species.

Development of such methods is helpful for understanding phage evolution in general. These studies also complete the ICTV classification system based on morphology and specific molecular markers.

5. Use of *Staphylococci* Phages for Phage Therapy and Other Bio-Technological Applications

Due to the renewed interest for phage therapy, an increasing number of phages have been isolated and characterized for their potential use against *Staphylococci* infections, both in humans and animals (see notably [26,51,88–91]). Besides phage therapy, staphylococcal phages are also attractive candidates for food preservation [92–95]. Most phages selected for phage therapy or food preservation are strictly lytic because of the complications caused by lysogeny (e. g., resistance of lysogen strains to phage infection and unexpected transduction of host genes) [2,88,89]. The direct use of lysins is often an alternative for the use of entire phages (see notably [19,96,97]).

Phage candidates are mainly virulent phages belonging to *Myoviridae* (e. g., phage K, [97,98] or the recently identified phiStau2A, [99]) and *Podoviridae* (e. g., phiSAP-2, [19]). They were isolated from diverse environments, including dairy products, farm environments (milk from cow infected mastitis), humans and medical devices implanted into patients (see notably [16,19,100]).

Another important characteristic of phages selected for therapeutic and bio-technological applications is their narrow host range [89]. As an example, phage K appears to be specific to particular *S. aureus* clinical isolates, as well as to particular CoNS strains, while phiStau2 shows a larger host range among *S. aureus* clinical isolates, but is inefficient against the CoNS strains tested [99]. Therefore, to set up an efficient treatment, it is mandatory to precisely determine the bacterial species/isolate responsible for the infection, as well as disposing of a battery of phages of known host range. Traditionally, phage cocktails are used to overcome these issues [89,101]. On the other hand, this property allows specific treatment against pathogens without affecting the commensal flora.

Nevertheless, phage therapy remains promising as a large number of phages—more than 10^{31} phage particles are estimated in the biosphere—await discovery, opening the perspectives for novel therapeutic approaches [89].

6. Concluding Remarks

The large number of *Staphylococci* phages sequenced and characterized so far revealed an extensive mosaicism, which indicates gene shuffling between phages of *Staphylococci* species, including *S. aureus* and CoNS. An important aspect of *Staphylococci* phages is their pivotal role in *S. aureus* pathogenesis. Horizontal transfer of phages is an efficient way to rapidly disseminate virulence determinants among pathogens. Such transfers have been reported between *S. aureus* clinical isolates during infections [6]. CoNS pathogenesis is less understood, and virulence genes appear to be absent in CoNS phage genomes. However, CoNS phage diversity remains to be explored, and their function in CoNS pathogenesis is likely to be underestimated. Identification and characterization of novel phages from pathogenic and non-pathogenic CoNS strains is therefore of crucial importance to further understand the evolutionary relationships connecting *S. aureus* phages with phages found in non-pathogenic commensal *Staphylococci*.

154

Acknowledgments

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Lindsay, J.A. Genomic variation and evolution of Staphylococcus aureus. *Int. J. Med. Microbiol.* **2010**, *300*, 98–103.
2. Brussow, H.; Canchaya, C.; Hardt, W.D. Phages and the evolution of bacterial pathogens: From genomic rearrangements to lysogenic conversion. *Microbiol. Mol. Biol. Rev.* **2004**, *68*, 560–602.
3. Kwan, T.; Liu, J.; DuBow, M.; Gros, P.; Pelletier, J. The complete genomes and proteomes of 27 Staphylococcus aureus bacteriophages. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 5174-5179.
4. Canchaya, C.; Proux, C.; Fournous, G.; Bruttin, A.; Brussow, H. Prophage Genomics. *Microbiol. Mol. Biol. Rev.* **2003**, *67*, 238–276.
5. Goerke, C.; Wirtz, C.; Fluckiger, U.; Wolz, C. Extensive phage dynamics in Staphylococcus aureus contributes to adaptation to the human host during infection. *Mol. Microbiol.* **2006**, *61*, 1673–1685.
6. Feng, Y.; Chen, C.J.; Su, L.H.; Hu, S.; Yu, J.; Chiu, C.H. Evolution and pathogenesis of Staphylococcus aureus: lessons learned from genotyping and comparative genomics. *FEMS Microbiol. Rev.* **2008**, *32*, 23–37.
7. Wentworth, B.B. Bacteriophage Typing of the Staphylococci. *Bacteriol. Rev.* **1963**, *27*, 253–272.
8. Rosenblum, E.D.; Tyrone, S. Serology, Density, and Morphology of Staphylococcal Phages. *J. Bacteriol.* **1964**, *88*, 1737–1742.
9. Otto, M. Molecular basis of Staphylococcus epidermidis infections. *Semin. Immunopathol.* **2012**, *34*, 201–214.
10. Otto, M. Staphylococcus epidermidis--the 'accidental' pathogen. *Nat. Rev. Microbiol.* **2009**, *7*, 555–567.
11. Von Eiff, C.; Peters, G.; Heilmann, C. Pathogenesis of infections due to coagulase-negative staphylococci. *Lancet Infect. Dis.* **2002**, *2*, 677–685.
12. Frank, K.L.; Del Pozo, J.L.; Patel, R. From clinical microbiology to infection pathogenesis: How daring to be different works for Staphylococcus lugdunensis. *Clin. Microbiol. Rev.* **2008**, *21*, 111–133.

13. Kuroda, M.; Yamashita, A.; Hirakawa, H.; Kumano, M.; Morikawa, K.; Higashide, M.; Maruyama, A.; Inose, Y.; Matoba, K.; Toh, H.; *et al*. Whole genome sequence of Staphylococcus saprophyticus reveals the pathogenesis of uncomplicated urinary tract infection. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 13272–13277.

14. EMBL phage database. Genomes pages – Phages. Available online: http://www.ebi.ac.uk/genomes/phage.html (accessed on 30 October 2012).

15. Patric database. Staphylococcus page. Available online: http://www.patricbrc.org/portal/portal/patric/Taxon?cType=taxon&cId=1279 (accessed on accessed on 30 October 2012).

16. Garcia, P.; Martinez, B.; Obeso, J.M.; Lavigne, R.; Lurz, R.; Rodriguez, A. Functional genomic analysis of two Staphylococcus aureus phages isolated from the dairy environment. *Appl. Environ. Microbiol.* **2009**, *75*, 7663–7673.

17. Kim, M.S.; Myung, H. Complete Genome of Staphylococcus aureus Phage SA11. *J. Virol.* **2012**, *86*, 10232.

18. Ackermann, H.W. Tailed bacteriophages: the order caudovirales. *Adv. Virus Res.* **1998**, *51*, 135–201.

19. Son, J.S.; Lee, S.J.; Jun, S.Y.; Yoon, S.J.; Kang, S.H.; Paik, H.R.; Kang, J.O.; Choi, Y.J. Antibacterial and biofilm removal activity of a podoviridae Staphylococcus aureus bacteriophage SAP-2 and a derived recombinant cell-wall-degrading enzyme. *Appl. Microbiol. Biotechnol.* **2010**, *86*, 1439–1449.

20. O'Flaherty, S.; Coffey, A.; Edwards, R.; Meaney, W.; Fitzgerald, G.F.; Ross, R.P. Genome of Staphylococcal Phage K: a New Lineage of Myoviridae Infecting Gram-Positive Bacteria with a Low G+C Content. *J. Bacteriol.* **2004**, *186*, 2862–2871.

21. Pedulla, M.L.; Ford, M.E.; Houtz, J.M.; Karthikeyan, T.; Wadsworth, C.; Lewis, J.A.; Jacobs-Sera, D.; Falbo, J.; Gross, J.; Pannunzio, N.R.; Brucker, W.; Kumar, V.; Kandasamy, J.; Keenan, L.; Bardarov, S.; Kriakov, J.; Lawrence, J.G.; Jacobs, W.R., Jr.; Hendrix, R.W.; Hatfull, G.F. Origins of highly mosaic mycobacteriophage genomes. *Cell* **2003**, *113*, 171–182.

22. Kwan, T.; Liu, J.; Dubow, M.; Gros, P.; Pelletier, J. Comparative genomic analysis of 18 Pseudomonas aeruginosa bacteriophages. *J. Bacteriol.* **2006**, *188*, 1184–1187.

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

24. Bae, T.; Baba, T.; Hiramatsu, K.; Schneewind, O. Prophages of Staphylococcus aureus Newman and their contribution to virulence. *Mol. Microbiol.* **2006**, *62*, 1035–1047.

25. Vybiral, D.; Takac, M.; Loessner, M.; Witte, A.; von Ahsen, U.; Blasi, U. Complete nucleotide sequence and molecular characterization of two lytic Staphylococcus aureus phages: 44AHJD and P68. *FEMS Microbiol. Lett.* **2003**, *219*, 275–283.

26. Lobocka, M.; Hejnowicz, M.S.; Dabrowski, K.; Gozdek, A.; Kosakowski, J.; Witkowska, M.; Ulatowska, M.I.; Weber-Dabrowska, B.; Kwiatek, M.; Parasion, S.; *et al*. Genomics of staphylococcal Twort-like phages--potential therapeutics of the post-antibiotic era. *Adv. Virus Res.* **2012**, *83*, 143–216.

27. Krisch, H.M.; Comeau, A.M. The immense journey of bacteriophage T4--from d'Herelle to Delbruck and then to Darwin and beyond. *Res. Microbiol.* **2008**, *159*, 314–324.

28. Holden, M.T.; Feil, E.J.; Lindsay, J.A.; Peacock, S.J.; Day, N.P.; Enright, M.C.; Foster, T.J.; Moore, C.E.; Hurst, L.; Atkin, R.; *et al.* Complete genomes of two clinical Staphylococcus aureus strains: evidence for the rapid evolution of virulence and drug resistance. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 9786–9791.

29. Lindsay, J.A.; Holden, M.T. Staphylococcus aureus: superbug, super genome? *Trends Microbiol.* **2004**, *12*, 378–385.

30. Suzuki, H.; Lefebure, T.; Bitar, P.P.; Stanhope, M.J. Comparative genomic analysis of the genus Staphylococcus including Staphylococcus aureus and its newly described sister species Staphylococcus simiae. *BMC Genomics* **2012**, *13*, 38.

31. Malachowa, N.; DeLeo, F.R. Mobile genetic elements of Staphylococcus aureus. *Cell. Mol. Life Sci.* **2010**, *67*, 3057–3071.

32. Coleman, D.C.; Sullivan, D.J.; Russell, R.J.; Arbuthnott, J.P.; Carey, B.F.; Pomeroy, H.M. Staphylococcus aureus bacteriophages mediating the simultaneous lysogenic conversion of beta-lysin, staphylokinase and enterotoxin A: molecular mechanism of triple conversion. *J. Gen. Microbiol.* **1989**, *135*, 1679–1697.

33. van Wamel, W.J.; Rooijakkers, S.H.; Ruyken, M.; van Kessel, K.P.; van Strijp, J.A. The innate immune modulators staphylococcal complement inhibitor and chemotaxis inhibitory protein of Staphylococcus aureus are located on beta-hemolysin-converting bacteriophages. *J. Bacteriol.* **2006**, *188*, 1310–1315.

34. Wagner, P.L.; Waldor, M.K. Bacteriophage control of bacterial virulence. *Infect. Immun.* **2002**, *70*, 3985–3993.

35. Zou, D.; Kaneko, J.; Narita, S.; Kamio, Y. Prophage, phiPV83-pro, carrying panton-valentine leukocidin genes, on the Staphylococcus aureus P83 chromosome: comparative analysis of the genome structures of phiPV83-pro, phiPVL, phi11, and other phages. *Biosci. Biotechnol. Biochem.* **2000**, *64*, 2631–2643.

36. Goerke, C.; Pantucek, R.; Holtfreter, S.; Schulte, B.; Zink, M.; Grumann, D.; Broker, B.M.; Doskar, J.; Wolz, C. Diversity of prophages in dominant Staphylococcus aureus clonal lineages. *J. Bacteriol.* **2009**, *191*, 3462–3468.

37. Deghorain, M.; Bobay, L.M.; Smeesters, P.R.; Bousbata, S.; Vermeersch, M.; Perez-Morga, D.; Dreze, P.A.; Rocha, E.P.; Touchon, M.; van Melderen, L. Characterization of novel phages isolated in coagulase-negative Staphylococci reveals evolutionary relationships with *S. aureus* phages. *J. Bacteriol.* **2012**, *194*, 5829–5839.

38. Goerke, C.; Koller, J.; Wolz, C. Ciprofloxacin and trimethoprim cause phage induction and virulence modulation in Staphylococcus aureus. *Antimicrob. Agents Chemother.* **2006**, *50*, 171–177.

39. Sumby, P.; Waldor, M.K. Transcription of the toxin genes present within the Staphylococcal phage phiSa3ms is intimately linked with the phage's life cycle. *J. Bacteriol.* **2003**, *185*, 6841–6851.

40. Novick, R.P.; Christie, G.E.; Penades, J.R. The phage-related chromosomal islands of Gram-positive bacteria. *Nat. Rev. Microbiol.* **2010**, *8*, 541–551.

41. Lindsay, J.A.; Ruzin, A.; Ross, H.F.; Kurepina, N.; Novick, R.P. The gene for toxic shock toxin is carried by a family of mobile pathogenicity islands in Staphylococcus aureus. *Mol. Microbiol.* **1998**, *29*, 527–543.

42. Tormo-Mas, M.A.; Mir, I.; Shrestha, A.; Tallent, S.M.; Campoy, S.; Lasa, I.; Barbe, J.; Novick, R.P.; Christie, G.E.; Penades, J.R. Moonlighting bacteriophage proteins derepress staphylococcal pathogenicity islands. *Nature* **2010**, *465*, 779–782.

43. Goerke, C.; Wolz, C. Adaptation of Staphylococcus aureus to the cystic fibrosis lung. *Int. J. Med. Microbiol.* **2010**, *300*, 520–525.

44. Goerke, C.; Wolz, C. Regulatory and genomic plasticity of Staphylococcus aureus during persistent colonization and infection. *Int. J. Med. Microbiol.* **2004**, *294*, 195–202.

45. McAdam, P.R.; Holmes, A.; Templeton, K.E.; Fitzgerald, J.R. Adaptive evolution of Staphylococcus aureus during chronic endobronchial infection of a cystic fibrosis patient. *PLoS One* **2011**, *6*, e24301.

46. Takeuchi, F.; Watanabe, S.; Baba, T.; Yuzawa, H.; Ito, T.; Morimoto, Y.; Kuroda, M.; Cui, L.; Takahashi, M.; Ankai, A.; *et al*. Whole-genome sequencing of staphylococcus haemolyticus uncovers the extreme plasticity of its genome and the evolution of human-colonizing staphylococcal species. *J. Bacteriol.* **2005**, *187*, 7292–7308.

47. Gill, S.R.; Fouts, D.E.; Archer, G.L.; Mongodin, E.F.; Deboy, R.T.; Ravel, J.; Paulsen, I.T.; Kolonay, J.F.; Brinkac, L.; Beanan, M.; *et al*. Insights on evolution of virulence and resistance from the complete genome analysis of an early methicillin-resistant Staphylococcus aureus strain and a biofilm-producing methicillin-resistant Staphylococcus epidermidis strain. *J. Bacteriol.* **2005**, *187*, 2426–2438.

48. Rosenstein, R.; Nerz, C.; Biswas, L.; Resch, A.; Raddatz, G.; Schuster, S.C.; Gotz, F. Genome analysis of the meat starter culture bacterium Staphylococcus carnosus TM300. *Appl. Environ. Microbiol.* **2009**, *75*, 811–822.

49. Rosenstein, R.; Gotz, F. Genomic differences between the food-grade Staphylococcus carnosus and pathogenic staphylococcal species. *Int. J. Med. Microbiol.* **2010**, *300*, 104–108.

50. Lina, B.; Bes, M.; Vandenesch, F.; Greenland, T.; Etienne, J.; Fleurette, J. Role of bacteriophages in genomic variability of related coagulase-negative staphylococci. *FEMS Microbiol. Lett.* **1993**, *109*, 273–277.

51. Gutierrez, D.; Martinez, B.; Rodriguez, A.; Garcia, P. Genomic characterization of two Staphylococcus epidermidis bacteriophages with anti-biofilm potential. *BMC Genomics* **2012**, *13*, 228.

52. Aswani, V.; Tremblay, D.M.; Moineau, S.; Shukla, S.K. Staphylococcus epidermidis bacteriophages from the anterior nares of humans. *Appl. Environ. Microbiol.* **2011**, *77*, 7853–7855.

53. Boussard, P.; Pithsy, A.; Devleeschouwer, M.J.; Dony, J. Phage typing of coagulase-negative staphylococci. *J. Clin. Pharm. Ther.* **1992**, *17*, 165–168.

54. Barcs, I.; Herendi, A.; Lipcsey, A.; Bognar, C.; Hashimoto, H. Phage pattern and antibiotic resistance pattern of coagulase-negative staphylococci obtained from immunocompromised patients. *Microbiol. Immunol.* **1992**, *36*, 947–959.

55. Bes, M.; Ackermann, H.W.; Brun, Y.; Fleurette, J. Morphology of Staphylococcus saprophyticus bacteriophages. *Res. Virol.* **1990**, *141*, 625–635.

56. Bes, M. Characterization of thirteen Staphylococcus epidermidis and S. saprophyticus bacteriophages. *Res. Virol.* **1994**, *145*, 111–121.

57. Gutierrez, D.; Martinez, B.; Rodriguez, A.; Garcia, P. Isolation and characterization of bacteriophages infecting Staphylococcus epidermidis. *Curr. Microbiol.* **2010**, *61*, 601–608.

58. Daniel, A.; Bonnen, P.E.; Fischetti, V.A. First complete genome sequence of two Staphylococcus epidermidis bacteriophages. *J. Bacteriol.* **2007**, *189*, 2086–2100.

59. Madhusoodanan, J.; Seo, K.S.; Remortel, B.; Park, J.Y.; Hwang, S.Y.; Fox, L.K.; Park, Y.H.; Deobald, C.F.; Wang, D.; Liu, S.; *et al*. An Enterotoxin-Bearing Pathogenicity Island in Staphylococcus epidermidis. *J. Bacteriol.* **2011**, *193*, 1854–1862.

60. Ackermann, H.W.; DuBow, M.S.; Jarvis, A.W.; Jones, L.A.; Krylov, V.N.; Maniloff, J.; Rocourt, J.; Safferman, R.S.; Schneider, J.; Seldin, L.; *et al*. The species concept and its application to tailed phages. *Arch. Virol.* **1992**, *124*, 69–82.

61. Lee, J.S.; Stewart, P.R. The virion proteins and ultrastructure of Staphylococcus aureus bacteriophages. *J. Gen. Virol.* **1985**, *66*, 2017–2027.

62. Stewart, P.R.; Waldron, H.G.; Lee, J.S.; Matthews, P.R. Molecular relationships among serogroup B bacteriophages of Staphylococcus aureus. *J. Virol.* **1985**, *55*, 111–116.

63. Pantucek, R.; Doskar, J.; Ruzickova, V.; Kasparek, P.; Oracova, E.; Kvardova, V.; Rosypal, S. Identification of bacteriophage types and their carriage in Staphylococcus aureus. *Arch. Virol.* **2004**, *149*, 1689–1703.

64. Brussow, H.; Desiere, F. Comparative phage genomics and the evolution of Siphoviridae: Insights from dairy phages. *Mol. Microbiol.* **2001**, *39*, 213–222.

65. Le Marrec, C.; van Sinderen, D.; Walsh, L.; Stanley, E.; Vlegels, E.; Moineau, S.; Heinze, P.; Fitzgerald, G.; Fayard, B. Two groups of bacteriophages infecting Streptococcus thermophilus can be distinguished on the basis of mode of packaging and genetic determinants for major structural proteins. *Appl. Environ. Microbiol.* **1997**, *63*, 3246–3253.

66. Lavigne, R.; Darius, P.; Summer, E.J.; Seto, D.; Mahadevan, P.; Nilsson, A.S.; Ackermann, H.W.; Kropinski, A.M. Classification of Myoviridae bacteriophages using protein sequence similarity. *BMC Microbiol.* **2009**, *9*, 224.

67. Lavigne, R.; Seto, D.; Mahadevan, P.; Ackermann, H.W.; Kropinski, A.M. Unifying classical and molecular taxonomic classification: analysis of the Podoviridae using BLASTP-based tools. *Res. Microbiol.* **2008**, *159*, 406–414.

68. Chibani-Chennoufi, S.; Dillmann, M.L.; Marvin-Guy, L.; Rami-Shojaei, S.; Brussow, H. Lactobacillus plantarum bacteriophage LP65: a new member of the SPO1-like genus of the family Myoviridae. *J. Bacteriol.* **2004**, *186*, 7069–7083.

69. Grossi, P.A. Early appropriate therapy of Gram-positive bloodstream infections: the conservative use of new drugs. *Int. J. Antimicrob. Agents* **2009**, *34*, Suppl. 4, S31–34.

70. Kahankova, J.; Pantucek, R.; Goerke, C.; Ruzickova, V.; Holochova, P.; Doskar, J. Multilocus PCR typing strategy for differentiation of Staphylococcus aureus siphoviruses reflecting their modular genome structure. *Environ. Microbiol.* **2010**, *12*, 2527–2538.

71. Rohwer, F.; Edwards, R. The Phage Proteomic Tree: a genome-based taxonomy for phage. *J. Bacteriol.* **2002**, *184*, 4529–4535.

72. Lima-Mendez, G.; Toussaint, A.; Leplae, R. A modular view of the bacteriophage genomic space: identification of host and lifestyle marker modules. *Res. Microbiol.* **2011**, *162*, 737–746.

73. Hatfull, G.F.; Hendrix, R.W. Bacteriophages and their genomes. *Curr. Opin. Virol.* **2011**, *1*, 298–303.

74. Ma, X.X.; Ito, T.; Kondo, Y.; Cho, M.; Yoshizawa, Y.; Kaneko, J.; Katai, A.; Higashiide, M.; Li, S.; Hiramatsu, K. Two different Panton-Valentine leukocidin phage lineages predominate in Japan. *J. Clin. Microbiol.* **2008**, *46*, 3246–3258.

75. Ma, X.X.; Ito, T.; Chongtrakool, P.; Hiramatsu, K. Predominance of clones carrying Panton-Valentine leukocidin genes among methicillin-resistant Staphylococcus aureus strains isolated in Japanese hospitals from 1979 to 1985. *J. Clin. Microbiol.* **2006**, *44,* 4515–4527.

76. Narita, S.; Kaneko, J.; Chiba, J.; Piemont, Y.; Jarraud, S.; Etienne, J.; Kamio, Y. Phage conversion of Panton-Valentine leukocidin in Staphylococcus aureus: molecular analysis of a PVL-converting phage, phiSLT. *Gene* **2001**, *268*, 195–206.

77. Iandolo, J.J.; Worrell, V.; Groicher, K.H.; Qian, Y.; Tian, R.; Kenton, S.; Dorman, A.; Ji, H.; Lin, S.; Loh, P.; *et al.* Comparative analysis of the genomes of the temperate bacteriophages phi 11, phi 12 and phi 13 of Staphylococcus aureus 8325. *Gene* **2002**, *289*, 109–118.

78. Zhang, M.; Ito, T.; Li, S.; Jin, J.; Takeuchi, F.; Lauderdale, T.L.; Higashide, M.; Hiramatsu, K. Identification of the third type of PVL phage in ST59 methicillin-resistant Staphylococcus aureus (MRSA) strains. *FEMS Microbiol. Lett.* **2011**, *323*, 20–28.

79. Belcaid, M.; Bergeron, A.; Poisson, G. Mosaic graphs and comparative genomics in phage communities. *J. Comput. Biol.* **2010**, *17*, 1315–1326.

80. Susskind, M.M.; Botstein, D. Molecular genetics of bacteriophage P22. *Microbiol. Rev.* **1978**, *42*, 385–413.

81. Clark, A.J.; Inwood, W.; Cloutier, T.; Dhillon, T.S. Nucleotide sequence of coliphage HK620 and the evolution of lambdoid phages. *J. Mol. Biol.* **2001**, *311*, 657–679.

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

160

83. Hatfull, G.F.; Pedulla, M.L.; Jacobs-Sera, D.; Cichon, P.M.; Foley, A.; Ford, M.E.; Gonda, R.M.; Houtz, J.M.; Hryckowian, A.J.; Kelchner, V.A.; *et al*. Exploring the mycobacteriophage metaproteome: phage genomics as an educational platform. *PLoS Genet.* **2006**, *2*, e92.

84. Glazko, G.; Makarenkov, V.; Liu, J.; Mushegian, A. Evolutionary history of bacteriophages with double-stranded DNA genomes. *Biol. Direct.* **2007**, *2*, 36.

85. Huson, D.H.; Bryant, D. Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* **2006**, *23*, 254–267.

86. Lawrence, J.G.; Hatfull, G.F.; Hendrix, R.W. Imbroglios of viral taxonomy: genetic exchange and failings of phenetic approaches. *J. Bacteriol.* **2002**, *184*, 4891–4905.

87. Lima-Mendez, G.; Van Helden, J.; Toussaint, A.; Leplae, R. Reticulate representation of evolutionary and functional relationships between phage genomes. *Mol. Biol. Evol.* **2008**, *25*, 762–777.

88. Mann, N.H. The potential of phages to prevent MRSA infections. *Res. Microbiol.* **2008**, *159*, 400–405.

89. Lu, T.K.; Koeris, M.S. The next generation of bacteriophage therapy. *Curr. Opin. Microbiol.* **2011**, *14*, 524–531.

90. Matsuzaki, S.; Yasuda, M.; Nishikawa, H.; Kuroda, M.; Ujihara, T.; Shuin, T.; Shen, Y.; Jin, Z.; Fujimoto, S.; Nasimuzzaman, M.D.; *et al*. Experimental protection of mice against lethal Staphylococcus aureus infection by novel bacteriophage phi MR11. *J. Infect. Dis.* **2003**, *187*, 613–624.

91. Kwiatek, M.; Parasion, S.; Mizak, L.; Gryko, R.; Bartoszcze, M.; Kocik, J. Characterization of a bacteriophage, isolated from a cow with mastitis, that is lytic against Staphylococcus aureus strains. *Arch. Virol.* **2012**, *157*, 225–234.

92. Hagens, S.; Loessner, M.J. Bacteriophage for biocontrol of foodborne pathogens: calculations and considerations. *Curr. Pharm. Biotechnol.* **2010**, *11*, 58–68.

93. Bueno, E.; Garcia, P.; Martinez, B.; Rodriguez, A. Phage inactivation of Staphylococcus aureus in fresh and hard-type cheeses. *Int. J. Food Microbiol.* **2012**, *158*, 23–27.

94. Garcia, P.; Madera, C.; Martinez, B.; Rodriguez, A.; Evaristo Suarez, J. Prevalence of bacteriophages infecting Staphylococcus aureus in dairy samples and their potential as biocontrol agents. *J. Dairy Sci.* **2009**, *92*, 3019–3026.

95. Garcia, P.; Martinez, B.; Obeso, J.M.; Rodriguez, A. Bacteriophages and their application in food safety. *Lett. Appl. Microbiol.* **2008**, *47*, 479–485.

96. Fischetti, V.A. Bacteriophage endolysins: a novel anti-infective to control Gram-positive pathogens. *Int. J. Med. Microbiol.* **2010**, *300*, 357–362.

97. O'Flaherty, S.; Coffey, A.; Meaney, W.; Fitzgerald, G.F.; Ross, R.P. The recombinant phage lysin LysK has a broad spectrum of lytic activity against clinically relevant staphylococci, including methicillin-resistant Staphylococcus aureus. *J. Bacteriol.* **2005**, *187*, 7161–7164.

98. O'Flaherty, S.; Ross, R.P.; Meaney, W.; Fitzgerald, G.F.; Elbreki, M.F.; Coffey, A. Potential of the polyvalent anti-Staphylococcus bacteriophage K for control of antibiotic-resistant staphylococci from hospitals. *Appl. Environ. Microbiol.* **2005**, *71*, 1836–1842.

99. Hsieh, S.E.; Lo, H.H.; Chen, S.T.; Lee, M.C.; Tseng, Y.H. Wide host range and strong lytic activity of Staphylococcus aureus lytic phage Stau2. *Appl. Environ. Microbiol.* **2011**, *77*, 756–761.

100. O'Flaherty, S.; Ross, R.P.; Flynn, J.; Meaney, W.J.; Fitzgerald, G.F.; Coffey, A. Isolation and characterization of two anti-staphylococcal bacteriophages specific for pathogenic Staphylococcus aureus associated with bovine infections. *Lett. Appl. Microbiol.* **2005**, *41*, 482–486.

101. Merabishvili, M.; Pirnay, J.P.; Verbeken, G.; Chanishvili, N.; Tediashvili, M.; Lashkhi, N.; Glonti, T.; Krylov, V.; Mast, J.; van Parys, L.; *et al*. Quality-controlled small-scale production of a well-defined bacteriophage cocktail for use in human clinical trials. *PLoS One* **2009**, *4*, e4944.

Review

Structural Aspects of the Interaction of Dairy Phages with Their Host Bacteria

Jennifer Mahony [1] **and Douwe van Sinderen** [1,2,*]

[1] Department of Microbiology, University College Cork, Western Road, Cork, Ireland;
     E-Mail: j.mahony@ucc.ie
[2] Alimentary Pharmabiotic Centre, Biosciences Institute, University College Cork, Western Road,
     Cork, Ireland

* Author to whom correspondence should be addressed: E-Mail: d.vansinderen@ucc.ie;
   Tel.: +353-21-4901365; Fax: +353-21-4903101.

Abstract: Knowledge of phage-host interactions at a fundamental level is central to the design of rational strategies for the development of phage-resistant strains that may be applied in industrial settings. Phages infecting lactic acid bacteria, in particular *Lactococcus lactis* and *Streptococcus thermophilus*, negatively impact on dairy fermentation processes with serious economic implications. In recent years a wealth of information on structural protein assembly and topology has become available relating to phages infecting *Escherichia coli*, *Bacillus subtilis* and *Lactococcus lactis*, which act as models for structural analyses of dairy phages. In this review, we explore the role of model tailed phages, such as T4 and SPP1, in advancing our knowledge regarding interactions between dairy phages and their hosts. Furthermore, the potential of currently investigated dairy phages to in turn serve as model systems for this particular group of phages is discussed.

Key words: bacteriophage; milk fermentation; receptor

1. Introduction

Phages are the most abundant biological entities on Earth [1] and are responsible, at least in part, for driving the evolution of their bacterial hosts [2]. The selective pressure imposed by phages on their hosts also requires an equivalent genomic plasticity and adaptability by the phages themselves in order to successfully produce progeny in a host that tries to escape its parasites. All characterized phages known to infect lactic acid bacteria (LAB) possess a tail and therefore this review will solely consider tailed phages.

One of the first physical phage-host interactions is that between the receptor on the host bacterium and the distal end of the tail of the phage. Many isolated phages have been characterised morphologically by electron microscopy analysis [3] and the information derived from such analyses may be useful to gain an insight into the position and nature of host receptor (see below). In recent years structural information has emerged related to phage proteins that are involved in binding to these receptors, usually referred to as anti-receptors, receptor binding proteins (RBPs) or adhesins. Depending on the macromolecular nature of the receptor, *i.e.*, being a protein or carbohydrate, the phage tail tip of *Siphoviridae* phages, which contains the receptor-binding protein, appears to display a particular morphology. For example, phages that interact with protein moieties generally possess either a pointed or "stubby" end, while those that interact with carbohydrate moieties more often display a larger, more obvious structure, which is called the baseplate. While there are exceptions to this generalisation, including many undefined phage-host interactions, this may perhaps be the primary indicator of the type of moiety required by a phage to interact with its host. Here we will explore the role of electron microscopy and structural analysis in defining the interactions between phages and their hosts, and examine the various types of receptor-binding proteins encoded by phages and their corresponding host-encoded receptor. The interactions and wealth of structural data for phage RBPs that now exists will be examined with a view to understand the interactions between dairy phages and their hosts. Strains of *Lactococcus lactis* and *Streptococcus thermophilus* are the most widely utilised LAB in the dairy industry, with certain species of *Lactobacillus* also being used intensively. Consequently, phages infecting these species dominate LAB phage research. Since the most frequently isolated, and hence the most problematic phages of dairy lactic acid bacteria (LAB) are those belonging to the *Siphoviridae* (isolated for *Lactococcus*, *Streptococcus* and *Lactobacillus* spp., among others) and, to a lesser extent, *Myoviridae* families (isolated for *Lactobacillus* spp.), we will focus on the former for purposes of clarity.

2. Phages that Recognize a Proteinaceous Receptor

Phages recognizing a proteinaceous receptor can achieve a strong phage-bacterium interaction through protein-protein interactions and therefore only a single phage protein or small complex is required due to the high affinity, avidity and specificity of this interaction [4]. In electron micrographic images such phages possess a pointed or stubby tail tip, which is where the RBP is located (See Table 1 for examples). A very well established example of this is the *Bacillus subtilis*-infecting

siphophage SPP1, which interacts irreversibly with its host-encoded receptor protein YueB via its pointed tail tip following an initial and reversible attraction to a carbohydrate moiety on the cell surface [4]. Additionally, the lactococcal c2-type siphophages interact with the phage infection protein (PIP) through a RBP that is presumed to be located at the blunt-ended tail tip [5]. It is noteworthy that both YueB and PIP are members of the Type VII secretion system and are closely related proteins and therefore functional analogies are perhaps not surprising. The tail adsorption protein of c2 is proposed to be encoded by orf *l10*, a 75 kDa protein identified through mass spectrophotometry and immune-gold electron microscopy [6]. The bacterial receptors of the *Siphoviridae* coliphages λ and T5 (both of these phages possess pointed tail tips) have been defined and isolated (LamB and FhuA, respectively), which has permitted the development of *in vitro* DNA ejection assays to positively identify these receptor proteins [7,8]. The coliphages belonging to the *Myoviridae* T4 superfamily possess six long tail fibers that interact reversibly with lipopolysaccharide (LPS) or a porin protein, such as OmpC [9,10]. Subsequently, the baseplate is lowered toward the cell surface and the six short tail fibers are released to allow reversible binding to occur, thereby allowing the pointed puncturing device to initiate DNA injection into the host after the baseplate is switched to the star conformation [11]. Of the above-mentioned phages, T4 is one of the best studied model phages with respect to the structural analysis of the tail fibers and injection devices. The 3D structure of the long tail fiber receptor-binding domain of gp37 has been resolved, which revealed that the aromatic and positive residues on the surface of the fiber tips are most likely receptor-binding determinants with the receptor, although the authors do not specify whether the receptor material is LPS or the OmpC porin protein [9]. The proteins constituting the T4 baseplate complex have been studied by means of cryo-electron microscopy and X-ray crystallography [11–13]. Through these analyses, at least twelve proteins have been identified as baseplate components: gp5, gp6, gp7, gp8, gp9, gp10, gp11, gp12, gp25, gp27, gp53, gp54. The six gp10 trimers are a crucial component of the baseplate and structural analysis of this protein has revealed its role in extension of the short tail fibers encoded by gp12 by means of a partner shift: the pre-infection state sees an interaction between the C-terminal end of gp10 and gp12, while infection requires an interaction between the baseplate protein gp9 and gp10 so as to release the short tail fibers, thus permitting interaction with the host receptor [13]. This fascinating and complex phage serves as an excellent model for the physical interactions between a phage and its host. However, this model is much more sophisticated than many of the phages of the *Siphoviridae* family and for this purpose phage SPP1 may be a better representative for this family of phages.

Table 1. Proven and putative phage receptors.

Host	Phage	Receptor type	Host receptor	Reference
Bacillus subtilis	SPP1	Protein	YueB	[14]
Escherichia coli	T4-like	Protein	OmpC/LPS	[10]
	T5	Protein	FhuA	[8]
	Λ	Protein	LamB	[7]
Lactococcus lactis	c2	Protein	PIP	[5]
	p2	Saccharide	Unknown	[15]
	bIL170	Saccharide	Unknown	[16]
	TP901–1	Saccharide	Unknown	[17]
	Tuc2009	Saccharide	Unknown	
Lactobacillus	LL-H	Lipoteichoic acid	Poly-Glycerophosphate LTAs	[18]
Streptococcus	OBJ	Saccharide	Glucosamine/Ribose	[19]
thermophilus	CYM	Saccharide	Glucosamine/Rhamnose	[19]

3. SPP1 as a Model for Protein-Interacting Siphophages of Gram-Positive Bacteria

SPP1 possesses a long non-contractile tail, which is the trademark property of the *Siphoviridae* phages. At the base of the tail is a narrow device, the tail tip, which interacts with a membrane-anchored protein (YueB) on the host cell surface [14]. Significant structural data has emerged in recent years relating to the capsid, tail, head-tail connector and the tail tip of SPP1, thus rendering it a well understood model for other siphophages infecting Gram-positive bacteria, including those that target LAB [20–27].

One of the first structural studies of phage SPP1 focused on the tail at pre- and post-infection stages, and detailed the loss of the tail tip following DNA ejection [14]. The tail of SPP1 is composed of an estimated 40 stacked rings (gp17.1 and gp17.1*), inside which is a channel containing the tail tape measure protein (gp18) and through which the phage genome is passed upon release from the capsid [14]. The outside of the tail rings exhibit a six-fold symmetry (gp17.1 and gp17.1*, see below), while the inner tube has a continuous appearance that is composed of several gp18 subunits [14]. The tail tip, which has no channel to allow DNA passage, interacts with YueB, which prompts its release from the tail cap, thereby triggering a sequence of conformational changes that result in the ejection of the viral genome from the capsid. Following DNA ejection, the internal channel diameter appears to undergo a size reduction from 56 Å to 42 Å due to rearrangements in the major tail proteins gp17.1 and gp17.1* and not negating the fact that the tail tape measure protein fills the channel prior to DNA ejection [14]. These gp17.1 and gp17.1* proteins are present in a 3:1 ratio and protein microsequencing has demonstrated that the product of gp17.1* contains an additional C-terminal 10 kDa region, which is thought to be exposed on the outer surface of the tail [28]. This feature also seems to be encoded by prophages of *Bacillus licheniformis* and *Bacillus halodurans*, while such extensions also appear to be present in the phage receptor-binding proteins and major tail protein-associated sequences of other *Siphoviridae* phages such as the lactococcal phage Q54 and *Listeria* phage A118 [28].

Figure 1. Schematic representation of a top view of the re-organisation of the p2 (top) and SPP1 (bottom) baseplate regions before (right) and after (left) binding to the host cell. The p2 baseplate, composed of a hexamer of ORF15 (orange), a trimer of ORF16 (green) and six trimers of ORF18 (blue) which represents the RBP. In parallel, the SPP1 distal tail region is composed of two back-to-back hexamers of gp19.1 (orange) and a trimer of gp21 (green) bound to each of the Dit (gp19.1) hexamers, one trimer in the open conformation and the other in the closed conformation. A representation of the genomic regions that encode the tail structural components are presented in the lowermost section, including the genes encoding the proposed major tail protein (pink), the tail tape measure protein (purple), the distal tail protein (Dit)/hub protein (orange), the tail spike (SPP1) or tail tip (p2) (green) and the RBP (blue, p2). Those in grey encode non-structural proteins or have not been functionally assigned.
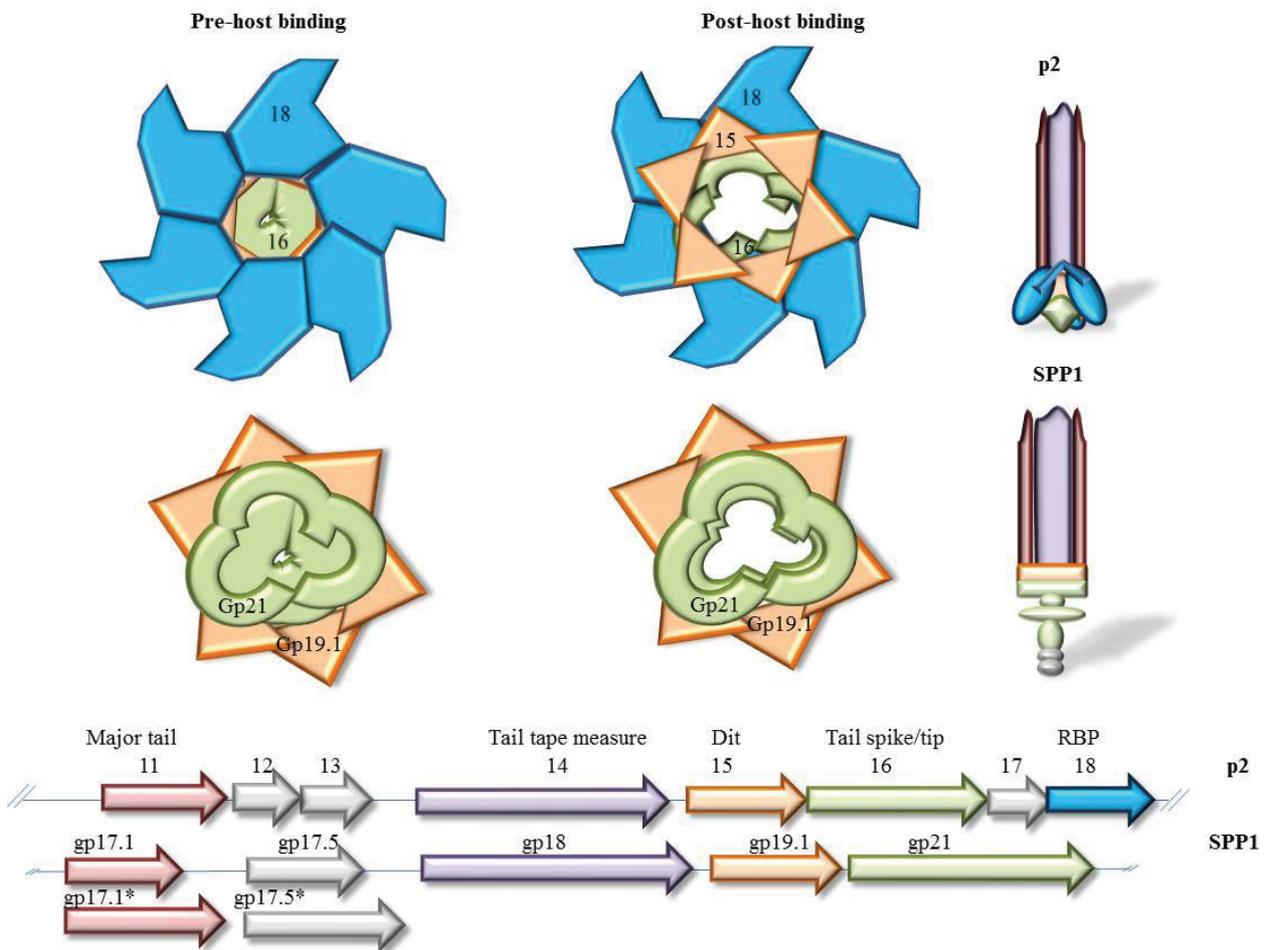
Other elements of the distal end of the tail of SPP1 have been characterised including the distal tail protein (Dit, gp19.1) and the N-terminal region of the tail spike protein (gp21) [23,25]. The SPP1 Dit appeared as a dimer of hexamers in a back-to-back orientation which is structurally similar to a component of the lactococcal phage p2 baseplate (ORF 15), however it is now clear that this apparent

dimer is a crystallographic artifact in the case of SPP1 [23]. Two trimers of the SPP1 truncated tail spike protein have been shown to interact with Dit, where one trimer is presumed to be in a "closed" conformation while the other trimer is in the "open" conformation (Figure 1) [25], a system which is consistent with observations made for the lactococcal phage p2 baseplate [25,29]. The Dit-tail spike complex appears to constitute the tail cap of SPP1 and the structural conservation of these features in lactococcal phages may indicate a common lineage of such systems [25]. Calcium ions stabilise the "open" conformation of tail cap, which is also consistent with the lactococcal phage p2 [25,29]. The Dit-tail complex acts as a hub upon which the baseplate (in the case of phages that recognize a saccharidic receptor) or the tail tip/spike (for phages that bind to a proteinaceous receptor) is hinged, at the distal end of the tail. Furthermore, (multimers of) the major tail protein (MTP) and tail tape measure protein (TMP) attach to this structure to form the tail tube. It is suggested that the interaction of the RBP with the host-encoded receptor induces tail base to change to the open conformation triggering the movement of the first ring of MTP which is then sequentially continued up through the MTP rings with ensuing release of DNA from the phage capsid [14,25]. The structural conservation of the Dit-tail cap complex and conformational arrangements in SPP1 and lactococcal phages endorses SPP1 as a model for *Siphoviridae* phages, and as such is a basis for understanding the interactions between dairy phages and their hosts.

4. Structural Approaches to Understand Dairy Phage Interactions and Carbohydrate-Recognizing Phages

Lactococcal phages have, among the dairy phages, received the most extensive scientific scrutiny with regards to the structural analysis of protein/protein complexes [17,29–32]. Isolated lactococcal phages are grouped into ten distinguishable species based on DNA homology and particle morphology [33]. Of these, the 936, c2 and P335 species are most frequently encountered in dairy facilities world-wide. Members of the 936 and P335 species are believed to bind a saccharide component that is present on the surface of the cell envelope, while the c2 phages are known to bind the so-called phage infection protein (PIP) [5]. Members of the 936 and P335 species have been studied at a structural level and constitute the backbone of emerging structural data relating to dairy phages.

4.1. Lactococcal P335 and 936 Phage-Host Interactions

Over the past decade, significant efforts have been made to determine the molecular nature and mechanism of the interactions between lactococcal phages and their hosts, and members of the P335 and 936 species have been central to such studies due to their industrial relevance. It has long been suggested that these phages recognize a carbohydrate moiety on their host, a notion that is consistent with carbohydrate inhibition assays [34]. Through mutational analyses several adjacent genes within an operon, which encodes the biosynthetic machinery for a cell envelope-associated polysaccharide (also called the pellicle), have been shown to be required for adsorption of 936-type phages. The involvement of this pellicle biosynthesis operon in phage infection has been demonstrated in two

separate studies for three 936-type phages, namely bIL170, 645 and sk1, although the precise role of the mutated genes as a phage receptor has not yet been established [35,36]. The 20–25 kb pellicle operon contains elements that are conserved among all sequenced lactococcal strains, while it also contains strain-specific regions that are expected to lead to the production of pellicles with a distinct saccharidic structure and composition. The so-called "adsorption genes", or receptor-encoding genes, involved in 936-type phage recognition and adsorption encode either a glycosyltransferase or a membrane-associated protein [35]. The involvement of these gene-products may imply two possibilities of the mode of interaction and DNA injection pathway of these phages. Firstly, that the glycosyltransferase inactivation may prevent the transfer and decoration of the required pellicle monosaccharide component to the surface thus negating the primary contact between the phage RBP and its saccharide receptor; or secondly, if primary contact is established and a membrane component is required for the translocation of the DNA across the membrane, the membrane-associated protein inactivation prevents internalisation of the phage DNA. Chemical mutagenesis of the lactococcal host strain 3107, which is sensitive to the P335 phages LC3 and TP901–1, resulted in the isolation of five genetically and phenotypically distinct mutants resistant to one or both phages [37]. This analysis revealed that the injection pathways employed by these phages of the same species are distinct. It is through studies such as these that more attention has been diverted towards the identification of phage RBPs on lactococcal phages in order to determine the relationship between a phage and its host(s).

The 936 phage RBPs has been identified through the isolation of a chimeric sk1/bIL170, which also was shown to exhibit a "host-range swapping" phenotype [38]. The identification of the RBPs of these phages presented the opportunity to define the structure of the corresponding proteins and their associated complexes. This study applied an approach previously used by Duplessis and Moineau (2001) in which the variable C-terminal VR2 region of ORF18 of DT1 was replaced by the corresponding fragment of MD4. The resultant chimeric DT1 infected the host for MD4, thereby elucidating the host specificity function of this protein [39]. Similarly, a chimeric derivative of phage TP901–1 (designated TP901–1C), harbouring the lower baseplate protein of phage Tuc2009 instead of its original TP901–1 equivalent, was shown to have acquired the host range of Tuc2009, thereby identifying the lower baseplate protein as the RBP of both of these P335 phages [40]. Such discoveries have been very useful in the assignment of RBPs of other P335 phages through homology and genomic positioning of the encoding gene [41]. Also noteworthy is the observation of the altered baseplate morphology in the case of the chimeric phage TP901–1C as compared to the parent phage [36]. TP901–1 has a clear double-disc baseplate while the chimeric phage TP901–1C was described to possess "hanging droplets" rather than a distinct lower baseplate and furthermore the lower disc appeared narrower than that of the parent phage (23 nm *versus* 28 nm in the parent phage) and intermediate between that of the wild-type TP901–1 and Tuc2009. The genetic plasticity of the lactococcal phages has permitted the development of recombinant phages of both the 936 and P335 species and may present a mechanism by which these phages alter their host specificity or enlarge their host spectrum as their environment dictates.

4.2. 936 Phage Baseplates

The first lactococcal phage RBP to be structurally defined was that of the 936 phage p2 [42]. The p2 RBP is a homotrimeric protein comprising three domains: the head, neck and shoulder. The actual receptor binding site was identified within the "head" domain through the application of the heavy chain neutralising llama antibody fragment $V_{HH}5$, which had previously been determined to prevent infection of the lactococcal host by phage p2 [15,42,43]. The shoulder domain possesses eight ß-strands interspersed with an α-helix between ß-strands one and two, and a coil between ß-strands two and three [27]. The neck is a ß-prism of three segments that yields a rigid structure to which analogy has been drawn to the gp12 short tail fiber structure of the coliphage T4. Finally, the head domain is a ß-barrel of seven anti-parallel ß-strands [42]. Seven charged or polar and five non-polar residues were identified in the RBP head domain interacting with the $V_{HH}5$.

The receptor binding site was further characterised through the isolation of escape mutants that could bypass the llama neutralising antibody fragments [15]. The majority of the mutations that enable the mutant phages to overcome the llama antibodies are within the interacting site. Just twelve out of fifty assessed lactococcal 936 phages were neutralised by $V_{HH}5$ antibodies, highlighting the apparent variability of RBPs found in members of this closely related phage species. Mutants overcoming the $V_{HH}5$ were readily isolated for eleven of the twelve phages neutralised and in all cases a single point mutation in the corresponding RBP-encoding gene was shown to be sufficient to prevent antibody recognition. This study also served to improve the resolution of the RBP structure of p2 from 2.3 to 1.7 Å and identified a glycerol molecule bound to the structure originating from the cryoprotectant. The binding of the glycerol molecule was observed to be tighter in the head domain than the shoulder domain where three hydrogen bonds are formed between head domain residues and the glycerol O-1 and O-2 atoms (*versus* one in the shoulder domain). Other saccharides were assessed for their ability to bind the RBP in solution and it was suggested that phosphoglycerol, which is a component of lipoteichoic acid and teichoic acid, represented a potential receptor. This was the first direct proof of carbohydrate binding of a lactococcal 936-type phage RBP, confirming biological analyses that suggested the role of specific carbohydrates as receptors for 936-type phages [15,34].

The crystal structure of the RBP head domain of another 936 phage, bIL170, was solved and represents one of a number of phages that were not neutralised by the $V_{HH}5$ antibodies due to its variable C-terminal sequence [16]. The phylogeny of the RBPs of 936 phages has demonstrated clear links to host range and $V_{HH}5$ antibody-neutralisation, and it is well established that while the N-terminal region is well-conserved, it is the C-terminal (head) domain that determines the specific host interaction [15,45–47]. Interestingly, while there is little sequence conservation between the head domains of p2 and bIL170, their structure is strikingly similar [16]. Furthermore, it bears structural similarity to the head domain of the lactococcal P335 phage TP901–1 [16]. Akin to the RBP of p2, the head domain of bIL170 is proposed to be a homotrimer that binds saccharides near a tyrosine residue (Tyr226) whose side chain forms the upper wall of the crevice between the trimer components and is accessible to solvents [16].

Expression and purification of ORFs 15, 16, 17 and 18 of p2 by affinity chromatography revealed that the baseplate of p2 is composed of three proteins, ORFs 15, 16 and 18. The structural data generated from crystals, obtained from the *in vitro* expressed recombinant baseplate complex, was mapped onto the baseplate structure, generated from electron microscopy images of the virion, provided comparative data to verify the accuracy of the expressed baseplate complex structure [29]. ORF15 and 16 readily mapped onto the virion structural data while ORF18 displayed some variation. This variation was proven to be the result of conformational swapping of the RBP head domains either upward, facing the capsid in an inactive form when calcium is absent, or in the "activated" form in the presence of calcium (or strontium) in which case the RBP head domains rotate 200° downward. Calcium is required for infection by many phages and therefore it is not surprising that it may be involved in the initial stages of the phage-host interaction, however, the conformational swapping of the RBP head domains is the first report of any such baseplate activation for *Siphoviridae* phages [29]. Such studies highlight the importance of electron microscopy analysis to define the native virion baseplate structures because, while the use of camelid antibodies provided data on the baseplate/RBP in the activated state, the state of the free phage baseplate revealed a previously unknown mechanism. While calcium is known to be required for the activation of the baseplate, it is unclear as yet if there is also as role for calcium in the physiology and organization of the bacterial envelope, thereby endorsing phage infection or if the effect is limited to the activation of the phage baseplate.

The application of these "block cloning" strategies to the expression and purification of baseplate complexes of the lactococcal phage p2 and TP901–1 is an important tool to facilitate the study of lactococcal base plates [47,48]. This block cloning also provides discerning information that may define which proteins of an operon form part of the complex and this approach may be applied to any operon encoding structural components of heteromeric organelles. A model for the construction and assembly of lactococcal P335 phage tails has been proposed in which the tail tape measure (TMP), distal tail protein (Dit) and tail-associated lysin (Tal) converge to act as the "initiation complex" upon which the baseplate is assembled [40,41]. Mass spectrometry has recently been used as an alternative method to define the assembly pathway of the baseplates of p2 and TP901–1 [49]. Using this approach, the p2 baseplate is proposed to assemble through the interaction of a hexamer of ORF15 with a preassembled ORF16 trimer followed by the capture of six ORF18 trimers [49]. This approach was also applied to the P335 phage TP901–1 and an assembly pathway was proposed which is outlined below.

4.3. P335 Phage Baseplates

The possible assembly pathway for TP901–1 that was proposed involves the interaction of homotrimers of ORF48 (upper baseplate protein) and ORF49 (lower baseplate protein) [49]. These then form tripods comprised of 3×ORF48 and 9×ORF49, and after interaction between ORF48 with ORF46 (distal tail protein or Dit), oligomerisation of ORF46 may proceed with concomitant baseplate formation with a proposed final composition of 6×ORF46, 18×ORF48 and 54×ORF49 [49]. Furthermore, through electron microscopic analysis of the TP901–1 baseplate together with that of

TP901–1 mutants lacking either BppU (upper baseplate) or BppL (lower baseplate), the structure and composition of the TP901–1 baseplate have been determined [31]. Structural analysis of the RBP of TP901–1 has revealed the similarity of the modularity of the RBP of TP901–1 and the previously studied p2, pointing to common ancestral origins of such structures (Figure 2) [17]. This is consistent with the development of a chimeric RBP with the N-terminal portion of TP901–1 and the C-terminal head domain of p2 [50]. The resulting structure presented domains almost indistinguishable from the parental structures highlighting the modular conservation of these lactococcal phages of distinct species [49]. The head domain of TP901–1 RBP was also observed to bind glycerol and fluorescence quenching further consolidated the carbohydrate affinity of this phage protein, identifying affinity for glycerol and muramyl-dipeptide [17].

Unlike the 936-type phage p2, the TP901–1 baseplate is maintained in a so-called "infection-ready conformation" [31,51]. This observation highlights the bipartite baseplate conformations either maintained in an infection-ready state or in a state requiring activation. It is suggested that the signalling to induce DNA ejection after adsorption of TP901–1 may require a more subtle conformational change than that of p2 and the T4 Myoviruses [51]. In this study, two potential infection pathway scenarios were proposed: first is the possibility that the tail-associated lysin (Tal) may interact with the cell wall peptidoglycan resulting in triggering of DNA release through a so-called cascade effect channelled up through the tail tube in a similar manner proposed for SPP1 [14,51]. The second proposed injection pathway involves the binding of the RBP tripod complexes to the carbohydrate moiety on the cell surface with mechanical alterations leading to signalling to the lowermost ring of the major tail protein with ensuing "un-screwing" of the tail to allow DNA ejection to occur [51]. Furthermore, comparisons have been drawn between the activation systems of p2 and the Myophage T4 upon interaction with a carbohydrate moiety, which is an interesting analogy and further corroborates theories of common ancestral elements and conservation of systems and modular structures in the case of RBPs [46].

Tuc2009 is a temperate P335 phage that is closely related to TP901–1 [52], although its baseplate-encoding operon contains an additional gene designated *bppA* (accessory baseplate protein). This appears to be a non-essential structural element of the Tuc2009 baseplate that is possibly involved in the determination or extension of host range. A low resolution model of the Tuc2009 baseplate has been determined in which the baseplate was proposed to consist of six tripods, each comprising 3×BppU, 3×BppA and 3×BppL, although this model is currently being refined that might align Tuc2009 more closely to TP901–1 in terms of its baseplate structure [30]. Structural analysis of this baseplate may uncover a potential role for BppA, for which a functional assignment remains to be provided. The presence of homologues of BppA in other lactococcal P335 phages (ul36 ORF303, accession no. NP_663688.1; P335 ORF45, accession no. ABI54248.1), indicates that there may be a biological significance and ecological advantage to phages possessing the protein in their baseplate. Furthermore, the recent finding that Tuc2009 is one of the few sequenced members of the P335 species that requires calcium to produce plaques may highlight the subtleties that distinguish individual phages within a species [51]. Although significant data is emerging relating to the structure

of P335 phage baseplates and individual components thereof, there remains a large gap between our knowledge of these phages and those of the model phages T4 and SPP1. While analogies may be drawn between these models and dairy phages it is imperative to also define relevant structural models within the dairy phage group.

Figure 2. A schematic representation of the distal tail region of Tuc2009, TP901–1 and p2. The stacks of tail rings composed of the major tail protein (MTP) (blue) with the distal tail (Dit) protein (orange) beneath, upon which the baseplate components are hinged (green and purple). In the case of Tuc2009 the green area represents the upper baseplate disc (BppU and BppA) while the TP901–1 upper disc is composed of BppU only. The lower baseplate (BppL) is presented in purple with a disc representation for TP901–1 and a so-called "petticoat" representation for Tuc2009. The protrusion from the baseplates of TP901–1 and Tuc2009 represent the tail associated lysin (Tal) (red). The schematic of the baseplate region of p2 highlights ORF15 (orange) at the base of the tail with ORF16 beneath (red) and the RBP in the active downward orientation (purple).
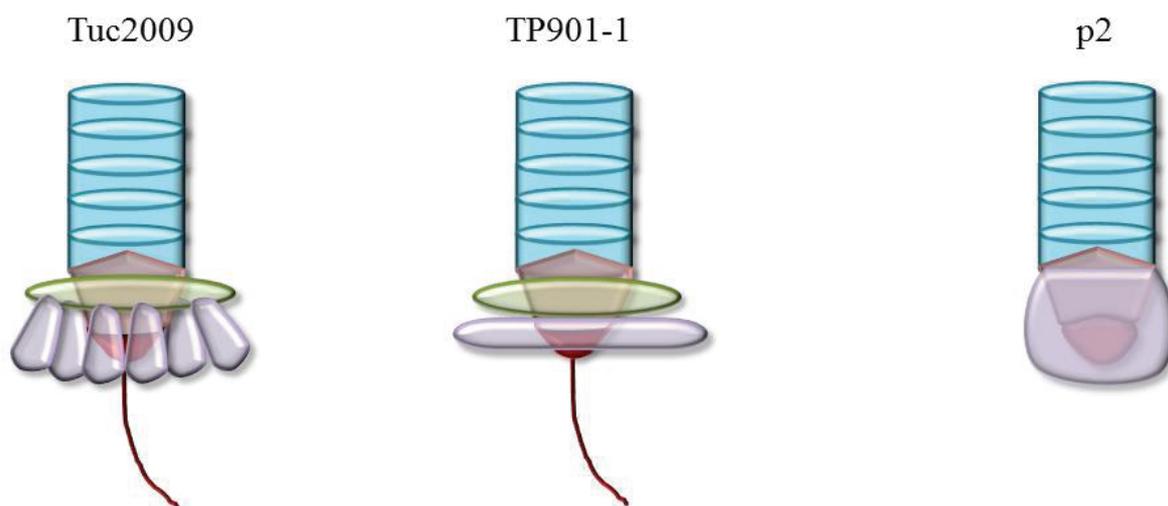
5. Conclusions and Future Perspectives

Structural studies relating to T4 in the 1990's and early 2000's paved the way for the structural analyses of many phage receptor binding proteins and their associated complexes and have permitted the detailed analysis of several phage-host interactions [9–12]. This, together with the isolation of the receptor for T5 and lambda in the 1990's and the 1970's prompted some of the early phage-host interaction studies for lactococcal phages [7,8,34]. Since then, numerous studies relating to the interactions of dairy phages and their hosts have been reported and have led to a surge of interest in structural analyses of these systems. The lactococcal phages p2, TP901–1 and Tuc2009 are emerging as model phages for the dairy *Siphoviridae* phages in terms of structural analyses. These data present a fresh perspective on studies attempting to define the host molecules with which dairy phages interact. The type of baseplate as defined by electron microscopy may be the initial indicator of the material

that acts as receptor for the individual phage although this is merely a first glimpse. Structural analysis and/or detailed electron microscopic analysis may provide detailed information on the baseplate organisation and structure, which may be applied to phage-host interaction studies. The analysis of such structures is paramount to understanding how dairy phages bind to and subsequently infect their hosts, and an improved knowledge of phage-host interactions will aid in developing novel and effective anti-phage strategies such as those already explored through the use of camelid antibodies and designed ankyrin repeat proteins (DARPins) to inactivate phages [15,32].

Since structural data relating to these model dairy phages continue to accrue, it is inevitable that these will become the dominant model for the analysis of other dairy phages including those of *S. thermophilus* and *Lb. delbrueckii*, typical cultures used in dairy fermentations. It has been suggested that the receptor material for the *Lb. delbrueckii* phage LL-H is polyglycerophosphate type lipoteichoic acids [18], which is reminiscent of affinity of the 936 phage p2, further demonstrating the usefulness of the lactococcal phage data as a suitable model for other dairy phages. However, while structural analyses are essential to furthering our knowledge of phage-host interactions, it is equally important to apply biological and functional characterisation to verify conclusions drawn from structural analyses. The marriage of molecular microbiology with structural biology represents one of the most powerful modern tools to understand complex interactions and pathways involved in host recognition and infection, and may be the key to developing next generation anti-phage systems with potential application in the dairy industry.

Acknowledgments

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Brussow, H.; Hendrix, R.W. Phage genomics: Small is beautiful. *Cell* **2002**, *108*, 13–16.
2. Stern, A.; Sorek, R. The phage-host arms race: Shaping the evolution of microbes. *Bioessays* **2011**, *33*, 43–51.
3. Ackermann, H.W. 5500 phages examined in the electron microscope. *Arch. Virol.* **2007**, *152*, 227–243.
4. Baptista, C.; Santos, M.A.; Sao-Jose, C. Phage SPP1 reversible adsorption to *Bacillus subtilis* cell wall teichoic acids accelerates virus recognition of membrane receptor YueB. *J. Bacteriol.* **2008**, *190*, 4989–4996.
5. Valyasevi, R.; Sandine, W.E.; Geller, B.L. A membrane protein is required for bacteriophage c2 infection of *Lactococcus lactis* subsp. *lactis* c2. *J. Bacteriol.* **1991**, *173*, 6095–6100.

6. Lubbers, M.W.; Waterfield, N.R.; Beresford, T.P.; Le Page, R.W.; Jarvis, A.W. Sequencing and analysis of the prolate-headed lactococcal bacteriophage c2 genome and identification of the structural genes. *Appl. Environ. Microbiol.* **1995**, *61*, 4348–4356.

7. Randall-Hazelbauer, L.; Schwartz, M. Isolation of the bacteriophage lambda receptor from *Escherichia coli*. *J. Bacteriol.* **1973**, *116*, 1436–1446.

8. Boulanger, P.; le Maire, M.; Bonhivers, M.; Dubois, S.; Desmadril, M.; Letellier, L. Purification and structural and functional characterization of FhuA, a transporter of the *Escherichia coli* outer membrane. *Biochemistry* **1996**, *35*, 14216–14224.

9. Bartual, S.G.; Otero, J.M.; Garcia-Doval, C.; Llamas-Saiz, A.L.; Kahn, R.; Fox, G.C.; van Raaij, M.J. Structure of the bacteriophage T4 long tail fiber receptor-binding tip. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 20287–20292.

10. Trojet, S.N.; Caumont-Sarcos, A.; Perrody, E.; Comeau, A.M.; Krisch, H.M. The gp38 adhesins of the T4 Superfamily: A complex modular determinant of the phage's host specificity. *Genome Biol. Evol.* **2011**, *3*, 674–686.

11. Kostyuchenko, V.A.; Leiman, P.G.; Chipman, P.R.; Kanamaru, S.; van Raaij, M.J.; Arisaka, F.; Mesyanzhinov, V.V.; Rossmann, M.G. Three-dimensional structure of bacteriophage T4 baseplate. *Nat. Struct. Biol.* **2003**, *10*, 688–693.

12. Leiman, P.G.; Shneider, M.M.; Kostyuchenko, V.A.; Chipman, P.R.; Mesyanzhinov, V.V.; Rossmann, M.G. Structure and location of gene product 8 in the bacteriophage T4 baseplate. *J. Mol. Biol.* **2003**, *328*, 821–833.

13. Leiman, P.G.; Shneider, M.M.; Mesyanzhinov, V.V.; Rossmann, M.G. Evolution of bacteriophage tails: Structure of T4 gene product 10. *J. Mol. Biol.* **2006**, *358*, 912–921.

14. Plisson, C.; White, H.E.; Auzat, I.; Zafarani, A.; Sao-Jose, C.; Lhuillier, S.; Tavares, P.; Orlova, E.V. Structure of bacteriophage SPP1 tail reveals trigger for DNA ejection. *EMBO J.* **2007**, *26*, 3720–3728.

15. Tremblay, D.M.; Tegoni, M.; Spinelli, S.; Campanacci, V.; Blangy, S.; Huyghe, C.; Desmyter, A.; Labrie, S.; Moineau, S.; Cambillau, C. Receptor-binding protein of *Lactococcus lactis* phages: Identification and characterization of the saccharide receptor-binding site. *J. Bacteriol.* **2006**, *188*, 2400–2410.

16. Ricagno, S.; Campanacci, V.; Blangy, S.; Spinelli, S.; Tremblay, D.; Moineau, S.; Tegoni, M.; Cambillau, C. Crystal structure of the receptor-binding protein head domain from *Lactococcus lactis* phage bil170. *J. Virol.* **2006**, *80*, 9331–9335.

17. Spinelli, S.; Campanacci, V.; Blangy, S.; Moineau, S.; Tegoni, M.; Cambillau, C. Modular structure of the receptor binding proteins of *Lactococcus lactis* phages. The RBP structure of the temperate phage TP901–1. *J. Biol. Chem.* **2006**, *281*, 14256–14262.

18. Raisanen, L.; Schubert, K.; Jaakonsaari, T.; Alatossava, T. Characterization of lipoteichoic acids as *Lactobacillus delbrueckii* phage receptor components. *J. Bacteriol.* **2004**, *186*, 5529–5532.

19. Quiberoni, A.; Stiefel, J.I.; Reinheimer, J.A. Characterization of phage receptors in *Streptococcus thermophilus* using purified cell walls obtained by a simple protocol. *J. Appl. Microbiol.* **2000**, *89*, 1059–1065.

20. Vinga, I.; Baptista, C.; Auzat, I.; Petipas, I.; Lurz, R.; Tavares, P.; Santos, M.A.; Sao-Jose, C. Role of bacteriophage SPP1 tail spike protein gp21 on host cell receptor binding and trigger of phage DNA ejection. *Mol. Microbiol.* **2012**, *83*, 289–303.

21. Lhuillier, S.; Gallopin, M.; Gilquin, B.; Brasiles, S.; Lancelot, N.; Letellier, G.; Gilles, M.; Dethan, G.; Orlova, E.V.; Couprie, J.; *et al.* Structure of bacteriophage SPP1 head-to-tail connection reveals mechanism for viral DNA gating. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 8507–8512.

22. Veesler, D.; Blangy, S.; Lichiere, J.; Ortiz-Lombardia, M.; Tavares, P.; Campanacci, V.; Cambillau, C. Crystal structure of *Bacillus subtilis* SPP1 phage gp23.1, a putative chaperone. *Protein Sci.* **2010**, *19*, 1812–1816.

23. Veesler, D.; Robin, G.; Lichiere, J.; Auzat, I.; Tavares, P.; Bron, P.; Campanacci, V.; Cambillau, C. Crystal structure of bacteriophage SPP1 distal tail protein (gp19.1): A baseplate hub paradigm in gram-positive infecting phages. *J. Biol. Chem.* **2010**, *285*, 36666–36673.

24. Veesler, D.; Blangy, S.; Spinelli, S.; Tavares, P.; Campanacci, V.; Cambillau, C. Crystal structure of *Bacillus subtilis* SPP1 phage gp22 shares fold similarity with a domain of lactococcal phage p2 RBP. *Protein Sci.* **2010**, *19*, 1439–1443.

25. Goulet, A.; Lai-Kee-Him, J.; Veesler, D.; Auzat, I.; Robin, G.; Shepherd, D.A.; Ashcroft, A.E.; Richard, E.; Lichiere, J.; Tavares, P.; *et al.* The opening of the SPP1 bacteriophage tail, a prevalent mechanism in Gram-positive-infecting siphophages. *J. Biol. Chem.* **2011**, *286*, 25397–25405.

26. Chagot, B.; Auzat, I.; Gallopin, M.; Petitpas, I.; Gilquin, B.; Tavares, P.; Zinn-Justin, S. Solution structure of gp17 from the *Siphoviridae* bacteriophage SPP1: Insights into its role in virion assembly. *Proteins* **2012**, *80*, 319–326.

27. White, H.E.; Sherman, M.B.; Brasiles, S.; Jacquet, E.; Seavers, P.; Tavares, P.; Orlova, E.V. Capsid structure and its stability at the late stages of bacteriophage SPP1 assembly. *J. Virol.* **2012**, *86*, 6768–6777.

28. Auzat, I.; Droge, A.; Weise, F.; Lurz, R.; Tavares, P. Origin and function of the two major tail proteins of bacteriophage SPP1. *Mol. Microbiol.* **2008**, *70*, 557–569.

29. Sciara, G.; Bebeacua, C.; Bron, P.; Tremblay, D.; Ortiz-Lombardia, M.; Lichiere, J.; van Heel, M.; Campanacci, V.; Moineau, S.; Cambillau, C. Structure of lactococcal phage p2 baseplate and its mechanism of activation. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 6852–6857.

30. Sciara, G.; Blangy, S.; Siponen, M.; Mc Grath, S.; van Sinderen, D.; Tegoni, M.; Cambillau, C.; Campanacci, V. A topological model of the baseplate of lactococcal phage Tuc2009. *J. Biol. Chem.* **2008**, *283*, 2716–2723.

31. Bebeacua, C.; Bron, P.; Lai, L.; Vegge, C.S.; Brondsted, L.; Spinelli, S.; Campanacci, V.; Veesler, D.; van Heel, M.; Cambillau, C. Structure and molecular assignment of lactococcal phage TP901–1 baseplate. *J. Biol. Chem.* **2010**, *285*, 39079–39086.

32. Veesler, D.; Dreier, B.; Blangy, S.; Lichiere, J.; Tremblay, D.; Moineau, S.; Spinelli, S.; Tegoni, M.; Pluckthun, A.; Campanacci, V.; *et al.* Crystal structure and function of a darpin neutralizing inhibitor of lactococcal phage TP901–1: Comparison of DARPin and camelid VHH binding mode. *J. Biol. Chem.* **2009**, *284*, 30718–30726.

33. Deveau, H.; Labrie, S.J.; Chopin, M.C.; Moineau, S. Biodiversity and classification of lactococcal phages. *Appl. Environ. Microbiol.* **2006**, *72*, 4338–4346.

34. Geller, B.L.; Ngo, H.T.; Mooney, D.T.; Su, P.; Dunn, N. Lactococcal 936-species phage attachment to surface of *Lactococcus lactis*. *J. Dairy Sci.* **2005**, *88*, 900–907.

35. Dupont, K.; Janzen, T.; Vogensen, F.K.; Josephsen, J.; Stuer-Lauridsen, B. Identification of *Lactococcus lactis* genes required for bacteriophage adsorption. *Appl. Environ. Microbiol.* **2004**, *70*, 5825–5832.

36. Chapot-Chartier, M.P.; Vinogradov, E.; Sadovskaya, I.; Andre, G.; Mistou, M.Y.; Trieu-Cuot, P.; Furlan, S.; Bidnenko, E.; Courtin, P.; Pechoux, C.; *et al.* Cell surface of *Lactococcus lactis* is covered by a protective polysaccharide pellicle. *J. Biol. Chem.* **2010**, *285*, 10464–10471.

37. Ostergaard-Breum, S.; Neve, H.; Heller, K.J.; Vogensen, F.K. Temperate phages TP901–1 and phiLC3, belonging to the P335 species, apparently use different pathways for DNA injection in *Lactococcus lactis* subsp. *cremoris* 3107. *FEMS Microbiol. Lett.* **2007**, *276*, 156–164.

38. Dupont, K.; Vogensen, F.K.; Neve, H.; Bresciani, J.; Josephsen, J. Identification of the receptor-binding protein in 936-species lactococcal bacteriophages. *Appl. Environ. Microbiol.* **2004**, *70*, 5818–5824.

39. Duplessis, M.; Moineau, S. Identification of a genetic determinant responsible for host specificity in *Streptococcus thermophilus* bacteriophages. *Mol. Microbiol.* **2001**, *41*, 325–336.

40. Vegge, C.S.; Vogensen, F.K.; Mc Grath, S.; Neve, H.; van Sinderen, D.; Brondsted, L. Identification of the lower baseplate protein as the antireceptor of the temperate lactococcal bacteriophages TP901–1 and Tuc2009. *J. Bacteriol.* **2006**, *188*, 55–63.

41. Mc Grath, S.; Neve, H.; Seegers, J.F.; Eijlander, R.; Vegge, C.S.; Brondsted, L.; Heller, K.J.; Fitzgerald, G.F.; Vogensen, F.K.; van Sinderen, D. Anatomy of a lactococcal phage tail. *J. Bacteriol.* **2006**, *188*, 3972–3982.

42. Spinelli, S.; Desmyter, A.; Verrips, C.T.; de Haard, H.J.; Moineau, S.; Cambillau, C. Lactococcal bacteriophage p2 receptor-binding protein structure suggests a common ancestor gene with bacterial and mammalian viruses. *Nat. Struct. Mol. Biol.* **2006**, *13*, 85–89.

43. Ledeboer, A.M.; Bezemer, S.; de Hiaard, J.J.; Schaffers, I.M.; Verrips, C.T.; van Vliet, C.; Dusterhoft, E.M.; Zoon, P.; Moineau, S.; Frenken, L.G. Preventing phage lysis of *Lactococcus lactis* in cheese production using a neutralizing heavy-chain antibody fragment from llama. *J. Dairy Sci.* **2002**, *85*, 1376–1382.

44. De Haard, H.J.; Bezemer, S.; Ledeboer, A.M.; Muller, W.H.; Boender, P.J.; Moineau, S.; Coppelmans, M.C.; Verkleij, A.J.; Frenken, L.G.; Verrips, C.T. Llama antibodies against a lactococcal protein located at the tip of the phage tail prevent phage infection. *J. Bacteriol.* **2005**, *187*, 4531–4541.

45. Mahony, J.; Deveau, H.; Mc Grath, S.; Ventura, M.; Canchaya, C.; Moineau, S.; Fitzgerald, G.F.; van Sinderen, D. Sequence and comparative genomic analysis of lactococcal bacteriophages jj50, 712 and p008: Evolutionary insights into the 936 phage species. *FEMS Microbiol. Lett.* **2006**, *261*, 253–261.

46. Dupont, K.; Vogensen, F.K.; Josephsen, J. Detection of lactococcal 936-species bacteriophages in whey by magnetic capture hybridization pcr targeting a variable region of receptor-binding protein genes. *J. Appl. Microbiol.* **2005**, *98*, 1001–1009.

47. Castro-Nallar, E.; Chen, H.; Gladman, S.; Moore, S.C.; Seemann, T.; Powell, I.B.; Hillier, A.; Crandall, K.A.; Chandry, P.S. Population genomics and phylogeography of an Australian dairy factory derived lytic bacteriophage. *Genome Biol. Evol.* **2012**, *4*, 382–393.

48. Campanacci, V.; Veesler, D.; Lichiere, J.; Blangy, S.; Sciara, G.; Moineau, S.; van Sinderen, D.; Bron, P.; Cambillau, C. Solution and electron microscopy characterization of lactococcal phage baseplates expressed in *Escherichia coli*. *J. Struct. Biol.* **2010**, *172*, 75–84.

49. Shepherd, D.A.; Veesler, D.; Lichiere, J.; Ashcroft, A.E.; Cambillau, C. Unraveling lactococcal phage baseplate assembly by mass spectrometry. *Mol. Cell Proteom.* **2011**, *10*, M111.009787.

50. Siponen, M.; Sciara, G.; Villion, M.; Spinelli, S.; Lichiere, J.; Cambillau, C.; Moineau, S.; Campanacci, V. Crystal structure of orf12 from *Lactococcus lactis* phage p2 identifies a tape measure protein chaperone. *J. Bacteriol.* **2009**, *191*, 728–734.

51. Veesler, D.; Spinelli, S.; Mahony, J.; Lichiere, J.; Blangy, S.; Bricogne, G.; Legrand, P.; Ortiz-Lombardia, M.; Campanacci, V.; van Sinderen, D.; *et al.* Structure of the phage TP901–1 1.8 MDa baseplate suggests an alternative host adhesion mechanism. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 8954–8958.

52. Seegers, J.F.; Mc Grath, S.; O'Connell-Motherway, M.; Arendt, E.K.; van de Guchte, M.; Creaven, M.; Fitzgerald, G.F.; van Sinderen, D. Molecular and transcriptional analysis of the temperate lactococcal bacteriophage Tuc2009. *Virology* **2004**, *329*, 40–52.

