

Review

Multisubunit RNA Polymerases of Jumbo Bacteriophages

Maria L. Sokolova ^{1,*} , Inna Misovets ¹ and Konstantin V. Severinov ²

¹ Center of Life Sciences, Skolkovo Institute of Science and Technology, 121205 Moscow, Russia; Inna.Misovets@skoltech.ru

² Waksman Institute of Microbiology, Rutgers, The State University of New Jersey, Piscataway, NJ 08854, USA; severik@waksman.rutgers.edu

* Correspondence: maria.sokolova@skolkovotech.ru

Received: 30 August 2020; Accepted: 21 September 2020; Published: 23 September 2020



Abstract: Prokaryotic viruses with DNA genome longer than 200 kb are collectively referred to as “jumbo phages”. Some representatives of this phylogenetically diverse group encode two DNA-dependent RNA polymerases (RNAPs)—a virion RNAP and a non-virion RNAP. In contrast to most other phage-encoded RNAPs, the jumbo phage RNAPs are multisubunit enzymes related to RNAPs of cellular organisms. Unlike all previously characterized multisubunit enzymes, jumbo phage RNAPs lack the universally conserved alpha subunits required for enzyme assembly. The mechanism of promoter recognition is also different from those used by cellular enzymes. For example, the AR9 phage non-virion RNAP requires uracils in its promoter and is able to initiate promoter-specific transcription from single-stranded DNA. Jumbo phages encoding multisubunit RNAPs likely have a common ancestor allowing making them a separate subgroup within the very diverse group of jumbo phages. In this review, we describe transcriptional strategies used by RNAP-encoding jumbo phages and describe the properties of characterized jumbo phage RNAPs.

Keywords: jumbo phages; RNA polymerase; multisubunit RNA polymerase; non-canonical RNA polymerases; transcription; promoter; RNA sequencing; pseudo-nucleus; uracil

1. Multisubunit RNA Polymerases of Cellular Organisms

Transcription, the synthesis of RNA from DNA template, is the first step of gene expression. In all cellular organisms from bacteria to humans, transcription of genomic DNA is catalyzed by evolutionarily-related multisubunit RNA polymerases (RNAPs) [1]. In bacteria and archaea, a single enzyme is employed for transcription of all genes, while in eukarya, there are at least three specialized RNAPs dedicated to transcription of different subsets of nuclear genes [1,2]. Archaeal RNAP is remarkably similar to eukaryal RNAP II, which synthesizes messenger RNAs [1,3,4].

Structural and functional studies revealed that the core of multisubunit RNAPs is highly conserved in all three domains of life (Figure 1a) [1,5]. The simplest bacterial RNAP core is composed of two large catalytic subunits β' and β , a dimer of α subunits (an assembly platform for large subunits), and the ω subunit (a chaperon for β') (Figure 1a) [6–8]. The active site is located at the interface of two double-psi β -barrel (DPBB) domains of β' and β (Figure 1b). The DPBB domain of the β' subunit carries the universally conserved amino acid motif DxDGD (where x represents a bulky residue) whose three aspartates coordinate Mg^{2+} ions required for catalysis [9,10]. The archaeal and eukaryal RNAPs core enzymes contain homologs of all bacterial RNAP core subunits as well as multiple additional subunits (Figure 1a) [1,5].

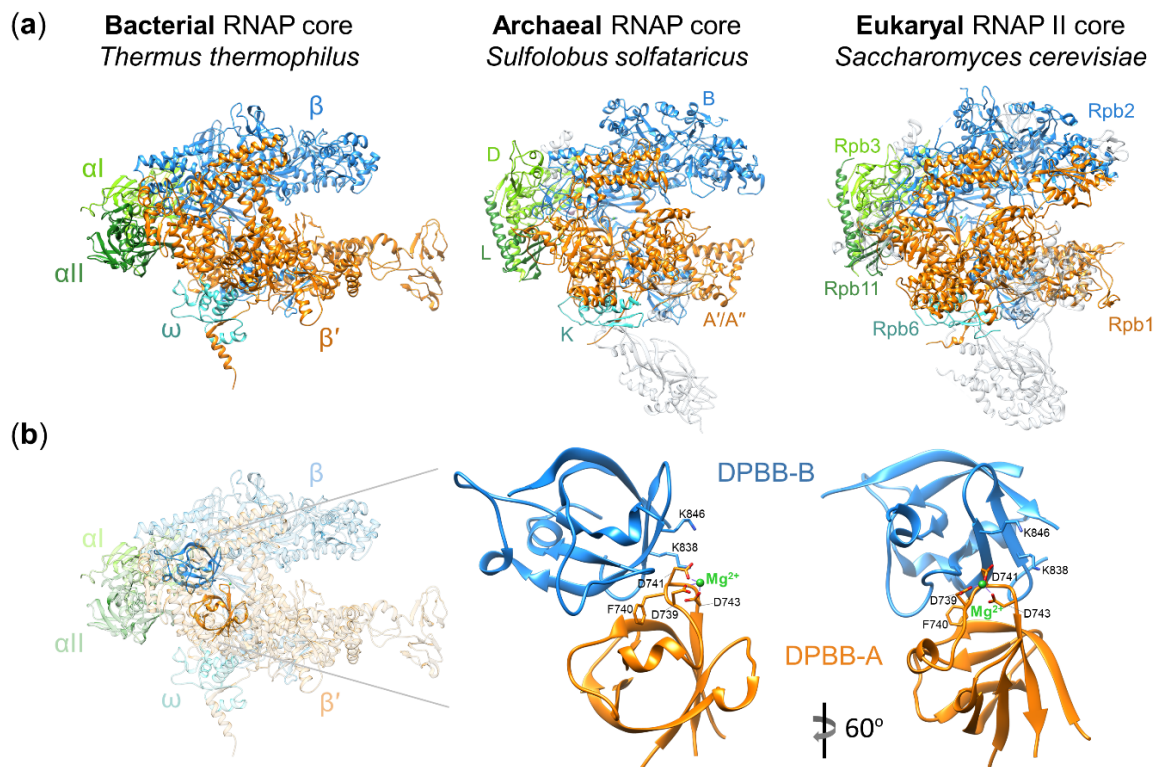


Figure 1. Architecture of multisubunit RNA polymerases (RNAPs). (a) Crystal structures of the bacterial RNAP core (PDB: 2O5J) [11], archaeal RNAP core (PDB: 2PMZ) [3], and eukaryal RNAP II core (PDB: 1Y1W) [12] are shown as ribbon diagrams. Homologous subunits are labeled and shown in the same colors in all three structures; archaeal/eukaryal specific subunits are grey and semitransparent; (b) left: crystal structure of the bacterial RNAP core (PDB: 2O5J) [11] with elements other than the double-psi β -barrel (DPBB) domains made semitransparent; right: an enlarged view of the two DPBB domains (DPBB-A and DPBB-B from the β' and β subunits, correspondingly). The active site is located at the interface of the two domains. Some of the universally conserved amino acid residues (aspartates that coordinate the Mg^{2+} ion and lysines that interact with the backbone of the RNA transcript at the $-1/-2$ positions in the transcribing RNAP [7]) are labeled and shown as stick representation.

RNAP synthesizes RNA in a template-dependent manner in the course of the transcription cycle, which can be subdivided into three stages: initiation, elongation, and termination of transcription. RNAP core enzyme is catalytically active and can operate alone during the elongation and termination stages but requires accessory factors for promoter-specific transcription initiation. For this purpose, bacteria employ one of several σ factors, each of which binds the RNAP core forming a holoenzyme able to recognize promoters with different consensus elements [13]. Archaeal and eukaryal RNAPs use a complex set of general and specific transcription factors, which are evolutionarily unrelated to bacterial σ factors [14,15]. Most archaeal and eukaryal transcription initiation factors first bind promoter DNA and then recruit the RNAP core [14,15].

