viruses

MDPI

Review

Recent Advances in Structural Studies of Single-Stranded RNA Bacteriophages

Jirapat Thongchol, Zachary Lill, Zachary Hoover and Junjie Zhang *

Center for Phage Technology, Department of Biochemistry and Biophysics, Texas A&M University, College Station, TX 77843, USA; jirapat.tho@tamu.edu (J.T.); zachlill@tamu.edu (Z.L.); zachhoov@tamu.edu (Z.H.)
* Correspondence: junjie.zhang@agnet.tamu.edu

Abstract: Positive-sense single-stranded RNA (ssRNA) bacteriophages (phages) were first isolated six decades ago. Since then, extensive research has been conducted on these ssRNA phages, particularly those infecting E. coli. With small genomes of typically 3–4 kb that usually encode four essential proteins, ssRNA phages employ a straightforward infectious cycle involving host adsorption, genome entry, genome replication, phage assembly, and host lysis. Recent advancements in metagenomics and transcriptomics have led to the identification of ~65,000 sequences from ssRNA phages, expanding our understanding of their prevalence and potential hosts. This review article illuminates significant investigations into ssRNA phages, with a focal point on their structural aspects, providing insights into the various stages of their infectious cycle.

Keywords: ssRNA phages; adsorption; replication; packaging; single-lysis protein

check for updates

Citation: Thongchol, J.; Lill, Z.; Hoover, Z.; Zhang, J. Recent Advances in Structural Studies of Single-Stranded RNA Bacteriophages. Viruses **2023**, 15, 1985. https://doi.org/10.3390/v15101985

Academic Editors: Wen Jiang and Matthew Lowell Baker

Received: 7 August 2023
Revised: 20 September 2023
Accepted: 21 September 2023
Published: 23 September 2023

1. Introduction

Positive-sense ssRNA phages are genetically and structurally simple RNA viruses that infect Gram-negative bacteria by utilizing bacterial retractile pili as receptors. The discovery of the first ssRNA phages occurred in 1960 [1], specifically those infecting E. coli through the F-pilus. Subsequently, other ssRNA phages that rely on different types of retractile pili as receptors have been isolated. Up until 2016, only a limited number of ssRNA phage species had been documented in databases. Table 1 provides an overview of the host range and receptors for those known ssRNA phages, highlighting their ability to evolve and target various types of retractile pili, such as the conjugative type IV secretion system (T4SS) pilus, the type IV pilus (T4P), and the tight adhesion (Tad) pilus.

ssRNA phages are lytic phages characterized by small genomes of approximately 3000–4000 nucleotides. These genomes usually encode four essential proteins: the maturation protein (Mat), the coat protein (CP), the -subunit of the replicase (Rep), and a single lysis protein (Lys) (Figure 1A). Mat and CP are responsible for the phage's structural integrity, while Rep replicates the viral genome within the host cell. Lys is necessary for lysing the host cell, allowing the release of newly formed viral progeny into the environment. The genomic architecture of ssRNA phages typically follows a similar pattern from the 5^0 to the 3^0 ends, with the mat gene located at the beginning, followed by cp and rep. However, the lys gene is an exception. It seems to have evolved to be distributed throughout the genomes of different ssRNA phages.

In 2016, ssRNA phage research entered the metagenomic and transcriptomic era, with 158 ssRNA phage sequences being identified [2]. From 2018 on, ssRNA sequences increased drastically, leading to the identification of thousands of new ssRNA genomes [3–5]. The identification of these phage genomes from metagenomic data of environmental samples is based on the detection of the rep gene sequence, which is highly conserved among ssRNA phages. This breakthrough prompted the taxonomic reclassification of ssRNA phages, resulting in the creation of two new orders (based on the Rep) and six families (based on

the CP), into a total of 428 genera (50% Rep pairwise amino acid sequence identity) and 882 species (80% Rep pairwise amino acid sequence identity) [6]. This not only highlights the prevalence of ssRNA phages in nature, but also suggests the potential existence of additional hosts and retractile pili that ssRNA phages can target. However, the hosts of these newly identified ssRNA phages remain unknown. By 2022, the number of identified ssRNA phage genomes had further expanded to 65,814 sequences [4]. Using the Hidden Markov Model (HMM), which is a statistical model to analyze protein sequence similarity and identify protein domains, from the 2020 literature [3], we revealed more possibly near-complete genomes (~12,288 genomes) in the 2022 dataset [4] that possess the three core genes: mat, cp, and rep.