Article

Genomic Sequences of two Novel *Levivirus* Single-Stranded RNA Coliphages (Family *Leviviridae*): Evidence for Recombination in Environmental Strains

Stephanie D. Friedman [1,*]**, Wyatt C. Snellgrove** [2] **and Fred J. Genthner** [1]

[1] US Environmental Protection Agency, Gulf Ecology Division, 1 Sabine Island Drive, Gulf Breeze, FL, 32561, USA; E-Mails: friedman.stephanie@epa.gov (S.D.F.); genthner.fred@epa.gov (F.J.G.)

[2] William Carey University College of Osteopathic Medicine, 498 Tuscan Avenue, Hattiesburg, MS 39401, USA; E-Mail: cliffsnellgrove@gmail.com

* Author to whom correspondence should be addressed; E-Mail: friedman.stephanie@epa.gov; Tel.: +1-850-934-2468; Fax: +1-850-934-9201.

Abstract: Bacteriophages are likely the most abundant entities in the aquatic environment, yet knowledge of their ecology is limited. During a fecal source-tracking study, two genetically novel *Leviviridae* strains were discovered. Although the novel strains were isolated from coastal waters 1130 km apart (North Carolina and Rhode Island, USA), these strains shared 97% nucleotide similarity and 97–100% amino acid similarity. When the novel strains were compared to nine *Levivirus* genogroup I strains, they shared 95–100% similarity among the maturation, capsid and lysis proteins, but only 84–85% in the RNA-dependent RNA polymerase gene. Further bioinformatic analyses suggested a recombination event occurred. To the best of our knowledge, this is the first description of viral recombinants in environmental *Leviviridae* ssRNA bacteriophages.

Keywords: male-specific coliphage; *Leviviridae*; viral recombinants; ssRNA virus; FRNA; bacteriophage

1. Introduction

Bacteriophages have played a major role contributing to our knowledge of molecular biology, not only in the role of model viruses but also as tools to investigate mRNA, genes, genetic codes and genomes. The first sequenced genomes were the RNA bacteriophage MS2 [1] and the DNA bacteriophage Φ-X174 [2]. As important as *Drosophila* was in shaping the field of genetics and Tobacco Mosaic Virus was in advancing the study of virology and biochemistry, the RNA phage MS2 (family *Leviviridae*) was fundamental in laying the foundation of molecular biology. Thus, it is important to continue adding to the basic understanding of phages. The observations presented in this study were rather serendipitous, in that the focus was not on searching for natural recombinant bacteriophages. Nonetheless, evidence of a recombination event was revealed during a ssRNA bacteriophage sequencing project [3].

Male-specific ssRNA (FRNA) coliphages belong to the family *Leviviridae*. They are classified into two genera (*Levivirus* and *Allolevivirus*) which are subdivided into four genogroups (genogroups I and II in *Levivirus* and genogroups III and IV in *Allolevivirus*). Investigating the genetic diversity of FRNA phages Vinjé *et al* [4] conducted a phylogenetic analysis of 32 *Levivirus* field strains using a 189 bp replicase gene fragment. This study revealed three main clusters: genogroup I, genogroup II and a potential novel group, designated JS, which clustered between genogroup I and genogroup II. The putative JS group, represented by phages, WWTP1_50 and 2GI13, had a >40% sequence diversity in the 189 bp replicase gene sequence when compared to strains from genogroups I and II. As these strains were isolated from widely separated geographical regions (Massachusetts and South Carolina) Vinjé *et al*., [4] proposed that JS may form a stable lineage. This report suggested further genomic sequencing and serological data were needed to confirm whether these strains formed a novel genogroup or whether they were the result of recombination or rearrangement events [5].

In its simplest form, recombination occurs when two disparate DNA or RNA strands exchange or merge stretches of their sequences whereas mutation involves the substitution, deletion or insertion of a nucleotide resulting in the change of the nucleotide sequence of a gene or an amino acid sequence of a protein. In some RNA viruses, RNA recombination events can occur when two or more strains infect the same host. Proposed models for the formation of novel RNA sequences include (i) cleavage and ligation in RNA molecules or RNA secondary structures [6]; (ii) replicative template switching whereby the RNA-dependent RNA polymerase (replicase) switches from one template to another RNA template, also known as copy choice [5,7]; and (iii) RNA transesterification which occurs when the polymerase adds a separate RNA fragment to the 3' terminus of the original RNA template [5].

Historically, experiments with ssRNA coliphage mutants failed to provide evidence for recombination and the investigators concluded that RNA phages would not undergo recombination [8]. A potential flaw in the conclusion may have been that the study occurred at the time when FRNA phages were thought to possess only three genes, not four. In all likelihood, laboratory-applied selective pressure failed to detect or generate a specific recombinant. This failure may not necessarily

reflect the lack of recombination or responsible mechanisms that could occur under actual environmental conditions encountered by ssRNA coliphages.

The first indication of RNA recombination in a male-specific FRNA phage was the report of small, non homologous, recombinant RNA molecules produced from a purified template-free Qβ replicase molecule [9]. The investigators noted similar RNA molecules were present in *E. coli* cells infected with phage Qβ. Chetverin *et al.*, [7] studied this phenomenon by observing the formation of novel sequences in RNA molecules which suggested that this recombination event occurred as a transesterification reaction catalyzed by a conformation acquired by Qβ replicase during RNA synthesis [5,7]. Nucleotide sequences of recombined RNA molecules non-homologous to the parent RNA were formed in the absence of DNA intermediates, demonstrating an RNA recombination mechanism in the presence of Qβ replicase [5]. Therefore, it was plausible to have recombination in environmental ssRNA male-specific coliphage (*Leviviridae*) isolates.

In the present study, two JS strains, DL52 and DL54, were isolated during an environmental genotyping study of *Leviviridae* FRNA phages [10,11]. As in the Vinjé study [4], strains DL52 and DL54 were isolated from separate coastal waters. These phages were placed into the putative JS subgroup using the genotyping methods of Vinjé *et al*. [4]. The objective of this study was to determine whether the existence of a novel JS-like subgroup representing a third *Levivirus* cluster as proposed by Vinjé *et al*., [4] could be verified. The approach taken here was to compare sequences from the JS strains to nucleotide and amino acid sequence data from entire genomes of 10 levivirus genogroup I strains and 5 levivirus genogroup II strains [3]. Analysis of the novel JS strains provided evidence to determine whether these *Levivirus* strains clustered to genogroup I, II, a combination of groups I and II or a unique genogroup. To further understand the phylogeny of these JS strains, complete genomic sequencing, amino acid composition, phylogenetic, bioinformatic and statistical analyses were performed.

2. Results

2.1. Sequence Analyses and Open Reading Frames

Preliminary analysis of nucleotide sequences from a replicase 189 bp amplicon placed the two novel strains, DL52 and DL54, into a "JS-like" subgroup [4]. Reverse-line blot hybridization failed to genotype the two strains into genogroups I or II [4].

A total of 17 strains (Table 1) were used to examine the relationships among nucleotides and amino acids in the *Levivirus* genus. The first 9 strains in genogroup I, Table 1, *i.e.*, MS2, ST4, DL1, DL2, DL13, DL16, R17, M12 and J20, were referred to as "MS2-like."

Genogroup I MS2-like strains Open Reading Frame (ORF) start and stop codons were located at identical or very similar nucleotide positions as previously reported for strain MS2. The JS strains also had identical ORF start and stop codon positions as the MS-2 like strains (Table 2).

Nucleotide pairwise comparisons of full-length genomes were made between all strains within the *Levivirus* genome, including strains within genogroups I, JS and genogroups II. Within the nine strains

of MS2-like genogroup I, full-length nucleotide sequence similarity was 91%–99% [3] whereas the two JS strains, DL52 and DL54, shared 96.73% sequence similarity to each other. In comparison, the JS nucleotide sequences were more similar to MS2-like genogroup I (80%–85%) than to the genogroup I strain fr (69%) or to genogroup II strains (52%–54%) (Table 3a).

Table 1. Male-specific ssRNA coliphages (FRNA), family *Leviviridae*, genus *Levivirus*, strain origins and identifications.

Strain	Genogroup	Source	Origin	Accession number
MS2	I	sewage	Berkeley, CA	**NC_001417**
M12	I	sewage	Germany	**AF195778**
DL1	I	river water	Tijuana River, CA	**EF107159**
DL2	I	bay water	Delaware Bay, DE	N/A
DL13	I	oyster	Whiskey Creek, NC	N/A
DL16	I	bay water	Great Bay, NH	**EF108464**
J20	I	chicken litter	South Carolina	**EF204939**
ST4	I	unknown	unknown	**EF204940**
R17	I	sewage	Philadelphia, PA	**EF108465**
fr	I	dung hill	Heidelberg, Germany	**X15031**
DL52	I-JS	bay water	Rachel Carson Reserve, NC	**JQ966307**
DL54	I-JS	bay water	Narragansett Bay, RI	**JQ966308**
GA	II	sewage	Ookayama, Japan	**NC_001426**
KU1	II	sewage	Kuwait	**AF227250**
DL10	II	mussel	Tijuana River, CA	**FJ483837**
DL20	II	clam	Narragansett Bay, RI	**FJ483839**
T72	II	bird	Talbert Marsh sandflats, CA	**FJ483838**

Table 2. Open Reading Frame positions and genome lengths of FRNA coliphage (family *Leviviridae*, genus *Levivirus*). Nucleotide positions are based on alignment. Number of amino acids for each gene is in parentheses [3].

Open Reading Frame Locations (amino acids)						
Strain	Group	Full length	ORF1	ORF2	ORF3	ORF4
MS2 [a]	I	3569	130-1311(393)	1335-1727(130)	1678-1905(75)	1761-3398(545)
M12 [a,b]	I	3340[b]	130-1311(393)	1335-1727(130)	1678-1905(75)	ND
DL1	I	3570	130-1311(393)	1335-1727(130)	1678-1905(75)	1761-3398(545)
DL2	I	3491[c]	130-1311(393)	1335-1727(130)	1678-1905(75)	1761-3398(545)
DL13	I	3491[c]	130-1311(393)	1335-1727(130)	1678-1905(75)	1761-3398(545)
DL16	I	3569	130-1311(393)	1335-1727(130)	1678-1905(75)	1761-3398(545)
J20	I	3569	130-1311(393)	1335-1727(130)	1678-1905(75)	1761-3398(545)
ST4	I	3569	130-1311(393)	1335-1727(130)	1678-1905(75)	1761-3398(545)
R17	I	3569	130-1311(393)	1335-1727(130)	1678-1905(75)	1761-3398(545)
fr [a]	I	3575	129-1310(393)	1336-1728(130)	1691-1906(71)	1762-3399(545)
DL52	JS	3525	130-1311(393)	1335-1727(130)	1678-1905(75)	1761-3398[d] (545)
DL54	JS	3398[c]	130-1311(393)	1335-1727(130)	1678-1905(75)	1761-3398[d] (545)

[a] previously published in GenBank; [b] strain M12 was not fully sequenced in the GenBank submission; [c] nearly full-length genome; [d] contains numerous deletions and insertions in ORF4; ND = not determined.

Table 3. (**a**) Pairwise nucleotide full-length genome percent similarity. (i) *Levivirus* JS strains DL52 and DL54 compared to genogroup I; (ii) *Levivirus* JS strains DL52 and DL54 compared to genogroup II. Pairwise alignments were performed in BioEdit with DAYHOFF similarity parameters.

Strain	DL52	DL54
(i) Genogroup I and JS strains		
DL52	100	
DL54	96.73	100
DL1	81.48	81.87
DL16	85.41	84.72
ST4	80.30	80.11
R17	80.55	80.53
J20	82.00	82.01
MS2	80.12	80.01
fr	69.18	69.06
(ii) Genogroup II and JS strains		
DL52	100	
DL54	96.73	100
T72	53.96	53.53
DL10	54.07	53.89
DL20	52.87	52.65
GA	52.44	52.29
KU1	52.94	52.66

Despite their sequence similarities, genome lengths for JS strains (3525 nt) were shorter than all genogroup I strains (3569–3575 nt) (Table 2) but longer than genogroup II (3458–3486 nt) [3]. Numerous deletions in the 3' untranslated region and a portion of ORF4 (replicase) in JS strains accounted for the decreased genome length (data not shown) but did not alter the ORF positions when the genogroup I strains were aligned (Table 2).

Analysis of the replicase gene revealed a 2 nt insertion at the 1374 nucleotide region when counting ORF4 start site as nucleotide 1 (Figure 1). This insertion occurred upstream from the ORF4 stop codon. Beginning approximately 40 nt downstream from the replicase ORF4 stop codon and continuing to the 3' termini, 53 nt deletions were present in the JS strains when aligned to MS2-like genomes. Nucleotide alignment of the replicase and nontranslated regions (NTR) revealed numerous nt deletions in the JS strains when compared to genogroup I strains accounting for the change in amino acid composition. However, JS strains shared the 3' terminal "signature", ACCACCCA, present in *Levivirus* genogroups I and II [3].

Figure 1. Replicase recombinant region in two JS strains when compared to genogroup I strains (family *Leviviridae*). Alignment (BioEdit v7.0.1) of the replicase nucleotide sequences from *Levivirus* genogroup I strains DL1, DL2, DL13, DL16, ST4, R17, J20, MS2 with JS strains DL52 and DL54. For clarity, only a portion of the alignment is shown. Alignment of each genogroup is depicted in discontinuous blocks. The numbers along the top are the nucleotide positions within the replicase gene with the start position of ORF4 assigned as nucleotide 1. Genome sequences read 5'-3' direction. Dots indicate identity with the consensus sequence. Degenerate bases are noted in the standard IUB codes. The replicase start codon and two nucleotide insertions are highlighted in red. Dashes denote a nucleotide sequence deletion from the consensus sequence.

2.2. Amino Acid Analysis

Initially, nucleotide pairwise analyses of full-length genomes were made comparing all strains within the *Levivirus* genome, including genogroups I, JS and II; an 80%–85% nucleotide similarity between JS strains and the MS2-like strains was observed (Table 3a). In comparison, the amino acid sequences of the maturation, capsid and lysis proteins of the JS strains were very similar to those of the MS2-like genogroup I strains, sharing 97%–100%, 98%–100% and 95%–100% sequence similarities, respectively (Table 3b). Genogroup I strain fr, when compared to MS2-like and JS genogroup I strains, only shared an amino acid similarity to the maturation, capsid and lysis proteins ranging from 75.73%–91.85% (Table 3b). In contrast, the replicase protein sequences of the JS strains were quite dissimilar to the replicase protein sequences of the MS2-like genogroup I strains, displaying a similarity range of 79%–85% (Table 3c). However, a similarity of 97%–99% was observed among the highly conserved replicase genes for the MS2-like strains. Strain fr shared a 79% replicase similarity

to JS strains and approximately 88%–89% similarity to MS2-like strains. Genogroup II replicase was approximately 52%–53% similar to JS strains, 50%–53% to MS2-like and fr strains and 92%–98% similar to other genogroup II strains (Table 3c).

Table 3. (b) Percent similarity in amino acid sequences between *Levivirus* JS strains and genogroup I maturation, capsid and lysis proteins. Amino acid pairwise computations were performed in Bionumerics.

Maturation Protein											
	ST4-I	MS2-I	R17-I	M12-I	DL1-I	J20-I	DL13-I	DL16-I	DL52-JS	DL54-JS	DL2-I	fr-I
ST4-I	100.0											
MS2-I	99.9	100.0										
R17-I	98.2	98.3	100.0									
M12-I	97.7	97.9	97.4	100.0								
DL1-I	97.7	97.8	97.2	97.2	100.0							
J20-I	97.9	98.0	97.1	97.1	98.3	100.0						
DL13-I	97.5	97.6	97.1	97.5	98.1	98.1	100.0					
DL16-I	97.5	97.6	97.1	97.5	98.1	98.1	100.0	100.0				
DL52-JS	97.5	97.6	97.1	97.5	98.1	98.1	100.0	100.0	100.0			
DL54-JS	97.5	97.6	97.1	97.5	98.1	98.1	100.0	100.0	100.0	100.0		
DL2-I	97.6	97.8	97.3	97.6	98.3	98.1	99.8	99.8	99.8	99.8	100.0	
fr-I	84.4	84.4	84.3	84.2	85.0	84.7	84.1	84.1	84.1	84.1	84.4	100.0

Capsid Protein											
	ST4-I	R17-I	MS2-I	M12-I	DL1-I	DL2-I	DL13-I	DL16-I	J20-I	DL52-JS	DL54-JS	fr-I
ST4-I	100.0											
R17-I	100.0	100.0										
MS2-I	100.0	100.0	100.0									
M12-I	98.7	98.7	98.7	100.0								
DL1-I	98.5	98.5	98.5	98.2	100.0							
DL2-I	98.5	98.5	98.5	98.2	100.0	100.0						
DL13-I	98.5	98.5	98.5	98.2	100.0	100.0	100.0					
DL16-I	98.5	98.5	98.5	98.2	100.0	100.0	100.0	100.0				
J20-I	98.5	98.5	98.5	98.2	100.0	100.0	100.0	100.0	100.0			
DL52-JS	98.5	98.5	98.5	98.2	100.0	100.0	100.0	100.0	100.0	100.0		
DL54-JS	98.5	98.5	98.5	98.2	100.0	100.0	100.0	100.0	100.0	100.0	100.0	
fr-I	91.2	91.2	91.2	90.9	91.9	91.9	91.9	91.9	91.9	91.9	91.9	100.0

Lysis Protein											
	ST4-I	MS2-I	R17-I	DL13-I	DL16-I	DL2-I	DL52-JS	DL54-JS	DL1-I	J20-I	M12-I	fr-I
ST4-I	100.0											
MS2-I	100.0	100.0										
R17-I	98.9	98.9	100.0									
DL13-I	95.0	95.0	95.0	100.0								
DL16-I	95.0	95.0	95.0	100.0	100.0							
DL2-I	95.0	95.0	95.0	100.0	100.0	100.0						
DL52-JS	95.0	95.0	95.0	100.0	100.0	100.0	100.0					
DL54-JS	95.0	95.0	95.0	100.0	100.0	100.0	100.0	100.0				
DL1-I	95.9	95.9	95.9	99.2	99.2	99.2	99.2	99.2	100.0			
J20-I	91.1	91.1	91.1	96.4	96.4	96.4	96.4	96.4	95.5	100.0		
M12-I	93.8	93.8	94.1	90.3	90.3	90.3	90.3	90.3	91.3	86.2	100.0	
fr-I	78.2	78.2	78.0	78.9	78.9	78.9	78.9	78.9	77.7	77.7	75.7	100.0

Table 3. (c) Amino acid percent similarity comparisons between *Levivirus* JS strains, DL52 and DL54, to *Levivirus* genogroup I and genogroup II RNA-dependent RNA polymerase (replicase) protein. Amino acid pairwise computations were performed in Bionumerics.

	ST4-I	R17-I	MS2-I	DL2-I	DL13-I	DL16-I	DL1-I	J20-I	fr-I	DL52-JS	DL54-JS	T72-II	KU1-II	DL10-II	DL20-II	GA-II
Replicase Protein																
ST4-I	100.0															
R17-I	99.1	100.0														
MS2-I	99.0	98.6	100.0													
DL2-I	97.5	97.1	97.5	100.0												
DL13-I	97.4	97.0	97.4	99.9	100.0											
DL16-I	97.2	96.8	97.3	99.7	99.6	100.0										
DL1-I	98.0	97.6	97.8	99.4	99.3	99.1	100.0									
J20-I	97.7	97.2	97.5	99.0	98.9	98.7	99.1	100.0								
fr-I	88.5	88.4	88.2	89.0	88.9	88.7	88.9	88.8	100.0							
DL52-JS	84.4	83.9	84.0	85.4	85.4	85.2	85.3	84.8	79.7	100.0						
DL54-JS	84.2	84.1	83.9	85.1	85.0	84.8	85.0	84.8	79.8	97.2	100.0					
T72-II	51.9	51.5	51.1	52.6	52.6	52.3	52.6	52.7	52.8	52.9	53.1	100.0				
KU1-II	51.2	50.8	50.4	51.9	51.9	51.6	51.9	52.0	52.5	52.5	52.6	97.1	100.0			
DL10-II	52.7	52.3	51.9	53.4	53.4	53.1	53.4	53.7	52.7	53.7	53.8	93.3	93.0	100.0		
DL20-II	52.7	52.3	51.9	53.4	53.4	53.1	53.2	53.7	53.2	53.4	53.5	93.8	93.4	98.6	100.0	
GA-II	51.3	50.9	50.5	52.0	52.0	51.7	51.8	52.3	51.8	52.5	52.6	92.4	92.3	97.7	98.6	100.0

All genogroup I strains, including fr, and the two JS strains had a replicase protein length of 545 amino acids (Table 2) [3]. However, JS replicase differed from genogroup I replicase as it had one amino acid insertion at replicase position 467 and one amino acid deletion at the 3' termini of the stop codon, but maintained a total of 545 amino acids (data not shown). Identical to genogroup I strains, the replicase catalytic domain in the JS strains occurred between amino acid positions 243–373, thereby adding confidence to placing the grouping of JS into genogroup I [3]. Beginning at amino acid number 455 within the replicase gene, JS strains were unique in amino acid composition and diverged from the MS2-like strains.

2.3. Pfam and Protein Sequence Motifs

Individual proteins from JS strains DL52 and DL54 were grouped to protein families by Pfam analysis. The maturation protein generated "phage_mat-A" domain including all *Leviviridae* strains plus three additional bacteriophage, PRR1, PP7 and AP205. The capsid protein resulted in a "Levi-coat" domain including all *Leviviridae* strains plus bacteriophage PRR1. The lysis protein only generated results in a PfamB search matching the genus *Levivirus* strains from both genogroups I and II including KU1, JP34, M12, FP501, MS2, JP500, fr, TH1, SD, GA, BO1, TL2 and ZR. Replicase protein matched "RNA_replicase_B" domain within the *Leviviridae* family plus the additional bacteriophages PRR1, PP7 and AP205.

Common protein motifs such as casein kinase II phosphorylation, cAMP and cGMP-dependent protein kinase phosphorylation and protein kinase C phosphorylation occurred in DL52 and DL54 when compared to the *Levivirus* strains [3]. Interestingly, every amino acid motif position in all four genes was identical among these two JS strains.

2.4. Phylogenetic and Recombination Analyses

Cophenetic correlations showed the genogroup I strains, the JS subgroup strains, and the genogroup II strains all formed faithful clusters with correlations of 100, 90 and 98, respectively. The cluster cutoff method, however, showed only two relevant clusters being the genogroup I strains, which included fr and JS, and genogroup II strains (Figure 2).

Figure 2. Cophenetic cluster analysis of *Levivirus* (family *Leviviridae*) genogroups I and II strains generated from pairwise similarities of the replicase amino acid sequences. Horizontal bars at three of the branches show the standard deviations of the average similarities of the clusters. Numbers at each branch are the cophenetic correlations which represent the faithfulness of the clusters. Two relevant clusters, as determined by the cluster Cutoff method, are grouped as dictated by the dashed lines. Analysis performed in Bionumerics.
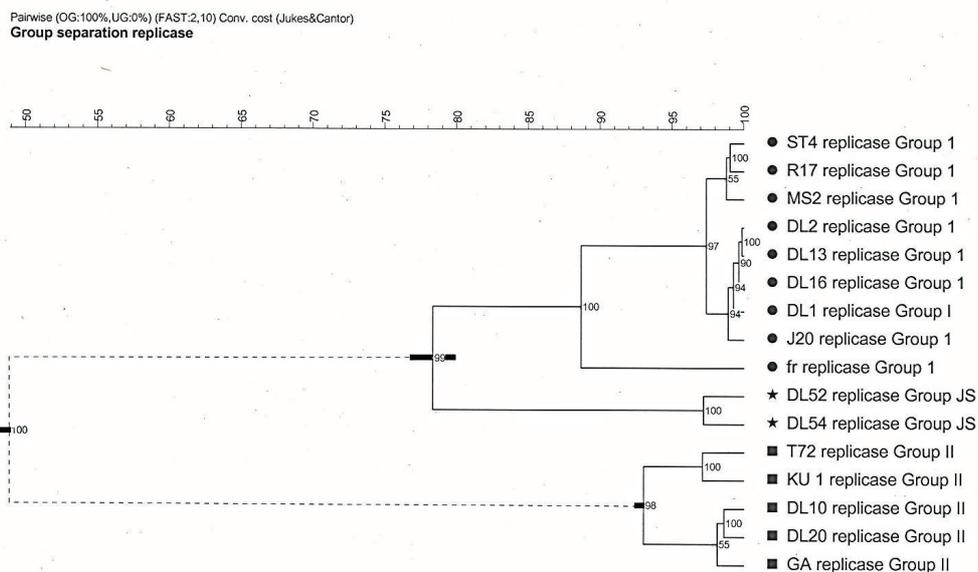
When referring to nucleotide or amino acid positions within the replicase gene, the numbering is in reference to the start codon as being position 1. In all analysis programs, the nucleotide or amino acid sequences were aligned to other strains and were therefore approximate positions on the replicase gene.

All recombination programs used, SimPlot, RAT, RDP3 and Recco, statistically predicted recombination in both JS strains, DL52 and DL54, when compared to genogroup I MS2-like strains. No recombination, however, was detected when DL52 and DL54 were compared to genogroup I strain fr and all genogroup II strains.

Figure 3. (a) The Simplot and bootscan analyses of the replicase nucleotides from JS strain DL52 queried to DL54, DL1, DL3, DL13, DL16, ST4, R17, J20 and MS2. The breakpoints are shown by the vertical red lines. The first recombination breakpoint occurred in the replicase gene in strain DL52 at nucleotide positions 787–818 where the χ^2 changed from 0.8 to 6.3 (sum χ^2 of 7.1). The second breakpoint occurred at nucleotide positions 979–1029 where the χ^2 changed from 0.6 to 7.0 (sum χ^2 of 7.6); **(b)** The Simplot and bootscan analyses of the replicase amino acids from JS strain DL52 queried to DL54, DL1, DL3, DL13, DL16, ST4, R17, J20 and MS2.
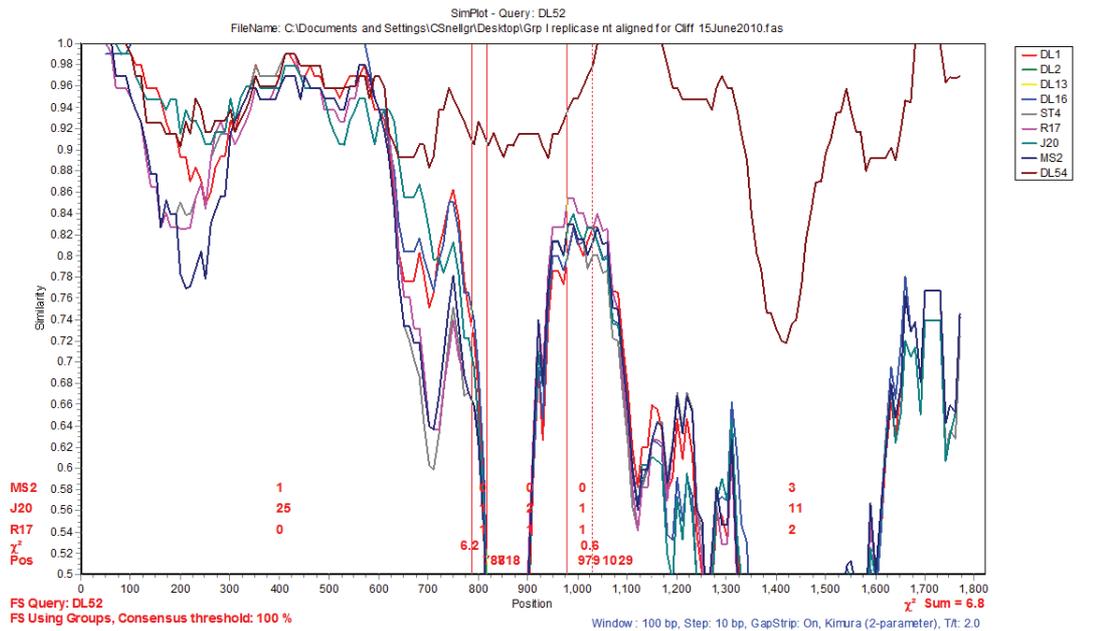
(a)

(b)
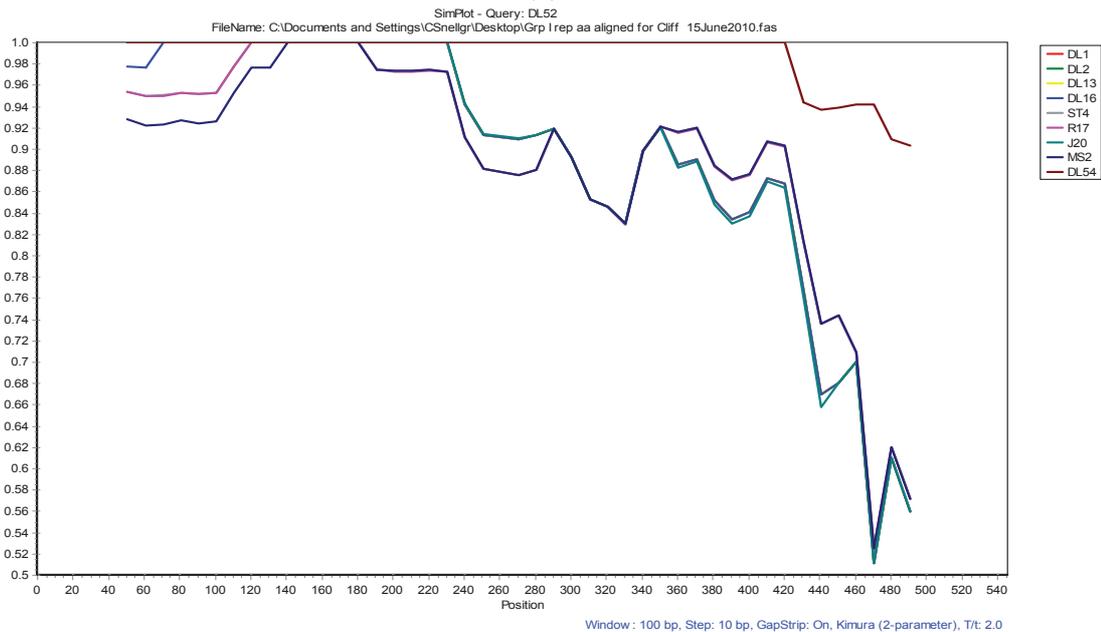
The Simplot and bootscan analyses of the replicase nucleotides from JS strains DL52 compared to *Levivirus* genogroup I strains DL54, DL1, DL3, DL13, DL16, ST4, R17, J20 and MS2 is shown in Figure 3A. Since the replicase nucleotide sequences in strain DL54 were 97% similar to strain DL52, DL52 was chosen as the query. The SimPlot analysis revealed the first recombination breakpoint occurred in the replicase from strain DL52 at nt positions 787–818 (approximate amino acid 262–273) where the χ^2 changes from 0.8 to 6.3 (sum χ^2 of 7.1). The second breakpoint occurred at nt positions 979–1029 (approximate amino acid 326–343) where the χ^2 changes from 0.6 to 7.0 (sum χ^2 of 7.6). However, Simplot amino acid analysis (Figure 3b) with strain DL52 showed a divergence at approximate amino acid position 460 region which is in agreement with the manual alignment (Figure 1).

Figure 4. Recombination analysis of the replicase nucleotide sequences from *Leviviridae* genogroup I strains DL13, DL16, ST4, R17, J20, MS2 and JS strains DL54, DL52 queried to DL1. Recombination Analysis Tool (RAT) was used to generate graphics with a window of 182 nt and step increments of 92 nt. The Y-axis represents the genetic distance and the X-axis is the sequence location along the genome. (**a**) The JS strains, depicted in green, diverged from the other genogroup I strains at approximate nucleotide (nt) position 660; (**b**) Recombination analysis of the replicase amino acid sequences from *Leviviridae* genogroup I strains DL13, DL16, ST4, R17, J20, MS2 and JS strains DL54, DL52 queried to DL1. Recombination Analysis Tool (RAT) was used to generate graphics with a window of 54 aa and step increments of 27 aa. The JS strains, DL52 and DL54, diverged from the other genogroup I strains at approximate amino acid 220 within the replicase gene.
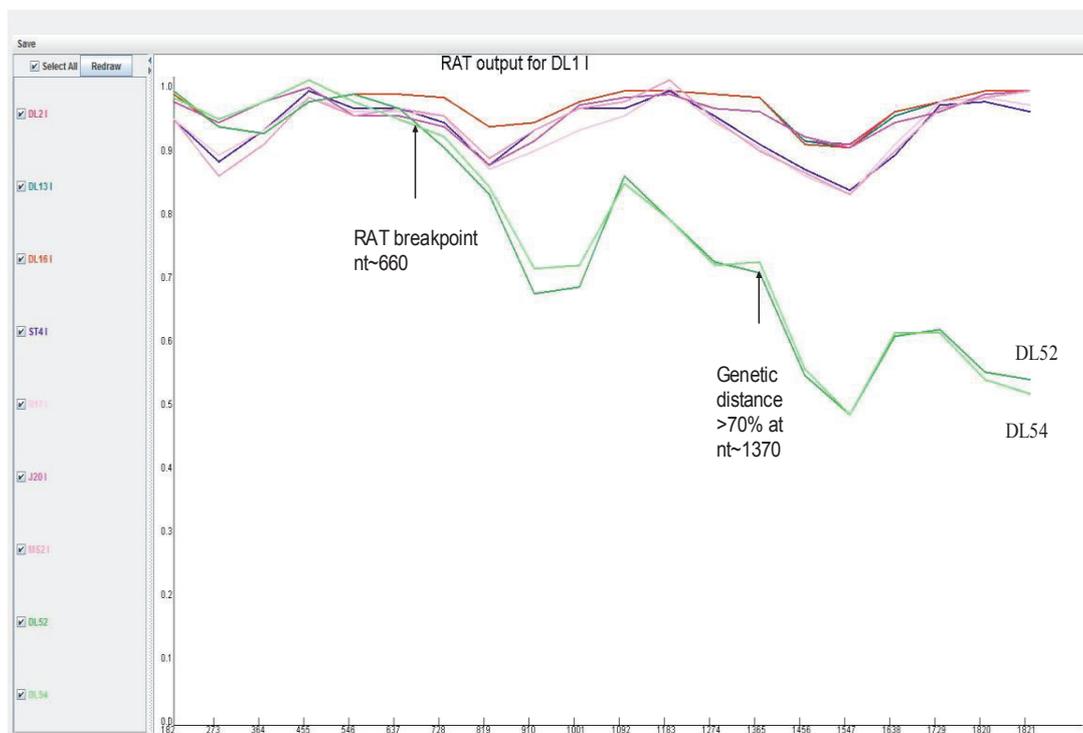
(**a**)

Figure 4. *Cont*.
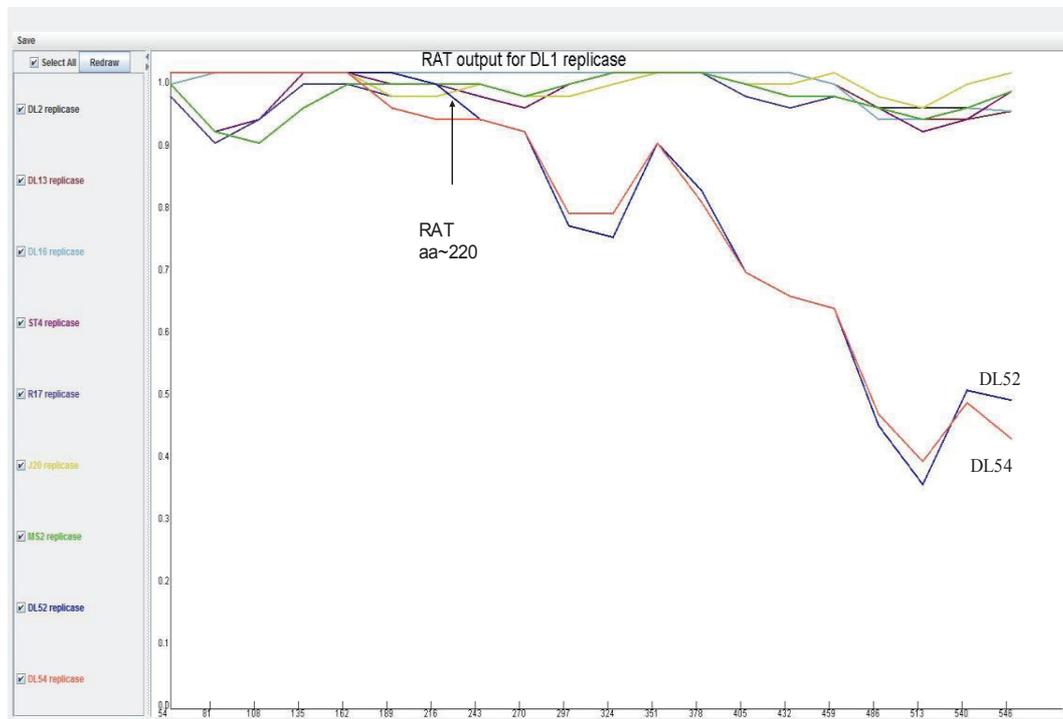
(b)

When analyzed with RAT, the nucleotide breakpoint (crossover) positions occurred at approximately nt 660 (Figure 4a) or amino acid 220 (Figure 4b) within the replicase gene. This crossover occurred when the two recombinant strains, DL52 and DL54, crossed the lines of the other MS-2 like strains and diverged by increasing genetic distances.

RDP3 predicted DL52 and DL54 as the recombinant strains using several detection methods and analysis algorithms (Table 4) and suggested DL16 as a minor parent strain. Breakpoint nucleotides for strains DL52 and DL54 (when aligned to genogroup I strains) occurred between nt 84–592 and 84–401, respectively (Figure 5a,b), corresponding to the approximate amino acid breakpoint positions of 133–197 within the replicase gene.

Table 4. Prediction of DL52 and DL54 as recombinant strains by analysis of *Levivirus* (family *Leviviridae*) genogroup I using Recombination Detection Program (RDP3).

Confirmation Table of Recombination Events		
Methods	**Events**	**Average p-value**
RDP	2	2.199×10^{-15}
GENECONV	1	3.031×10^{-27}
Bootscan	2	7.867×10^{-19}
MaxChi	2	1.445×10^{-10}
Chimaera	2	3.536×10^{-11}
SiScan	1	1.168×10^{-13}
3Seq	1	4.486×10^{-8}

190

Figure 5. (**a**) RDP3 analyses prediction of DL52 as a recombinant strain. Recombination area within the replicase gene is shown in pink beginning at nucleotide 84 and crossing over at 592, upstream from the catalytic domain. DL52 was queried to all *Levivirus* (family *Leviviridae*) genogroup I FRNA *Levivirus* strains. RDP3 suggested DL16 as the minor parental strain; (**b**) RDP3 analyses predicted DL54 as a recombinant strain. Recombination area within the replicase gene is shown in pink beginning at nucleotide 84 and crossing over at 401, upstream from the catalytic domain. RDP3 suggested DL16 as the minor parental strain.
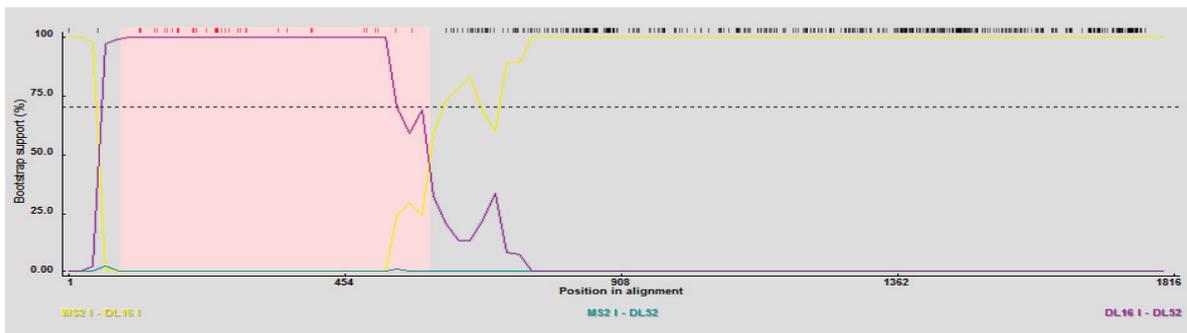
(**a**)
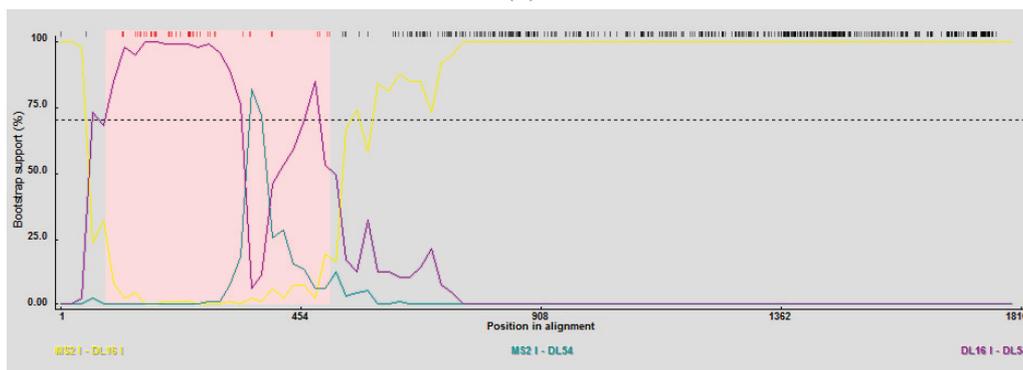
(**b**)

Manual alignment in BioEdit of the replicase nucleotides, counting the ATG start codon of the replicase gene as nt 1, showed an insertion of the nucleotides YA beginning at position 1374 (Figure 1) whereas the amino acid composition of the JS strains diverged from the other genogroup I strains slightly upstream from this insertion at amino acid position 455 (nucleotide 1366). Alignment also revealed numerous nt deletions as discussed in the "Sequence analyses and ORF" section.

The Recco p-value inspector predicted strain DL52 had recombined with strain DL1 (Figure 6a). In DL52, the recombinant region spanned from amino acids 181–212 whereas the DL1 region spanned from 396–457 with resulting sequence p-values of 0.000999 and 0.004995, respectively. Recco parametric cost curves predicted the highest preference for recombination in strains DL52 and DL54 (cost of 12.5–13) whereas the remaining genogroup I strains did not show a preference for recombination (cost of 0–3) (Figure 6b).

Figure 6. (**a**) Recco analysis of the RNA-dependent RNA polymerase (replicase) amino acid sequences in *Levivirus* genogroup I male-specific coliphages (FRNA). Recombination events are displayed by downward peaks in the graphics dataset. The upper graph represents the p-value for recombination at each position along the replicase gene. The lower graph is the breakpoint p-values for the entire set of *Levivirus* genogroup I and JS strains DL52 and DL54; (**b**) Recco parametric cost curve analysis of the RNA-dependent RNA polymerase (replicase) amino acid sequences for each FRNA strain in *Levivirus* genogroup I and JS strains DL52 and DL54. The y-axis corresponds to the cost curve and the x-axis represents α (0–1).
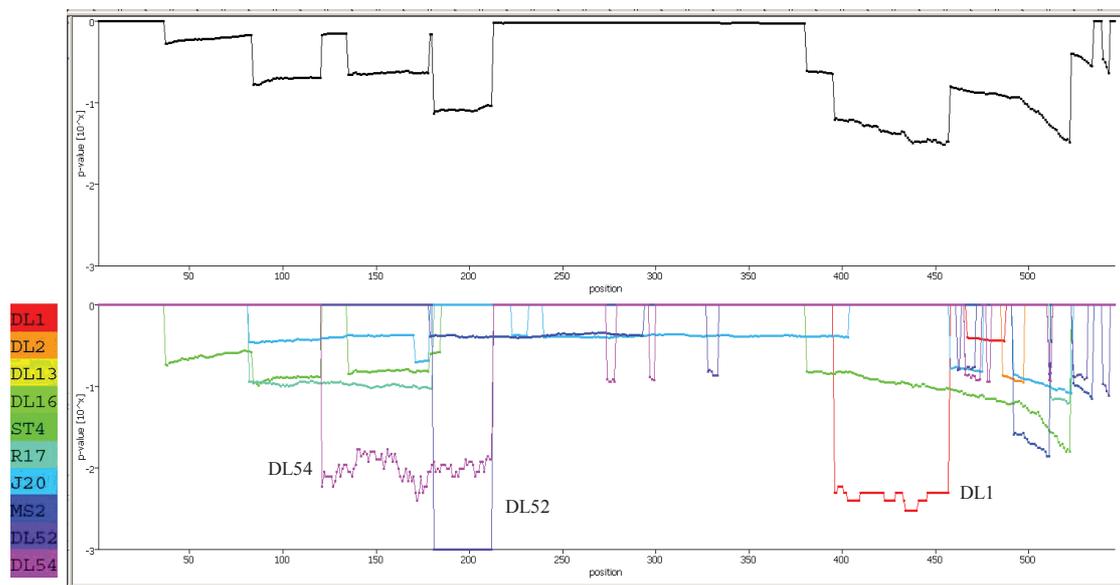
(**a**)
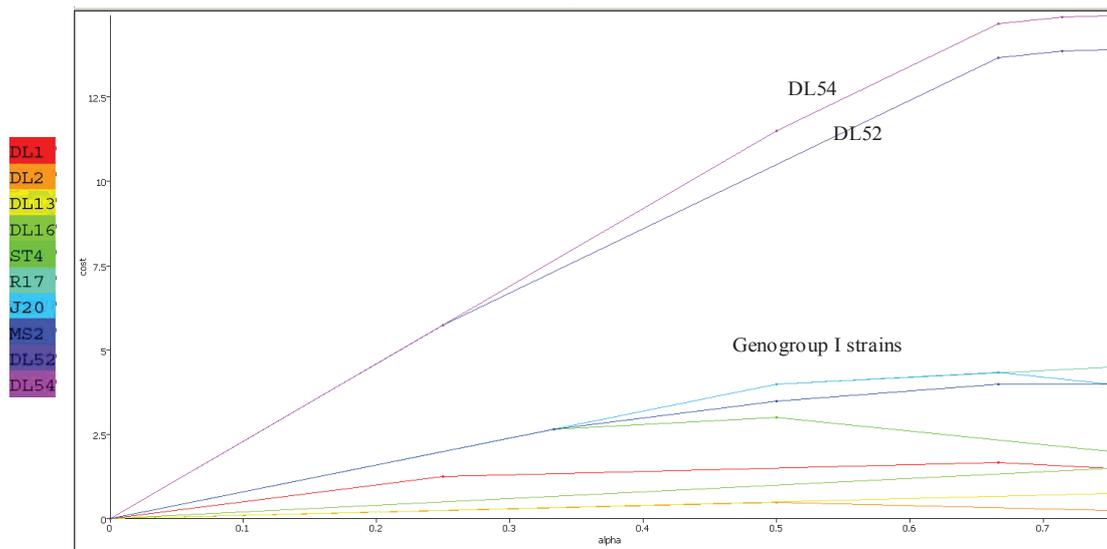
(**b**)

RAT, RDP3 and Recco all predicted recombination breakpoints ranging from amino acid positions 181–252 whereas Simplot agreed most closely with the manual alignment of 460 and 455, respectively. Also in agreement with the manual alignment was the crossover region between DL52 and DL1 occurring in the approximate amino acid region of 396–457 (Figure 6a). The predicted breakpoint regions occurred either upstream or downstream from the highly conserved catalytic domain amino acid positions 243–373 in *Levivirus* genogroup I [3].

3. Discussion

Reported here are whole genome sequence data and bioinformatic analyses supporting the hypothesis that two novel FRNA isolates, DL52 and DL54, were the result of natural recombination. Initially classified as a JS subgroup of genogroup I within the *Levivirus* genus (family *Leviviridae*), these strains were isolated from seawater approximately 1130 km apart in the Rachel Carson Reserve, Beaufort, NC, and Narragansett Bay, RI. Findings that JS strains were highly similar to three out of four genes (maturation, capsid and lysis) in genogroup I MS2-like strains, shared the catalytic site location in the RNA-dependent RNA polymerase (replicase) gene, had an identical 3' signature [3] and then greatly diverged along a stretch of the replicase gene all supported the occurrence of a recombination event.

In this study, two JS strains shared >95% amino acid identity in three (maturation, capsid and lysis) *Levivirus* genogroup I MS2-like genes but only an 84%–85% amino acid identity to the otherwise, highly conserved replicase protein. In comparison, genogroup I strain fr was uniformly different from all other genogroup I strains in all four proteins [3]. Cophenetic correlations and bootstrap analysis strengthen the possibility that JS strains were recombinants as the JS strains were only a subgroup of genogroup I and not a novel genogroup. Throughout the *Leviviridae* family, subgroups emerge within genogroups, however, subgroup strains differ in all four genes from the parent genogroup [3]. It is therefore plausible to propose natural recombination in these two novel JS-like FRNA coliphages as data presented here suggested a specific genetic rearrangement or recombination event in the replicase gene.

Interestingly, different *Leviviridae* subgrouped strains originating from across the globe display high amino acid similarity among subgrouped strains [3]. Recombination may explain why *Leviviridae* strains circulate as discrete subgroups independent of geographical location. Although the two unique JS-like strains were isolated from NC and RI, they shared 96.73% nucleotide similarity across the entire genome. Thus, either a single, natural recombination event occurred as a *de novo* mutation in each strain or identical, natural recombinations along a hot spot in these genomes formed these strains. In either case discovering that geographically-separated JS-like strains acquired the same recombination event is intriguing.

Largely responsible for the diversity of RNA viruses [12] RNA-RNA recombination was observed in several positive-sense, ssRNA human and animal viral taxa including caliciviruses, coronaviruses, hepatitis, dengue, enteroviruses and astroviruses [13–21]. For example, genetic exchange in ssRNA viruses was first demonstrated in polioviruses [22,23].

Recombination events frequently alter the RNA-dependent RNA polymerase region. Human Noroviruses, a positive sense ssRNA virus with a genome length of 7400–8300 nt, are considered to

belong to a prototype strain if they share approximately 85% overall nucleotide sequence identity and a high amino acid sequence identity (>95%) in the polymerase gene [20]. A naturally occurring human Norovirus strain shared 95% amino acid sequence identity with the capsid sequences from a Mexico cluster and 95% amino acid identity to the polymerase in a Lordsdale virus cluster. Sequences from the natural strain were obtained from one viral isolate. The combination of sequences in the one strain being complementary to two distinct human Norovirus clusters led to the proposition that this strain was a naturally occurring recombinant [20].

Genetic recombination is known to occur in certain Enteroviruses, a positive ssRNA virus having an approximate 7500 nt genome. Poliovirus recombination occurs in vaccine-derived strains [24] in the human population as a single infected individual excretes a high proportion of recombinants [18]. To determine if other enteroviruses undergo natural recombination, isolates of echoviruses were collected from a meningitis outbreak. Nucleotide sequences were clustered based on a capsid protein (VP1) and RNA-dependent RNA polymerase (3D). Dendrogram relatedness of the echovirus strains grouped the VPI sequences to the prototype strains. However, the RNA polymerase sequences did not cluster to the prototype strains, suggesting genetic recombination among the outbreak strains [18].

Human astroviruses are positive sense, ssRNA with a genome length of approximately 6,800 nucleotides [19] and a polyadenylated 3' tail [21]. Two sets of strains were investigated for recombination; one set was identified from a child care center in Houston, TX, and the two other strains were found in stool samples from two children in Mexico City. The pool of strains shared >97% nucleotide sequence similarity in two out of three genomic regions. The novel strain clustered to one group based on the capsid region. When the RNA-dependent RNA polymerase was analyzed, the novel strain clustered to a separate human astrovirus group. The strains were identified as naturally occurring recombinants on the evidence of high sequence similarity to a few genes of one prototype and similarity to different genes in a second prototype. A total of 64 additional human astroviruses lacked these novel traits [19].

An enteric turkey astrovirus is a non-enveloped, positive sense ssRNA virus with a polyadenylated 3' tailed genome of approximately 7 kb. The most conserved gene in the avian and mammalian astrovirus is the RNA-dependent RNA polymerase or replicase. Genetic analysis of capsid and polymerase sequences from twenty-three turkey astrovirus strains resulted in 8 clusters for the capsid gene and two phylogenetic clusters for the RNA polymerase gene. Computer-generated analyses identified polymerase gene recombination in strains of turkey astrovirus [14].

In this study, four different recombination detection programs along with manual alignment predicted strains DL52 and DL54 as recombinants although the exact amino acid and/or nucleotide breakpoint varied somewhat along the RNA-dependent RNA polymerase gene. As expected the breakpoints did not occur within the catalytic-site domain. The sliding window approach as used with many recombination programs is based on an arbitrarily chosen window length, thus affecting the sensitivity and accuracy when pinpointing the precise breakpoint [25]. Overall, the use of a variety of recombination algorithms provided a stronger, more rigorous scientific case. When comparing the *Levivirus* strains, manual alignment provided a more accurate picture of where the recombination

event occurred along the genome but it did not provide a statistical analysis. Therefore, statistical analysis in combination with manual alignment resulted in a more confident assessment of recombination. Evidence for recombination among positive ssRNA viruses exists within the RNA-dependent RNA polymerase (replicase) gene for numerous viruses as described here and supports the data that natural recombination can occur within the *Leviviridae* family.

4. Experimental Section

4.1. FRNA Coliphage Strains and RNA Extraction

FRNA phage strains CICEET 29 and CICEET 24 were isolated and placed into the putative JS subgroup [10,11] using the genotyping methods of Vinjé *et al*., [4]. CICEET 29, renamed DL52, was isolated from estuarine waters in Rachel Carson W Reserve (Beaufort), NC, and CICEET 24, renamed DL54, and was isolated from Narragansett Bay, RI (Table 1).

Each strain was plaque purified and further enriched using *Escherichia coli* HS(pFamp)R as host [4]. Approximately 1–2 mL aliquots of the purified viral supernatant were frozen at &75 °C. Coliphage RNA was extracted from purified virus as described [26] using a QIAamp viral RNA mini kit (Qiagen, Valencia, CA, USA). Purified RNA was stored frozen at &20 °C.

4.2. Sequencing and Analysis

Full-length genome sequencing was performed by the "primer walking" approach as described [3]. Nucleotide and amino acid sequences from JS strains DL52 and DL54 were compared to nucleotide and amino acid sequences from 10 genogroup I strains (MS2, DL1, DL2, DL13, DL16, ST4, R17, J20, M12, fr) and 5 genogroup II strains (T72, DL10, DL20, GA, KUI) [3].

Nucleotide sequences from three individual clones were imported and aligned using BioEdit v7.0.1 [27] followed by Basic Local Alignment Search Tool (BLAST, National Center for Biotechnology Information) analyses for sequence and phylogenetic confirmation. Completed sequences from all strains were aligned with full-length prototype strains (GenBank) using BioEdit ClustalW application. For each strain, the Open Reading Frames (ORFs) were mapped using BioEdit.

4.3. Amino Acid Analysis

Deduced amino acid sequences for each of the four genes were determined using a computer-generated DNA-to-protein translation tool, ExPASY (http://ca.expasy.org/). Prediction of protein sequence motifs were identified by PROSITE (http://ca.expasy.org/) and protein families and domains were modeled in Pfam (http://pfam.janelia.org).

4.4. Phylogenetic, Statistical and Recombination Analyses

Sequence data were analyzed using BioNumerics Software v.3.5 (Applied Maths, Saint-Martens-Latem, Belgium). Phylogenetic trees were built by global cluster analysis performed on multiple aligned

sequences and clustered by unweighted pair group method using arithmetic averages (UPGMA). A bootstrap analysis, based on 10,000 substitutions, was used to measure cluster significance. The reliability of each cluster was expressed on a percentage basis [3].

Nucleotide percent similarity and dendrograms were constructed using BioNumerics Software v.3.5 (Applied Maths, Saint-Martens-Latem, Belgium). Phylogenetic trees were built by global cluster analysis performed on multiple aligned sequences and clustered by UPGMA using the Jukes and Cantor correction [28]. Cophenetic correlations and cluster Cutoff method were employed to measure faithfulness and relevancy of the clusters (Applied Maths, Saint-Martens-Latem, Belgium). Average similarities with standard deviations were calculated for the relevant clusters.

Various approaches were used to examine recombination in *Levivirus* FRNA strains, using aligned nucleotides and aligned amino acids, as follows: (i) manual alignment using BioEdit; (ii) bootscan analysis in SimPlot v3.5.1 [29] (iii) Recombination Analysis Tool v1.0 [30], (iv) Recombination Detection Program v3.44 [31]; and (v) Recombination Analysis Using Cost Optimization (Recco) v0.93 [32] FRNA strains used in analyses were genogroup I strains DL1, DL2, DL13, DL16, ST4, R17, J20, MS2, fr; JS strains DL52 and DL54; and genogroup II strains GA, KU1, DL10, DL20 and T72.

SimPlot analyses was determined with a sliding window of 100 bp wide and a step size between plots of 10 bp when comparing reference strains to the queried sequences. Recombination events by SimPlot bootscan analysis occurred when the χ^2 value changes signifying a breakpoint position.

Aligned replicase amino acids and/or nucleotide sequences were analyzed in Recombination Analysis Tool (RAT), Recombination Detection Program (RDP3) and Recco. RAT uses a distance-based method of recombination in both DNA and protein multiple alignments [30]. Unless stated otherwise, default settings were used with each program. The RAT default settings were window size of 10% of the sequence length and an increment size being half of the window size. Both settings of "auto search" and "test sequence search" were used with RAT.

Recombination Detection Program v3 (RDP3) uses a number of recombination detection algorithms such as RDP, Bootscan, GENECONV, Maximum Chi Square, CHIMAERA, Sister Scanning (SISCAN) and 3SEQ [31]. The RDP3 program sorts the analyses from these various algorithms and statistical data to determine the unique recombination events. RDP3 used an alignment of all genogroup I and JS replicase nucleotide sequences and queried to DL16, MS2, DL52 and DL54.

The Recco p-value inspector was set at 3 and the permutation was set at 1000. Recco uses an algorithm that locates putative recombination points based on cost minimization. Recco compares the cost of mutation relative to recombination as represented by α.

4.5. Nucleotide Sequence Accession Numbers

The accession numbers for DL52 and DL54 are JQ966307 and JQ966308, respectively.

5. Conclusions

The results of this study provide genetic evidence, bioinformatic and statistical analyses suggesting a natural recombination event in the formation of a genogroup I subgroup JS-like *levivirus,* represented by two strains, DL52 and DL54. There was high nucleotide and amino acid identity in three genes, the maturation, capsid and lysis genes (≥95%) but a lack of similarity in the replicase gene (84%–85%) when JS strains were compared to genogroup I MS2-like strains. Four different recombination programs demonstrated one or two breakpoint regions in the replicase gene, signifying a recombination event. The recombination event occurred downstream of the replicase catalytic site thereby maintaining viral integrity and replication function. Thus, primers for oligonucleotide hybridization probes targeting the replicase beyond the catalytic site would not hybridize to JS strains. In contrast, molecular assays targeting the maturation, capsid or lysis sequences would presumptuously place JS strains as an MS-2 like genogroup.

Phylogenetic tree analysis produced a cophenetic correlation which showed (i) ten genogroup I strains, including strain fr; (ii) the JS subgroup strains; and (iii) the genogroup II strains all formed faithful clusters with correlations of 100, 90 and 98, respectively. The cluster cutoff method, however, revealed only two relevant clusters, (i) genogroup I strains, which included fr and JS; and (ii) genogroup II strains. Therefore, the novel JS strains are not a unique *Levivirus* genogroup. The proposed classification of JS strains is genogroup I subgroup "JS-like".

Although both JS strains were prepared for sequencing in the same laboratory, these strains were field-collected by different investigators and shipped to another location where they were plaque-purified and preliminarily classified. Therefore, the possibility that contamination resulted in false recombinants seems unlikely. Likewise, the possibility of cloning and/or PCR errors contributing to the nucleotide and amino acids changes would not have led to both JS strains being almost identical in the non recombinant regions as seven genogroup I strains (DL1, DL2, DL13, DL16, R17, J20 and ST4), three genogroup II strains (DL10, DL20, T72) and two JS strains (DL52 and DL54) were sequenced in this lab in no certain strain or fragment order using the same methods and sequencing company [3]. Finally, to the best of our knowledge, this is the first description of recombinant viruses from natural isolates in ssRNA *Leviviridae* bacteriophages.

The information in this document has been funded wholly (or in part) by the U.S. Environmental Protection Agency. It has been subjected to review by the National Health and Environmental Effects Research Laboratory and approved for publication. Approval does not signify that the contents reflect the views of the Agency, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Fiers, W.; Contreras, R.; Duerinck, F.; Haegeman, G.; Iserentant D.; Merregaert, J.; Min Jou, W.; Molemans, F.; Raeymaekers, A.; van den Berghe, A.; *et al*. Complete nucleotide sequence of bacteriophage MS2 RNA: Primary and secondary structure of the replicase gene. *Nature* **1976**, *260*, 500–507.
2. Sanger, F.; Air, G.M.; Barrell, B.G.; Brown, N.L.; Coulson, A.R.; Fiddes, J.C.; Hutchinson, C.A. III; Slocombe, P.M.; Smith M. Nucleotide sequence of bacteriophage Φ-X174 DNA. *Nature* **1977**, *265*, 687–695.
3. Friedman, S.D.; Genthner, F.J.; Gentry, J.; Sobsey, M.D; Vinjé, J. Gene mapping and phylogenetic analysis of the complete genome from 30 single-stranded RNA male-specific coliphages (family *Leviviridae*). *J. Virol.* **2009a**, *83*, 11233–11243.
4. Vinjé, J.; Oudejans, S.J.G.; Stewart, J.R.; Sobsey, M.D.; Long, S.C. Molecular detection and genotyping of male-specific coliphages by reverse transcription-PCR and reverse line blot hybridization. *Appl. Environ. Microbiol.* **2004**, *70*, 5996–6004.
5. Chetverin, A.B. The puzzle of RNA recombination. *FEBS Lett.* **1999**, *460*, 1–5.
6. Lutay, A.V.; Zenkova, M.A.; Vlassov, V.V. Nonenzymatic recombination of RNA: Possible mechanisms for the formation of novel sequences. *Chem. Biodiv.* **2007**, *4*, 762–767.
7. Chetverin, A.B.; Kopein, D.S.; Chetverina, H.V.; Demidenko A.A.; Ugarov V.I. Viral RNA-directed RNA polymerases use diverse mechanisms to promote recombination between RNA molecules. *J. Biol. Chem*. **2005**, *280*, 8748–8755.
8. Horiuchi, K. Genetic Studies of RNA Phages. In *RNA Phages*; Zinder, N.D., Ed.; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY, USA, 1975; pp. 29–50.
9. Munishkin, A.V.; Voronin, L.A.; Chetverin, A.B. An *in vivo* recombinant RNA capable of autocatalytic synthesis by Qβ replicase. *Nature* **1988**, *333*, 473–475.
10. Sobsey, M.D.; Love, D.C.; Lovelace, G.L. F+RNA coliphages as source tracking viral indicators of fecal pollution. A final report submitted to the NOAA/UNH Cooperative Institute for Coastal and Estuarine Environmental Technology (CICEET). **2006**.

11. Love, D.C.; Vinjé, J.; Khalil, S.M.; Murphy, J.; Lovelace, G.L.; Sobsey, M.D. Evaluation of RT-PCR and reverse line blot hybridization for detection and genotyping F+ RNA coliphages from estuarine waters and molluscan shellfish. *J. Appl. Microbiol*. **2008**, *104*, 1203–1212.

12. Lai, M.M.C. RNA recombination in animal and plant viruses. *Microbiol. Rev*. **1992**, *56*, 61–79.

13. Cristina, J.; Colina, R. Evidence of structural genomic region recombination in Hepatitis C virus. *Virol. J.* **2006**, *3*, 53–60.

14. Pantin-Jackwood, M.J.; Spackman, E.; Woolcock, P.R. Phylogenetic analysis of turkey astroviruses reveals evidence of recombination. *Virus Genes* **2006**, *32*, 187–192.

15. Holmes, E.C.; Worobey, M.; Rambaut, A. Phylogenetic evidence for recombination in Dengue virus. *Mol. Biol. Evol.* **1999**, *16*, 405–409.

16. Oberste, M.S.; Maher, K.; Pallansch, M.A. Evidence for frequent recombination within species *Human Enterovirus B* based on complete genomic sequences of all thirty-seven serotypes. *J. Virol.* **2004**, *78*, 855–867.

17. Banner, L.R.; Lai, M.M.C. Random nature of coronavirus RNA recombination in the absence of selection pressure. *Virolgy* **1991**, *185*, 441–445.

18. Oprisan, G.; Combiescu, M.; Guillot, S.; Caro, V.; Combiescu, A.; Delpeyroux, F.; Crainic, R. Natural genetic recombination between co-circulating heterotypic enteroviruses. *J. Gen. Virol.* **2002**, *83*, 2193–2200.

19. Walter, J.E.; Briggs, J.; Guerrero, M.L.; Matson, D.O.; Pickering, L.K.; Ruiz-Palacios, G.; Berke, T.; Mitchell, D.K. Molecular characterization of a novel recombinant strain of human astrovirus associated with gastroenteritis in children. *Arch. Virol.* **2001**, *146*, 2357–2367.

20. Jiang, X.; Espul, N.; Zhong, W.M.; Cuello, H.; Matson, D.O. Characterization of a novel human calicivirus that may be a naturally occurring recombinant. *Arch. Virol.* **1999**, *144*, 2477–2387.

21. Belliot, G.; Laveran, H.; Monroe, S.S. Detection and genetic differentiation of human astroviruses: Phylogenetic grouping varies by coding region. *Arch. Virol.* **1997**, *142*, 1323–1334.

22. Hirst, G.K. Genetic recombination with Newcastle disease virus, polioviruses, and influenza. *Cold Spring Harb. Symp. Quant. Biol.* **1962**, *27*, 303–309.

23. Ledinko, N. Genetic recombination with poliovirus type 1 studies of crosses between a normal horse serum-resistant mutant and several guanidine-resistant mutants of the same strain. *Virology* **1963**, *20*, 107–119.

24. Kew, O.; Morris-Glasgow V.; Landaverde, M.; Burns C.; Shaw J.; Garib, Z.; André J.; Blackman, E.; Freeman C.J.; Jorba, J.; *et al*. Outbreak of poliomyelitis in Hispaniola associated with circulating type 1 vaccine-derived poliovirus. *Science* **2002**, *296*, 356–359.

25. Lee, W.H.; Sung, W.K. RB-finder: An improved distance-based sliding window method to detect recombination breakpoints. *J. Comput. Biol.* **2008**, *15*, 881–898.

26. Stewart, J.R.; Vinjé, J.; Oudejans, S.J.G.; Scott, G.I. Sequence variation among group III F-specific RNA coliphages from water samples and swine lagoons. *Appl. Environ. Microbiol*. **2006**, *72*, 1226–1230.

27. Hall, T.A. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* **1999**, *41*, 95–98.

28. Jukes, T.H.; Cantor, C.R. Evolution of Protein Molecules. In *Mammalian Protein Metabolism*; Munro, H.N., Ed.; Academic Press: New York, NY, USA, 1969; pp. 21–132.

29. Lole, K.S.; Bollinger, R.C.; Paranjape, R.S.; Gadkari, D.; Kulkarni, S.S.; Novak, N.G.; Ingersoll, R.; Sheppard, H.W.; Ray S.C. Full-length human immunodeficiency virus type 1 genomes from subtype c-infected seroconverters in India, with evidence of intersubtype recombination. *J. Virol.* **1999**, *73*, 152–160.

30. Etherington, G.J.; Dicks, J.; Roberts, I.N. Recombination analysis tool (RAT): A program for the high-throughput detection of recombination. *Bioinformatics* **2005**, *21*, 278–281.

31. Martin, D.P.; Williamson, C.; Posada, D. RDP2: Recombination detection and analysis from sequence alignments. *Bioinformatics* **2005**, *21*, 260–262.

32. Maydt, J.; Lengauer, T. Recco: Recombination analysis using cost optimization. *Bioinformatics* **2006**, *22*, 1064–1071.

Review

Interaction of Bacteriophage λ with Its *E. coli* Receptor, LamB

Sujoy Chatterjee and Eli Rothenberg *

Department of Biochemistry and Molecular Pharmacology, NYU Medical School, 550 First Avenue, New York, NY 10016, USA; E-Mail: sujoy.chatterjee@nyumc.org

* Author to whom correspondence should be addressed; E-Mail: Eli.Rothenberg@nyumc.org; Tel.: +1-212-263-5622.

Abstract: The initial step of viral infection is the binding of a virus onto the host cell surface. This first viral-host interaction would determine subsequent infection steps and the fate of the entire infection process. A basic understating of the underlining mechanism of initial virus-host binding is a prerequisite for establishing the nature of viral infection. Bacteriophage λ and its host *Escherichia coli* serve as an excellent paradigm for this purpose. λ phages bind to specific receptors, LamB, on the host cell surface during the infection process. The interaction of bacteriophage λ with the LamB receptor has been the topic of many studies, resulting in wealth of information on the structure, biochemical properties and molecular biology of this system. Recently, imaging studies using fluorescently labeled phages and its receptor unveil the role of spatiotemporal dynamics and divulge the importance of stochasticity from hidden variables in the infection outcomes. The scope of this article is to review the present state of research on the interaction of bacteriophage λ and its *E. coli* receptor, LamB.

Keywords: bacteriophage; lambda phage; LamB receptor; protein J; single-virus tracking

1. Introduction

Viral infections are initiated through a binding process, which involves a specific interaction between the virus and the host cell surface [1–3]. The nature of initial virus host interaction greatly varies amongst different systems, and is mediated by various cell surface receptors, co-receptors and

other molecules, depending on the specific virus and host [4–6]. Nevertheless, the underlining mechanism of this interaction relies on a diffusion limited viral-receptor finding process. This viral receptor-finding process is crucial for the propagation of the infection process as it determines the fate of infection [4,6]. Here we review the studies of this interaction focused on the *Escherichia coli* bacterium and its virus, bacteriophage λ as a virus-host model system.

Since the seminal discovery of bacteriophage λ, it has drawn paramount interest by the molecular biologists and served as an excellent genetic tool for studying fundamental principles of biology such as gene regulation, DNA replication, homologous and site-specific recombination [7–9]. Even today, the λ system continues to yield new insights into its gene regulatory circuits. The best-characterized feature of phage λ biology is the genetic switch that decides whether a phage propagates by cell lysis, or integrates into host genomes to become a prophage [10–12]. Other well studied aspects of phage λ biology is the adsorption of the phages and subsequent DNA delivery into the host *E. coli* cell [13–16]. Phage λ uses the bacterial maltose pore LamB (λ-receptor) for delivery of its genome into the bacterial cell. The interaction between λ and its receptor has been extensively studied both biochemically using purified components and genetically by either mutated phage or bacterial receptor [17–25]. However, the quantitative understanding of the dependence of viral target-finding on virus-receptor interactions and cellular architecture came from recent single virus tracking studies [26,27]. In this review we focus on the interaction of phage λ and its receptor LamB. We will summarize previous biochemical and structural studies of phage λ and its receptor (LamB), and provide an overview of a recent study of virus target searching mechanisms at single molecule level.

2. Structure of the LamB Receptor and Interaction with Bacteriophage λ

Bacteriophage λ is one of the well studied models in molecular biology. Although a vast amount of information is available about the gene regulation network of λ [10–12], the quantitative depiction of host cell infection was determined only recently [26]. Bacteriophage λ consists of an icosahedrally symmetric (5,3,2 rotational symmetries) head of diameter 60 nm encapsulating the 48,502 bp double strand DNA molecule and a flexible tail through which the viral DNA expel during infection (Figure 1) [28,29]. Following infection, the invading phage DNA can either replicate within the host, forming new phages and propagate by lysing the host cell (lytic pathway), or it can become a prophage by integrating its DNA into the host chromosome, which then replicate as a part of host chromosome (lysogenic pathway). However, a switch from lysogenic to lytic pathway can be induced where the prophages can replicate independently, assemble the head and tail, forming new viruses and promoting lysis for further propagation [10–12]. The *in vivo* head and tail assembly are complex process and beyond the limit of present discussion [29,30]. Early genetic experiments showed that most of the *E. coli* K-12 mutations resistant to λ phage are located in two genetic regions *malA* and *malB* [14,16,21,22]. This *malB* region contains a gene *lamB* whose product, LamB, involves in the λ receptor synthesis, a component of *E. coli* outer membrane. However, a recent study shows that mutant form of bacteriophage λ can target alternative receptor. When phage cI26 (a strictly lytic derivative of phage λ) was cultured with *E. coli* in condition that suppressed the expression of LamB, mutant phage changed

their specificity from LamB to a new receptor, OmpF [31]. A combination of four mutants in phage tail protein J were required for targeting this new receptor. It is noteworthy that some host mutations prevented phage from evolving this new function, demonstrating the complexity of interactions in a co-evolving population. The adsorption of λ phages onto bacterial surface is the first step in the infection process [15]. At this stage, phages can either dissociate from the host cell, known as desorption or alternatively bind irreversibly to the host cell [14]. However, once the phage irreversibly bind to the cell surface it triggers a series of poorly understood events and finally delivers its DNA into the bacterial cytoplasm through the channel formed by its tail, leaving the phage protein capsid behind. The tail fibers of bacteriophages are also important to make specific contacts with receptor molecules on the surface of the bacterial cell. The common laboratory strain of bacteriophage λ, so called λ wild type carries a frameshift mutation in *stf* gene relative to Ur-λ, the original isolate. The Ur-λ phages have thin tail fibers which are absent in λ wild type and the Ur-λ has expanded receptor specificity and adsorbs to host cells more rapidly, suggesting the importance of tail fibers [32]. The process between phage adsorption and DNA injection can be sub divided in three steps: 'lag', 'trigger' and 'uptake' [33]. It has been shown that the free tail can itself adsorb to the cell surface, however, the head attachment is required for the lag reaction [9,29,34].

Figure 1. A cartoon representation of λ phage.
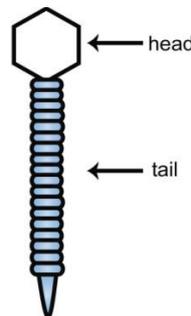
Gram-negative bacteria have two membranes [35]; the outer monolayer is composed of lipopolysaccharide (LPS) as its major lipid, and the inner leaflet contains mostly phosphatidylethanolamine, small amounts of phosphatidylglycerol and cardiolipin [36]. The outer membranes of Gram negative bacteria protect cells from harmful agents by slowing down their penetration while allowing uptake of nutrients [37–39]. The major proteins of outer membrane are β-barrel proteins, which play crucial roles in adhesion and virulence. In Gram negative bacteria β-barrel proteins are found exclusively in the outer membranes and contributed ~2%–3% of the Gram negative proteome [40]. However, their occurrence is not restricted to Gram negative bacteria and also found in Mycobacteria, mitochondria and chloroplasts [41–44]. Transmembrane β-barrels also found in several microbial toxins such as heptameric pore-forming α-haemolysin from Staphylococcus aureus, toxin aerolysin and the anthrax-protective antigen [45–47].

Well characterized constitutive outer membrane proteins are porins. They form a general diffusion pore with a defined exclusion limit for hydrophilic solutes within the outer membrane. In addition to

the constitutive porins, the outer membrane may contain porins that are induced only under special growth conditions. The most abundant porins in *Escherichia coli*, the OmpF and OmpC porins, are called general porins, allowing passage of hydrophilic solutes up to a size limit determined by the constrictions of their channels [37]. The biogenesis of β-barrel proteins is complex owing to their intrinsic structure. The insertion of individual β-strands into the hydrophobic core of the bilayer is thermodynamically restricted as they cannot assemble from previously inserted transmembrane segments due to the lack of hydrogen bonds of individual β-strands of β-barrel proteins. Henceforth, membrane insertion must be orchestrated with acquisition of both secondary and tertiary structure to produce the contiguous membrane-spanning barrel held by inter-strand hydrogen bonds [37–39]. Furthermore, the porins must also attain their quaternary structure, presumably by associations formed in the membrane.

The LamB protein is a well-characterized example of a porin, termed as a maltoporin, because it is required for growth on limiting concentrations of maltose. The protein coded by gene *lamB* of the maltose operon also serves as receptors for several phages, such as λ, K10 and TP1 [48–50]. The molecular weight of LamB is 135.6 kDa, looking like a half-open tulip, formed by 3 identical subunits, each one having a molecular weight of 45.9 kDa [51–53]. A major contribution to understanding the molecular basis of the λ phage interaction with LamB receptor has come from determination of the crystal structure of LamB [54]. This study showed that each subunit of the trimeric protein formed by an 18-stranded antiparallel β-barrel, which form a wide channel with a diameter of about 2.5 nm (Figure 2A,B). Loops are found at the end of barrel. Three loops, L1, L3 and L6 (in Figure 2C, they are colored as green, red and yellow respectively) interact with the Loop 2 from an adjacent subunit and packed against the inner wall of the barrel and line the channel. The other loops form a compact structure at the cell surface. Six aromatic residues (Y6, Y41, W74, W358, W420 and F227) lining up the channel interior, forms the 'greasy slide', and actively participate in carbohydrate (maltooligosaccharides are in apolar van der Waals contact with the "greasy slide") transport. Tyrosine 118 (Y118), located opposite of the greasy slide has a major impact on ion and carbohydrate transport through LamB [54,55]. When all six residues of the greasy slide are mutated to alanine, the mutation Y118W is sufficient to confer to LamB maltopentaose transport *in vivo* and maltopentaose binding *in vitro*. How the mutation in LamB could correlate with the resistance to phage λ infection? We have summarized the mutants of LamB that conferred resistance for phage λ in Table 1 [56]. Only about half of the phage λ resistant mutants are surface exposed, and are labeled in yellow in Figure 2D [54]. The remaining mutations (labeled in red in Figure 2D) might have indirect effect. They may alter or the cause structural change on the surface, or they might alter the dynamic behavior of the loops. For instance, one mutation, G18V, is known to affects the stability of trimers and may thus have long-range effects [56]. Similar to the LamB receptor, a number of studies had been carried to understand how λ utilizes the LamB receptor to inject its DNA inside the host cell [24,25]. Phage λ tail contains a hollow tube that consists of 32-stacked disks, where each disk is formed by six subunits of the major tail protein gpV, arranged such that each disk has a central, 3 nm hole. Phage λ uses this channel to eject its DNA, though the tail by itself (without the head) can attach to the host cell. Genetic

evidence indicates that gpJ directly interacts with the outer membrane protein LamB during the attachment of the bacteriophage to the surface of the cell. When the J gene from phage λ was substituted with the tail fiber gene from a closely related bacteriophage 434, the resulting phage was found to bind to a different membrane receptor OmpC [16], which phage 434 uses for infection [57]. Further studies indicate that the C-terminal of gpJ protein determines the host specificity of the phage [24,25]. Electron microscopy imaging of soluble LamB receptor and LamB protein incorporated into liposomes revealed the presence of two different types of bacteriophage λ—LamB complexes. In one type of the complexes binding occurs near the end of the tail fiber (J protein), while in the other type of complexes the distal end of the tail tube was directly and irreversibly attached to the receptor particles or the liposome. Genetic studies showed that the mutations in λ phages that have compensatory effect with the LamB mutants tightly blocks the phage λ adsorption are located in the C-terminal portions of J [24]. In fact, using a chimeric protein comprising the last 249 amino acids of J in fusion to the C-terminal end of the carrier maltose binding protein (MBP) could bind to LamB trimers and inhibited recognition by anti-LamB antibody [17]. Electron microscopy study showed that this chimeric J protein could also bind to the LamB at the cell surface and this interaction prevented λ adsorption. In this same study, when this chimeric protein was reconstituted with either LamB or the loop deletion mutant LamB 4+6+9v, both of them showed similar blocking of ion current in the lipid bilayer experiment, which indicated that the phage λ binding includes not only the extracellular loops.