During the past two decades, extensive genome sequencing revealed genes coding for distant homologs of cellular RNAPs in the genomes of some eukaryotic viruses, bacteriophages, prophages and likely mobile elements located in the genomes of some *Firmicutes* and *Cyanobacteria*, and in fungal killer plasmids [16–22]. Some of these genes were shown to encode functional RNAPs [23–28], while the products of others remain uncharacterized. These partially characterized and non-characterized putative RNAPs are collectively referred to as “non-canonical RNAPs” since they are highly divergent from multisubunit RNAPs of cellular organisms. While none of these enzymes have been studied to the extent “canonical” enzymes were, investigation of non-canonical RNAPs will be required to

achieve a comprehensive view of the function, structure, and evolution, including the evolutionary origins, of multisubunit RNAPs.

2. A Historical Expose: The First Multisubunit Phage RNA Polymerase

The story of multisubunit phage RNAPs starts back in the early nineteen-seventies, when it had been shown that infection of *Bacillus subtilis* by phage PBS2 was insensitive to the addition of rifampicin, a drug targeting bacterial RNAP, even if the drug was administered before the infection [29,30]. This result is in contrast to those obtained with *Escherichia coli* T4 and *B. subtilis* SP01 phages, whose development was strongly inhibited by rifampicin because their genes are transcribed by host RNAP [31,32]. The *E. coli* phage T7 was sensitive to rifampicin (and therefore required the host RNAP) only during the first minutes of infection [33]; at later stages, T7 infection was resistant to the drug because the phage utilizes its own single-subunit rifampicin-resistant RNAP [34].

Inspired by the unusual property of PBS2 infection, the group of Richard Losick and Janice Pero purified a rifampicin-resistant RNAP from *B. subtilis* cells infected with PBS2 [35]. The PBS2 RNAP turned out to be a multisubunit enzyme composed, depending on the purification procedure, of either four or five subunits [35,36]. Based on their molecular weights, not a single subunit of the PBS2 RNAP corresponded to the multisubunit RNAP of the host bacterium. Judging from RNA/DNA hybridization experiments, the PBS2 RNAP was specifically transcribing late PBS2 genes in vitro [36]. Despite it being distinct from both multisubunit bacterial RNAP and single-subunit RNAP of phage T7 and its relatives [35,36], the PBS2 RNAP remained largely forgotten and its properties and evolutionary origins stayed unknown for several decades.

3. Jumbo Phages Encoding RNA Polymerases and Their Features

At the beginning of the 2000s, several independent studies suggested that jumbo phages related to the *Pseudomonas aeruginosa* phage phiKZ encode two sets of unusual proteins with similarities to different parts of the largest subunits of cellular RNAPs (β' and β subunits in bacterial RNAP) [37–40]. The genome of jumbo phage AR9 infecting *B. subtilis*, a close relative of the PBS2 phage, was sequenced and shown to contain genes orthologous to the RNAP genes of phiKZ-like phages [18], suggesting that the PBS2 RNAP purified a long time ago [35,36] in fact belongs to this new group of non-canonical jumbo phage RNAPs. At the time of this writing, several dozens of jumbo phages encoding distant homologs of the β' and β subunits are known [18,37–62]. They infect diverse bacterial hosts including members of *Pseudomonas*, *Salmonella*, *Yersinia*, *Erwinia*, *Vibrio*, *Ralstonia*, *Bacillus*, *Aeromonas*, *Serratia*, *Klebsiella*, *Escherichia*, and other genera [18,37–62]. Remarkably, hosts of jumbo phages encoding two RNAPs include both Gram-positive and Gram-negative bacteria. It is thus possible that the ancestor of phages of this group could have acquired bacterial RNAP genes before the divergence of Gram-positive and Gram-negative bacteria. In addition to RNAPs, these jumbo phages encode a set of core proteins, including phosphoesterase, DnaB-like replicative helicase, large terminase subunit, the split family B DNA polymerase, the SbcCD complex ATPase, RNA helicase, and ribonuclease H [18]. Jumbo phages encoding their own RNAPs likely rely on a common transcriptional strategy during the infection. Jumbo phages that do not encode their own RNAPs are evolutionarily unrelated to jumbo phages with RNAPs and usually encode clearly recognizable homologs of bacterial σ factors likely interacting with the host RNAP core and directing it to phage promoters [63–67].

While most jumbo phages encoding RNAPs have unmodified DNA genomes, the genomes of some, for example, PBS2 and AR9 infecting *B. subtilis*, vB_BpuM_BpSp infecting *B. pumilus* and phiR1-37 infecting *Yersinia enterocolitica*, contain uracil in place of thymine [18,45,49,68]. Another jumbo phage with a genome in which thymine has been replaced by uracil, is staphylococcal phage S6 [69]. The genome of S6 has not been sequenced yet although it has been suggested that this phage is related to PBS2 [18,69]. To our knowledge, no other viruses except these jumbo phages are known to have DNA genomes with uracils.

Another distinguishing feature of jumbo phages encoding RNAPs (which, however, relates only for jumbo phages with thymine-containing DNA genomes) is that most of them form a unique nucleus-like compartment within the infected cell, which separates phage DNA, located inside, from the host cell cytoplasm [70,71]. This “pseudo-nucleus” is positioned at the center of the cell by phage-encoded tubulin-like PhuZ proteins [71]. The pseudo-nucleus provides defense against bacterial anti-phage systems such as CRISPR-Cas and restriction–modification systems of different types (except the Type III and Type VI CRISPR-Cas systems that target RNA) [62,72]. The absence of pseudo-nucleus in jumbo phages with uracil-containing genomes may reflect the fact that modified DNA itself is sufficient to protect their genomes from host nucleases and defense systems.

In the course of infection, phage and host proteins are partitioned between the pseudo-nucleus and cytoplasm according to their functions. For example, phage 201phi2-1 proteins involved in DNA replication and transcription (specifically, two proteins that constitute the β' subunit of the non-virion RNAP) were localized inside the pseudo-nucleus [70]. The host DNA topoisomerase I was also found within the pseudo-nucleus [70]. In contrast, host proteins involved in translation and phage proteins involved in nucleotide metabolism are localized in the cytoplasm and are excluded from the pseudo-nucleus [70]. The mechanism of sorting the proteins between pseudo-nucleus and cytoplasm is unknown.

4. Transcriptional Strategy of Jumbo Phages Encoding RNA Polymerases

Phage phiKZ replication was shown to be resistant to rifampicin, indicating that similarly to PBS2, its development is fully independent of host transcription machinery [73]. Bacteriophage AR9, a close relative of PBS2, also replicates without the involvement of host RNAP, though rifampicin leads to a decreased yield of phage progeny [74]. The decreased AR9 yield in the presence of rifampicin was explained by the detrimental effect of rifampicin on *B. subtilis* cell integrity [74,75]. Indeed, the AR9 yield was dependent on *B. subtilis* strain used for infection and correlated with a number of cells that survived the treatment with the drug [74].

The effect of rifampicin on infection was also investigated for four jumbo phages infecting *Ralstonia solanacearum*-phiRP12, phiRP31, RSL2, and RSF1 [41,53]. Interestingly, despite the presence of all jumbo phage RNAP genes in each genome, only phiRP12 and phiRP31 produced phage progeny in the presence of rifampicin, while the multiplication of RSL2 and RSF1 was abolished [41,53]. In comparison to RSL2 and RSF1, the phiRP12 and phiRP31 genomes possess extra regions coding for proteins with unknown functions [53]. It might be that these proteins are somehow associated with lesser dependence of these phages on cell resources compromised by rifampicin. Yet, it cannot be excluded that host RNAPs transcribe some genes of AR9, RSL2, and RSF1 required for efficient infection.