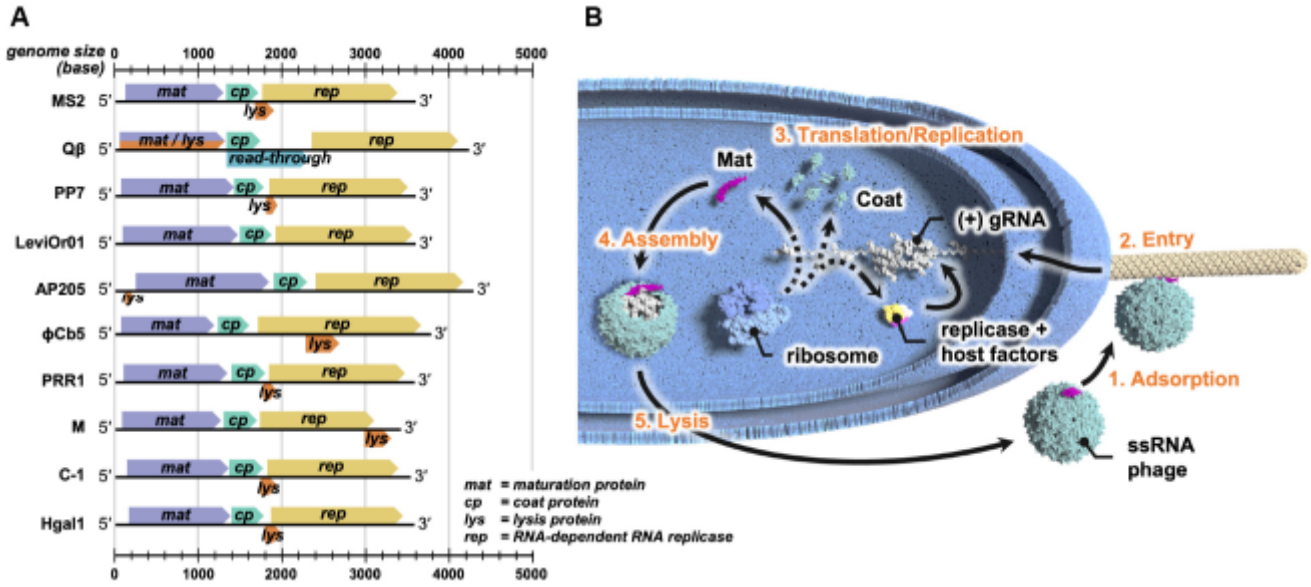


Figure 1. The genome architecture and the infection cycle of ssRNA phages. (**A**) The genome architecture of some culturable ssRNA phages. Their genomes contain three core genes: mat (encoding maturation protein), cp (encoding coat protein), and rep (encoding -subunit of the replicase). (**B**) The infection cycle of ssRNA phages. The cycle starts with (1) the adsorption of ssRNA phage to the side of the retractile pilus; (2) the retraction of the pilus promotes the entry of gRNA into the host cell; (3) the genome is translated by host ribosomes to manufacture the viral structural proteins, as well as the replicase, which with host factors, synthesizes more of the viral genome; (4) the gRNA is encapsidated by CPs and interacts with the Mat to assemble into mature virions; and (5) the Lys encoded in the gRNA of ssRNA phages induces host lysis to release new viral progeny.

Table 1. Examples of host ranges and receptors of previously known ssRNA phages [7–16].

Phages	Genus	Hosts	Receptors
MS2	Emesvirus	E. coli	Conjugative F pili
Q	Qubevirus	E. coli	Conjugative F pili
PP7	Pepevirus	P. aeruginosa	Type IV pili
LeviOr01	Pepevirus	P. aeruginosa	Type IV pili
AP205	Apeevirus	Acinetobacter spp.	Type IV pili
Cb5	Cebevirus	C. crescentus	Type IV Tad pili
PRR1	Perrunavirus	Pseudomonas, Salmonella, Vibrio, Escherichia	Conjugative P pili
M	Empivirus	Escherichia, Salmonella, Klebsiella, Proteus and Serratia	Conjugative M pili
C-1	Cunavirus	Escherichia, Salmonella, Proteus and Serratia	Conjugative C pili
Hgal1	Hagavirus	Escherichia, Citrobacter, Klebsiella, Enterobacter	Conjugative H pili