Table 1. Mutation in LamB receptor that confers resistance to lambda phage infection (From [56]).

Residue(s)	Substitution
18	Gly → Val
148	Glu → Lys
151	Gly → Asp
152	Ser → Phe
154	Ser → Phe
155	Phe → Ser
163	Tyr → Asp
164	Thr → Pro
245	Gly → Arg
245	Gly → Val
247	Ser → Leu
249	Gly → Asp
250	Ser → Phe
259	Phe → Val
259	Phe → Val
382	Gly → Asp
382	Gly → Val
401	Gly → Asp

Figure 2. Crystal structure of 18 antiparallel β-strands barrel trimer LamB. (**A**) Side view and (**B**) Top view. Three monomers are shown by different colors (red, yellow and green). (**C**) Interaction of loops in barrel. Three loops L1 (green), L3 (red) and L6 (yellow) interact with L2 from an adjacent subunit and packed against the inner wall of the barrel. (**D**) The mutation(s) of LamB confer resistance to phage infection. The surface exposed mutant shown in yellow and the others in red. Mutation G18V is known to affect the stability of trimer and shown in green. The Adopted and modified after [54,56].
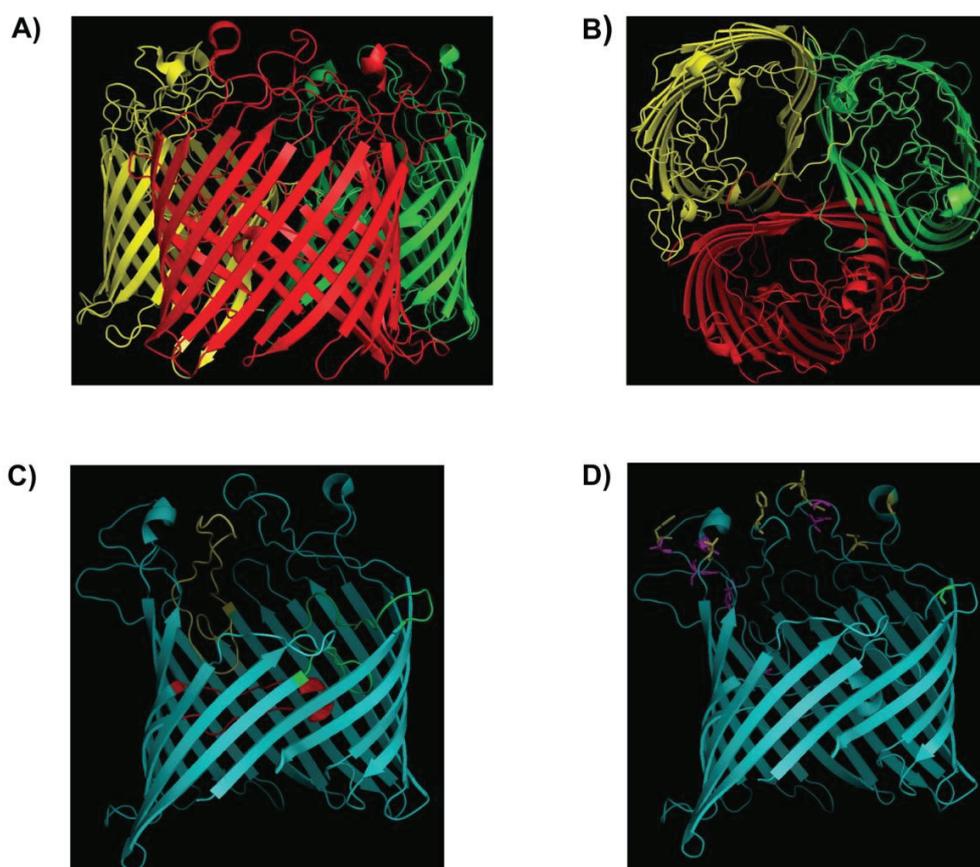
3. The λ Phage LamB Target Finding Process

Numerous genetic and biochemical studies had been carried out to provide the nature of interaction of the LamB receptor and protein J. Recently, a study of the spatiotemporal dynamics by which a phage λ, initially diffusing in bulk, arrives at a specific LamB site on the cell surface, provided a fundamental understanding of the initial viral-host target finding process.

The classical problem of diffusion limited target finding on a cell surface was first addressed by the concept of reaction rate enhancement by dimensional reduction (RREDR), introduced by Adam and Delbrück [58]. To explain the enhancement of adsorption rates they proposed a two-stage capture model , where free phages diffuse in three dimensions (3D), but once in contact with the cell surface this 3D motion is replaced by a two-dimensional (2D) 'random walk' on the cell surface, until it is

captured by the receptor. This model of reduction in spatial dimension, from 3D to 2D, implementing a 3D+2D searching strategy initially explained the accelerated rate of the process of target finding. In later years, Berg and Purcell showed that to make this 2D diffusion to be advantageous, the adsorption energy has to be strong enough to keep the phage on the surface of the bacterial cell while weak enough to allow the 2D diffusion [59]. Nonetheless, this model was also the first to predict the adsorption rate k as a function of receptors N on a bacterium and the calculated maximum adsorption rate (k_{max}) based on this model was noticeably less than experimental findings, which leads to hypothesize the possible contribution of bacterial swimming to the search process. Further understanding of the adsorption process comes from a recent work [60], which showed that upon incubation of phage λ with *E. coli* strain Ymel (this strain carries wild type λ receptor) in a solution of 10 mM $MgSO_4$ (pH = 7.4), the decreases of population of free phage in bulk obeys a double-exponential function with a fast and a slow decay time. Both the fast and slow processes are specific to interactions between phage λ and its receptor and henceforth the interaction is an on-and-off process followed by an irreversible binding. It is also noteworthy that the reversible and the irreversible binding rate in nearly independent of temperature, suggesting the entropic nature of phage retention by the receptor. However, this simplified model is based on uniform distribution of receptor sites in the cell surface of the bacteria, while studies using fluorescently labeled λ tail showed that spatial distribution of LamB in the outer membrane is not uniform, and rather LamB accumulated in irregular and spiral patterns that are dynamic and depend on cell length [61]. In another study, fluorescently labeled λ phages adsorbed on individual cells revealed a preferential binding of phages to the bacterial poles rather than cover the cell surface uniformly [62]. Interestingly, this preference for polar sites is not restricted to λ. When similar experiments were carried with other *E. coli* phages such as P1 [63], virulent coliphages T4 [64] and T7 [65] or with other bacteria than *E. coli* K-12, like *Yersinia pseudotuberculosis* (and with T7-like Yersinia phage A1122) [66] and *Vibrio Cholera* (with T4 like vibrio phage KVP40) [67], similar preference to bacterial poles had been documented (See Table 2) [62]. Taken together, these findings indicated that a model based on uniform 2D diffusion search process is lacking.

Table 2. Localization of different phages on cell surface. QDots label phages are compared for their localization on cell surface (Adopted and modified from [62]).

Phage: Host	Foci at the pole and mid-cell (%)	Foci in other location (%)
λ: *E. coli*	69	31
T7: *E. coli*	71	29
P1: *E. coli*	78	22
T4: *E. coli*	95	5
λØ80: *E. coli*	68	32
KVP40: *V. cholera*	73	27
ØA1122: *Y. pseudotuberculosis*	68	32

An accurate depiction of the search process was recently established using single-particle tracking experiments of fluorescently labeled phages. This experimental approach allowed to monitor the early

stages of infection in live bacteria at single-virus:single-cell level with nanometer localization accuracy and ~30-ms time resolution [26]. In this study live *E. coli* cells were first attached to the surface of a microfluidic chamber followed by addition of fluorescent viruses enabling to monitor their adsorption onto cells. The motion of individual phages was categorized in three different modes, free diffusion, motion on the surface of the host cells, and attachment (Figure 3A,B). Similar to the classical pictures [58], free phages initial diffuse in 3D until they encounter a bacterial cell, hence their motion will transition into a 2D diffusion on the cell surface. The phages will continue to diffuse on the surface of the cell until they either irreversibly bind to a receptor, or fall off and continue their free 3D motion. It is noteworthy that the diffusion coefficient value for each mode of motion is strikingly different (Figure 3D). Both free viruses and viruses moving on the cell surface followed normal Fick's laws of diffusion but with an order of magnitude difference in their diffusion coefficient. On the other hand, viruses that were attached to the cells exhibited slow local motion. However, in contrary to the classic view of uniform 2D movement of phages on the cell surface, phages exhibited a distinctly anisotropic motion pattern, with a tendency to move along the short axis of the cell (Figure 3C). Their 2D motion on cell surface was also spatially heterogeneous, showed a spatial focusing along the cell. The bound phages also showed a distinct preference for the poles of the cells (Figure 3F).

These observations were hypothesized to be linked to an ordered pattern of phage receptors, LamB, on the surface of the cell. This hypothesis is also in good agreement with the previous results [61] which showed, using fluorescently labeled phage tails that a spatial arrangement of LamB receptors on the *E. coli* surface was reminiscent of the helices and rings found for other bacterial surface proteins [68,69]. To test that notion, the arrangement of LamB receptors on the cell surface was examined using an *E. coli* strain (S2188:pLO16) with an inducible expression for a modified biotinylated version of the LamB protein [70], enabling to specifically label LamB receptors using SA-conjugated fluorophores. Labeling multiple receptors with a high concentration of QDs (10 nM) resolved the spatial organization of LamB on the cell surface, clearly showing various striped patterns reminiscent of rings and helices (Figure 4A). The distribution of receptors along individual cells showed distinct peaks corresponding to the observed rings and helices with high receptor concentration around the cell pole. The striking resemblance of organization of the LamB receptors with the features exhibited by viruses moving on the cell surface, suggest a unique virus-receptor interaction resulting in an increased viral residence in receptor rich regions. This idea is further supported by: (1) a similar angular distribution for viral trajectories and LamB bands (Figure 4C); (2) an increased viral affinity for polar localization; and (3) distinct LamB bands at poles (Figure 4B). Further proof that the viral motion of phages on the cell surface is governed by interaction with LamB receptors was provided by comparing the dwell times of viruses on wild-type cells and receptor-deleted cells (Figure 5B) [70,71]. Cells lacking receptors exhibited a >15-fold decrease in dwell time, indicating that LamB receptors are required for prolonged interaction between viruses and cell surface. Along this line of evidence, co localization experiments of moving viruses and receptors (Figure 5A,C) showed that the viruses spent a mean of 73.6 ± 3.7% of their total trajectory time in receptor-rich regions.

Figure 3. The target-finding process of individual viruses. (**A**) Time-lapse images of a single fluorescently label λ-phage virus (green spot) diffusing near and on an *E. coli* host cell (black) (images scale: height = 10 μm, width = 11 μm). (**B**) Cartoon of the observed stages for virus receptor-finding process: (I) Virus initially diffuses freely until it encounters a cell, followed by (II) motion on the host cell and (III) binding to a receptor (or detachment from the host cell and continued free diffusion). (**C**) Four representative single-virus trajectories plotted in normalized bacterial coordinates (X_L, long axis: length normalized coordinate; X_T, short axis: width normalized coordinate [26]), showing a tendency for motion along the short axis X_T (perpendicular to the long axis). (**D**) Normalized distribution of the angle between the momentary displacement vector and the short axis, X_T, (138 viral trajectories). The histogram shows a predominant inclination for motion along the short axis (X_T) with a mean angle = 29.6 ± 0.4° (mean ± SE, SD = 24.38 degrees). (**E**) The calculated MSD as a function of lag time for individual viral trajectories in each of the observed regimes, forming distinct groups with more than an order of magnitude separation in MSD values: off-cell diffusion trajectories (green), on-cell diffusion (red), and attachment (blue). Trajectories for both off-cell diffusion and on-cell diffusion yielded a log-log slope of ~1, indicative of normal diffusion. Viruses bound to the host showed either small, confined movement or no movement, with a 0.5–0 slope range. (**F**) Distributions of virus positions along the cell (X_L) throughout the spatial focusing process. Error bars: mean ± SE. (i) Initial random point of encounter (top panel, 47 viral trajectories). (ii) Unbound viruses moving on the surface of the host (middle panel, 138 viral trajectories). (iii) Distribution of final infection sites with a clear trend for polar localization (bottom panel, 59 viral trajectories). The area of all three distributions was normalized to facilitate comparison between the different distributions. Adopted and modified after [26].
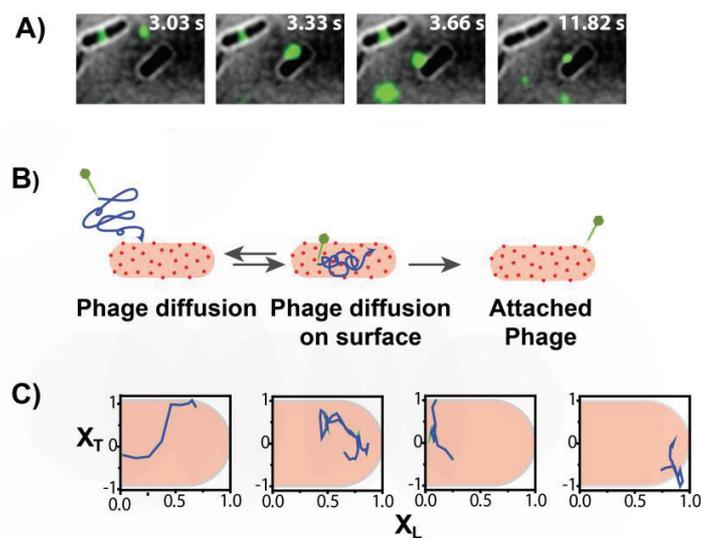
Figure 3. *Cont.*
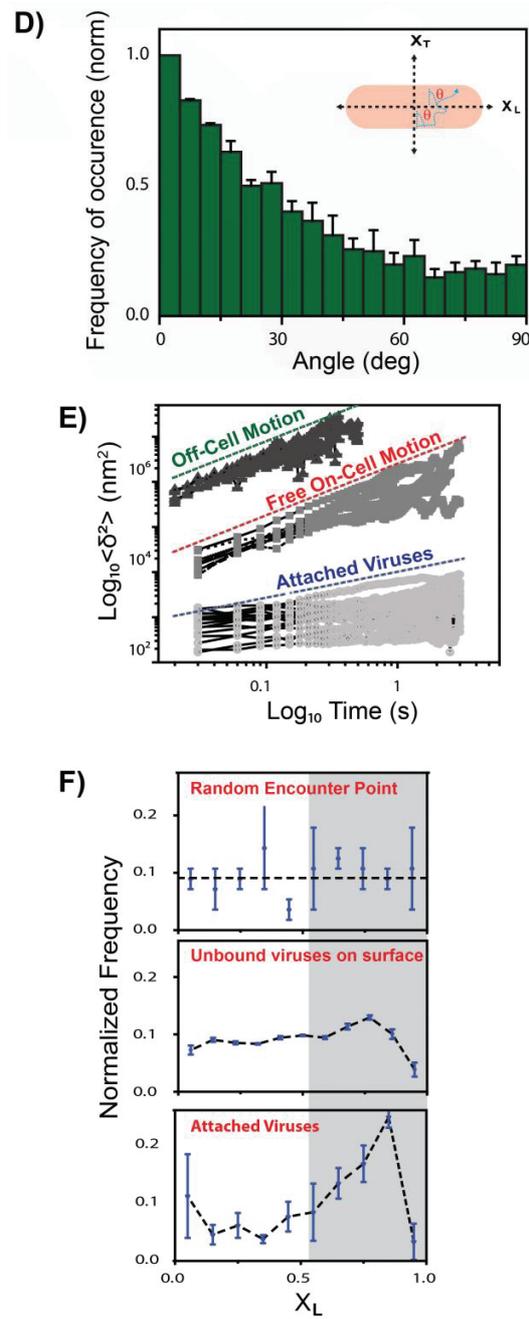
Figure 4. Arrangement of the LamB receptors network on cell surface. (**A**) One cell is shown. Rendered 3D image obtained by sectioning epifluorescence microscopy of cells with labeled receptors, showing nearly continuous helices (scale bar = 2 μm). Quantum dots (green) were used to label the receptor and cell outline is in red. The pattern is described by two helices in blue. (**B**) Normalized LamB distribution along X_L averaged for 50 cells. A clear peak at the cellular poles indicates of a high concentration of receptors. (**C**) Angular distribution of LamB bands from 98 cells, showing a tendency for band orientation along the short axis X_T (perpendicular to the long axis). Adopted and modified after [26].
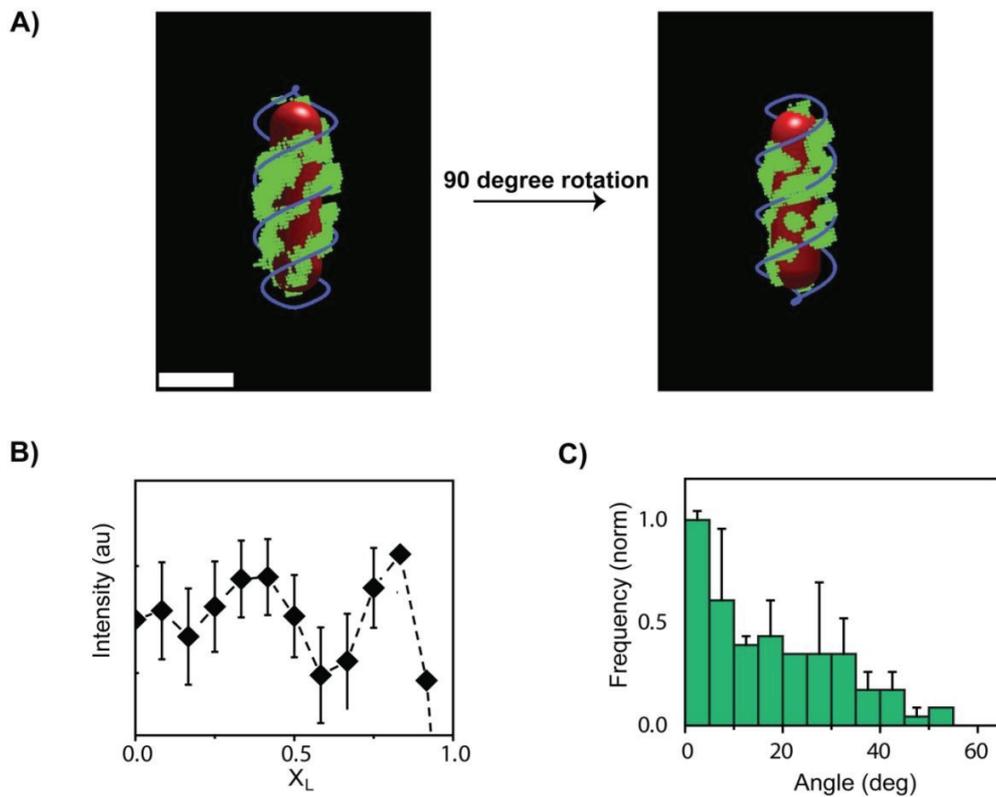
Figure 5. Viral motion on the cell surface is dominated by the interaction with LamB receptors. (**A**) Time-lapse images of the colocalization of a free virus (green) moving on a cell (blue) with LamB receptor bands (purple). The virus is observed to be predominantly moving on the LamB receptor bands (scale bar = 2 μm). (**B**) The mean dwell times of viruses on host cell with (+LamB) and without (−LamB) LamB. (**C**) Colocalization dwell-time analysis of 70 unbound viruses moving on cells with labeled bio LamB, showing the distribution of the fraction of a trajectory the viruses spent on LamB regions, with strong prevalence for residing in highly labeled, receptor-rich regions. Inset: Distribution of the cell area (fraction of entire area) with labeled receptors. Adopted and modified after [26].
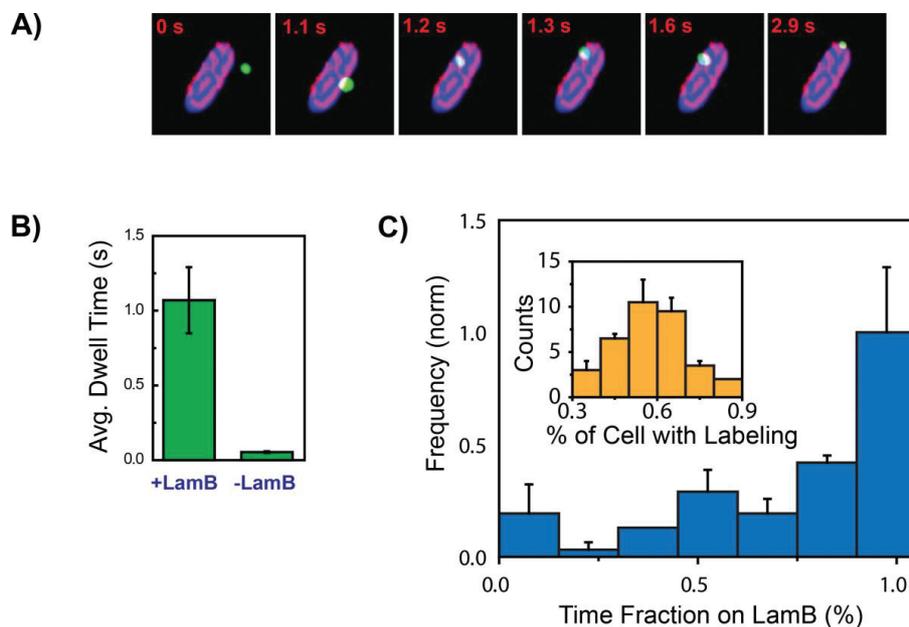
4. Theoretical Model for Viral Target Finding

A quantitative understanding of the dynamics of viral receptor-finding process came from a theoretical model obtained by simulating the motion of 10,000 viruses using an algorithm for curved surface diffusion (Figure 6) [26,72,73]. In this model, *E. coli* cells were considered as cylinders with spherical caps. For each cell, the cell size defined the pitch of a double helix. This double helix was defined as a receptor-rich area, and the rest of the cell surface was defined as receptor-free. A virus located in a receptor-rich zone may diffuse randomly within the receptor-rich area, it may become attached to the surface, or it may move into a receptor-free area, whereas a virus located in the receptor-free area may diffuse within this area, diffuse into the receptor-rich area, or fall off the cell surface. The resulting spatial focusing features generated by the model and the resulting trajectories showed to be in agreement with the experiments. Viruses arriving at random places along the cell followed by gradually concentrate at receptor-rich regions and the final attachment sites show a pronounced preference for the receptor-rich areas, and especially the cell poles.

Figure 6. Theoretical model of phage target finding. (**A**) A schematic description of the model, showing a typical phage trajectory (green) on the surface of an *E. coli* cell. Cell shape was modeled as a cylinder with hemispherical caps (radius R_c = 0.4 µm). The cell surface was divided into receptor-free (transparent pink) and receptor-rich (solid red) areas. The receptors form a 50-nm-thick double helical or multi-ring pattern along the cell surface. The zoomed-in area shows the kinetic scheme in detail. A phage located in a receptor-rich zone can either move within the receptor area with a stepsize consistent with the diffusion coefficient D_{RR}, become attached with a probability P_{Att} (chosen to match the experimentally observed dwell time to attachment), or move into a receptor-free zone. This move can be rejected with a probability $1 - P_{NR \rightarrow R}$, in which case the move will be repeated. On the other hand, a phage located in a receptor-free area can move with a diffusion coefficient D_{NR}. If the phage attempts to enter the receptor zone, the move will always be allowed. The phage can also fall off the cell with a probability $P_{Fall} = \tau sim/\langle\tau_{Fall}\rangle$, where τ_{sim} is the time step of the simulation and $\langle\tau_{Fall}\rangle$, is the experimentally observed dwell time to fall-off in a LamB-*E. coli* strain. (**B**) A simplified kinetic scheme of the model, shown as a four-state, discrete-time Markov chain. The nonreceptor (NR) and receptor (R) states are transient, and the attachment (A) and fall-off (F) states are absorbing. (**C**) Generating population heterogeneity. Cell lengths were randomly chosen from a log-normal distribution spanning the experimentally observed cell sizes. The cell size defines the pitch of receptor double helix according to the relationship obtained experimentally. For short cells with helix angles of <10, a multi-ring pattern was used instead of a double helix. (**D**) Representative 2D projections of phage trajectories in normalized units (green lines). The initial landing site is shown as a gray circle. Panels in the first two rows display trajectories ending in attachment (blue circle), and panels in the last two rows display trajectories ending with the phage falling off the cell (at a position denoted with a yellow circle). Adopted and modified after [26].
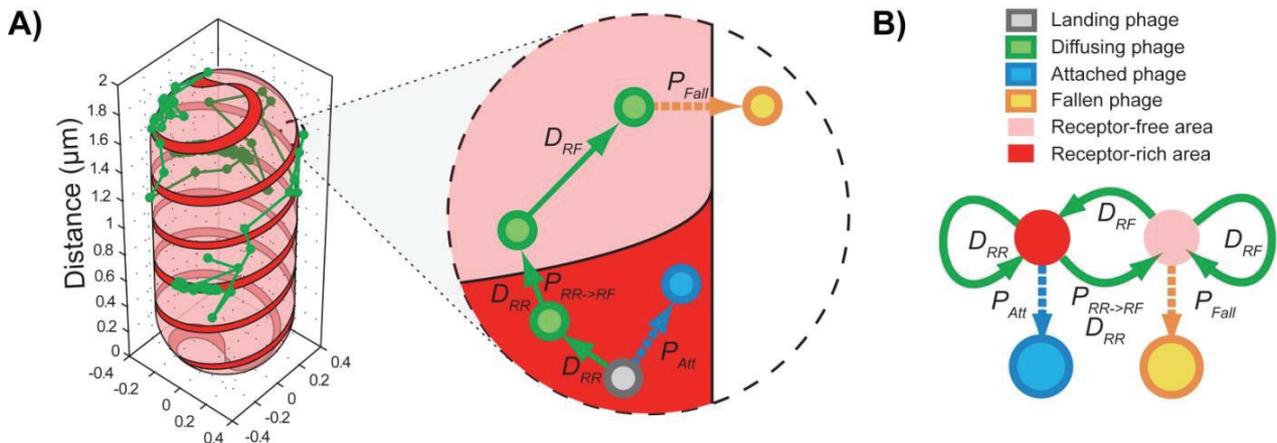
Figure 6. *Cont.*
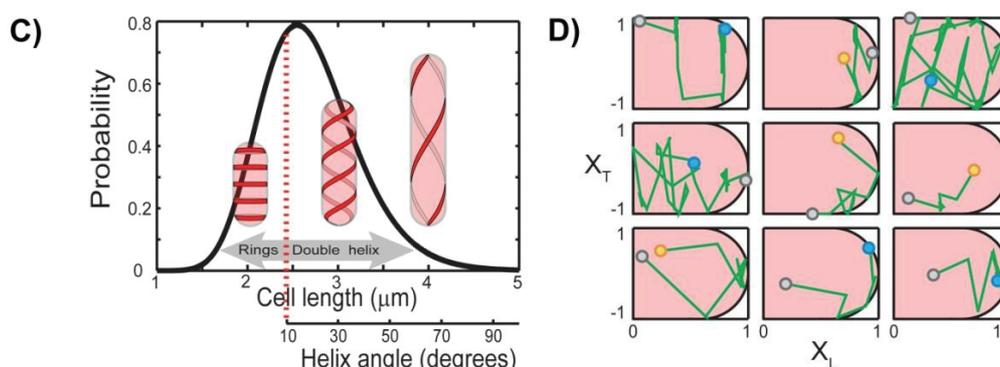
5. Summary

The mechanisms by which molecules, viruses and cells find their targets in biological systems are crucial for a fundamental understanding all biological processes. The interaction of bacteriophage λ with its cell surface receptor, LamB, and its receptor finding process serves an excellent model for this purpose. A plethora of theoretical and experimental studies had been done to address this problem [14–16,18–26]. The classic RREDR model [58], a reduction of dimensionality from 3D motion in bulk to 2D motion on cell surface, was initially considered to be in good agreement with the experimental observation [14,15,58–60,74]. However, finding that the interaction of viruses with a localized network of receptors on the cell surface limits their sampling motion to a fraction of the surface, in effect rendering it quasi-one-dimensional, indicates that the searching process is more complex than what was previously thought and paved the way for further modification of the existing model. We are concluding our present discussion by speculating that the connection between target finding processes and the spatial distribution of the target may be universal, and could be better addressed by considering fractal dimensional motion.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Marsh, M.; Helenius, A. Virus entry: Open sesame. *Cell* **2006**, *124*, 729–740.
2. Mudhakir, D.; Harashima, H. Learning from the viral journey: How to enter cells and how to overcome intracellular barriers to reach the nucleus. *AAPS J.* **2009**, *11*, 65–77.
3. Smith, A.E.; Helenius, A. How viruses enter animal cells. *Science* **2004**, *304*, 237–242.
4. Dimitrov, D.S. Virus entry: Molecular mechanisms and biomedical applications. *Nat. Rev. Microbiol.* **2004**, *2*, 109–122.
5. Mercer, J.; Helenius, A. Virus entry by macropinocytosis. *Nat. Cell Biol.* **2009**, *11*, 510–520.

214

6. Skehel, J.J.; Wiley, D.C. Receptor binding and membrane fusion in virus entry: The influenza hemagglutinin. *Annu. Rev. Biochem.* **2000**, *69*, 531–569.

7. Friedman, D.I.; Court, D.L. Bacteriophage lambda: Alive and well and still doing its thing. *Curr. Opin. Microbiol.* **2001**, *4*, 201–207.

8. Hendrix, R.W. Bacteriophage genomics. *Curr. Opin. Microbiol.* **2003**, *6*, 506–511.

9. Hendrix, R.; Roberts, J.; Stahl, F.; Wesberg, R. *Lambda II*; Cold Spring Harbor Laboratory Press: Plainview, NY, USA, 1983.

10. Ptasne, M. *A Genetic Switch, Gene Control and Phage Lambda*; Blackwell Scientific and Cell Press: Palo Alto, CA, USA, 1986.

11. Ptasne, M. *A Gentic Switch: Phage Lambda Revisited*; Cold Spring Harbor: New York, NY, USA, 2004.

12. Oppenheim, A.B.; Kobiler, O.; Stavans, J.; Court, D.L.; Adhya, S. Switches in bacteriophage lambda development. *Annu. Rev. Genet.* **2005**, *39*, 409–429.

13. Randall-Hazelbauer, L.; Schwartz, M. Isolation of the bacteriophage lambda receptor from *Escherichia coli*. *J. Bacteriol.* **1973**, *116*, 1436–1446.

14. Schwartz, M. Reversible interaction between coliphage lambda and its receptor protein. *J. Mol. Biol.* **1975**, *99*, 185–201.

15. Schwartz, M. The adsorption of coliphage lambda to its host: Effect of variations in the surface density of receptor and in phage-receptor affinity. *J. Mol. Biol.* **1976**, *103*, 521–536.

16. Schwartz, M. *Interaction of Phages with Their Receptor Proteins*; Chapman and Hall: London, UK, 1980.

17. Berkane, E.; Orlik, F.; Stegmeier, J.F.; Charbit, A.; Winterhalter, M.; Benz, R. Interaction of bacteriophage lambda with its cell surface receptor: An *in vitro* study of binding of the viral tail protein gpJ to LamB (Maltoporin). *Biochemistry* **2006**, *45*, 2708–2720.

18. Charbit, A.; Clement, J.M.; Hofnung, M. Further sequence analysis of the phage lambda receptor site. Possible implications for the organization of the lamB protein in *Escherichia coli* K12. *J. Mol. Biol.* **1984**, *175*, 395–401.

19. Charbit, A.; Werts, C.; Michel, V.; Klebba, P.E.; Quillardet, P.; Hofnung, M. A role for residue 151 of LamB in bacteriophage lambda adsorption: Possible steric effect of amino acid substitutions. *J. Bacteriol.* **1994**, *176*, 3204–3209.

20. Clement, J.M.; Lepouce, E.; Marchal, C.; Hofnung, M. Genetic study of a membrane protein: DNA sequence alterations due to 17 lamB point mutations affecting adsorption of phage lambda. *EMBO J.* **1983**, *2*, 77–80.

21. Hofnung, M.; Hatfield, D.; Schwartz, M. malB region in *Escherichia coli* K-12: Characterization of new mutations. *J. Bacteriol.* **1974**, *117*, 40–47.

22. Hofnung, M.; Jezierska, A.; Braun-Breton, C. lamB mutations in E. coli K12: Growth of lambda host range mutants and effect of nonsense suppressors. *Mol. Gen. Genet.* **1976**, *145*, 207–213.

23. Scandella, D.; Arber, W. An *Escherichia coli* mutant which inhibits the injection of phage lambda DNA. *Virology* **1974**, *58*, 504–513.

24. Wang, J.; Hofnung, M.; Charbit, A. The C-terminal portion of the tail fiber protein of bacteriophage lambda is responsible for binding to LamB, its receptor at the surface of *Escherichia coli* K-12. *J. Bacteriol.* **2000**, *182*, 508–512.

25. Wang, J.; Michel, V.; Hofnung, M.; Charbit, A. Cloning of the J gene of bacteriophage lambda, expression and solubilization of the J protein: First *in vitro* studies on the interactions between J and LamB, its cell surface receptor. *Res. Microbiol.* **1998**, *149*, 611–624.

26. Rothenberg, E.; Sepulveda, L.A.; Skinner, S.O.; Zeng, L.; Selvin, P.R.; Golding, I. Single-virus tracking reveals a spatial receptor-dependent search mechanism. *Biophys. J.* **2011**, *100*, 2875–2882.

27. Golding, I. Decision making in living cells: Lessons from a simple system. *Annu. Rev. Biophys.* **2011**, *40*, 63–80.

28. Sanger, F.; Coulson, A.R.; Hong, G.F.; Hill, D.F.; Petersen, G.B. Nucleotide sequence of bacteriophage lambda DNA. *J. Mol. Biol.* **1982**, *162*, 729–773.

29. Hendrix, R. *Lambda II*; Cold Spring Harbor Laboratory: New York, NY, USA, 1985.

30. Weigle, J. Assembly of phage lambda *in vitro*. *Proc. Natl. Acad. Sci. USA* **1966**, *55*, 1462–1466.

31. Meyer, J.R.; Dobias, D.T.; Weitz, J.S.; Barrick, J.E.; Quick, R.T.; Lenski, R.E. Repeatability and contingency in the evolution of a key innovation in phage lambda. *Science* **2012**, *335*, 428–432.

32. Hendrix, R.W.; Duda, R.L. Bacteriophage lambda PaPa: Not the mother of all lambda phages. *Science* **1992**, *258*, 1145–1148.

33. Roa, M.; Scandella, D. Multiple steps during the interaction between coliphage lambda and its receptor protein *in vitro*. *Virology* **1976**, *72*, 182–194.

34. Weigle, J. Studies on head-tail union in bacteriophage lambda. *J. Mol. Biol.* **1968**, *33*, 483–489.

35. Smit, J.; Kamio, Y.; Nikaido, H. Outer membrane of Salmonella typhimurium: Chemical analysis and freeze-fracture studies with lipopolysaccharide mutants. *J. Bacteriol.* **1975**, *124*, 942–958.

36. Kamio, Y.; Nikaido, H. Outer membrane of Salmonella typhimurium: Accessibility of phospholipid head groups to phospholipase c and cyanogen bromide activated dextran in the external medium. *Biochemistry* **1976**, *15*, 2561–2570.

37. Nikaido, H. Porins and specific channels of bacterial outer membranes. *Mol. Microbiol.* **1992**, *6*, 435–442.

38. Nikaido, H. Molecular basis of bacterial outer membrane permeability revisited. *Microbiol. Mol. Biol. Rev.* **2003**, *67*, 593–656.

39. Nikaido, H.; Vaara, M. Molecular basis of bacterial outer membrane permeability. *Microbiol. Rev.* **1985**, *49*, 1–32.

40. Zhai, Y.; Saier, M.H., Jr. The beta-barrel finder (BBF) program, allowing identification of outer membrane beta-barrel proteins encoded within prokaryotic genomes. *Protein Sci.* **2002**, *11*, 2196–2207.

41. Engelhardt, H.; Heinz, C.; Niederweis, M. A tetrameric porin limits the cell wall permeability of *Mycobacterium smegmatis*. *J. Biol. Chem.* **2002**, *277*, 37567–37572.

42. Heinz, C.; Engelhardt, H.; Niederweis, M. The core of the tetrameric mycobacterial porin MspA is an extremely stable beta-sheet domain. *J. Biol. Chem.* **2003**, *278*, 8678–8685.

216

43. Model, K.; Prinz, T.; Ruiz, T.; Radermacher, M.; Krimmer, T.; Kuhlbrandt, W.; Pfanner, N.; Meisinger, C. Protein translocase of the outer mitochondrial membrane: Role of import receptors in the structural organization of the TOM complex. *J. Mol. Biol.* **2002**, *316*, 657–666.

44. Schleiff, E.; Eichacker, L.A.; Eckart, K.; Becker, T.; Mirus, O.; Stahl, T.; Soll, J. Prediction of the plant beta-barrel proteome: A case study of the chloroplast outer envelope. *Protein Sci.* **2003**, *12*, 748–759.

45. Parker, M.W.; Buckley, J.T.; Postma, J.P.; Tucker, A.D.; Leonard, K.; Pattus, F.; Tsernoglou, D. Structure of the Aeromonas toxin proaerolysin in its water-soluble and membrane-channel states. *Nature* **1994**, *367*, 292–295.

46. Petosa, C.; Collier, R.J.; Klimpel, K.R.; Leppla, S.H.; Liddington, R.C. Crystal structure of the anthrax toxin protective antigen. *Nature* **1997**, *385*, 833–838.

47. Song, L.; Hobaugh, M.R.; Shustak, C.; Cheley, S.; Bayley, H.; Gouaux, J.E. Structure of staphylococcal alpha-hemolysin, a heptameric transmembrane pore. *Science* **1996**, *274*, 1859–1866.

48. Hancock, R.E.; Reeves, P. Lipopolysaccharide-deficient, bacteriophage-resistant mutants of *Escherichia coli* K-12. *J. Bacteriol.* **1976**, *127*, 98–108.

49. Roa, M. Interaction of bacteriophage K10 with its receptor, the lamB protein of *Escherichia coli*. *J. Bacteriol.* **1979**, *140*, 680–686.

50. Wandersman, C.; Schwartz, M. Protein Ia and the lamB protein can replace each other in the constitution of an active receptor for the same coliphage. *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 5636–5639.

51. Ishii, J.N.; Okajima, Y.; Nakae, T. Characterization of lamB protein from the outer membrane of *Escherichia coli* that forms diffusion pores selective for maltose-maltodextrins. *FEBS Lett.* **1981**, *134*, 217–220.

52. Nakae, T.; Ishii, J.N. Molecular weights and subunit structure of LamB proteins. *Annales de Microbiologie* **1982**, *133A*, 21–25.

53. Neuhaus, J.M. The receptor protein of phage lambda: Purification, characterization and preliminary electrical studies in planar lipid bilayers. *Annales de Microbiologie* **1982**, *133A*, 27–32.

54. Schirmer, T.; Keller, T.A.; Wang, Y.F.; Rosenbusch, J.P. Structural basis for sugar translocation through maltoporin channels at 3.1 A resolution. *Science* **1995**, *267*, 512–514.

55. Dutzler, R.; Wang, Y.F.; Rizkallah, P.; Rosenbusch, J.P.; Schirmer, T. Crystal structures of various maltooligosaccharides bound to maltoporin reveal a specific sugar translocation pathway. *Structure* **1996**, *4*, 127–134.

56. Charbit, A.; Gehring, K.; Nikaido, H.; Ferenci, T.; Hofnung, M. Maltose transport and starch binding in phage-resistant point mutants of maltoporin. Functional and topological implications. *J. Mol. Biol.* **1988**, *201*, 487–496.

57. Fuerst, C.R.; Bingham, H. Genetic and physiological characterization of the J gene of bacteriophage lambda. *Virology* **1978**, *87*, 437–458.

58. Adam, G.; Delbruck, M. Reduction of dimensionality in biological diffusion processes. In *Structural Chemistry and Molecular Biology*; W.H. Freeman & Company: San Francisco, NC, USA, 1968.

59. Berg, H.C.; Purcell, E.M. Physics of chemoreception. *Biophys. J.* **1977**, *20*, 193–219.

60. Moldovan, R.; Chapman-McQuiston, E.; Wu, X.L. On kinetics of phage adsorption. *Biophys. J.* **2007**, *93*, 303–315.

61. Gibbs, K.A.; Isaac, D.D.; Xu, J.; Hendrix, R.W.; Silhavy, T.J.; Theriot, J.A. Complex spatial distribution and dynamics of an abundant *Escherichia coli* outer membrane protein, LamB. *Mol. Microbiol.* **2004**, *53*, 1771–1783.

62. Edgar, R.; Rokney, A.; Feeney, M.; Semsey, S.; Kessel, M.; Goldberg, M.B.; Adhya, S.; Oppenheim, A.B. Bacteriophage infection is targeted to cellular poles. *Mol. Microbiol.* **2008**, *68*, 1107–1116.

63. Kaiser, D.; Dworkin, M. Gene transfer to myxobacterium by *Escherichia coli* phage P1. *Science* **1975**, *187*, 653–654.

64. Hashemolhosseini, S.; Holmes, Z.; Mutschler, B.; Henning, U. Alterations of receptor specificities of coliphages of the T2 family. *J. Mol. Biol.* **1994**, *240*, 105–110.

65. Steven, A.C.; Trus, B.L.; Maizel, J.V.; Unser, M.; Parry, D.A.; Wall, J.S.; Hainfeld, J.F.; Studier, F.W. Molecular substructure of a viral receptor-recognition protein. The gp17 tail-fiber of bacteriophage T7. *J. Mol. Biol.* **1988**, *200*, 351–365.

66. Garcia, E.; Elliott, J.M.; Ramanculov, E.; Chain, P.S.; Chu, M.C.; Molineux, I.J. The genome sequence of Yersinia pestis bacteriophage phiA1122 reveals an intimate history with the coliphage T3 and T7 genomes. *J. Bacteriol.* **2003**, *185*, 5248–5262.

67. Miller, E.S.; Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Durkin, A.S.; Ciecko, A.; Feldblyum, T.V.; White, O.; Paulsen, I.T.; Nierman, W.C.; *et al.* Complete genome sequence of the broad-host-range vibriophage KVP40: Comparative genomics of a T4-related bacteriophage. *J. Bacteriol.* **2003**, *185*, 5220–5233.

68. Osborn, M.J.; Rothfield, L. Cell shape determination in *Escherichia coli*. *Curr. Opin. Microbiol.* **2007**, *10*, 606–610.

69. Vats, P.; Shih, Y.L.; Rothfield, L. Assembly of the MreB-associated cytoskeletal ring of Escherichia coli. *Mol. Microbiol.* **2009**, *72*, 170–182.

70. Oddershede, L.; Dreyer, J.K.; Grego, S.; Brown, S.; Berg-Sorensen, K. The motion of a single molecule, the lambda-receptor, in the bacterial outer membrane. *Biophys. J.* **2002**, *83*, 3152–3161.

71. Brown, S. Engineered iron oxide-adhesion mutants of the *Escherichia coli* phage lambda receptor. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 8651–8655.

72. Golding, I.; Cox, E.C. Physical nature of bacterial cytoplasm. *Phys. Rev. Lett.* **2006**, *96*, 098102.

73. Holyst, R.; Plewczynski, D.; Aksimentiev, A.; Burdzy, K. Diffusion on curved, periodic surfaces. *Phys. Rev. E* **1999**, *60*, 302–307.

74. Axelrod, D.; Wang, M.D. Reduction-of-dimensionality kinetics at reaction-limited cell surface receptors. *Biophys. J.* **1994**, *66*, 588–600.

Article

Phage Lambda P Protein: Trans-Activation, Inhibition Phenotypes and their Suppression

Sidney Hayes *, Craig Erker, Monique A. Horbay, Kristen Marciniuk, Wen Wang and Connie Hayes

Department of Microbiology and Immunology, College of Medicine, University of Saskatchewan, Saskatoon, S7N 5E5 Canada; E-Mails: sidney.hayes@usask.ca (S.H.); craigaerker@gmail.com (C.E.); mah134@mail.usask.ca (M.A.H.); kdm449@mail.usask.ca (K.M.); wew153@mail.usask.ca (W.W.); clh127@mail.usask.ca (C.H.)

* Author to whom correspondence should be addressed; sidney.hayes@usask.ca; Tel.: +1-306-966-4307; Fax: +1-306-966-4298.

Abstract: The initiation of bacteriophage λ replication depends upon interactions between the *oriλ* DNA site, phage proteins O and P, and *E. coli* host replication proteins. P exhibits a high affinity for DnaB, the major replicative helicase for unwinding double stranded DNA. The concept of P-lethality relates to the hypothesis that P can sequester DnaB and in turn prevent cellular replication initiation from *oriC*. Alternatively, it was suggested that P-lethality does not involve an interaction between P and DnaB, but is targeted to DnaA. P-lethality is assessed by examining host cells for transformation by ColE1-type plasmids that can express P, and the absence of transformants is attributed to a lethal effect of P expression. The plasmid we employed enabled conditional expression of *P*, where under permissive conditions, cells were efficiently transformed. We observed that ColE1 replication and plasmid establishment upon transformation is extremely sensitive to P, and distinguish this effect from P-lethality directed to cells. We show that alleles of *dnaB* protect the variant cells from *P* expression. P-dependent cellular filamentation arose in *ΔrecA* or *lexA*[Ind⁻] cells, defective for SOS induction. Replication propagation and restart could represent additional targets for P interference of *E. coli* replication, beyond the *oriC*-dependent initiation step.

1. Introduction

The mechanism for bi-directional initiation of bacteriophage λ DNA replication involves a complex interaction of λ proteins O and P with the *E. coli* host DNA replication machinery. The arrangement of genes *O* and *P* on the λ genome is shown in Figure 1. Their expression depends upon *pR-cro-cII-O-P* mRNA synthesis from promoter *pR* which is negatively regulated by the CI repressor, made from gene *cI*, binding to *oR* operator sequences overlapping *pR*. The gp*O* (= O) acts to bind *ori*λ, the λ origin of replication, situated midway within the *O* sequence [1]. The gp*P* = P protein facilitates replication initiation through recruitment of an *E. coli* host protein(s) to form a preprimosomal initiation complex at *ori*λ. In agreement with earlier genetic studies, P was found to physically interact [2] with DnaB the major replicative helicase for unwinding double-stranded (ds) DNA. DnaB promotes the advancement of a growing replication fork [3] by using energy provided by ATP hydrolysis; it functions as a "mobile promoter" in the general priming reaction [4,5] to aid DnaG primase in producing RNA primers for extension by DNA polymerase III; and it is able to promote progression of Holliday Junctions in the repair of DNA damage [6,7]. DnaB is a hexameric homomultimer with up to six bound ATP molecules [8,9]. Its intrinsic ssDNA binding activity is inhibited when the free form of DnaB forms a complex with the host replication initiation protein DnaC ([10] and included references). P can commandeer DnaB away from DnaC [11]. P:DnaB participates in loading DnaB onto DNA during formation of the DnaB:P:O:*ori*λ preprimosomal complex. The interaction between P and DnaB inactivates the helicase activity of DnaB. Restoration of DnaB activity, coupled with activation of *ori*λ-dependent replication initiation involves the dissociation of P bound to DnaB by *E. coli* heat shock proteins DnaK, DnaJ, and GrpE [12,13]. The ability of P to dissociate preformed cellular DnaB-DnaC complexes, coupled with a low cellular concentration of DnaB (~20 hexamers/cell) [14,15], suggests the hypothesis that λ *P* expression can inhibit *E. coli* DNA replication initiation from *oriC*, alluded to herein as P-lethality.

Figure 1. Relevant λ genes in expression plasmids and prophage. (**A**) Arrangement of synthetic plasmid pcIpR-[GOI]-timm. Plasmids employed include the precise coding sequence for GOI (gene of interest) genes *O, P,* or alleles of *P*, including P^π, *P*-SPA, $P^{\Delta76}$, each inserted, in precisely the same orientation and position as gene *cro* in λ, directly downstream of *pR*. The coding sequence for the GOI is terminated by an ochre stop codon inserted just ahead of the powerful *timm* termination sequence [16,17], previously named *ti* [18,19]. The regulatory regions of pcIpR-(GOI)-timm were described (Figure 1 plus supplemental sequence file in [20]). The expression of the GOI from promoter *pR* is negatively regulated by the lambda CI[Ts] repressor binding to the wild type *oR* operator sites overlapping *pR*. These plasmids do not contain the *oL* operator sequences so that the tightest repression of transcription from the *pR* promoter, requiring CI-mediated DNA looping via CI dimers binding both operator sites [21], is not possible. Transcriptional read through beyond the GOI is prevented by *timm*, shown in its natural map position between the left operators, *oL* and the C-terminal end of *rexB* (see gene map drawn in Figure 1B). The synthetic sequence from the translational termination sequence for GOI through *timm* to downstream *EcoR*I-*Sal*I is: *TAATCGAT*cccggGGtcagcCccgggttttcttt*TGAATTCGTCGAC*, where the bases that were modified from the wild type lambda DNA sequence are in capitalized italics. Shifting cells with a pcIpR-[GOI]-timm plasmid that were grown at 25 or 30 °C, to above 39 °C induces expression of the GOI from the plasmid [20]. (**B**) Cryptic prophage in strain Y836. The λ phage genes *cro-cII-O-P-ren* are transcribed from promoter *pR* that is embedded within the rightward operator sequence, *oR*, between genes *cI* and *cro*. Gene *cI*, encoding a temperature sensitive repressor is transcribed from promoter *pM*; and *cro* is transcribed in the opposite direction from *pR*. C. At temperatures where CI remains active, *i.e.*, at or below 38 °C, λ replication initiation is prevented, and the λ fragment is replicated as part of the *E. coli* chromosome by forks arising from *oriC*. At about 39 °C, CI becomes fully denatured, *pR* transcription is induced, and λ undergoes a few replication initiation events from *ori*λ [22].

A pcIpR-[gene of interest]-timm plasmids

Figure 1. *Cont.*

B Cryptic lambda (λ) prophage fragment in *E. coli* strain designated Y836

C Thermal induction of replication from cryptic λ prophage in strain Y836
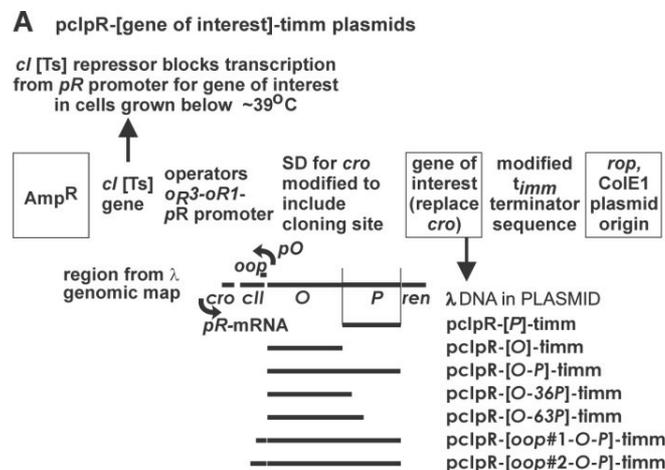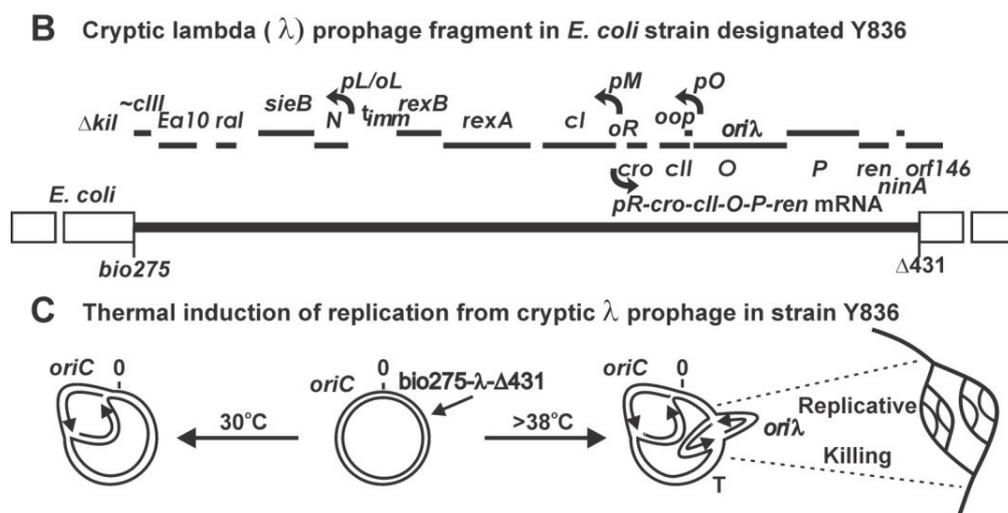
Klinkert and Klein [23] examined the effect of cloning the λ fragment sequence bp's 39168–44972 (part of *O*, *P-ren*-and downstream *orf's 290-146-57-56-60-204-68-64*) into a plasmid expression vector regulated by the *lac* promoter. They showed that inducing *P* expression correlated with effects of shifting cells with temperature sensitive slow-stop mutations in replication initiation genes to the non-permissive temperature, *i.e.*, an eventual inhibition in *E. coli* replication was seen after about a 50 min induction. The effect paralleled the addition of chloramphenicol, and was opposite the effect of adding nalidixic acid, which produced an immediate (fast-stop) inhibition in chromosomal replication (their Figure 5). Maiti *et al.* [24] cloned λ genes into pBR322 and followed the survival of bacteria after transformation with a plasmid that constitutively expressed *pR-cro-O-P-ren-tR2*, measuring transformants (Amp^R colony forming units, cfu) per µg DNA. They observed that plasmids expressing active *P* did not yield survivor cfu, concluding that *P* expression is lethal to the host, *i.e.*, P-lethality. They found that three types of *groP* mutants of *E.coli*, which do not allow λ DNA replication, possibly due to the lack of interaction of P protein with the altered DnaB (*gro*PA15, *gro*PB558) or DnaJ protein (*gro*PC259), were equally susceptible to killing by P protein. They concluded that P-lethality *does not involve interaction of* P with the host DnaB, DnaJ, or DnaK proteins, which are all essential for λ DNA replication [25-27] but is targeted to the *E. coli* DNA replication initiation protein DnaA, inhibiting its *oriC* DNA binding [28]. Datta *et al.* [29] identified mutations within the gene *dnaA* that confer cellular resistance to P-lethality. An alternative hypothesis for the absence of transformants formed by plasmids expressing *P* was not considered in prior studies, *i.e.*, that *P* expression from the transformed pBR322-derived plasmids employed interfered with / blocked the initiation of plasmid replication and copy increase within the newly transformed cell. Since the replication origin of pBR322 was derived from pMB1, closely related to ColE1 [30] we refer herein to the plasmids employed as ColE1-type plasmids.

In the present manuscript we examine if the interaction of P and DnaB provides an alternative mechanism for P-lethality. We assess the influence of P expression upon the maintenance of ColE1

plasmids, and we examine whether P expression, shown to stimulate cellular filamentation, depends upon triggering a cellular SOS response.

2. Result and Discussions

2.1. Transformation of Cells by Plasmids Expressing P, Complementation and Immunity Assays

Plasmid pHB30 (Table 10), encoding an intact lambda *P-ren* sequence whose expression is regulated by the CI[Ts] repressor, yielded similar transformation results as those reported by Maiti *et al.* [24], *i.e.* no transformants were observed at 42 °C for *dnaB*[+] hosts, or for several host strains isolated as being defective for λ replication, *i.e.*, *groP*A15, *groP*558, *grpE*280, *grpC*2 [31]. However two mutants employed in this study, *grpA*80 and *grpD*55 [32] supported pHB30 transformation at 42 °C at about the same efficiency as at 30 °C [31]. The *grpD*55 mutation (initial reported *E. coli* linkage-map position ~71.5 min) was re-mapped as an allele of *dnaB* and was co-transduceable with malF3089::Tn*10* at ~91.5 min [33]. (We were unable to move the grpA80 mutation into another host.) These results were considered preliminary because pHB30 included several genes (a *cro-O* gene fusion, *P* and *ren*), and because the *grpA80* and *grpD55* mutations were not then confirmed to fall within *dnaB*. Table 1 shows that both *grpA*80 and *grpD*55 mutations each possess two missense changes within *dnaB*, one being in common, E426K.

Table 1. Sequencing analysis of alleles of *dnaB*.

dnaB allele	bp *E. coli* mutation	AA changed
grpD55	4,263,102 G to A	V256I
	4,263,612 G to A	E426K
grpA80	4,263,349 G to A	G338E
	4,263,612 G to A	E426K

The ability of the *dnaB-grpD55* allele within *E. coli* strain 594 to inhibit λ vegetative growth is shown in Table 2.

Table 2. Inhibition of λ*cI*72 vegetative growth on *dnaB-grpD55* host.

Host strains	EOP on host cells incubated at temperature [a]			
	30	37	39	42 °C
594	1.0	nd [b]	nd	0.94
594 *grpD55*	0.0147	4.4×10^{-7}	4.4×10^{-7}	4.6×10^{-7}

[a] Efficiency of plating (EOP) of λ*cI*72 was determined by dividing the average phage titer on 594 *dnaB-grpD55* host cells by the titer on 594 host at 30 °C (= 6.8×10^9 pfu/mL). [b] nd = not determined.

Plasmid pcIpR-*P*-timm (Figure 1A) includes λ gene *cI*[Ts]857, encoding a thermolabile repressor that is expressed from promoter *pM*, and a precise copy of wild type gene *P* DNA, situated downstream from λ promoter *pR* in exactly the same configuration as is occupied by gene *cro* in the λ genome. The expression of *P* from this plasmid is regulated by the CI[Ts] repressor binding to the *oR*

sequences overlapping *pR*. The expression of *P* from pcIpR-*P*-timm was initially assumed similar to that shown for D-CAP or D-SPA gene fusions, as in pcIpR-D-CAP-timm (Figure 4 in [20]), where there was no detected expression from the non-induced gene fusion. The expression of genes *P* and *cI*[Ts]857 from pcIpR-*P*-timm was assessed by complementation, measuring the biological activity of P, and a CI-immunity assay to assess CI[Ts] repression of *imm*λ phage plating. The complementation assay determined whether P expressed from the plasmid could support the plating of a phage with an amber mutation in *P*, *i.e.*, the heteroimmune infecting phage λ*imm*434*P*am3, which is insensitive to repression by CI$^\lambda$. The ability of cells with pcIpR-*P*-timm to produce CI[Ts] at a sufficient level to turn off *pR* transcription from both plasmid or an infecting phage was assessed by its inhibition of λ*cI*72 plating.

Table 3 shows the transformation of nonlysogenic 594 *dnaB*$^+$ cells (*dnaB* was sequenced [34]) and of 594(λnin5) lysogenic cells. The CI repressor made from the λnin5 prophage is not Ts. The lysogenic cells expressing CI$^+$ were transformable by pcIpR-*P*-timm at 37 °C. The nonlysogenic cells were transformed by pcIpR-*P*-timm only between 25-34 °C, where CI[Ts] retains sufficient repressor activity to shut off *P* expression in transformed cells. CI[Ts] is reported [35] marginally active at ≥37 °C and would not fully repress *P* expression, explaining the absence of transformants arising in nonlysogenic cells when the transformation temperature is raised above a temperature where CI[Ts] repressor retains sufficient DNA binding activity.

Table 3. Cellular transformation by pcIpR-*P*-timm plasmid.

Cells Transformed	Plating Temp. °C	Transformation Frequency [a]	Transformants per µg DNA
594	25	5.2×10^{-5}	1.5×10^5
	37	$<1.7 \times 10^{-8}$	0
594(λ wt nin5)	25	7.2×10^{-5}	4.7×10^4
	37	8.5×10^{-6}	5.5×10^3

[a] Transformation frequency represents the average cell titer on TB+Amp$_{50}$ plates per average cell titer on TB plates.

The 594[pcIpR-*P*-timm] cells complement for P, at 30 °C, and very marginally (in comparison to the Δ*P* plasmid) complement at 25 °C (Table 4), suggesting a low level of leaky *P* expression, even when CI[Ts] is active.

Table 4. Complementation of λ *P*am mutation by plasmid cloned *P* gene.

Plating host	Incubation temperature	EOP [a] of λ*imm*434*P*am3
594[pcIpR-*P*-timm]	25 °C	0.0008
	30	0.82
	37	0.93
594[pcIpR-*P*$^{\Delta76}$-timm]	30	<0.00003
	37	2×10^{-7}

[a] EOP of 1.0 determined on host TC600 *supE* at 30 °C.

The noninduced P-SPA fusion complements for P in cells grown at 30 and 37 °C (data not shown). 594[pcIpR-$P^{\Delta76}$-timm] transformants were unable to complement for P, suggesting that the region of P removed by the in-frame deletion of 76 codons, localized to the N-teminal end of the 233 amino acid P protein (fusing part of codon 9 with 86) is essential for the role of P in λ replication. The immune state, *i.e.*, *imm*λ[Ts] phenotype, of the pcIpR-[ ]-timm plasmids was examined (Table 5, and [36]) by plating λ*cI*72 on hosts with plasmids encoding intact P or O genes.

Table 5. EOP [a] of λ*cI*72 on host strains with *imm*λ[TS] or hybrid *imm* plasmids [b].

Host strains	Plating temperature (°C)					
	25	30	35	37	39	42
594	0.5	0.9	0.9	1.0	1.0	0.4
594[pcIpR-*P*-timm]	$<4 \times 10^{-10}$	$<4 \times 10^{-10}$	0.2	0.5	0.7	1.0
594[pcIpR-*O*-timm]	$<4 \times 10^{-10}$	$<4 \times 10^{-10}$	$<4 \times 10^{-10}$	9×10^{-7} [c]	2×10^{-6} [c]	0.8
594[p434'pR-*O*-timm]	0.8	0.9	1.0	1.0	1.0	1.0

[a] EOP (efficiency of plating) of 1.0 determined for λ*imm*λ*cI*72 on host 594 at 37 °C. [b] The CI[Ts] repressor made from the pcIpR plasmids blocks the vegetative growth of an infecting *imm*λ phage as λ*cI*72, at or below ~39 °C, as well as the expression of genes O or P, inserted downstream of the *pR* promoter. Plasmid p434'pR-*O*-timm has a hybrid *cI* gene fusion made from λ and 434 *cI* genes and the resulting repressor does not repress transcription from *pR*, allowing constitutive O expression at all temperatures [36]. [c] Rare min plaques can only be seen under stereo microscope.

The plasmid control p434'pR-*O*-timm with a hybrid λ/434 CI repressor does not block *pR* transcription, expresses O constitutively at all temperatures, and showed no immunity to an infecting *imm*λ phage. Plasmid pcIpR-*O*-timm confers high *imm*λ immunity in cells grown between 25 and 39 °C, and loses it at 42 °C where the CI[Ts]857 repressor is fully inactive. Additional immune characterization of these plasmids was reported (Table 4 in [36]). Although pcIpR-*O*-timm only carries *oR* operators for repressor binding, CI immune repression is maintained through 39 °C, but lost at 42 °C where the CI repressor is denatured. Thus, the pcIpR-(GOI)-timm plasmids, even though they carry only *oR* site, can maintain CI immune repression, and presumably block the expression of GOI positioned downstream of *pR*, between 25-39 °C. However, the analogous plasmid pcIpR-*P*-timm exhibits *imm*λ immunity at 25-30 °C but the immune response is ineffective at and above 35 °C. Phage λ*cI*72 forms tiny to small plaques on 594[pcIpR-*P*-timm] cells at 35 and 36 °C where the CI[Ts]857 repressor is somewhat active [35], and larger plaques at 42 °C where the repressor is inactive (Figure 2).

Figure 2. Plaque type / formation by λ*cI*72 on 594[pcIpR-*P*-timm] and 594 host cells. The ability of λ*cI*72 to form plaques on 594[pcIpR-*P*-timm] host cells at 35-42 °C is attributed to loss, or considerable reduction, in plasmid copy number per cell. The *imm*λ interference phenotype (compare EOP, Table 5, 25 and 30 °C *vs.* 35 °C) is dependent upon CI[Ts] repressor expressed from the plasmid. The cellular loss of *imm*λ interference correlates with plasmid loss and the observed increase in plaque size between 35 and 42 °C; whereas, the plaque size was essentially constant between 30 to 42 °C when λ*cI*72 was plated on 594 cells without the plasmid. The photos of the individual plaques shown in the top row were taken through the lens of a stereo microscope, from agar overlay plates (middle row) incubated at the indicated temperatures.
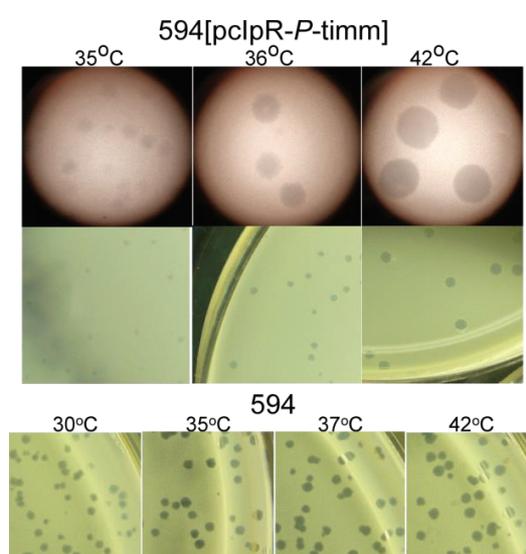
2.1.1. "Trans-activation" of P

The gene *P* function of an *imm*λ prophage was previously shown to be *trans*-activatable [37,38], *i.e.*, cells with a repressed λ prophage could fully complement and support the vegetative growth of an infecting *imm*434 phage defective for *P* (by a factor of 10^3), even though the expression of the prophage copy of *P* was repressed by the *imm*λ CI repressor blocking *pR* transcription in prophage with both *oL* and *oR* sites. Thomas [38] argued that "a small but significant amount of P product was synthesized by the prophage", which he rationalized as being explained by the existence of a minor constitutive promoter mapping between *pR* and *P* on the prophage genome. We saw *trans*-activation of *P* from repressed pcIpR-*P*-timm (slightly at 25 °C and nearly full complementation at 30 °C), which is explained by leaky transcription from *pR*, since there is no room for an unaccounted promoter upstream of *P* as only 18 bases, including the ribosomal binding site, separates the start site for *pR* mRNA transcription and the AUG for *P*. We conclude that trace, certainly not full, *P* expression is sufficient for *ori*λ dependent replication initiation.

2.2. Influence of P-Expression on Plasmid Retention.

The *in*ability of plasmids expressing *P* to form stable transformants is used as a measure of "P-lethality", based on an assumption that P made from the transforming plasmid kills the transformed cells. The influence of *P* expression from pcIpR-*P*-timm in 594 *dnaB*⁺ cells was examined for plasmid replication / maintenance / stability (Figure 3). Plasmid loss was observed in cultures of 594[pcIpR-*P*-timm] grown at 34 or 35 °C (in comparison to cultures grown at 30 °C), but curing was extensive in cell cultures grown at or above 36 °C (Figure 3B). Essentially identical results were seen for 594 *clpP*[pcIpR-*P*-timm] cultures (Figure 3B) showing that inactivation of the *clpP* protease, which participates along with ClpX in degrading λ O protein [39,40], did not significantly influence P-dependent plasmid loss relative to the *clpP*⁺ strain 594; the ClpP protease defect did not amplify plasmid loss, which is expected if its absence increased the P concentration per cell. Secondary cultures prepared from 594 *grpD55*[pcIpR-*P*-timm], or 594[*pcIpR-P*$^{\Delta76}$-timm] cells retained their plasmids, even when grown at 42 °C for at least seven doublings while constitutively expressing *P* (Figure 3C). The *grpD55* allele blocks λ vegetative growth (Table 2) and permits cell survival (at some level) in the presence of constitutive *P* expression at 42 °C, suggesting that allelic variations in DnaB *can* protect cells from *P* expression and its consequent negative influence on cell metabolism. This result is in disagreement with the hypothesis made by Maiti *et al.* [24], which in essence states that allelic variants of *dnaB* will not protect cells from P-lethality.

Figure 3. *P*-induced plasmid loss. (**A**) Cultures of 594[pcIpR-*P*-timm], 594 *clpP*::kan[pcIpR-*P*-timm], 594 *grpD55*[pcIpR-*P*-timm], and 594[pcIpR-*P*Δ76-timm] were grown to stationary phase in TB plus 50 μg/ml ampicillin for 48 hr at 25 °C. Cell aliquots were diluted into fresh TB medium (no ampicillin) as shown in the outline (A) and incubated for about 20 hr in shaking bath between 30 to 42 °C (refer to Experimental Section *3.8*). (**B**,**C**) Plasmid retention by culture cells (described in (**A**)) grown between 30 to 42 °C.

A.

Inoculate broth to 0.008-0.01 A575nm from 25°C culture → Grow at indicated temp. to saturation → Dilute, spread for single colonies, incubate 25°C

→ Extract plasmid DNA

→ Assay colonies, 25°C (AmpR cfu / total cfu)

B. pcIpR-*P*-timm retention in culture cells

Culture growth temp. (°C)	% cells retaining plasmid (AmpR cfu / total cfu)	
	594[pcIpR-*P*-timm]	594 *clpP*[pcIpR-*P*-timm]
30	100 (375/375)	100 (126/126)
34	71.4 (30/42)	77.5 (45/48)
35	77.9 (282/362)	51.4 (37/72)
36	0 (0/50)	0 (0/58)
37	0.3 (1/362)	0 (0/72)
42	0 (0/56)	nd

Figure 3. *Cont.*

C. **Plasmid retention in culture cells**

2.3. P-Lethality Suppression

The results in Figure 3 do not account for P-lethality, only plasmid loss, and so their interconnection was examined (Table 6). Whereas cells within a starting culture that become cured of a plasmid expressing P can grow and eventually fully populate a culture, isolating individual cells by spreading dilutions on an agar plate will permit determination of cell viability at the incubation temperature. Less than 1% viability was observed for spread 594[pcIpR-*P*-timm] cells incubated at or above 37 °C, and any surviving cfu had lost the plasmid. *P* expression was not lethal in spread 594 *grpD55*[pcIpR-*P*-timm] cells incubated between 37 and 42 °C, where the cfu arising at or above 37 °C retained the plasmid. We examined whether the plasmids extracted from these survivor cfu retained P-lethality and plasmid-loss properties, when moved back into 594 *dnaB*$^+$ cells. The reclaimed plasmids retained both the P-lethality and plasmid loss phenotypes (bottom lines, Table 6).

Table 6. Influence of modifying *P* or *dnaB* on P lethality and plasmid loss.

Host cells and plasmids	Cell viability [a] and (plasmid retention/cfu; %) [b] at growth temperature (°C)		
	25	**Ave 37 & 39**	**42**
594[pcIpR-*P*-timm]	1.0 (195/196; ~100)	0.0049 (0/57; 0)	0.001 (0/28; 0)
594[pcIpR-*P*^{SPA}-timm]	1.0 (14/14; 100)	0.006 (0/150; 0)	0.004 (0/39; 0)
594[pcIpR-*P*$^{\pi}$-timm]	1.0 (103/103; 100)	0.635 (789/793; ~100)	0.003 (0/294; 0)
594[pcIpR-*P*$^{\Delta76}$-timm]	1.0 (166/168; 99)	0.96 (334/336; 99)	0.12 (166/168; 99)
594 *grpD55* [pcIpR-*P*-timm]	1.0 (35/35; 100)	0.99 (87/88; 99)	1.0 (14/14; 100)
594 *grpD55* [pcIpR-*P*^{SPA}-timm]	1.0 (42/42; 100)	0.75 (88/94; 94)	0.57 (18/19; 95)
594 *grpD55* [pcIpR-*P*$^{\pi}$-timm]	1.0 (54/54; 100)	0.08 (434/435; ~100)	0.01 (5/108; 5)
594 *grpD55* [pcIpR-*P*$^{\Delta76}$-timm]	1.0 (nd)	0.45 (nd)	0.001 (nd)
Re-claim pcIpR-*P*-timm from 594 *grpD55* cultures and transform into 594 *dnaB* $^{+}$ cells			
Re-claim from 25°C cultures [c]	1.0 (54/56; 96)	0.08 (0/119; 0)	0.01 (0/91; 0)
Re-claim from 42°C cultures [d]	1.0 (130/130; 96)	0.059 (1/230; 0.4)	0.026 (0/201; 0)

[a] The cell viability shown in each column entry, in top line, was determined by dividing the cell titer obtained at each given incubation temperature by the cell titer at 25 °C.) Refer to Experimental Section *3.5*. [b] The values in parentheses in each column entry show the number of Amp^{R} cfu / number of survivor cfu assayed per indicated temperature; the value following represents the percentage of Amp^{R} cfu with plasmids, and was rounded up. All data show the results for two or more independently transformed single colonies. [c] Exp.'s for sc1, sc2. [d] Exp.'s for sc's 3,4,5, 6.

A stark contrast was noted between the *O* and *P* versions of pcIpR-timm plasmids regarding their ability to prevent the plating of *imm*λ phage λ*cI*72. The *O*-version retained the *imm*λ phenotype between 25 and 39 °C, losing it when the CI[Ts] repressor became denatured; the *P*-version lost it between 30 and 35 °C and supported plaque formation on the cell lawns at 35 and 36 °C. The simplest explanation is that the *P*-version plasmid was lost from the cells grown above 30 °C, and coordinately, *cI* repressor gene expression was lost. Gel analysis showed that 100% of cfu from cultures grown at 30 °C retained the *P*-version plasmid, but cultures grown at 34 and 35 °C were reduced in plasmid copies, and cultures grown at or above 36 °C were fully cured. Since cultures grown at 30 °C could complement for *P*, plasmid maintenance was very sensitive to trace *P* expression, which apparently increased with an increase in culture temperature from 30 to 36 °C. In support of this claim, we obtained 1.5×10^{5} transformants per μg of pcIpR-*P*-timm DNA at 25 °C and none at 37 °C. The partial

to full loss between 34–37 °C (Figure 3) of *P*-encoded plasmids, argues for trace levels of P causing ColE1 plasmid curing. This result suggests a new dimension to the concept of P-lethality. It appears that ColE1 plasmid replication is extremely sensitive to P^{λ} protein and that P-inhibition of plasmid replication prevents plasmid establishment upon transformation. This provides a better explanation for the lack of transformants (found by us and other authors) by plasmids capable of expressing P than does cell killing by P (*trans* P lethality).

P^{π} proteins exhibit lower affinity for DnaB, compared to P [41], which may explain why elevated P^{π} expression at 42 °C was needed to evoke the normal P phenotypes in the *grpD55* strain in order for it to compete for and sequester DnaB away from DnaC [11], effectively creating the possibility for a slow-stop type inhibition in *E. coli* replication initiation. A cautionary note here is that the dissociation of DnaB-DnaC complexes by P was only demonstrated *in vitro*.

As well as interacting with λ P [42], DnaB interacts with many cellular proteins, e., g., with DnaA [43], DnaC [44], DnaG [45], SSB [46], the Tau subunit of DNA polymerase III [47], and RNA polymerase [48], and as noted the functional DnaB hexamer is present in limiting amounts, *i.e.*, about 20 hexamers per cell [14,15]. Introducing two missense mutations into *dnaB* will permit the cells to survive constitutive *P* expression (P-lethality), and can suppress P-inhibition of ColE1 replication, yet these mutations do not overtly impede cellular growth/replication. A single missense mutation in *P* can suppress P-lethality and plasmid loss at 37 to 39 °C. Therefore, we suggest that an interaction between P and DnaB is an important component of both P-lethality and P-inhibition of ColE1 replication.

Datta *et al.*, [28,29] could transform *dnaA rpl* mutants (resistance to *P l*ethality) of the same strain we employed, namely 594, with plasmids expressing P, concluding that an interaction between P and DnaA, and not with DnaB, was important for suppressing P-lethality. Their *dnaA*-rpl8 [29] mutant did not inhibit λ plating [49], and thus has a different phenotype from *E. coli* mutants arising in selections where λ development is inhibited, such as the one used by Saito and Uchida [32]. Sequence analysis revealed that our 594 strain is wild type for *dnaA* [34] and therefore, the *grpD55* mutation in *dnaB* appears sufficient for cellular resistance to P lethality; however, we have not compared the two 594 strains by whole genome sequence analysis. We suggest that these possibly contradictory results are explained by the participation of DnaA in mechanistic events shared with DnaB, allowing the *dnaA-rpl* allele(s) to modulate DnaB:P dependent P-lethality and P-inhibition. Alternatively, the reported effect of P protein on DnaA [28] is not at the level of DNA replication, but is in the context of transcription of genes regulated by DnaA.

2.3.1. Model for Plasmid Loss

How can P evoke plasmid loss? The expression and accumulation of lambda P protein, even in seemingly trace levels, interferes with ColE1-type plasmid maintenance replication (Figure 4).