Global analysis of gene expression in cells infected with several RNAP-encoding jumbo phages was performed. As is typical for most lytic phages, the phiKZ genes are expressed in three temporal classes: early, middle, and late [73]. Transcription profiling of *B. subtilis* cells infected with AR9 revealed early, late and continuously expressed classes of genes [74]. AR9 genes from the unusual latter class are transcribed throughout the infection and have both early and late promoters upstream of them [74]. PhiR1-37, another jumbo phage with a uracil-containing DNA genome, also has a large group of continuously expressed genes [76]. Transcription of early phiKZ and AR9 genes was shown to be independent of protein synthesis in the infected cells [73,74]. Together with the independence of these genes' transcription from host RNAP, this observation points out that the early genes of AR9 and phiKZ are transcribed by the virion-packaged RNAP that must be injected into the host cell along with phage DNA (Figure 2) [73,74]. Indeed whenever it has been investigated, one of the two sets of RNAP β/β' homologs of jumbo phages is found in virions and thus should constitute the virion RNAP (vRNAP) [18,41,45,77,78]. Transcription of middle and late phiKZ genes, and late genes of the AR9 phage, requires protein synthesis and thus must be performed by an enzyme that is fully or partially synthesized *de novo* after infection [73,74]. In accordance with this expectation, the second set of RNAP β/β' homologs is encoded by early phiKZ and AR9 genes [73,74] and must therefore

form their non-virion RNAP (nvRNAP) synthesized during the infection (Figure 2). Overall, it follows from these observations that RNAP-encoding jumbo phages rely on vRNAP (present in virions) for transcription of early genes and nvRNAP (synthesized during the infection) for transcription of late (and, possibly, middle) genes (Figure 2).

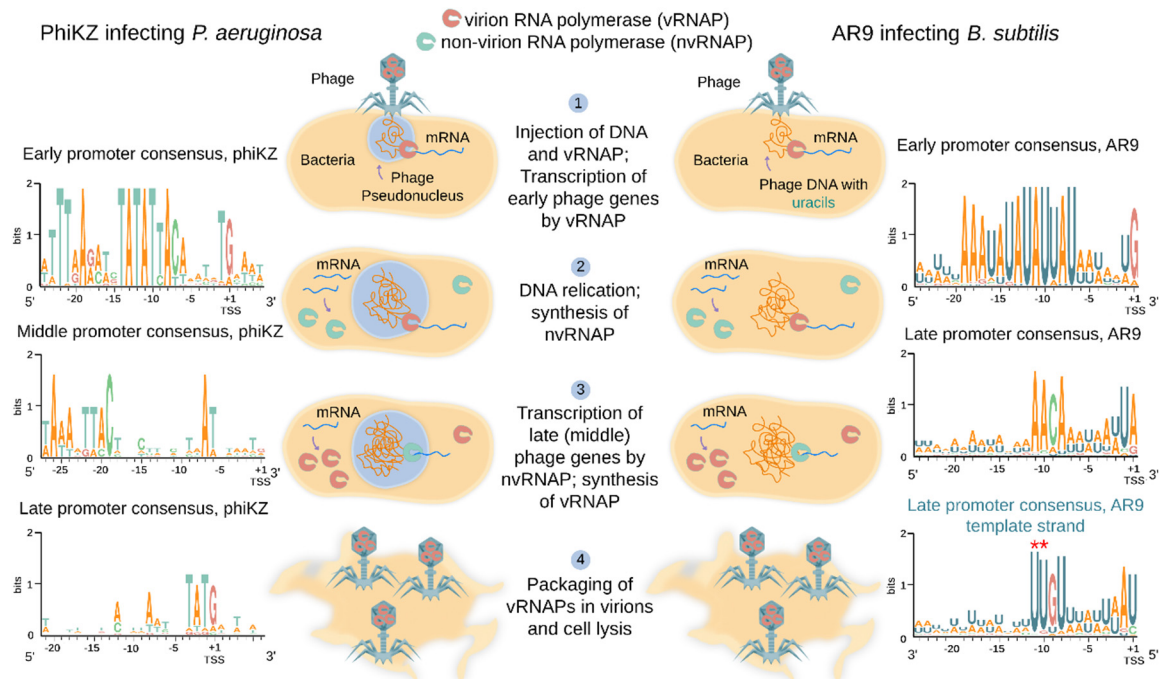


Figure 2. Transcription strategies of jumbo phages encoding two RNAPs. PhiKZ (left) and AR9 (right) phages are, respectively, representatives of thymine-containing DNA genome phages encoding a pseudo-nucleus, and uracil-containing DNA genome phages that do not encode a pseudo-nucleus. The principal steps of the infection cycle are listed in the middle of the figure. WebLogos of phiKZ early, middle, and late promoters are shown on the left. WebLogos of AR9 early and late promoters are shown on the right. Experimentally identified transcription start sites (TSSs) are marked. For the AR9 late promoter, a WebLogo of the template strand is shown in the right bottom corner. The two uracils crucial for promoter recognition are marked by red asterisks (see explanation in Section 6).

In each jumbo phage encoding RNAPs, a conserved motif can be located bioinformatically upstream of presumed early operons [18]. For all phages for which transcript profiling was performed, the predicted motif was indeed shown to be an early promoter consensus [73,74,76]. While different from phage to phage, these motifs are characterized by ~10-nucleotide AT-rich highly conserved sequences centered approximately at position -10 with respect to the transcription start site (TSS) of early genes (Figure 2) [18,73,74]. They must define early phage promoters recognized by vRNAPs.

Bioinformatics analysis fails to predict middle or late promoters of RNAP-encoding jumbo phages. The 5' ends of middle and late phiKZ transcripts were identified from RNA sequencing and primer extension experiments [73]. The only apparent commonality found in DNA located upstream of likely TSSs of middle promoters was a weak AT-rich motif centered at position -24 with respect to TSS (Figure 2) [73]. For late promoters, no sequence conservation upstream of 5' ends of late transcripts could be detected apart from a $5'-T^{-3}ATG^{+1}-3'$ motif overlapping the TSS (Figure 2) [73]. Differential RNA sequencing allowed researchers to determine a consensus motif $5'-A^{-11}ACA-(6N)-UA/G^{+1}-3'$ upstream of late AR9 promoters (Figure 2) [74]. This motif is distinct from either the middle or late promoter consensus motifs of phiKZ.

5. Analysis of RNA Polymerases Genes of Jumbo Phages and Their Origins

Sequence and structural features of multisubunit RNAPs in all three domains of life were described in detail by Lane and Darst [7,79]. The sequences of the jumbo phage β'/β -like subunits were compared with those of corresponding subunits of *Thermus thermophilus* RNAP and most of the universally conserved RNAP regions were identified including the two DPBB domains and the amino acid motifs that comprise the catalytic center [18,28]. Initially, the protein corresponding to the very end of the β' subunit of vRNAP was not found in the genomes of jumbo phages. However, subsequent investigation of essential genes of *Salmonella* jumbo phage SPN3US identified a small virion protein gp244 that corresponds to the C-terminal-most bacterial RNAP β' conserved region [80]. Homologs of SPN3US gp244 can be identified in all RNAP-encoding jumbo phages [80]. Thus, the core of the jumbo phage vRNAP should contain at least five polypeptides that together correspond to full-length β'/β subunits of bacterial RNAP, while the core of nvRNAP contains at least four polypeptides corresponding to full-length β'/β (Figure 3).