The infection cycle of ssRNA phages commences with their adsorption to the host retractile pilus by the Mat (Figure 1B). After adsorption, the genomic RNA (gRNA) of the ssRNA phage enters the host cell through an unknown mechanism. Inside the host, the positive-sense ssRNA genome acts as mRNA and is translated by host ribosomes, producing the phage-encoded proteins. The expressed Rep hijacks and assembles with host proteins, which include the ribosomal protein S1 and the elongation factors Tu and Ts (EF-Tu and EF-Ts), forming a holoenzyme that drives the synthesis of additional gRNA [17,18]. This process amplifies the translation of phage proteins and production of more gRNA. The newly synthesized phage capsid proteins (Mat and CP) and gRNA undergo assembly to form mature virions. Cell lysis is then achieved by the lysis protein to release the newly assembled ssRNA virions.

In this review article, we highlight recent studies of ssRNA phages, covering these infection steps, with a particular focus on their related structures.

2. The Known Structures of ssRNA Phage Capsids: VLPs and Mature Virions

Prior to the metagenomic era of ssRNA phage, extensive studies were conducted on the capsid structures of known ssRNA phages using both X-ray crystallography and cryo-electron microscopy (cryo-EM) (Appendix A). These structures, known as virus-like particles (VLPs), were solved as non-infectious forms, mostly through the overexpression of CPs, which form capsids with an icosahedral symmetry (triangulation number, T = 3). Later, Rumnieks J. et al. overexpressed and solved 22 VLP structures of pre-2018 unculturable ssRNA phages [19]. Some 10 out of the 22 phages had CP structures that consisted of the typical MS2 CP fold, while the remaining phages had substantial differences in the CP, especially in the N-terminal region. This work revealed deviations in wild-type CPs of ssRNA phages, suggesting classification of CP proteins into different groups. After the arrival of new metagenomic data, the CPs were eventually classified into eight distinct groups: A through H [3,6].

While these structures offer valuable insight into the fold of the CP and the symmetry of the capsids, they do not provide information regarding how the gRNA is packaged, as well as how Mat proteins are assembled in the mature virion. Not until 2016 (~55 years since the discovery of ssRNA phages) were the mature structures of E. coli ssRNA phages (coliphages) reported at high resolutions. These F-specific coliphages are MS2 [20] and Q [21].

The cryo-EM structures of MS2 have provided insights into the composition of mature ssRNA phage virions [20,22,23]. These virions consist of a capsid that encapsidates a single strand of gRNA, along with a single copy of Mat. The gRNA is folded into a defined tertiary structure inside the virion capsid. The capsids are near-icosahedral, with 89 CP dimers and the Mat disrupting the perfect T = 3 icosahedral symmetry at one of the 2-fold axes of the particle (Figure 2A).

Given the similar genome architectures and high sequence similarity among different CPs, it seemed reasonable to speculate that other ssRNA phages might possess a comparable structure to that of MS2. While the cryo-EM structure of mature Q exhibits similarities to MS2, it also varies from MS2 [21]. Surprisingly, Q contained 90 copies of CP dimers, with 89 forming the capsid and one becoming encapsidated alongside the gRNA (an internalized CP dimer) [24] (Figure 2B).

Mat is a crucial protein that plays a role in determining the maturity of ssRNA phages [25]. Despite its significant role in binding to gRNA and the host receptor, limited structural information is available for the Mat due to the challenges of purifying it; Mat proteins tend to be insoluble when overexpressed [26]. So far, the structures of Mat proteins have only been elucidated for ssRNA coliphage MS2 and Q. Due to the difficulties in the purification of Mat alone, cryo-EM has proven to be a powerful tool for exploring the structure of Mat within mature virions. The structure of Mat^{MS2} was first successfully revealed using cryo-EM [20]. Mat^{Q}, also known as A_2, was the first successfully purified

Mat in vitro through an MBP-tagged approach [27], and was later characterized using X-ray crystallography [28].