Figure 4. The expression and accumulation of lambda P protein, even in trace levels, interferes with ColE1-type plasmid maintenance replication.
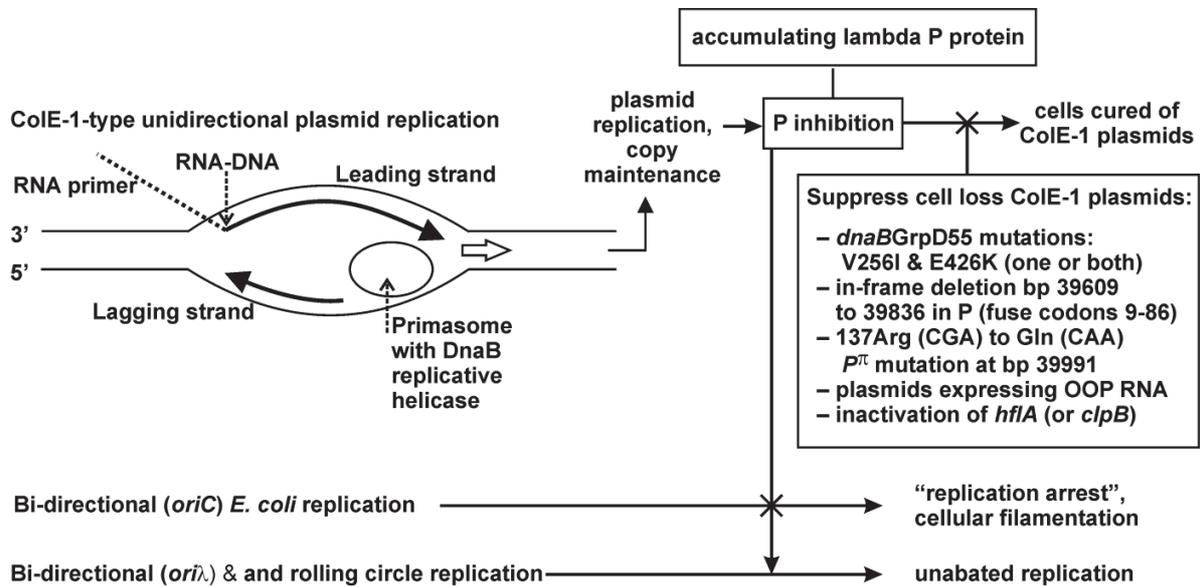
Although ColE1 is considered a theta-type replicating plasmid [50], at an early stage, leading-strand synthesis proceeds in the absence of lagging-strand synthesis, yielding unidirectional replication [51,52], and requires DNA polymerase I (Pol I) to initiate continuous leading-strand replication [53–55] (shown by open arrow, Figure 4). Replication initiation requires primer RNA II for leading strand synthesis [52,56], stable hybridization of RNA II to DNA [57,58], and processing by RNase H to generate the primer for leading strand synthesis [57,59–61]. Pol I extension of the primer unwinds the DNA exposing a primasome assembly signal (n' pas) or single-strand initiation A sequence (*ssiA*) that allows assembly of the primosome through recruitment and activation of the PriA protein [62]. *This represents a distinct form of DNA replication initiation associated with DNA repair*, in contrast to DnaA-dependent replication initiation at *oriC* [63,64]. Following PriA-primosome assembly, DnaB helicase and DnaG primase are loaded and work coordinately to initiate the discontinuous priming of lagging-strand [58,65], which opens a sequence for termination of lagging-strand synthesis, *terH*, effectively establishing unidirectional replication [53,66]. There is a *dnaA* box that is close to n' pas, and potentially serves as a DnaA-dependent DnaB-DnaC assembly site [67,68]. The simplest model for ColE1 plasmid replication inhibition and P-dependent plasmid curing is that formation of a cellular P:DnaB complex limits the availability, or perturbs the activity of DnaB to participate in ColE1 replication initiation, perhaps in a step very similar to replication restart of a stalled replication fork, which can involve DnaA. Replication restart proteins form multiple pathways to restart repaired replication forks (Figure 1 in [69]) and they function equivalently to DnaA in forming protein-DNA complexes so that DnaB helicase is loaded onto the DNA [70,71]. Of note, using *in vitro* ΦX174 DNA synthesis assays to monitor DnaB participation, Wickner [2] observed that P inhibited DnaB protein activity in ΦX replication, making a similar proposal to ours for P inhibition of ColE1 replication.

2.4. Influence of P-alteration, or O and OOP RNA Co-Expression on P-lethality and Plasmid Retention

Linking P with a 66 amino acid SPA tag [72] at its C-terminal end did not suppress P lethality, nor plasmid loss, but the lethality of P^{SPA} strongly decreased and plasmid retention sharply increased in the *grpD55* host at 37–42 °C (Table 6). The 76 codon in frame N-terminal deletion, $P^{\Delta76}$, suppressed both plasmid loss and P-lethality at 37 or 39 °C. Thus, amino acids 9–86 of P include a critical domain. Its removal knocks out P initiation of replication (our complementation data), the P-lethality phenotype, and P-induced ColE1 plasmid loss.

The allele $P^{\pi39991}$ was used to make pcIpR-$P^{\pi39991}$-timm. The mutation was identified in λ*cI*72 mutants that were able to form plaques on the 594 *grpD55* host, and is identical to the πA7 mutation previously sequenced [73]. This R137Q single base change partially suppressed P-lethality and fully suppressed plasmid loss at 37-39 °C (but not at 42 °C where P^π was fully expressed) in the *dnaB*⁺ host. The lethality of P^π significantly increased at 37-39 °C in the *grpD55* host even though plasmid retention remained near 100%. Cells with pcIpR-$P^{\pi39991}$-timm (or pcIpR-P^{SPA}-timm) could complement the amber mutation in λ*imm*434*P*am3.

We asked if co-expression of *O-P*, or if the introduction into the plasmid of the coding sequence for the CII antisense micro-RNA termed OOP, was able to suppress P-lethality and plasmid loss (Table 7).

Table 7. Influence of *O*, *P* and OOP-RNA expression on cell killing and plasmid loss [a].

Plasmids in 594 host cells	Cell viability and (plasmid retention/cfu; %) at growth temperature (°C)				
	30	35	37	39	42
pcIpR-*P*-timm	1.0 (35/35; 100)	0.32 (33/33; 100)	0.01 (0/35; 0)	0.008 (0/35; 0)	0.07 (0/35; 0)
pcIpR-*O-P*-timm	1.0 (62/70; 89)	0.345 (28/70; 40)	0.12 (0/70; 0)	0.12 (0/70; 0)	Nd
pcIpR-*O*-36*P*-timm	1.0 (30/30; 100)	0.825 (30/30; 100)	0.793 (30/30; 100)	0.012 (1/30; 3)	Nd
pcIpR-*O*-63P-timm	1.0 30/30; 100)	0.895 (30/30; 100)	0.895 (30/30; 100)	0.055 (14/40; 35)	Nd
pcIpR-*O*-timm	1.0 (30/30; 100)	1.0 (30/30; 100)	1.0 (30/30;100)	1.0 (30/30; 100)	0.608 (29/30; 97)
p434'pR-*O*-timm [b]	1.0 (26/30; 87)	1.0 (29/30; 97)	1.0 (28/30; 93)	1.0 (26/30; 87)	0.615 (26/30; 87)
pcIpR-*oop*#1-*O-P*-timm [c]	1.0 (120/120; 100)	0.938 (120/120; 100)	0.20 (98/101; 97)	0.005 (115/120; 96)	Nd
pcIpR-*oop*#2 *O-P*-timm [c]	1.0 (117/120; 98)	0.988 (60/60; 100)	0.055 (76/154; 49)	0.048 (62/120; 52)	Nd

[a] As described for Table 6, except stationary phase cultures were grown up at 30 and not 25 °C. [b] Plasmid has constitutive expression of *O*. [c] The viability results represent the average for four independent plasmid isolates, each single experiments. The results in parentheses, for plasmid retention by survivor cfu, sums results for all the cfu's assayed from the four isolates. The cell viabilities at 30 and 35 °C were at or very near unity. pcIpR-*oop*#1-*O-P*-timm: 37°C (0.34, 0.49, 0.006, 0.008), 39°C (0.01, 0.006, 0.0006, 0.0015); pcIpR-*oop*#2-*O-P*-timm: 37°C (0.01, 0.01, 0.12, 0.08), and 39°C (0.01, 0.01, 0.09, 0.08).

The co-expression of *O* would provide an interactive partner for P, possibly lowering the cellular level of unbound P. The 77 nucleotide antisense OOP RNA, which is transcribed from promoter *pO* in the opposite orientation to the *pR-O-P* mRNA (Figure 1A), can hybridize between *pR* and *O*, forming

an OOP-RNA: *pR-O-P*-mRNA hybrid that can provide a target for *P* mRNA degradation downstream from the OOP-binding site. The co-expression of *O-P* increased cell viability at 37–39 °C, but not plasmid retention. The constitutive or inducible expression of *O* without *P* did not influence cell viability or plasmid loss between 30–39 °C. Early studies [74] attempting to determine the coding sequence for *O*, reported a longer protein (O') made by read-through of the normal stop codon, terminating downstream somewhere within the N-terminal sequence of *P*. Plasmids were made that included the wild type λ sequence between the start of *O* through the UGA stop codon [75], plus 35 or 62 in-frame codons downstream into the *P* DNA sequence, so that if the stop codon for *O* was designated "1", the next in-frame stop codons would occur at positions 36 or 63 in frame with *O*, even though the downstream reading frame for *P* differs from *O*. While the plasmids expressing the intact *O* sequence were not toxic, the plasmids with O-36P or O-63P, when expressed at 39 °C, were toxic and caused plasmid loss. It is unclear if their toxicity depends on O-read-through downstream of its stop codon, or to polypeptides made from the N-terminal end of *P*. Plasmids pcIpR-*oop*#1-*O-P*-timm and pcIpR-*oop*#2-*O-P*-timm (Figure 1A, Table 10), when compared to versions expressing only *P* or *O-P* significantly suppressed plasmid loss at 37–39 °C, and partially suppressed P-lethality at 37 °C, with the *oop*#1 plasmid showing stronger suppressing effects than the slightly larger #2 version. This data supports a hypothesis that *P* expression from the λ genome is reduced by OOP antisense micro RNA, possibly via a OOP:*pR*-mRNA complex serving as a target for mRNA degradation.

We observed that the co-expression of *O* and *P*, with the potential for formation of an O:P complex, did not suppress the P phenotypes. In contrast, placing the natural *oop-O-P* sequence orientation downstream from *pR* in the pcIpR-[ ]-timm plasmids markedly suppressed P-inhibition of ColE1 replication as monitored by assaying for plasmid curing. This is explained, possibly, by OOP micro-RNA serving in some manner as a competitor to RNA I made by the plasmid, resulting in the loss of plasmid copy control and higher plasmid copies per cell. We view this explanation as unlikely. Alternatively, OOP RNA may limit the adverse effects of *P* expression. It can bind to *pR*-promoted *cII* mRNA and target its degradation in an RNaseIII-dependent reaction [76,77], suggested to limit downstream *O-P* gene expression.

What is the possibility that inefficient termination of *pR*-GOI transcription and consequent run-on transcription could cause problems with plasmid maintenance and explain the loss of pcIpR-*P*-timm from cells, rather than a direct effect of P on plasmid maintenance? We have observed that the wild type *timm* terminator prevents read through from the *cI-rex* transcript into the *oL* operator region or through into *N* when CII stimulated transcription from *pE* occurs at 30 to 100-fold the level of *cI-rex* transcription from the *cI* maintence promoter *pM* (see [17] and contained references). This was the rationale for incorporating *timm* into the design of pcIpR-(GOI)-timm plasmids. Table 7 shows the influence of expression of lambda gene *O* from plasmids pcIpR-*O*-timm and p434'pR-*O*-timm on plasmid loss at 30, 35, 37, 39 and 42 °C. The immunity properties of these two plasmids was shown in Table 4 of [36]. The expression of gene O is constitutive at all temperatures from plasmid p434'pR-*O*-timm. It can be seen that 100% cellular maintenance / retention occurred for plasmid pcIpR-*O*-timm between 30–39 °C and 97% at 42 °C. Even with constitutive expression of *O* from

p434'pR-*O*-timm over the range of cell growth temperatures, between 87%–97% of the cells retained the plasmid. A similar result is also seen in Table 6 for plasmid pcIpR-$P^{\Delta76}$-timm, where 99% of the cells grown between 25 and 42 °C retained the plasmid. These results strongly support an argument that *timm* in pcIpR-(GOI)-timm is likely a powerful terminator that prevents inefficient termination and consequent run-on transcription, which might otherwise cause problems with plasmid maintenance.

2.5. Replicative (cis) killing; P (trans) Lethality / Inhibition

We compared the kinetics of cell death resulting from *P* expression in cells with pcIpR-*P*-timm, *i.e.*, P-lethality (Table 8A) with Replicative Killing, resulting from initiating *ori*λ replication from the Y836 chromosome (Figure 1C) or in cells where this fragment was moved by transduction (Table 8B). Shifting these cell cultures to 42 °C denatures the reversibly-denaturable CI[Ts]857 repressor [17,78] resulting in transcription from promoter *pR*, and *P* expression from the plasmid, or *cro-cII-O-P-ren* expression from the chromosomally inserted λ fragment. The latter event results in the initiation of bidirectional replication from *ori*λ. After the period of gene de-repression the cells were swirled in an ice bath and then plated for survivors at 30 °C, which serves to shut off transcription from *pR* via the renaturation of CI, which, in turn stimulates *cI* expression.

Table 8. Contrasting *trans* P-lethality/inhibition and *cis* Replicative Killing.

A.	**Cells tolerate short-term exposure to P (short-term interference, not lethality)** [a]		
Strain with plasmid	Incubation at 37 °C	Cell viability (Amp^R cfu / total cfu)	Outcome of P expression from plasmid
594[pcIpR-*P*-timm]	1 h	0.87 (232/242)	most cells recover
	2 h	0.49 (258/281)	many cells recover
	6 h	0.13 (2/173)	high plasmid loss
B.	**Inducing replication from a trapped cryptic prophage causes Replicative Killing** [a]		
Strains with cryptic prophages	Prophage Induction time	Cell viability	Outcome of prophage Induction
Y836[~*cIII-cI*857-*O-P-ren*]	5 min	0.33	rapid cell killing
	20 min	0.13	rapid cell killing
	3 h	0.00018	extensive cell killing
	5 h	0.00008	extensive cell killing
594[~*cIII-cI*857-*O-P-ren*]	5 min	0.55	rapid cell killing
	20 min	0.25	rapid cell killing
	3 h	0.0022	extensive cell killing
	5 h	0.0022	extensive cell killing
Y836[~*cIII-cI*857-*O-P::kan-ren*]	3 h	5.1 [b]	cell growth
	5 h	6.1 [b]	cell growth

[a] Refer to Experimental Section *3.6*. [b] The increase in viability by for example 5.1 indicates somewhat more than two and less than three cell doublings.

The λ fragment is not excised from the chromosome of strain Y836 or from the (*cIII-ren*)λ-transduced 594 variant since each inserted λ fragment lacks the genes for *int-xis-kil* and is deleted for both *attL* and *attR* sites. The *ori*λ initiation event from these integrated λ fragments exerts a

rapid *cis*-active Replicative Killing effect, likely causing interference with *E. coli* replication fork progression. In *cis*-killing 75%–87% of the cells die (*i.e.*, are not rescued upon spreading / incubating the cells at 30 °C) by 20 min after prophage induction (Table 8B). But the *trans*-P lethality/inhibition effect is reversible, to a point, and much slower as only 14% of the cells are killed after 1 h (55% after 2 h) of *P* expression, even though >90% of the cells retained the plasmid (Table 8A).

2.5. 1. Contrasting *trans* P-lethality / inhibition, and *cis* Replicative Killing

We showed that shifting culture cells transformed with pcIpR-*P*-timm and grown at 25 °C up to 37 °C for two hrs reduced cell viability to about half, and thus cells are capable of tolerating or metabolizing some level of P, with a reported half-life of up to an hr [79,80]. However, initiating irreversible *ori*λ replication from a trapped cryptic prophage, Figure 1C, reduced cell viability between 4 and 10-fold after only 20 min of induction. The Replicative Killing (RK[+] phenotype) survivor frequency of 10^{-6} to 10^{-8} is distinguished by rare RK[-] mutants in host or prophage (not merely in *P*) that block some aspect of λ replication initiation [78,81–83]. We suggest the terms *cis* Replicative Killing and *trans* P lethality to distinguish mechanistically these the two ideas. The RK[+] phenotype seems uniquely dependent upon multiple, non-repairable λ replication forks arising from *ori*λ (drawn in Figure 1C), but there are likely multiple possibilities for P-lethality. Several experiments suggest that the encoded gene *ren* downstream of *P* is not responsible for P-lethality. A modified version of pHB30 (*i.e.*, pHB31) encoding *P-ren,* but deleted for bases 39609–39836 within *P* (equivalent to $P^{\Delta 76}$ used herein) was fully capable of transforming 594 cells at 42 °C, suggesting that P, not Ren blocked transformation at 42 °C [31]. RK[-] mutants of Y836 with insertions or deletions (recombinant or natural) within/inactivating *O* (ilr208b, ilr223a, ilr541c, ilr200b, ilr203b, ilr207d, ilr201b), or *P* (ilr566a, Bib11t), yet which sequenced to be *ren*[+], lost the Replicative Killing competence phenotype when shifted from 30 to 42 °C, suggesting that replication initiation from *ori*λ, and not *ren* expression, is responsible for Replicative Killing; transduction of the λ fragments into 594 did not alter the phenotype observed for Y836 [RK[+]], or the RK[-] mutants. Ren is not required for P-lethality phenotype, nor for Replicative Killing, *i.e.*, in the absence of *ori*λ replication initiation, *ren* expression will not produce the RK[+] phenotype, however, our experiments do not rule out some ancillary role for Ren.

Several examples reported suggest that P-lethality is a separate mechanism from P-inhibition of ColE1 replication: i) *grpD55* cells with pcIpR-P^{π}-timm were killed at 37–39 °C, yet 100% of the surviving cells retained the plasmid; ii) most of the cells with pcIpR-*oop-O-P*-timm plasmids retained the plasmids at 39 °C while suffering high P-lethality. Additionally (see Figure 4), cells defective in host protease genes *clpB* or *hflA* were sensitive to P lethality at 37–39 °C, yet the majority of the survivors retained the pcIpR-*P*-timm plasmid (data not shown; available from authors).

Is P-inhibition due only to a decrease in available DnaB, or does it influence / perturb ongoing replication, or replication restart? Our results do not indicate if P fully sequesters DnaB or if a single P monomer bound to DnaB is sufficient. Moreover, the genetic evidence does not show whether this interaction is persistent or transient. There seems little doubt that competing DnaB away from DnaC is an important function of P; but the longer term consequence is that the interaction of P and DnaB

yields a P-DnaB-ATP dead-end ternary complex [10], and the only known role for P is in early *ori*λ-dependent initiation events, of which only a few are required. We propose that the full sequestration of DnaB by P is not necessary in order for P to interfere with ongoing plasmid replication, because of the sensitivity of ColE1 replication / copy maintenance to leaky (sub-induction) levels of P (arising from a plasmid encoding *P*). The simple explanation for plasmid loss is that while the cells continue to replicate and divide, plasmid replication is differentially inhibited. For example, *grpD55* cells very poorly support vegetative growth of λ at 30 °C, and not at all between 37–42 °C, whereas cellular growth is not noticeably perturbed. A low level of P, tolerated by cells, could fully inhibit ColE1 replication, in turn reducing plasmid copy number with each cell division, which could explain plasmid curing for cells grown between 34–36 °C. Less clear is how cells with pcIpR-*P*-timm grown at 30 °C are cured when shifted to 37–42 °C, where division is inhibited and cells elongate forming filaments (Sect 2.8.).

2.6. Does P Expression from pcIpR-P-timm perturb λ Vegetative Growth?

As previously noted, although present in very limited amounts in the cell, DnaB is a multifunctional protein involved in replication fork movement [3], serving as a mobile promoter in priming reactions [4,5], in the progression of Holliday junctions [6,84], and in replication restart reactions [70,71], some of which are likely involved in λ replication beyond the *ori*λ-O-P-dependent initiation step. Since the interaction of P and DnaB yields a P-DnaB-ATP dead-end ternary complex [10], we wondered if the expression of P from fully induced pcIpR-*P*-timm in cells shifted to 42 °C could serve to limit λ replication, as monitored by phage burst. Lysogenic cultures with a λ*cI*[Ts]857 *S*am7 prophage defective for natural cell lysis (because of the nonsense mutation in *S*) were synchronously induced by shifting the cells from 30 to 42 °C. One of the parallel cultures contained pcIpR-*P*-timm. Following induction, the cells were artificially lysed. Relative phage burst from each of the cultures was determined by dividing the released phage titer by the cell titer pre-induction. The cells with pcIpR-*P*-timm showed an increased relative burst by 60 min post induction over parallel cells without the plasmid (*i.e.*, bursts of 248 (cells with pcIpR-*P*-timm) and 42). Clearly, λ replication/maturation is not curtailed (and appears enhanced) by combined expression of *P* from both induced prophage and plasmid, which is dramatically opposite to the effect of *P* expression on ColE1 plasmid replication/maintenance.

2.6.1. λ Replication and Phage Maturation

Our results suggest that P is not inhibitory to λ replication, which, considering its influence on *E. coli* and ColE1 replication, raises some interesting questions. Doesn't λ need to deal with replication restart, likely one of several possible targets for P-DnaB interaction, or does it have an alternative mechanism? Does P or a P-complex have an unrecognized DNA helicase activity? While some *P* analogues of lambdoid-type phages encode their own DNA helicase activity (see suppl. Figure S2

in [36], and [85] for comparison of 20 *P*-like genes in lambdoid phages), P is not recognized as having this property and those P-like proteins with putative helicase activities are larger than λ *P*.

2.7. P-Induced Cellular Filamentation

We induced *P* expression from plasmids pHB30 and pcIpR-*P*-timm and followed filament formation (Table 9) in SOS$^+$ strain 594, and in SOS-defective variants [86] of 594 made by transducing in the alleles of *lexA*3[Ind$^-$], where the alanine-glycine protease cleavage site [87] is changed to alanine-aspartate [88] (Figure 5, Table 9), or Δ*recA* (Table 9)

Figure 5. Cellular filamentation resulting from induced P-expression. The photos are representative of data shown for one of the three sets of photos per strain per assay condition that were used for cell measurements in Table 9.

The results in Table 9 suggest that *P* expression negatively impacts *E. coli* cell division. The cellular filamentation observed upon shifting 594 *lexA*3[pcIpR-*P*-timm] cells to 37 °C (Table 9) supports the concept (e.g., from Figure 3) that some leaky *P* expression arises from this plasmid at 37 °C in sufficient level to prevent ColE1 plasmid replication maintenance.

Table 9. SOS-independent P-induced cellular filamentation. [a]

Strain [plasmid]	Time (Temp)	Relative cell length										Sum cells
		1X	2X	3X	4X	5X	6X	7X	8X	9X	≥10X	
594	0 (25)	26	19	0	0	0	0	0	0	0	0	45
"	1 (42)	15	11	11	8	0	0	0	0	0	0	45
"	3 (42)	28	15	2	0	0	0	0	0	0	0	45
"	5 (42)	29	14	2	0	0	0	0	0	0	0	45
594 *lexA3*	0 (25)	40	32	2	0	0	0	0	0	0	0	74
"	1 (42)	23	16	6	0	0	0	0	0	0	0	45
"	3 (42)	31	13	1	0	0	0	0	0	0	0	45
"	5 (42)	42	3	0	0	0	0	0	0	0	0	45
594 *lexA3*[pcIpR-*P*-timm]	0 (25)	71	19	0	0	0	0	0	0	0	0	90
"	1 (37)	21	23	1	0	0	0	0	0	0	0	45
"	3 (37)	7	9	9	6	4	3	2	1	2	2	45
"	5 (37)	6	20	0	0	2	0	3	3	3	8	45
594 *lexA3*[pcIpR-*P*-timm]	1 (42)	9	19	13	3	1	0	0	0	0	0	45
"	3 (42)	11	1	3	4	4	9	8	2	2	1	45
"	5 (42)	8	6	7	1	1	0	2	1	2	14	42
594 [pHB30]	0 (30)	6	20	4	0	0	0	0	0	0	0	30
"	1 (42)	0	2	10	9	3	4	1	1	0	0	30
"	3 (42)	2	4	8	3	2	5	2	1	2	1	30
"	5 (42)	2	8	10	2	2	2	3	0	0	1	30
594 *lexA* [pHB30]	0 (30)	5	17	8	0	0	0	0	0	0	0	30
"	1 (42)	0	6	6	8	4	1	4	0	1	0	30
"	3 (42)	3	5	2	6	4	0	5	3	1	1	30
"	5 (42)	1	7	5	6	5	2	1	1	0	2	30
594 Δ*recA*	0 (30)	16	12	2	0	0	0	0	0	0	0	30
"	1 (42)	4	12	7	4	1	2	0	0	0	0	30
"	3 (42)	9	16	2	4	2	0	0	0	0	0	30
"	5 (42)	12	11	4	2	0	0	1	0	0	0	30
594 Δ*recA* [pHB30]	0 (30)	8	17	4	1	0	0	0	0	0	0	30
"	1 (42)	2	5	8	10	3	1	1	0	0	0	30
"	3 (42)	5	4	5	3	6	4	2	0	0	1	30
"	5 (42)	3	5	4	8	0	4	2	0	1	3	30

[a] Refer to Experimental Section *3.7*.

2.7. 1. P-Dependent Cellular Filamentation

We previously observed extensive cellular filamentation upon the de-repression of λ*cI*[Ts]857*O*am8 prophage, but not for a λ*cI*[Ts]857*P*am3 prophage in lysogenic cells [31]. Klinkert and Klein [23] demonstrated that expression of *P* from a plasmid can cause cellular filamentation, which they attributed to an ability of P to impair bacterial DNA synthesis. Agents that block cellular DNA

synthesis can stimulate a cellular SOS response [86], which in turn elevates expression of one of the SOS response genes, *sulA*, whose product prevents the action of FtsZ at the site of septum formation, causing a cell to stop dividing [89,90], and allows time for DNA repair (see [91]). Cells inhibited for FtsZ activity form filaments that are longer than dividing cells. Induction defective mutations in *lexA* encoding the repressor for the SOS response, or certain *recA* mutations prevent the induction of a cellular SOS response [86].

The hypothesis that P-lethality does not involve an interaction between P and DnaB, but is instead targeted to DnaA, implies that the P:DnaA interaction can prevent *E. coli* replication initiation at *oriC* by removing available DnaA. Previous studies have shown that a *dnaA*46[Ts] mutant exhibits a slow-stop effect on DNA synthesis when the cells are shifted to 42 °C, blocking re-initiation at *oriC*; however, these cells continue to divide, forming anucleate cells, not filaments [92,93]. Other cellular interactions beyond those between DnaA or DnaB and P are important. For example, as noted, P interacts with DnaK and DnaJ. The inactivation of *dnaK*, and apparently *dnaJ*, results in multiple cellular defects, including formation of cell filaments, abnormally segregated chromosomes, and loss of plasmid maintenance [94].

Does P function in replication restart? It was reported [69] that neither *lexA* nor *sulA* deletions abolished filamentation in a *priA sulB* mutant, suggesting models whereby repair/restart of chromosomal replication is essential for completing a round of replication, and that replicative forms are resolved to monomers before cell division can take place. Since DnaB is reloaded onto DNA for replication restart, P accumulation could serve to limit the availability of DnaB to replication restart pathways, or it may directly interact with the elongating replisome, neither previously suggested. The need for replication restart is influenced by collisions between transcription and replication forks. RNA polymerase mutations in *rpoB* and *rpoC* [95,96] facilitate replication progression. Indeed, CH has shown that two of 22 mutants obtained in screens for rifampicin-resistant *E. coli* cfu had significant resistance to P-lethality (remainder did not; Hayes lab, unpublished results). However, among the group of Rif[R] isolates, 30 were localized by DNA sequence analysis to two regions within *rpoB* (because *rpoB* is 4029 bp, we did not sequence the whole gene), but the site of mutation in the two resistant isolates remains undetermined, and the existence of additional suppressor mutations arising in these two mutants remains a possibility. We predict that mutations in *E. coli* which decrease conflicts between transcription and replication would make cells more resistant to P, and mutations which increase conflicts would make cells more sensitive.

3. Experimental Section

3.1. Strains Employed

The bacteria, plasmids (Figure 1A) and phages employed are listed in Table 10.

Table 10. Bacteria, plasmids and phages employed.

Bacterial strains	Characteristics or genotype	Source/Ref.'; Hayes lab #[a]
594	F[-] *lac*-3350 *galK*2 *galT*22 *rpsL*179 IN(*rrnD-rrnE*)1; see [97]; called R594	[97], SH lab; B10
TC600	*supE*, Pm[+]	SH lab, B8
Ymel	*supF*, Pm[+]	SH lab, B71
DE407	*lexA*3[Ind[-]] *malB*::Tn*9* Tet[R] *sulA*211 *sfiA*11, UV[S]	D. Ennis; B142
FC40 (=SMR624)	Δ(*srlR-recA*)306::Tn*10* Tet[R] UV[S]	SM Rosenberg [98]; Y921
AB2834 *aroE*	*grpD*55, *thi* tsx[R] λ[R] at 42 °C from K552	H. Uchida [32,33]; NB83
W3874 *malB*5	*dnaB grpA*80 *lac*[-] Str[R] λ[R] at 42 °C	[32], NB81
W3350 *dnaB- grpD*55	*grpD*55 *malF*3089::Tn*10* Tet[R] λ[R] at 42°C, λ*rep*P22[S]	[33], NB15
594 *dnaB-grpD*55	*grpD*55 allele *malF*3089::Tn*10*; Tet[R], λ[R] at 42°C, λ*rep*P22[S]	[34], NB295
594 *lexA*3[Ind[-]] *malB*::Tn*9*	LexA repressor induction defective	CE, NB293
594 Δ(*srlR-recA*)306::Tn*10*	deletion of *recA* Tet[R] UV[S]	CE, B318
W3350	F[-] *lac*[-]3350 *galK*2 *galT*22 *rpsL*179 IN(*rrnD-rrnE*)1	SH lab, B12
Y836	SA500(λ*bio*275*cI*[Ts]857 Δ431) *his*[-]	[78,82], NY1049
594::*nadA*::Tn*10* [~*cIII-ren*][λ]	Tn*10* [zbh29 at 16.8 min] *bio*[+] transductant = 594 *bio*275 (λ*cIII-cI*[Ts]857-O-P-*ren*) Δ431	A. Chu, SH lab, NY1057
Y836 *P*::kan (Bib11t)	SA500 (λ*bio*275 *cI*[Ts]857 O[+]*P*::kan Δ431) *his*[-] Kan[R]	SH, NY1153
594(λ*cI*857*S*am7)	λ lysogen defective for cell lysis	C. Marek, SH lab, Y1163
594(λ*cI*857*S*am7)[pcIpR-*P*-timm]	as above with transformed plasmid	SH lab, P509
594 *clpP*::kan	*clpP*[-], Kan[R] from SG22159	S. Gottesman; [99], NB276
Plasmids	**Transformed into strain 594**	**Source/Ref.'; Hayes lab #[a]**
pUC19	Wild type Amp[R] (New England Biolabs)	NP188
pcIpR-*P*-timm	*Bam*HI-*Cla*I PCR fragment from λcI857, replacing D-CAP in P459 with λ bp's 39582-40280	CH, P466
pcIpR-*P*::kan-timm	PCR *Bam*HI-*Cla*I fragment from Y836 *P*::kan (Bib11t) strain NY1153	KM, P510
PcIpR-*P*-SPA-timm [b]	Replace D in P462 between *Bam*HI and AscI sites with *Bam*HI-P(λ bp's 39582-40280)-*Asc*I PCR fragment	KM, P467
pcIpR-*P*[Δ76]-timm	In-frame deletion76 codons: λbp 39609-39836 in pcIpR-*P*-timm with HpaI, ligate	KM, P515
pcIpR-*P*[π]-timm	*Bam*HI-*Cla*I PCR fragment from λcI72π Lysate #3a, replacing D-CAP in P459 with λ bp's 39582-40280	KM, P505

Table 10. *Cont.*

Plasmids	Transformed into strain 594	Source/Ref.'; Hayes lab #[a]
pcIpR-*O*-timm	*Bam*HI-*Cla*I PCR fragment from λcI857, replacing D-CAP in P459 with λ bp's 38686-39582	[36], CH, P465
P434'pR-*O*-timm	Constitutive *O* expression; *Bam*HI-*Cla*I PCR fragment from λcI857, replacing D-CAP in P459 with λ bp's 38686-39582	[36], CH, P494
pcIpR-*O-P*-timm	*Bam*HI-*Cla*I PCR fragment from λcI857, replacing D-CAP in P459 with λ bp's 38686-40280	CH, P569
pcIpR-*O*-36*P*-timm	*Bam*HI-*Cla*I PCR fragment from λcI857, replacing D-CAP in P459 with λ bp's 38686-39687	CH, P565
pcIpR-*O*-63*P*-timm	*Bam*HI-*Cla*I PCR fragment from λcI857, replacing D-CAP in P459 with λ bp's 38686-39768	CH, P566
pcIpR-*oop*#1-*O-P*-timm	*Bam*HI-*Cla*I PCR fragment from λcI857, replacing D-CAP in P459 with λ bp's 38559-40280	CH, P567
pcIpR-*oop*#2-*O-P*-timm	*Bam*HI-*Cla*I PCR fragment from λcI857, replacing D-CAP in P459 with λ bp's 38546-40280	CH, P568
pHB30	λ bases 34499-34696, 36965-38103, 38814-40806 (see Section 3.2.)	[31,34], SH lab, P8
Bacteriophage	**Genotype**	**Hayes lab lysate #**
λ wild type (wt)	λpapa	[78], 944,1001
λ*cI*72	*cI*	[78], 951, 999
λnin5	made from λ wt	[78], CH, 698
Λvir	λ*v2v1v3*	[78], 260
λ*cI*857	*cI*[Ts]857	[100], 1002
λ*cI*857*S*am7	defective for cell lysis	[101], 963
λ*imm*434*P*am3	*imm*434, sequenced *P*am3 mutation C to T, λ base 39786 (CAG to TAG)	[83], SH lab, 518, 664
λ*imm*434nin5	*imm*434, Δnin5 region, forms very turbid plaques at 37°C	[22], CH, 963

[a] The strain numbers are from the Hayes laboratory collections. All gene inserts within the pcIpR-[ ]-timm plasmids were sequenced to confirm the genetic integrity of the inserted fragment. [b] Plasmid pcIpR-D-SPA-timm [20] (strain P462) was prepared from pcIpR-D-CAP-timm (strain P459), replacing 318 bp CAP from P459 by digestion with *Asc*I and *Cla*I and replacing with 239 bp SPA tag from pMZS3F [72] (from J. Greenblatt) isolated via PCR with primers L-Asc-CBP & R-ClaI-FLAG. SPA is a 66 amino acid tag with 3X FLAG sequences.

3.2. Construction of Expression Vectors for λ Genes, Gel Analysis of Plasmids, Insertion Localization

The plasmids were extracted from cells using Qiagen plasmid mini preps. They were separated by agarose electrophoresis on gels made and run using 1X TBE buffer (10X = 1M Tris, 1M Boric acid, 0.02 M Na$_2$EDTA, pH 8). The precursor plasmid pcIpR was made by degrading pBR322 with *Eco*RI and *Bam*HI and purifying the large *Eco*RI-*bla*-rep/rop-*Bam*HI fragment. This was ligated with an 833 bp PCR fragment derived from λ*cI*857 DNA amplified using primers that added MfeI and *Bam*HI tags to the ends of lambda bases 37203 and 38036 to produce plasmid pcIpR, where the ligation of the MfeI end to the *Eco*RI site removed both sites in the resulting construct. pcIpR was digested with *Bam*HI and SalI and the large fragment of 4817 bp was separated from the 276 bp region between the *Bam*HI and SalI sites. A synthetic DNA sequence including λ oR/pR region, a *Bam*HI site, gene fusion D-CAP-*Cla*I-timm-*Eco*RI-SalI sequence provided by IDT, Coralville IA, was digested with *Bam*HI and SalI and the 710 bp fragment was ligated to the 4817 bp fragment to produce pcIpR-D-CAP-timm [20]. This permitted positioning the AUG for the D-CAP fusion protein immediately to the right of the *Bam*HI sequence, and blended into the consensus Shine-Delgarno sequence, so that the D-fusion orf was positioned 18 bp downstream from the mRNA start site for the λ *pR* promoter (see Figure 1, [20]). The D-CAP orf was removed from pcIpR-D-CAP-timm λ by digestion with *Bam*HI and *Cla*I and the resulting large fragment was ligated with gene *P*, produced by generating a *Bam*HI-*P*-*Cla*I PCR fragment, including the precise sequence of *P* (λ bp's 39582 – 40280) from λ*cI*857 DNA to yield pcIpR-*P*-timm, where a TAA stop codon was added that immediately followed the *P* insertion. Plasmid pcIpR-$P^{\pi-39991}$-timm was generated by inserting a *Bam*HI-P^{π}-*Cla*I PCR fragment made from λ*cI*72-$P^{\pi-39991}$ phage. The plasmid pcIpR-$P^{\Delta76}$-timm was constructed by using restriction endonuclease *Hpa*I to delete λ bp 39609–39836, *i.e.*, bp 28 through 255 within the N-terminal end of *P*, then ligating to fuse codon 9 with codon 86. The plasmid pcIpR-*P*-SPA-timm was constructed by removing the *Bam*HI to AscI fragment from the 5155 bp plasmid pcIpR-D-SPA-timm [20], and ligating with the remaining 4807 bp fragment a *Bam*HI-*P*-AscI PCR fragment encoding lambda bases 39582 to 40280 to produce the 5524 bp plasmid pcIpR-*P*-SPA-timm. SPA is a 66 amino acid tag sequence [72] with both calmodulin and 3XFLAG binding recognition sequences. Plasmids pcIpR-*O*-timm and p434'pR-*O*-timm have the precise *O* sequence (ATG=38686-39582) plus TAA stop codon. In p434'pR-*O*-timm, the SD differed by one bp compared to SD in pcIpR-*O*-timm because of the slightly different sequence ahead of *cro* in *imm*434 DNA [102]. Plasmids pcIpR-*O*-36P-timm and pcIpR-*O*-63P-timm, respectively, have λ DNA sequences 38686–39687, or 38686–39768, each including an intact *O* sequence plus an extension comprising the N-terminal portion of *P*, followed by TAA stop at the end of inserted partial *P* sequence. Plasmids pcIpR-*O*-*P*-timm (with precise *O*-*P* sequence), pcIpR-*oop*#1-*O*-*P*-timm and pcIpR-*oop*#2-*O*-*P*-timm (each with DNA from within *cII* through *P*) have inserted sequences, respectively: 38686-40280, 38559-40280, and 38546-40280, followed by TAA stop codon. All GOI inserts within the pcIpR-[GOI]-timm plasmids were sequenced to confirm the genetic integrity of the inserted fragment. pHB30 [31] contains λ genes *cI*[Ts]857, a *cro-O* in frame fusion,

P-ren; *i.e.*, the pBR322 bases from 375–4286 and λ bases (*Bam*HI)34499-34696(*Cla*I)-(*Cla*I)36965-38103 (BglII)-(BglII)38814-40806(AatII) and was re-characterized and sequenced [34].

3.3. Plasmid Transformation; Phage and Culture Assays

Cells from a single colony of the *E. coli* strain being transformed with a plasmid were inoculated into 20 ml fresh LB (5 g NaCl, 10g Bacto Tryptone, 10 g Bacto Yeast Extract per liter), grown overnight, subcultured into fresh LB medium and grown at 30 °C to A_{575} = 0.4, which equals about 4×10^8 cfu (colony forming units) per ml. The cells (1 ml) were centrifuged in 1.5ml microtubes for 1 min at 12×10^3 rpm in an Eppendorf 5424 microcentrifuge. The supernatant was decanted and cell pellet washed with 750ul 0.01M NaCl. The cells were again pelleted, suspended in 750 μl ice cold 0.03 M $CaCl_2$, incubated on ice for 30 min, pelleted, and resuspended in 150 ul 0.03 M $CaCl_2$ = competent cells. 200 ng DNA of the plasmid in TE* buffer (0.01 M Tris, 0.001 M Na_2EDTA, pH 7.6) was combined with the competent cells and mixed gently. The mixture was held on ice for 60 min, the tubes were heat shocked at 42 °C for 90 seconds in a heating block and plunged on ice for 2 min. 850 ul of room temperature LB was added to each sample tube and these were incubated with gentle shaking in a 25 °C water bath for 90 min. At the end of incubation time the cell samples were diluted in ø80 buffer (1.2 g Tris, 5.8 g NaCl per liter, pH 7.6). Then aliquots (0.1 ml) were spread on LB or TB (10 g Bacto tryptone, 5 g NaCl per liter) agar plates. The TB or LB medium used for the solid support agar plates included the addition of 11 g Bacto agar per liter prior to autoclaving. The agar plates used for screening Amp^R cfu were supplemented with 50 μg/ml of ampicillin (=Amp50) added after the autoclaved agar medium had cooled, prior to pouring the plates. Molten TB top agar (10 g Bacto tryptone, 5 g NaCl and 6.5 g Bacto agar per liter) was used for plating phage.

3.4. Sequence Analysis of Alleles of dnaB

The *dnaB* genes were amplified with primers DnaB-1 and DnaB-6. The PCR fragments were sequenced with overlapping primer pairs DnaB-1 and DnaB-2, DnaB-3 and DnaB-4, and DnaB-5 and DnaB-6 [34]. The *grpA80* and *grpD55* GenBank accession nos. are DQ324464 and DQ324465.

3.5. Assessing Influence of Modifying P or dnaB on P Lethality and Plasmid Loss

The cultures were grown up to stationary phase in TB culture medium plus 50μg/ml ampicillin for 48 hr at 25 °C, diluted, spread on TB agar plates that were incubated at 25, 30, 37, 39, or 42 °C for 48 h and cfu per ml was determined. Survivor cfu arising on the plates were stabbed to TB and TB+Amp50 plates to estimate the proportion of cfu retaining the plasmid. We tried to assay all cfu per dilution plate sector(s) to avoid colony size discrimination. Individual cfu arising from 594 *grpD55* [pcIpR-*P*-timm] from the 25 or 42 °C plates (single colonies – sc1 through sc6) were inoculated into TB+Amp50 plates, incubated 48 hrs at 25 °C, and plasmid DNA was extracted. The extracted plasmid preparations from the six individual cultures were each transformed into 594 culture cells and Amp^R cfu were selected on agar plates incubated at 25 °C. Individual single colonies (sc's 1-6) from these

plates were inoculated into TB with 50 µg/ml ampicillin and grown 48 h at 25 °C, then diluted and spread on TB agar plates that were incubated at 25, 37, 39, or 42 °C for 48 hr and cfu per ml was determined. Survivor cfu arising on the plates were stabbed to TB and TB+Amp50 plates to estimate the proportion of cfu retaining the plasmid. The experiment with the *re-claimed* plasmids was undertaken to determine if the survivor plasmids extracted from the 594 *grpD55* [pcIpR-*P*-timm] cultures grown at 42 °C retained the P-lethality phenotype.

3.6. Contrasting trans *P-Lethality / Inhibition and* cis *Replicative Killing*

The influence of transient *P* expression (*trans* P lethality) from pcIpR-*P*-timm on cell viability and plasmid loss was assessed by incubating cells diluted and spread on pre-heated agar plates, that were held at 37 °C for 1, 2, or 6 h, and then incubated at 25 °C for about 72 hrs for survivor cfu. The survivors were stabbed to TB+Amp50 agar plates to assess for plasmid loss. Replicative Killing, or *cis* killing, involves the irreversible effect of inducing gene expression from a trapped (nonexcisable) cryptic fragment of λ [~*cIII-ren* genes in Y836] inserted within bacterial chromosome. The result is the formation of an onion-skin replication bubble at *ori*λ site in bacterial chromosome as shown in Figure 1C. The λ fragment in strain Y836 was transduced into 594 in two steps by P1 transduction: a) moving *nadA*::Tn*10* into Y836, and b) moving the *nadA*::Tn*10* [~*cIII-cI*857-*O-P-ren*]^λ fragment into 594. Gene *P* expression in strain Y836 was inactivated by recombineering involving use of primers L-P-stop-kan [5'-gaccgtgagcagatgcgtcggatcgccaacaacatgactaactagctctgatgttacattgcacaag**]** and R-kan-stop-P [5'-ggtcgattctgccgacgggctacgcgcattcctgcgctagttagtcagtcagcgtaatgctctgcca] and the insertion of Kan^R (same as in pBR322::kan) into *P* to make strain Y836 [~*cIII-cI*857-*O-P*::kan-*ren*] (isolate Bib11t), where Kan substituted the bases 39651-39838 of *P*.

3.7. SOS-Independent P-Induced Cellular Filamentation

Overnight cultures of each strain were prepared in TB (plus 50 µg/ml ampicillin for strains with plasmids) and grown at 30 °C for the controls and strains with pHB30, and at 25 °C for strains with pcIpR-*P*-timm. Subcultures were made into 50 ml TB or TB-Amp50 and grown at 25 or 30 °C, as shown for 0-time, to an A575nm of 0.01 to 0.15. An aliquot was removed for the 0-time assay and the flasks were transferred to 37 or 42 °C shaking water baths. For each assay point culture aliquots were removed (5 µl, and 1 µl then diluted with 4 µl φ80 buffer) and placed on glass slides. The samples were allowed to air dry and gently fixed over a flame. The slides were prepared by Gram stain, and then examined by light microscopy, taking three pictures per slide of areas with lower density cells. The digital pictures were projected onto a large screen and the length per cell was manually measured, 15 cells per slide. Representative culture absorbance 0 time, 1, 3 and 5 hrs at inducing temperature: 594 (0.09, 0.22, 0.74, 0.95); 594 *lexA*3 (0.01, 0.25, 0.68, 0.95); 594 *lexA*3[pcIpR-*P*-timm] induced to 37 °C (0.10, 0.26, 0.54, 0.76); and 594 *lexA*3[pcIpR-*P*-timm] induced to 42 °C (0.10, 0.25, 0.50, 0.67). The value for 1X represented the average of ten smallest cell measurements for 594 cells, representing

35.6 mm. All values were rounded up into next category, e.g., cells with length >53.4 mm were presented as 2X average length.

3.8. P-Induced Plasmid Loss

The volume of cells utilized for plasmid extraction from each of the cultures was normalized to achieve a final A_{575} of 1.0. Aliquots were removed to extract plasmid DNA and isolate cfu on TB agar plates. The survivor cfu arising on the TB plates were stabbed to TB and TB+Amp50 plates to estimate the proportion of cfu retaining the Amp^R plasmid. Plasmid DNA was extracted from aliquots (5 ml) of the cultures and mixed with 0.5 ml of a stationary phase culture of 594[pUC19] (serving as an internal plasmid extraction / gel loading control). DNA was extracted from the cell pellet(s) using QIAgen spin mimiprep kits and suspended in 0.05 ml elution buffer. Aliquots from all DNA preparations were digested to completion with restriction endonuclease *Eco*RI, which cuts each plasmid once. The digests were run on 0.8% agarose gels in TBE buffer at 90 volts for 90 min and then stained with ethidium bromide for 10 min. DNA band sizes were estimated using a 1Kb DNA ladder (left gel lanes in B and C). Plasmid pcIpR-*P*-timm is 5292bp, pcIpR-*P*$^{\Delta76}$-timm is 5064bp, and the high copy pUC19 is 2686bp.

4. Conclusions

Our complementation results suggest that only trace levels of *P* expression are needed to catalyze the initiation of λ replication, providing an explanation for early observations about "trans-activation." It appears that ColE1 replication is extremely sensitive to P, and cells with repressed *P*-encoded plasmids (*i.e.*, with sub-induction levels of *P* expression) can lead to ColE1 plasmid curing. Both P-lethality to cells, and the observed P-dependent cell loss of ColE1 plasmids, were fully suppressed by dual missense mutations altering *dnaB*, or an in frame deletion near the N-terminal end of P. The P-inhibitory phenotypes were partially suppressed by a π missense mutation in *P*, and plasmids expressing the λ encoded OOP antisense micro-RNA. P-dependent cellular filamentation was observed in *ΔrecA* or *lexA*[Ind⁻] cells, considered defective for SOS induction. These studies suggest the hypothesis that cellular levels of P can directly interfere not only with *E. coli* replication initiation, but subsequent steps involving DNA propagation and replication restart.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Tsurimoto, T.; Matsubara, K. Purified bacteriophage lambda O protein binds to four repeating sequences at the lambda replication origin. *Nucleic Acids Res.* **1981**, *9*, 1789–1799.

2. Wickner, S.H. DNA replication proteins of *Escherichia coli* and phage lambda. *Cold Spring Harbor Symp. Quant. Biol.* **1979**, *43 Pt 1*, 303–310.

3. LeBowitz, J.H.; McMacken, R. The *Escherichia coli* DnaB replication protein is a DNA helicase. *J. Biol. Chem.* **1986**, *261*, 4738–4748.

4. McMacken, R.; Ueda, K.; Kornberg, A. Migration of *Escherichia coli* DnaB protein on the template DNA strand as a mechanism in initiating DNA replication. *Proc. Natl. Acad. Sci. USA* **1977**, *74*, 4190–4194.

5. Zyskind, J.W.; Smith, D.W. Novel *Escherichia coli dnaB* mutant: Direct involvement of the *dnaB*252 gene product in the synthesis of an origin–ribonucleic acid species during initiaion of a round of deoxyribonucleic acid replication. *J. Bacteriol.* **1977**, *129*, 1476–1486.

6. Bujalowski, W. Expanding the physiological role of the hexameric DnaB helicase. *Trends Biochem. Sci.* **2003**, *28*, 116–118.

7. Kaplan, D.L.; O'Donnell, M. DnaB drives DNA branch migration and dislodges proteins while encircling two DNA strands. *Mol. Cell* **2002**, *10*, 647–657.

8. Arai, K.; Kornberg, A. Mechanism of dnab protein action. Ii. ATP hydrolysis by DnaB protein dependent on single– or double–stranded DNA. *J. Biol. Chem.* **1981**, *256*, 5253–5259.

9. Reha–Krantz, L.J.; Hurwitz, J. The *dnaB* gene product of *Escherichia coli*. I. Purification, homogeneity, and physical properties. *J. Biol. Chem.* **1978**, *253*, 4043–4050.

10. Biswas, S.B.; Biswas, E.E. Regulation of *dnaB* function in DNA replication in *Escherichia coli* by *dnaC* and lambda P gene products. *J. Biol. Chem.* **1987**, *262*, 7831–7838.

11. Mallory, J.B.; Alfano, C.; McMacken, R. Host virus interactions in the initiation of bacteriophage lambda DNA replication. Recruitment of *Escherichia coli* DnaB helicase by lambda P replication protein. *J. Biol. Chem.* **1990**, *265*, 13297–13307.

12. Alfano, C.; McMacken, R. Ordered assembly of nucleoprotein structures at the bacteriophage lambda replication origin during the initiation of DNA replication. *J. Biol. Chem.* **1989**, *264*, 10699–10708.

13. Zylicz, M.; Ang, D.; Liberek, K.; Georgopoulos, C. Initiation of lambda DNA replication with purified host– and bacteriophage–encoded proteins: The role of the DnaK, DnaJ and GrpE heat shock proteins. *EMBO J.* **1989**, *8*, 1601–1608.

14. Nakayama, N.; Arai, N.; Bond, M.W.; Kaziro, Y.; Arai, K. Nucleotide sequence of *dnaB* and the primary structure of the DnaB protein from *Escherichia coli*. *J. Biol. Chem.* **1984**, *259*, 97–101.

15. Ueda, K.; McMacken, R.; Kornberg, A. DnaB protein of *Escherichia coli*. Purification and role in the replication of phix174 DNA. *J. Biol. Chem.* **1978**, *253*, 261–269.

246

16. Hayes, S.; Bull, H.J.; Tulloch, J. The Rex phenotype of altruistic cell death following infection of a lambda lysogen by T4rII mutants is suppressed by plasmids expressing OOP RNA. *Gene* **1997**, *189*, 35–42.

17. Hayes, S.; Slavcev, R.A. Polarity within *pM* and *pE* promoted phage lambda *cI–rexA–rexB* transcription and its suppression. *Can. J. Microbiol.* **2005**, *51*, 37–49.

18. Hayes, S.; Szybalski, W. Control of short leftward transcripts from the immunity and *ori* regions in induced coliphage lambda. *Mol. Gen. Gen.* **1973**, *126*, 275–290.

19. Landsmann, J.; Kroger, M.; Hobom, G. The *rex* region of bacteriophage lambda: Two genes under three–way control. *Gene* **1982**, *20*, 11–24.

20. Hayes, S.; Gamage, L.N.; Hayes, C. Dual expression system for assembling phage lambda display particle (LDP) vaccine to porcine circovirus 2 (PCV2). *Vaccine* **2010**, *28*, 6789–6799.

21. Revet, B.; von Wilcken–Bergmann, B.; Bessert, H.; Barker, A.; Muller–Hill, B. Four dimers of lambda repressor bound to two suitably spaced pairs of lambda operators form octamers and DNA loops over large distances. *Curr. Biol.* **1999**, *9*, 151–154.

22. Hayes, S.; Asai, K.; Chu, A.M.; Hayes, C. NinR– and Red–mediated phage–prophage marker rescue recombination in *Escherichia coli*: Recovery of a nonhomologous *imm*lambda DNA segment by infecting lambda*imm*434 phages. *Genetics* **2005**, *170*, 1485–1499.

23. Klinkert, J.; Klein, A. Cloning of the replication gene P of bacteriophage lambda: Effects of increased P–protein synthesis on cellular and phage DNA replication. *Mol. Gen. Gen.* **1979**, *171*, 219–227.

24. Maiti, S.; Mukhopadhyay, M.; Mandal, N.C. Bacteriophage lambda P gene shows host killing which is not dependent on lambda DNA replication. *Virology* **1991**, *182*, 324–335.

25. Georgopoulos, C.P.; Herskowitz, I. *Escherichia coli* mutants blocked in lambda DNA synthesis. In *The bacteriophage lambda*, Hershey, A.D., Ed. Cold Spring Harbor Laboratory: Cold Spring Harbor, NY, USA, 1971; pp. 553–564.

26. Georgopoulos, C.P. A new bacterial gene (*groPC*) which affects lambda DNA replication. *Mol. Gen. Gen.* **1977**, *151*, 35–39.

27. Sunshine, M.; Feiss, M.; Stuart, J.; Yochem, J. A new host gene (*groPC*) necessary for lambda DNA replication. *Mol. Gen. Gen.* **1977**, *151*, 27–34.

28. Datta, I.; Sau, S.; Sil, A.K.; Mandal, N.C. The bacteriophage lambda DNA replication protein P inhibits the *oriC* DNA– and ATP–binding functions of the DNA replication initiator protein DnaA of *Escherichia coli*. *J. Biochem. Mol. Biol.* **2005**, *38*, 97–103.

29. Datta, I.; Banik–Maiti, S.; Adhikari, L.; San, S.; Das, N.; Mandal, N.C. The mutation that makes *Escherichia coli* resistant to lambda P gene–mediated host lethality is located within the DNA initiator gene *dnaA* of the bacterium. *J. Biochem. Mol. Biol.* **2005**, *38*, 89–96.

30. Bolivar, F.; Betlach, M.C.; Heyneker, H.L.; Shine, J.; Rodriguez, R.L.; Boyer, H.W. Origin of replication of pBR345 plasmid DNA. *Proc. Natl. Acad. Sci. USA* **1977**, *74*, 5265–5269.

31. Bull, H.J. Bacteriophage lambda replication–coupled processes: Genetic elements and regulatory choices. University of Saskatchewan, Saskatoon, SK, Canada, 1995.

32. Saito, H.; Uchida, H. Initiation of the DNA replication of bacteriophage lambda in *Escherichia coli* K12. *J. Mol. Biol.* **1977**, *113*, 1–25.

33. Bull, H.J.; Hayes, S. The *grpD*55 locus of *Escherichia coli* appears to be an allele of *dnaB*. *Mol. Gen. Gen.* **1996**, *252*, 755–760.

34. Horbay, M.A. Inhibition phenotype specific for *ori*–lambda replication dependent phage growth, and a reappraisal of the influence of lambda *P* expression on *Escherichia coli* cell metabolism: P–interference phenotype. University of Saskatchewan, Saskatoon, SK, Canada, 2005.

35. Mandal, N.C.; Lieb, M. Heat–sensitive DNA–binding activity of the cI product of bacteriophage lambda. *Mol. Gen. Gen.* **1976**, *146*, 299–302.

36. Hayes, S.; Horbay, M.A.; Hayes, C. A cI–independent form of replicative inhibition: Turn off of early replication of bacteriophage lambda. *PLoS One* **2012**, *7*, e36498.

37. Herskowitz, I.; Signer, E. Control of transcription from the *r* strand of bacteriophage lambda. *Cold Spring Harbor Symp. Quant. Biol.* **1970**, *35*.

38. Thomas, R. Control of development in temperate bacteriophages. 3. Which prophage genes are and which are not *trans*–activable in the presence of immunity? *J. Mol. Biol.* **1970**, *49*, 393–404.

39. Bejarano, I.; Klemes, Y.; Schoulaker–Schwarz, R.; Engelberg–Kulka, H. Energy–dependent degradation of lambda O protein in *Escherichia coli*. *J. Bacteriol.* **1993**, *175*, 7720–7723.

40. Wegrzyn, A.; Czyz, A.; Gabig, M.; Wegrzyn, G. ClpP/clpX–mediated degradation of the bacteriophage lambda O protein and regulation of lambda phage and lambda plasmid replication. *Arch. Microbiol.* **2000**, *174*, 89–96.

41. Konieczny, I.; Marszalek, J. The requirement for molecular chaperones in lambda DNA replication is reduced by the mutation pi in lambda P gene, which weakens the interaction between lambda P protein and DnaB helicase. *J. Biol. Chem.* **1995**, *270*, 9792–9799.

42. Klein, A.; Lanka, E.; Schuster, H. Isolation of a complex between the P protein of phage lambda and the *dnaB* protein of *Escherichia coli*. *Eur. J. Biochem.* **1980**, *105*, 1–6.

43. Sutton, M.D.; Carr, K.M.; Vicente, N.; Kaguni, J.M. *Escherichia coli* DnaA protein – the N–terminal domain and loading of DnaB helicase at the *E. coli* chromosomal origin. *J. Biol. Chem.* **1998**, *273*, 34255–34262.

44. Wickner, S.; Hurwitz, J. Interaction of Escherichia coli *dnaB* and *dnaC*(*D*) gene products *in vitro*. *Proc. Natl. Acad. Sci. U.S.A.* **1975**, *72*, 921–925.

45. Lu, Y.B.; Ratnakar, P.; Mohanty, B.K.; Bastia, D. Direct physical interaction between DnaG primase and DnaB helicase of *Escherichia coli* is necessary for optimal synthesis of primer RNA. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 12902–12907.

46. Biswas, E.E.; Chen, P.H.; Biswas, S.B. Modulation of enzymatic activities of *Escherichia coli dnaB* helicase by single–stranded DNA–binding proteins. *Nucleic Acids Res.* **2002**, *30*, 2809–2816.

47. Gao, D.X.; McHenry, C.S. Tau binds and organizes *Escherichia coli* replication proteins through distinct domains – domain IV, located within the unique C terminus of Tau, binds the replication fork helicase, DnaB. *J. Biol. Chem.* **2001**, *276*, 4441–4446.

48. McKinney, M.D.; Wechsler, J.A. RNA polymerase interaction with DnaB protein and lambda P protein during lambda replication. *J. Virol.* **1983**, *48*, 551–554.

49. Maiti, S.; Das, B.; Mandal, N.C. Isolation and preliminary characterization of *Escherichia coli* mutants resistant to lethal action of the bacteriophage lambda P gene. *Virology* **1991**, *182*, 351–352.

50. Del Solar, G.; Giraldo, R.; Ruiz–Echevarria, M.J.; Espinosa, M.; Diaz–Orejas, R. Replication and control of circular bacterial plasmids. *Microbio. Mol. Biol. Rev.* **1998**, *62*, 434–464.

51. Tomizawa, J.; Sakakibara, Y.; Kakefuda, T. Replication of colicin E1 plasmid DNA in cell extracts. Origin and direction of replication. *Proc. Natl. Acad. Sci. U.S.A.* **1974**, *71*, 2260–2264.

52. Tomizawa, J.I.; Sakakibara, Y.; Kakefuda, T. Replication of colicin E1 plasmid DNA added to cell extracts. *Proc. Natl. Acad. Sci. USA* **1975**, *72*, 1050–1054.

53. Dasgupta, S.; Masukata, H.; Tomizawa, J. Multiple mechanisms for initiation of ColE1 DNA replication: DNA synthesis in the presence and absence of ribonuclease H. *Cell* **1987**, *51*, 1113–1122.

54. Kogoma, T. Absence of Rnase H allows replication of pBR322 in *Escherichia coli* mutants lacking DNA polymerase I. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 7845–7849.

55. Naito, S.; Uchida, H. Rnase H and replication of ColE1 DNA in *Escherichia coli*. *J. Bacteriol.* **1986**, *166*, 143–147.

56. Tomizawa, J. Two distinct mechanisms of synthesis of DNA fragments on colicin E1 plasmid DNA. *Nature* **1975**, *257*, 253–254.

57. Itoh, T.; Tomizawa, J. Formation of an RNA primer for initiation of replication of ColE1 DNA by ribonuclease H. *Proc. Natl. Acad. Sci. U.S.A.* **1980**, *77*, 2450–2454.

58. Masukata, H.; Tomizawa, J. A mechanism of formation of a persistent hybrid between elongating RNA and template DNA. *Cell* **1990**, *62*, 331–338.

59. Hillenbrand, G.; Staudenbauer, W.L. Discriminatory function of ribonuclease H in the selective initiation of plasmid DNA replication. *Nucleic Acids Res.* **1982**, *10*, 833–853.

60. Bird, R.E.; Tomizawa, J. Ribonucleotide–deoxyribonucleotide linkages at the origin of DNA replication of colicin E1 plasmid. *J. Mol. Biol.* **1978**, *120*, 137–143.

61. Tomizawa, J.I.; Ohmori, H.; Bird, R.E. Origin of replication of colicin E1 plasmid DNA. *Proc. Natl. Acad. Sci. USA* **1977**, *74*, 1865–1869.

62. Masai, H.; Nomura, N.; Kubota, Y.; Arai, K. Roles of phi x174 type primosome– and G4 type primase–dependent primings in initiation of lagging and leading strand syntheses of DNA replication. *J. Biol. Chem.* **1990**, *265*, 15124–15133.

63. Allen, J.M.; Simcha, D.M.; Ericson, N.G.; Alexander, D.L.; Marquette, J.T.; van Biber, B.P.; Troll, C.J.; Karchin, R.; Bielas, J.H.; Loeb, L.A.; *et al*. Roles of DNA polymerase I in leading and lagging–strand replication defined by a high–resolution mutation footprint of ColE1 plasmid replication. *Nucleic Acids Res.* **2011**, *39*, 7020–7033.

64. Masai, H.; Arai, K. Dnaa– and pria–dependent primosomes: Two distinct replication complexes for replication of *Escherichia coli* chromosome. *Front. Bios.* **1996**, *1*, d48–58.

65. Nomura, N.; Ray, D.S. Expression of a DNA strand initiation sequence of ColE1 plasmid in a single–stranded DNA phage. *Proc. Natl. Acad. Sci. U.S.A.* **1980**, *77*, 6566–6570.

66. Minden, J.S.; Marians, K.J. Replication of pBR322 DNA *in vitro* with purified proteins. Requirement for topoisomerase I in the maintenance of template specificity. *J. Biol. Chem.* **1985**, *260*, 9316–9325.

67. Seufert, W.; Dobrinski, B.; Lurz, R.; Messer, W. Functionality of the *dnaA* protein binding site in DNA replication is orientation–dependent. *J. Biol. Chem.* **1988**, *263*, 2719–2723.

68. Seufert, W.; Messer, W. DnaA protein binding to the plasmid origin region can substitute for primosome assembly during replication of pBR322 *in vitro*. *Cell* **1987**, *48*, 73–78.

69. McCool, J.D.; Sandler, S.J. Effects of mutations involving cell division, recombination, and chromosome dimer resolution on a *priA*2::Kan mutant. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 8203–8210.

70. Bramhill, D.; Kornberg, A. Duplex opening by DnaA protein at novel sequences in initiation of replication at the origin of the *E. coli* chromosome. *Cell* **1988**, *52*, 743–755.

71. Marians, K.J. Pria–directed replication fork restart in *Escherichia coli*. *Trends Biochem. Sci.* **2000**, *25*, 185–189.

72. Zeghouf, M.; Li, J.; Butland, G.; Borkowska, A.; Canadien, V.; Richards, D.; Beattie, B.; Emili, A.; Greenblatt, J.F. Sequential peptide affinity (SPA) system for the identification of mammalian and bacterial protein complexes. *J. Proteome. Res.* **2004**, *3*, 463–468.

73. Reiser, W.; Leibrecht, I.; Klein, A. Structure and function of mutants in the *P* gene of bacteriophage lambda leading to the pi phenotype. *Mol. Gen. Gen.* **1983**, *192*, 430–435.

74. Yates, J.L.; Gette, W.R.; Furth, M.E.; Nomura, M. Effects of ribosomal mutations on the read–through of a chain termination signal: Studies on the synthesis of bacteriophage lambda *O* gene protein *in vitro*. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 689–693.

75. Tsurimoto, T.; Hase, T.; Matsubara, H.; Matsubara, K. Bacteriophage lambda initiators: Preparation from a strain that overproduces the O and P proteins. *Mol. Gen. Gen.* **1982**, *187*, 79–86.

76. Krinke, L.; Wulff, D.L. Rnase iii–dependent hydrolysis of lambda *cII–O* gene mRNA mediated by lambda OOP antisense RNA. *Genes Dev.* **1990**, *4*, 2223–2233.

77. Krinke, L.; Wulff, D.L. The cleavage specificity of Rnase III. *Nucleic Acids Res.* **1990**, *18*, 4809–4815.

78. Hayes, S.; Hayes, C. Spontaneous lambda *oR* mutations suppress inhibition of bacteriophage growth by nonimmune exclusion phenotype of defective lambda prophage. *J. Virol.* **1986**, *58*, 835–842.

79. Miwa, T.; Akaboshi, E.; Matsubara, K. Instability of bacteriophage lambda initiator O and P proteins in DNA replication. *J. Biochem.* **1983**, *94*, 331–338.

80. Wyatt, W.M.; Inokuchi, H. Stability of lambda O and P replication functions. *Virology* **1974**, *58*, 313–315.

250

81. Hayes, S. Mutations suppressing loss of replication control. Genetic analysis of bacteriophage lambda–dependent replicative killing, replication initiation, and mechanisms of mutagenesis. In *DNA replication and mutagenesis*, Moses, R.E., Summers, W.C., Ed. American Society for Microbiology: Washington, D.C., 1988; pp 367–377.

82. Hayes, S.; Duncan, D.; Hayes, C. Alcohol treatment of defective lambda lysogens is deletionogenic. *Mol. Gen. Gen.* **1990**, *222*, 17–24.

83. Hayes, S.; Hayes, C.; Bull, H.J.; Pelcher, L.A.; Slavcev, R.A. Acquired mutations in phage lambda genes *O* or *P* that enable constitutive expression of a cryptic lambda–*N*⁺*cI*[Ts]*cro*[−] prophage in *E. coli* cells shifted from 30 degreesC to 42 degreesC, accompanied by loss of *imm*lambda and Rex⁺ phenotypes and emergence of a non–immune exclusion–state. *Gene* **1998**, *223*, 115–128.

84. Kaplan, D.L.; O'Donnell, M. DnaB drives DNA branch migration and dislodges proteins while encircling two DNA strands. *Mol. Cell* **2002**, *10*, 647–657.

85. Horbay, M.A.; McCrea, R.P.E.; Hayes, S. OOP RNA: A regulatory pivot in temperate lambdoid phage development. In *Modern bacteriophage biology and biotechnology*, Wegrzyn, G., Ed. Research Signpost: Kerala, India, 2006; pp 37–57.

86. Walker, G.C. Mutagenesis and inducible responses to deoxyribonucleic acid damage in *Escherichia coli*. *Microbiol. Rev.* **1984**, *48*, 60–93.

87. Horii, T.; Ogawa, T.; Nakatani, T.; Hase, T.; Matsubara, H.; Ogawa, H. Regulation of SOS functions: Purification of *E. coli* LexA protein and determination of its specific site cleaved by the RecA protein. *Cell* **1981**, *27*, 515–522.

88. Markham, B.E.; Little, J.W.; Mount, D.W. Nucleotide sequence of the *lex*A gene of *Escherichia coli* K–12. *Nucleic Acids Res.* **1981**, *9*, 4149–4161.

89. Gottesman, S.; Halpern, E.; Trisler, P. Role of SulA and SulB in filamentation by *lon* mutants of *Escherichia coli* K–12. *J. Bacteriol.* **1981**, *148*, 265–273.

90. Schoemaker, J.M.; Gayda, R.C.; Markovitz, A. Regulation of cell division in *Escherichia coli*: SOS induction and cellular location of the SulA protein, a key to Lon–associated filamentation and death. *J. Bacteriol.* **1984**, *158*, 551–561.

91. Hill, T.M.; Sharma, B.; Valjavec–Gratian, M.; Smith, J. Sfi–independent filamentation in *Escherichia coli* is LexA dependent and requires DNA damage for induction. *J. Bacteriol.* **1997**, *179*, 1931–1939.

92. Hirota, Y.; Jacob, F.; Ryter, A.; Buttin, G.; Nakai, T. On the process of cellular division in *Escherichia coli*. I. Asymmetrical cell division and production of deoxyribonucleic acid–less bacteria. *J. Mol. Biol.* **1968**, *35*, 175–192.

93. Mulder, E.; Woldringh, C.L. Actively replicating nucleoids influence positioning of division sites in *Escherichia coli* filaments forming cells lacking DNA. *J. Bacteriol.* **1989**, *171*, 4303–4314.

94. Bukau, B.; Walker, G.C. Delta *dnaK*52 mutants of *Escherichia coli* have defects in chromosome segregation and plasmid maintenance at normal growth temperatures. *J. Bacteriol.* **1989**, *171*, 6030–6038.

95. Baharoglu, Z.; Lestini, R.; Duigou, S.; Michel, B. RNA polymerase mutations that facilitate replication progression in the *rep uvrD recF* mutant lacking two accessory replicative helicases. *Mol. Microbiol.* **2010**, *77*, 324–336.

96. Trautinger, B.W.; Lloyd, R.G. Modulation of DNA repair by mutations flanking the DNA channel through RNA polymerase. *EMBO J.* **2002**, *21*, 6944–6953.

97. Bachmann, B.J. Derivations and genotypes of some mutant derivatives of *Escherichia coli* K–12. In *Escherichia coli and Salmonella typhimurium: Cellular and molecular biology*, Neidhardt, F.C., Ingraham, J.I., Low, K.B., Magasanik, B., Schaechter, M., Umbargr, H.E., Ed. American Society for Microbiology: Washington, D.C., 1987; Vol. 2, pp 1192–1219.

98. Harris, R.S.; Longerich, S.; Rosenberg, S.M. Recombination in adaptive mutation. *Science* **1994**, *264*, 258–260.

99. Slavcev, R.A.; Hayes, S. Blocking the T4 lysis inhibition phenotype. *Gene* **2003**, *321*, 163–171.

100. Hayes, S.; Hayes, C. Control of lambda repressor prophage and establishment transcription by the product of gene *tof*. *Mol. Gen. Gen.* **1978**, *164*, 63–76.

101. Hayes, S. Control of the initiation of lambda replication, oop, lit and repressor establishment RNA synthesis. In *DNA synthesis, present and future*, Molineux, I.; Kohiyama, M., Eds. Plenum Press: New York, 1978; pp 127–142.

102. Grosschedl, R.; Schwarz, E. Nucleotide sequence of the *cro–cII–oop* region of bacteriophage 434 DNA. *Nucleic Acids Res.* **1979**, *6*, 867–881.

Article

Bacteriophage 434 Hex Protein Prevents RecA-Mediated Repressor Autocleavage

Paul Shkilnyj, Michael P. Colon and Gerald B. Koudelka *

Department of Biological Sciences, University at Buffalo, Buffalo, NY 14260, USA;
E-Mails: shkilnyj@gmail.com (P.S.); mpcolon@buffalo.edu (M.P.C.)

* Author to whom correspondence should be addressed; E-Mail: koudelka@buffalo.edu;
 Tel.: +1-716-645-4940; Fax: +1-716-645-2975.

Abstract: In a λ^{imm434} lysogen, two proteins are expressed from the integrated prophage. Both are encoded by the same mRNA whose transcription initiates at the P_{RM} promoter. One protein is the 434 repressor, needed for the establishment and maintenance of lysogeny. The other is Hex which is translated from an open reading frame that apparently partially overlaps the 434 repressor coding region. In the wild type host, disruption of the gene encoding Hex destabilizes λ^{imm434} lysogens. However, the *hex* mutation has no effect on lysogen stability in a *recA⁻* host. These observations suggest that Hex functions by modulating the ability of RecA to stimulate 434 repressor autocleavage. We tested this hypothesis by identifying and purifying Hex to determine if this protein inhibited RecA-stimulated autocleavage of 434 repressor *in vitro*. Our results show that *in vitro* a fragment of Hex prevents RecA-stimulated autocleavage of 434 repressor, as well as the repressors of the closely related phage P22. Surprisingly, Hex does not prevent RecA-stimulated autocleavage of phage lambda repressor, nor the *E. coli* LexA repressor.

Keywords: bacteriophage; RecA; lysogeny

1. Introduction

Upon infection of a host cell, the lambdoid phages choose between two developmental fates, opting either to replicate and lyse the cell, or to enter the latent, or lysogenic phase in which the phages' chromosome is integrated into that of the host and is replicated along with it [1,2]. Maintenance of the lysogenic state requires the activity of the phage's cI repressor protein. The repressor functions to maintain lysogeny by repressing transcription from promoters P_R and P_L, thereby preventing synthesis of proteins needed for phage lytic development. At the same time, the repressor activates transcription of its own gene by stimulating transcription from P_{RM} and thereby maintains appropriate repressor concentrations inside the cell [1].

The survival of the integrated prophage is linked to the fitness and survival of its host. Thus to ensure their survival, the lambdoid phages employ mechanisms that allow prophages to enter the lytic growth pathway and thus escape from hosts whose survival is in doubt. The best understood regulator of lambdoid prophage induction is the RecA protein [3–5]. As the master regulator of the SOS response, RecA inactivates LexA protein, the repressor of the SOS response genes, by stimulating its nascent autocleavage activity [5] thereby eliminating LexA protein's DNA binding activity. Similarly, RecA also stimulates autoproteolysis of the lambdoid phage repressors [6,7]. When the repressor is cleaved into its constituent carboxyl terminal and amino terminal domains, it is incapable of binding DNA with high affinity DNA binding or forming higher order oligomers (dimers, tetramers, *etc.*) Consequently transcription from P_R and P_L are no longer repressed, and the prophage enters the lytic growth pathway.

Transcription from P_{RM} requires the stimulatory activity of DNA bound repressor. Hence this promoter is only active in the presence of repressor, *i.e.*, in the lysogenic state. In addition to the repressor gene many lambdoid bacteriophages encode additional proteins on the transcript initiated from P_{RM}. Since these genes are not essential for lysogenic or lytic growth, they are considered accessory [8]. However in some cases they apparently provide an advantage to both the host and the phage. For example, the RexA and RexB proteins, which are produced by bacteriophage λ lysogens, excludes superinfection by bacteriophage T4rII⁻ [9]. In bacteriophage 933W a eukaryotic-like tyrosine kinase is cotranscribed with cI repressor [10] and its activity aborts infection of the 933W lysogen by a superinfecting HK97 phage [11]. Further examples of proteins that are expressed during lysogeny of temperate bacteriophages are found in the tripartite immunity system of phages P1 and P7 [12], and phage K139 [13]. These proteins also prevent superinfection by other phages.

In a lysogen, bacteriophage 434 expresses, in addition to cI repressor, a second protein Hex, which is encoded on the mRNA transcript initiated from P_{RM}. Susskind and Botstein demonstrated that *S. typhimurium* lysogenized with λ^{imm434} excluded growth of superinfecting phage P22 [14]. However, *S. Typhimurium* lysogenized with $\lambda^{imm434:Hex-}$ did not exclude the growth of superinfecting phage P22. The name Hex, therefore, stands for heterologous exclusion.

Since the heterologous phage exclusion functions of each of the proteins encoded along with repressor is apparently highly specific for particular phage(s), we wondered whether these proteins

may have a more general role in bacteriophage biology. Our results show that deletion of the open reading frame of Hex increases the spontaneous induction frequency of λ^{imm434} phage. We also show that in a RecA$^-$ strain of *E. coli*, the spontaneous induction of $\lambda^{imm434:hex-}$ and wild type λ^{imm434} are identically low. The results of genetic and biochemical experiments suggest that, when present, Hex increases the stability of bacteriophage 434 by interfering with RecA-mediated autocleavage of the phage 434 repressor. Consistent with this finding, we found that purified Hex blocks RecA-mediated stimulation of 434 repressor autocleavage. We found that Hex also prevents RecA stimulated autoproteolysis of the repressor encoded by bacteriophage P22, but not the repressor of bacteriophage λ or *E. coli* LexA. All of which are susceptible to RecA stimulated autoproteolysis. This finding suggests that Hex does not block RecA activity by directly interacting with it.

2. Results

Computer analysis of the three reading frames in the mRNA initiated from P$_{RM}$ of the bacteriophage 434 identified a a potential open reading frame with a capacity to code for a 158 amino acid protein located 3' to the 434 repressor gene (Figure 1). The predicted amino acid sequence of the identified ORF is shown in Figure 2, beginning with a methionine encoded by the 5'-most AUG of the transcript. To ensure that the transcript initiated at P$_{RM}$ extends through the putative ORF downstream from the repressor gene, we isolated total bacterial RNA from MG1655(λ^{imm434}) lysogens, reverse transcribed this RNA using primer 1 (Figure 1) and the resulting cDNA was analyzed by PCR using a series of primers complementary to this region (Figure 1). Since the RNA was isolated from a phage lysogen, RNA encoding the repressor must be present. As expected, a product of ~627 bp, the predicted size, was obtained when primer set complementary to this region (Primers 2 + 3) of the RNA was used (Figure 3, lane 1). When PCR was performed using primers 1 + 3, a DNA product of about 1100 base pairs was obtained (Figure 3, lane 2). This observation shows that in the MG1655::λ^{imm434} lysogen, transcripts initiated at P$_{RM}$ continue through the predicted Hex ORF.