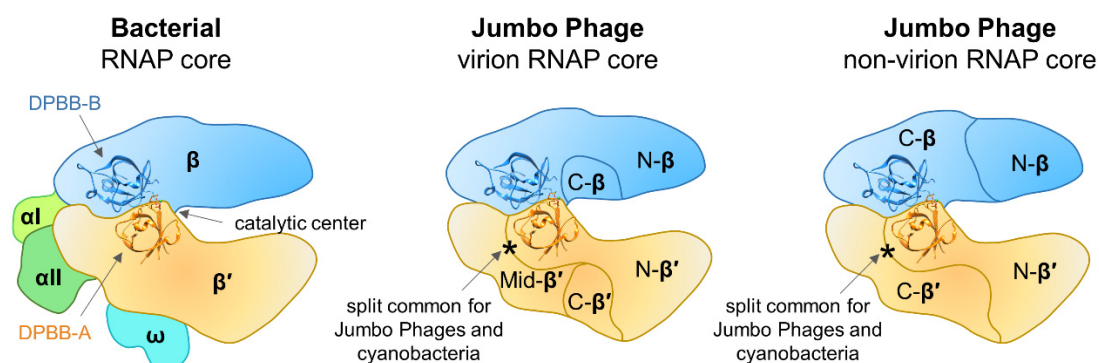


Figure 3. Schematic representation of bacterial and jumbo phage RNAP core enzymes. Structures of DPBB domains are taken from the *T. thermophilus* RNAP crystal structure (PDB: 2O5J) [11]. Jumbo phage RNAPs do not contain homologs of bacterial RNAP α and ω subunits. The virion RNAP (vRNAP) core is composed of at least five polypeptides: N- β' , Mid- β' , C- β' , and N- β , C- β , homologous to the N-terminal, middle, and C-terminal parts of bacterial β' , and N-terminal and C-terminal parts of bacterial β , correspondingly. The non-virion RNAP (nvRNAP) core is composed of four polypeptides: N- β' , C- β' , and N- β , C- β , homologous to the N-terminal and C-terminal parts of β' , and N-terminal and C-terminal parts of β , correspondingly. The DPBB-A is present in the N-terminal parts of β' in both vRNAP and nvRNAP. The split site after DPBB-A (marked by asterisk) is common for vRNAP, nvRNAP, and cyanobacterial RNAP.

Virion RNAPs of *Bacillus* jumbo phages (AR9, PBS1, vB_BpuM_BpSp) have a unique N-terminal HD nuclease domain fused to subunits corresponding to C-terminal fragment of bacterial RNAPs β subunits [18,49]. This domain is present in enzymes that are either known or predicted to possess phosphohydrolase activity and appear to be involved in nucleic acid metabolism or signal transduction in bacteria, archaea, and eukaryotes [81]. The role that the HD domain plays in vRNAP function and phage transcription is unknown.

With available data, it is impossible to trace whether the vRNAP and nvRNAP emerged as a result of gene duplication in the ancestral jumbo phage genome or there were two independent acquisitions of RNAP genes from the host. It was proposed that split jumbo phage RNAP genes could have appeared due to enhanced intron mobility in the ancestral phage followed by the deletion of the introns and genome rearrangements [18]. However, the β' -like subunits of vRNAP and nvRNAP have a common split-site (following the DPBB domain), whose position coincides with a split found in cyanobacterial homologs (Figure 3) [82] but not with that found in archaeal RNAPs [4]. Based on the location of split sites in β' -like subunits of both vRNAP and nvRNAP, it can be hypothesized that these RNAPs were acquired from cyanobacteria, though the alternative hypothesis of their independent appearance can

not be excluded either. The jumbo phages β -like subunits of vRNAP and nvRNAP have split sites in different positions (Figure 3) that do not match that found in archaea [4]. Thus, these splits must have appeared independently in the ancestral jumbo phage and were not inherited from a cellular ancestor.

The uniquely organized multisubunit jumbo phage RNAPs should show notable differences from their cellular counterparts in terms of function. First, they lack compulsory components of multisubunit RNAP core enzymes (α , ω) (Figure 3). Second, jumbo phage genomes do not encode recognizable homologs of any known transcription initiation factors. Thus, it is likely that RNAPs of jumbo phages have unique properties: they must rely on an alternative mechanism of assembly of the core complex and may utilize novel transcription initiation strategies; other aspects of transcription may also significantly differ.

6. In Vitro Properties of Non-Virion RNA Polymerases of Jumbo Phages

Though no vRNAP has been purified to date, nvRNAPs encoded by PBS2, phiKZ, and AR9 phages have been obtained and characterized to different extents [26,28,35,36]. The phiKZ nvRNAP was shown to consist of four predicted phage proteins jointly comprising the full-length β' - and β -like subunits and a fifth subunit gp68 with no sequence similarity to functionally characterized proteins [28]. Homologs of phiKZ gp68 are found in all other jumbo phage genomes encoding β' - and β -like subunits. The phiKZ nvRNAP did not transcribe from DNA fragments containing middle phiKZ promoters in vitro. It was suggested that an alternative form of phiKZ nvRNAP may exist and transcribe from the middle promoters [28]. In contrast, phiKZ nvRNAP efficiently transcribes in vitro from DNA fragments containing phiKZ late promoters containing the 5'-T⁻³ATG⁺¹-3' consensus element [28]. All four conserved nucleotides are strictly required for transcription initiation [28]. Since neighboring nucleotides can be substituted without affecting the in vitro function, it is not clear what determines the specificity of late promoter recognition since the short consensus motif frequently present in the genome can not be sufficient for promoter-specific transcription on its own. It was shown that the region downstream of TSS is important but no conserved motifs were identified there [28]. Since the phiKZ phage as many other jumbo phages with RNAPs and thymine-containing genomes forms a pseudo-nucleus shielding its DNA during infection [62,70–72], it is possible that phage DNA has a specific arrangement within the compartment that affects promoter recognition by phiKZ nvRNAP in vivo.

Similarly to the phiKZ nvRNAP, the AR9 nvRNAP consists of β' and β homologs and the fifth protein gp226 (a homolog of phiKZ gp68) [26]. In contrast to phiKZ nvRNAP it was possible to purify the four-subunit AR9 nvRNAP core complex composed of β' and β homologs only [26,83]. The AR9 nvRNAP is catalytically active but cannot initiate transcription from promoters [26]. The initiation-competent AR9 nvRNAP holoenzyme contains the fifth protein, gp226, which is thus formally equivalent to bacterial RNAP σ factors despite the lack of sequence similarity. The holoenzyme recognizes late phage promoters with a template strand sequence 3'-U⁻¹¹UGU-(6N)-AU/C⁺¹-5' (Figure 2) [26]. The same sequence containing thymines instead of uracils is not recognized and uracils at positions -11 and -10 are strictly required for promoter recognition [26]. It was proposed that the requirement for uracils in promoter consensus element is a strategy that should allow AR9 (and perhaps other jumbo phages with uracil-containing DNA genomes) to avoid unnecessary transcription from host DNA, which contains multiple matches to the simple consensus of phage late promoter [26].

While the requirement for an additional core-binding factor to initiate transcription is similar to that of bacterial RNAPs, in contrast to the σ^{70} -RNAP holoenzymes which bind sequence specifically to the non-template strand of their -10 promoter element (T⁻¹²ATAAT⁻⁷) [26,84,85] the AR9 nvRNAP holoenzyme recognizes the template strand of its promoter (3'-U⁻¹¹UGU-(6N)-AU/C⁺¹-5') (Figure 2) and is even capable of efficient promoter-specific transcription from single-stranded DNA [26]. The molecular mechanism of non-template strand recognition by σ^{70} -RNAP holoenzymes is known: the A⁻¹¹ and T⁻⁷ bases of the -10 element are flipped and bound by the σ_2 domain of σ factors in RNAP holoenzymes [85,86] (C⁻¹⁰ of the consensus sequence G⁻¹²TC⁻¹⁰ plays this role for σ^E -RNAP

holoenzyme [87]). The U^{-11} and U^{-10} of the template strand of AR9 late promoters may be specifically recognized by AR9 nvRNAP in a similar way, through base flipping and burying into protein pockets. In support of this hypothesis, it was shown that U^{-11} and U^{-10} bases become protected from modification by $KMnO_4$ in AR9 nvRNAP-promoter complexes formed on single-stranded DNA, indicating their tight interaction with the enzyme [83]. The low sensitivity to $KMnO_4$ treatment was previously observed for single-stranded T^{-7} of the -10 promoter element in the open promoter complex formed by the σ^{70} -RNAP holoenzyme [88]. The fact that AR9 nvRNAP does not bind and does not melt thymine-only late promoter templates suggests that 5-methyl groups of thymines at positions -11 and -10 of the template strand interfere with the recognition by AR9 nvRNAP [26].