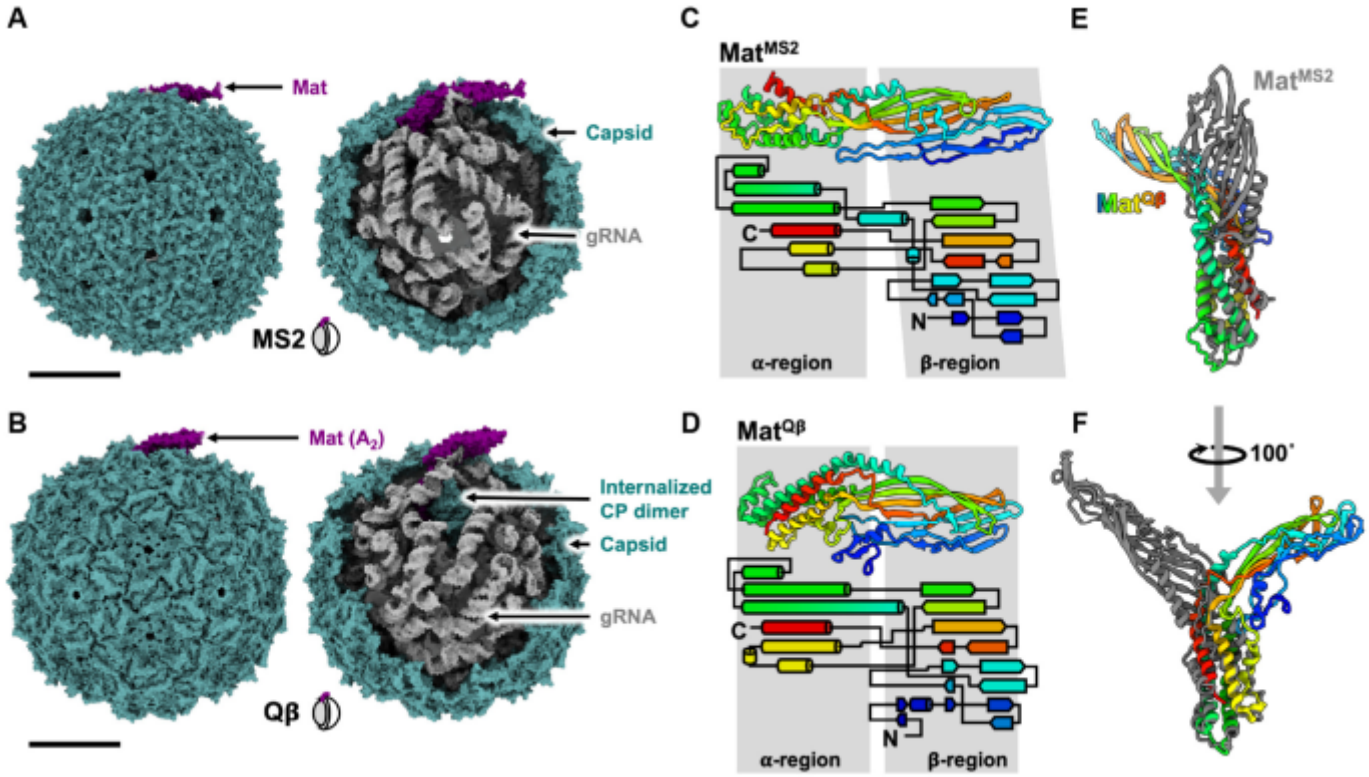


Figure 2. The known structures of F-specific ssRNA coliphage MS2 and Q. (**A**) The atomic model of MS2 (capsid combined from PDB 6NM5 and 5MSF, and RNA modeled using conformation 1 from Chang et al. [29]). The scale bar denotes 100 Å. (**B**) The atomic model of Q (from PDB 7LHD). The scale bar denotes 100 Å. (**C**) The ribbon model and secondary structural topology of Mat^{MS2}. The tip of the -sheet, colored blue and cyan, only has a backbone model in PDB 6NM5. Therefore, we used AlphaFold to predict the secondary structures and sidechains of this region (**D**) The ribbon model and secondary structural topology of Mat^{Q}. (**E**) The structures of Mat^{MS2} and Mat^{Q} aligned through -helical region. (**F**) A 100-degree rotational turn from that of panel (**E**).

Mat^{Q} (NCBI_001890) and Mat^{MS2} (NCBI_001417) share ~19% sequence identity, ~30% sequence similarity, and fold into similar tertiary structures comprised of an -helical region and a -sheet region (Figure 2C,D). However, these two Mat proteins display two major differences. First, Mat^{Q} serves an additional function, distinct from that of MS2, as a lysis protein. Second, Mat^{Q} exhibits a structural perturbation in comparison to Mat^{MS2} (Figure 2E,F): When aligning -helical regions, the -sheet regions of Mat^{Q} and Mat^{MS2} are oriented in opposite directions. It is worth noting that both MS2 and Q infect the same host through the same receptor.