Figure 1. Schematic representation of the immunity region of bacteriophage 434. The N- and C-terminal domains of the cI repressor dimers are represented by black and gray circles, respectively bound to the operator sites. In above configuration shown, transcription of the promoters P$_R$ and P$_L$ are repressed and transcription from P$_{RM}$ is activated on. The Hex gene is located on the same transcript initiated at P$_{RM}$. Primers designated 1, 2, and 3 were used to analyze the P$_{RM}$ transcript (see text).
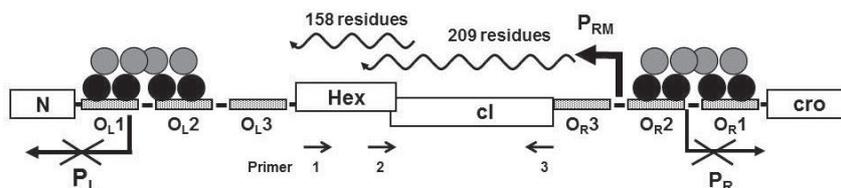
Figure 2. The putative amino acid sequence of Hex. The predicted amino acid sequence starts from the 5'-most AUG, in the hex transcript. Additional potential methionine initiators highlighted. The red amino acids highlight the beginning and end of the sHex polypeptide (see also Figure 5).

```
                10              20              30
M A Y D S Y Q R E L Q D Y R C C R G S E G K I R M I R I A A
                40              50              60
L L S I L L T T S A N S E C W I V T N L H G Y G A M N G D R
                70              80              90
Y E F T K D S T E D S V F N V T I N G D K S S V Y E S V S G
               100             110             120
V Y P E M K Y T A L S S N T M V G E Y Q S G G G I T V E T W
               130             140             150
S I T T D K K A L Y S K V M N I P G M Q Q L T S T K S F V G

D V V G T C N Q
```

Figure 3. RT-PCR analysis of total bacterial RNA isolated from MG1655::λ^{imm434} lysogens. Total RNA was reverse transcribed using Primer 1 (see Figure 1). The cDNA was PCR separately amplified with primers 2 + 3 (lane 1) or 1 + 3 (lane 2), and products separated on an agarose gel. The numbers represent sizes of molecular weight standards.
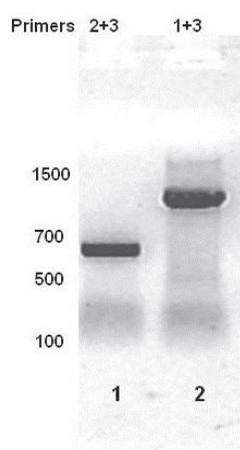
2.1. Effect of Hex Mutation on λ^{imm434} Prophage Induction

E. coli strain MG1655 was lysogenized with $\lambda^{imm434hex-}$ and the amount of phage produced spontaneously by *E. coli* strain MG1655 separated lysogenized with λ^{imm434} and $\lambda^{imm434hex-}$ was determined as described in the Experimental Section. We find that the MG1655::$\lambda^{imm434hex-}$ lysogens produced 30-fold more phage during a five hour incubation then did the wild type (MG1655::$\lambda^{imm434)}$ lysogens (Figure 4). This finding suggests that Hex may function in regulating the lysis-lysogeny of integrated λ^{imm434} prophage. Alternatively, the increased in number of phage produced during the incubation may be a result of an effect of *hex* on the number of phage released/cell. We tested this idea by measuring phage burst sizes for both of these bacteriophages (see Experimental Section). We find that the burst sizes of the λ^{imm434} and $\lambda^{imm434hex-}$ phages are indistinguishable from each other; under the conditions of our measurements, each cell infected by these phages produces 104 phage.

Figure 4. Effect of *hex* mutation on stability of MG1655::λ^{imm434} lysogens. MG1655(λ^{imm434}) and MG1655($\lambda^{imm434hex-}$) were grown to saturation, washed and resuspended in fresh media for 4 hours. The amount of phage release into the culture was determined by plating various amounts of supernatant on MG1655 cells that did not contain λ^{imm434} prophage. The concentration of lysogen cells was determined by plating various dilutions of cells on LB agar plates. Error bars are standard deviation of twenty four separate repeats.
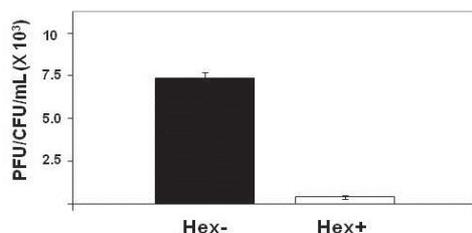
2.2. Role of RecA and Hex in Regulating Prophage Stability

Since RecA affects the level of spontaneous phage induction in a RecA$^+$ host [15], we tested the idea that Hex may affect spontaneous induction by influencing RecA's co-protease anti-repressor activity. To test this idea, we compared the spontaneous induction frequency of λ^{imm434} and $\lambda^{imm434hex-}$ prophages in wild type and RecA$^-$ MG1655 strains. The spontaneous induction frequency of MG1655recA938::cat ($\lambda^{imm434hex-}$) is $1.8 \pm 0.08 \times 10^{-5}$, which is nearly identical to the spontaneous induction rate of MG1655recA938::cat::(λ^{imm434}) (Table 1). In contrast, the spontaneous induction frequency of $\lambda^{imm434hex-}$ and λ^{imm434} in wild type MG1655 are remarkably different. In the wild type host, the amount of induction of MG1655($\lambda^{imm434hex-}$) is nearly 30-fold higher than the induction frequency of MG1655(λ^{imm434}) (Table 1) Although the frequency of spontaneous induction in a RecA$^-$ host is much lower than in the wild type MG1655 host, the nearly identical frequency of induction in RecA$^-$ host of both $\lambda^{imm434hex-}$ and wild type λ^{imm434} suggests that Hex increases the stability of λ^{imm434} by interfering with RecA mediated autocleavage of the cI repressor.

Table 1. Effect of *hex* Deletion on Spontaneous Induction of RecA$^-$ MG1655(λ^{imm434}) Lysogens. The values (+/− standard deviation) were determined as described in the Experimental Section. Data is an average based on twenty four separate repeats.

Strain	PFU/mL/viable cell (X1000)		
Wild-type	0.275 ± 0.06		
*rec*A-	$1.8 \pm 0.08 \times 10^{-5}$		
hex-	7.34 ± 0.1		
*rec*A-/*hex*-	$1.8 \pm 0.23 \times 10^{-5}$		

Figure 5. Expression and purification of Hex polypeptide. Full length *hex* gene sequence was inserted into plasmid pMAL-c2X to generate a fusion protein between the genes encoding MBP and Hex. *E. coli* extracts containing the fusion peptide were passed through an amylase column and the eluate digested by Factor Xa (see Experimental Section). Shown is a Coomassie stain of the products of the Factor Xa digest of the eluate from amylase column, fractionated by SDS-PAGE (left panel). Bacterial strain BL21(DE3)[pLysS] was transformed with plasmid psHex containing the gene coding for the identified sHex (see Figure 2). The sHex was purified from cell extracts using ammonium sulfate, GuHCl, followed by size exclusion chromatography. The right panel shows Coomassie stain of an SDS-PAGE of the the pooled protein-containing fractions from the sizing column. The positions of sHex, maltose binding protein (MBP) and molecular weight standards are indicated.
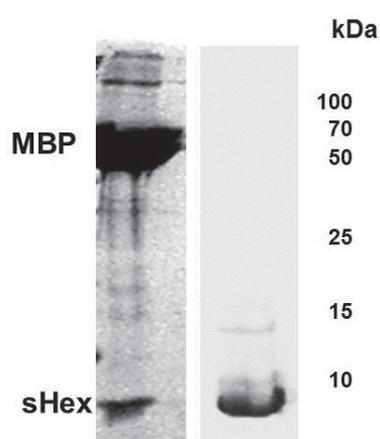
2.3. Expression and Purification of Hex

We placed the *hex* gene under control of the T7 RNA polymerase promoter and attempted to express this full length gene product. We repeatedly observed that upon induction of Hex expression, cultures bearing plasmids encoding full-length Hex stopped growing, and subsequently showed a decrease in optical density at 600 nm. Therefore we initially attempted to purify Hex as a fusion with maltose binding protein (MBP). We reasoned that as a fusion with MBP (Maltose Binding Protein), the Hex peptide may be less detrimental to cell viability. When a plasmid encoding a MBP-Hex fusion protein was transformed into *E. coli* strain BL21(DE3)[pLysS], over-expression of the MBP-Hex protein was induced upon addition of IPTG. The expressed peptide was purified on an amylose resin and fractions containing the MPB-Hex were pooled and digested with factor Xa to separate the two proteins and then fractioned on SDS-PAGE (Figure 5). Sequence analysis shows that the MBP-Hex fusion protein contains a single Factor Xa target sequence, located at the fusion site between MBP and Hex. Thus we would have anticipated this treatment would have released a ~17.5 kDa protein-the size of full-length Hex. Instead, treatment of the MBP-Hex protein with Factor Xa produced two fragments, one the approximate size of MBP, and the other roughly 6 kDa (Figure 5). Repeated

attempts to generate the full Hex (MW ~16 kDa) from the MBP-Hex fusion by cleaving the fused peptide with the Factor Xa under various conditions, or sources of Factor Xa failed. We are unsure as to why we are unable to obtain full length Hex by Factor Xa treatment. However we hypothesized that this 6 kDa protein fragment represented a stable fragment of Hex and proceeded to analyze this Hex fragment.

2.4. Characterization of sHex

To determine whether the 6 kDa protein fragment that results from factor Xa cleavage of the MBP-Hex fusion is a fragment of Hex, we cut this fragment out of a 15% SDS-polyacrylamide Tris-Tricine gel, and determined its sequence by MALDI-TOF mass spectrometry. The resulting sequence confirmed this peptide derives from full length Hex. The shortened Hex polypeptide (hereafter referred to as sHex) sequence begins with glutamic acid 43 and ends with asparagine 103 (see Figure 1).

Figure 6. Effect of sHex expression on the stability of MG1655::$\lambda^{imm434\ hex-}$ lysogens in MG1655. MG1655::$\lambda^{imm434\ hex-}$ lysogens were transformed with a control plasmid (black bars) or a plasmid encoding the sHex only (hatched bar). For comparison, the amount of spontaneously released from MG1655(λ^{imm434}) lysogens (denoted as WT) under these conditions is also shown (white bar). The amount of phage spontaneously released was determined as described in the Experimental Section. Error bars are standard deviation of twenty four separate experiments.
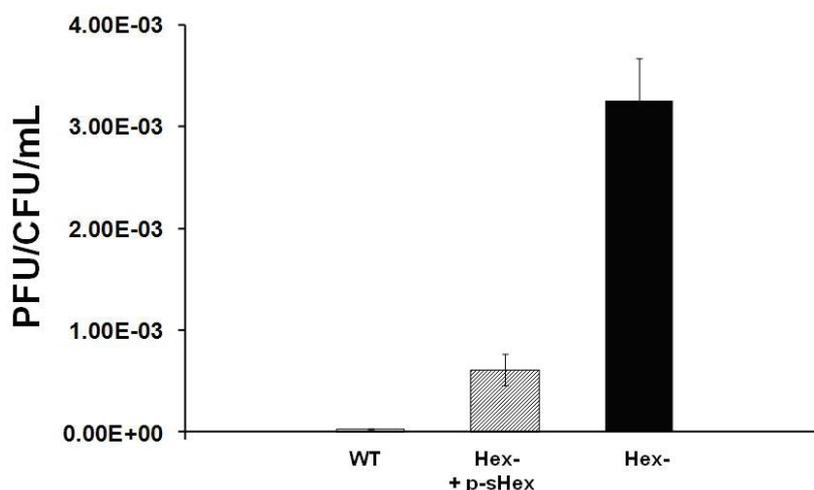
We determined whether sHex is capable of complementing the effects of Hex deletion. To accomplish this, we amplified DNA encoding the sHex sequence and placed expression of sHex gene under the control of the T7 RNA polymerase promoter in pET17b [16], creating the plasmid p-sHex. The sequence of this protein is as given in Figure 2, except it contains an initiator methonine precedingE43. MG1655($\lambda^{imm434hex-}$) lysogens were transformed with p-sHex together with pGP1-2

which encodes a T7 RNA polymerase under the control of the temperature sensitive *cI857* mutant λ repressor [16]. As a control, MG1655(λ$^{imm434hex-}$) cells were transformed with pGP1-2, and pET17b lacking the sHex insert. Cultures of these cells were grown to saturation overnight at 32 °C, washed, and resuspended in fresh medium and incubated at 37 °C for four hours. After 4 hours, the spontaneous induction frequencies of λ$^{imm434hex-}$ prophages in MG1655 with or without the sHex encoding plasmid were determined as described (Experimental Section). Expression of sHex reduces the spontaneous induction frequency of Hex$^-$ lysogens by ~7-fold compared to cells transformed with control plasmid (Figure 6). This finding indicates that sHex can act in *trans* to partially complement the *hex-* defect in λ$^{imm434hex-}$. This observation suggests that sHex constitutes the 'active' fragment of Hex.

2.5. sHex Inhibits RecA-Mediated Autocleavage of 434 Repressor *in Vitro*

The results in Table 1 suggest that Hex may alter the spontaneous induction frequency of λ^{imm434} prophages by interfering with RecA mediated autocleavage of the 434 repressor, thereby reducing the amount bacteriophage 434 produced. To test this hypothesis we purified sHex as described in the Experimental Section and examined the effect of this protein on RecA-mediated autocleavage of 434 repressor. Mixing sHex (5 μM) with 300 nM 434 repressor does not result in the formation of any 434 repressor cleavage products (Figure 7A, lane 2), whereas mixing 434 repressor with RecA (see Experimental Section) causes the formation of two lower molecular weight antibody reactive species (Figure 7A, lane 3). The molecular weights of these products correspond to the 434 repressor N- and C-terminal domain fragments, consistent with the previous observation that RecA stimulates 434 repressor autocleavage [17]. When increasing concentrations of sHex are added into mixtures of 434 repressor and RecA, the amount of repressor cleavage products observed progressively decreases (Figure 7A, lanes 5–8). In an identical set of experiments in which increasing concentrations of BSA instead of sHex were added into the mixtures of 434 repressor and RecA, BSA did not interfere with the reaction (Figure 7B) Therefore the findings in Figure 7 indicate that sHex interferes with the ability of RecA to catalyze autoproteolysis of 434 repressor *in vitro*.

2.6. Effect of sHex on RecA-Mediated Autocleavage of Related Repressors P22, λ, and the Bacterial LexA Repressor

To determine whether or not the sHex inhibition of RecA stimulated autoproteolysis is specific to the 434 cI repressor, we examined the effect of sHex on RecA stimulated autocleavage of related cI repressors from lambdoid phages P22 and λ, as well as the SOS repressor LexA, which also undergoes RecA stimulated autocleavage. In an identical procedure as described above, we examined the effect of adding sHex on RecA stimulated autocleavage of P22 repressor (Figure 8A), λ repressor (Figure 8B) and LexA repressor (Figure 8C). Similar to the case with 434 repressor, added sHex interferes with the RecA stimulated autocleavage of P22 repressor (compared Figure 8A with Figure 7A). In contrast, added sHex does not inhibit RecA mediated autocleavage of the λ repressor (Figure 8B) under these

conditions. Similar to the λ repressor case, sHex is also incapable of blocking RecA stimulated autocleavage of the SOS repressor LexA (Figure 8C) under these conditions.

Figure 7. Effect of Hex on RecA-stimulated autocleavage of 434 Repressor. (**A**) 434 repressor [300 nM] (lane1) was incubated with sHex (lane2), 8 µM RecA (lanes 3) or 8 µM RecA plus the indicated sHex concentrations (lanes 5–8). (**B**) In identical procedure as in (A), except BSA was substituted for sHex. The figure is a composite immunoblot of the reaction products reacted with anti-434R antibodies and visualized by chemiluminescence.
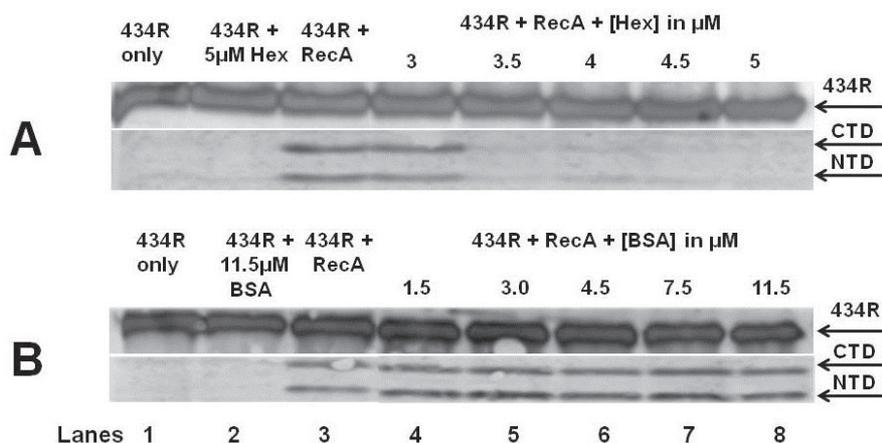
Figure 8. Effect of Hex on RecA-stimulated autocleavage of bacteriophage and bacterial repressors. Bacteriophage repressors from phage P22 (**A**), λ (**B**), and the bacterial SOS repressor LexA (**C**) were incubated in the absence (lane1) or presence of indicated concentrations of sHex and/or 5 µM RecA. The figure is a composite immunoblot of the reaction products reacted with anti-protein antibodies and visualized by chemiluminescence.
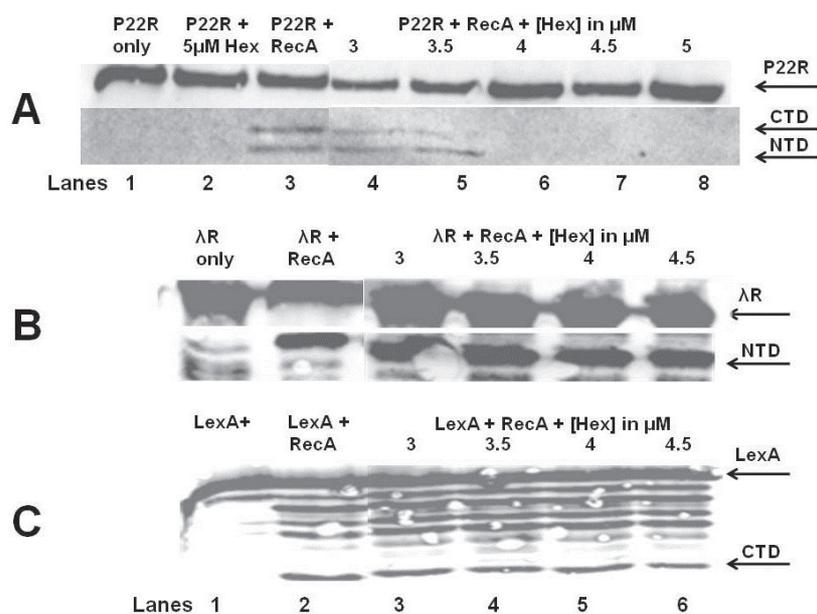
3. Discussion

Our results show that the Hex protein of bacteriophage 434 reduces spontaneous induction of the 434 prophage. Based on the results of *in vitro* biochemical data, Hex apparently reduces prophage spontaneous induction by inhibiting RecA mediated autoproteolysis of the cI repressor. The *hex* gene is cotranscribed along with that of the cI repressor gene on mRNA that is initiated at the promoter for repressor maintenance (P_{RM}) in a *P_{RM}-cI-Hex* operon (Figures 1 and 3). Many temperate prophages produce accessory proteins that benefit the phage, the *E. coli* host, or both, from the same message as their repressor. The most common advantage conferred by many accessory proteins is exclusion of other superinfecting phages [9,12,13,18]. Our data show that the Hex protein of bacteriophage 434, in addition to conferring immunity to phage P22 when expressed in *S. typhimurium*, also reduces induction of λ^{imm434} by interfering with RecA stimulated autocleavage of phage 434 repressor. When the RecA protein is not present in *E. coli* host, $\lambda^{imm434Hex-}$ induces with same frequency as the wild type phage (Table 1), however in a wild type *E. coli* host, $\lambda^{imm434Hex-}$ induction is 30 fold higher. Producing sHex from a plasmid in $\lambda^{imm434hex-}$ lysogens reduces the spontaneous induction about 7 fold (Figure 6). These results suggest that Hex is an advantageous adaptation to bacteriophage 434 in that it "fine tunes" the phage lysis-lysogeny decision.

The host SOS response to DNA damage causes the lysogenic phage to enter lytic growth; once this cycle initiated, it is irreversible. Under normal conditions, (aerated rich media and lack of DNA damaging agents) SOS genes are expressed at a very low basal level [12]. Furthermore, under normal conditions there is enough repressor made to reduce the transcription of P_R roughly 1,000 fold [1]. However, since phage 434 cI repressor's cooperatively binds DNA, a 5 fold drop in repressor concentration causes repressor to only partially occupy O_R1 and O_R1, and thus increase the activity of P_R by 50% [19]. The small decrease in repressor concentration, combined with a small amount of active RecA, which may lead to further autocleavage of the repressor. This situation could result in unwanted spontaneous induction. Bacteriophage 434 may use Hex to prevent this "false alarm" induction.

The *in vitro* autocleavage experiments shown in Figures 7 and 8 demonstrate that sHex blocks RecA mediated autocleavage of the phage 434 repressor and P22 repressor, but not λ repressor or *E. coli* LexA repressor (a protein central in the SOS response). Although the 434 repressor, P22 repressor, λ repressor and LexA are structurally similar, they exhibit differences in their mechanisms of RecA-mediated autocleavage. For example, 434 repressor undergoes autocleavage most efficiently as a DNA-bound dimer [17], as opposed to lambda and LexA repressors which can be efficiently autocleaved as monomers [20,21]. P22 repressor autocleavage increases with increase repressor concentrations, which suggests that P22 repressor is a better substrate for RecA-catalyzed inactivation as a dimer [4]. The different mechanisms of intra or intermolecular cleavage are also prevalent in other RecA targets such as UmuD [22]. These differences in RecA mediated autocleavage of various repressors may account for the selectivity of Hex for the repressors of phages 434 and P22.

RecA function can be inhibited by competing for DNA binding by DNA binding proteins, filament formation, or preventing RecA from hydrolyzing ATP [23]. For example, RecA dependent proteolysis

of P22 repressor is blocked when P22 antirepressor-repressor complex is formed, which prevents access of RecA to the P22 repressor [24]. Phage P22 *virA* produces the antirepressor protein [24]. The specific manner in which Hex prevents RecA mediated autocleavage of 434 and P22 repressors suggests that Hex does not function by directly competing with RecA for DNA, does not prevent ATP hydrolysis by RecA, nor does it directly interact with RecA active site for repressor interaction. Rather, Hex may interact with 434 repressor, and its close homolog the P22 repressor (and with repressor-DNA complexes), thereby preventing RecA mediated autocleavage. Regardless of the precise mechanism, it is clear that Hex reduces RecA-mediated repressor autocleavage. We further speculate that by being selective only to phages 434 & P22 repressors, Hex does not interfere with bacterial SOS repair. A number of DNA repair cycles may be necessary throughout the bacterial life cycles in order to repair minor DNA damage. These repairs will require a small amount of SOS gene products, active RecA being one of them. However the phage must be able to distinguish life threatening DNA damage from more minor effects, and Hex may function in this capacity.

Outside of the biological function of Hex, for puzzling, and as of yet unidentified reasons, overexpression of full-length Hex is lethal to cells. We were also unable to isolate the MBP-Hex fusion. Treatment of the fusion protein with factor Xa, surprisingly, yielded a much shorter peptide then the full length Hex. Neither long nor short Hex contain factor Xa cleavage sites, nor are there any basic residues present at the ends of the sHex which sometimes can be cleaved by factor Xa. However a presence of cryptic Xa sites cannot be excluded. The possibility that full length Hex is processed to sHex by the activity of residual bacterial proteases unrelated to factor Xa present during the purification process also cannot be excluded. A more intriguing possibility is that full length Hex has an intrinsic autoprotease activity, and undergoes spontaneous proteolytic cleavage. Regardless of the precise mechanism of the full length Hex modification, we conclude that we have identified the active Hex peptide, and demonstrated its role in bacteriophage 434 lysis-lysogeny modulation.

4. Experimental Section

4.1. Media and Growth Conditions

Liquid cell cultures were grown in Luria broth or M9 minimal media [25], each of which were supplemented with 100 μg/mL ampicillin and 30 μg/mL kanamycin, where needed, to select for the presence of plasmids.

4.2. Bacterial Strains, Phages, and Plasmids

The host strain used for all spontaneous induction analysis was MG1655 [26,27]. The MG1655 *recA* mutants were created by P1 transduction using a lysate from *E. coli* GW4212, which bears the *recA938::cat* allele [28] (a gift from Mark Sutton, University at Buffalo, Buffalo, NY, USA). Both wild type and the *recA* mutant MG1655 were lysogenized with λ^{imm434} or λ$^{imm434: hex-}$ as previously described [8]. The *hex* mutant version of λ^{imm434} was a gift from Lynn Thomason, Center for Cancer Research, National Cancer Institute at Frederick, Frederick, MD, USA. This phage was generated

completely replacing the Hex coding sequence with that encoding the gene for chloramphenicol acyltransferase flanked by its promoter and transcription terminator. The replacement was done by recombineering.

The Hex DNA sequence ATGGCCTATGACTCCTATCAACGGGAACTGCAAGATTATCG GTGTTGTCGTGGAAGCGAGGGTAAAATTCGTATGATCAGGATTGCGGCGCTACTCTCAAT ACTCTTAACTACCAGCGCCAATTCTGAATGCTGGATTGTCACAAACCTGCACGGGTACGG GGCAATGAATGGCGATCGTTACGAGTTTACAAAAGACAGCACGGAAGATTCCGTTTTCAA CGTAACAATAAATGGCGATAAATCATCAGTTTATGAATCAGTTTCTGGCGTCTATCCAGA GATGAAATACACTGCTTTGTCATCGAACACTATGGTAGGAGAATACCAGTCTGGAGGAGG CATAACCGTTGAAACTTGGTCAATCACTACAGACAAAAAAGCTCTTTACTCCAAAGTAAT GAACATCCCAGGTATGCAACAACTTACATCAACCAAATCCTTTGTTGGTGATGTAGTCGG AACCTGCAATCAGTAAT was PCR amplified and cloned into XbaI and HindIII sites of pMAL-c2X (New England Biolabs), generating a MBP-Hex fusion. This construct encodes a Factor Xa cleavage site in the linker between Hex and MBP. Expression of this fusion protein is under control of the IPTG (isopropyl-β-$_D$-thiogalactopyranoside) inducible *lac* promoter. A truncated version of Hex (short hex—sHex) was created by PCR amplifying the following sequence from the central portion of the hex gene: GAATGCTGGATTGTCACAAACCTGCACGGGTACGGGGCAATGAATGGCGATCGTTACGA GTTTACAAAAGACAGCACGGAAGATTCCGTTTTCAACGTAACAATAAATGGCGATAAATC ATCAGTTTATGAATCAGTTTCTGGCGTCTATCCAGAGATGAAATACACTGCTTTGTCATCG AAC using forward primer 5'-GGCCATCATATGGAATGCTGGATTGTC-3', and a reverse primer 5'-CCAGCGCTCGAGTTAGTTCGATGACAAAGC-3' (IDT Technologies). This DNA fragment was cleaved with NdeI and XhoI (New England Biolabs), purified and ligated into pET17b (Novagen) which was cleaved with the same enzymes.

4.3. Analysis of RNA Originating from P_{RM}

RNA was isolated from MG1655$^{\lambda imm434}$ as described in reference [29]. Briefly, cell cultures were grown to OD$_{600}$ = 0.6, collected by centrifugation at 12,000 × *g* at 4 °C, and resuspended in protoplasting buffer (15 mM Tris-HCl pH 8.0, 0.45 M Sucrose, 8 mM EDTA) containing 50 mg/mL lysozyme, and incubated 15 minutes on ice. Protoplasts were collected by centrifugation for 5min at 5,900 × *g* 4 °C, resuspended in 0.5 mL gram-negative lysing buffer containing DEPC, and incubated for 5 min at 37 °C. The RNA was precipitated from the preparation by adding saturated NaCl (34 g/100 mL of water) and 100% ethanol, in an overnight incubation at −20 °C. The RNA was collected by centrifugation and washed with ice-cold 70% ethanol, dried, and dissolved in DEPC-treated water. The RNA was digested for 10 minutes with RNAse-free DNase I (Promega Madison, WI, USA) which was subsequently deactivated at 65 °C for 10 minutes. The RNA was immediately used in RT-PCR reactions. Synthesis of first-strand cDNA was performed using AffinityScript^{TM} QPCR cDNA Synthesis Kit (Qiagen Valencia, CA, USA). The primer 5'-CGAGATTGAGGTGGGGATTAC-3' was incubated with 3 μg total RNA, first strand master mix,

AffinityScript RT/RNase block enzyme mixture, and incubated at 25 °C for 5 minutes, followed by 15 minutes in 42 °C. The reaction was terminated by placing the mixture at 95 °C for 5 minutes, chilled on ice, and immediately used in PCR with the following primer combinations: 5'-CGCGTCCAGTATCTACTAACA-3' with 5'-CAGCTTGGACTTAACCAGGCT-3', or 5'-CGAGATTGAGGTGGGGATTAC-3' with 5'-CAGCTTGGACTTAACCAGGCT-3'. The PCR products were fractionated by electrophoresis on agarose gels.

4.4. Analysis of Spontaneous Induction

Measurements of spontaneous induction frequency were performed as described previously [30]. Briefly, cultures of MG1655 lysogenized with λ^{imm434} or $\lambda^{imm434\ hex-}$ were grown to saturation overnight in LB or M9 minimal media at 37 °C. Phage that spontaneously produced overnight were removed from the stationary cells by washing three times with fresh medium. During the last wash the cells were resuspended in a volume of fresh medium that was equal to the starting culture volume. The cultures were incubated with shaking at 37 °C for 5 hours. Cell cultures were then centrifuged at 8,000 × g, and the phage-containing supernatant was sterilized by addition of chloroform. To determine the number of CFU (colony forming units), an aliquot of the remaining phage culture was evenly spread over Luria broth (LB) solidified with Luria agar (LA) in a Petri dish. The amount of phage released into the supernatant was determined by evenly distributing the released phage on nonlysogenic MG1655 as described previously [8]. The amount of phage released per cell in culture was calculated as the plaque forming units PFU/CFU ratio.

4.5. Measurement of Burst Size

Standard methodologies were used for determining the mean number of phage particles per bacterium [31,32] Briefly, phage particles at a concentration of 1×10^7 pfu mL^{-1} were mixed with 1×10^8 cfu mL^{-1} MG1655 in LB supplemented with 10 mM MgSO$_4$ and 5 mM CaCl$_2$ for 5 minutes at 37 °C. The mixture was diluted 10,000-fold with pre-warmed LB to give a final culture volume of 10 mL. At intervals up to 75 minutes, two 0.2 mL samples were acquired to determine amount of unadsorbed (U) phage particles and total (T) phage particles. The U phage particles were additionally treated with 2% CHCl$_3$. The burst size is calculated as $b = \frac{F}{(T-U)}$ where F is the final phage count at time point 75 minutes and T and U are total and unadsorbed phage, respectively.

4.6. Analysis of Complementation of the Hex Mutation

The MG1655 $\lambda^{imm434\ hex-}$ strain was transformed with pET17b bearing the short Hex gene under the T7 RNA polymerase promoter, together with plasmid pGP1-2 which contains temperature inducible T7 RNA polymerase [16]. Control MG1655 $\lambda^{imm434\ hex-}$ lysogens were transformed with plasmids pGP1-2 and pET17b that does not express any Hex derivatives. The cultures were grown overnight at 32 °C, and the cells collected by centrifugation as described above, washed three times and resuspended in a volume that was equal to the starting culture volume, and incubated at 37 °C for three

hours. Cell cultures were then centrifuged at 8,000 × *g* and the phage containing supernatant was sterilized by addition of chloroform. To determine the number of CFU, an aliquot of the remaining phage culture was evenly spread on LB/LA Petri dish as described above. The amount of phage released into the supernatant was determined by infecting nonlysogenic MG1655 as described previously [8]. The amount of phage released per cell in culture was calculated as the PFU/CFU ratio.

4.7. RecA-Mediated in Vitro Autoproteolysis

The standard buffer used in all assays contained 50 mM KCl, 15 mM Tris (pH 7.5), 2 mM MgCl$_2$, 0.1 mM EDTA, 2 mM dithiothreitol. Concentrations of repressors from phage 434, λ, P22, and LexA in each reaction was 300 nm. λ repressor and anti-lambda repressor antibodies were a gift from Ann Hochschild (Harvard University Medical School, Cambridge MA, USA). LexA and anti-LexA antibodies were a gift from John Little, (University of Arizona, Tuscon, AZ, USA). Reaction mixtures were incubated 10 minutes at room temperature with different repressors, 0.5 mM γ-S-ATP, 2 mM DTT, and 3 μM oligo(dT$_{20}$). After incubation, RecA protein was added to a final concentration of 8 μM, and the mixture was then incubated for 2 hours at 37 °C. Following incubation, Tris/Tricine loading dye (0.1 M Tris, pH 6.8, 1% sodium dodecyl sulfate (SDS), 20 mM 2-mercaptoethanol, 20% glycerol, 0.01% bromophenol blue) were added, and the samples were boiled for 5 min. The products were fractionated by electrophoresis on 15% SDS-polyacrylamide Tris-Tricine gels [33]. The repressor and its cleaved products were visualized by chemiluminescent detection of Western blots by a Storm Imager using an ECL-Plus kit (both obtained from GE Life Sciences, Piscataway, NJ, USA) with horseradish peroxidase secondary antibody.

4.8. Expression and Purification of Hex and sHex

A saturated, overnight culture of BL21(DE3)::pLysS with pMAL-c2X (New England Biolabs, Beverly, MA, USA), bearing the full Hex sequence, was diluted 1:50 in three liters of LB broth supplemented with 100 μg of ampicillin/mL and 20 μg of chloramphenicol/ml. The culture was grown for 2 hours at 37 °C. Cultures were induced to produce Hex by the addition of 0.5 mM IPTG. After additional growth for 4 hours at 37 °C, the cells were harvested by centrifugation at 10,000 × *g* for 10 minutes, and the cell pellet was resuspended in 30 mL of lysis buffer (100 mM Tris [pH 7.5], 200 mM NaCl, and 10 mM EDTA). All subsequent procedures were performed at 4 °C. The cells were lysed by using a French press, and the cell lysate was diluted to 100 mL with lysis buffer. The lysate was centrifuged at 10,500 × g for 20 minutes to remove cellular debris. The resulting supernatant was passed through a 2 mL Amylose resin column, washed three times with 5 bed volumes of resin buffer (100 mM Tris [pH 7.5], 500 mM NaCl), and eluted with resin buffer plus 10mM maltose in 500 μL fractions. The purification progress was monitored by analyzing fractions collected at each step by SDS-PAGE. MBP-Hex fractions were pooled and digested with factor Xa (Qiagen, Valencia, CA, USA) as specified by this manufacturer. Expected size of Hex peptide is ~17.5 kDa. The products were analyzed on 15% SDS-polyacrylamide Tris-Tricine gels. Unexpectedly, no full length Hex

protein was observed. Instead, a ~6 kDa product was observed to accumulate. This product was cut out from the gel and its amino acid sequence was determined using MALDI-TOF MS analysis. Subsequently, DNA encoding this sHex (short Hex) was PCR amplified, cloned and expressed as described above. Purification of sHex was performed as described for the MBP-Hex fusion protein, except after breaking the cells and centrifuging the lysate at $10,500 \times g$ for 20 minutes to remove cellular debris, sHex was precipitated from solution with addition of ammonium sulfate to the final concentration of 10%, and harvested by centrifugation at $10,000 \times g$ for 10 minutes. The insoluble pellet was dissolved in final concentration of 4 M Guanidine-HCl (GuHCl) in 100 mM Tris [pH 7.5], 100 mM NaCl, and passed through 20 cm Sephadex G50 column equilibrated with the same buffer. The sHex containing fractions were pooled, concentrated, and dialyzed in buffers of successively lower concentrations of GuHCl in order to remove GuHCl. Purity of sHex was monitored by SDS-PAGE.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Ptashne, M. *A Genetic Switch*; Blackwell PRESS: Palo Alto, CA, USA, 1986.
2. Pirrotta, V. Operators and promoters in the OR region of phage 434. *Nucleic Acids Res.* **1979**, *6*, 1495–1508.
3. Janion, C. Inducible SOS response system of DNA repair and mutagenesis in *Escherichia coli*. *Int. J. Biol. Sci.* **2008**, *4*, 338–344.
4. Phizicky, E.M.; Roberts, J.W. Induction of SOS functions: Regulation of proteolytic activity of *E. coli* RecA protein by interaction with DNA and nucleoside triphosphate. *Cell* **1981**, *25*, 259–267.
5. Butala, M.; Zgur-Bertok, D.; Busby, S.J. The bacterial LexA transcriptional repressor. *Cell. Mol. Life Sci.* **2009**, *66*, 82–93.
6. Phizicky, E.M.; Roberts, J.W. Kinetics of RecA protein-directed inactivation of repressors of phage lambda and phage P22. *J. Mol. Biol.* **1980**, *139*, 319–328.
7. Sauer, R.T.; Nelson, H.C.; Hehir, K.; Hecht, M.H.; Gimble, F.S.; DeAnda, J.; Poteete, A.R. The lambda and P22 phage repressors. *J. Biomol. Struct. Dyn.* **1983**, *1*, 1011–1022.

8. Arber, W.; Enquist, L.; Hohn, B.; Murray, N.E.; Murray, K.; Hendrix, R.W.; Roberts, J.W.; Stahl, F.W.; Weisberg, R.A. *Lambda II*; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY, USA, 1983; pp. 433–466.

9. Slavcev, R.A.; Hayes, S. Stationary phase-like properties of the bacteriophage lambda Rex exclusion phenotype. *Mol. Genet. Genom.* **2003**, *269*, 40–48.

10. Tyler, J.S.; Mills, M.J.; Friedman, D.I. The operator and early promoter region of the Shiga toxin type 2-encoding bacteriophage 933W and control of toxin expression. *J. Bacteriol.* **2004**, *186*, 7670–7679.

11. Friedman, D.I.; Mozola, C.C.; Beeri, K.; Ko, C.-C.; Reynolds, J.L. Activation of a prophage-encoded tyrosine kinase by a heterologous infecting phage results in a self-inflicted abortive infection. *Mol. Microbiol.* **2011**, *82*, 567–577.

12. Heinrich, J.; Velleman, M.; Schuster, H. The tripartite immunity system of phages P1 and P7. *FEMS Microbiol. Rev.* **1995**, *17*, 121–126.

13. Nesper, J.; Blass, J.; Fountoulakis, M.; Reidl, J. Characterization of the major control region of Vibrio cholerae bacteriophage K139: Immunity, exclusion, and integration. *J. Bacteriol.* **1999**, *181*, 2902–2913.

14. Susskind, M.M.; Botstein, D. Superinfection exclusion by lambda prophage in lysogens of Salmonella typhimurium. *Virology* **1980**, *100*, 212–216.

15. Livny, J.; Friedman, D.I. Characterizing spontaneous induction of Stx encoding phages using a selectable reporter system. *Mol. Microbiol.* **2004**, *51*, 1691–1704.

16. Tabor, S.; Richardson, C.C. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 1074–1078.

17. Pawlowski, D.R.; Koudelka, G.B. The preferred substrate for RecA-mediated cleavage of bacteriophage 434 repressor is the DNA-bound dimer. *J. Bacteriol.* **2004**, *186*, 1–7.

18. Kliem, M.; Dreiseikelmann, B. The superimmunity gene sim of bacteriophage P1 causes superinfection exclusion. *Virology* **1989**, *171*, 350–355.

19. Bushman, F.D. The bacteriophage 434 right operator. Roles of O R 1, O R 2 and O R 3. *J. Mol. Biol.* **1993**, *230*, 28–40.

20. Little, J.W. Autodigestion of lexA and phage lambda repressors. *Proc. Natl. Acad. Sci. USA* **1984**, *81*, 1375–1379.

21. Slilaty, S.N.; Rupley, J.A.; Little, J.W. Intramolecular cleavage of LexA and phage lambda repressors: Dependence of kinetics on repressor concentration, pH, temperature, and solvent. *Biochemistry* **1986**, *25*, 6866–6875.

22. McDonald, J.P.; Frank, E.G.; Levine, A.S.; Woodgate, R. Intermolecular cleavage by UmuD-like mutagenesis proteins. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 1478–1483.

23. Cox, M.M. Regulation of bacterial RecA protein function. *Crit. Rev. Biochem. Mol. Biol.* **2007**, *42*, 41–63.

24. Klinge, S.; Voigts-Hoffmann, F.; Leibundgut, M.; Arpagaus, S.; Ban, N. Crystal structure of the eukaryotic 60S ribosomal subunit in complex with initiation factor 6. *Science* **2011**, *334*, 941–948.

25. Sambrook, J.; Fritsch, E.F.; Maniatis, T. *Molecular Cloning — A Laboratory Manual*; Cold Spring Harbor Press: Cold Spring Harbor, NY, USA, 1989.

26. Guyer, M.S.; Reed, R.E.; Steitz, T.; Low, K.B. Identification of a sex-factor-affinity site in *E. coli* as gamma delta. *Cold Spr. Harb. Symp. Quant. Biol.* **1981**, *45*, 135–140.

27. Blattner, F.R.; Plunkett, G., III; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Collado-Vides, J.; Glasner, J.D.; Rode, C.K.; Mayhew, G.F.; *et al.* The complete genome sequence of *Escherichia coli* K-12. *Science* **1997**, *277*, 1453–1474.

28. Winans, S.C.; Elledge, S.J.; Krueger, J.H.; Walker, G.C. Site-directed insertion and deletion mutagenesis with cloned fragments in *Escherichia coli*. *J. Bacteriol.* **1985**, *161*, 1219–1221.

29. Engberg, J.; Nielsen, H.; Lenaers, G.; Murayama, O.; Fujitani, H.; Higashinakagawa, T. Comparison of primary and secondary 26S rRNA structures in two Tetrahymena species: Evidence for a strong evolutionary and structural constraint in expansion segments. *J. Mol. Evol.* **1990**, *30*, 514–521.

30. Shkilnyj, P.; Koudelka, G.B. Effect of salt shock on the stability of λ^{imm434} lysogens. *J. Bacteriol.* **2007**, *189*, 3115–3123.

31. Delbrück, M. The growth of bacteriophage and lysis of the host. *J. Gen. Physiol.* **1940**, *23*, 643–660.

32. Wang, I.N. Lysis timing and bacteriophage fitness. *Genetics* **2006**, *172*, 17–26.

33. Schagger, H.; Von Jagow, G. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* **1987**, *166*, 368–379.

Article

Insights into the Functions of a Prophage Recombination Directionality Factor

Gaël Panis [†], Nathalie Franche, Vincent Méjean and Mireille Ansaldi *

Laboratoire de Chimie Bactérienne CNRS UMR7283, Institut de Microbiologie de la Méditerranée, Aix-Marseille University, 31 chemin Joseph Aiguier, 13402 Marseille cedex 20, France; E-Mails: Gael.Panis@unige.ch (G.P.); nfranche@imm.cnrs.fr (N.F.); mejean@imm.cnrs.fr (V.M.)

[†] Present address: Department of Microbiology and Molecular Medicine, Faculty of Medicine, University of Geneva, CH-1211 Geneva, Switzerland.

* Author to whom correspondence should be addressed; E-Mail: ansaldi@imm.cnrs.fr; Tel.: +33-491-164-585; Fax: +33-491-718-914.

Abstract: Recombination directionality factors (RDFs), or excisionases, are essential players of prophage excisive recombination. Despite the essentially catalytic role of the integrase in both integrative and excisive recombination, RDFs are required to direct the reaction towards excision and to prevent re-integration of the prophage genome when entering a lytic cycle. KplE1, HK620 and numerous (pro)phages that integrate at the same site in enterobacteria genomes (such as the *argW* tRNA gene) all share a highly conserved recombination module. This module comprises the *attL* and *attR* recombination sites and the RDF and integrase genes. The KplE1 RDF was named TorI after its initial identification as a negative regulator of the *tor* operon. However, it was characterized as an essential factor of excisive recombination. In this study, we designed an extensive random mutagenesis protocol of the *torI* gene and identified key residues involved in both functions of the TorI protein. We show that, in addition to TorI-TorR protein-protein interaction, TorI interacts in solution with the IntS integrase. Moreover, *in vitro*, TorR and IntS appear to compete for TorI binding. Finally, our mutagenesis results suggest that the C-terminal part of the TorI protein is dedicated to protein-protein interactions.

1. Introduction

The name bacteriophage encompasses all bacterial viruses, including temperate phages which have the particularity to integrate their genomes into their hosts, becoming prophages. The physiological state of the host determines the type of infection, either lytic or lysogenic. For example, in lambda, a starved host or a multiplicity of infection (MOI) higher than two, favors lysogeny [1,2]. The prophage is passively replicated as part of the host chromosome as long as conditions are not threatening to the host, in which case the prophage shifts to a lytic development [3–5]. Site specific recombination (SSR) constitutes a key step in lysogenic development since it is required for integration as well as for excision of the prophage genome [6]. This reaction is mediated in both directions by a specific recombinase, called integrase, that belongs either to the tyrosine or the serine recombinase families [7,8]. Whereas host factors can modulate the efficiency of the integrase mediated reactions, most of the time directionality is driven by recombination directionality factors (RDF) or excisionases [9–12]. In lambda, the RDF protein not only directs the reaction towards excision but also prevents reintegration of the excised phage genome [13–15].

Bacterial genomes are parasitized by prophages and prophage remnants, which can constitute up to 20% of the host genome. A pan-genomic study of 20 *Escherichia coli* genomes revealed that unique prophage genes are more abundant than core genes present in all 20 genomes [16]. Prophages are thus an important vector of bacterial genome evolution. While integrated into the host genome prophages also undergo rapid evolution, occurring mostly through homologous recombination and frequently leading to the loss of lytic genes [17,18]. The resulting prophages are not infectious anymore, however, they may conserve features that suit the host. In *Escherichia coli* K12, 10 prophage regions have been identified [19]. We studied one of them in particular, the KplE1 (or CPS53) prophage. This latter is inserted into the *argW* tRNA gene at 2,474 kb on the *E. coli* chromosome and contains 16 ORFs. Most have unknown functions, whereas we previously characterized the role of the first gene *intS* and the last one *torI* in site-specific recombination [20,21]. Interestingly, various (pro)phages that insert at the same tRNA gene locus share a highly conserved recombination module that comprises the *attL* and *attR* sites and the IntS integrase and the TorI RDF genes [21]. The identity even reaches 100% with the RDF proteins of HK620 and Sf6, named HkaC and P18, respectively [22,23].

The TorI protein (for Tor Inhibition) was originally identified using a genetic screen as a negative regulator of the *torCAD* operon that encodes the trimethylamine oxide reductase respiratory system in *E. coli* [24]. Despite its role as an inhibitor of the TorR response regulator, TorI was then genetically and structurally characterized as the RDF of the KplE1 prophage [20]. Recently, we also identified the host-encoded stress-responsive molecular chaperone DnaJ as an active participant in

KplE1 prophage excision. DnaJ is recruited by TorI and stabilizes its tridimensional structure which has for consequence to increase TorI affinity for its specific binding sites on *attL* [25,26].

In this study, we designed an extensive random mutagenesis protocol of the *torI* gene to identify critical residues involved in the anti-response regulator (anti-RR) and/or the excisive recombination activity functions of the TorI protein.

2. Results and Discussion

2.1. Tester Strain and Random Mutagenesis of the torI Gene

In order to identify critical residues involved in the anti-RR and/or the excisive recombination activity of the TorI RDF, we designed a tester strain that can report both activities. Strain LCB995 contains a *torA'-'lacZ* fusion as well as a *cat* cassette inserted into a non-coding region of the KplE1 prophage [20,27]. When this reporter strain was plated onto MacConkey lactose plates in the presence of 10 mM TMAO to induce the *tor* operon promoter, the colonies formed turned red, as the β-galactosidase was produced [27]. In contrast, when this strain was transformed with a multicopy plasmid encoding a wild-type version of the *torI* gene (pJFi plasmid), which allows inducible TorI production in the presence of 1 mM IPTG, the resulting colonies remained white on the same medium due to the anti-RR activity of TorI. On the other hand, the RDF activity of TorI was monitored with the same tester strain containing pJFi plated onto chloramphenicol (Cm) containing plates in the presence of 1 mM IPTG. As described before [20], expression of the *torI* gene is sufficient to promote KplE1 excision, and thus renders the cells chloramphenicol sensitive. Since colonies on agar plate arose from single cells, colonies are not forming from cells that have excised the KplE1 prophage. We thus have set up an experimental procedure to screen a mutant library of TorI that may be affected in two different activities (Figure 1).

Figure 1. Strategy to study TorI functions.
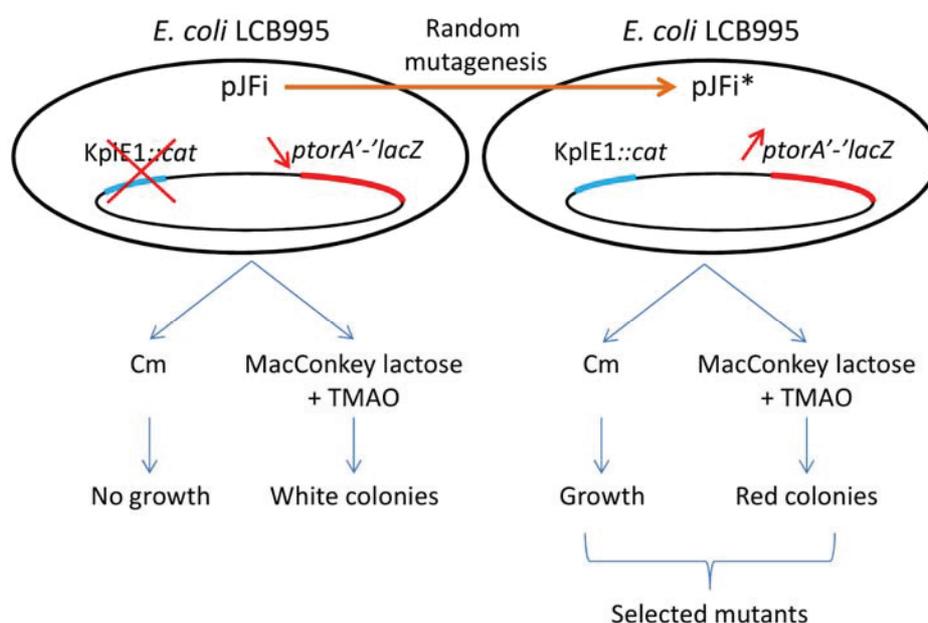
The mutant library of *torI* alleles was generated by extensive random PCR-mutagenesis of the *torI* gene (201 base pairs) carried by plasmid pJFi (see the Experimental Section). About 3,000 colonies were screened on reporter (MacConkey plates containing lactose and TMAO) or selective media (LB plates containing chloramphenicol), and colonies that presented a red color on MacConkey lactose plates containing TMAO or that could grow in the presence of chloramphenicol, were selected. A final set of 42 clones reproducibly altered in one or both TorI activities was selected for further investigation (Figure 1). To make sure that the observed phenotypes were not due to chromosomal mutations, all plasmids were extracted and used to transform the tester strain. Direct Sanger sequencing of the plasmids carrying the mutated alleles allowed the identification of the mutations. In total we isolated 18 distinct mutations, of which 16 were due to a point mutation leading to a single amino-acid substitution (Table 1). Overall, the mutagenesis showed no obvious mutational bias with mutations dispersed all along the primary sequence and a wide diversity in the mutations that we obtained with 15 transitions, four transversions, one deletion and two extensions due to a frameshift on the stop codon. No redundant mutations were observed, in total 20 different positions out of 66 were affected, most of them were changes in residues located on the surface of the protein (with the exception of Ile_{16}, Trp_{47} and Phe_{56}), and only one substitution has impacted the overall structure of the protein (L5P, Table 1 and text below). Thereafter, pJFi* designates a plasmid carrying a mutated allele, and TorI* indicates a mutant protein.

A preliminary *in vivo* quantification of TorI activities was performed (Table 1). The anti-RR activity of the TorI mutants was estimated by monitoring β-galactosidase activities of strain LCB995 containing pJFi*. This activity reports the expression of the *torA'-'lacZ* fusion in the presence of TMAO. On the other hand, the RDF activity of TorI mutants was measured in the same strain by counting colonies able to grow on chloramphenicol plates relative to colonies counted on ampicillin plates. Results are expressed as the ratio of ampicillin-resistant/chloramphenicol-resistant colonies, and reflect the ability of the TorI mutants to promotes KplE1 excision. We identified two classes of mutants: one class contained mutants affected in the anti-RR activity only, whereas the second class was composed of mutants affected in both activities. As a result, we did not isolate mutants affected in the RDF activity only. In recent studies [20,26], we also designed several mutants by site-directed mutagenesis which were included in Table 1 ([d] mark) and were further analyzed together with the mutants generated by random mutagenesis.

Table 1. Substitutions in TorI mutants.

Plasmid [a]	Mutation	Substitution	RDF activity [b]	Anti-TorR activity [c]
pJFi	N/A	N/A	+ + + +	+
pJFi-L5P	CTA → C**C**A	Leu5 → Pro	+ + + +	−
pJFi-S9L	TCA → T**T**A	Ser9 → Leu	−	−
pJFi-D12Y-D35G	GAT → **T**AT GAC → G**G**C	Asp12 → Tyr Asp35 → Gly	−	−
pJFi-F15L	TTC → **C**TC	Phe15 → Leu	−	−
pJFi-I16V	ATC → **G**TC	Ile16 → Val	+ +	−
pJFi-M17V	ATG → **G**TG	Met17 → Val	+ +	−
pJFi-F22I	TTT → **A**TT	Phe22 → Ile	−	−
[d] pJFi-Y28F	TAT → T**T**T	Tyr28 → Phe	−	−
[d] pJFi-Y28S	TAT → T**C**T	Tyr28 → Ser	−	−
pJFi-P37L	CCA → C**T**A	Pro37 → Leu	+	−
pJFi-H43Y-C54R	CAC → **T**AC TGT → **C**GT	His43 → Tyr Cys54 → Arg	+	−
pJFi-R45STOP	CGA → **T**GA	Arg45 → STOP	−	−
[d] pJFi-R45Q	CGA → C**A**A	Arg45 → Gln	+	−
[d] pJFi-R45K	CGA → **AA**A	Arg45 → Lys	+ +	−
pJFi-A46V	GCA → G**T**A	Ala46 → Val	−	−
pJFi-A46T	GCA → **A**CA	Ala46 → Thr	−	−
pJFi-W48R	TGG → **C**GG	Trp48 → Arg	−	−
pJFi-E55G	GAA → G**G**A	Gln55 → Gly	+ + + +	−
pJFi-F56L	TTC → **C**TC	Phe56 → Leu	+ + + +	−
[d] pETsI-L61S	TTA → T**C**A	Leu61 → Ser	ND	ND
[d] pETsI-R63C-A64S	CGC → **T**GC GCC → **T**CC	Arg63 → Cys Ala64 → Ser	ND	ND
pJFi-N65Y	AAT → **T**AT	Asn65 → Tyr	+ + + +	−
pJFi+18	TAA → T**C**A	+18 residues [e]	+	−
pJFi+24	TAA → **-**AA	+24 residues [f]	+ + + +	−

[a] All *torI* alleles of the pJFi* plasmid series have been sub-cloned into the pET-22(+) vector leading to the pETsi* series; [b] RDF activities were estimated by the ratio of colonies Ap^R/Cm^R and indicated as follows: ++++ 80 to 100% of the activity measured in the presence of the WT *torI* allele , ++ 1 to 10%, + 0,1 to 1% , and − <0,1% (corresponds to the activity observed in the presence of the empty vector pJF119EH); [c] Anti-RR activities were estimated by β-galactosidase activity measurements of cells grown in the presence of 10 mM TMAO and 1 mM IPTG. + indicates the activity of the WT *torI* allele, − refers to a null anti-RR activity; [d] Mutants obtained by site-directed mutagenesis [20,26]; [e] Additional residues: SGSSRVDLQACKLGCFGG; [f] Additional residues: KDPLESTCRHASLAVLADERRFSA.

2.2. Mapping of the Mutations that Affect TorI Activities

To our surprise, all the mutations we obtained impaired the anti-RR activity of TorI, and these mutations map all over the surface of the protein (Figure 2A). In a previous work, we showed that TorI inhibited the transcriptional activator TorR through a direct protein-protein interaction without preventing TorR binding to its specific DNA targets, and we hypothesized that TorI was binding to a region of TorR that is important for RNA polymerase recruitment [24]. All the surface mutations that we obtained altered the anti-RR activity, suggesting that more than one face of the protein is involved in this function and that one face of the protein could bind to TorR and another to RNA

polymerase. A vast majority of the mutants isolated affect surface residues, which are probably involved in protein-protein interactions. One may have expected a more subtle effect on the anti-RR activity, however, our screen failed to select those.

In contrast to the anti-RR activity, only a subset of mutations altered the RDF activity of TorI and most of these mutations map on the surface of the protein although the mutations are scattered along the primary sequence (Figure 2B). By this approach, we identified two regions of the protein involved in the excisionase function: one centered on helix 1 and another one comprising the helix 2 and the wing motif. The wing-helix is an atypical helix-turn-helix motif involved in DNA-binding commonly found in transcription factors and particularly in the OmpR family of response regulators [28,29]. Based on the work done on the lambda Xis protein, two separate functions are [22] required for the excisionase activity: (i) Xis interacts with DNA to position itself on the *X* sites on the *attR* recombination region, (ii) Xis interacts with Int to direct a proper positioning of the integrase for the excisive reaction [15]. In a previous work, residues involved in DNA binding were identified by NMR titration with a short DNA sequence and mapped onto the wing-helix motif [20]. Therefore, two positions of the TorI wing-helix motif were mutated (Tyr$_{28}$ and Arg$_{45}$) and the resulting proteins proved to be impaired in KplE1 excision, although Tyr$_{28}$ mutants were more affected than Arg$_{45}$ ones, suggesting a role for this region in DNA binding activity.

Figure 2. Mapping of the mutations that affect the anti-RR function (**A**) or the recombination directionality factors (RDF) activity (**B**) of TorI; Secondary structures of the TorI protein (**C**). Color code is as follows: *blue*, no anti-RR activity; *green*, 80 to 100% of the WT activity; *yellow*, 1 to 10%, *orange*, 0.1 to 1%, and *red*, no RDF activity.
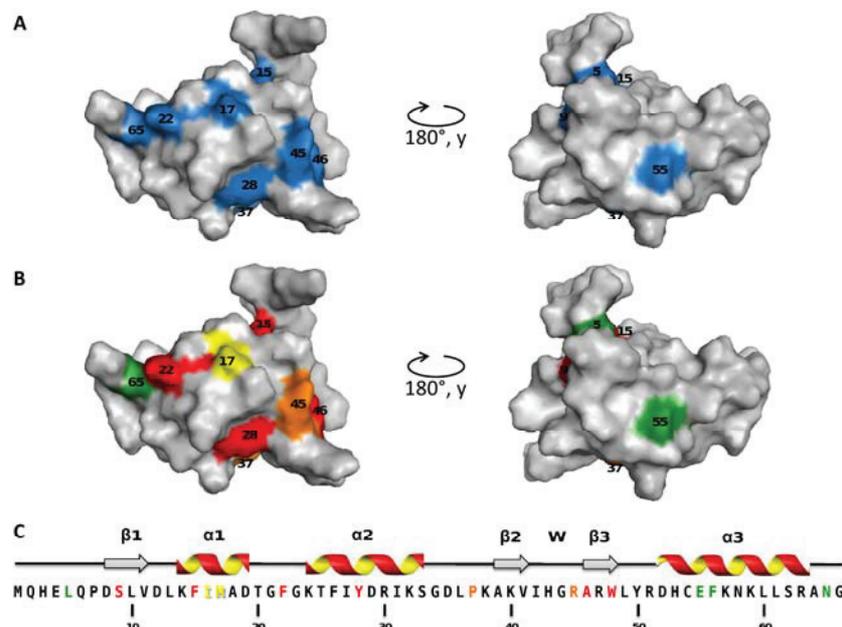
2.3. Characterization of the TorI Mutants

2.3.1. Production and Stability of the Mutants for *in Vitro* Studies

As a preliminary quality check, we selected 14 single substitution mutants (only one substitution at position 28 and 45 were selected) and analyzed their respective production and stability by

western-blot. All the mutants were produced from the pJFi* plasmids in the presence of 1 mM IPTG. Crude extracts were run on a 16% SDS Tricine-PAGE and TorI* production was revealed using an anti-TorI polyclonal serum (see Experimental Section). Accordingly, all mutant proteins were produced at similar levels and the antibodies failed to reveal bands below the full size proteins indicating that all mutants have a similar stability *in vivo* as the wild-type protein (data not shown).

2.3.2. *In Vitro* RDF Activity of TorI Mutants

Most mutants isolated in this study, as well as those designed elsewhere [20,26], were produced from pETsI* plasmids, purified near homogeneity and analyzed *in vitro*. Only a few of them proved to be unstable upon purification (TorI-S9L and TorI-W48R). Since the RDF activity of TorI relies on its DNA binding activity, we first checked the ability of the mutants to bind to *attL*. Electrophoretic mobility shift assay (EMSA) were performed essentially as described [30], except that two different concentrations of DNA were used in this assay (Figure 3).

Figure 3. DNA binding activity of the TorI mutants. Electrophoretic mobility shift assay (EMSA) experiments were performed using the KplE1 *attL* recombination region that contains five cooperative binding sites for TorI [30]. Proteins were all used at a concentration of 10 μM whereas DNA concentration was either 20 nM (top) or 10 nM (bottom). When two *torI* alleles were available, only one was chosen for *in vitro* characterization, namely TorI-Y28F, TorI-R45Q, and TorI-A46V.
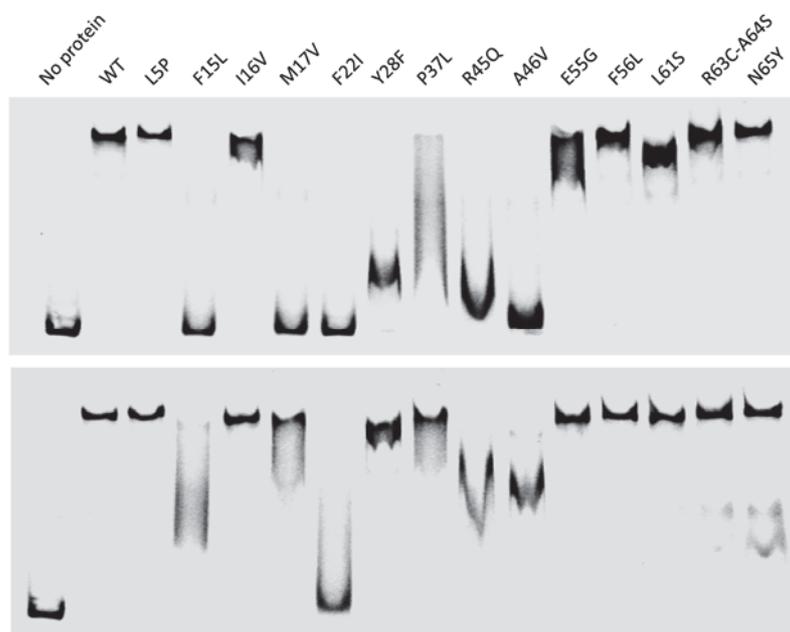
At high DNA concentration (top panel), several mutants displayed reduced DNA binding activity. The mutations mapped either in the helix-turn-helix motif (F15, M17, F22, Y28), or in the wing region (R45, A46). Interestingly when the protein:DNA ratio was increased (Figure 3, bottom panel), most of these mutants were able to shift *attL* to a certain extent, with the exception of F22I. As TorI binding to the 5 TorI sites in *attL* is highly cooperative [30], smeary patterns were observed in some cases rather than intermediate shifts. This result indicates that point mutations in the DNA binding

region of TorI affect binding to *attL* by reducing the affinity of the protein for its DNA substrate, and this effect can be often overcome by increasing the protein:DNA ratio.

2.3.3. *In Vitro* Excisive Recombination

We further investigated the effect of the TorI mutations on the ability of the protein to promote excisive recombination *in vitro*. Under the *in vitro* conditions we used (see the Experimental Section), all mutants generally behaved the same way as *in vivo*, also some discrepancies could be observed probably due to the highest sensitivity of the *in vitro* assay (Figure 4). Four mutants (TorI-F22I, TorI-F28F, TorI-F45Q, and TorI-A46V) were totally inactive *in vitro*, and the corresponding mutations all lie in the winged-helix structural motif (Figure 2). Interestingly, these same mutants were also largely impaired in binding to the *attL* substrate (Figure 3). In addition, three mutants (TorI-F15L, TorI-M17V, and TorI-L61S) showed significantly reduced *in vitro* activities. Among these, TorI-F15L and TorI-M17V were clearly affected in DNA binding to *attL* as well and the corresponding mutations lie in the first α-helix of the helix-turn-helix motif. In contrast, TorI-L61S exhibited an almost wild-type DNA binding activity and the substitution occurred in the last α-helix of the protein, therefore excluding this region from the DNA binding domain. Together, these results are consistent with DNA binding to *attL* being the primary activity required for optimal RDF activity.

Figure 4. *In vitro* excisive recombination assay. Reactions were conducted as indicated in the Experimental Section with the same protein variants as in Figure 3. A, the agarose gel electrophoresis shows the relative migration of the substrates (*attL* and *attR*) and the products (*attP* and *attB*). B, the gel shown in A was scanned and the *attP* product quantified. Activities of the mutants are expressed as a percentage of the WT protein activity.
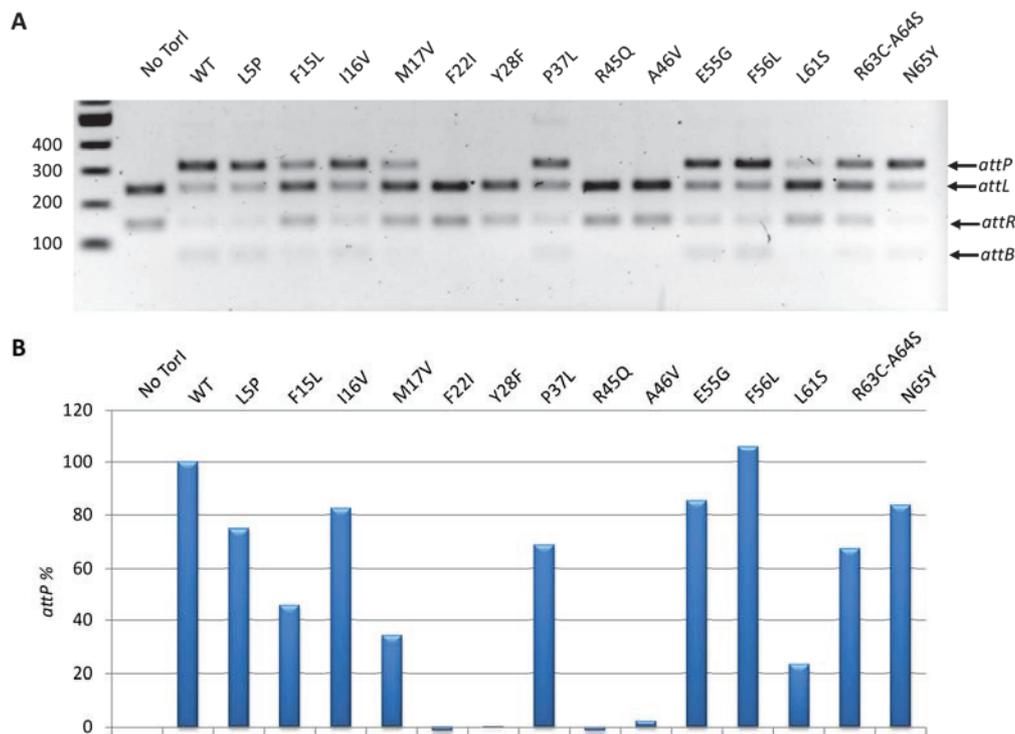
2.3.4. Competition for TorR and IntS Binding on TorI Protein

Regarding the two functions in which TorI was shown to be involved, we then asked if the proteins involved in these functions, namely TorR and IntS could compete for the same TorI target. We have previously shown that TorI was able to bind to the C-terminal domain of the TorR response regulator [24], however an interaction with IntS, although suspected, was not yet identified. Indeed, when incubated in the presence of IntS, TorI was revealed in a band in which migration (~50 kDa) was compatible with a heterodimer TorI-IntS (IntS, 42.5 kDa and TorI, 7.7 kDa) (Figure 5, lane TorI+IntS10). Moreover, the presence of the IntS protein in this extra band of ~50 kDa was confirmed by mass spectrometry analysis and by using α-IntS antiserum (data not shown). According to what has been shown earlier [24], in the presence of TorR a major band corresponding to a 1:1 TorI:TorR ratio (~34 kDa) was detected with α-TorI antiserum. Additional bands may correspond to complexes with different TorI:TorR ratios (2:1, ~42 kDa; 4:1, ~57 kDa), and suggest that TorR can bind multiple forms of TorI. Remarkably, when IntS and TorR were incubated together with TorI and at equimolar concentrations (10 µM), only a faint band corresponding to the major complex TorI:TorR was detected, indicating that IntS has probably more affinity for TorI than TorR does. However, in the presence of molar excess of TorR (20, 30 or 40 µM), TorR was able to displace IntS, although not totally. On the other hand, when IntS was added in excess in the reaction compared to TorR, no more TorI:TorR complex was observed. Together, these results suggest that IntS and TorR bind to a similar region on TorI and that the complex formed between TorI and IntS is more stable than with TorR.

Figure 5. IntS and TorR bind to TorI and replace each other. A binding assay of TorI (10 µM) in the presence of various IntS and/or TorR concentrations, as indicated, was performed using BMH as a crosslinker (see the Experimental Section). Major TorI-IntS and TorI-TorR complexes are pointed by arrows.
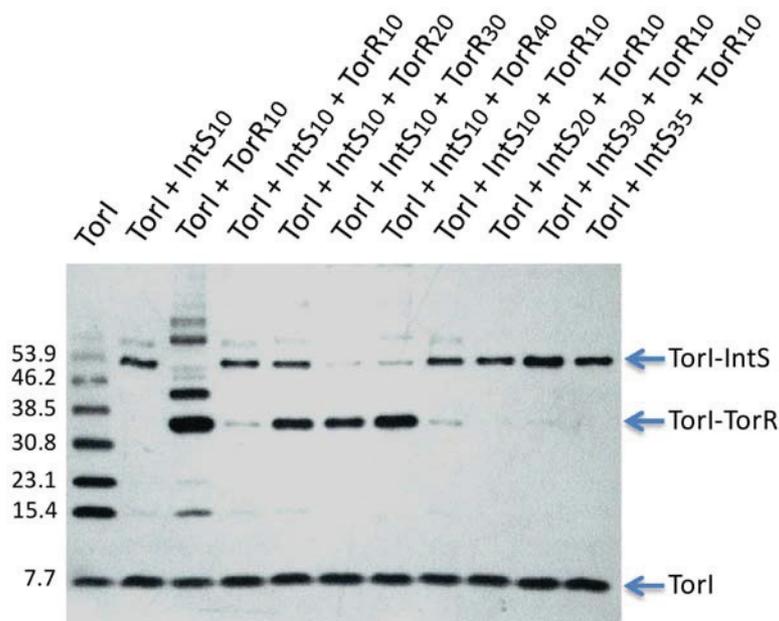
2.3.5. TorI Mutants Binding to TorR and IntS

As suggested by the competition experiments, TorR and IntS may bind a similar region on TorI. Our TorI mutant collection was thus assayed for binding with the two proteins (Figure 6). When assayed in the presence of IntS (Figure 6A), all TorI mutants were able to crosslink with IntS although, some, such as TorI-A46V, TorI-E55G and TorI-N65Y exhibited reduced binding. In contrast, binding with TorR (Figure 6B) indicated that TorI-N65Y, which has the mutation in the last amino-acid of the protein, did not bind anymore TorR. Together, these results suggest that the C-terminal part of the TorI protein is involved in both IntS and TorR binding.

Figure 6. IntS and TorR binding to TorI mutants. A, binding assay of TorI mutants (10 µM) in the presence of IntS (10 µM). B, binding assay of TorI mutants (10 µM) in the presence of TorR (10 µM). Arrows indicate the major complexes.
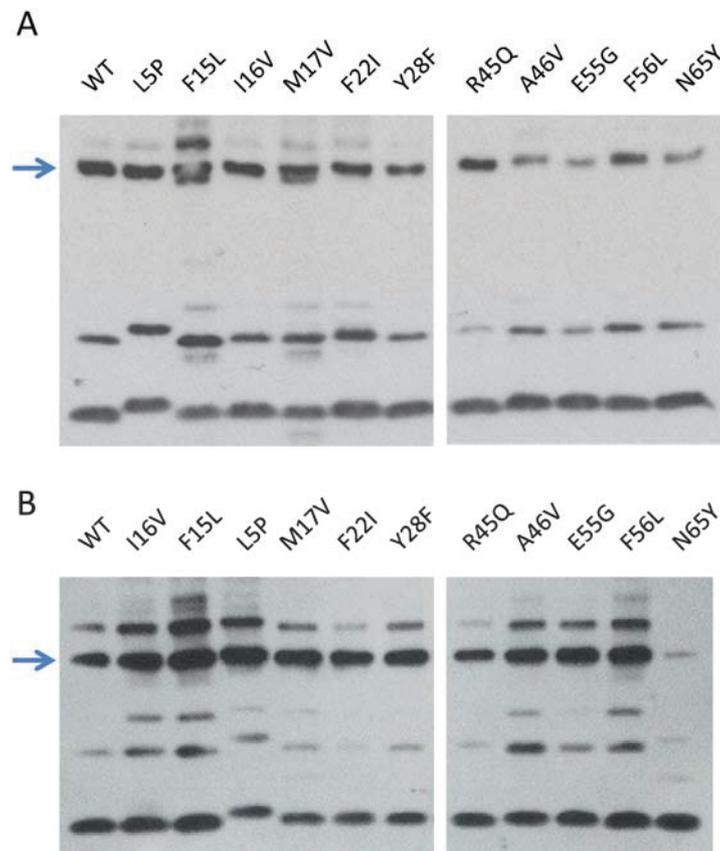
3. Experimental Section

3.1. Strains and Media Used in This Study

Strains and plasmids used in this study are listed in Table 2. Strain LCB995 was constructed by transduction of the *cat* marker inserted in the KplE1 prophage between *yfdO* and *yfdP* (strain LCB970 [20]) into strain LCB620 carrying a *torA'-'lacZ* fusion [27]. Strains were grown in LB medium supplemented when necessary with ampicillin (50 µg.mL^{-1}), chloramphenicol (25 µg.mL^{-1}), TMAO (10% v/v) or IPTG (1 mM).

Table 2. Strains and plasmids.

Strains and plasmids	Characteristics	Sources
Strains		
MC4100	*araD139* (Δ*lacIPOZYA-argF*) *U169 rpsL thi*	Casadaban
LCB620	MC4100 *torA8*::MudII 1734 (*torA'-'lacZ*, Km^R)	[27]
LCB970	MC4100 *yfdO-cat-yfdP*	[24]
LCB995	MC4100 *torA8*::MudII 1734, *yfdO-cat-yfdP*	This work
LCB984	MC4100 *yfdO-kan-yfdP*	This work
BL21(DE3)	*E. coli* B F- *[lon] dcm ompT hsdS* (r_B⁻m_B⁺) gal λ(DE3)	Novagen
C41(DE3)	Derived from BL21(DE3)	[31]
Plasmids		
pBAD33	pACYC184 (ori p15A) vector containing a P_BAD promoter (Cm^R)	[32]
pBtorR	*torR* coding sequence cloned into pBAD33	[33]
pJF119EH	pBR322 (ori *col*E1) containing the IPTG inducible promoter p*tac* (Ap^R)	[34]
pJFi	*torI* coding sequence cloned into pJF119EH BamHI and EcoRI sites	[24]
pJFi* series	pJFi derived plasmids carrying *torI* mutated alleles	This work [20,26]
pET-22(+)	Promoter T7 containing vector (Ap^R)	Novagen
pETsi	*torI* coding sequence with a Stop codon cloned into pET-22(+) NdeI and XhoI sites	[24]
pETsi* series	pETsi derived plasmids carrying *torI* mutated alleles	This work [20,26]

> NB, all plasmids derived from pJFi and pETsi plasmids and carrying mutated *torI* alleles are mentioned in Table 1. Otherwise indicated, pJFi and pETsi derivatives were isolated or constructed in this study.

3.2. Random Mutagenesis

Error prone PCR was performed to generate mutated alleles of the *torI* gene. PCR was conducted with the primer pair torI_MunI (5'-TAC AAT TGC GGA GAT AGC ACT CAT GCA ACA C)/torI_BamHI (5'-TTG GAT CCT TAC CCA TTG GCG CGG CTT AAG AG), plasmid pJFi as a template and a classical Taq polymerase (GoTaq Promega, error rate 10^{-6}). Three rounds of PCR were performed using as a template a 10^6 dilution of the product generated by the previous round of PCR. After three rounds of PCR, products were purified, hydrolyzed with MunI and BamHI enzymes and ligated into the pJF119EH vector cut with EcoRI and BamHI. The plasmid library was then transformed into the tester strain LCB995.

3.3. Protein Production and Purification

IntS, TorI and IHF proteins were overproduced and purified to homogeneity as previously described [24,30]. TorI mutants were purified as the wild-type protein. All proteins were dialyzed in

Tris-HCl buffer (40 mM, pH 7.6) containing 50 mM KCl and 10% glycerol. The protein concentrations were measured by densitometry with the wild-type TorI protein as a reference.