Although we do not know yet the biological role of promoter-specific transcription of single-stranded DNA by AR9 nvRNAP, the ability of AR9 nvRNAP to transcribe single-stranded DNA implies that its elongation and termination properties are different from those of canonical multisubunit RNAPs. The non-template DNA strand was shown to play a crucial role in RNA displacement from the RNA-DNA hybrid during transcription by cellular multisubunit RNAPs [89,90]. If AR9 nvRNAP also separates RNA from single-stranded DNA during transcription elongation, it must accomplish this by a unique mechanism that is absent in cellular enzymes. It was shown that in *in vitro* single-stranded DNA transcription reactions, a significant amount of AR9 nvRNAP-synthesized transcripts are present in a free form when the concentration of DNA template is lower than that of RNAP, indicating the ability of the enzyme to separate RNA from single-stranded DNA, i.e., in the absence of non-template strand [83]. The RNA transcript may be separated from the single-stranded DNA template in two ways. The first way is the direct displacement of the nascent RNA from the RNA-DNA hybrid achieved through an unknown structural feature that should distinguish the AR9 nvRNAP from other cellular RNAPs. Alternatively, it is possible that the first round of transcription by AR9 nvRNAP leads to the formation of RNA-DNA fork-junctions with nascent RNA occupying the position of non-template DNA strand downstream of TSS. This annealed RNA can be displaced from the hybrid not during its synthesis but in the course of a subsequent round of transcription. This scenario is only possible because AR9 nvRNAP recognizes the template strand of its promoter: it would be impossible for other known RNAPs requiring the non-template promoter DNA strand. The AR9 DNA is very AU-rich (72.25%) [18] and thus it was proposed that phage DNA may be present in a partially single-stranded form in infected cells, especially during phage DNA replication and the unique properties of phage nvRNAP may allow it to utilize such partially single-stranded DNA for specific transcription of late genes [26].

The AR9 nvRNAP core is unable to bind promoter DNA but is catalytically active [26]. How the β' - and β -like subunits of the AR9 nvRNAP (and nvRNAPs of other jumbo phages) are kept together in the absence of additional subunits (especially considering that both counterparts of the β' and β are split into several polypeptides) remains unknown. The arrangement where catalytic subunits of a multisubunit RNAP alone constitute an active RNAP core in the absence of additional subunits is a unique one. The α subunits forming the assembly platform likely were present in the multisubunit RNAP of the Last Universal Common Ancestor since all cellular multisubunit RNAPs contain them (Figure 1a). The absence of the assembly platform from jumbo phage nvRNAP core (and, likely, from vRNAPs) likely means that RNAPs of jumbo phages derived from cellular enzymes must have evolved in a way that strengthened the interactions between the catalytic subunits which allowed the loss of the assembly platform.

7. Concluding Remarks

In this contribution, we described transcriptional strategies used by RNAP-encoding jumbo phages and properties of jumbo phage nvRNAPs characterized to date. The vRNAP has not been yet purified from any jumbo phage. In addition to potentially unique transcription properties, the vRNAPs may possess remarkable folding capabilities because they are likely delivered into the cell through a narrow

tail channel in an unfolded state and their subunits must efficiently find each other upon the delivery into the host. These properties, if confirmed, can be of interest for biotechnological applications.

The investigation of jumbo phage RNAPs is important for the understanding of multisubunit RNAPs evolution and tracing the possibly ancient origins of RNAP-encoding jumbo phages. Recently, a distinct group of multisubunit RNAP-encoding phages was discovered by analyzing metagenomics data [91]. Thorough phylogenetic analyses placed the newly identified RNAPs as a separate branch on the phylogenetic tree and these RNAPs were proposed to be acquired from an ancient cell before the split of eubacteria from other kingdoms [91]. The authors excluded the jumbo phage RNAPs from their analysis because it was impossible to confidently place them on the phylogenetic tree likely due to the Long Branch Attraction effect caused by their rapid evolution [91]. Thus, the origins of jumbo phage RNAPs remain a mystery. It is possible, however, that once available sequence information is supplemented with jumbo phage RNAPs structures, building a phylogenetic tree, which includes jumbo phage RNAPs, will become possible. This structural work is currently ongoing in our laboratories.

Author Contributions: Writing—original draft preparation, M.L.S., I.M.; writing—review and editing, K.V.S.; funding acquisition, M.L.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Russian Science Foundation, grant number 19-74-00011 to M.L.S.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Werner, F.; Grohmann, D. Evolution of multisubunit RNA polymerases in the three domains of life. *Nat. Rev. Microbiol.* **2011**, *9*, 85–98. [[CrossRef](#)] [[PubMed](#)]
2. Zhou, M.; Law, J.A. RNA Pol IV and V in gene silencing: Rebel polymerases evolving away from Pol II's rules. *Curr. Opin. Plant Biol.* **2015**, *27*, 154–164. [[CrossRef](#)] [[PubMed](#)]
3. Hirata, A.; Klein, B.J.; Murakami, K.S. The X-ray crystal structure of RNA polymerase from Archaea. *Nature* **2008**, *451*, 851–854. [[CrossRef](#)] [[PubMed](#)]
4. Werner, F. Structure and function of archaeal RNA polymerases. *Mol. Microbiol.* **2007**, *65*, 1395–1404. [[CrossRef](#)] [[PubMed](#)]
5. Werner, F. Structural evolution of multisubunit RNA polymerases. *Trends Microbiol.* **2008**, *16*, 247–250. [[CrossRef](#)]
6. Ishihama, A. Subunit of assembly of Escherichia coli RNA polymerase. *Adv. Biophys.* **1981**, *14*, 1–35.
7. Lane, W.J.; Darst, S.A. Molecular evolution of multisubunit RNA polymerases: Structural analysis. *J. Mol. Biol.* **2010**, *395*, 686–704. [[CrossRef](#)]
8. Minakhin, L.; Bhagat, S.; Brunning, A.; Campbell, E.A.; Darst, S.A.; Ebright, R.H.; Severinov, K. Bacterial RNA polymerase subunit omega and eukaryotic RNA polymerase subunit RPB6 are sequence, structural, and functional homologs and promote RNA polymerase assembly. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 892–897. [[CrossRef](#)]
9. Vassylyev, D.G.; Sekine, S.; Laptenko, O.; Lee, J.; Vassylyeva, M.N.; Borukhov, S.; Yokoyama, S. Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature* **2002**, *417*, 712–719. [[CrossRef](#)]
10. Zhang, G.; Campbell, E.A.; Minakhin, L.; Richter, C.; Severinov, K.; Darst, S.A. Crystal structure of Thermus aquaticus core RNA polymerase at 3.3 Å resolution. *Cell* **1999**, *98*, 811–824. [[CrossRef](#)]
11. Vassylyev, D.G.; Vassylyeva, M.N.; Zhang, J.; Palangat, M.; Artsimovitch, I.; Landick, R. Structural basis for substrate loading in bacterial RNA polymerase. *Nature* **2007**, *448*, 163–168. [[CrossRef](#)] [[PubMed](#)]
12. Kettenberger, H.; Armache, K.J.; Cramer, P. Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIS. *Mol. Cell.* **2004**, *16*, 955–965. [[CrossRef](#)] [[PubMed](#)]
13. Paget, M.S. Bacterial Sigma Factors and Anti-Sigma Factors: Structure, Function and Distribution. *Biomolecules* **2015**, *5*, 1245–1265. [[CrossRef](#)]
14. Blombach, F.; Smollett, K.L.; Grohmann, D.; Werner, F. Molecular Mechanisms of Transcription Initiation-Structure, Function, and Evolution of TFE/TFIIE-Like Factors and Open Complex Formation. *J. Mol. Biol.* **2016**, *428*, 2592–2606. [[CrossRef](#)] [[PubMed](#)]