3.4. In Vivo Excision Assay

Strain LCB970 carrying *torI* encoding plasmids pJFi (7) was grown in LB medium until the OD_{600} reached 0.5 units (0.5×10^9 cells.mL^{-1}), and IPTG (1 mM) was added for 2 h at 37 °C under agitation. Culture dilutions were prepared and plated onto rich medium containing either 50 µg.mL^{-1} ampicillin or 25 µg.mL^{-1} chloramphenicol. Numeration of the colonies plated on both antibiotics was performed and the ratio of ampicillin-resistant/chloramphenicol-resistant colonies was calculated. Values represent the average of at least three independent determinations.

3.5. β-Galactosidase Assay

β-Galactosidase activities were measured on whole cells according to the method of Miller (1972); values represent the average of at least three determinations with a variation of no more than 10% from the mean.

3.6. In Vitro *Excisive Recombination*

Linear *att* sites were amplified by PCR with primer pairs attL-SpeI (5'-GAC TAG TTT CAA TCT GCT TAA CGG TGA GCA T)/attL-KpnI (5'-GGG GTA CCG CTA ATT GCA GGT TCG ATT CC) for *attL* (220 bp) and attR-XbaI (5'-GCT CTA GAG GTT TTA GGG ATA AAC ACA CAA GGA TG)/attR-IHF2 (5'-CTC TTA AGC CGC GCC AAT GG) for *attR* (135 bp), and then purified using Qiaquick PCR purification kit protocol (Qiagen). Reaction mixtures (25 µL) included linear *att* DNA sites (28 nM) in buffer containing 33 mM Tris-HCl pH 7.6, 33 mM KCl, 9 mM spermidine, 4 mM EDTA, 0.9 mg/mL^{-1} acetylated BSA and 7 % glycerol. IHF (0.3 µM), IntS (0.6 µM), and TorI (2.1 µM) were added as indicated in the figures legends. The reactions were carried out in optimized conditions at 30 °C for 2 h at an IHF:IntS:TorI protein ratio of 1:3:7. Reaction products were purified (Qiaquick kit, Qiagen) and analyzed on a 2% agarose gel electrophoresis. The gel was then scanned and the data analyzed using AlphaView software (Protein simple).

3.7. Electrophoretic Mobility Shift Assays (EMSA)

EMSA were carried out using purified proteins and fluorescently Cy5-labelled *attL* DNA fragment that was amplified by PCR using MG1655 chromosomal DNA as a template with the primer pair attL-pro (5'-AAT GGA TAT AAC GAG CCC CTC C)/attL-ter-Cy5 (5'-CAT CGA GAA GGC GGT ATG GTT TTT C). DNA and purified proteins were mixed together at different concentrations (as indicated in figure legends) in the presence of 4 mg.mL^{-1} BSA and 0.5 mg.mL^{-1} calf thymus DNA (CT-DNA) in binding buffer (40 mM Tris pH 7.6, 85 mM KCl, 19% glycerol). Reactions were incubated for 30 min at 30 °C. DNA-protein complexes were then separated using a 6% non-denaturing polyacrylamide gel (37:1 acrylamide:bisacrylamide ratio). A pre-migration step (1 hour at 160 V) was carried out to reduce ionic charges which may have destabilized the

DNA-protein complex. Samples were then loaded and left to migrate at 80 V during 30 min and then at 160 V for 2 more hours in 0.5× TBE (50 mM Tris, 45 mM Boric acid, 0.5 mM EDTA) running buffer. The gel was scanned using a FLA5100 (Fuji) scanner, using excitation wavelength of 635 nm (800 V scanning intensity) and emission wavelength of 665 nm. Data was analyzed using Multi Gauge (Version 2.3) software [35].

3.8. Cross-Linking Analysis

TorI protein was pre-incubated 10 min at 25 °C in the presence or absence of IntS and/or TorR proteins. Then, homobifunctional sulhydryl reactive agent bis(maleimido)hexane (BMH, 1mM, Pierce) was added and the reaction continued for another 30 min. Samples were ran on a 7%–16% Tricine-SDS PAGE, transferred onto a nitrocellulose membrane and revealed by immunodetection with TorI antiserum.

4. Conclusions

In this work we describe an extensive mutational analysis of a small protein involved in prophage excision. Together, our results suggest that the C-terminal part of the TorI protein is somehow dedicated to protein-protein interactions, since this region seems to bind at least three different molecular partners, the TorR response regulator, the IntS integrase and, as shown before, the DnaJ cochaperone [25,26]. This is particularly relevant to the tridimensional structure of the TorI family of RDF proteins that contain a long and well defined α-helix at the C-terminus of the protein, which is not found in other RDF proteins in solutions [20].

Acknowledgments

Conflict of Interest

The authors declare no conflict of interest.

References and Notes

1. Kourilsky, P.; Knapp, A. Lysogenization by bacteriophage lambda. III. Multiplicity dependent phenomena occuring upon infection by lambda. *Biochimie* **1974**, *56*, 1517–1523.
2. Edgar, R.; Rokney, A.; Feeney, M.; Semsey, S.; Kessel, M.; Goldberg, M.B.; Adhya, S.; Oppenheim, A.B. Bacteriophage infection is targeted to cellular poles. *Mol. Microbiol.* **2008**, *68*, 1107–1116.
3. Williamson, S.J.; Houchin, L.A.; McDaniel, L.; Paul, J.H.: Seasonal variation in lysogeny as depicted by prophage induction in Tampa Bay, Florida. *Appl. Environ. Microbiol.* **2002**, *68*, 4307–4314.

4. McDaniel, L.; Paul, J.H. Effect of nutrient addition and environmental factors on prophage induction in natural populations of marine synechococcus species. *Appl. Environ. Microbiol.* **2005**, *71*, 842–850.

5. Rokney, A.; Kobiler, O.; Amir, A.; Court, D.L.; Stavans, J.; Adhya, S.; Oppenheim, A.B. Host responses influence on the induction of lambda prophage. *Mol. Microbiol.* **2008**, *68*, 29–36.

6. Nash, H. Site-specific recombination: Integration, excision, resolution, and inversion of defined DNA segments. In *Escherichia coli and Salmonella: Cellular and Molecular Biology*; Neidhardt, F., Ed.; ASM Press:Washington, DC, USA, 1996; pp. 2363–2376.

7. Argos, P.; Landy, A.; Abremski, K.; Egan, J.B.; Haggard-Ljungquist, E.; Hoess, R.H.; Kahn, M.L.; Kalionis, B.; Narayana, S.V.; Pierson, L.S.; *et al.* The integrase family of site-specific recombinases: Regional similarities and global diversity. *EMBO J.* **1986**, *5*, 433–440.

8. Nunes-Duby, S.E.; Kwon, H.J.; Tirumalai, R.S.; Ellenberger, T.; Landy, A. Similarities and differences among 105 members of the Int family of site-specific recombinases. *Nucleic Acids Res.* **1998**, *26*, 391–406.

9. Gottesman, S.; Abremski, K. The role of HimA and Xis in lambda site-specific recombination. *J. Mol. Biol.* **1980**, *138*, 503–512.

10. Bushman, W.; Yin, S.; Thio, L.L.; Landy, A. Determinants of directionality in lambda site-specific recombination. *Cell* **1984**, *39*, 699–706.

11. Gardner, J.F.; Nash, H.A. Role of *Escherichia coli* IHF protein in lambda site-specific recombination. A mutational analysis of binding sites. *J. Mol. Biol.* **1986**, *191*, 181–189.

12. Lewis, J.A.; Hatfull, G.F. Control of directionality in integrase-mediated recombination: Examination of recombination directionality factors (RDFs) including Xis and Cox proteins. *Nucleic Acids Res.* **2001**, *29*, 2205–2216.

13. Numrych, T.E.; Gumport, R.I.; Gardner, J.F. Characterization of the bacteriophage lambda excisionase (Xis) protein: The C-terminus is required for Xis-integrase cooperativity but not for DNA binding. *EMBO J.* **1992**, *11*, 3797–3806.

14. Sam, M.D.; Papagiannis, C.V.; Connolly, K.M.; Corselli, L.; Iwahara, J.; Lee, J.; Phillips, M.; Wojciak, J.M.; Johnson, R.C.; Clubb, R.T. Regulation of directionality in bacteriophage lambda site-specific recombination: Structure of the Xis protein. *J. Mol. Biol.* **2002**, *324*, 791–805.

15. Warren, D.; Sam, M.D.; Manley, K.; Sarkar, D.; Lee, S.Y.; Abbani, M.; Wojciak, J.M.; Clubb, R.T.; Landy, A. Identification of the lambda integrase surface that interacts with Xis reveals a residue that is also critical for Int dimer formation. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 8176–8181.

16. Touchon, M.; Hoede, C.; Tenaillon, O.; Barbe, V.; Baeriswyl, S.; Bidet, P.; Bingen, E.; Bonacorsi, S.; Bouchier, C.; Bouvet, O.; *et al.* Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. *PLoS Genet.* **2009**, *5*, e1000344.

17. Canchaya, C.; Proux, C.; Fournous, G.; Bruttin, A.; Brussow, H. Prophage genomics. *Microbiol. Mol. Biol. Rev.* **2003**, *67*, 238–276.

18. Casjens, S. Prophages and bacterial genomics: What have we learned so far? *Mol. Microbiol.* **2003**, *49*, 277–300.

19. Campbell, A.M. Cryptic prophages. In *Escherichia coli and Salmonella: Cellular and Molecular Biology*; Neidhardt, F., Ed.; ASM Press: Washington, DC, USA 1996; pp. 2041–2046.

20. ElAntak, L.; Ansaldi, M.; Guerlesquin, F.; Méjean, V.; Morelli, X. Structural and genetic analyses reveal a key role in prophage excision for the TorI response regulator inhibitor. *J. Biol. Chem.* **2005**, *280*, 36802–36808.

21. Panis, G.; Méjean, V.; Ansaldi, M. Control and regulation of KplE1 prophage site-specific recombination: A new recombination module analyzed. *J. Biol. Chem.* **2007**, *282*, 21798–21809.

22. Clark, A.J.; Inwood, W.; Cloutier, T.; Dhillon, T.S. Nucleotide sequence of coliphage HK620 and the evolution of lambdoid phages. *J. Mol. Biol.* **2001**, *311*, 657–679.

23. Casjens, S.; Winn-Stapley, D.A.; Gilcrease, E.B.; Morona, R.; Kuhlewein, C.; Chua, J.E.; Manning, P.A.; Inwood, W.; Clark, A.J. The chromosome of *Shigella flexneri* bacteriophage Sf6: Complete nucleotide sequence, genetic mosaicism, and DNA packaging. *J. Mol. Biol.* **2004**, *339*, 379–394.

24. Ansaldi, M.; Théraulaz, L.; Méjean, V. TorI, a response regulator inhibitor of phage origin in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 9423–9428.

25. Champ, S.; Puvirajesinghe, T.M.; Perrody, E.; Menouni, R.; Genevaux, P.; Ansaldi, M. Chaperone-assisted excisive recombination, a solitary role for DnaJ (Hsp40) chaperone in lysogeny escape. *J. Biol. Chem.* **2011**, *286*, 38876–38885.

26. Puvirajesinghe, T.M.; Elantak, L.; Lignon, S.; Franche, N.; Ilbert, M.; Ansaldi, M. DnaJ (HSP40) binding to a folded substrate impacts KplE1 prophage excision efficiency. *J. Biol. Chem.* **2012**, doi:10.1074/jbc.M111.331462.

27. Méjean, V.; Iobbi-Nivol, C.; Lepelletier, M.; Giordano, G.; Chippaux, M.; Pascal, M.C. TMAO anaerobic respiration in *Escherichia coli:* Involvement of the *tor* operon. *Mol. Microbiol.* **1994**, *11*, 1169–1179.

28. Martinez-Hackert, E.; Stock, A.M. The DNA-binding domain of OmpR: Crystal structures of a winged helix transcription factor. *Structure* **1997**, *5*, 109–124.

29. Aravind, L.; Anantharaman, V.; Balaji, S.; Babu, M.M.; Iyer, L.M. The many faces of thehelix-turn-helix domain: Transcription regulation and beyond. *FEMS Microbiol. Rev.* **2005**, *29*, 231–262.

30. Panis, G.; Duverger, Y.; Champ, S.; Ansaldi, M. Protein binding sites involved in the assembly of the KplE1 prophage intasome. *Virology* **2010**, *404*, 41–50.

31. Miroux, B.; Walker, J.E. Over-production of proteins in *Escherichia coli*: Mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J. Mol. Biol.* **1996**, *260*, 289–298.

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

33. Bordi, C.; Ansaldi, M.; Gon, S.; Jourlin-Castelli, C.; Iobbi-Nivol, C.; Méjean, V. Genes regulated by TorR, the trimethylamine oxide response regulator of *Shewanella oneidensis*. *J. Bacteriol.* **2004**, *186*, 4502–4509.

34. Furste, J.P.; Pansegrau, W.; Frank, R.; Blocker, H.; Scholz, P.; Bagdasarian, M.; Lanka, E. Molecular cloning of the plasmid RP4 primase region in a multi-host-range *tacP* expression vector. *Gene* **1986**, *48*, 119–131.

35. *Multigauge*, version 2.3; Fujifilm: Tokyo, Japan, 2004.

Review

Utility of the Bacteriophage RB69 Polymerase gp43 as a Surrogate Enzyme for Herpesvirus Orthologs

Nicholas Bennett [1] and **Matthias Götte** [1,2,3,*]

[1] Department of Microbiology and Immunology, McGill University, 3775 University Street, Montreal, Quebec H3A 2B4, Canada

[2] Department of Biochemistry, McGill University, 3655 Sir William Osler Promenade, Montreal, Quebec H3G1Y6, Canada

[3] Department of Medicine, Division of Experimental Medicine, McGill University, 1110 Pine Avenue West, Montreal, Quebec H3A 1A3, Canada

* Author to whom correspondence should be addressed; E-Mail: matthias.gotte@mcgill.ca; Tel.: +1-514-398-1365; Fax: +1-514-398-7052.

Abstract: Viral polymerases are important targets in drug discovery and development efforts. Most antiviral compounds that are currently approved for treatment of infection with members of the *herpesviridae* family were shown to inhibit the viral DNA polymerase. However, biochemical studies that shed light on mechanisms of drug action and resistance are hampered primarily due to technical problems associated with enzyme expression and purification. In contrast, the orthologous bacteriophage RB69 polymerase gp43 has been crystallized in various forms and therefore serves as a model system that provides a better understanding of structure–function relationships of polymerases that belong the type B family. This review aims to discuss strengths, limitations, and opportunities of the phage surrogate with emphasis placed on its utility in the discovery and development of anti-herpetic drugs.

Keywords: DNA polymerase; T4 DNA polymerase; gp43; *herpesviridae*; UL30; UL54; HSV1; HCMV; RB69 DNA polymerase

Nomenclature

HSV1	Herpes Simplex Virus 1
HSV2	Herpes Simplex Virus 2
VZV	Varicella Zoster Virus
EBV	Epstein–Barr Virus
HCMV	Human cytomegalovirus
HHV6	Human Herpesvirus 6
HHV7	Human Herpesvirus 7
KSHV	Kaposi's sarcoma-associated herpesvirus
PFA	Phosphonoformic acid
PAA	Phosphonoacetic acid
ACV	Acyclovir
GCV	Ganciclovir
CDV	Cidofovir

1. Introduction

The eukaryotic viruses of the *herpesviridae* family are important human pathogens. In all, there are eight different human herpesviruses; Herpes simplex virus 1 and 2 (HSV1: HHV1 and HSV2: HHV2), varicella zoster virus (VZV: HHV3), Epstein-Barr virus (EBV: HHV4), Human cytomegalovirus (HCMV: HHV5), Human Herpes virus 6 and 7 (HHV6 and HHV7) and Kaposi's sarcoma-associated herpesvirus (KSHV, HHV8). Human herpesviruses cause a spectrum of diseases ranging from relatively benign cutaneous lesions to serious conditions like encephalitis and cancer. Viruses that belong to the *herpesviridae* family are characterized by their ability to establish lifelong, latent infections. Thus, a *substantial proportion of the global population is seropositive for one or more herpesviridae viruses. Although individuals with a functioning immune system can generally keep the virus suppressed, the ability to form latent infections, and the fact that the virus is widespread in the human population means that herpesvirus reactivation is a major source of disease and morbidity in immunocompromised individuals.*

The majority of approved antiviral drugs have been shown to inhibit the herpesvirus-specific DNA polymerase, reducing viral DNA replication, and, in turn, viral load [1]. However, although herpesvirus polymerases are all structurally related, they are not highly homologues. As a consequence, most drugs do not show broad antiviral activities against the various members of the *herpesviridae*. The nucleoside analog acyclovir (ACV) and its pro-drug valacyclovir are utilized to treat infection with HSV1, HSV2 or VZV, while the nucleotide analog ganciclovir (GCV) (or valganciclovir) and cidofovir (CDV) are approved to manage HCMV infection. The pyrophosphate analog phosphonoformic acid (PFA, foscarnet) provides an option to treat HSV1, HSV2, VZV and HCMV, if first-line drugs have failed to lower the viral burden. Like all current antiviral treatments,

long-term treatment can lead to the development of drug resistance. Severe side effects and complicated treatment schedules represent other problems in the management of herpesvirus infection.

Unfortunately, the development of assays to screen for novel anti-herpetic DNA polymerase inhibitors has been limited by technical problems. For the purpose of biochemical screens, *herpesviridae* DNA polymerases are difficult to overexpress in heterologous expression systems and have limited solubility. Hence, it has been difficult to characterize structural and functional details of these polymerases [2–8]. Of the eight human herpesvirus DNA polymerases, the best-studied is perhaps UL30 from HSV1. This enzyme has been characterized extensively biochemically and has been successfully crystallized [9]. Progress has also been made in characterizing HCMV UL54 [10,11]

In contrast to *herpesviridae* DNA polymerases, the orthologues enzymes of bacteriophage T4 (T4gp43) and "T4 like" bacteriophage RB69 (RB69gp43) are well studied. T4gp43 has been studied extensively using genetic, molecular biology, and biochemistry. Research into T4gp43 has been key to our current understanding of the dynamics of DNA replication [12]. RB69gp43 has been crystallized in various forms and therefore provides an important structural model for polymerases that belong to the same family [13–16]. It is here attempted to discuss the general aspects of structure and function of these related enzymes and the utility of RB69gp43 as a surrogate system for *herpesviridae* DNA polymerases in efforts to provide a better understanding of mechanisms of drug action and resistance.

2. Structure and Function of B Family Polymerases

DNA dependant DNA polymerases can be subdivided into five different families based on sequence and structural homology [17]. The DNA polymerases of bacteriophage RB69 and the *herpesviridae* are classified as B family polymerases (Figure 1a,b) [18]. B family polymerases have been identified in all domains of life and are primary involved in genome replication [19]. Unlike other polymerase families, the B family polymerases form part of a multi-subunit complexes, sometime referred to as the DNA replisome, which can co-ordinate both leading and lagging strand replication [17]. However, the polymerase catalytic activity of B family DNA polymerases is encoded by a single gene, which is sometimes referred to as the DNA polymerase catalytic subunit [20]. The catalytic subunit also often encodes an intrinsic 3'–5' exonuclease activity which provides proofreading. This substantially increases the accuracy of DNA synthesis [21,22]. The B family catalytic subunit, in the presence of the polymerase accessory proteins, is both high faithful in replicating DNA and are highly processive [12].

RB69 and each of the members of the *herpesviridae* family encode a B family polymerase (Figures 1 and 2). The virally encoded polymerase serves to replicate the viral genome. Both RB69gp43 [13] and HSV1 UL30 [9] have been studied using X-ray crystallography. Both polymerases are composed of five conserved structural domains, referred to as N-terminal, 3'–5' exonuclease, palm, fingers and thumb subdomains. In addition to these five conserved domains, the x-ray crystal structure of HSV1 UL30 showed an extra domain at the N-terminal end of the protein, which is called the pre N-terminal domain (Figure 1b).

Figure 1. (**a**) Domain structure of HSV1 UL30 (pdb 2GV9) [9]. The pre N-terminal domain is shown in white, the N-terminal domain is yellow, the exonuclease domain is red, the palm domain is magenta, the fingers domain is blue and the thumb domain is green. (**b**) The structure of the RB69gp43 apo form (pdb file 1IH7) [15]. Both structures show the fingers subdomain the open conformation. Images were generated using Pymol [23].

(a)

HSV1:UL30

(b)

RB69:gp43
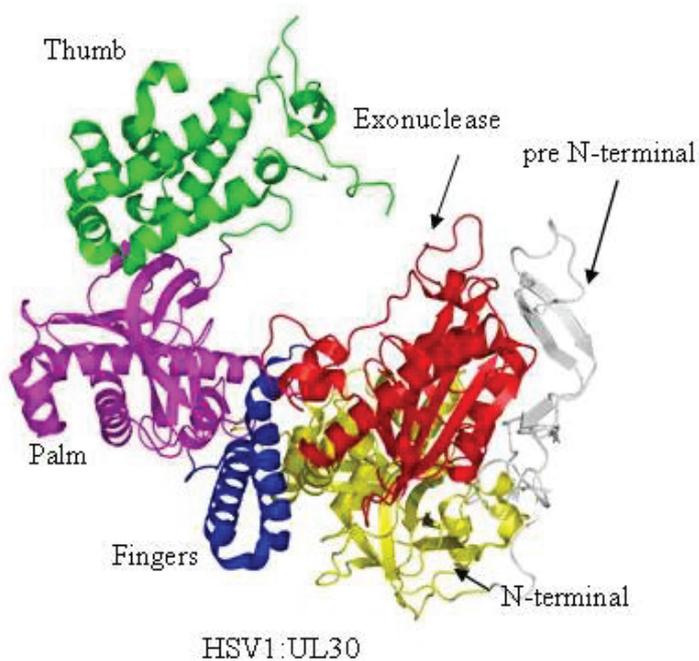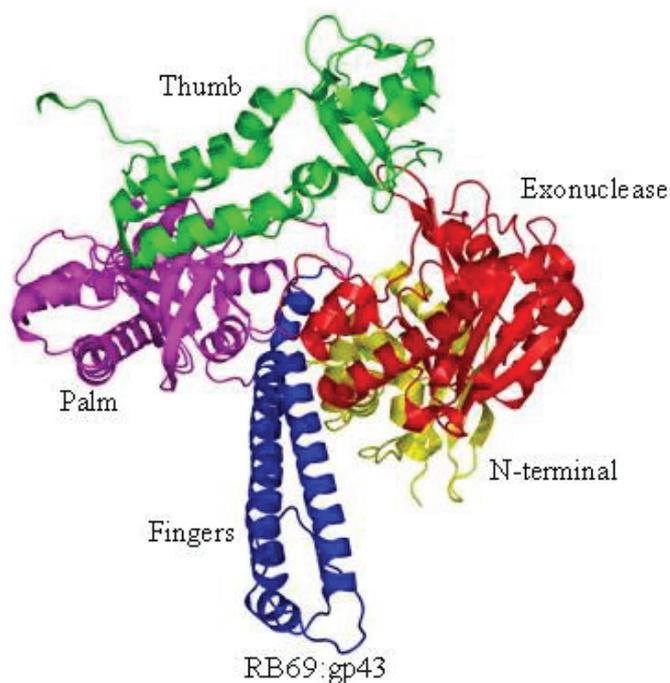
Figure 2. Protein sequence alignment of herpesvirus and bacteriophage polymerase finger domains. Herpesvirus and bacteriophage polymerase were aligned individually using the Muscle algorithm within Geneious [24,25]. The bacteriophage and herpesvirus sequences were then structurally aligned by RAPIDO [26]. Blocks above sequence highlight structural domains of polymerase. The palm domain is in pink. The fingers domain is in blue. The known conserved regions are shown in magenta blocks above sequence [27–31]. Secondary structural elements of HSV1 UL30 are indicated and are number according to Liu *et al.* (2006). Secondary structural elements of RB69 gp43 are indicated and are numbered according to Wang *et al.* (1997). Structural motifs are highlighted in sequence. N helix residues are in yellow. Motif B is in green. Mutations that have been associated with resistance to current anti-herpetic drugs are shown below the corresponding residue [32,33]. Resistance mutations are colored using the following scheme: Red: Pyrophosphate^R (Resistant), Blue: Nucleotide^R, Green: Pyrophosphate^R and Nucleotide^R. Purple: Pyrophosphate^HS (Hypersensitive), Brown: Nucleotide^R but Pyrophosphate^HS.
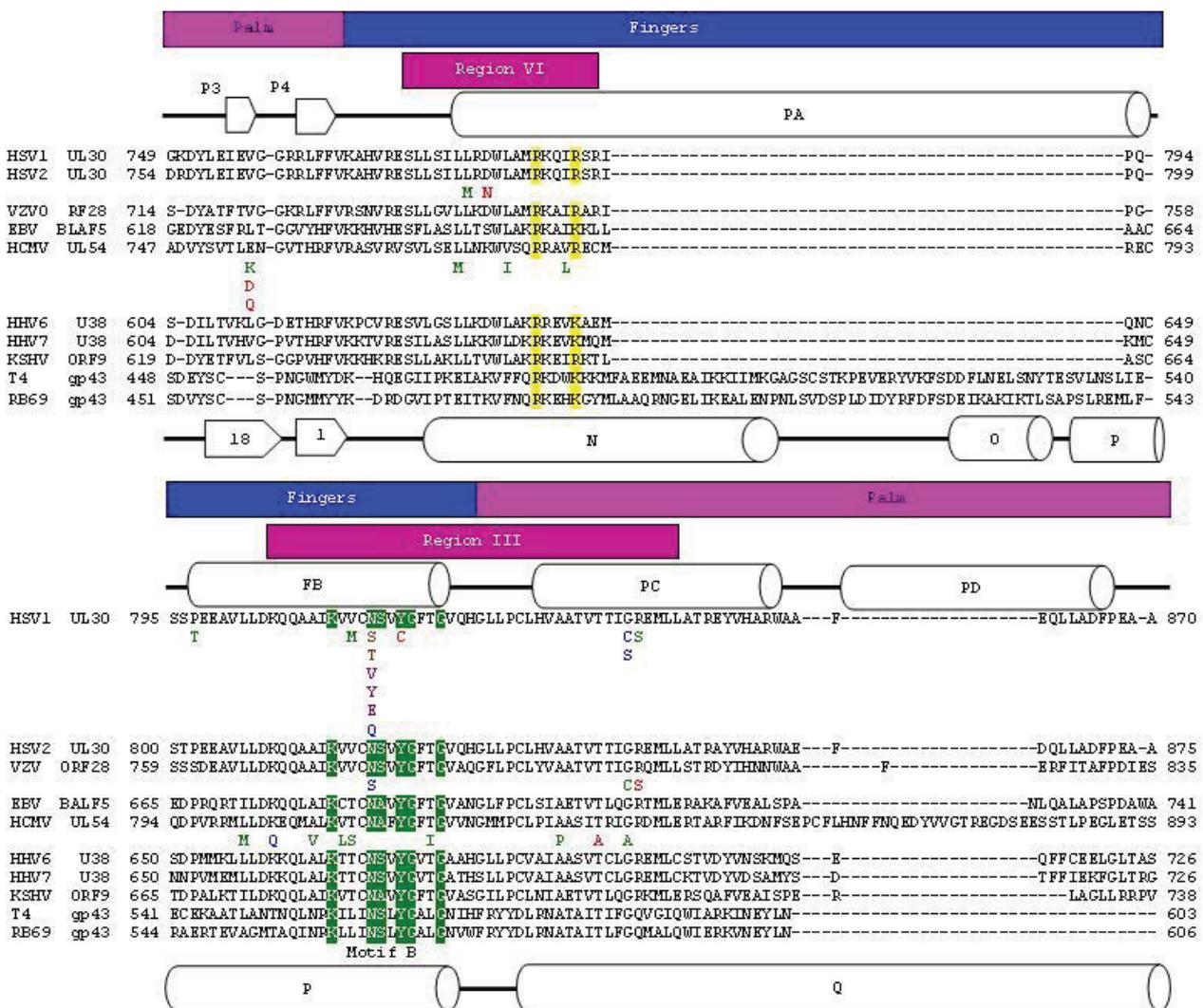
Structures of RB69gp43 in various forms in the absence and presence of substrates provided a detailed insight into distinct events involved in nucleotide incorporation, as well as the dynamics of exonuclease function. The structure of the HSV1 UL30 apo enzyme shows a similar domain structure as seen with RB69gp43; however, the exact structural requirements for DNA and nucleotide binding have yet to be established in this case. HSV1 UL30 contains the particular conserved motifs shared by all B family polymerases (Appendix Figure A1).

3. Subdomains of B Family Polymerase

Protein sequence alignments of HSV1 UL30 to other *herpesviridae* DNA polymerase show that it is likely that all Human *herpesviridae* DNA polymerases contain a pre-N terminal domain [9]. The exact function of the pre-N terminal this domain remains elusive. The pre-N terminal domain also contains the FYNPYL motif specific to *herpesviridae* family polymerases (Appendix Figure A1) [9]. It has recently been shown that this motif is required for efficient replication of viral DNA synthesis *in vivo*; a mutant polymerase lacking the FYNPYL motif showed a substantial reduction in viral DNA synthesis [34]. However, the purified FYNPYL deletion mutant showed no reduction in polymerase activity, suggesting that this motif may have a function in the formation of the viral DNA replisome.

The N-terminal domain shows a βαββαβ fold, which has been found in some RNA binding proteins [9,35]. In addition, the crystal structure of the RB69 gp43 contains a rGMP bound to the N-terminal domain [15], and some mutations in the N-terminal domain of T4gp43 decrease the expression of the polymerase leading to the suggestion that the N-terminal domain may be involved in expression regulation [36]. However, in spite of these observations, the functional role of this domain remains to be defined.

The sequence of the 3'–5' exonuclease domain of B family polymerases is not highly conserved. However, all 3'–5' exonuclease domains currently characterized adopt a ribonuclease H-like (RNase H-like) fold. The RNase H fold brings four highly conserved negatively charged residues together to form the active site. In both RB69gp43 and HSV1 UL30 these residues have been identified as three aspartic acids and a glutamic acid. These residues are essential for the binding of two divalent, catalytic metal ions. Structural elements, that harbor active site residues in RB69gp43, are referred to as exoI (D114, E116), exoII (D222), and exoIII (D327) (Table 1, Appendix Figure A1). In *herpesviridae* polymerases the equivalent regions are referred to as ExoI, region IV and Delta (δ) C, respectively. In HSV1 UL30 these residues are; ExoI, D368 and E370, region IV or ExoII, D471 and delta C or ExoIII, D581 (Table 1, Appendix Figure A1).

Figure 3. (**a**) Superpositioning of open and closed structures of RB69 gp43 showing finger domain movement. This diagram is composed of RB69 gp43 in the fingers closed position (pdb 3LDS) [37] and in the fingers opened position (pdb 1IH7) [15]. (**b**) Polymerase active site of RB69 gp43 showing interactions between conserved residues of motif A and C, metal ions A and B and dNTP and interactions between K560 and dNTP. This image is an aligned composite image of pdb 3LDS [37] and 3SCX [38]. (**c**) Structural alignment of the polymerase active site of RB69gp43 (pdb 3LDS) and HSV1 UL30 (pdb 2GV9). RB69 gp43 backbone is in light blue while the HSV1 UL30 backbone is in light orange. Active site residues of RB69 gp43 are indicated. RB69 gp43 motif A is in magenta, motif C is in orange and KKRY is in purple. HSV1 UL30 motif A is in pink, motif C is in light orange and KKRY is in light purple. Images were generated using Pymol [23]. (**d**) Generalized diagram of the polymerase catalytic cycle showing steps at which inhibitors can act. Nucleotide inhibitors, once incorporated, prevent further extension of the DNA primer, by inhibiting nucleotidyl transfer. Whereas, pyrophosphate inhibitors mimicking the pyrophosphate leaving group, stabilizing the pre-translocation complex and prevent translocation.
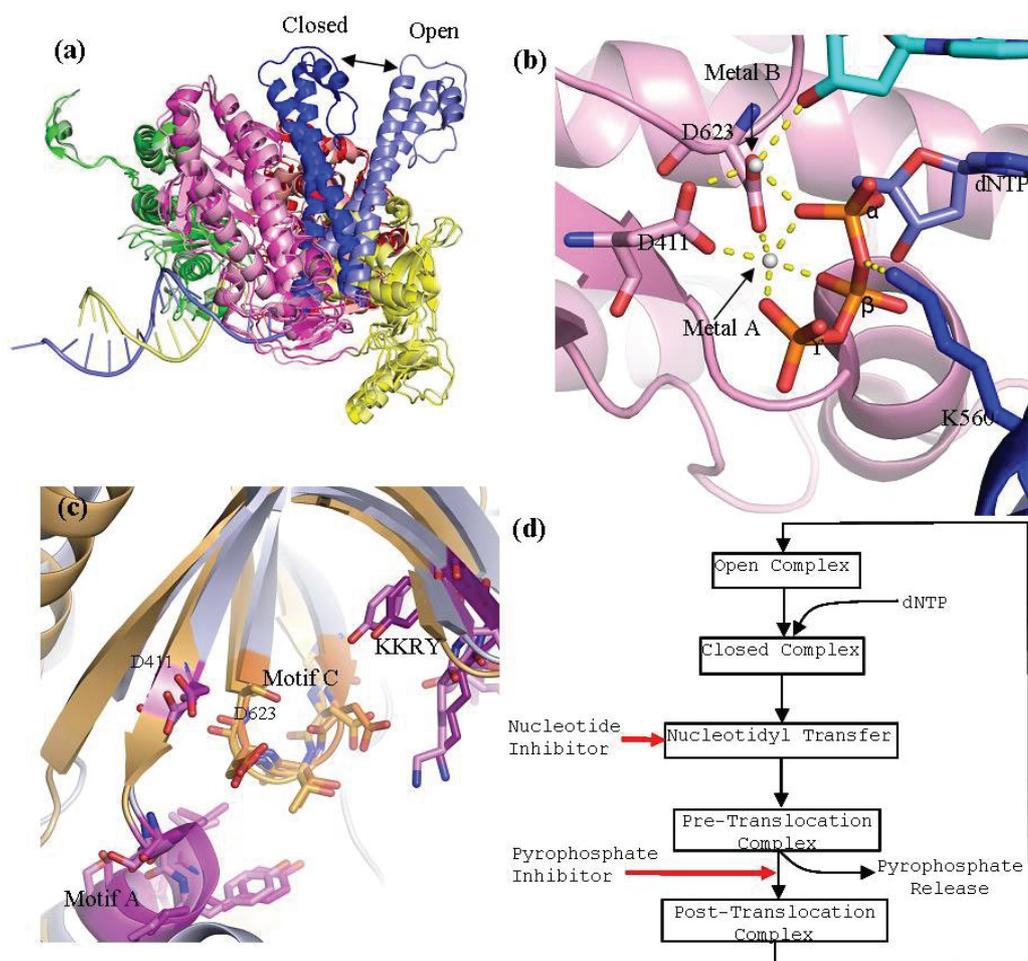
The polymerase active site of B family polymerases is made up of the three domains; the palm, fingers and thumb domains. Together they adopt the classic right hand conformation seen in all available structures of viral polymerases (Figure 1). Two highly conserved motifs in the palm domain, one in the fragment of the palm domain prior to the fingers domain called motif A and one in the fragment post fingers domain called motif C (Table 1, Appendix Figure A1) are likewise seen as signature motifs. In *herpesviridae* polymerases these domains are referred to as region II and region I, respectively. These motifs are; motif A; DXXLYPS and motif C; DTDS (Table 1, Appendix Figure A1). Structurally, these motifs fold together to form a three-strand anti-parallel β sheet (Figure 3c) [13]. Two conserved aspartic acid residues, D411 from motif A and the D625 from motif C, are required to form critical interactions which help co-ordinate the two divalent metal cation that are critical for DNA polymerization (Figure 3b) [15]. In addition, to motif A and C there are two other conserved motifs which also form parts the active site, motif B (KXXXNSXYG), which is known as Region III in *herpesviridae*, which is located on the helix P of the fingers domain, and the KKRY motif which is in the palm domain sequentially after motif C (Appendix Figure A1).

Table 1. Comparison of size, weight and position of conserved motifs or active site residues of RB69 and *herpesviridae* DNA polymerases. The position conserved motifs where assigned based on alignments generated using geneious [24].

Virus (Gene)	Amino Acids (aa)	Weight (kDa)	FYNPYL motif	Exo 1	Exo2/ region IV	Exo3/ delta (ä) C	Motif A/ region II	Motif B/ region III	Motif C/ region I	KKRY motif
RB69 (gp43)	903	104.47	N/A	113–117	222	327	411–420	560–571	621–624	804–807
HSV1 (UL30)	1236	136.42	167–173	367–371	471	581	717–726	811–822	886–889	938–941
HSV2 (UL30)	1241	137.32	166–172	368–372	472	582	722–731	816–827	891–894	943–946
VZV (Orf28)	1195	134.05	7–13	348–352	452	562	682–691	775–786	851–854	903–906
EBV (BALF5)	1016	113.43	6–12	295–299	384	497	584–593	681–692	755–758	807–810
HCMV (UL54)	1242	137.21	2–8	300–304	413	542	717–726	811–822	910–913	962–965
HHV6 (UL38)	1013	115.67	6–12	281–285	369	482	572–581	666–677	740–743	792–795
HHV7 (UL38)	1014	115.91	6–12	280–284	368	480	572–581	666–677	740–743	792–795
KSHV (Orf9)	1013	113.33	3–9	295–299	383	498	585–594	681–692	752–755	804–807

Residues K560 and N564 of motif B are important in coordinating the tri-phosphate tail of the incoming dNTP during catalysis [15]. Lysine 560 also serves as a proton donor during catalysis (Figure 3b) [39]. Residue Y567 of motif B has been shown to be involved in forming an important interaction with the minor groove of the DNA [40]. This interaction has been shown to be highly important in maintaining polymerase fidelity [40]. In addition to motif B of helix P, there are also conserved positively charged residues within helix N (R482 and K486) that form important interactions with the tri-phosphate tail of the incoming dNTP during catalysis.

The KKRY motif is primarily involved in stabilizing the B form of DNA. Residue Y708 forms a hydrogen bond with the 3' terminus of the primer while the K705 and R707 form interactions with the

phosphate backbone. Together, these interactions help stabilize the interactions between base pairs of the primer and template strands [15].

4. Catalytic Cycle of B Type Polymerase RB69gp43

Both RB69gp43 and HSV1 UL30 have been crystallized in the so-called open conformation [9,13], but only RB69 has been crystallized in the closed conformation [15]. The closed conformation of a ternary complex contains a DNA primer-template pair and a trapped nucleotide. The most striking difference between the open and closed formation of RB69gp43 is the movement of the fingers domain. Structural alignment of the open and close conformation of RB69gp43 shows that, upon dNTP binding, the fingers domain rotates 60° inwards relative to the palm domain, with the tip of the fingers moving approximately 30 Å (Figure 3a) [15]. Closure of the fingers domain moves the residues on helix N and P that are involved in binding the tri-phosphate tail of the incoming dNTP, 4–8 Å closer to the polymerase active site. This action traps the dNTP in the active site and allows the nucleotidyl transfer to take place. By contrast, the overall structure of the thumb and palm domain between the open and closed conformation remains relatively unchanged. The thumb domain moves approximately 8° toward the palm domain. This action wraps the minor groove of primer-template duplex [15].

B family polymerases employ a two divalent metal ion mechanism for the nucleotidyl transfer [41], in conjunction with a concurrent two proton transfer reaction [39,42]. The catalytic ions are bound by D411 of motif A and D623 of motif C, and form an extensive network of interactions with the dNTP aligning it in the correct orientation for polymerization (Figure 3b). Metal ion A is required to activate the 3'-OH of the primer terminus. Interaction between metal ion A and the primer terminus attract the primer terminus closer to the α-phosphorus atom of the incoming dNTP. This lowers the pKa of the 3'-OH group allowing it to be deprotonated, which facilitates the nucleophilic attack on the α-phosphorous atom of the nucleotide substrate [43]. Metal ion B orientates the dNTP triphosphate tail and helps stabilizes the transition states; it has also been suggested that it assists in pyrophosphate release [38]. In RB69gp43, lysine 560 acts as a proton donor [42] and is required to protonates the pyrophosphate leaving group, which may facilitate its release from the complex. This step formally ends the catalysis, leaving the complex now in the pre-translocational state. The fingers of the polymerase rotate away from the active site allowing the release of the pyrophosphate, which allows the DNA substrate to translocate relative to the enzyme. This movement shifts the new 3'-OH terminus into the 1+ position forming a post-tranlocated complex with the polymerase reset for a new catalytic cycle. Polymerases are in general able to discriminate between correct and incorrect nucleotides [44]. Incorrect binding of a nucleotide destabilizes the closed ternary complex, which enables the fingers domain to return to the open form releasing the incorrect nucleotide. Effective discrimination against the incorrect nucleotide at the level of substrate binding raises the fidelity of DNA synthesis significantly[45]. However, in B family polymerases, if an incorrect nucleotide is indeed incorporated, these enzymes can switch into the 3'–5' exonuclease mode and remove the misincorporated base. The

3'–5' exonuclease activity has been shown to increases the fidelity of the RB69 polymerase from the Exo$^-$ rate of 2.8 errors per genome (μ_g) to the Exo$^+$ error rate of 4×10^{-3} μ_g [40,46]. The excision of an incorrectly incorporated terminal nucleotide is also dependant on two divalent metal ions. Structures of RB69gp43 have shown that the exonuclease active site is located approximately 40 Å from the polymerase active site [47]. Thus, for 3'–5' exonuclease activity to occur the DNA primer-template terminus must be translocated from the polymerase active site to the exonuclease site. The details of this process remain to be defined. It has been suggested that DNA replication accessory proteins, particularly the sliding clamp gp45, may be involved in the process of translocating DNA from polymerase to exonuclease active site [15]. It has also been shown that a β hairpin loop between residues 251–262 is important for exonuclease function [48,49]. This loop is involved in stabilizing the frayed base pair at the exonuclease active site allowing the removal of the incorrect nucleotide. During translocation from the polymerisation active site to the exonuclease active site the primer-template pair is partially melted, producing three unpaired bases. The three unpaired bases of the primer strand are then sequestered into the exonuclease active site, which facilitates the excision. Structures of RB69gp43 poised with its primer-template in the exonuclease mode are available [47,48].

5. Base Selectivity in RB69 DNA Polymerase

A large number of mutant RB69 and T4 DNA polymerases that affect both efficiency and fidelity of DNA replication have been isolated and characterized. In the case of RB69, X-ray crystallography has firmly established the nucleotide binding site at the atomic level. Crucial residues include K560 (Motif B), L415 and Y416 (Motif A) and L561 (Not conserved), Y567 and G568 (Motif B) [50].

Conserved residue Motif A has been shown to be important for sugar selectivity. Y416 forms a stacking interaction with deoxyribosyl moiety of the incoming dNTP [15]. It has been proposed that the 2'-OH group of a mismatched rNTP would cause a steric conflict with Y416 preventing formation of a stable complex [51]. Biochemical experiments with the Y416A mutant enzyme corroborated this notion. Unlike the wild type, the mutant was able to incorporate rCTP, ddCTP, and dCTP at similar rates [51].

Residue Y567 of motif B has been shown to be important in maintaining the fidelity of base selection [40,52]. This residue forms an interaction—via a water molecule—with the minor groove of the terminal primer-template pair. This interaction is important for sensing the geometry of the newly form base pair and thus detecting distortions caused by incorrect base pairing [15]. Mutations to residue 567 increase the size of the nascent base-pair-binding pocket allowing the misincorporation of nucleotides.

Residue L561 protrudes into the major groove of the templating base. It has been proposed that this residue is involved in detecting mismatches that lead to distortion of the major groove. L561A mutant confers a mutator phenotype [53]. Interestingly, the equivalent residue in herpesvirus-associated polymerases are not conserved (Figure 4b).

Figure 4. (**a**) Diagram of RB69 Fingers domain showing location of ABC block mutation relative to dNTP binding residues. Motif B residues are in orange, Tri-phosphate interacting residues on Helix N are in yellow, Block A is in red, Block B is in dark blue and Block C is in yellow. Aligned image of pdb 3LDS (dNTP) and pdb 3KD5 (N and P helix). (**b**) Sequence alignment of RB69 and *herpesviridae* sequences showing location of block mutations. (**c**) Diagram showing clash between W478 of Block A and W365. RB69 ABC5 is in white (3KD5), RB69 WT is in blue (pdb 1IH7) and HSV1 UL30 is in orange (pdb 2GV9). (**d**) Diagram of RB69 ABC5 block mutations active site showing phosphonoformic acid binding in β and γ phosphate position and with acyclovir in the pre-translocation position (pdb 3KD5). Images were generated using Pymol [23]. Alignment was generated using Geneious [24].
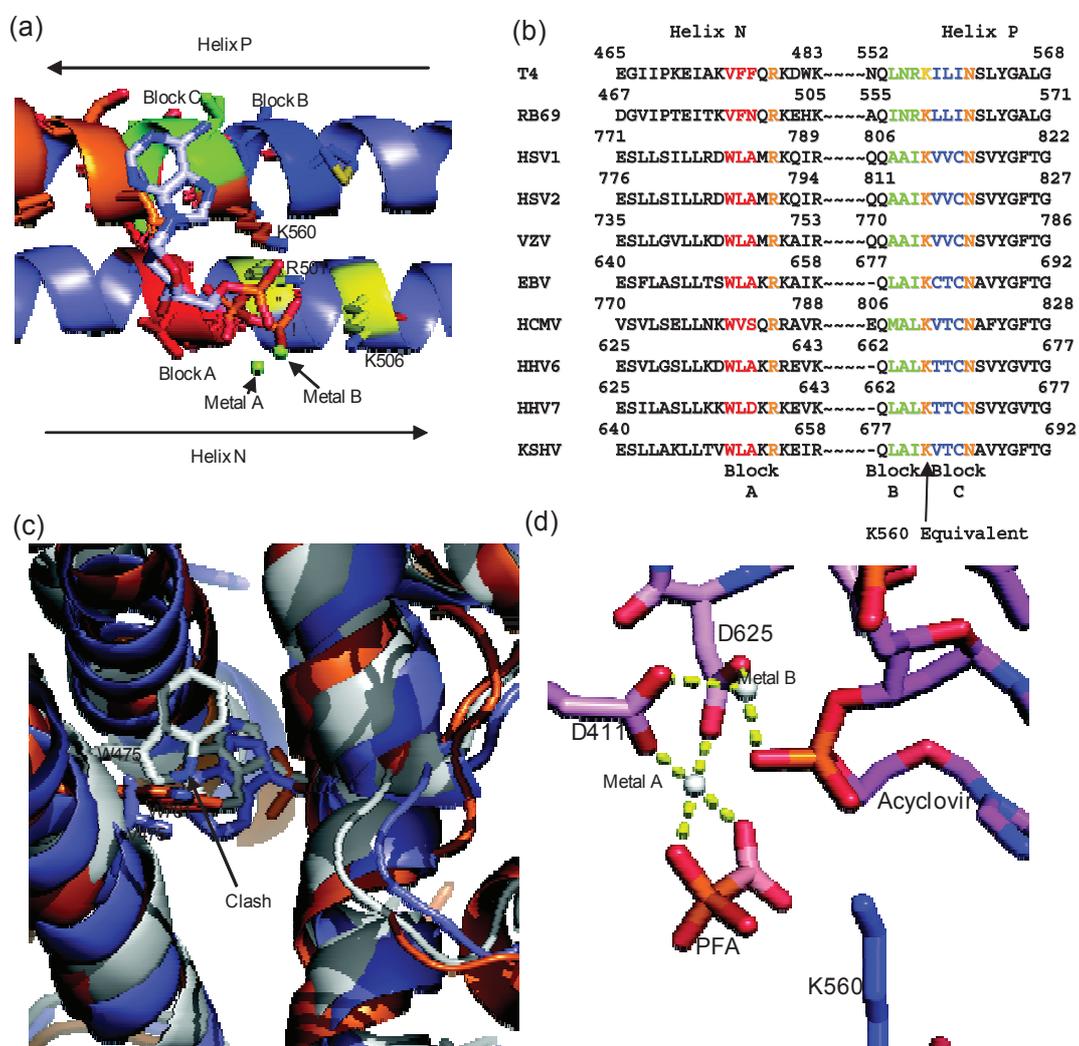
6. The Active Site of HSV1 UL30 and RB69gp43

At both the nucleotide and amino acid level, *herpesviridae polymerases* are not highly conserved (Appendix Figure A1). They range in length from 1013–1236 aa, and range in molecular weight from 113 to 137 kDa (Table 1). RB69gp43 is relatively small with 903 aa in length and a molecular weight of 104 kDa (Table 1). In terms of homology, there is very low homology between RB69gp43 and *herpesviridae* DNA polymerases sequences. However, both RB69gp43 and *herpesviridae* contain all the conserved motifs associated with B family polymerases (Appendix Figure A1).

Unlike herpesvirus DNA polymerases, there is a wealth of information on the structure and function of RB69gp43. Thus, this enzyme is often used as a model for herpesviridae DNA polymerase. Because of the shared conserved motifs between B family polymerase it is reasonable to assume that both herpesviridae and RB69gp43 have a similar, if not an identical catalytic mechanisms. The residues of motif A and C in the HSV1 UL30 and RB69gp43 structure are superimposable, supporting the notion that the function of these residues is similar (Figure 3c). In contrast to the metal cation-binding portion of the active site, the residues of helix N and P of the fingers domain vary greatly between HSV1 UL30 and RB69 gp43. On helix N, the only conserved residues are two basic amino acids R482 and K486 in RB69 gp43, and R785 and R789 in HSV1 UL30. Likewise, on the helix P, the only conserved residues are those of motif B: K560 to N565 in RB69gp43 and K811 to N815 in HSV1 UL30. Thus, the non-conserved residues of helix N and P account for the major differences in the nucleotide binding site, which could in turn account for the difference in sensitivity to antiviral drugs when herpesviridae polymerases are compared with RB69gp43. This notion is supported by biochemical studies with mutant enzymes derived from HCMV UL54 (see below) [11].

7. Mutations of Bacteriophage T4 that Induce PAA Sensitivity

There are several known natural mutations of bacteriophage T4 DNA polymerase (T4gp43) that affect sensitivity to the pyrophosphate analog phosophonoacetic acid (PAA). T4gp43 is highly homologues to RB69gp43, and, like RB69gp43, is naturally resistant to PAA. However, mutants of bacteriophage T4 with reduced plaque formation in the presence of PAA have been identified [54,55]. One of these mutations, L412M, is located in the conserved motif A. L412M confers increased sensitivity to PAA and interestingly also confers mutator properties to the polymerase [54]. Another interesting feature of the L412M mutant phage is that it can replicate in *E. coli* strains with restricted dGTP pools (*optA1*). The ability to grow under these conditions suggests that the mutant polymerase can make more efficient use of the nucleotide substrates. Because fidelity of DNA synthesis is in part controlled by the rate with which a frayed primer switches from polymerase to exonuclease activity, a bias toward stabilizing the frayed primer in the polymerase active site would cause an overall reduction in enzyme fidelity. When this hypothesis was tested with purified T4gp43 L412M enzyme, it was found that T4gp43 L412M exhibited less exonulease activity in relation to the wild type supporting the notion that the L412M mutation was in fact changing the partitioning between

exonuclease and polymerase activity away from exonuclease activity. The equivalent residue to L412 in RB69gp43 is L415. Mutations at this residue show increased rates of mis-incorporation [56]. Interestingly, several suppressor mutants of L412M were also isolated. When subjected to a similar analysis, they were found to be antimutator mutations and unable to grow on *E. coli optA1* [55]. Mutations R335C and S345F are both located within the delta C region of the exonuclease domain (Appendix Figure A1). It has been proposed, that these mutation may affect pyrophosphate sensitivity by increasing the opportunity for exonuclease activity [54]. Since the primer terminus needs to be physically transferred to the exonuclease site, any increase in stability of the primer-terminus in the exonuclease active site or decrease of stability of the primer-terminus in the polymerase active site would increase the disruption of the pyrophosphate analog inhibited complex and henceforth increase the polymerase resistance to inhibition by pyrophosphate analogs. Interestingly in the RB69gp43 crystal structure the equivalent residue to R335 is R338, which is positioned at the very C terminus of the exonuclease domain and points into the cleft in which the finger domain rotates into during catalysis. Thus, the R335 mutation may not directly affect exonuclease function, but may instead inhibit finger domain movement [57].

8. Chimeric Enzymes

Tchesnokov *et al.* (2009) engineered an RB69gp43-UL54 chimeric enzyme, by mutating the active site of RB69gp43 to include the non-conserved elements from helix N and P of the HCMV enzyme. Swapping the polymerase active sites produced an enzyme that can be expressed in *E. coli*, and is soluble and easily purified. Biochemical assays have shown that the chimeric enzyme is sensitive to the nucleotide inhibitor acyclovir and the pyrophosphate analog PFA [58].

Three blocks of non-conserved amino acid residues were considered to engineer the chimera. Block A is located on helix N, and consists of residues 478–480 from RB69gp43 (VFN), these residue were replaced with equivalent residues of HCMV UL54: residues 779 to 781 (WVS). Block B and Block C are both located on Helix P. Block B consist of residues 557–559 of RB69gp43 (INR) and Block C 561–563 (LLI) of RB69gp43 where replaced with residues 808–810 (MAL) and 812–814 (VTC) from HCMV UL54, respectively. Block B and C flank residue K560, which is the conserved basic amino acid that likely donates a proton to the pyrophosphate leaving group [39,42]. Previous studies had already shown that several amino acids within this region can affect sensitivity to PFA (Figure 2) [11]. The chimeric polymerase reproduces drug sensitive and drug resistant phenotypes in cell-free biochemical assays, which validates this enzyme as a model system for polymerase active site inhibitors.

The structure of the chimera provides a detailed understanding of the mechanism of action of PFA [59]. The enzyme was co-crystallized in complex with a primer-template terminated with acyclovir in the presence and absence of PFA. PFA is bound at the polymerase active site and traps the enzyme in the pre-translocational state. The compound interacts with metal ion B and residue R482 of helix N, similar to the interactions formed by the β- and γ-phosphate of a bound dNTP in the post-translocational conformation. It appears that W478 of block A is critical in mediating sensitivity

to PFA. Although no additional contacts are formed, this residue likely reduces the population of complexes that exist in the open conformation due to steric interference with W365 on helix J. Generally binary complex structures of B family polymerase are found to be in the open conformation; however, in the case of the chimeric enzyme the fingers are in the closed conformation even in the absence of PFA [59]. The predicted steric clash was confirmed with enzymes containing amino acid substitutions at residues 478 and 365 [59].

9. Resistance to Antiviral Drugs

Resistance-conferring mutations in the polymerase of HSV1, HSV2, VZV and HCMV UL54 have been identified *in vivo* and *in vitro* [10,32,33,60]. Drug resistance is measured as an increase in the inhibitory concentrations of a given drug required to block 50% viral replication (IC_{50}). In HSV1, HSV2 and VZV an $IC_{50} \geq 4.4$ μM confers significant levels of resistance to acyclovir [32], and in HCMV an $IC_{50} \geq 12$μM confers significant levels of resistance to ganciclovir, and $IC_{50} \geq 400$ μM to PFA [60]. Because of the difficulties of purifying Herpesvirus DNA polymerase the many resistance mutation have only been assessed in cell-based phenotypic assays [60].

All known resistance mutations can be roughly divided into two groups: 1. mutations within the exonuclease domain, which may affect 3'–5' exonuclease function, and 2. mutations within the domains that make up the polymerase active site, which may therefore affect polymerase function directly (Appendix Figure A1). Within the 3'–5' exonuclease domain there are several resistance mutations around the active site residues exo 1, exo2 and exo 3 respectively (Appendix Figure A1). Many of these mutations have been characterized to impair 3'–5' exonuclease activity [61,62]. Interestingly some of these resistance mutations have been characterized to confer resistance to pyrophosphate inhibitors but hypersensitivity to nucleotide inhibitors. An example of this phenotype is the HSV1 mutations Y577H and D581A [62] (Appendix Figure A1). Both mutations are located within the delta C region close to the exo 3 residue. Both mutations have been shown to impair exonuclease activity [62]. Logically a polymerase impaired in 3'–5' exonuclease activity would be unable to remove an incorporated nucleotide inhibitor from viral DNA, which would increase the viruses sensitivity to nucleotide inhibitor, thus explaining the hypersensitivity phenotype. However, the mutations also affect pyrophosphate inhibitors potency and this effect is harder to reconcile. Since pyrophosphate inhibitors mimic the pyrophosphate leaving group, and thus are not incorporated into DNA, the presence or absence of exonuclease activity should not directly affect pyrophosphate analog inhibitors susceptibility. The analysis of T4gp43 PAA sensitive mutants provides some possible insight into a mechanism [54,55]. While working with the T4gp43 L412M mutant they identified several suppressors of PAA sensitivity [54]. These suppressor mutants were shown to be antimutator polymerases, implying that L412M suppressor mutant polymerase were likely to have an altered rate of exonuclease activity compared to the L412M mutants. The authors suggested that because pyrophosphate analogs competitively inhibit polymerase activity by mimicking the pyrophosphate leaving group, that the transition to exonuclease activity could potentially bypass inhibition. Thus,

altering the rate of exonuclease activity could potentially affect pyrophosphate inhibitor potency. Unfortunately this hypothesis has not fully been tested in either T4gp43 or herpesviridae DNA polymerases.

Resistance mutations within the polymerase active site can be arbitrarily split into two groups: 1. Fingers domain mutations and palm domain mutations. In HSV1, HSV2, VZV and HCMV several resistance mutations within the finger domain associated with region VI within helix PA and region III within helix PB. Helix PA and PB are equivalent to helix N and P of RB69gp43 (Figure 2) [32,33]. These two regions make up the finger domain contribution to the polymerase active site. Helix N contains several conserved residues important for nucleotide binding, while Helix P contains conserved residue K560 that is required for proton transfer during catalysis and N564, which is required for nucleotide binding. Crystal structures of RB69gp43 have shown that residue N564 interacts with the β phosphate of the nucleotide via a water molecule [52,63]. There are also several resistance mutations associated with the motif A, C and KKRY within the palm domain (Appendix Figure A1). In Motif A or region II there are several mutations prior to Motif A, which cause a pyrophosphate inhibitor hypersensitive, nucleotide inhibitor resistance phenotype. Mutations in motif C or region I and the KKRY motif or region VII (Appendix Figure A1) have all been characterized as inducing both pyrophosphate and nucleotide inhibitor resistance. Being that motif C and the KKRY motif is involved in aligning the 3'-OH nucleophile during catalysis any subtle change to the positioning of these motifs could change the binding and catalytic constants of the polymerase.

Several mutations remain to be confirmed as resistance-conferring amino acid substitutions [33]. Because of the difficulties in working with herpesviridae these mutations have not been tested in a defined genetic background to determine the phenotypic effect on resistance or susceptibility.

10. Conclusions

RB69gp43 provides an excellent model for the study of structure–function relationships of B family polymerases. However, there are limitations for the study of orthologous *herpesviridae* polymerases. Most importantly, the phage enzyme is not inhibited by approved drugs that bind to the polymerase active site in either post- or pre-translocational states. Chimeric enzymes composed of a RB69gp43 backbone and important elements of the active site of *herpesviridae* DNA polymerases can potentially address this problem. These findings warrant further investigation in such enzymes as novel tools in future drug discovery and development efforts.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Elion, G.B.; Furman, P.A.; Fyfe, J.A.; de Miranda, P.; Beauchamp, L.; Schaeffer, H.J. Selectivity of action of an antiherpetic agent, 9-(2-hydroxyethoxymethyl) guanine. *Proc. Natl. Acad. Sci. USA* **1977**, *74*, 5716–5720.
2. Ertl, P.F.; Thomas, M.S.; Powell, K.L. High level expression of DNA polymerases from herpesviruses. *J. Gen. Virol.* **1991**, *72*, 1729–1734.
3. Tsurumi, T.; Kobayashi, A.; Tamai, K.; Daikoku, T.; Kurachi, R.; Nishiyama, Y. Functional expression and characterization of the epstein-barr virus DNA polymerase catalytic subunit. *J. Virol.* **1993**, *67*, 4651–4658.
4. Lin, J.C.; De, B.K.; Mar, E.C. Functional characterization of partially purified epstein-barr virus DNA polymerase expressed in the baculovirus system. *Virus Genes* **1994**, *8*, 231–241.
5. Tsurumi, T.; Daikoku, T.; Nishiyama, Y. Further characterization of the interaction between the epstein-barr virus DNA polymerase catalytic subunit and its accessory subunit with regard to the 3'-to-5' exonucleolytic activity and stability of initiation complex at primer terminus. *J. Virol.* **1994**, *68*, 3354–3363.
6. De Bolle, L.; Manichanh, C.; Agut, H.; De Clercq, E.; Naesens, L. Human herpesvirus 6 DNA polymerase: Enzymatic parameters, sensitivity to ganciclovir and determination of the role of the a961v mutation in hhv-6 ganciclovir resistance. *Antivir. Res.* **2004**, *64*, 17–25.
7. Picard-Jean, F.; Bougie, I.; Bisaillon, M. Characterization of the DNA- and dntp-binding activities of the human cytomegalovirus DNA polymerase catalytic subunit ul54. *Biochem. J.* **2007**, *407*, 331–341.
8. Dorjsuren, D.; Badralmaa, Y.; Mikovits, J.; Li, A.Q.; Fisher, R.; Ricciardi, R.; Shoemaker, R.; Sei, S. Expression and purification of recombinant kaposi's sarcoma-associated herpesvirus DNA polymerase using a baculovirus vector system. *Protein Expr. Purif.* **2003**, *29*, 42–50.
9. Liu, S.; Knafels, J.D.; Chang, J.S.; Waszak, G.A.; Baldwin, E.T.; Deibel, M.R., Jr.; Thomsen, D.R.; Homa, F.L.; Wells, P.A.; Tory, M.C.; *et al.* Crystal structure of the herpes simplex virus 1 DNA polymerase. *J. Biol. Chem.* **2006**, *281*, 18193–18200.
10. Ducancelle, A.; Gravisse, J.; Alain, S.; Fillet, A.M.; Petit, F.; Pors, M.J.; Mazeron, M.C. Phenotypic characterisation of cytomegalovirus DNA polymerase: A method to study cytomegalovirus isolates resistant to foscarnet. *J. Virol. Meth.* **2005**, *125*, 145–151.
11. Tchesnokov, E.P.; Gilbert, C.; Boivin, G.; Gotte, M. Role of helix p of the human cytomegalovirus DNA polymerase in resistance and hypersusceptibility to the antiviral drug foscarnet. *J. Virol.* **2006**, *80*, 1440–1450.

12. Mueser, T.C.; Hinerman, J.M.; Devos, J.M.; Boyer, R.A.; Williams, K.J. Structural analysis of bacteriophage t4 DNA replication: A review in the virology journal series on bacteriophage t4 and its relatives. *Virol. J.* **2010**, *7*, 359.

13. Wang, J.; Sattar, A.K.; Wang, C.C.; Karam, J.D.; Konigsberg, W.H.; Steitz, T.A. Crystal structure of a pol alpha family replication DNA polymerase from bacteriophage rb69. *Cell* **1997**, *89*, 1087–1099.

14. Yang, G.; Lin, T.; Karam, J.; Konigsberg, W.H. Steady-state kinetic characterization of rb69 DNA polymerase mutants that affect dntp incorporation. *Biochemistry* **1999**, *38*, 8094–8101.

15. Franklin, M.C.; Wang, J.; Steitz, T.A. Structure of the replicating complex of a pol alpha family DNA polymerase. *Cell* **2001**, *105*, 657–667.

16. Hogg, M.; Wallace, S.S.; Doublie, S. Crystallographic snapshots of a replicative DNA polymerase encountering an abasic site. *EMBO J.* **2004**, *23*, 1483–1493.

17. Garcia-Diaz, M.; Bebenek, K. Multiple functions of DNA polymerases. *Crit. Rev. Plant Sci.* **2007**, *26*, 105–122.

18. Braithwaite, D.K.; Ito, J. Compilation, alignment, and phylogenetic relationships of DNA polymerases. *Nucleic Acids Res.* **1993**, *21*, 787–802.

19. Garg, P.; Burgers, P.M. DNA polymerases that propagate the eukaryotic DNA replication fork. *Crit. Rev. Biochem. Mol. Biol.* **2005**, *40*, 115–128.

20. De Waard, A.; Paul, A.V.; Lehman, I.R. The structural gene for deoxyribonucleic acid polymerase in bacteriophages t4 and t5. *Proc. Natl. Acad. Sci. USA* **1965**, *54*, 1241–1248.

21. Goulian, M.; Lucas, Z.J.; Kornberg, A. Enzymatic synthesis of deoxyribonucleic acid. Xxv. Purification and properties of deoxyribonucleic acid polymerase induced by infection with phage t4. *J. Biol. Chem.* **1968**, *243*, 627–638.

22. Kornberg, A. Active center of DNA polymerase. *Science* **1969**, *163*, 1410–1418.

23. Schrodinger, L.L.C. *The Pymol Molecular Graphics System*, version 1.3r1; Schrödinger K.K.; http://www.pymol.org/, 2010.

24. Drummond, A.; Ashton, B.; Buxton, S.; Cheung, M.; Cooper, A.; Duran, C.; Field, M.; Heled, J.; Kearse, M.; Markowitz, S.; *et al. Geneious v5.4*; Biomatters Ltd; http://www.geneious.com/, 2010

25. Edgar, R.C. Muscle: Multiple sequence alignment with high accuracy and high throughput./Zh *Nucleic Acids Res.* **2004**, *32*, 1792–1797.

26. Mosca, R.; Brannetti, B.; Schneider, T.R. Alignment of protein structures in the presence of domain motions. *BMC Bioinformatics* **2008**, *9*, 352.

27. Hwang, C.B.; Ruffner, K.L.; Coen, D.M. A point mutation within a distinct conserved region of the herpes simplex virus DNA polymerase gene confers drug resistance. *J. Virol.* **1992**, *66*, 1774–1776.

28. Wong, S.W.; Wahl, A.F.; Yuan, P.M.; Arai, N.; Pearson, B.E.; Arai, K.; Korn, D.; Hunkapiller, M.W.; Wang, T.S. Human DNA polymerase alpha gene expression is cell proliferation dependent and its primary structure is similar to both prokaryotic and eukaryotic replicative DNA polymerases. *EMBO J.* **1988**, *7*, 37–47.

29. Zhang, J.; Chung, D.W.; Tan, C.K.; Downey, K.M.; Davie, E.W.; So, A.G. Primary structure of the catalytic subunit of calf thymus DNA polymerase delta: Sequence similarities with other DNA polymerases. *Biochemistry* **1991**, *30*, 11742–11750.

30. Simon, M.; Giot, L.; Faye, G. The 3' to 5' exonuclease activity located in the DNA polymerase delta subunit of saccharomyces cerevisiae is required for accurate replication. *EMBO J.* **1991**, *10*, 2165–2170.

31. Blanco, L.; Bernad, A.; Blasco, M.A.; Salas, M. A general structure for DNA-dependent DNA polymerases. *Gene* **1991**, *100*, 27–38.

32. Gilbert, C.; Bestman-Smith, J.; Boivin, G. Resistance of herpesviruses to antiviral drugs: Clinical impacts and molecular mechanisms. *Drug Resist. Updates* **2002**, *5*, 88–114.

33. Lurain, N.S.; Chou, S. Antiviral drug resistance of human cytomegalovirus. *Clin. Microbiol. Rev.* **2010**, *23*, 689–712.

34. Terrell, S.L.; Coen, D.M. The pre-nh2-terminal domain of the herpes simplex virus 1 DNA polymerase catalytic subunit is required for efficient viral replication. *J. Virol.* **2012**, *86*, 11057–11065.

35. Burd, C.G.; Dreyfuss, G. Conserved structures and diversity of functions of rna-binding proteins. *Science* **1994**, *265*, 615–621.

36. Hughes, M.B.; Yee, A.M.; Dawson, M.; Karam, J. Genetic mapping of the amino-terminal domain of bacteriophage t4 DNA polymerase. *Genetics* **1987**, *115*, 393–403.

37. Hogg, M.; Rudnicki, J.; Midkiff, J.; Reha-Krantz, L.; Doublie, S.; Wallace, S.S. Kinetics of mismatch formation opposite lesions by the replicative DNA polymerase from bacteriophage rb69. *Biochemistry* **2010**, *49*, 2317–2325.

38. Xia, S.; Wang, M.; Blaha, G.; Konigsberg, W.H.; Wang, J. Structural insights into complete metal ion coordination from ternary complexes of b family rb69 DNA polymerase. *Biochemistry* **2011**, *50*, 9114–9124.

39. Castro, C.; Smidansky, E.D.; Arnold, J.J.; Maksimchuk, K.R.; Moustafa, I.; Uchida, A.; Gotte, M.; Konigsberg, W.; Cameron, C.E. Nucleic acid polymerases use a general acid for nucleotidyl transfer. *Nat. Struct. Mol. Biol.* **2009**, *16*, 212–218.

40. Bebenek, A.; Dressman, H.; Carver, G.; Ng, S.; Petrov, V.; Yang, G.; Konigsberg, W.; Karam, J.; Drake, J. Interacting fidelity defects in the replicative DNA polymerase of bacteriophage rb69. *J. Biol. Chem.* **2001**, *276*, 10387–10397.

41. Steitz, T.A. A mechanism for all polymerases. *Nature* **1998**, *391*, 231–232.

42. Castro, C.; Smidansky, E.; Maksimchuk, K.R.; Arnold, J.J.; Korneeva, V.S.; Gotte, M.; Konigsberg, W.; Cameron, C.E. Two proton transfers in the transition state for nucleotidyl transfer catalyzed by rna- and DNA-dependent rna and DNA polymerases. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 4267–4272.

43. Fothergill, M.; Goodman, M.F.; Petruska, J.; Warshel, A. Structure-energy analysis of the role of metal-ions in phosphodiester bond hydrolysis by DNA-polymerase-i. *J. Am. Chem. Soc.* **1995**, *117*, 11619–11627.

44. Patel, S.S.; Wong, I.; Johnson, K.A. Pre-steady-state kinetic analysis of processive DNA replication including complete characterization of an exonuclease-deficient mutant. *Biochemistry* **1991**, *30*, 511–525.

45. Wong, I.; Patel, S.S.; Johnson, K.A. An induced-fit kinetic mechanism for DNA replication fidelity: Direct measurement by single-turnover kinetics. *Biochemistry* **1991**, *30*, 526–537.

46. Drake, J.W. A constant rate of spontaneous mutation in DNA-based microbes. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 7160–7164.

47. Shamoo, Y.; Steitz, T.A. Building a replisome from interacting pieces: Sliding clamp complexed to a peptide from DNA polymerase and a polymerase editing complex. *Cell* **1999**, *99*, 155–166.

48. Aller, P.; Duclos, S.; Wallace, S.S.; Doublie, S. A crystallographic study of the role of sequence context in thymine glycol bypass by a replicative DNA polymerase serendipitously sheds light on the exonuclease complex. *J. Mol. Biol.* **2011**, *412*, 22–34.

49. Hogg, M.; Aller, P.; Konigsberg, W.; Wallace, S.S.; Doublie, S. Structural and biochemical investigation of the role in proofreading of a beta hairpin loop found in the exonuclease domain of a replicative DNA polymerase of the b family. *J. Biol. Chem.* **2007**, *282*, 1432–1444.

50. Zhang, H.; Beckman, J.; Wang, J.; Konigsberg, W. Rb69 DNA polymerase mutants with expanded nascent base-pair-binding pockets are highly efficient but have reduced base selectivity. *Biochemistry* **2009**, *48*, 6940–6950.

51. Yang, G.; Franklin, M.; Li, J.; Lin, T.C.; Konigsberg, W. A conserved tyr residue is required for sugar selectivity in a pol alpha DNA polymerase. *Biochemistry* **2002**, *41*, 10256–10261.

52. Yang, G.; Franklin, M.; Li, J.; Lin, T.C.; Konigsberg, W. Correlation of the kinetics of finger domain mutants in rb69 DNA polymerase with its structure. *Biochemistry* **2002**, *41*, 2526–2534.

53. Zhang, H.; Rhee, C.; Bebenek, A.; Drake, J.W.; Wang, J.; Konigsberg, W. The l561a substitution in the nascent base-pair binding pocket of rb69 DNA polymerase reduces base discrimination. *Biochemistry* **2006**, *45*, 2211–2220.

54. Reha-Krantz, L.J.; Wong, C. Selection of bacteriophage t4 antimutator DNA polymerases: A link between proofreading and sensitivity to phosphonoacetic acid. *Mutat. Res.* **1996**, *350*, 9–16.

55. Reha-Krantz, L.J.; Nonay, R.L.; Stocki, S. Bacteriophage t4 DNA polymerase mutations that confer sensitivity to the ppi analog phosphonoacetic acid. *J. Virol.* **1993**, *67*, 60–66.

56. Zhong, X.; Pedersen, L.C.; Kunkel, T.A. Characterization of a replicative DNA polymerase mutant with reduced fidelity and increased translesion synthesis capacity. *Nucleic Acids Res.* **2008**, *36*, 3892–3904.

57. Li, V.; Hogg, M.; Reha-Krantz, L.J. Identification of a new motif in family b DNA polymerases by mutational analyses of the bacteriophage t4 DNA polymerase. *J. Mol. Biol.* **2010**, *400*, 295–308.

58. Tchesnokov, E.P.; Obikhod, A.; Schinazi, R.F.; Gotte, M. Engineering of a chimeric rb69 DNA polymerase sensitive to drugs targeting the cytomegalovirus enzyme. *J. Biol. Chem.* **2009**, *284*, 26439–26446.

59. Zahn, K.E.; Tchesnokov, E.P.; Götte, M.; Doublié, S. Phosphonoformic acid inhibits viral replication by trapping the closed form of the DNA polymerase. *J. Biol. Chem.* **2011**, *286*, 25246–25255.

60. Landry, M.L.; Stanat, S.; Biron, K.; Brambilla, D.; Britt, W.; Jokela, J.; Chou, S.; Drew, W.L.; Erice, A.; Gilliam, B.; *et al.* A standardized plaque reduction assay for determination of drug susceptibilities of cytomegalovirus clinical isolates. *Antimicrob. Agents Chemother.* **2000**, *44*, 688–692.

61. Kuhn, F.J.; Knopf, C.W. Herpes simplex virus type 1 DNA polymerase. Mutational analysis of the 3'-5'-exonuclease domain. *J. Biol. Chem.* **1996**, *271*, 29245–29254.

62. Hwang, Y.T.; Smith, J.F.; Gao, L.; Hwang, C.B. Mutations in the exo iii motif of the herpes simplex virus DNA polymerase gene can confer altered drug sensitivities. *Virology* **1998**, *246*, 298–305.

63. Wang, M.; Xia, S.; Blaha, G.; Steitz, T.A.; Konigsberg, W.H.; Wang, J. Insights into base selectivity from the 1.8 a resolution structure of an rb69 DNA polymerase ternary complex. *Biochemistry* **2011**, *50*, 581–590.

Appendix Figure A1. Protein sequence alignment of herpesvirus and bacteriophage polymerases based on structural data. Herpesvirus and bacteriophage polymerase were aligned individually using the muscle algorithm within Geneious [24,25]. Then the bacteriophage and herpesvirus sequences were aligned using an alignment based on a structural alignment of HSV1 UL30 (2GV9) and RB69 gp43 (1IH7) produced by RAPIDO [26]. Blocks above sequence highlight structural domains of polymerase. Regions unresolved in the HSV1 structural model are shown in green. The pre N terminal domain is in white. The N terminal domain is in yellow. The 3'–5' exonuclease domain is in red. The palm domain is in pink. The fingers domain is in blue and thumb domain is in green. The known conserved regions are shown in magenta blocks above sequence [27–31]. Secondary structural elements of HSV1 UL30 are indicated. Elements are number according scheme provided in Liu *et al.* (2006). Secondary Structural elements of RB69 gp43 are indicated. Elements are numbered according scheme provided in Wang *et al.* (1997). Structural motifs involved in polymerase and exonuclease activity are highlighted in sequence. Herpes virus specific motif is in blue. Exonuclease conserved residues are in red. Motif A is in magenta. N helix residues are in yellow. Motif B is in green. Motif C is in brown. KKRY motif is in purple. Mutations that have been associated with resistance to anti-herpetic drugs are shown below the corresponding residue [32,33]. Resistance conferring mutations are colored using the following scheme: Red: Pyrophosphate^R (Resistant), Blue: Nucleotide^R, Green: Pyrophosphate^R and Nucleotide^R. Purple: Pyrophosphate^{HS} (Hypersensitive). Pink: Pyrophosphate^R but Nucleotide^{HS}, Brown: Nucleotide^R but Pyrophosphate^{HS}.