15. Hantsche, M.; Cramer, P. Conserved RNA polymerase II initiation complex structure. *Curr. Opin. Struct. Biol.* **2017**, *47*, 17–22. [[CrossRef](#)]
16. Berdygulova, Z.; Westblade, L.F.; Florens, L.; Koonin, E.V.; Chait, B.T.; Ramanculov, E.; Washburn, M.P.; Darst, S.A.; Severinov, K.; Minakhin, L. Temporal regulation of gene expression of the *Thermus thermophilus* bacteriophage P23-45. *J. Mol. Biol.* **2011**, *405*, 125–142. [[CrossRef](#)]
17. Iyer, L.M.; Aravind, L. Insights from the architecture of the bacterial transcription apparatus. *J. Struct. Biol.* **2012**, *179*, 299–319. [[CrossRef](#)]
18. Lavysh, D.; Sokolova, M.; Minakhin, L.; Yakunina, M.; Artamonova, T.; Kozyavkin, S.; Makarova, K.S.; Koonin, E.V.; Severinov, K. The genome of AR9, a giant transducing *Bacillus* phage encoding two multisubunit RNA polymerases. *Virology* **2016**, *495*, 185–196. [[CrossRef](#)]
19. Mirzakhanyan, Y.; Gershon, P.D. Multisubunit DNA-Dependent RNA Polymerases from Vaccinia Virus and Other Nucleocytoplasmic Large-DNA Viruses: Impressions from the Age of Structure. *Microbiol. Mol. Biol. Rev.* **2017**, *81*, e00010–e00017. [[CrossRef](#)]
20. Ruprich-Robert, G.; Thuriaux, P. Non-canonical DNA transcription enzymes and the conservation of two-barrel RNA polymerases. *Nucleic Acids Res.* **2010**, *38*, 4559–4569. [[CrossRef](#)]
21. Sykora, M.; Pospisek, M.; Novak, J.; Mrvova, S.; Krasny, L.; Vopalensky, V. Transcription apparatus of the yeast virus-like elements: Architecture, function, and evolutionary origin. *PLoS Pathog.* **2018**, *14*, e1007377. [[CrossRef](#)]
22. Yutin, N.; Makarova, K.S.; Gussow, A.B.; Krupovic, M.; Segall, A.; Edwards, R.A.; Koonin, E.V. Discovery of an expansive bacteriophage family that includes the most abundant viruses from the human gut. *Nat. Microbiol.* **2018**, *3*, 38–46. [[CrossRef](#)]
23. Ahn, B.Y.; Moss, B. RNA polymerase-associated transcription specificity factor encoded by vaccinia virus. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 3536–3540. [[CrossRef](#)]
24. Forrest, D.; James, K.; Yuzenkova, Y.; Zenkin, N. Single-peptide DNA-dependent RNA polymerase homologous to multi-subunit RNA polymerase. *Nat. Commun.* **2017**, *8*, 15774. [[CrossRef](#)]
25. Guarino, L.A.; Dong, W.; Jin, J. In vitro activity of the baculovirus late expression factor LEF-5. *J. Virol.* **2002**, *76*, 12663–12675. [[CrossRef](#)]
26. Sokolova, M.; Borukhov, S.; Lavysh, D.; Artamonova, T.; Khodorkovskii, M.; Severinov, K. A non-canonical multisubunit RNA polymerase encoded by the AR9 phage recognizes the template strand of its uracil-containing promoters. *Nucleic Acids Res.* **2017**, *45*, 5958–5967. [[CrossRef](#)]
27. Spencer, E.; Shuman, S.; Hurwitz, J. Purification and properties of vaccinia virus DNA-dependent RNA polymerase. *J. Biol. Chem.* **1980**, *255*, 5388–5395.
28. Yakunina, M.; Artamonova, T.; Borukhov, S.; Makarova, K.S.; Severinov, K.; Minakhin, L. A non-canonical multisubunit RNA polymerase encoded by a giant bacteriophage. *Nucleic Acids Res.* **2015**, *43*, 10411–10420. [[CrossRef](#)]
29. Rima, B.K.; Takahashi, I. The synthesis of nucleic acids in *Bacillus subtilis* infected with phage PBS 1. *Can. J. Biochem.* **1973**, *51*, 1219–1224. [[CrossRef](#)]
30. Price, A.R.; Frabotta, M. Resistance of bacteriophage PBS2 infection to rifampicin, an inhibitor of *Bacillus subtilis* RNA synthesis. *Biochem. Biophys. Res. Commun.* **1972**, *48*, 1578–1585. [[CrossRef](#)]
31. Geiduschek, E.P.; Sklar, J. Continual requirement for a host RNA polymerase component in a bacteriophage development. *Nature* **1969**, *221*, 833–836. [[CrossRef](#)] [[PubMed](#)]
32. Haselkorn, R.; Vogel, M.; Brown, R.D. Conservation of the rifampicin sensitivity of transcription during T4 development. *Nature* **1969**, *221*, 836–838. [[CrossRef](#)] [[PubMed](#)]
33. Summers, W.C.; Siegel, R.B. Control of template specificity of *E. coli* RNA polymerase by a phage-coded protein. *Nature* **1969**, *223*, 1111–1113. [[CrossRef](#)]
34. Chamberlin, M.; McGrath, J.; Waskell, L. New RNA polymerase from *Escherichia coli* infected with bacteriophage T7. *Nature* **1970**, *228*, 227–231. [[CrossRef](#)] [[PubMed](#)]
35. Clark, S.; Losick, R.; Pero, J. New RNA polymerase from *Bacillus subtilis* infected with phage PBS2. *Nature* **1974**, *252*, 21–24. [[CrossRef](#)]
36. Clark, S. Transcriptional specificity of a multisubunit RNA polymerase induced by *Bacillus subtilis* bacteriophage PBS2. *J. Virol.* **1978**, *25*, 224–237. [[CrossRef](#)] [[PubMed](#)]