Appendix Figure A1. *Cont.*

Not Visible | Pre N Terminal | pN1

		Sequence	
HSV1	UL30	1 MFSGGGGGPLSPGGKSAAPAAASGFFAPAGPRGAS-RGPPPCLRQNFYNFYLAPVGTQQKFTGPTQR-----HTYYSECDE-FRFIAPRVLDEDAPPEKRA	92
HSV2	UL30	1 MFCAAGGPASPGGKPAARAAASGFFAPHNPRGATIQTAPPCRRQNFYNFHLAQTGTQPKALGPAQR-----HTYYSECDE-FRFIAPRSLDEDAPAEQRT	93
VZV	ORF28	1 -------------MAIRTGFCNPFLTQASGIKYNFRTGRGSNREFLHSYKTTMSS-FQFLAPKCLDEDVPMEERK	61
EBV	BALF5	1 -------------MSGGLFYNFFLR------------------PNKGLLKKPDKEYLRLIPKC	32
HCMV	UL54	1 -------------MFFNFYLSGGVTGGAVAGG--------RQRSQPGSAQGSGKRPQPQFLQIVPRG	48
HHV6	U38	1 -------------MDSVSFFNFYLEANR-LKK--------------KSRSSYIRILPRG	31
HHV7	U38	1 -------------MDLVSFFNFYLENVRI-KK--------------KTKSTFLRIFPRG	31
KSHV	ORF9	1 -------------MDFFNFFIDPTRGGPR--------------NTVRQPTPSQSPT	29
T4	gp43	1 ----------------------------------	
RB69	gp43	----------------------------------	

pN2 | Pre N Terminal | Herpes Specific Motif | N1 | N2 | N Terminal | NA

		Sequence	
HSV1	UL30	93 GVHDGHLKR----A-----P----KVYCGGDERDVLRVGS---GGFWPRRSRLWGGVDH-APAGFNPTVTVFHVYDILENVEHAYGMRAAQFHARFMDAITPT	178
HSV2	UL30	94 GVHDGRLRR----A-----P----KVYCGGDERDVLRVGP---EGFWPRRLRLWGGADH-APEGFDPTVTVFHVYDILEHVEHAYSMRAAQLHERFMDAITPA	179
VZV	ORF28	62 GVHVGTLSR--------P--PKVYCNGKEVPILDFRC----SSPWPRRVNIWGEIDF-RGDKFPDRFNTFHVYDIVETTEAASNGDVS---RFATATRPL	143
EBV	BALF5	33 FQTPGAAGV---VDVRGPQ--PPLCFYQDSLTVVGGDEDGKGMWWRQRAQEGTAR--PEADTHGSPLDFHVYDILETVYTHEKCAVIPS--DKQGYVVPC	123
HCMV	UL54	48 VMFDGQTGL------------I----KHKTGRLPLMFYRE---IKHLLSHDMVWPCFWRETLVGRVVGPIRFHTVDQTDAVLFFDSPENVSP--RYRQHLVPS	130
HHV6	U38	31 IMHDGAAGL------I-----KDVCDSEPR--MFYRD---RQYLLSKEMTWPSLDRVRSKDYDHTRMKFHIYDAVETLMFTDSIENLFF--QYRHFVIPS	113
HHV7	U38	31 IMHDGAPGLM---------K----TLCDSEPR--MFYQD---KQYILKNDMTWPSLSQVAEKELR-APLKFHIYDASESLLFTDSIENIPF--QYRHFVIPS	112
KSHV	ORF9	29 VPSETRVCRLIPACFQTPGRPGVVAVDTTFPPTYFQGPKRGEVFAGETGSIWKTRRGQARNAPMSHLIFHVVDIVETTYTADRCEDVPFS---FQTDIIPS	127
T4	gp43	1 --MKEFYISIETVG----	12
RB69	gp43	1 --MQEFYLIVEQIG----	12

N Terminal

N3 N4 N5 NB NC N6

```
HSV1  UL30  179 GTVITLLGLTP-EGHRVAVHVYGTRQYFYMNKEEVDRHLQCRAPRDLCERMAAALRES-----------PGASFRGISADHFEAEVVERTDVYYY 261
HSV2  UL30  180 GTVITLLGLTP-EGHRVAVHVYGTRQYFYMNKAEVDRHLQCRAPRDLCERLAAALRES-----------PGASFRGISADHFEAEVVERADVYYY 263
VZV   ORF28 144 GTVITLLGMSR-CGKRVAVHVYGICQYFYINKAEVDTACGIRSGSELSVLLAECLRSSMITQNDATLNGDKNAFHGTSFKSASPESFRVEVIERTDVYYY 242
EBV   BALF5 124 GIVIKLLGRRKADGASVCVNVFGQQAYFY-------------------ASAPQGLDVEFAVLSALKAS----------TFDRRTPCRVSVEKVTRRSIMGY 195
HCMV  UL54  131 GNVLRFFGATE-HGYSICVNVFGQRSYFY-------------------CEYSDTDRLREVIASVGELVP----------EPRTPYAVSVTPATKTSIYGY 200
HHV6  U38   114 GTVIRMFGRSE-DGEKICVNVFGQEQYFY-------------------CECVDGKSLKATINNIMLTG----------EVKMSCSFVIEPADKLSLYGY 182
HHV7  U38   113 GNVIKLFGKTE-CGKKVCINVFGQNSYFY-------------------CEYQCKKELNSRICSLLNSS----------EIKMSCSFSIESVTKYNFYGY 181
KSHV  ORF9  128 GTVLKLLGRTL-DGASVCVNVFRQRCYFY-------------------TLAPQGVNLTHVLQQALQAG----------FGRASCGFSTEPVRKKILRAY 196
T4    gp43   13 NNIVERYIDEN--GKERTREVE-YLPTMFRHCKE----------------ES--KYKD  49
RB69  gp43   13 DNILERYIDKN--GKERSREVP-YKPTMFMHANP----------------EQAVKYID  51
```

2 3 4

N Terminal

N7 ND N8 NE N9 N10 N11

```
HSV1  UL30  262 ETRPALFYRVYVRSGRVLSYLCDNFC--PAIKKYEGGVDATTRFILD-NPGFVTFGWYRLKPGRNNTLAQPAAPMAFGTSSDVEFNCTADNLAIEGGMSD 358
HSV2  UL30  264 ETRPTLYYRVFVRSGRALAYLCDNFC--PAIRKYEGGVDATTRFILD-NPGFVTFGWYRLKPGRGNAPAQPRPPTAFGTSSDVEFNCTADNLAVEGAMCD 359
VZV   ORF28 243 DTQPCAFYRVYSPSSKFTNYLCDNF--HPELKKYEGRVDATTRFLMD-NPGFVSFGWNYQLKPGVDGERVRPASRQLTLSDVEIDCMSDNLQAIPNDDS 339
EBV   BALF5 196 GNHAGDYHKITLSHPNSVCHVATWLQDKHGCRIFEANVDATRRFVLD-N-DFVTFGWYSCRRAIPRLQHRD------SYAELEYDCEVGDLSVRREDSS 286
HCMV  UL54  201 GTRPVPDLQCVSISNWTMARKIGEYLLEQGFPVYEVRVDPLTRLVID-RR-ITTFGWCSVNRYDWRQQGRAS------TCDIEVDCDVSDLVAVPDDSS 291
HHV6  U38   183 NANTVVNLFKVSFGNFYVSQRIGKILQNEGFVVYEIDVDVLTRFFVD-N-GFLSFGWYNVKKYIPQDMGKG------SNLEVEINCHVSDLVSL-ENVN 272
HHV7  U38   182 NTEPIKNLFKLSFSNFYISNRIGKILLNEGVSVYEAEVEILNRFFID-N-NLKSFGWYQINYLSIQEFAKS------SNVEIELNCHVSDLFLLKED-S 271
KSHV  ORF9  197 DTQQYAVQKITLSSSPMMRTLSDRLTT-CGCEVFESNVDAIRRFVLD-H-GFSTFGWYECSNPAPRTQARDS------WTELEFDCSWEDLKFIPERTE 286
T4    gp43   50 IYGKNCAPQKFPSMKDARDWMKRMED--IGLEALG-MNDFKLAYISDTYGSEI-----------------------VYD 102
RB69  gp43   52 IYGKGCVKKEFDSMRDASQWMKRMDD--MGLEALG-MDDYKLAYLSDTYRKEI-----------------------KYD 104
```

5 A B

Appendix Figure A1. *Cont.*

3' → 5' exonuclease

```
                    Exo I        E1              E2         E3         EA

HSV1  UL30  359  LPAYKLMCF LEC KAGGEDELAFP-VAGHPEDLVIQISCLLYDLSTTA-LEHVL------LFSLG-----SCDLPESHINEL-----AARG--  431
HSV2  UL30  360  LPAYKLMCF LEC KAGGEDELAFP-VAERPEDLVIQISCLLYDLSTTA-LEHIL------LFSLG-----SCDLPESHLSDL------ASRG--  432
VZV   ORF28 340  WPDYKLLCF LEC KSGG-SNELAFP-DATHLEDLVIQISCLLYSIPRQS-LEHIL------LFSLG-----SCDLPQRYVQEM------KDAGL-  413
EBV   BLAF5 287  WPSYQALAF LEC LGEEG----FP-TATNEADLILQISCVLMSTGEEAGRYRRI-------LLTLG-----TCEDI---------------  345
HCMV  UL54  292  WPRYRCLSF LEC MSGEGG---FP-CAEKSDDIVIQISCVCYETGGNTAVDQGIPNGNDGRGCTSEG------IFGHSGLIHLFTI-----GTCGQV  372

                 N

HHV6  U38   273  WPLYGCWSF LEC LGQNGN---FP-DAENLGDIVIQISVVSFDTEGDRDERH------------LFTLG----TCEQI---------------  330
HHV7  U38   272  WPLYDCCSF LEC LSQNGN---FP-DAEQMGDIIQISVIDFDSEGNYQSKH------------LFTLG----TCEQI---------------  329
KSHV  ORF9  287  WPPYTILSF LEC MGEKG----FP-NATQEDEMIIQISCVLHTVGNDKPYTRM-------LIGLG----TCDPL---------------  344
T4    gp43  103  RKFVRVANC DIEVTG------DKFPDPMKAEYEIDAITH-----FDLINSMYGSVSKWD--AKLAAKLDCEGGDEVPQEI---  179
RB69  gp43  105  SSKIRIANF DIEVTS-----PDGFPEPAQAKHPIDAITH----YDSID-DKFYV------FDLLVSPYGRVKEWS--IKIAEKLQDEGGDEVPQEI--  182

                 exo I
                 6      7        8         Ba   Bb      Bc
```

Appendix Figure A1. *Cont.*


```
                    3' - 5' exonuclease

              Region IV

        E4   EB        E5   EC

HSV1  UL30  432  LPTPVLEFDSEFEMLLAFMTLIVKQYGPEFVTGYNIINFDWPFLLAKLTDIYKVPL-DGYGRM--N----GRGVFRVWDI------G 505
HSV2  UL30  433  LPAPVVLEFDSEFEMLLAFMTFVKQYGPEFVTGYNIINFDWPFVLTKLTEIYKVPL-DGYGRM--N----GRGVFRVWDI------G 506
VZV   ORF28 414  -PEPTVLEFDSEFELLIAFMTLIVKQYAPEFATGYNIVNFDWAFIMEKINSIYSLKL-DGYGSI--N----RGGLFKIWDV-G----  486

EBV   BLAF5 346  -EGVEVYEFPSELDMLYAFFQLIRDLSVEIVTGYNVANFDMPYILDRARHIYSINP-ASLGKI-RA----GGVCEVRRP--H---DAG 421
HCMV  UL54  373  GPDVDVYEFPSEYELLLGFMFFQRYAPAFVTGYNINSFDKYILTRLEYLYKVDS-QRFCKLPTA-----QGGRFFLHSPAVGFKRQYAAAFPSASHN 465
                                                    K
                                           D K CA
                                           K  VE

HHV6  U38   331  -DGVHIYEFASEFELLLGFFIFLRIESPEFITGYNINNFDLKYLCIRMDKIYHYEI-GCFSKLKNG-----KIGISVPHEQ---------YK 406
HH7   U38   330  -EGVFIYEFASEFELLYAFFLFLKCKSPEILTGYNIINFDLKYLCTRMEKIYNFEI-GGFSKLKRG----------H------- 404
KSHV  ORF9  345  -PGVEVFEFPSEYDMLAAFLSMLRDYNVEFITGYNIANFDLPYIIARATQVYDFKL-QDFTKI--K----TGSVFEVHQPRGG-----SDG 422
T4    gp43  180  LDRVIYMPFDNERDMIMEYINLWEQKRPAIFTGWNIEGFDVPYIMNRVKMILGERSMKRFSPI--GRVKSKLIQNMYGSKE------------ 258
RB69  gp43  183  IDKIVYLPFNSEEMMLEYLNFWQQKTPVILTGWNVESFDVPYVYNRLKNLFGENTAKRLSPH--RKTRVKVIENMYGARE------------ 261

                                                    exo 2

        9   C         10   D    Da    11   12
```

Appendix Figure A1. *Cont.*

```
HSV1  UL30  506  QSHFQRKSKIKVNGMVNIDMYGIITDK---IKLSSYKLNAVAEAVLKDK---KKDLSYRDIPAYYAAGPAQRGVIGEYCLQSLLVGQLFFKFLPHLELSA  600
HSV2  UL30  507  QSHFQRKSKIKVNGMVNIDMYGIITDK---VKLSSYKLNAVAEAVLKDK---KKDLSYRDIPAYYASGPAQRGVIGEYCVQDSLLVGQLFFKFLPHLELSA  601
VZV   ORF28 487  KSGFQRRSKVKINGLLISLDMYAIATEK---LKLSSYKLDSVAREALNES---KRDLPYKDIPGYYASGPNTRGIIGEYCIQDSALVGKLFFKYLPHLELSA  581
EBV   BLAF5 422  KGFLRANTKVRITGLIPIDMYAVCRDK---LSLSDYKLDTVARHLLGAK---KEDVHYKEIPRLFAAGPEGRRRLGMYCVQDSALVMDLLNHFVIHVEVAE  516
HCMV  UL54  466  NPASTAATKVYIAGSVVIDMYPVCMAK---TNSPNYKLNTMAELYLRQR---KDDLSYKDIPRCFVANAEGRAQVGRYCLQDAVIVRDLFNTINFHYEAGA  560

HHV6  U38   407  KGFLQAQTKVFTSGVLYLDMYPVYSSK---ITAQNYKLDTIAKICLQQE---KEQLSYKEIPKKFISGPSGRAAVGKYCLQDSVIVVRLFKQINYHYEVAE  501
HHV7  U38   405  KKFLNSLTKVNMSGILCFDMYNVYSSK---ISAQNYKLDTIAKLCLNQE---KKNLSYKEIPKKFIDGSKGRAAVGRYCIQDSLLVWQLFTKINYHYEMAF  499
KSHV  ORF9  423  GNFMRSQSKVKISGIVPIDMYQVCREK---LSLSDYKLDTVAKQCLGRQ---KDDISYKDIPPLFKSGPDGRAKVGNYCVIDSVIVMDLLLRFQTHVEISE  517
T4    gp43  259  ------IYSIDGVSILDYLDLYKKFAFTNLPSFSLESVAQHETKKGKLPYDGPINKLRET------NHQRYISYNIIDVESVQAIDKIRGFIDLVLS  343
RB69  gp43  262  ------IITLFGISVLDYIDLYKKFSFTNQPSYSLDYVSEYELKVGKLKYDGPISKLRET------NHQRYISYNIIDVRVVQIDMKRQFILLSLD  346
```

Appendix Figure A1. *Cont.*

N Terminal | delta C | Non visible

```
                                                                                      NG      NH
HSV1  UL30 601 VARLAGINITRTIYDGQQIRVFTCLLRLIADQKGFILPDTQ--GRFRGAGGEAPKRPAA-----------------A---------------REDEERPEEGE  669
                                                    V
HSV2  UL30 602 VARLAGINITRTIYDGQQIRVFTCLLRLIAGQKGFILPDTQ--GRFRGLDKEAPKRPAV-----------------P---------------RGEGERPGDGNG 670
VZV   ORF28 582 VARLARITLTKAIYDGQQVRIYTCLLGLASSRGFILPDGGYPATFEYKDVIPDVGDVE------------------------------------EEMDEDESVSPTGTSS 655
EBV   BALF5 515 IAKIAHIPCRRVLDDGQQIRVFSCLLAAAQKENFILP-M---------P------------------------------------------SASDRD--------- 561
HCMV  UL54 561 IARLAKIPLRRVIFDGQQIRIYTSLLDECACRDFILPNHYSKGTTVPETNSVAVSPNAAIISTAAVPGDAGSVAAMFQMSPPLQSAPSSQDGVLPGSGSN 660
                                                    N
                                                    E

HHV6  U38 502 VARLAHVTARCVVFEGQQKKIFPCILTEAKRRNMILPSM-------------------------V----------SSHNRQG------ 548
HHV7  U38 498 VASLAYITIRCAVFEGQQKKIFPCILHEAKNLNMILPSM-------------------------N----------TNFNKGKEN- 548
KSHV  ORF9 518 IAKLAKIPTRRVLTDGQQIRVFSCLLEAAATEGYILPVP-------------------------K----------GDAVS------ 562
T4    gp43 344 MSYYAKMPFS-GVMS-PIKTWDAIIFNSLKGEHKVIPQ------------------------------------------------- 379
               F                                   I    J        K
RB69  gp43 347 MGYYAKMQIQ-SVFS-PIKTWDAIIFNSLKEQGRVIPQ------------------------------------------------- 382
```

Appendix Figure A1. *Cont.*


```
                                                                        Non visible        Region II        Palm
                                                                                        P1  P2  PA        PB
                                                                                                        Motif A
                                                                                                14 15 16 L 17 M

HSV1 UL30 670 DEDERE-------EGGGEREPEGAREIAGRH-------VGYQGARVLDPTSGFHVNPVVVFDFASLYPSIIQAHNLCFSTLS-------L--RADAVAHLEA 748
                                                                 H G
                                                                 M
                                                                 V
                                                                 D
                                                                 Q
                                                                 E
                                                                 A
                                                                 K
                                                                 T
                                                                 H
HSV2 UL30 671 DEDKDDDEDGDEDGDEREEVARETGGRH-------VGYQGARVLDPTSGFHVDEVVVFDFASLYPSIIQAHNLCFSTLS-------L--RPEAAVAHLEA 753
                                                                     H
                                                                     T
                                                                     N
VZV ORF28 656 GRN-------V-------GYKGARVFDPDTGFYIDEVVVLDFASLYPSIIQAHNLCFTLLT-------L--NFETVKRLNP 713
                                                              GL
                                                              R
EBV BALF5 562 -------GYQGATVIQPLSGFYNSFVILVVDFASLYPSIIQAHNLCYSTMI-------TPGEEHRIAGIRP 617
HCMV UL54 661 SSSSVGVFSVGSGSSGGVGVSNDSHGAGG-------TAAVSYQGATVFEFEVGYYNDFVAVFDFASLYPSIIMAHNLCYSTLL-------VPGGEYPV---DP 746
                                                                           A
                                                                           M
HHV6 U38 549 -------IGYKGATVLERKTGYYAVFPTVVVFDFQSLYPSIYMAHNLCYSTLV-------L--DERQIAGLSE 603
HHV7 U38 549 -------VGYKGATVLEPKIGYYATFPTVVVFDFQSLYPSIYMAHNLCYSTLV-------V--DENAVIGLHA 603
                                                                   M
KSHV ORF9 563 -------GYQGATVISPSPGFYDDFVIWLDFASLYPSIIQAHNLCYSTLI-------PGDSLHLHPHLSP 618
T4 gp43 380 -------QGSHVKQSFPGAFVFEFPKPIAR-RYIMSFLTELYFSIIRQVNISPETIRGQFKVHPIH--EYIAGTAPKP 447
RB69 gp43 383 -------ARSHVVQPYPGAFVKEPIPNAY-KYVMSFLTSLYFSIIRQVNISPETIAGTFANAPIH--EYIAGTAPRP 450
```

Appendix Figure A1. *Cont.*

```
                                                                    P3    P4                           Region VI                    Palm      Finger

                                                                    PA

HSV1  UL30  749 GKDYLEIEVG-GRRLFFVKAHVRESLLSILLRDWLAMRKQIRSRI-------------------------PQ- 794
HSV2  UL30  754 DRDYLEIEVG-GRRLFFVKAHVRESLLSILLRDWLAMRKQIRSRI-------------------------PQ- 799
                                                             M N
VZVO  RF28  714 S-DYATFTVG-GKRLFFVRSNVRESLLGVLLKDWLAMRKATRARI-------------------------PG- 758
EBV   BLAF5 618 GEDYESFRLT-GGVYHFVKKHVHESFLASLLTSWLAKRKAIKKLL-------------------------AAC 664
HCMV  UL54  747 ADVYSVTLEN-GVTHRFVRASVRSVLSELLNKWVSQRRAVRECM--------------------------REC 793
                            K         M   I   L
                            D
                            Q
HHV6  U38   604 S-DILTVKLG-DETHRFVKPCVRESVLGSLLKDWLAKRREVKAEM------------------------QNC 649
HHV7  U38   604 D-DILTVHVG-PVTHRFVKKTVRESILASLLKKWLDKRKEVKMQM------------------------KMC 649
KSHV  ORF9  619 D-DYETFVLS-GGPVHFVKKHKRESLLAKLLTVWLAKKEIRKTL------------------------ASC 664
T4    gp43  448 SDEYSC---S-PNGWMYDK--HQEGIIPKEIAKVFFQRKDWKKKMFAEEMNAEAIKKIIMKGAGSCSTKPEVERYVKFSDDFLNELSNYTESVLNSLIE- 540
RB69  gp43  451 SDVYSC---S-PNGMMYYK--DRDGVIPTEITKVFNQRKEHKGYMLAAQRNGELIKEALENPNLSVDSPLDIDYRFSDEIKAKIKTLSAPSLREMLF- 543

                                                    18   19                 N                         O          P
```

Appendix Figure A1. *Cont.*

```
HSV1  UL30  795  SSPEEAVLLDKQQAAIKVVQNSVYGFTGVQHGLLPCLHVAATVTTIGREMLLATREYVHARWAA---F----------EQLLADFPEA-A  870
HSV2  UL30  800  STPEEAVLLDKQQAAIKVVQNSVYGFTGVQHGLLPCLHVAATVTTIGREMLLATRAYVHARWAE---F----------DQLLADFPEA-A  875
VZV   ORF28 759  SSSDEAVLLDKQQAAIKVVQNSVYGFTGVAQGFLPCLYVAATVTTIGRQMLLSTRDYIHNNWAA---F----------ERFITAFPDIES  835
EBV   BALF5 665  EDPQRTILDKQQLAIKCTQNAVYGFTGVANGLFPCLSIAETVTLQGRTMLERAKAFVEALSPA---------------NLQALAPSPDAWA  741
HCMV  UL54  794  QDPVRRMLLDKEQMALKVTQNAFYGFTGVVNGMMPCLPIAASITRIGRDMLERTARFIKDNFSEPCFLHNFFNQEDYVVGTREGDSEESSTLPEGLETSS  893
HHV6  U38   650  SDPMMKLLLDKKQLAIKTTQNSVYGVTGAAHGLLPCVAIAASVTCLGREMLCSTVDYVVNSKMQS--E---------------QFFCEELGLTAS  726
HHV7  U38   650  NNFVMEMLLDKKQLAIKTTQNSVYGVTGATHSLLPCVAIAASVTCLGREMLCKTVVDVDSAMYS--D---------------TFFIEKFGLTRG  726
KSHV  ORF9  665  TDPALKTILDKQQLAIKVTQNAVYGFTGVASGILPCLNIAETVTLQGRKMLERSQAFVEAISPE--R---------------LAGLLRRPV  738
T4    gp43  541  ECEKAAATLANTNQLNRKILLINSLYGALGNTHFRYYDLRNATAITIFGQVGIQWIARKINEYLN--------------------------  603
RB69  gp43  544  RAERTEVAGMTAQINRKLLLINSLYGALGNVWFRYYDLRNATAITLFGQMALQWIERKVNEYLN---------------------------  606
```

Appendix Figure A1. *Cont.*

Palm

Region I   Region VII

P5 P6   PE   P7 P8 P9

HSV1 UL30 871 DMRAPGPYSMRIIYGDTDSIFVLCR----------GLTAAGLTAVGDKMASHISRALFLPPIKLECEKTFT------KLLLIAK 938

HSV2 UL30 876 GMRAPGPYSMRIIYGDTDSIFVLCR----------GLTAAGLIVAMGDKMASHISRALFLPPIKLECEKTFT------KLLLIAK 943
VZV   ORF28 836 SVLSQKAYEVKVIYGDTDSVFIRFK----------GVSVEGIAKIGEKMAHIISTALFCPPIKLECEKTFI------KLLLITK 903

EBV   BALF5 742 PLNPEGQL--RVIYGDTDSLFIECR----------GFSESETLRFAEALAAHTTRSLFVAPISLEAKTFS------CLMLITK 807
HCMV UL54 894 GGSDERRVEARVIYGDTDSVFVRFR----------GLTPQALVARGPSLAHYVTACLFVEPVKLEFEKVFV------SIMMICK 961
HHV6 U38   727 DFT--GDLKVEVIYGDTDSIFMSVR----------NMANESLRRIAPMIAKHITDRLFKSPIKLEFEKILC------PLILICK 792
HHV7 U38   727 DFS--GTFGIEVIYGDTDSLFVTFK----------NVCPKALKRIAPSIASHITNTLFKHPIKLEFEKILF------PLILICK 792
KSHV ORF9 739 DVSPDARF--KVIYGDTDSLFICCMGFNMDSVSDFAEELASITT----------NTLFRSPIKLEABKIFK------CLLLLTK 804
T4   gp43 604 -KVCGTNDEDFIAAGDTDSVYVCVDKVIEKVGLDRFKEQNDLVEFMNQFGKKKMEPMIDVAYRELCDYMNNREHLMHDREAISCPPLGSKGVGGFWKAK 702
RB69 gp43 607 --SVCKTENQKFVLYGDTDSIYVKADAIINKVGEDKFRDTNHWVDFLDKFARERMEPAIDAGFREMCEYMNNKQHLMFMDREAIAGPPLGSKGLGGFWTGK 705

Motif C

Q 20 21 R S 22 23

Appendix Figure A1. *Cont.*

Palm | Thumb

Region VII — Region V

P10 · P11 · TA · TB · TC · 24 · 25 · T · U · V · 26

KKRY Motif

```
                          H                    K                              G
HSV1  UL30   939  KKYIGVIY------GGKML-IKGVD-LVRKNNCAF INRTSRALVDLLFYDDTVSGAAAALAERPAEEWLARPLPEGLQAFGAVIVDAHRRITDPERDI  1300
HSV2  UL30   944  KKYIGVIC------GGKML-IKGVD-LVRKNNCAF INRTSRALVDLLFYDDTVSGAAAALAERPAEEWLARPLPEGLQAFGAVIVDAHRRITDPERDI  1028
VZV   ORF28  904  KKYIGVIY------GGKVL-MKGVD-LVRKNNCQF INDYARKLVELLLYDDTVSRAAAEASCVSIAEWNRRAMPSGMAGFGRIIADAHRQITSPKLDI   993
EBV   BALF5  808  KRYVGVLT------DGKTL-MKGVE-LVRKTACKFVQTRCRRVLDLVIADARVKEAASLLSHRPFQESFTQGLPVGFLPVIDILNQAYTDLREGRVFM   897
HCMV  UL54   962  KRYIGKVE------GASGLSMKGVD-LVRKTACEFVKGVTRDVLSLLFEDREVSEAAVRLSRLSLIDEVKKYGVPRGFWRILRRLVQARDDLYLHRVRV  1052
HHV6  U38    793  KRYIGR-Q------DDSLLIFKGVD-LVRKTSCDEVKGVVKDLVDLLFFDEEVQTAAVEFSHMTQTQLREQGVPVGIHKILRRLCKAREELFQNRADV   882
HHV7  U38    793  KRYIGKL-------DDATLVFKGVE-LVRKTSCDEVKVVVKDIIDLLFWDVDVDQKSAEKLSNMTIQEIYENGVPDGIQKIIKKLCDARDALFINRVNI   882
KSHV  ORF9   805  KRYVGVLS------DDKVL-MKGVD-LIRKTACRFVQEKSSQVLDLIIREPSVKAAAKLISGQATDWVYREGLPEGFVKIIQVLNASHRELCERSVPV   894
T4    gp43   703  KRYALNVYDMEDKRFAEFHLK-IMGMETQQSSTPKAVQEALEESTRRILQEGEE------------------SVQEYYKNFEKEYRQLDYKVIAEVKTANDIA   786
RB69  gp43   705  KRYALNVWDMEGTRYAEFHLK-IMGLETQKSSTPKAVQKALKECTRRMLQEGEA---------------------SLQEYFKQFNKEFNELDYVSIAAVSSANNIE   722
```

Thumb

HSV1	UL30	1029	QDFVLTAELSRHPRAYTNKRLAHLTVYYKLMARRAQVPSIKDRIPYVIVAQTREVEETVARLAALRELDAAAPGDEPA-PPAALPSPAKRPRETPSPADP	1127
HSV2	UL30	1034	QDFVLTAELSRHPRAYTNKRLAHLTVYYKLMARRAQVPSIKDRIPYVIVAQTREVEETVARLAALRELDAAAPGDEPA-PPAALPSPAKRPRETPSHADP	1132
VZV	ORF28	994	NKFVMTAELSRPPSAYINRRLAHLTVYYKLVMRQGQIPNVRERIPYVIVAPTDEVE--------------ADAKSVALLRGDPL------QNTAGK	1069
EBV	BALF5	898	GELCFSTELSRKLSAYKSTQMPHLAVYQKFVERNEELPQIHDRIQYVFVEPKGGVKGAR-KTEMAEDPAYAERHGVPVAVDHYFDKLLQGAANILQCLFD	996
HCMV	UL54	1053	EDIVLSSVLSKDISLYRQSNLPHIAVIKRLAARSEELPSVGDRVFYVLTAPGVRTAPQ-GSSDHGDSVTTGVVSRSDAIDGTDDDADGGGVEESNRRGGE	1151
HHV6	U38	883	RHLMLSSVLSKEVAAYKQPNLAHLSVIRRLAQRKEEIPNVGDRIMYVLIAPSTGNKQTH-NYELAEDPNYVLEHKIPIHAEKYFDQIIKAVTNAISPIFP	981
HHV7	U38	883	KSIVLSSVLSKDISAYKQANLPHLAVVKRLAQRKEELPNIGDRVMYILIAPTETVKKTFHNYELAEDPAYAIENNLKINADKYFEQIMKAVTNAISPIFP	982
KSHV	ORF9	895	DKLTFTTELSRPLADYKTQNLPHLTVYQKLQARQEELPQIHDRIPYVFVDAPGSLRS-----ELAEHPEYVKQHGLRVAVDLYFDKLVHAVANIIQCLFQ	989
T4	gp43	787	KYDDKGWPGFKCPFHIRGVLTYRRAVSG-LGVAPILDGNKVMVLPLREGNPFGDKCIAWPSGTELPKEIRSDVLSWIDHSTLFQKSFVKPLAGMCESAGM	885
RB69	gp43	790	KYNEGGFPGSKCPYHIKGCLAYMRASAGDLTMPRIVDGEKVYVLPLKEGNPFGEACMAWPSGIELPVQIRDQVLRWMDYNVLFQKTFIKPLTGFTESAKI	889

Structural elements: T1, TD, T2, TE, W, 27, 28, X, Y

Appendix Figure A1. *Cont.*

Thumb   Non Visible

TF   TG   TH

```
HSV1  UL30  1128  PGGASKPRKLLV---SEIAEDPAYAIAHGVAINTDYYFSHLLGAACVTFKALFGNNAKITESLIKRFIPEVWHPDDVAARLRTAGFGAVGAGATAEETR  1224
HSV2  UL30  1133  PGGASKPRKLLV---SEIAEDPGYAIARGVPINTDYYFSHLLGAACVTFKALFGNNAKITESLIKRFIPEVWHPDDVAARLRAAGFGAVGAGATAEETR  1229
VZV   ORF28 1070  RCGEAKRLLIS----DLAEDPIHVTSHGLSINIDYYFSHLIGTASVTFKALFGNDTKLTERLIKRFIPETRVVNVKRMINRLQAAGFVCIHAPCWDNKMN 1165
EBV   BALF5  997  NNSGAALSVLQNFTARPF*----                                                                             1016
HCMV  UL54  1152  PAKKRARKPPSAVCNYEVEAEDPSYVREHGVPIHADKYFEQVLKAVTNVLSPVFPGGETARKDKFLHMVLPRRLHLEPAFLPYSVKAHECC*--------- 1242
HHV6  U38    982  KTDIKKEKLLLYLLPMKVYLDETFSAIAEVM*---------                                                            1013
HHV7  U38    983  KTGIKKEKFLLSTLPLKVYVDQSFCDLTDVL*---------                                                            1014
KSHV  ORF9   990  NNTSATVAILYNFLDIPVTFPTP*----                                                                        1013
RB69  gp43   890  DYEKKATLFDMFDF*----                                                                                  903
T4    gp43   886  DYEEKASL--DFLF*----                                                                                  897
```

Non Visible

```
HSV1  UL30  1225  RMLHRAFDTLA*---------  1236
HSV2  UL30  1230  RMLHRAFDTLA*---------  1241
VZV   ORF28 1166  TEAEITEEEQSHQIMRRVFCIPKAILHQS*  1195
EBV   BALF5        -------------
HCMV  UL54         -------------
HHV6  U38          -------------
HHV7  U38          -------------
KSHV  ORF9         -------------
RB69  gp43         -------------
T4    gp43         -------------
```

Review

A Genetic Approach to the Development of New Therapeutic Phages to Fight *Pseudomonas Aeruginosa* in Wound Infections

Victor Krylov *, Olga Shaburova, Sergey Krylov and Elena Pleteneva

Laboratory for Bacteriophages Genetics. Mechnikov Research Institute of Vaccines and Sera, RAMS, 5a, Maliy Kazenniy per., Moscow 105064, Russia; E-Mails: oshabs@mail.ru (O.S.); sergeykrylovv@gmail.com (S.K.); f2600@yandex.ru (E.P.)

* Author to whom correspondence should be addressed; E-Mail: krylov.mech.inst@mail.ru; Tel.: +7-495-9170782.

Abstract: *Pseudomonas aeruginosa* is a frequent participant in wound infections. Emergence of multiple antibiotic resistant strains has created significant problems in the treatment of infected wounds. Phage therapy (PT) has been proposed as a possible alternative approach. Infected wounds are the perfect place for PT applications, since the basic condition for PT is ensured; namely, the direct contact of bacteria and their viruses. Plenty of virulent ("lytic") and temperate ("lysogenic") bacteriophages are known in *P. aeruginosa.* However, the number of virulent phage species acceptable for PT and their mutability are limited. Besides, there are different deviations in the behavior of virulent (and temperate) phages from their expected canonical models of development. We consider some examples of non-canonical phage-bacterium interactions and the possibility of their use in PT. In addition, some optimal approaches to the development of phage therapy will be discussed from the point of view of a biologist, considering the danger of phage-assisted horizontal gene transfer (HGT), and from the point of view of a surgeon who has accepted the Hippocrates Oath to cure patients by all possible means. It is also time now to discuss the possible approaches in international cooperation for the development of PT. We think it would be advantageous to make phage therapy a kind of personalized medicine.

---

1. Introduction

Gram-negative bacteria of species *Pseudomonas aeruginosa* may be found in different natural habitats, because they easily adapt to different conditions. The capability for quick adaptation is the main reason that identifies them as opportunistic pathogens. They cause infections in immune compromised patients or patients with cystic fibrosis—a frequently occurring hereditary disease in Caucasians. *P. aeruginosa* strains are common components in microbial communities of different origins. They have acquired the status of hospital pathogens, and may be isolated from clinical samples taken from the wounds, sputum, bladder, urethra, vagina, ears, eyes and respiratory tract. The emergence of resistance to the most powerful new antibiotics in such clinical *P. aeruginosa* strains, occurring even during treatment [1–9], makes the fight with *P. aeruginosa* hospital pathogens a great problem.

Genomes in most hospital strains of *P. aeruginosa* contain pathogenic islands, where genes, coding many factors of pathogenicity and virulence of this bacterial species such as phospholipase C elastase, protease, siderophore, DNAse, pyocyanin *etc*., are located simultaneously with genes controlling multiple drug resistance. One of these large genomic islands, PAPI-1, can be transformed into an extrachromosomal circular form, plasmid, after precise excision from a bacterial chromosome [10]. Sometimes such plasmid acquires the ability to be transferred into other recipient *P. aeruginosa* strains by a conjugative mechanism, via a type IV pilus [11]. Such migration quickly disseminates an antibiotic resistance into new strains, thus making the use of antibiotics useless. This is the reason for a quite unexpected renaissance of phage therapy, the use of bacterial viruses in the treatment of bacterial infections, which was proposed by Felix D'Herelle in 1917, immediately after the finding of bacteriophages [12]. The bacteriophage treatment was applied in medical practice with varying success until the introduction of antibiotics; then, because of the great success and simplicity in use of antibiotics, phages were no longer regarded as a serious tool in anti-infective therapy. However, in Russia, Poland and Georgia the use of phage therapy has not ceased and continues to the present day. Given the frequent epidemics of food borne diseases and the increase of many enteric pathogens resistant to antibiotics [13–17], it is useful to remember that some specific phage compositions introduced by F. D'Herelle (but permanently updated) are used with success in the treatment and in the prevention of intestinal infections (see Figure 1).

Figure 1. Commercially produced mix of phages "Intesti" (ImBio Nizhny Novgorod, Russia).

In Eastern Europe, bacteriophages are widely used in surgical wards [18–20]. Apparently, this trend in phage therapy of *P. aeruginosa* caused infections will be used further, considering the failure of the development of safe vaccines against *P. aeruginosa* [21–23].

Bacteriophages were an excellent model of genetic research. Many of the basic concepts of modern biology and the different elements of the methodology in biological and medical sciences have emerged as a result of studying the genetics of phages from the 1940s to 1970s. Being used in other areas, they have become a powerful boost to research of various pro- and eukaryotic systems. In the last two decades, interest in bacteriophages increased significantly again. One reason is the above mentioned occurrences of multiple antibiotic resistant bacteria and the hope that the use of live phages or their products—structured or molecular bacteriocins—antibacterial peptide phage origin—can help in the treatment of bacterial infections. In our opinion, there is another, no less significant reason for a detailed study of bacteriophages. As it has turned out, bacteriophages (both temperate and virulent) are actively involved in the evolution of bacteria, including pathogens, accomplished through different kinds of transduction (general and specialized) horizontal gene transfer (HGT). Studies of the structure of pathogenic islands of different bacterial species confirmed the presence in them of complete or fragmented genomes of temperate phages. Many of the new forms of infective diseases (so-called "emerging diseases") are caused by the appearance in pathogenic islands of previously known pathogens of new genes introduced by phages, plasmids, and transposons (which can be considered as largely related genetic elements that form a common pool of evolutionary active genes). The increased interest in the use of bacteriophages as a replacement for antibiotics in the West has led to an increase in the number of publications illustrating the possibility of actual use of phages in human and veterinary medicine, in finding and describing new phages. Here are the numerous review papers related somehow to bacteriophages that were published recently [24–30]. Extremely bold proposals have been made of the possible use of bacteriophages as antiseptic additives in ready to eat foods, which may potentially be infected with pathogenic bacteria. There are ideas for use of bacteriophages and their products to retard spoilage of foodstuffs during the time of storage or for the prevention of

endemic infections and food poisoning associated with the consumption of potentially contaminated food (phage preparations of the «Intesti» type) (see above). The need to quickly find, some alternative to antibiotics is evident. However, in the case of phages there is an evident problem: how to eliminate the possible undesirable activity of phages with the emergence of new pathogens in the process of introducing phage therapy. Some other problems must also be solved if there is an intention to raise the level of phage therapy up to accepted standards in modern medicine. Among such problems are ensuring reproducible results of phage therapy, its long-term use and achieving the required standards for therapeutic phages and their mixtures. Can the introduction of phage therapy be a cause of the emergence of pathogens with new properties? There are also some questions about the prerequisite of high level cooperation, including a quick exchange with bacteriophages capable of stopping infectional outbreaks. We consider some of these issues later, but first we discuss the natural limitations in the use of bacteriophages that distinguish them from antibiotics. This assesses the real prospects of live phage therapy and also helps understand why phage therapy is not the main procedure of treatment for some diseases in the countries where its use was not interrupted after introduction of antibiotics.

2. The Goal, Objectives and Content of the Present Review

Many studies related in one way or another to bacteriophages have been published in the last 10–15 years. They concern, the numerous advantages of phage therapy and specific features of a large number of newly isolated phages have been discussed. The list of phages with sequenced and annotated genomes is in constant enlargement in NCBI.

Currently, there has been established a certain unofficial algorithm in the studies of newly isolated bacteriophages: the isolation of the phage itself, extraction of genomic nucleic acid (in most cases, DNA), electron microscopy of phage particles for classification purposes, genome sequencing and annotation, estimation of evolutionary relationships with other phages and prospects for practical use. Previously initiated as well as new companies are in action, producing and advertising their numerous commercial phage mixtures; newly organized specialized therapeutic bacteriophage treatment centers are operating. We have not provided links to these activities: all this can be found on the Internet.

The purpose of this review is to draw attention to the need for a close examination of the possible features in the behavior of different types of bacteriophages in their actual application, which can be found on the stage of laboratory studies [31], in order to understand the possible consequences and the necessary modifications of the bacteriophages and the methods of their use. The system of phage therapy is fundamentally different from the organization of antibiotic therapy. In supporting the use of bacteriophages in medical practice, we want first of all to justify the need for the introduction of such a system. We start by looking at the natural limitations in the use of phages as therapeutic agents.

3. Limitations in the Use of Live Bacteriophages in a Treatment

(1) Phage therapy with living phages is applicable only in cases where it is possible to provide direct contact with the virus and bacteria. However, the direct introduction of even highly purified

phage particles in the blood, such as in the case of septicemia (although there are described unique cases of this kind) [32] is very risky. The view that oral administration of phages may cause them to penetrate into the bloodstream, and then into the urine (the concept of "phages are absorbed into blood and cure"), is not justified by rigorous research (although such cases have also been described in the literature) [33]. Our own laboratory experiments did not confirm this. Requirements for direct contact limits the use of phage therapy with wound infections (other than pressure ulcers), urogenital infections, intestinal infections, eye infections and infections of the ear, nose and throat organs. Cystic fibrosis, where bacteria of the species P. aeruginosa are in contact with different phages cannot yet be attributed to the list of infections that are acceptable for phage therapy, although research in this area (in a mouse model) continues [34].

(2) Another reason for the limited use of phage therapy is the inability to provide long-term effectiveness of a particular mixture of therapeutic phages. In the manufacturing companies, producing commercial phage mixtures, it is generally accepted that included in such products are phages with the broadest spectrum of lytic activity for each bacterial species involved. However, regardless of the spectrum of lytic activity of phages in the mixtures, bacterial mutants of a species that reveal resistance to all phages in the mixture may arise even after a few days after the start of treatment. Thus, a single mutation in the genome of Gram-negative bacteria such as *P. aeruginosa*, leading to resistance to adsorption, may prevent the growth and lytic activity several unrelated phages [35]. It is known from personal contact with surgeons of departments of purulent infections, that phage-resistant clones often arise after short time use of polyvalent phage mixtures. Such resistant clones can then replace the previous hospital pathogens. We consider the properties of these pathogens and their relation to phages later. Usually renewal of phages in commercial mixtures is quite rare, a few times a year. Therefore, the rapid accumulation of phage-resistant mutants can lead very quickly to loss of activity of the applied phage mixture. The only possible solution to the problem is the replacement of such an ineffective phage mixture with the mixture from another manufacturer, in the hope that the phages in the new mixture have a different spectrum of lytic activity. This is a temporary albeit not very reliable solution.

(3) The obvious aim of each producer of therapeutic phages is to create a preparation with the highest range of final lytic activity, which can be used without replacement for as long as possible. Therefore, since the time of F. D'Herelle, in order to ensure maximal activity of therapeutic phages against pathogenic strains, phages for therapy have been produced with the use of pathogenic clinical isolates. Another important reason for such use of clinical (pathogenic) isolates is that it is not always possible to find appropriate non-pathogenic variants with sensitivity to a particular phage, which must be incorporated into the therapeutic mix. This is the positive side of the use in the batch phage production of clinical isolates as hosts. On the other hand, the real pathogen used as a production strain, even when non-lysogenic, obviously contains genes with harmful effects. As a result, phages obtained with the use of such a pathogen, when introduced into the therapeutic mixture, may in the course of treatment, with any kind of transduction, transfer such genes into clinical strains with the resulting emergence of bacteria that cause the disease with symptoms of "emerging diseases". The

only solution is to use only well studied bacterial hosts. In principal, it is necessary to create production strains that are sensitive to different phages, based on well-studied model strains that do not contain dangerous genes in their genomes.

(4) Some weaknesses in planning and use of phage therapy also hinder the successful introduction as a standard treatment or preventive procedure. There is a need to create a well thought out organization at all levels from the isolation of a pathogenic strain to the fast selection of the best phage for treatment. The creation of such an organization requires support on an inter-state level, because its development goes beyond the possibilities of individual researchers or research laboratories, and apparently requires coordination on a number of different levels. The existence of such a system could play a positive role in quick reactions to the occurrence of local epidemics, such as the emergence of the outbreak caused by an enterohemorrhagic (EHEC) strain of *E. coli* serotype O104: H4 in early May 2011 in northern Germany, or in the case of the outbreak of listeriosis in September and October 2011 in the U.S.A. In the U.S.A., for instance, there is a phage mixture active against *Listeria* (integrated drug ListShield™; [36]) which is permitted by the US Food and Drug Administration for use in «ready-to-eat» products. It is possible that the timely distribution of therapeutic phage mixtures amongst the population in areas threatened by epidemiological characteristics could reduce the number of subsequent cases. In the same hypothetical system of organization of phage therapy should be included laboratories involved in research of phages for therapy as well as international collections of bacteriophages [37]. An optimal situation is the case when only such phages active against local pathogens, but not polyvalent mixtures are used in clinics. It is good in the case when there is the possibility to choose such phages *in situ*, in the hospital, and then treatment must be accompanied by constant monitoring to check for arising phage-resistant variants. The choice of inappropriate phages and use of polyvalent mixtures without checking for the effectiveness of their component phages against particular wound pathogens will affect the efficiency of treatment and, as a consequence, will discredit the idea of phage therapy among doctors and patients.

Therapeutic phages, which are intended for personal use, should definitely go through the standard procedure of classification up to species. This significantly lowers the probability of HGT, revealing the presence in the genomes of phages containing any features that pose a potential danger. Perhaps, phage therapy at some future time could be included into the arsenal of personalized medicine.

4. Non-Canonical Interaction of Phages and Bacteria

The original idea of phage therapy is based on the assumption that interactions of phages and bacteria are similar with predator-prey relationships and the development of virulent phage proceeds within the framework of the lytic cycle. This ideal ("canonical") lytic cycle is as follows: adsorption of the phage to the surface cell structures (adsorption receptors specific for each type of phage); injection of the phage genome into the cell (using the natural pores or after local melting of the cell wall by a lytic enzyme incorporated into the structure of the tail of the phage particle); the implementation of phage development programs written in phage genome; the destruction of the cell envelope from

inside by specific phage enzymes and release progeny phage particles; readiness to repeat the same cycle. The ability of cells to divide normally disappears soon after the injection of the phage genome.

In the case of infection cells with temperate phage there can be two outcomes: a typical lytic cycle or lysogenization of the infected cell with a phage genome in prophage condition; prophage can be integrated into the bacterial chromosome or be in a state of plasmid, whose division and distribution in each of the daughter bacterial cells are strictly coordinated with the division of bacteria and replication of bacterial chromosome. A stable state of bacterial cell with prophage is supported through the blockade of lytic genes expression in prophage with a special protein repressor. Sometimes the repressor undergoes inactivation, and the phage starts the lytic cycle of development. However, under certain conditions, there is a deviation from such a course of events that could significantly change the nature of the phage behavior and, as a consequence, the final result of phage therapy. These deviations can result in a pseudolysogenic state of infected bacterial cells. Developmental disorders in the case of temperate phages can prevent the establishment of a lysogenic state (pseudovirulent state).

A good example of the lytic cycle is the development of phage T4. As a model for temperate phage, the development of bacteriophage lambda can be accepted. Nevertheless, it is necessary to recognize that the properties, even of these models, are still far from being fully explored. For example, the functions of 20 proteins coded by the lambda genome were not possible to assign in comparison with the known database. Their functions were studied with specially designed experiments [38]. Authors predicted functions for 12 of the mentioned proteins. So even in the case of phage for which careful research has lasted several decades, the functions of many of the genes are still unknown.

5. Pseudolysogeny and Pseudovirulence

5.1. Pseudolysogenic Conditions

5.1.1. Some Examples of Reasons for Pseudolysogenic Conditions

Studies of bacteriophages in laboratories are usually accomplished under conditions which are optimal for bacterial host growth. However, interactions of bacteria and phages at the place of infection may occur in different situations. To check their influence on deviations in the canonical interaction of phage and bacteria before introduction of a phage into a therapeutic mixture, it is necessary to conduct special studies under different conditions. The importance of such studies is confirmed by the existence of particular types of phage development, which differ from the canonical lytic and temperate infections. We do not mention here specific bacterial viruses with chronic type infection as their use in therapy or occasional occurrence in phage therapeutic mixtures in our opinion may be very dangerous [37]. Let us now consider some different examples of pseudolysogeny:

(a) Frequent reasons for establishment of pseudolysogeny and its maintenance are changes of cultural conditions in the background of specific bacterial and phage genotypes. Typical results were obtained in the study of *P. aeruginosa*, slowly growing in the presence of its

phage in a chemostat continuous culture model. The frequency of pseudolysogens increased as cells were starved. According to the authors, pseudolysogeny is specific phage strategy to survive in periods of starvation of their hosts [39].

(b) Transition of lysogenic cells from active multiplication into stationary phase can also stimulate high frequency of phage loss, as has been found in the study of the gram-positive bacteria *Propionibacterium acnes* with a group of closely related temperate inducible phages. Bacterial clones surviving such an event were susceptible to new infection by the same phages. Most likely, the phage genome in this case was not integrated into the bacterial chromosome [40].

(c) Bacteria may acquire a capability for lysogeny as result of mutation. Phage VHS1 produces clear plaques on its host *Vibrio harveyi* (strain VH1114). Bacterial clones have been isolated which could be lysogenized but being lysogens they produced cured cells after approaching the stationary growth phase. These bacteria had inherited differences proving their mutational origin [41].

(d) Sometimes pseudolysogeny is a result of several different genetical modifications in phage genome and/or in a bacterial genome. Neurotoxins C and D of *Clostridium botulinum* are encoded by genes of bacteriophages. The genome of converting phage c-st was sequenced and annotated. As it turned out, c-st prophage is present as a circular plasmid. Plasmids (and plasmid prophages) code special proteins, necessary to resolve plasmid multimer partition, and segregation, but it is possible that their activity is not sufficient to ensure the stable inheritance of c-st as plasmid. Besides, there is another important reason for the plasmid instability. A remarkable feature of the c-st genome is the abundance of IS elements (altogether 12 copies). The presence of such a number of IS is unexpected for a viable phage. The authors suggest that possible recombination between these multiple elements is the basic reason for the plasmid prophage instability. This case of pseudolysogeny may be considered as a typically weakened impaired lysogenic state [42].

Detailed studies of the regulation systems of the temperate phage lambda *E. coli* demonstrated the dependence of the stability of the lysogenic state of functioning of a group of genes that affect the expression of the gene encoding the repressor. However, apparently with new phages allocated from the natural environment there are many cases that cannot be explained by reference to known patterns.

5.1.2. Pseudolysogeny and Therapeutic Phages for *P. aeruginosa*: General Considerations

Our laboratory, for a number of years, has studied different bacteriophages, active in pseudomonades of different species. In the course of such work, we have met some situations which evidenced the "non-canonical" behavior of phages and phage-sensitive bacteria for *P. aeruginosa*. In relation to phage therapy, it may be of interest to consider some cases of non-canonical interactions for phages from existing commercial mixtures. Besides, it is necessary to take into account the special role of temperate phages in horizontal genetic transfer. Such observations can have certain practical

significance and can be taken into consideration to enhance phage therapy of infected wounds or hospital infections. The studies of non-canonical interactions of phages and bacteria are intended primarily to assess the possibility and probability of occurrence during the treatment of hybrid phages capable of transposition, conversion, and lysogenization, *i.e.*, participating in the process of HGT. The need for such studies has been proved because even the sequencing and annotation of the genome cannot always predict the behavior of bacteriophages in the real world. Similarly, *in silico* experiments on the interaction of controlled phage proteins cannot confirm the safety of their functions or prove the significance of the function of these proteins for the viability or lytic activity of the phage. In relation to *P. aeruginosa* caused infections, there are detailed studies of genomes of several phage groups which may be considered as favorites for phage therapy, such as KMV-like, PB1-like, N4-like and phiKZ-like phages, which are be found in some commercial mixtures. Different phages of these four species are the frequent sources used to replenish commercial mixtures. It has been assumed that the representatives of individual phages belonging to certain species are not very different from each other. However, it is impossible to suggest the identity of the behavior of all phages of the same species without their detailed comparison. Certainly, there are some limitations in the attempts for immediate applications of results obtained in the laboratory ("Petri dish studies") applied to conditions existing in a real infected wound. Nevertheless the basic processes in relation of *P. aruginosa* and its bacteriophages in the case of surface infections reveal a good similarity in both cases (biofilm formation, adhesion, alginate production, basic features of phage infections as transpositions, plasmid and phage migrations, *etc.*). We consider here several different deviations from canonical interactions of phages and *P. aeruginosa*, which may manifest themselves in the infected wound. In addition, we discuss the possible use of temperate phages in therapy and the influence of wound bacterial *P. aeruginosa* strains on phage growth. Some *P. aeruginosa* temperate transducing phages ((B3, G101, F116) [43] reveal a natural ability to infect strains as *Burkholderia cepacia* which may be considered a proof of common origin of the species [44].

5.1.3. Pseudolysogeny in the Case of phiKZ-Like Phages

PhiKZ (access number NC_004629) is a very special giant phage, which has the specific structure of capsid with a spiral inner body [45,46]. The phiKZ-like phages active on *P. aeruginosa* were classified into three species (phiKZ, Lin68, EL) [47,48]. Different phiKZ-lke phages are permanent components of commercial phage therapeutic polivalent mixtures. One of the attractive features of this phage species when used in commercial preparations is its capability to produce high final yields. The phiKZ-like phages are a very common component of phage biota in water and soil, and may be isolated in different regions [49]. PhiKZ was found to contain several genes coding for orthologs of proteins of pathogenic prokaryotes, unrelated phages, and some eukaryotes [50]. It is not known up to now, whether these genes belong to the phage genome and code for proteins essential for phage development or whether they represent "genetic noise", captured by the phage during migration and potentially playing a detrimental role.

These issues are of interest for several reasons. First, commercial mixtures used in phage therapy sometimes show a regional specificity. However, phages of species phiKZ, isolated from geographically distant regions, have just small genetic differences from phage phiKZ. Each of the phages has a quite broad (more than 20%) lytic activity spectrum. Such phages are extremely useful to kill specific bacterial strains resistant to other phages. We have found that the drastic differences in interactions of phiKZ-like phages with their host are dependent on the multiplicity of infection (m.o.i.). In a one step growth cycle experiment (m.o.i. in a range of 1–5 particles), phiKZ behaves as a typical virulent phage. All infected cells were killed with liberation of a not very high number of phage particles [45]. However, there was a significant increase of m.o.i. of sensitive bacteria with phiKZ-like phages, which arises when they are grown in biofilm conditions. As a result, these phages produce huge amounts of progeny particles, which exceed phage concentrations many times in the case of low m.o.i. In the genomes of all phiKZ-like phages that have been sequenced up to now, no gene coding DNA-polymerases have been found as closely related with the ones described earlier [50]. Thus, at first it was suggested that phiKZ-like phages could use bacterial DNA polymerase for their replication. Recently, with the use of modern computational analysis, sequences have been found supposedly coding some proteins, which have a distant similarity with some DNA-polymerases domains, which were found in genomes of several different phiKZ-like phages [51,52].

It is assumed that the mechanism of replication of phiKZ-like phages requires two components, and probably the transition of multiply infected cells into the pseudolysogenic state may be associated with the change of conditions in DNA polymerase activity. For instance, this may be due to deficit of a bacterial component which is necessary for replication. Anyway, the bacterial cell recognizes multiple infection with phiKZ-like phage particles and stops servicing phage replication. This is one of the possible explanations. However, there is another opportunity to develop an adequate explanation for induction in a cell's pseudolysogenic state after multiple infection with phiKZ and EL-like phages. The genomes of these phages are coded proteins similar to repressors of some temperate phages [50,53]. It is possible that the activity of the gene's products expressed from a small number of genomes (in low m.o.i.) is not enough to stop or to slow down the lytic development of these phages. An increase in the multiplicity of infection leads to an increase in repressor concentration (more template DNA), which stimulates transfer to a lysogenic condition. In this model, the cell estimates the level of repressor activity and prevents phage development, not with a common mechanism (blocking of a specific site in the phage genome), but with turn-off of bacterial functions necessary for lytic development. In one way or in another, it is evident that in conditions that arise after the infection of cells with high multiplicity, intracellular development of virulent phage is turned off for a while.

The phiKZ-like wild type phages are not appropriate components for phage therapeutic mixtures. These phages can lead to HGT (because pseudolysogenic cells can support development of other phages, including temperate ones, transfer of plasmids, and transduction). We have isolated mutants of phiKZ and EL, unable to induce pseudolysogeny and, moreover, actually killing pseudolysogenic cells. The mutants have properties of classical virulent phages. Interestingly, phiKZ and EL behave like phages with repressors of different specificity (with phiKZ revealing dominance) [54], see Figure

2. Probably, the substitution of virulent mutants of phiKZ-like phages for wild type phages in therapeutic mixtures decreases the possibility of HGT [55,56].

Figure 2. Phylogenetically related bacteriophages phiKZ and EL have repressors of different specificity. Virulent mutant ELvir5 cannot lyse phiKZ wild type infected pseudolysogenic cells (plated on a lawn of bacteria *P. aeruginosa* PAO1).
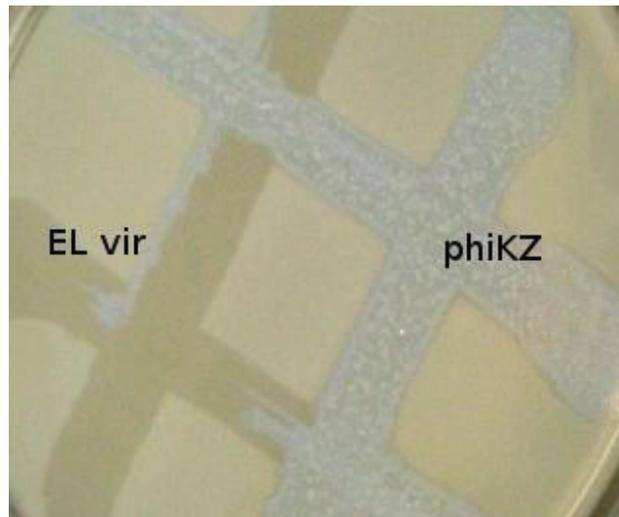
In Figure 3, the pseudolysogenic growth of *P. aeruginosa* PAO1 cells, infected by phiKZ with high m.o.i. in a Petri dish with nutrient media is shown. Besides the opalescent "bluish" growth of colonies with cells in pseudolysogenic conditions, there are some other additional features, which are of interest. First, material taken from the "bluish" (opalescent) initial growth (Figure 3a), after replanting, produces highly viscous colonies (Figure 3b). One of the possible explanations for such unexpected results may be the conversion of infected bacteria into an alginate producing state. Indeed, as it has been shown, the genes encoding the alginate biosynthetic enzymes are clustered in a single operon, which is under transcriptional control [57]. It may be very important for the activator of the alginate operon is AlgZ, a proposed ribbon-helix-helix DNA binding protein, which reveals some similarity with the repressors of phiKZ and phiEL phages (ORF 196 for phiKZ and ORF 163 for phiEL). Thus, overproduction of phage repressor proteins can be a reason for the activation of AlgZ operon with corresponding alginate overproduction. The second interesting conclusion that can be made from Figure 3c is that after additional prolongation of incubation for several days around the initial growth of opalescent colonies, a growth of phage sensitive colonies with a bluish border can be seen. This means that the sensitive bacteria and phage are able to move a certain distance from the place of initial planting. Phage particles are incapable of active movement. We suggest that phage particles were transported by bacteria resistant (at least for a time) to phage killing ability. One possible reason may be the blocking of intracellular phage development or interference with phage DNA injection (transfer of phages particles on the cell, as a "rider"). The most sensitive bacteria are killed in the process

of moving to the border of the colony, as evidenced by a high concentration of phages in the intermediate area.

Figure 3. Properties of pseudolysogens which arise after infection of *P. aeruginosa* PAO1 cells with phage phiKZ. (**a**) the growth of phage phiKZ after two days of incubation, the appearance of opalescence in growth of pseudolysogens; (**b**) replanting of the material taken from growth seen on "a" on the surface of the cultural solid medium; single colonies formed with pseudolysogens after two days of incubation. The colonies produced a lot of mucous material; (**c**) segregation of phage sensitive bacteria and their "runaway" of pseudolysogenic colonies.
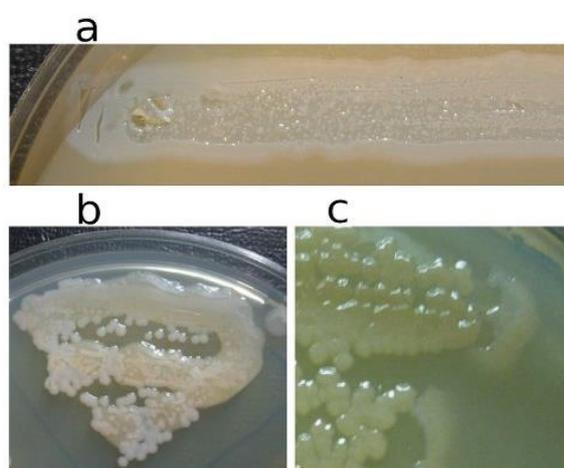
Pseudolysogenic condition for cells infected with phiKZ-like phages at high m.o.i. in natural conditions may have a biological sense of increasing final phage production. However, the use of these phages for phage therapy in infected wounds may increase viscosity by release of large amounts of DNA. This may obstruct the penetration of other phages or of antibiotics.

5.2. Pseudovirulence

The properties of infected bacteria have the ability to influence the development of temperate phages. In some cases temperate phages of P. aeruginosa cannot accomplish lysogenization. For instance Les-phenotype *P. aeruginosa* cells (mutations in control of the recombination genes) cannot be lysogenized with phage D3 [58]. This is a general effect and it can be found in the temperate phages of other bacterial species. For example, the temperate phage GIL01 *Bacillus thuringiensis* has no ordinary C1 repressor to maintain a lysogenic state. Presumably the LexA protein of host bacteria binds to specific 14-bp palindromic sequences within the promoter region of the phage. This prevents phage genes expression in lysogenic bacteria and provides the switch necessary to enter lytic development. Thus any damage of LexA protein creates the condition for lytic development of the phage [59].

5.2.1. Pseudovirulence of Phages with Mosaic Genomes and Migration of Non-Transposable
Temperate Phage

In the course of classification of phages growing on *P. aeruginosa* PAO1 based on the evaluation
levels of genome homology, all temperate phages were assigned to two groups [60]. One group
included four species of transposable phages and the other one six species of non-transposable phages.
The genomes of temperate non-transposable phages of different species showed a low level of DNA
homology in pairwise comparison. The level of homology varied when comparing different phages
from different species. Sequencing the genomes of several phages of this group confirmed the
presence of homology between them [61,62]. A mosaic structure of genomes was found in general,
fairly common in phages of different species of bacteria and usually found in phages active in related
species of bacteria [63]. However, there are some examples wherein genomic mosaicism arises
through the exchange of genes or blocks of genes between phage of unrelated bacterial species. This
suggests the possibility of migration of phage genomes or their fragments between distant species of
bacteria; although this transfer could happen before full speciation of the hosts.

We consider several such cases to support this point as well as to look from a different point of
view. In study [64], evidence was presented of significant relation at the level of genomes of two
temperate phages active on unrelated bacteria, namely *P. aeruginosa* phage phi CTX and *E. coli* phage
P2. There are not a large number of genome fragments with complete homology in phiCTX and P2
genomes but pronounced similarity at 28.9%–65.8% is observed in some ORFs. As shown by
comparison of the total structure of the genomes of these phages, a high level of similarity is also
exhibited. Besides, very importantly, phi CTX and P2 have some similar phenotypic traits, including
the similar capsid structure, the lack of response to the effect of inducing agents, and the important
feature to accomplish interspecies migration, which is the use of similar lipopolysaccharides for
adsorption. Authors of the study suggested that in this case the introduction of intestinal phage P2 in
pseudomonade has occurred and considered the case as a clear proof for the ability of phages to
overcome interspecific and intergeneric barriers in bacterial hosts. The phage phiCTX is one of the few
phage species that is related to the function CRISPR mechanism in *P. aeruginosa* (see Chapter
"Temperate *P. aeruginosa* phages and CRISPR effects"). Evidence supporting interspecies migration
appears as a result of the isolation and study of new bacteriophages. The phages of *P. aeruginosa*
described in [65] are the first representatives of a novel kind of *P. aeruginosa* phages having a
similarity in genome structure with N4-like viruses active in *E. coli*. The finding supports the notion of
interspecies migrations of bacteriophages.

Another example of a mosaic genome that arose due to exchange with fragments of the genomes of
phages, which belong to different species, including phages active on an unrelated bacterial species, is
the genome of phage phi297 (NCBI access number NC_016762) [66]. A significant part of the genome
phi297 exhibits a high level of homology to the genomes of phages F116 and D3 of *P. aeruginosa*. At
the same time, some special properties F116 and D3 are only partially manifested in phi297. The
bacteriophage D3 performs lysogenc conversion of PAO1 surface antigens (serotype O5 in serotype

O16) [67]. Such conversion requires the activity of three genes, located in a single fragment of the genome [68]. These ORFs encode three proteins: alpha-polymerase inhibitor; O-beta-polymerase and O-acetylase. Their sequential action leads to a change in the surface structure of lysogenic bacteria and in adsorption specificity. As it turned out, phi297 contains only gene O-beta-polymerase from this group of genes [66]. Apparently, this is not sufficient to ensure complete conversion, and as a result, the phage D3 can be adsorbed onto the surface of the cells of *P. aeruginosa* (phi297) and produce small plaques on its lawn. Further, the two regions homologous to F116, with the coordinates of 1–544 bp, and 47,119–47,438 bp in the genome phi297 match phage integrase gene F116 and its fragments [62]. At the same time, F116 prophage is not integrated into the bacterial chromosome and is present in the cell as a plasmid [69]. Usually, interaction between homologues of the bacterial partitioning proteins ParA and ParB, coded with a plasmid of low copy number, is necessary for stable maintenance and distribution of copies of plasmids in the division of bacterial cells. Phage phi297 genome contains a homologue of parA gene. However, parB is not found in the phi297 genome. Perhaps, lack of parB is the cause of an unstable lysogenization by phi297. Having said that, there is a view that the ParB function is not required for proper distribution of plasmids to daughter cells [70]. Usually, in the newly isolated phages, after sequencing the genome, it is commonly found that plenty of genes coding gene products have no similarities with functions of other gene products in databases. In genome of phi297 there were just eight of such genes. However, even in those cases where phi297 gene products show similarity with some of the gene products available in databases, their real functions are unknown. This is the common case. It can be assumed that in such cases, to determine a value of a gene product for phage viability, it would be useful to apply, along with other approaches, direct genetic research; for example, the selection of mutants of phage with subsequent allocation of the mutations to ORFs. The confirmation of phage genes flowing between these bacterial species has been found in the comparison of phiKZ-like bacteriophages active on different pseudomonades and in the recently discovered large bacteriophage SPN3US active on *Salmonella* [71]. Perhaps this indicates the presence of open-channel exchange between *Salmonella* and *Pseudomona*s. In the structure of phi297 genome there is also definite confirmation of the possibility of migration of large blocks of genes between *Salmonella* and *Pseudomonas*. It would be important to find out the mechanism and ways of such exchanges (participation of unique bacteria, specific bacteriophages, plasmids, migration of native phage genomes, transduction, *etc*.) (see Figure 4).

5.2.2. The Behavior of Wild-Type Phage phi297 on Different *P. aeruginosa* Pathogenic Strains of Different Origin

When comparing the growth of the various phages on a group of clinical isolates from a burn center, it was found that wild-type phage phi297 grows on the lawns of some clinical isolates in a manner of a virulent phage (see Figure 5). One such strain, Che1, shows resistance to most phages used for the treatment of burn wound infections. Properties of *P. aeruginosa* Che1 have similar properties to Les mutants of *P. aeruginosa* PAO1 (see above), as described in phage D3. However in

the case when we accept that *P. aeruginosa* Che1 is indeed a variant of the Les-phenotype, it is necessary to recognize that its Les phenotype is repressor dependant because mutants of phi297 (as phi297ci, phi297vir) with diminished or absent lysogenization capability do not grow on this strain at all. The very existence of these bacterial isolates suggests that such generally accepted concepts, such as temperance and virulence of phages, may lose their ultimate significance in medical practice. Indeed, in the selection of phages needed to treat real infected wounds, it is better to abandon the dogmas that the use, in treatment, of temperate phages is impossible or undesirable because of the danger of HGT. If a temperate phage acts as a virulent in relation to a particular wound pathogen it is obvious that the probability of HGT in such a case will not be higher than for inherent genetic virulent phages.

Figure 4. Graphical qualitative layout of the comparative degree of relatedness among a group of phages which reveal some similarity with phiKZ, active on *Psedomonas aeruginosa* (phiKZ and EL), *Pseudomonas putida* (phage Lu11), *Pseudomonas fluorescens* (phage OBP), *Salmonella enterica* (phage SPN3US), *Pseudomonas chlororaphis* (phage 201φ2-1) and *Escherichia coli* (contigs PA3). The thickness and length of connection lines characterize levels of genome relatedness.
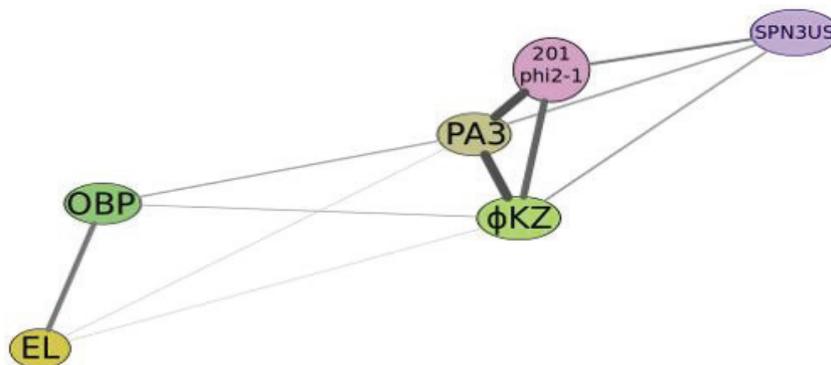
Figure 5. The growth of different phages on *P. aeruginosa* PAO1 and on clinical isolate *P. aeruginosa* Che1. Temperate phage phi 297w does not lysogenize cells of Che1 and behaves as a virulent phage.
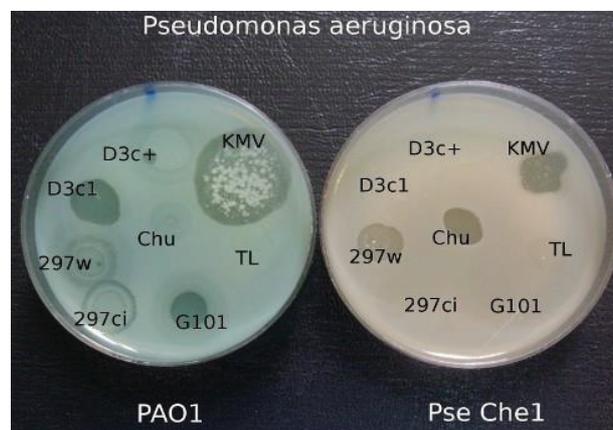
6. Growth of Phages in the Presence of Plasmids

The other examples of disturbances in canonical phage growth types of bacteriophages may be related to extremely diverse and specific effects of different plasmids in clinical isolates of *P. aeruginosa*. Conjugative plasmids can migrate between different strains of the same species and also between strains of unrelated species. The plasmids are active participants in the modification of pathogenic islands of bacterial genome transfer gene cassettes, controlling multiple antibiotic resistance, phage resistance or growth modification and genes influencing virulence and pathogenicity. We discuss here some plasmid-phage interactions and their significance for the choice of phages for therapy. The most differentiated activity in phage development inhibition is demonstrated by plasmids of IncP2 group of incompatibility. These plasmids, being specific to pseudomonades, are capable of recombination with IncP1 plasmids of a wide host range and thus can easily migrate to other species of gram-negative bacteria [72]. Such interplasmid recombination provides a powerful means for genetic diversity in plasmids. All IncP2-group plasmids inhibit phage growth, blocking intracellular phage development [73–75]. With the use of interspecies crosses between *P. aeruginosa* cells with IncP2 plasmids and *P. putida* it was possible to select plasmid variants with different inhibitive activity. As an example, plasmid pMG53 produced variants with at least six different types of phage growth inhibition. The other IncP2 plasmid, RpL11 also revealed high specificity in phage growth inhibition. It selectively reduces the frequency of lysogenization for temperate transducing phage G101 and transposable phages B39 and D3112 but not for transposable phage B3. It was found that phage D3112 functions, controlling establishment of lysogeny and the lytic cycle, were not expressed after infection of cells with RPL11. However, in the case when the cell carries both plasmid RPL11 and D3112 as prophage, there is no interference with repressor synthesis or with vegetative phage development after prophage induction. Thus, plasmid activity differentiates between the processes of primary integration after infection and that of reintegration of DNA after prophage induction [76]. Moreover, plasmids may recognize fine phenotypic differences among closely related phages with identical repressor immunity. In the presence in cells of plasmid Rms163 (IncP5), transposable phage B39 increase its lysogenization efficiency and as a result has efficiency of plating (e.o.p.) near 0.1, while mutants in c1 gene plate as clear plaques have e.o.p. of 1.0. Two other transposable phages with immB39 reveal a much more profound reaction on Rms163 and practically all infected cells become lysogenized. Mutations in their c1 genes of both phages restore lytic growth (e.o.p. 1.0). The site responsible for such differences in plasmid reaction for immB39 type repressor is located within the interval 1.1–3.9 kb of genome, being closely linked to gene cI [77]. Integration of *P. aeruginosa* transposable phages into plasmid RP4 induces different stable mutations [78,79]. Such hybrid plasmids can migrate among different gram-negative bacterial species, such as *Escherichia coli*, *Pseudomonas putida*, *Alcaligenes eutrophus* [80,81], they can be recipients for transposable phages of others species, such as Mu [82,83], and can influence the development of phages in other bacterial species [84]. Transfer of plasmid and phage genes capable of causing complicated genome rearrangements and dissemination among the strains as in the environment or as in the microbial wound community, may frequently be the cause of

unpredictable results (production of filaments by *E. coli* (RP4::D3112) at 30 °C, stability and growth of *E. coli* (D3112) with a single phage genome copy in chromosome) [85–87].

In a general sense, it is possible to consider plasmids and phages as a common gene pool, organized in different structures, designed by nature for the fine regulation of evolutionary events. It is evident that plasmid effects can restrict the possibility of using some phages for the purposes of therapy. However, it is possible to obtain variants of bacteriophages of particular species (mutants or natural isolates), which overcome the inhibitory effect of plasmids. One such example is an isolation of a new phage phiPMG1 (NCBI access number NC_016765), capable to grow in the presence of pMG1, IncP2 plasmid. However, as it happens, phage phiPMG1 can also lyse bacteria with other IncP2 plasmids. It was found that the phage has some other unusual features. Being closely related with temperate phage D3, it cannot make lysogenic bacteria. Results of phiPMG1 genome sequencing [88] have shown (see Figure 6) that in the central part of its genome some very complicated restructuring events have occurred. In spite of its repressor gene being undamaged, the phage cannot accomplish stable lysogenization as D3 does, but only induce a temporary lysogenic state in infected cells.

Figure 6. Comparison of phage genomes phiPMG1 and D3 (with Dot-plot program GEPARD) shows the complicated rearrangements in phiPMG1 genome regulatory region. phiPMG1 produces repressor but there are no stable lysogens.
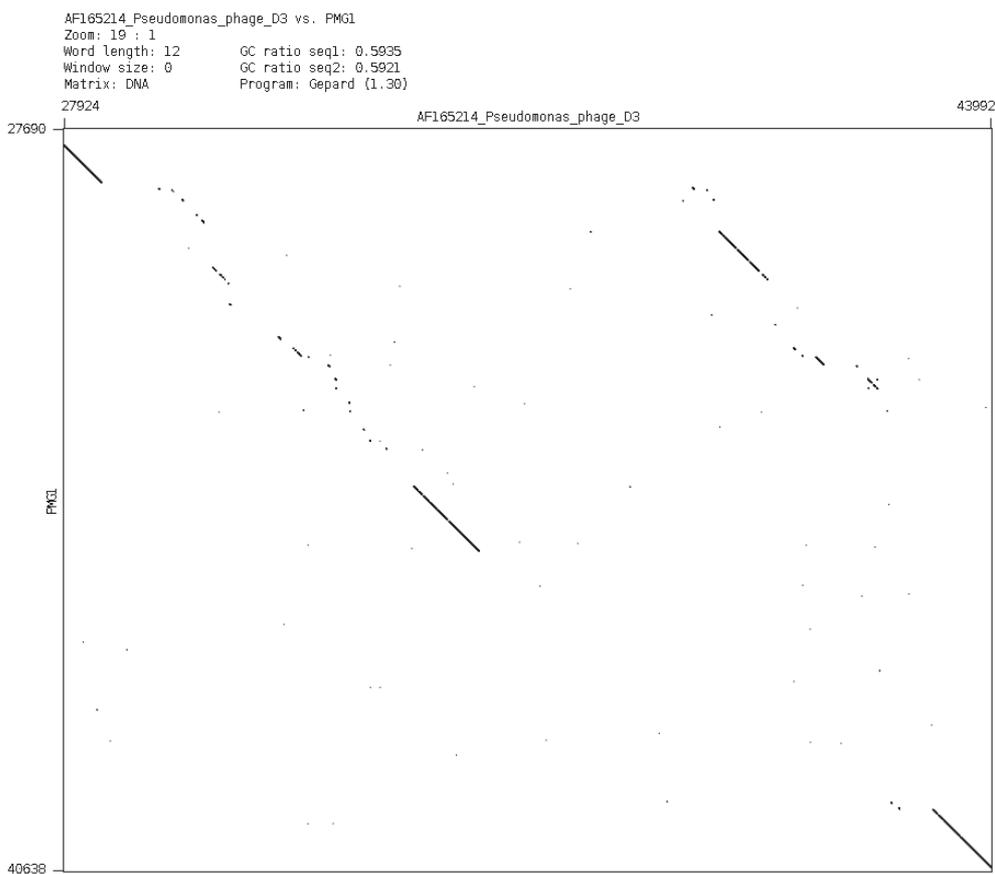
Such observed virtual virulence of the phage has arisen as a result of gene rearrangement in the regulatory region of the genome. Presence of plasmids can block lytic growth of recognized therapeutic phages. For example, plasmid RMS148 interferes with lytic development of phage phiKZ of *P. aeruginosa*. This fact has permitted the revelation of the transducing activity of phiKZ. It means that in the course of phage therapy that particular virulent giant phage may transduce in its own particle up to 5% of *P. aeruginosa* bacterial chromosome [89]. It is more than the largest of known pathogenic islands in genome of *P. aeruginosa* and more than the sizes of temperate prophages able to participate in construction or modification of pathogenic islands.