37. Mesyanzhinov, V.V.; Robben, J.; Grymonprez, B.; Kostyuchenko, V.A.; Bourkaltseva, 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. [[CrossRef](#)]
38. 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**, *354*, 536–545. [[CrossRef](#)]
39. Thomas, J.A.; Rolando, M.R.; Carroll, C.A.; Shen, P.S.; Belnap, D.M.; Weintraub, S.T.; Serwer, P.; Hardies, S.C. Characterization of *Pseudomonas chlororaphis* myovirus 201varphi2-1 via genomic sequencing, mass spectrometry, and electron microscopy. *Virology* **2008**, *376*, 330–338. [[CrossRef](#)]
40. 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. [[CrossRef](#)]
41. Bhunchoth, A.; Blanc-Mathieu, R.; Mihara, T.; Nishimura, Y.; Askora, A.; Phironrit, N.; Leksomboon, C.; Chatchawankanphanich, O.; Kawasaki, T.; Nakano, M.; et al. Two asian jumbo phages, phiRSL2 and phiRSF1, infect *Ralstonia solanacearum* and show common features of phiKZ-related phages. *Virology* **2016**, *494*, 56–66. [[CrossRef](#)] [[PubMed](#)]
42. Monson, R.; Foulds, I.; Foweraker, J.; Welch, M.; Salmond, G.P.C. The *Pseudomonas aeruginosa* generalized transducing phage phiPA3 is a new member of the phiKZ-like group of ‘jumbo’ phages, and infects model laboratory strains and clinical isolates from cystic fibrosis patients. *Microbiology* **2011**, *157 Pt 3*, 859–867. [[CrossRef](#)]
43. Lee, J.H.; Shin, H.; Kim, H.; Ryu, S. Complete genome sequence of *Salmonella* bacteriophage SPN3US. *J. Virol.* **2011**, *85*, 13470–13471. [[CrossRef](#)] [[PubMed](#)]
44. 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. [[CrossRef](#)] [[PubMed](#)]
45. Skurnik, M.; Hyytiainen, H.J.; Happonen, L.J.; Kiljunen, S.; Datta, N.; Mattinen, L.; Williamson, K.; Kristo, P.; Szeliga, M.; Kalin-Manttari, L.; et al. Characterization of the genome, proteome, and structure of yersiniophage varphiR1-37. *J. Virol.* **2012**, *86*, 12625–12642. [[CrossRef](#)]
46. Domotor, D.; Becsagh, P.; Rakhely, G.; Schneider, G.; Kovacs, T. Complete genomic sequence of *Erwinia amylovora* phage PhiEaH2. *J. Virol.* **2012**, *86*, 10899. [[CrossRef](#)]
47. Meczker, K.; Domotor, D.; Vass, J.; Rakhely, G.; Schneider, G.; Kovacs, T. The genome of the *Erwinia amylovora* phage PhiEaH1 reveals greater diversity and broadens the applicability of phages for the treatment of fire blight. *FEMS Microbiol. Lett.* **2014**, *350*, 25–27. [[CrossRef](#)]
48. Yagubi, A.I.; Castle, A.J.; Kropinski, A.M.; Banks, T.W.; Svircev, A.M. Complete Genome Sequence of *Erwinia amylovora* Bacteriophage vB_EamM_Ea35-70. *Genome Announc.* **2014**, *2*, e00413–e00414. [[CrossRef](#)]
49. Yuan, Y.; Gao, M. Proteomic Analysis of a Novel *Bacillus* Jumbo Phage Revealing Glycoside Hydrolase As Structural Component. *Front. Microbiol.* **2016**, *7*, 745. [[CrossRef](#)]
50. Danis-Wlodarczyk, K.; Vandenheuvel, D.; Jang, H.B.; Briers, Y.; Olszak, T.; Arabski, M.; Wasik, S.; Drabik, M.; Higgins, G.; Tyrrell, J.; et al. A proposed integrated approach for the preclinical evaluation of phage therapy in *Pseudomonas* infections. *Sci. Rep.* **2016**, *6*, 28115. [[CrossRef](#)]
51. Lee, J.H.; Bai, J.; Shin, H.; Kim, Y.; Park, B.; Heu, S.; Ryu, S. A Novel Bacteriophage Targeting *Cronobacter sakazakii* Is a Potential Biocontrol Agent in Foods. *Appl. Environ. Microbiol.* **2016**, *82*, 192–201. [[CrossRef](#)] [[PubMed](#)]
52. Esplin, I.N.D.; Berg, J.A.; Sharma, R.; Allen, R.C.; Arens, D.K.; Ashcroft, C.R.; Bairett, S.R.; Beatty, N.J.; Bickmore, M.; Bloomfield, T.J.; et al. Genome Sequences of 19 Novel *Erwinia amylovora* Bacteriophages. *Genome Announc.* **2017**, *5*. [[CrossRef](#)]
53. Matsui, T.; Yoshikawa, G.; Mihara, T.; Chatchawankanphanich, O.; Kawasaki, T.; Nakano, M.; Fujie, M.; Ogata, H.; Yamada, T. Replications of Two Closely Related Groups of Jumbo Phages Show Different Level of Dependence on Host-encoded RNA Polymerase. *Front. Microbiol.* **2017**, *8*, 1010. [[CrossRef](#)] [[PubMed](#)]
54. Wojtus, J.K.; Fitch, J.L.; Christian, E.; Dalefield, T.; Lawes, J.K.; Kumar, K.; Peebles, C.L.; Altermann, E.; Hendrickson, H.L. Complete Genome Sequences of Three Novel *Pseudomonas fluorescens* SBW25 Bacteriophages, Noxifer, Phabio, and Skulduggery. *Genome Announc.* **2017**, *5*. [[CrossRef](#)] [[PubMed](#)]

55. Latz, S.; Kruttgen, A.; Hafner, H.; Buhl, E.M.; Ritter, K.; Horz, H.P. Differential Effect of Newly Isolated Phages Belonging to PB1-Like, phiKZ-Like and LUZ24-Like Viruses against Multi-Drug Resistant *Pseudomonas aeruginosa* under Varying Growth Conditions. *Viruses* **2017**, *9*, 315. [[CrossRef](#)] [[PubMed](#)]
56. Kim, S.G.; Jun, J.W.; Giri, S.S.; Yun, S.; Kim, H.J.; Kim, S.W.; Kang, J.W.; Han, S.J.; Jeong, D.; Park, S.C. Isolation and characterisation of pVa-21, a giant bacteriophage with anti-biofilm potential against *Vibrio alginolyticus*. *Sci. Rep.* **2019**, *9*, 6284. [[CrossRef](#)]
57. Wojtus, J.K.; Frampton, R.A.; Warring, S.; Hendrickson, H.; Fineran, P.C. Genome Sequence of a Jumbo Bacteriophage That Infects the Kiwifruit Phytopathogen *Pseudomonas syringae* pv. *actinidiae*. *Microbiol. Resour. Announc.* **2019**, *8*. [[CrossRef](#)]
58. Korf, I.H.E.; Meier-Kolthoff, J.P.; Adriaenssens, E.M.; Kropinski, A.M.; Nimtz, M.; Rohde, M.; van Raaij, M.J.; Wittmann, J. Still Something to Discover: Novel Insights into *Escherichia coli* Phage Diversity and Taxonomy. *Viruses* **2019**, *11*, 454. [[CrossRef](#)]
59. Price, L.; Rohren, M.; Newkirk, H.; Liu, M.; Ramsey, J. Complete Genome Sequence of *Serratia marcescens* Myophage Moabite. *Microbiol. Resour. Announc.* **2019**, *8*. [[CrossRef](#)]
60. Lewis, R.; Clooney, A.G.; Stockdale, S.R.; Buttmer, C.; Draper, L.A.; Ross, R.P.; Hill, C. Isolation of a Novel Jumbo Bacteriophage Effective Against *Klebsiella aerogenes*. *Front. Med.* **2020**, *7*, 67. [[CrossRef](#)]
61. Ong, S.P.; Azam, A.H.; Sasahara, T.; Miyana, K.; Tanji, Y. Characterization of *Pseudomonas* lytic phages and their application as a cocktail with antibiotics in controlling *Pseudomonas aeruginosa*. *J. Biosci. Bioeng.* **2020**, *129*, 693–699. [[CrossRef](#)] [[PubMed](#)]
62. Malone, L.M.; Warring, S.L.; Jackson, S.A.; Warnecke, C.; Gardner, P.P.; Gumy, L.F.; Fineran, P.C. A jumbo phage that forms a nucleus-like structure evades CRISPR-Cas DNA targeting but is vulnerable to type III RNA-based immunity. *Nat. Microbiol.* **2020**, *5*, 48–55. [[CrossRef](#)] [[PubMed](#)]
63. Simoliunas, E.; Kaliniene, L.; Truncaite, L.; Zajanckauskaite, A.; Staniulis, J.; Kaupinis, A.; Ger, M.; Valius, M.; Meskys, R. *Klebsiella* phage vB_KleM-RaK2-a giant singleton virus of the family Myoviridae. *PLoS ONE* **2013**, *8*, e60717. [[CrossRef](#)]
64. Abbasifar, R.; Griffiths, M.W.; Sabour, P.M.; Ackermann, H.W.; Vandersteegen, K.; Lavigne, R.; Noben, J.P.; Alanis Villa, A.; Abbasifar, A.; Nash, J.H.; et al. Supersize me: *Cronobacter sakazakii* phage GAP32. *Virology* **2014**, *460–461*, 138–146. [[CrossRef](#)] [[PubMed](#)]
65. Pan, Y.J.; Lin, T.L.; Chen, C.C.; Tsai, Y.T.; Cheng, Y.H.; Chen, Y.Y.; Hsieh, P.F.; Lin, Y.T.; Wang, J.T. *Klebsiella* Phage PhiK64-1 Encodes Multiple Depolymerases for Multiple Host Capsular Types. *J. Virol.* **2017**, *91*. [[CrossRef](#)]
66. Buttmer, C.; Hendrix, H.; Oliveira, H.; Casey, A.; Neve, H.; McAuliffe, O.; Ross, R.P.; Hill, C.; Noben, J.P.; O'Mahony, J.; et al. Things Are Getting Hairy: Enterobacteria Bacteriophage vB_PcaM_CBB. *Front. Microbiol.* **2017**, *8*, 44. [[CrossRef](#)]
67. Evseev, P.; Sykilinda, N.; Gorshkova, A.; Kurochkina, L.; Ziganshin, R.; Drucker, V.; Miroshnikov, K. *Pseudomonas* Phage PaBG-A Jumbo Member of an Old Parasite Family. *Viruses* **2020**, *12*, 721. [[CrossRef](#)]
68. Takahashi, I.; Marmur, J. Replacement of thymidylic acid by deoxyuridylic acid in the deoxyribonucleic acid of a transducing phage for *Bacillus subtilis*. *Nature* **1963**, *197*, 794–795. [[CrossRef](#)]
69. Uchiyama, J.; Takemura-Uchiyama, I.; Sakaguchi, Y.; Gamoh, K.; Kato, S.; Daibata, M.; Ujihara, T.; Misawa, N.; Matsuzaki, S. Intragenus generalized transduction in *Staphylococcus* spp. by a novel giant phage. *ISME J.* **2014**, *8*, 1949–1952. [[CrossRef](#)]
70. Chaikerasitak, V.; Nguyen, K.; Khanna, K.; Brilot, A.F.; Erb, M.L.; Coker, J.K.; Vavilina, A.; Newton, G.L.; Buschauer, R.; Pogliano, K.; et al. Assembly of a nucleus-like structure during viral replication in bacteria. *Science* **2017**, *355*, 194–197. [[CrossRef](#)]
71. Chaikerasitak, V.; Nguyen, K.; Egan, M.E.; Erb, M.L.; Vavilina, A.; Pogliano, J. The Phage Nucleus and Tubulin Spindle Are Conserved among Large *Pseudomonas* Phages. *Cell Rep.* **2017**, *20*, 1563–1571.
72. Mendoza, S.D.; Nieweglowska, E.S.; Govindarajan, S.; Leon, L.M.; Berry, J.D.; Tiwari, A.; Chaikerasitak, V.; Pogliano, J.; Agard, D.A.; Bondy-Denomy, J. A bacteriophage nucleus-like compartment shields DNA from CRISPR nucleases. *Nature* **2020**, *577*, 244–248.
73. Ceysens, P.J.; Minakhin, L.; Van den Bossche, A.; Yakunina, M.; Klimuk, E.; Blasdel, B.; De Smet, J.; Noben, J.P.; Blasi, U.; Severinov, K.; et al. Development of giant bacteriophage varphiKZ is independent of the host transcription apparatus. *J. Virol.* **2014**, *88*, 10501–10510.

74. Lavysh, D.; Sokolova, M.; Slashcheva, M.; Forstner, K.U.; Severinov, K. Transcription Profiling of *Bacillus subtilis* Cells Infected with AR9, a Giant Phage Encoding Two Multisubunit RNA Polymerases. *MBio* **2017**, *8*. [[CrossRef](#)]
75. Coote, J.G.; Wood, D.A.; Mandelstam, J. Lethal effect of rifampicin in *Bacillus subtilis* as a complicating factor in the assessment of the lifetime of messenger ribonucleic acid. *Biochem. J.* **1973**, *134*, 263–270.
76. Leskinen, K.; Blasdel, B.G.; Lavigne, R.; Skurnik, M. RNA-Sequencing Reveals the Progression of Phage-Host Interactions between phiR1-37 and *Yersinia enterocolitica*. *Viruses* **2016**, *8*, 111.
77. Lecoutere, E.; Ceysens, 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.
78. Thomas, J.A.; Weintraub, S.T.; Hakala, K.; Serwer, P.; Hardies, S.C. Proteome of the large *Pseudomonas myovirus* 201 phi 2-1: Delineation of proteolytically processed virion proteins. *Mol. Cell. Proteom.* **2010**, *9*, 940–951.
79. Lane, W.J.; Darst, S.A. Molecular evolution of multisubunit RNA polymerases: Sequence analysis. *J. Mol. Biol.* **2010**, *395*, 671–685.
80. Thomas, J.A.; Benitez Quintana, A.D.; Bosch, M.A.; Coll De Pena, A.; Aguilera, E.; Coulibaly, A.; Wu, W.; Osier, M.V.; Hudson, A.O.; Weintraub, S.T.; et al. Identification of Essential Genes in the Salmonella Phage SPN3US Reveals Novel Insights into Giant Phage Head Structure and Assembly. *J. Virol.* **2016**, *90*, 10284–10298.
81. Aravind, L.; Koonin, E.V. The HD domain defines a new superfamily of metal-dependent phosphohydrolases. *Trends Biochem. Sci.* **1998**, *23*, 469–472. [[CrossRef](#)]
82. Xie, W.Q.; Jager, K.; Potts, M. Cyanobacterial RNA polymerase genes rpoC1 and rpoC2 correspond to rpoC of *Escherichia coli*. *J. Bacteriol.* **1989**, *171*, 1967–1973. [[CrossRef](#)]
83. Sokolova, M. Functional and Structural Analysis of a Non-Canonical Multisubunit RNA Polymerase Encoded by Giant Bacteriophage AR9. Ph.D. Thesis, Skolkovo Institute of Science and Technology, Moscow, Russia, 2018.
84. Feklistov, A. RNA polymerase: In search of promoters. *Ann. N. Y. Acad. Sci.* **2013**, *1293*, 25–32. [[CrossRef](#)]
85. Feklistov, A.; Darst, S.A. Structural basis for promoter-10 element recognition by the bacterial RNA polymerase sigma subunit. *Cell* **2011**, *147*, 1257–1269. [[CrossRef](#)]
86. Bae, B.; Feklistov, A.; Lass-Napiorkowska, A.; Landick, R.; Darst, S.A. Structure of a bacterial RNA polymerase holoenzyme open promoter complex. *Elife* **2015**, *4*, e08504.
87. Campagne, S.; Marsh, M.E.; Capitani, G.; Vorholt, J.A.; Allain, F.H. Structural basis for -10 promoter element melting by environmentally induced sigma factors. *Nat. Struct. Mol. Biol.* **2014**, *21*, 269–276.
88. Zhilina, E.; Esyunina, D.; Brodolin, K.; Kulbachinskiy, A. Structural transitions in the transcription elongation complexes of bacterial RNA polymerase during sigma-dependent pausing. *Nucleic Acids Res.* **2012**, *40*, 3078–3091.
89. Touloukhonov, I.; Landick, R. The role of the lid element in transcription by *E. coli* RNA polymerase. *J. Mol. Biol.* **2006**, *361*, 644–658.
90. Naryshkina, T.; Kuznedelov, K.; Severinov, K. The role of the largest RNA polymerase subunit lid element in preventing the formation of extended RNA-DNA hybrid. *J. Mol. Biol.* **2006**, *361*, 634–643.
91. Weinheimer, A.R.; Aylward, F.O. A distinct lineage of Caudovirales that encodes a deeply branching multi-subunit RNA polymerase. *Nat. Commun.* **2020**, *11*, 4506.