7. Possibilities for the Emergence of New Phages in the Course of Phage Therapy

It may be interesting to estimate the probability of the emergence of new recombinant phages as a result of the application of phage therapy to wounds infected with pathogens of different distantly related species. For example, *Pseudomonas aeruginosa* and species of *Burkholderia* complex often coexist in places of common infections (wounds, CF). For both bacterial groups similar temperate transducing phages are described. There are temperate phages specific for *Burkholderia sp.* growing on *Pseudomonas aeruginosa* [90] and there are strains of *Burkholderia*, which support the growth of temperate phages *P. aeruginosa* [44]. We see examples of interspecies migration for phages of several different types, including (1) transposable phages *P.aeruginosa*; (2) converting temperate phages related D3; (3) phage F116.

7.1. Transposable Phages

In our previous studies, we found plenty of different transposable phages of *P. aeruginosa* (all of them are temperate), which were distributed into groups of D3112-like phages (14 phages, one species) and B3-like phages (six phages, three species) [91–94].

D3112-like and B3-like phage genomes are composed in a similar pattern, but mostly from different non-homologous blocks. Phages belonging to different groups reveal only traces of DNA homology along most parts of their genomes. Some essential DNA homology can be found on the right ends of their genomes where genes responsible for specific adsorption to bacterial IV type pili are located [95]. Nevertheless, it has been possible to isolate several D3112/B3 hybrid phages with very strange genome structures. They have arisen through an even number of crossovers with low frequencies. The comparison of D3112 and B3 genomes by dot-plots has found just several of the 30–50 bp long DNA homology sites distributed along the genomes. We suggest that these small sites of homology are used in interspecies recombination. The compulsory even number of crossovers may be explained by the possible incompatibility of specific gene modules in D3112 and B3 [96,97]. D3112-like and B3-like phages have a difference in their transposition specificity. Although phage D3112 has multiplicity of the integration sites in *P. aeruginosa* chromosome and plasmids [79], nevertheless, it does not induce auxotrophic mutations in lysogenized bacterial cells. However, phage B3 and its relative PM105 induce such mutations with different nutritional requirements [98]. We have suggested that D3112 has

a high transposition specificity and mostly uses sites that are located in the part of the *P. aeruginosa* chromosome controlling catabolism. Indeed, Rehmat and Shapiro [99] have localized different D3112-induced mutations in the amidase gene.

It was found that genome of D3112 may be expressed in different species of gram-negative bacterial species (*Escherichia coli*, *Pseudomonas putida*, *Alcaligenes eutrophus*) when delivered as a part of hybrid plasmid RP4::D3112 [80,81,85,86]. In the new hosts, the expression of D3112 can change some of the bacterial properties. Thus, the hybrid strain *E. coli* (RP4::D3112) grows well only at 42 °C but dies quickly at 30 °C (forming long filaments as a result of blockade of cell division).

The real reason is that D3112 in *E. coli* cannot establish stable lysogenic conditions but its transposase is active. However, at 42 °C D3112 genome transposition is blocked, and cells survive and form colonies [86]. It is possible to isolate quite stable *E. coli* strains where phage genome is inserted into bacterial chromosome with loss of RP4 plasmid. Such *E. coli* (D3112) clones produce a very low number of mature D3112 phage particles. It means that in such clones, a foreign phage becomes a component of *E. coli* genome and its stability is ensured by rare expression in single cells. It may be considered one of the examples of phage migrations. Such phage interspecies migrations (including migrations of plasmids with complete phage genomes or their fragments) may lead to emergence of new pathogens.

There are some other effects related to activities of transposable phages of *P. aeruginosa*. It has been shown that after thermal induction of *P. aeruginosa* cells carrying heat inducible prophage D3112 cts15, it is possible to observe plenty of different survivors producing highly mucous colonies [100]. Most of them quickly (after 1–2 replatings) lost their mucous phenotype but some of the survivors showed good stability.

Frequently among related transposable phages the effects of mutual influence on vegetative growth may be found. This is usually found in the interaction of phages of the same species, but with different specificities of repressors; for example, the suppression of the vegetative growth of phages with immB39 on the lawns of bacteria carrying prophage with immD3112. The suppression was found to be determined by the activity of locus *cip*. The *cip* is located in the genome of D3112 in the range of 1.3–2.45 kb and reveals its inhibitive effect only in lysogenic bacteria. In the case of multiple lysogeny the suppression increases substantially. There have been isolated mutants of phage B39 which can not be inhibited with Cip-activity. Mutations in gene cip which lost the ability to suppress B39 also can be selected [101,102].

As was found in the study of the genomes in different *P. aeruginosa* transposable phages, all of them have a pronounced modular structure that shows evidence of the possibility of free exchange with gene modules between the genomes of phages of the same species. Interestingly, in some individual genes, their composite nature can be detected. So, the number of phage repressor genes is composed of three sub-modules, which are supposed to control the various functions of repressors [103,104]. Apparently, the genomes of *P. aeruginosa* transposable phages, even those which now show no homology along their genomes genes (phages of D3122 and B3 groups), have the single pattern of a genome of an ancient predecessor. This is confirmed firstly by a similar sequence in

groups of genes controlling similar functions; secondly, by the presence of a considerable number of genes with good homology which control the flexible structure of the tail and its ability to adsorb to bacterial pili. However, transposable phage BcepMu, active on *Burkholderia cenocepacia*, has a different origin, because it reveals some common features with phage Mu of *E. coli* (they have a contractible tail), although they have differences in specificity transposase compared with the phage Mu [105]. Likely divergences in the evolution of different types of phage-transposons for gram-negative bacteria occurred at the moment of selecting the main host. It is possible to suggest that most of the time the genome evolution of BcepMu like phages was in bacteria belonging to the family Enterobacteriaceae. In the genome BcepMu are present homologs of many genes of different strains of salmonella prophages.

We can assume that the differences in the genomes of modern *P. aeruginosa* transposable phages belonging to different groups arose as a result of their numerous migrations. This is confirmed by the results of genome annotation of transposable phages. Thus, in D3112 phage genome there were found genes which are orthologs and not only genes of pseudomonads, but of other bacteria or phages. For example, ORFs, 18, 19, 24 and 26 of D3112 encode proteins similar to protein defective prophages PNM1 PNM2 and two different strains of *Neisseria meningitidis* MC58 and Z2491, the next group ORFs 27, 28, 29, 32, 34 encodes proteins similar to proteins of the phage Mu, *etc*. [106]. In case of phage B3 it was found that its proteins have mostly another origin, revealing relatedness with proteins of phage Mu, BcepMu or *Salmonella typhimurium*. In accordance with results of phage B3 genome annotation, not less than 10 genes have their origin from genes of other phages of different bacterial species: five genes of salmonella phage P22, five genes of two BsepMu-like phages and one gene is a homologue of the gene in salmonella phage Sti3. Genes found in phage B3 genome reveal different levels of relatedness (on the level of controlled products) with such potential intermediate bacterial hosts as *Bordetella bronchiseptica*, *Burkholderia cenocepacia*, *Escherichia coli*, *Haemophilus ducreyi*, *Haemophilus influenzae*, *Neisseria meningitidis*, *Salmonella enterica*, *Vibrio cholerae*, soil species and plant pathogens of different families including non-pathogenic and pathogenic strains with various ecological niches [107]. In this respect, it should be interesting to continue the study of two other B3-like species of transposable phages, active on *P. aeruginosa*, but with a significant difference of B3 in the level of DNA homology [93,94]. It may be especially interesting considering the significance of *P. aeruginosa* transposable phages in CRISPR-effects (see later).

The origin of several phage genes in D3112 and B3 cannot be determined through annotation. To elucidate their functional significance in phage development it is necessary to accomplish detailed phenogenetical studies (directed mutagenesis of such genes, looking for mutants with a specific phenotype and study of their behavior under different conditions).

7.2. Converting Temperate Phages of P. aeruginosa

Temperate *P. aeruginosa* phage phiCTX carries a cytotoxin gene which is expressed in prophage state. The phage has a large genome homology with *E. coli* phage P2 as a result of highly probable

interspecies migration. Both phages are identical in particle morphology and their genomes have similar structures. However, the fact that genome of phiCTX is similar in GC-content with *P. aeruginosa* suggests that the divergence of P2 and phiCTX occurred a long time ago [64].

D3 is the other converting temperate phage of *P. aeruginosa*, which in the prophage state changes the structure of the surface lipopolysaccharides, and prevents adsorption of several bacteriophages, including D3 itself [61,68]. Phage phi297, showing DNA homology to DNA of D3 and of F116 [61,108], also has converting activity. The distribution of DNA homology regions in genome of phi297 with genomes D3 and F116 is shown in Figure 7. The mosaic structure in the phi297 genome is evident. Phi297 inherited from phages D3 and F116 just some of the genes that give these phages unique features, including the ability of D3 to perform bacterial surface modification and, possibly, the ability of phage F116 to be plasmid in prophage conditions. As a result, phi297 carries only a partial modification of cell covers (as shown by only partial blockade of phage adsorption). At the same time, as it is proposed, inheriting only one of the F116 genes, necessary to support prophage as plasmid, phi297 phage does not form stable lysogens.

Figure 7. Genome of phage phi297 has homology regions of different sizes and locations with D3 and F116 genomes (yellow: phi297 own genes; green: fragments of homology with D3 DNA; blue: fragments of homology with F116 DNA).

We have found that phi297 is capable of recombination with the bacterial PAO1 strain chromosome, producing lytic non-reverting variants, incapable of accomplishing lysogenisation. The resulting hybrids exhibit lytic properties towards certain groups of clinical isolates and phage resistant mutants arising after use of commercial phage mixtures [66,88]. We consider that the addition of such lytic derivates of temperate phages into commercial phage mixtures for the cure of infected wounds may be useful for enlarging their spectrum of lytic activity. The lytic activity of the hybrid phi297vir will stimulate the search for other similar possibilities. It will require more detailed studies of temperate phages of different species and the isolation of new species of temperate phages.

7.3. Phage F116: Interspecies Wanderer?

Bacteriophage F116 is of special interest for the assessment of the possibility of experimental and inter-specific migration and study of its mechanisms. The fact that F116 genome fragments (or an intact genome of other phages with great similarity to the genome of F116) were found in an integrated state in *Neisseria* support the need for special attention on the further study of F116-like phages. Indeed, pathogenic *Neisseria gonorrhoeae* strains do not support growth of any of the presently known tailed phages. However, as a result of careful studies in bacterial genomes, several regions were found

showing similarity to genes of different phage species. One of these regions, NgoPhi2 is unique in showing a very high level of similarity with the genome of the phage F116 [109]. Genes of prophage, which reveal relatedness with F116, are apparently intact, and exhibit multiple effects, including the inhibition of the growth of *E. coli* and the propagation of phage lambda. The repressor in the NgoPhi2 region was able to inhibit transcription genes of *N. gonorrhoeae* and of phage *Haemophilus influenzae* HP1, and the gene for choline replaces the function of the homologous gene of phage lambda. After induction of cells with mitomycin C and microscopy of the supernatant, complete phage particles were found. It can be assumed that the phage genome exhibiting affinity with F116 was introduced into *N. gonorrhoeae* by an act of HGT (in the course of direct infection with phage particles, or in the state of prophage plasmid in the conjugation process or further being integrated into a transmissible plasmid).

7.4. Evidence for Migrations of phiKZ-Like Phages

Bacteriophage phiKZ active to *P. aeruginosa* reveals such features as large particle size, large genome and a spiral formation, the "inner body", and a unique packaging of DNA as spools of thread [45,46]. Later two other species were described of giant phages of *P. aeruginosa*, EL and Lin68, which are not distinguishable from phiKZ in particle morphology [47,48]. Recently, structural proteomes of phiKZ and EL were studied in detail [110]. However, all new phages exhibit significant differences from phiKZ. Phage EL has no detectable genome homology with phages of the other two species, and differs from them in GC-composition and in size of the genome [53]. Phages of another species, Lin68, show DNA homology of a low level with phiKZ in one of the restrictional DNA fragments [47]. In addition, phages of species Lin68 are unable to grow at 42 °C, and their mature particle reveals more instability than phiKZ particles at 60 °C. The phages of this species exhibit weak lytic activity on the lawn of psychrophilic bacteria *P. fluorescens*. It is of interest to compare the genomes of phages in the species Lin68 with phiKZ-like phages of other psychrophiles. The phiKZ-like phages can persist in pseudolysogenic clinical isolates in wounds. Thus, phage Che, a new EL-like phage, has been found in a clinical isolate from the Chelyabinsk burn center. *P. aeruginosa*, at least in other bacteria; similar phages were not found for a long time. However, as has been shown, bacteriophages of similar morphotype, showing signs of kinship with phiKZ, can be found in other bacterial hosts [48,51,111–113]. It is possible too that phiKZ-like phages will be found among other large phages of distant bacterial species. A recently described giant bacteriophage specific to *Salmonella* [71] and another one for *Ralstonia solanocearum* [114] have similarities with phiKZ-like phages. Apparently, the ability of all phiKZ-like phages to establish a pseudolysogenic state in infected cells has evolved from an ancient temperate phage as a result of successive multiple migrations between different hosts. If it is possible to confirm specific DNA packaging for these phages as inner bodies, as in phiKZ, it will prove not only the existence of a new phage superfamily, but also the existence of the open gene exchange channel between *Pseudomonas*, *Salmonella* and soil types of bacteria for this phage superfamily.

8. Optimizing the Selection of Phages for Therapy

The detailed phenogenetical study of a phage is a compulsory condition for its acceptance for direct use in phage therapy or inclusion into phage therapeutic mixtures. Although in each case the approach for this study may be different, the first aim is to show that a selected phage is genuinely virulent. In case the phage virulence is derivative of a temperate phage, it is necessary to prove its incapability to revert into a temperate condition in different situations (as multiplicity of infection, temperature, presence of plasmids or temperate phages, *etc*.). For example, the giant phiKZ-like phages of *P. aeruginosa* are used in various commercial mixtures [108] due to their wide spectrum of lytic activity. These phages in standard single-stage growth cycle experiments behave as typical virulent phages by lysing all infected bacteria, but at a high multiplicity of infection, bacteria continue dividing over several days (see above). During this time, the infected cells can be recipients of HGT through the activity of plasmids or temperate phages, which often occurs in wound microbial communities.

The number of genuinely virulent phage species for *P. aeruginosa* is quite limited. There are only nine species of different families, and, moreover, some species may have limited usefulness. In the list below are the phage species (in parentheses are the minimum and maximum size of genomes of the species, where it is known). Family Podoviridae: (1) the species of phiKMV-like (42,954–43,548 bp); (2) species Luz24-like (45,503–45,625 bp); (3) species N4-like (72,544–74,901 bp); (4) species LUZ7/PEV2*; family Syphoviridae: species (5) M6-like (58,663–59,446 bp); family Myoviridae: (6) species PB-1-like (64,427–66,530 bp); (7) species phiKZ-like (280,334 bp); (8) species Lin68 (is not sequenced); species (9) EL-like (211,215 bp). Different manufactures can use one or the other species. However, it is evident that the number of species is not unlimited.

Of this list phiKZ-like phages (two species) can be considered only as conditionally virulent. Virulent phages of other species—Luz24 and TL—exhibit high relatedness with temperate phage PaP3 [115,116]. So far the most promising species for phage therapy of P. aeruginosa infections are phages of species PB1-like phages, which exhibit significant variability on the basis of the specificity of adsorption [117–119], and species of phiKMV-, N4-, LUZ7- and M6-like phages. Bacteriophages of species Lin68 can have only limited application, being sensitive to the development of human body temperatures. We can conclude that the phage potential of *P. aeruginosa* is limited. Variations in lytic activities for phages inside each of the virulent species are also limited. A sequenced set of phages and their number are altogether insufficient to maintain long-term phage therapy. Obviously, the need to expand efforts to find new species of virulent phages exists. Moreover, there must be some coordination of specialists working with phages and interest in the development of phage therapy to overcome the randomness in the selection of phages for sequencing and phenogenetical studies.

For each prospective therapeutic phage it is desirable to estimate the probabilities and diversity of phage-resistant mutants. According to our observations both evaluations vary greatly in dependance on phage species and (obviously) on specific hosts: clinical isolates and laboratory variants of the standard host.

9. At First Glance, the Phages with a Broad Spectrum of Lytic Activity Are Preferred for Therapeutic Blends, but only at First Glance

The use of sophisticated phage cocktails, including many strains of phages with a broad spectrum of lytic activity against different bacterial pathogens, has as its purpose the fast interruption of infection without prior isolation of the pathogen. This is justified only in cases where the infection and the loss of time may create a life threatening situation. However, it may be that the long-term presence of pathogenic bacteria in the wound are caused by the unique bacterial strains that are resistant to most of the phages in the therapeutic mix. Because there is no previous information about strains in the wound, before applying emergency treatment with a phage mix, in fact, the attempt at treatment iscarried out blindly and if the first application of phages does not give positive results, it means that time is lost. However, in the majority of cases of wound or nosocomial infections due to *P. aeruginosa*, there is usually no such extreme urgency.

Typically, the clinical laboratory identifies pathogenic bacteria bv determining their sensitivity to various antibiotics and only if antibiotic application is ineffective, will the decision be made on the use of bacteriophages. Thus in the case of wound and nosocomial infections in different locations due to isolated strains, more rational is the use of a minimal set of phages, even a single phage with a narrow spectrum of lytic activity capable to lyse strains isolated from the patient. The choice of a set of phages or a single phage against a particular pathogen is not a more labor-consuming procedure than a standard check of bacterial sensitivity to antibiotics with paper discs. Such a restriction in the number of phage species is intended to limit the accumulation of mutant pathogens that are resistant to the most valuable phages with a broad spectrum of lytic activity and save them for use in really extreme cases. Thus bacteriophages with a narrow spectrum of lytic activity can also be very useful in phage therapy. We consider the properties of several of these phages.

10. Description of the Properties of New Phage in the Possible Application of Phage Therapy

10.1. The Properties of a New Bacteriophage CHU Pseudomonas aeruginosa

Phage CHU is one of the newly discovered phages for *P. aeruginosa*. This is a relatively small (diameter of the head is near 50 nm) DNA containing tailed Syphoviridae phage. The unique feature of the phage is an unusual spectrum of lytic activity (trace growth on PAO1 strain and good growth on several strains producing alginate). There are two such strains, Pse163 and CF013A, isolated from CF patients (from the collection of Prof. M. Vaneechoutte, Laboratory Bacteriology & Virology University Hospital Ghent, 9000, Gent, Belgium). These strains differ in their phage sensitivity and other properties (Pse163 is lysogenic for a transposable phage related with D3112). Phage CHU grows on lawns of washed alginate Pse163. The phage has no alginase activity, as there is no visible destruction of slime around the spots of phage growth. There is, however, a difference in the character of phage CHU growth on lawn of strain CF013A. After a prolonged (three days) incubation around a lysis spot CHU on the lawn of CF013A was formed, (see Figure 8). The reason for this effect may be related to the difference in properties of alginate producing strains. It was confirmed in the study of four additional stable P. aeruginosa alginate producers exhibiting selective sensitivity to phage CHU, which

were found among clinical isolates of the burn center of the Chelyabinsk Regional Hospital. All these four strains grow similarly, being plated on the surface of the agar medium. However, when plated in semisolid agar layers, two of these strains do not produce a visible layer of alginate but the other two strains continue to produce alginate, visible on the surface of the upper agar layer. Bacterial strains, which do not produce visible alginate, on being placed into semisolid agar of the upper layer, reveal alginate around phage CHU spots of lysis. However, the reasons for the difference in properties from two kinds of alginate producers, remains unknown. As the last four strains from the burn center were isolated at different times from different patients they may be regarded as hospital pathogens. It is possible the differences between them arose in the burn center.

Figure 8. The raised slimy ring on the CF013A lawn around the phage CHU lysis spot.

A feature of phage CHU is the high level of instability. The best differentiation of plaque types was achieved with use as lawns of strains *P. aeruginosa* 8–20 and 2–10 (obtained from Professor C. Pourcel, GPMS Institut de Genetique et Microbiologie Bat 400 Universite Paris-Sud 91405 Orsay cedex, France), see Figure 9. Up to now, despite repeated attempts, we could not isolate stable plaque derivatives of phage. In this relation CHU reveals some similarity with phage TL, described below.

Figure 9. Growth of CHU phage variants on the lawn *P. aeruginosa* 8–20.

The use of this phage can be useful in open infected wounds against strains producing alginate.

10.2. Properties of a New Bacteriophage TL

Bacteriophage TL genome sequence (carried out in the laboratory of Professor L. Kulakov, (The School of Biological Sciences, The Queen's University of Belfast, Belfast BT9 7BL, Northern Ireland, United Kingdom), annotation made in our laboratory shows a high degree of relationship, with phage Luz24 [115,116]. TL reveals a relationship (as Luz 24) with the temperate phage PaP3. This leaves no doubt as to the common origin of the three phages. TL has poor growth on lawns of strain PAO1. A PAO1 mutant ELR2 has been isolated which is resistant to many phages, but supports good growth of TL. We suggest the loss or modification of surface adsorptional receptors led to the availability of receptor for phage TL [35,117–119]. Phage TL is produced on the lawn of ELR2 plaques with a different appearance, which cannot be stabilized, see Figure 10.

Figure 10. Growth of TL phage variants on the lawn *P. aeruginosa* ELR2.
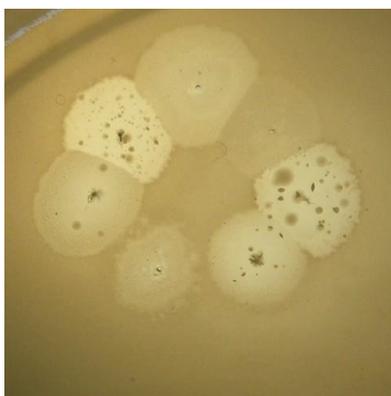
After the genome annotation, it was found that the TL genome contains gene coding for transposase gene in the region 5,000–6,000 bp. The gene was not detected in genomes of related phages Luz24 (AM910650), see Figure 11, and PaP3. The appearance of different plaque morphology variants and their instability is may be due to the activity of the phage transposase. As TL does not produce lysogens, like Luz24, it can be attributed to the lytic phages. TL is capable of lysing bacteria resistant to other phages used in anti-*P. aeruginosa* preparations. Therefore, the introduction of TL into any standard phage mixture against *P. aeruginosa* may be very useful. Indeed the phage will reveal its activity when the wound *P. aeruginosa* pathogens become resistant to other phages in the mixture. Although the features of this phage confirm the importance of the proper selection of phages in the preparation of the phage therapeutic mixture, nevertheless the use of Luz 24 related phages in phage therapy requires a more thorough study.

Figure 11. Genomes of phages TL and Luz24 have a high level of DNA homology (DotPlot on UGENE program parameters: 60 bp, 85% similarity) but TL genome contains transposase gene in the region of non-homology with Luz24.
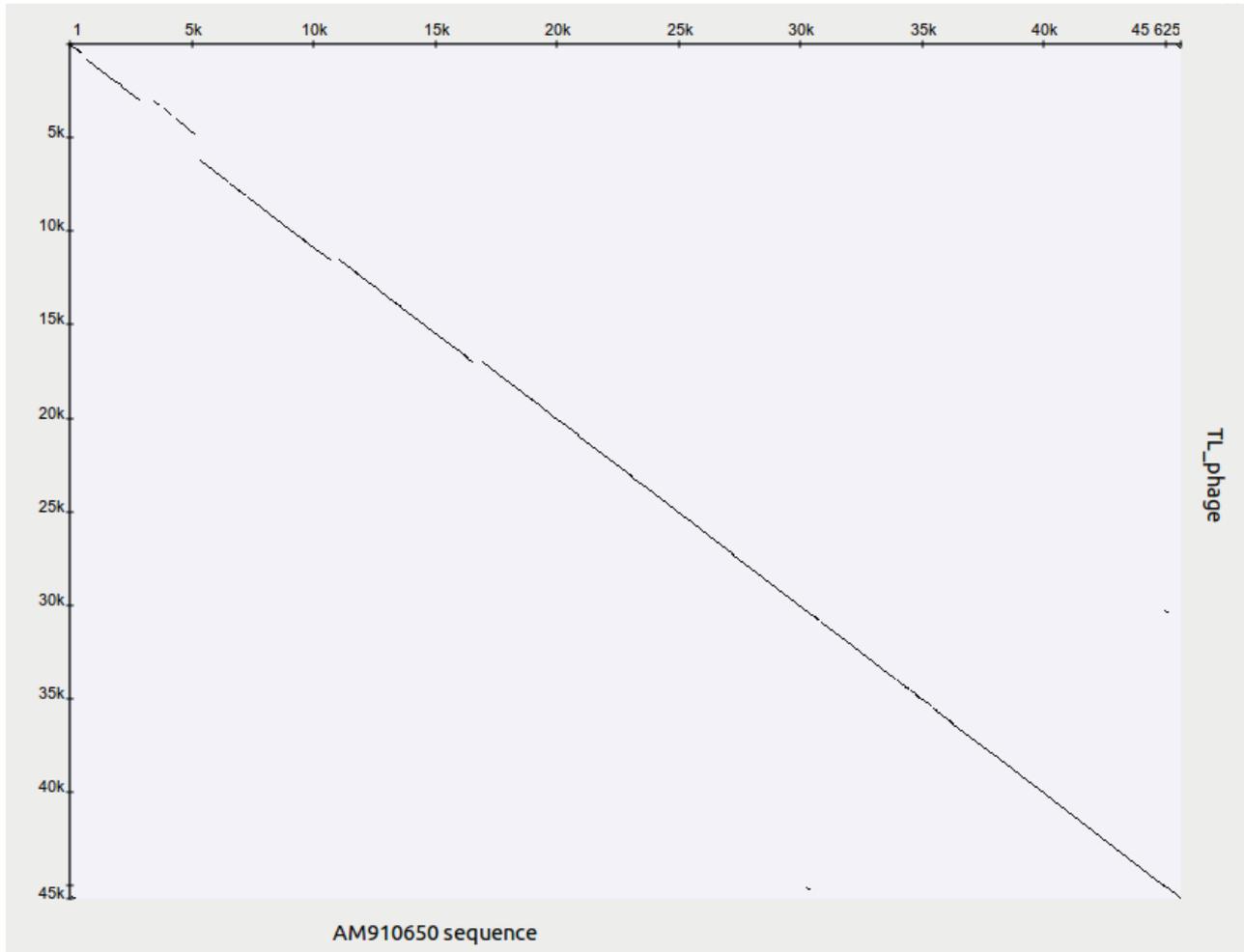
11. Significance of Specific Phage-Bacteria Interactions in Transfer to Personalized Phage Therapy

The mixed microbial community includes different types of bacteria (pathogenic and nonpathogenic) which may be in complicated mutual relations. For example, in the case of cystic fibrosis the frequent companion to *P. aeruginosa* is the bacterial species of the *Burkholderia* complex. One of the reasons for the persistence of *Burkholderia* is the presence of alginate [120]. There are described cases of acquisition, of the ability to grow at 37 °C, adhesiveness and cytotoxicity comparable with the corresponding features of *P. aeruginosa,* by psychrophilic soil pseudomonades as *P. fluorecens*, which were previously considered as harmless [121,122]. Furthermore, strains of *P. putida* with cassettes controlling antibiotic resistance in plasmid-borne integron from an *Achromobacter xylosoxidans* were isolated from several patients in the general intensive care unit,

which suggests the possibility of evolutionary interactions between commensals in the clinical conditions [123]. It can be expected that the frequency of finding such new species from commensals will expand as will their significance in wound bacterial communities (carrying with them a set of new genetic elements—phages, plasmids, transposons—that can perform HGT). Thus, there is a need to continuously increase the number and diversity of bacteriophages that are acceptable for the treatment of infected wounds.

Accordingly, the number of truly virulent phage species is usually small in the case of different bacterial species, including *P. aeruginosa*, which obviously limits the possibilities of phage therapy. Moreover, it is clear that the continued use of the same phage will quickly lead to the emergence of multiple phage-resistant bacterial pathogens. It is impossible to predict the time of active use of the described phage species in the case of widespread introduction of phage therapy. It is not evident that we must expect to reveal a sufficient number of new virulent phage species. In this regard, there is a need to understand and accept decisions on the possibility of using specially modified temperate phages in phage therapy.

12. Let Us Discuss the Possibility to Use *P. aeruginosa* Temperate Phages in Phage Therapy

Temperate phages active on *P. aeruginosa* make up a significant part of the phage potential of this species. Refusal to use temperate phage of wild types capable to lysogenize infected bacteria is understandable and evident. However, as we have shown, among the recognized phage species acceptable for therapy, there are phiKZ-like species whose behavior (lytic against pseudolysogenic) may greatly depend on cotrolled or uncontrolled conditions in their application. In the case of these phages, special virulent mutants incapable of pseudolysogenization have been isolated.

We consider that it is possible in certain cases to use some temperate phages after their inheritable modifications or even to use phages showing temperance in infection of standard laboratory hosts, but having obvious virulence for real wound strain. Of course this can be done in conditions of individual selection of specific phage only under the supervision of physicians and clinical laboratories. Quite a lot of temperate *P. aeruginosa* phages have sequenced genomes. The most studied phages belong to the Syphoviridae family, including species: (1) transducing cytotoxic phiCTX-like: phiCTX (35,580–36,415 bp); (2) D3-related: D3 (56,425 bp), PAJU2 (46,872 bp), phi297 (NCBI HQ_711984), phiPMG1 (NCBI NC_016765); (3) F116-(65,195 bp); (4) Transposable phages: (a) D3112-like (13 different isolates): D3112 (37,611 bp), MP22 (36,409 bp); DMS3; (b) B3-like phages (six phages) B3 (38,439 bp), HW12, PM105 [93]. In addition, there is a group of phages of different families and species, which should be studied in more detail as Myoviridae, P2-like-MP29, MP38, F10; Podoviridae 119X, PaP2; Syphoviridae: 73, G101, 160 and unclassified phage PA11.

Phage phiPMG1 may be an example of an earlier temperate (D3-like) phage, which became acceptable (in our opinion) for use in therapy as a result of spontaneous genetical modifications [88]. Our attempts to isolate revertants of the phage similar to D3-like wild-type phage were unsuccessful. We tested the activity of this phage compared to the activity of two commercial mixtures of different

manufacturers and showed that this phage successfully lyses spontaneous mutants of strain PAO1, resistant to phages in both commercial mixtures [108]. Phage phi297vir, which is a recombinant of phi297 wild type with bacterial chromosome also lyses resistant to commercial mixtures of bacteria, including bacteria in the pseudolysogenic state after multiple infections with phiKZ-like phages, such as bacteriophages.

The study of new species of temperate phages expands our understanding of their evolution [64]. For instance, the mosaic structure of phi297 genome occurred as a result of directly borrowing gene encoding products that are similar to the products of bacteria and phages of different taxonomic groups (including *Salmonella* and other enteric bacteria). This clearly supports the idea that genomes of different natural phages have arisen as a result of a combination of pre-engineered functionally related genes. Introduction to the temperate phage genome modifications, leading to irreversible virulence, can be used to expand the collection of therapeutic phages.

Of course, the likelihood of HGT using even non revertible virulent variants of temperate phages will be higher than in the case of truly virulent phages (because, for example, of the possibility of such a mutant recombining in infected wounds with related prophage in pathogenic bacteria). However, the immediate danger to the patient in the course of phage therapy will not be raised because the duration of phage therapy is usually not more than a few days. If during this time, the therapeutic effect is not achieved, it means that the mixture of phages in this particular case is not efficient and should be replaced [37]. Phage phi297vir behaves as lytical but it is not virulent, since it does not grow on bacteria which contain wild type prophage phi297w. Apparently phi297vir remains a functionally active operator.

13. Temperate *P. aeruginosa* Phages and CRISPR Effects

We have discussed the above possible migrations of intact phage genomes or their fragments between bacteria of different species. One of the obvious ways is the migration of phage genomes incorporated in the conjugative plasmids, as shown for transposable phages of *P. aeruginosa* (see above) and *E. coli* [123]. Another possible means for migration may be acceptable for phages by using for their adsorption type IV pili. Type IV pili are highly specialized structures on bacterial surfaces having immense significance in bacterial life in the expression of many phenotypes, such as motility, sensitivity to bacteriophages, natural genetic transformation, adherence and speciation for multicellular behavior. Modified bacterial pili of IV type may provide adaptation to specific tissues expanding the possibilities of the particular species of pathogenic bacteria and increasing their virulence [124–126]. The pili, being frequent appendages of different pathogens and adsorptional receptors for numerous phages (as F116, transposable phages of *P. aeruginosa*, possibly transposable phage specific for *Neisseria meningitidis*), may be used by such phages in interspecies migrations [127,128]. It is possible to assume that this interspecies migration path made the bacterial cells develop CRISPR as a specific mechanism of protection against expression of foreign genetic material, plasmids and phages. Loci CRISPR, control the synthesis of specific small RNA molecules that, under invasion by viruses or

other genetic elements block their expression, binding the foreign DNA. The number and variety of such phage spicers evidences the number of previous contacts of bacteria with different foreign DNAs. At the same time, there is evidence that the CRISPR mechanism, a multifunctional system response, manifests itself not through the blockade of phage or plasmid penetration to save individual cells subjected to attack, but by modifying other essential features. Infection bacteria *P. aeruginosa* PA14 with phage DMS3 inhibits biofilm formation and swarming motility, both very important bacterial features associated with the activity of type IV pili. After introduction, mutations in the CRISPR region of both biofilm formation and swarming in DMS3 lysogenized strains recovered [129].

Reports of the suppression of the growth of temperate phages are rather contradictory. Some studies clearly stated the suppression of the growth group of temperate phages, which (according to the results of the comparison with database) are mainly related to the species transposable D3112-or B3-like phages and some other species as F116 and phiCTX. In other works, the inhibition of the growth of phages is denied. Thus, in two moderate phages, MP29 and MP42, whose genomes are similar to those of *P. aeruginosa* transposable temperate phages DMS3, there was no change of swarming [130]. Such discrepancy in results from different research groups may reflect subtle differences in the structure of phage genomes, as it is believed even point mutation in the genome of the phage can eliminate the effect of its CRISPR-suppression mechanism [131,132]. Strains in different laboratories are used under one name, which derived from the same parental strain, can rapidly accumulate point mutations, with unpredictable effects. Therefore, the hopes of using such a labile system for practical purposes (e.g., in the case of cystic fibrosis) may be considered premature. There is evidence that *P. aeruginosa* PA14, being capable of preventing replication of six newly isolated temperate phages can also acquire a new spacer content as a result of some lytic phage infection. The authors consider it as evidence of the adaptive nature of this CRISPR/Cas system [133]. CRISPR-bearing loci of the strains of *P. aeruginosa* are quite common among clinical isolates [134]. For example, among 122 clinical isolates sampled in different geographical areas 45 strains were found carrying CRISPR with 132 viral spacers matched to temperate bacteriophages/prophages capable of inserting into the host chromosome, but not to extrachromosomally replicating lytic *P. aeruginosa* bacteriophages. It is unclear exactly what the authors define as lytic phage. Because none of the tested temperate phages showed deterioration in growth or lysogenization in the presence of CRISPR, described in this paper, the authors actually reject the dogma for CRISPR protective effect.

The examination of the spacers (inserts) in CRISPR loci, found in all strains of *P. aeruginosa* studied to date that this system is aware of, only several definite temperate phages. At least for most of the phages, bacterial pili are compulsory receptors of adsorption. These pili are very important for the establishment of relationships in the microbial community. Bacterial conjugation is a form of such relations, carried by IV type pili in which DNA and proteins are transferred to the appropriate recipient cells [135].

A common feature of these phages is their high mobility. An existing relationship between the bacteria and most temperate phages are relations in a community where all interactions maintain a common gene pool and adapt to changing environmental conditions through natural selection

(including humankind with its various activities as an integral part of the environment). There is an opinion that CRISPR activity is not designed to preserve the functional capabilities and integrity of the bacterial genome in a single bacterial cell. Moreover, it is protection from the expansion of destructive effects on all cells in a population, possibly in a state of biofilm. On the other hand, CRISPR can be viewed as a mechanism of active accumulation in the bacterial genomes of different phage genome fragments of different phages. It may create the possibility of re-infecting bacteriophages for further evolution through recombination. This may be of general interest, as there are many clinical isolates supporting the lytic cycle of temperate phages but not lysogenization. In the strict sense, it can also be seen as a way to protect the bacterial population from the invasion of new genetic elements into genomes.

Indeed, there is no described mechanism for protection of bacteria from virulent phages, similar to the mechanism of CRISPR. In this case, there may be situations where the death of one separate population of bacteria may be a more effective way to protect the entire population from the spread of genes that are incompatible with the survival of the bacteria. In this respect, it would be interesting to search among the known CRISPR strains for such options, which would respond to infection by certain temperate phages just as to infection by virulent lytic phages.

14. Future Therapeutic Phages: Different Directions

If phage therapy is destined to become an accepted and widely used medical procedure, its maintenance and development will depend entirely on the success of the constant expansion of the collection of therapeutic bacteriophages with lytic activities against newly arising phage resistance of bacterial populations. Apparently, there will be the need for permanent development and modification of genetically engineered phages with required properties. There are possible several different approaches to reach this purpose. For instance, such unified phages could be engineered on the basis of certain existing species with an attempt to create a unified assembly of phages having in their genomes individual blocks of genes which can improve the efficiency of infection and lysis of different bacterial species. It is known that in some cases of removing genes of phage, genome can improve the phage growth, as in the case of *P. aeruginosa* phage F116 [136]. There are other examples of loss of a substantial part of the genome *P. aeruginosa* phages without sacrificing growth efficiency.

Temperate phage SM is capable of growing in P. aeruginosa PAO1 cells with addition insertion in its intentionally deleted shortened genome up to 12 kb DNA insertion [137]. Even more striking is the difference in the amount of phage DNAs in phages phiKZ and EL (NCBI access number NC_007623). Capsids of these phages have the same size, but their genomes show a great difference (280,334 bp for phiKZ and 211,216 bp for EL). Thus, the redundant capacity of DNA to be packed in an EL capsid is not less than 69 kb (!).

One of the possible areas of work in the creation of modified phages would be affording different phage species with the ability to recognize the type IV pili of different bacterial species. In this regard,

there is a promising group of species showing relations with phage F116, whose genome is present in the cell as plasmids and that use pili of IV type for adsorption.

In phage therapeutic mixtures, different phages of the same species are often used, with different activities against various clinical isolates of *P. aeruginosa* [108]. It would be interesting to investigate the possibility of combining their specific features within a single phage genome. This will also help to start the work on unification of therapeutic phages.

Another possible trend for creating unified phages may be related to a deeper investigation of the origin of similar genes in genomes of phages active against bacteria of different species as *Salmonella enterica* and *Pseudomonas aeruginosa*, and possible mutual channels for migration. In relation to the hypothetical work on construction of artificial bacteriophages, looking for the elaboration of their industrial hosts will be required.

We consider that some principal changes related to the further development and application of therapy with live bacteriophages should be accomplished in the near future, as long as society has a high level of confidence in their use.

Acknowledgments

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Tejedor, C.; Foulds, J. Zasloff, M. Bacteriophages in sputum of patients with bronchopulmonary pseudomonas infections. *Infect. Immun*. **1982**, *36*, 440–441.
2. Lombardi, G.; Luzzaro, F.; Docquier, J.D.; Riccio, M.L.; Perilli, M.; Coli, A.; Amicosante, G.; Rossolini, G.M.; Toniolo, A. Nosocomial infections caused by multidrug-resistant isolates of *Pseudomonas putida* producing VIM-1 metallo-Гυ-lactamase. *J. Clin. Microbiol.* **2002**, *40*, 4051–4055.

3. Yomoda, S; Okubo, T; Takahashi, A; Murakami, M; Iyobe, S. Presence of *Pseudomonas putida* strains harboring plasmids bearing the Metallo-Гʋ-Lactamase gene blaIMP in a hospital in Japan. *J. Clin. Microbiol.* **2003**, *9*, 4246–4251.

4. Finnan, S.; Morrissey, JP.; O'Gara, F.; Boyd, E.F. Genome diversity of *Pseudomonas aeruginosa* isolates from cystic fibrosis patients and the hospital environment. *J. Clin. Microbiol*. **2004**, *42*, 5783–5792.

5. Ehrlich, N.; Jocz, J.; Kropp, L.; Wong, R.; Wadowsky, RM.; Slifkin, M.; Preston, R.A.; Erdos, G.; Post, J.C.; Ehrlich, G.D.; Hu, F.Z. Extensive genomic plasticity in *Pseudomonas aeruginosa* revealed by identification and distribution studies of novel genes among clinical isolates. *J. Infect. Immun.* **2006**, *74*, 5272–5283.

6. Barbier, F.; Wolff, M. Multi-drug resistant *Pseudomonas aeruginosa*: Towards a therapeutic dead end? *Med. Sci.* **2010**, *26*, 960–968.

7. Riou, M.; Carbonnelle, S.; Avrain, L.; Mesaros, N.; Pirnay, J.P.; Bilocq, F.; de Vos, D.; Simon, A; Pierard, D.; Jacobs, F.; *et al. In vivo* development of antimicrobial resistance in *Pseudomonas aeruginosa* strains isolated from the lower respiratory tract of Intensive Care Unit patients and receiving antipseudomonal therapy. *Int. J. Antimicrob. Agents* **2010**, *36*, 513–522.

8. Oliver, A.; Mena, A. Bacterial hypermutation in cystic fibrosis, not only for antibiotic resistance. *Clin. Microbiol. Infect.* **2010**, *16*, 798–808.

9. Madi, A.; Lakhdari, O.; Blottiere, H.M.; Guyard-Nicodeme M.; Le Roux, K.; Groboillot, A.; Svinareff, P.; Dore, J.; Orange, N.; Feuilloley, M.G.J.; *et al.* The clinical *Pseudomonas fluorescens* MFN1032 strain exerts a cytotoxic effect on epithelial intestinal cells and induces Interleukin-8 via the AP-1 signaling pathway. *BMC Microbiol.* **2010**, *10*, 215.

10. Qiu, X.; Gurkar, A.U.; Lory, S. Interstrain transfer of the large pathogenicity island (PAPI-1) of *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. U. S. A*. **2006**, *26*, 19830–19835.

11. Carter, M.Q.; Chen, J.; Lory, S. The *Pseudomonas aeruginosa* pathogenicity island PAPI-1 is transferred via a noveltype IV pilus. *J. Bacteriol.* **2010**, *192*, 3249–3258.

12. D'Herelle, F. *Le Bacteriophage*: *Son role dans l'immunite* (in French); Masson et cie: Paris, France, 1921. Russian translation, GIZ, 1926, p. 223.

13. Kudva, I.T.; Jelacic, S.; Tarr, P.I.; Youderian, P.; Hovde, C.J. Biocontrol of *Escherichia coli* O157 with O157-Specific Bacteriophages. *Appl. Environ. Microbiol.* **1999**, *9*, 3767–3773.

14. Buzby, J.C.; Roberts, T. The economics of enteric infections: human foodborne disease costs. *Gastroenterology* **2009**, *136*, 1851–1862.

15. Santos, S.B.; Fernandes, E.; Carvalho, C.M.; Sillankorva, S.; Krylov, V.N.; Pleteneva, E.A.; Shaburova, O.V.; Nicolau, A.; Ferreira, E.C.; Azeredo, J. Selection and characterization of a multivalent salmonella phage and its production in a non-pathogenic *E. coli* strain. *Appl. Environ. Microbiol*. **2010**, *21*, 7338–7342.

16. Patel, J.; Sharma, M.; Millner, P.; Calaway, T.; Singh, M. Inactivation of *Escherichia coli* O157:H7 attached to spinach harvester blade using bacteriophage. *Foodborne Pathog. Dis.* **2011**, *4*, 541–546.

17. Maura, D.; Debarbieux, L. Bacteriophages as twenty-first century antibacterial tools for food and medicine. *Appl. Microbiol. Biotechnol*. **2011**, *3*, 851–859.

18. Lakhno, V.M.; Bordunovski, V.N. Use of bacteriophage therapy in surgical practice. *Vestn. Khir. Im. II Grek*. **2001**, *160*, 122–125.

19. Lazareva, E.B.; Smirnov, S.V.; Khvatov, V.B.; Spiridonova, T.G.; Bitkova, E.E.; Darbeeva, O.S.; Mayskaia, L.M.; Parfeniuk, R.L.; Men'shikov, D.D. Efficacy of bacteriophages in compl treatment of patients with burn wounds. *Antibiot. Khimioter*. **2001**, *1*, 10–14.

20. Aslanov, B.I.; Yafaev, R.Ch.; Zueva, L.P. The ways of rational use of pyocianea bacteriophages in medical and anti-epidemic practice. *Zh. Mmikrobiol*. **2003**, *5*, 72–76.

21. Gilleland, H.E.; Gilleland, L.B.; Fowler, M.R. Vaccine efficacies of elastase, exotoxin A, and outer-membrane protein F in preventing chronic pulmonary infection by *Pseudomonas aeruginosa* in a rat model. *J. Med. Microbiol.* **1993**, *38*, 79–86.

22. Johansen, H.K.; Getzsche, P.C. Vaccines for preventing infection with *Pseudomonas aeruginosa* in cystic fibrosis. *Cochrane Database Syst. Rev.* **2008**, *8*, CD001399.

23. Kamei, A.; Coutinho-Sledge, Y.S.; Goldberg, J.B.; Priebe, G.P.; Pier, G.B. Mucosal vaccination with a multivalent, live-attenuated vaccine induces multifactorial immunity against *Pseudomonas aeruginosa* acute lung infection. *Infect. Immun*. **2011**, *79*, 1289–1299.

24. Chibani-Chennoufi, S.; Sidoti, J.; Bruttin, A.; Kutter, E.; Sarker, S.; Brussow, H. *In vitro* and *in vivo* bacteriolytic activities of *Escherichia coli* phages: Implications for phage therapy. *Antimicrob. Agents Chemother*. **2004**, *48*, 2558–2569.

25. Kropinski, A.M. Phage therapy — Everything old is new again. *Can. J. Infect. Dis Med. Microbiol*. **2006**, *17*, 297–306.

26. McVay, C.S.; Velasquez, M.; Fralick, J.A. Phage therapy of *Pseudomonas aeruginosa* infection in a mouse burn wound model. *Antimicrob. Agents Chemother*. **2007**, *6*, 1934–1938.

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

28. Ahiwale, S.; Tamboli, N.; Thorat, K.; Kulkarni, R.; Ackermann, H.; Kapadnis, B. *In vitro* management of hospital *Pseudomonas aeruginosa* biofilm using indigenous T7-Like lytic phage. *Curr. Microbiol*. **2011**, *62*, 335–340.

29. Garbe, J.; Wesche, A.; Bunk, B.; Kazmierczak, M.; Selezska, K.; Rohde, C.; Sikorski, J.; Rohde M.; Jahn, D.; Schobert, M. Characterization of JG024, a *Pseudomonas aeruginosa* PB1-like broad host range phage under simulated infection condition. *BMC Microbiol*. **2010**, *10*, 301.

30. Connerton, P.L.; Timms, A.R.; Connerton, I.F. Campylobacter bacteriophages and bacteriophage therapy. *J. Appl. Microbiol*. **2011**, *111*, 255–265.

31. Hens, D.K.; Chatterjee, N.C.; Kumar, R. New temperate DNA phage BcP15 acts as a drug resistance vector. *Arch. Virol*. **2006**, *151*, 1345–1353.

32. Stone, R. Bacteriophage therapy: Stalin's forgotten cure. *Science* **2002**, *298*, 728–731.

33. Weber-Dabrowska, B.; Dabrowski, M.; Slopek, S. Studies on bacteriophage penetration in patients subjected to phage therapy. *Arch. Immunol. Ther. Exp. (Warsz)* **1987**, *5*, 563–568.

34. Morello, E.; Saussereau, E.; Maura, D.; Huerre, M.; Touqui, L.; Debarbieux, L. Pulmonary bacteriophage therapy on *Pseudomonas aeruginosa* cystic fibrosis strains: first steps towards treatment and prevention. *PLoS One* **2011**, *6*, e16963.

35. Pleteneva, E.A.; Shaburova, O.V.; Krylov, V.N. A formal scheme of adsorpbtional receptors in *Pseudomonas aeruginosa* and possibilities for its practical implementation. *Russ. J. Genet.* **2009**, *1*, 35–40.

36. Intralytix, Phages for Food Safety. Available online: http://www.intralytix.com/ (accessed on 19 January 2011).

37. Krylov, V.N. Use of live phages for therapy on a background of co-evolution of bacteria and phages. *Int. Res. J. Microbiol.* **2011**, *2*, 315–332.

38. Rajagopala, S.V.; Casjens, S.; Uetz, P. The protein interaction map of bacteriophage lambda. *BMC Microbiol.* **2011**, *26*, 213.

39. Ripp, S.; Miller, R.V. Dynamics of the pseudolysogenic response in slowly growing cells of *Pseudomonas aeruginosa*. *Microbiology* **1998**, *8*, 2225–2232.

40. Lood, R.; Collin, M. Characterization and genome sequencing of two Propionibacterium acnes phages displaying pseudolysogeny. *BMC Genomics* **2011**, *12*, 198.

41. Khemayan, K.; Pasharawipas, T.; Puiprom, O.; Sriurairatana, S.; Suthienkul, O.; Flegel, T.W. Unstable lysogeny and pseudolysogeny in Vibrio harveyi siphovirus-like phage 1. *Appl. Environ. Microbiol.* **2006**, *2*, 1355–1363.

42. Sakaguchi, Y.; Hayashi, T.; Kurokawa, K.; Nakayama, K.; Oshima, K.; Fujinaga, Y.; Ohnishi, M.; Ohtsubo, E.; Hattori, M.; Oguma, K. The genome sequence of *Clostridium botulinum* type C neurotoxin-converting phageand the molecular mechanisms of unstable lysogeny. *Proc. Natl. Acad. Sci. USA* **2005**, *29*, 17472–17477.

43. Holloway, B.W.; Egan, J.B.; Monk, M. Lysogeny in *Pseudomonas aeruginosa*. *Aust. J. Exp. Biol. Med. Sci.* **1960**, *38*, 321–329.

44. Nzula, S.; Vandamme, P.; Govan, J.R.; Sensitivity of the *Burkholderia cepacia* complex and *Pseudomonas aeruginosa* to transducing bacteriophages. *FEMS Immunol. Med. Microbiol.* **2000**, *28*, 307–312.

45. Krylov, V.N.; Zhazykov, I.Zh. Pseudomonas bacteriophage phiKZ — A possible model for studying the genetic control of morphogenesis. *Genetika* **1978**, *14*, 678–685.

46. Krylov, V.N.; Smirnova, T.A.; Minenkova, I.B.; Plotnikova, T.G.; Zhazikov, I.Z.; Khrenova, E.A. Pseudomonas bacteriophage phiKZ contains an inner body in its capsid. *Can. J. Microbiol.* **1984**, *30*, 758–762.

47. Bourkal'tseva, M.V.; Krylov, V.N.; Pleteneva, E.A.; Shaburova, O.V.; Krylov, S.V.; Volkart, G.; Sykilinda, N.N.; Kurochkina, L.P.; Mesianzhinov, V.V. Phenogenetic characterization of a group of giant Phi KZ-like bacteriophages of *Pseudomonas aeruginosa*. *Genetika* **2002**, *38*, 1470–1479.

48. Krylov, V.N.; Dela Cruz, D.M.; Hertveldt, K.; Ackermann, H.W. "phiKZ-like viruses", a proposed new genus of myovirus bacteriophages. *Arch. Virol.* **2007**, *152*, 1955–1959.

49. Krylov, V.N., Bourkal'tseva, M.V.; Sykilinda, N.N.; Pleteneva, E.A.; Shaburova, O.V.; Kadykov, V.A.; Miller, S.; Biebl, M. Comparison of genomes of new gigantic *Pseudomonas aeruginosa* phages from native populations from different regions *Russ. J. Genet.* **2004**, *40*, 363–368.

50. Mesyanzhinov, V.V.; Robben, J.; Grymonprez, B.; Kostyuchenko, V.A.; Bourkal'tseva, M.V.; Sykilinda, N.N.; Krylov, V.N.; Volckaert, G. The Genome of Bacteriophage phiKZ of *Pseudomonas aeruginosa*. *J. Mol. Biol*. **2002**, *317*, 1–19.

51. Cornelissen, A.; Hardies, S.C.; Shaburova, O.V.; Krylov, V.N.; Mattheus, W.; Kropinski, A.M.; Lavigne, R. Complete genome sequence of the giant virus OBP and comparative genome analysis of the diverse phiKZ-related phages. *J. Virol*. **2012**, *86*, 1844–1852.

52. Kazlauskas, D.; Venclovas, C. Computational analysis of DNA replicases in double-stranded DNA viruses: relationship with the genome size. *Nucleic Acids Res.* **2011**, *39*, 8291–305.

53. Hertveldt, K.; Lavigne, R.; Pleteneva, E.; Sernova, N.; Kurochkina, L.; Korchevskii, R.; Robben, J.; Mesyanzhinov, V.; Krylov, V.N.; Volckaert, G. Genome comparison of *Pseudomonas aeruginosa* large phages. *J. Mol. Biol.* **2005**, *2*, 536–545.

54. Pleteneva, E.A.; Krylov, S.V.; Shaburova, O.V.; Bourkal'tseva, M.V.; Miroshnikov, K.A.; Krylov, V.N. Pseudolysogeny of *Pseudomonas aeruginosa* bacteria infected with phiKZ-like bacteriophages. *Russ. J. Genet.* **2010**, *46*, 20–25.

55. Krylov, S.V.; Pleteneva, E.A.; Bourkal'tseva, M.V.; Shaburova, O.V.; Miroshnikov, K.A.; Lavigne, R.; Cornelissen, A.; Krylov, V.N. Genome instability of *Pseudomonas aeruginosa* phages of the EL species: examination of virulent mutants. *Russ. J. Genet.* **2011**, *2*, 162–167.

56. Krylov, V.N.; Miroshnikov, K.A.; Krylov, S.V.; Veyko, V.P.; Pleteneva, E.A.; Shaburova, O.V.; Bourkal'tseva, M.V. Interspecies migration and evolution of bacteriophagesof the genus phiKZ: The purpose and criteria of the search for new phiKZ-like bacteriophages. *Russ. J. Genet.* **2010**, *46*, 138–145.

57. Ramsey, D.M.; Baynham, P.J.; Wozniak, D.J. Binding of *Pseudomonas aeruginosa* AlgZ to sites upstream of the algZ promoter leads to repression of transcription. *J. Bacteriol.* **2005**, *187*, 4430–4443.

58. Mondello F.J.; Miller, R.V. Identification of *Pseudomonas* plasmids able to suppress the lysogeny-establishment-deficiency (Les-) phenotype. *Plasmid* **1984**, *11*, 185–187.

59. Fornelos, N.; Bamford, J.K.; Mahillon, J. Phage-borne factors and host LexA regulate the lytic switch in phage GIL01. *J. Bacteriol.* **2011**, *193*, 6008–6019.

60. Krylov, V.N.; Tolmachova, T.O.; Akhverdian, V.Z. DNA homology in species of bacteriophages active on *Pseudomonas aeruginosa*. *Arch. Virol.* **1993**, *1–2*, 141–151.

61. Kropinski, A.M. Sequence of the genome of the temperate, serotype-converting, *Pseudomonas aeruginosa* bacteriophage D3. *J. Bacteriol.* **2000**, *182*, 6066–6074.

62. Byrne, M.; Kropinski, A.M. The genome of the *Pseudomonas aeruginosa* generalized transducing bacteriophage F116. *Gene* **2005**, *346*, 187–194.

354

63. Casjens, S.R.; Gilcrease, E.B.; Winn-Stapley, D.A.; Schicklmaier, P.; Schmieger, H.; Pedulla M.L.; Ford, M.E.; Houtz, J.M.; Hatfull, G.F.; Hendrix, R.W. The generalized transducing salmonella bacteriophage ES18: Complete genome sequence and DNA packaging strategy. *J. Bacteriol.* **2005**, *187*, 1091–1104.

64. Nakayama, K.; Kanaya, S.; Ohnishi, M.; Terawaki, Y.; Hayashi, T. The complete nucleotide sequence of, a cytotoxin-converting phage of *Pseudomonas aeruginosa*: Implications for phage evolution and horizontal gene transfer via bacteriophages. *Mol. Microbiol.* **1999**, *31*, 399–419.

65. Ceyssens, P.J.; Brabban, A.; Rogge, L.; Lewis, M.S.; Pickard, D.; Goulding, D.; Dougan, G.; Noben, J.P.; Kropinski, A.; Kutter, E.; Lavigne, R. Molecular and physiological analysis of three *Pseudomonas aeruginosa* phages belonging to the "N4-like viruses". *Virology* **2010**, *15*, 26–30.

66. Krylov, S.V.; Kropinski, A.; Shaburova, O.V.; Chesnokova, E.N.; Miroshnikov, K.A.; Krylov, V.N. Features of the genome structure of phi297—Temperate phage of new species active on *Pseudomonas aeruginosa. Russ. J. Genet.* **2013**, in press.

67. Holloway, B.W.; Cooper, G.N. Lysogenic conversion in *Pseudomonas aeruginosa. J. Bacteriol.* **1962**, *84*, 1321–1324.

68. Kuzio, J.; Kropinski, A.M. O-antigen conversion in *Pseudomonas aeruginosa* PAO1 by bacteriophage D3. *J. Bacteriol.* **1983**, *155*, 203–212.

69. Miller, R.V.; Pemberton, J.M.; Clark, A.J. Prophage F116: Evidence for extrachromosomal location in *Pseudomonas aeruginosa* strain PAO. *J. Virol.* **1977**, *22*, 844–847.

70. Rodionov, O.; Yarmolinsky, M. Plasmid partitioning and the spreading of P1 partition protein ParB. *Mol. Microbiol.* **2004**, *52*, 1215–23.

71. Lee, J.H.; Shin, H.; Kim, H.; Ryu, S. Complete genome sequence of salmonella bacteriophage SPN3US. *J. Virol.* **2011**, *85*, 13470–13471.

72. Jacoby, G.A.; Jacob, A.E.; Hedges, R.W. Recombination between plasmids of incompatibility groups P-1 and P-2. *J. Bacteriol.* **1976**, *127*, 1278–1285.

73. Freizon, E.V.; Kopylova, Iu.I.; Cheremukhina, L.V.; Krylov, V.N. The effect of the IncP-2-group plasmid on the growth of *Pseudomonas aeruginosa* bacteriophages. *Genetika* **1989**, *25*, 1168–1178.

74. Jacoby, G.A.; Sutton, L. Properties of plasmids responsible for production of extended-spectrum beta -lactamases. *Antimicrob. Agents Chemother.* **1991**, *35*, 164–169.

75. Philippon, A.M.; Paul, G.C.; Thabaut, A.P.; Jacoby, G.A. Properties of a novel carbenicillin-hydrolyzing beta-lactamase (CARB-4) specified by an IncP-2 plasmid from *Pseudomonas aeruginosa. Antimicrob. Agents Chemother.* **1986**, *29*, 519–520.

76. Krylov, V.N.; Eremenko, E.N.; Bogush, V.G.; Kirsanov, N.B. Interaction of *Pseudomonas aeruginosa* plasmids and mu-like phages: the suppression of the early stages of cell infection by phage D3112 in the presence of plasmid RPL11. *Genetika* **1982**, *18*, 743–752.

77. Gerasimov, V.A.; Eremenko, E.N.; Khrenova, E.A.; Gorbunova, S.A.; Krylov, V.N. Stimulation of the lysogenization of *Pseudomonas aeruginosa* cells by phage transposon B39 induced by plasmid Rms163. *Genetika* **1984**, *20*, 1080–1087.

78. Krylov, V.N.; Plotnikova, T.G.; Kulakov, L.A.; Fedorova, T.V.; Eremenko, E.N. Integration of the genome of the Mu-like *Pseudomonas aeruginosa* bacteriophage D3112 into plasmid RP4 and its hybrid plasmid transfer into *Pseudomonas putida* and *Escherichia coli* C600 bacteria. *Genetika* **1982**, *18*, 5–12.

79. Plotnikova, T.G.; Akhverdian, V.Z.; Reulets, M.A.; Gorbunova, S.A.; Krylov, V.N. Multiplicity of the integration sites of Mu-like bacteriophages into *Pseudomonas aeruginosa* chromosome and plasmids. *Genetika* **1983**, *19*, 1604–1610.

80. Gorbunova, S.A.; Ianenko, A.S.; Akhverdian, V.Z.; Reulets, M.A.; Krylov, V.N. Expression of *Pseudomonas aeruginosa* transposable phages in *Pseudomonas putida* cells. I. The establishment of a lysogenic state and the effectiveness of lytic development. *Genetika* **1985**, *9*, 1455–1463.

81. Krylov, V.; Merlin, C.; Toussaint, A. Introduction of *Pseudomonas aeruginosa* mutator phage D3112 into *Alcaligenes eutrophus* strain CH34. *Res. Microbiol.* **1995**, *146*, 245–250.

82. Kaplan, A.M.; Akhverdian, V.Z.; Reulets, M.A.; Krylov, V.N. Compatibility of transposable phages of *Escherichia coli* and *Pseudomonas aeruginosa*. I. Co-development of phages Mu and D3112 and integration of phage D3112 into RP4::Mu plasmid in *Pseudomonas aeruginosa* cells. *Genetika* **1988**, *4*, 634–640.

83. Kaplan, A.M.; Akhverdian, V.Z.; Krylov, V.N. Integration of phage Mu into the RP4::D3112A-plasmid in *Escherichia coli* cells. *Genetika* **1989**, *25*, 1384–1390.

84. Kopylova, I.I.; Gorbunova, S.A.; Krylov, V.N. Interaction of heterologous bacteriophages: Growth suppression of temperate PP56 phage of *Pseudomonas putida* by the transposable phage D3112 of *Pseudomonas aeruginosa*. *Genetika* **1988**, *5*, 803–807.

85. Plotnikova, T.G.; Kulakov, L.A.; Eremenko, E.N.; Fedorova, T.V.; Krylov, V.N. Expression of the genome of Mu-like phage D3112 specific for *Pseudomonas aeruginosa* in *Escherichia coli* and *Pseudomonas putida* cells. *Genetika* **1982**, *7*, 1075–1084.

86. Plotnikova, T.G.; Ianenko, A.S.; Kirsanov, N.B.; Krylov, V.N. Transposition of the phage D3112 genome in *Escherichia coli* cells. *Genetika* **1983**, *10*, 1611–1615.

87. Trenina, M.A.; Akhverdian, V.Z.; Kolibaba, L.G.; Rebentish, B.A.; Krylov, V.N. Features of genome expression of phage transposon D3112 of *Pseudomonas aeruginosa* in *Escherichia coli* bacteria: Dependence of bacterial phenotype on copy number of D3112 genome. *Genetika* **1991**, *27*, 1324–1335.

88. Krylov, S.V.; Kropinski, A.M.; Pleteneva, E.A.; Shaburova, O.V.; Bourklal'tseva, M.V.; Miroshnikov, K.A.; Krylov, V.N. Features of new D3-like phage phiPMG1 of *Pseudomonas aeruginosa*: Genome structure and prospects for use the phage in phage therapy. *Russ. J. Genet.* **2012**, *48*, 902–911.

89. Dzhusupova, A.B.; Plotnikova, T.G.; Krylov, V.N. Detection of transduction by virulent bacteriophage phiKZ of *Pseudomonas aeruginosa* chromosomal markers in the presence of plasmid RMS148. *Genetika* **1982**, *18*, 1799–1802.

90. Langley, R.; Kenna, D.T.; Vandamme, P.; Ure, R.; Govan, J.R. Lysogeny and bacteriophage host range within the *Burkholderia cepacia* complex. *J. Med. Microbiol.* **2003**, *52*, 483–490.

356

91. Krylov, V.N.; Bogush, V.G.; Shapiro, J. Bacteriophages of *Pseudomonas aeruginosa* with DNA similar in structure to that of phage Mu1. I. General description, localization of sites sensitive to endonucleases in DNA, and structure of homoduplexes of phage D3112. *Genetika* **1980**, *16*, 824–832.

92. Akhverdian, V.Z.; Khrenova, E.A.; Reulets, M.A.; Gerasimova, T.V.; Krylov, V.N. The properties of transposable phages of *Pseudomonas aeruginosa* belonging to two groups distinguished by DNA-DNA homology. *Genetika* **1985**, *21*, 735–747.

93. Akhverdian, V.Z.; Khrenova, E.A.; Lobanov, A.O.; Krylov, V.N. Role of divergence in evolution of group B3 *Pseudomonas aeruginosa* transposable phage evolution. *Russ. J. Genet.* **1998**, *34*, 699–701.

94. Akhverdian, V.Z.; Lobanov, A.O.; Khrenova, E.A.; Krylov, V.N. Recombinational origin of natural transposable phages of related species. *Russ. J. Genet.* **1998**, *34*, 697–700.

95. Bogush, V.G.; Plotnikova, T.G.; Kirsanov, N.B.; Rebentish, B.A.; Permogorov, V.I.; Krylov V.N. Bacteriophages of *Pseudomonas aeruginosa* with DNA similar in structure to that of phage Mu1. III. Isolation and analysis of hybrid phages D3112 and B39: Localization of the immunity region and some genetic factors. *Genetika* **1981**, *17*, 967–976.

96. Krylov, V.N.; Akhverdian, V.Z.; Khrenova, E.A.; Cheremukhina, L.V.; Tiaglov, B.V. Two types of molecular structure (composition) of a genome in one species of transposable bacteriophages of *Pseudomonas aeruginosa*. *Genetika* **1986**, *22*, 2637–2648.

97. Krylov, V.N.; Bogush, V.G.; Ianenko, A.S.; Kirsanov, N.B. Bacteriophages of *Pseudomonas aeruginosa* with DNA similar in structure to that of phage Mu1. II. Evidence of relatedness of bacteriophages D3112, B3 and B39: analysis of DNA digestion with restrictional endonucleases, isolation of recombinant of phages D3112 and B3. *Genetika* **1980**, *16*, 975–984.

98. Zemlyanaya, N.Y.; Kozma, A.R.; Krylov, V.N. Differences in mutator acrivity of phage-transposons of *Pseudomonas aeruginosa*. *Genetika* **1992**, *28*, 160–163.

99. Rehmat, S.; Shapiro, J.A. Insertion and replication of the *Pseudomonas aeruginosa* mutator phage D3112. *Mol. Gen. Genet.* **1983**, *192*, 416–423.

100. Krylov, V.N.; Solov'eva, T.I.; Bourkal'tseva, M.V. Mucoid clones of *Pseudomonas aeruginosa* PAO1, surviving after induction of prophage transposons. *Russ. J. Genet.* **1995**, *31*, 1170–1174.

101. Gerasimov, V.A.; Ianenko, A.S.; Akhverdian, V.Z.; Krylov, V.N. Interaction between *Pseudomonas aeruginosa* phages-transposons: Genetic analysis of the trait of inhibition by prophage D3112 of phage B39 development. *Russ. J. Genet.* **1996**, *32*, 1068–1073.

102. Gerasimov, V.A.; Ianenko, A.S.; Akhverdian, V.Z.; Krylov, V.N. Phage-transposon interaction: The cip locus of prophage D3112 responsible for the inhibition of integration and transposition of related phage B39 of *Pseudomonas aeruginosa*. *Genetika* **1985**, *21*, 1634–1642.

103. Krylov, V.N.; Akhverdian, V.Z.; Bogush, V.G.; Khrenova, E.A.; Reulets, M.A. Modular structure of the genes of phages-transposons of *Pseudomonas aeruginosa*. *Genetika* **1985**, *21*, 724–734.

104. Ianenko, A.S.; Bekkarevich, A.O.; Gerasimov, V.A.; Krylov, V.N. Genetic map of the transposable phage D3112 of *Pseudomonas aeruginosa*. *Genetika* **1988**, *24*, 2120–2126.

105. Summer, E.J.; Gonzalez, C.F.; Carlisle, T.; Mebane, L.M.; Cass, A.M.; Savva, C.G.; LiPuma, J.; Young, R. *Burkholderia cenocepacia* phage BcepMu and a family of Mu-like phages encoding potential pathogenesis factors. *J. Mol. Biol.* **2004**, *340*, 49–65.

106. Wang, P.W.; Chu, L.; Guttman, D.S. Complete sequence and evolutionary genomic analysis of the *Pseudomonas aeruginosa* transposable bacteriophage D3112. *J. Bacteriol.* **2004**, *186*, 400–410.

107. Braid, M.D.; Silhavy, J.L.; Kitts, C.L.; Cano, R.J.; Howe, M.M. Complete genomic sequence of bacteriophage B3, a Mu-like phage of *Pseudomonas aeruginosa*. *J. Bacteriol.* **2004**, *186*, 6560–6574.

108. Bourkaltseva, M.V.; Krylov, S.V.; Kropinski, A.M.; Pleteneva, E.A.; Shaburova, O.V.; Krylov, V.N. Bacteriophage phi297, a New Species of *Pseudomonas aeruginosa* Temperate phages with a mosaic genome: potential use in phage therapy. *Russ. J. Genet.* **2011**, *47*, 794–798.

109. Piekarowicz, A.; Kłyz, A.; Majchrzak, M.; Adamczyk-Popławska, M.; Maugel, T.K.; Stein, D.C. Characterization of the dsDNA prophage sequences in the genome of *Neisseria gonorrhoeae* and visualization of productive bacteriophage. *BMC Microbiol.* **2007**, *7*, 66.

110. Lecoutere, E.; Ceyssens, P.J.; Miroshnikov, K.A.; Mesyanzhinov, V.V.; Krylov, V.N.; Noben, J.P.; Robben, J.; Hertveldt, K.; Volckaert, G.; Lavigne, R. Identification and comparative analysis of the structural proteomes of phiKZ and EL, two giant *Pseudomonas aeruginosa* bacteriophages. *Proteomics* **2009**, *9*, 3215–3219.

111. Serwer, P.; Hayes, S.J.; Thomas, J.A.; Demeler, B.; Hardies, S.C. Isolation of novel large and aggregating bacteriophages. *Methods Mol. Biol.* **2009**, *501*, 55–66.

112. Adriaenssens, E.M.; Mattheus, W.; Cornelissen, A.; Shaburova, O.; Krylov, V.N.; Kropinski, A.M.; Lavigne, R. Complete genome sequence of the giant *Pseudomonas* phage Lu11. *J. Virol.* **2012**, *86*, 6369–6370.

113. Shaburova, O.V.; Hertveldt, K.; de la Cruz, D.M.A.; Krylov, S.V.; Pleteneva, E.A.; Bourkaltseva, M.V.; Lavigne, R.; Volckaert, G.; Krylov, V.N. Comparison of new giant bacteriophages OBP and Lu11 of soil pseudomonads with bacteriophages of the phiKZ-supergroup of *Pseudomonas aeruginosa. Russ. J. Genet.* **2006**, *42*, 877–885.

114. Yamada, T.; Satoh, S.; Ishikawa, H.; Fujiwara, A.; Kawasaki, T.; Fujie, M.; Ogata, H. A jumbo phage infecting the phytopathogen *Ralstonia solanacearum* defines a new lineage of the *Myoviridae* family. *Virology* **2010**, *398*, 135–147.

115. Ceyssens, P.J.; Hertveldt, K.; Ackermann, H.W.; Noben, J.P.; Demeke, M.; Volckaert, G.; Lavigne, R. The intron-containing genome of the lytic *Pseudomonas* phage LUZ24 resembles the temperate phage PaP3. *Virology* **2008**, *377*, 233–238.

116. Kulakov, L. The School of Biological Sciences, The Queen's University of Belfast, Northern Ireland, UK. The genome sequence of phage TL. Personal communication, 2012.

117. Pleteneva, E.A.; Shaburova, O.V.; Sykilinda, N.N.; Miroshnikov, K.A.; Krylov, S.V.; Mesianzhinov, V.V.; Krylov, V.N. Study of the diversity in a group of phages of *Pseudomonas aeruginosa* species PB1 (*Myoviridae*) and their behavior in adsorbtion-resistant bacterial mutants. *Genetika* **2008**, *44*, 185–194.

118. Ceyssens, P.J.; Miroshnikov, K.; Mattheus, W.; Krylov, V.; Robben, J.; Noben, J.P.; Vanderschraeghe, S.; Sykilinda, N.; Kropinski, A.M.; Volckaert, G.; *et al*. Comparative analysis of the widespread and conserved PB1-like viruses infecting *Pseudomonas aeruginosa. Environ. Microbiol*. **2009**, *11*, 2874–2883.

119. Pleteneva, E.A.; Bourkal'tseva, M.V.; Shaburova, O.V.; Krylov, S.V.; Pechnikova, E.V.; Sokolova, O.S.; Krylov, V.N. TL, the new bacteriophage of *Pseudomonas aeruginosa* and its application for the search of halo-producing bacteriophages. *Russ. J. Genet*. **2011**, *47*, 5–9.

120. Chattoraj, S.S.; Murthy, R.; Ganesan, S.; Goldberg, J.B.; Zhao, Y.; Hershenson, M.B.; Sajjan U.S. *Pseudomonas aeruginosa* alginate promotes *Burkholderia cenocepacia* persistence in cystic fibrosis transmembrane conductance regulator knockout mice. *Infect. Immun*. **2010**, *78*, 984–993.

121. Chapalain, A.; Rossignol, G.; Lesouhaitier, O.; Merieau, A.; Gruffaz, C.; Guerillon, J.; Meyer, J.M.; Orange, N.; Feuilloley, M.G. Comparative study of 7 fluorescent pseudomonad clinical isolates. *Can. J. Microbiol*. **2008**, *54*, 19–27.

122. Chapalain, A.; Rossignol, G.; Lesouhaitier, O.; Le Roux, K.;Groboillot, A.; Svinareff, P.; Dore, J.; Orange, N.; Feuilloley, M.G.J.; Connil, N. The clinical *Pseudomonas fluorescens* MFN1032 strain exerts a cytotoxic effect on epithelial intestinal cells and induces Interleukin-8 via the AP-1 signaling pathway. *BMC Microbiol*. **2010**, *10*, 215.

123. Groisman, E.A.; Casadaban, M.J. Cloning of genes from members of the family *Enterobacteriaceae* with mini-Mu bacteriophage containing plasmid replicons. *J. Bacteriol*. **1987**, *169*, 687–693.

124. Ishiwa, A.; Komano, T. Thin pilus PilV adhesins of plasmid R64 recognize specific structures of the lipopolysaccharide molecules of recipient cells. *J. Bacteriol*. **2003**, *185*, 5192–5199.

125. Rabel, C.; Grahn, A.M.; Lurz, R.; Lanka, E. The VirB4 family of proposed traffic nucleoside triphosphatases: Common motifs in plasmid RP4 TrbE are essential for conjugation and phage adsorption. *J. Bacteriol*. **2003**, *185*, 1045–1058.

126. Stewart, R.M.; Wiehlmann, L.; Ashelford, K.E.; Preston, S.J.; Frimmersdorf, E.; Campbell, B.J.; Neal, T.J.; Hall, N.; Tuft, S.; Kaye, S.B.; Winstanley, C. Genetic characterization indicates that a specific subpopulation of *Pseudomonas aeruginosa* is associated with keratitis infections. *J. Clin. Microbiol*. **2011**, *49*, 993–1003.

127. Masignani, V.; Giuliani, M.M.; Tettelin, H.; Comanducci, M.; Rappuoli, R.; Scarlato, V. Mu-like Prophage in serogroup B *Neisseria meningitidis* coding for surface-exposed antigens. *Infect. Immun*. **2001**, *4*, 2580–2588.

128. Winther-Larsen, H.C.; Wolfgang, M.C.; van Putten, J.P.; Roos, N.; Aas, F.E.; Egge-Jacobsen, W.M.; Maier, B.; Koomey, M. *Pseudomonas aeruginosa* Type IV pilus expression in *Neisseria gonorrhoeae*: Effects of pilin subunit composition on function and organelle dynamics. *J. Bacteriol*. **2007**, *189*, 6676–6685.

129. Zegans, M.E.; Wagner, J.C.; Cady, K.C.; Murphy, D.M.; Hammond, J.H.; O'Toole, G.A. Interaction between bacteriophage DMS3 and host CRISPR region inhibits group behaviors of *Pseudomonas aeruginosa*. *J. Bacteriol.* **2009**, *191*, 210–219.

130. Chung, I.Y.; Cho, Y.H. Complete genome sequences of two *Pseudomonas aeruginosa* temperate phages MP29 and MP42, which lack the phage-host CRISPR interaction. *J. Virol.* **2012**, *86*, 8336.

131. Barrangou, R.; Horvath, P. CRISPR: New horizons in phage resistance and strain identification. *Annu. Rev. Food Sci. Technol.* **2012**, *3*, 143–162.

132. Cady, K.C.; O'Toole, G.A. Non-identity-mediated CRISPR-bacteriophage interaction mediated via the Csy and Cas3 proteins. *J. Bacteriol.* **2011**, *193*, 3433–3445.

133. Cady, K.C.; Bondy-Denomy, J.; Heussler, G.E.; Davidson, A.R.; O'Toole, G.A. The CRISPR/Cas adaptive immune system of *Pseudomonas aeruginosa* mediates resistance to naturally occurring and engineered phages. *J. Bacteriol.* **2012**, doi:10.1128/JB.01184-12.

134. Cady, K.C.; White, A.S.; Hammond, J.H.; Abendroth, M.D.; Karthikeyan, R.S.; Lalitha, P.; Zegans, M.E.; O'Toole, G.A. Prevalence, conservation and functional analysis of *Yersinia* and *Escherichia* CRISPR regions in clinical *Pseudomonas aeruginosa* isolates. *Microbiology* **2011**, *157*, 430–437.

135. Lang, S.; Kirchberger, P.C.; Gruber, C.J.; Redzej, A.; Raffl, S.; Zellnig, G.; Zangger, K.; Zechner, E.L. An activation domain of plasmid R1 TraI protein delineates stages of gene transfer initiation. *Mol. Microbiol.* **2011**, *82*, 1071–1085.

136. Caruso, M.; Shapiro, J.A. Interactions of Tn7 and temperate phage F116L of *Pseudomonas aeruginosa*. *Mol. Gen. Genet.* **1982**, *188*, 292–298.

137. Kul'ba, A.M.; Gorelyshev, A.S.; Evtushenkov, A.N.; Fomichev, Iu.K. Temperate SM phage from *Pseudomonas aeruginosa* as a vector for cloning genetic information. *Mol. Gen. Mikrobiol. Virusol.* **1991**, *12*, 26–29.

MDPI