Notably, the proportions of mature virions containing Mat proteins and gRNA in the purified particles vary significantly between MS2 and Q. According to the published cryo-EM structures, MS2 efficiently assembles ~85–90% viral particles into mature virions, while Q only exhibited ~25–30% mature virions, a ~3 fold lower Mat incorporation than MS2 (Table 2). This might explain the observation that Q adsorbs to the F-pilus less frequently than MS2 [30].

Table 2. Mature virion percentages of ssRNA phages from published cryo-EM data.

Phages	Total Particles	Mature Virion Particles *	Reference
MS2	22,441	18,977 (~85%)	Konning R.I., 2016 [22]
	48,276	44,897 (~93%)	Meng R., 2019 [31]
Q	51,815	12,975 (~25%)	Gorzelnik K.V., 2016 [21]
	248,445	76,843 (~31%)	Cui Z., 2017 [24]

* Mature virion particles are the particles that contain Mat proteins and have defined gRNA density.

3. The Interactions between Host Receptors and ssRNA Phages

Early studies showed that ssRNA phages readily adsorb to the side of their host pilus [32,33]. However, there is debate as to whether pilus binding alone is sufficient to trigger the gRNA release from the phage particle [34,35]. With the emergence of cryo-EM, the pioneering structural investigation of the interaction between ssRNA phages and their host receptors focused on the MS2 and F-pili pair [31,36]. The structure of the MS2/F-pilus complex shows a complete MS2 virion with intact gRNA inside, indicating the gRNA has not yet been released upon pilus binding.

In 2013, Dent et al. employed cryo-electron tomography (cryo-ET) and sub-tomogram averaging to explore the MS2/F-pilus complex, which resulted in an asymmetric reconstruction of the complex at an approximate resolution of 39 Å [36]. While this density map did not offer detailed information on residue–residue interactions, it revealed additional density, which may correspond to the Mat protein, at the binding interface. This observation provided supporting evidence that the Mat disrupts the perfect icosahedral T = 3 capsid symmetry and is crucial for F-pilus binding.

Six years later (2019), Meng et al. employed single-particle cryo-EM to unravel the structural basis of MS2^{0}s interaction with its receptor, the F-pilus, at a resolution of ~5–7 Å [31]. This study demonstrated that the Mat^{MS2} utilizes its -sheet region to interact with four F-pilin subunits spanning two adjacent helical turns (Figure 3A,B). The inter-face area of the interaction was approximately 1063 Å^2 (Figure 3B). The binding between Mat^{MS2} and the pilins involved electrostatic and hydrophobic interactions (Figure 3C,E). The first mode of electrostatic interaction occurred between Mat residue R99 and pilin chain A residue D23 (Figure 3C). The second mode involved Mat residue R36 and pilin chain C residue D7 (Figure 3D). Additionally, the hydrophobic residues F94 and F92 of Mat interacted with the N-terminus of pilin chain D (Figure 3E). The binding of MS2 to the F-pilus occurred in a specific orientation, with Mat pointing its tip away from the cell surface. This orientation is proposed to facilitate gRNA entry upon retraction, especially when brought closer to the basal body of the Type IV secretion system [31].

Unexpectedly, in 2020, Harb et al. demonstrated that both MS2 and Q trigger detachment of F-pili during entry, as fragments of pili appeared in the media, and this occurs in a phage concentration-dependent manner [37]. Even in the absence of the Type IV coupling protein, TraD (a protein which is part of the bacterial Type IV secretion system and required for conjugation), the mutant T4SS F-pilus was still able to be detached by MS2. However, the gRNA of MS2 failed to enter the host cell's cytoplasm [37]. Notably, TraD has been shown to be necessary for the MS2-related phages R17 and f2, but not Q [38]. The exact discrepancy in the requirement of TraD for gRNA entry and infection between MS2 and Q remains elusive. Although it is unknown how the gRNA of ssRNA phages enters the host cell, these results have emphasized the requirement of pilus retraction. Based on Harb et al.'s results, pilus detachment is necessary, and gRNA penetration into the host cell possibly starts once the phages have been brought close to the cell surface via retraction of the pilus assembly machinery. This mechanism seems to be conserved among ssRNA phages, since PP7 has also been shown to result in a 50% reduction in length of the type IV pilus during infection [32]. It is not clear what the exact mechanism of ssRNA phage-induced pilus detachment is. It may be that the first ssRNA phage entering the cell

causes the pilus to detach, leaving the other pilus-adsorbed phages unable to infect the cell. This process resembles superinfection exclusion in other phages.

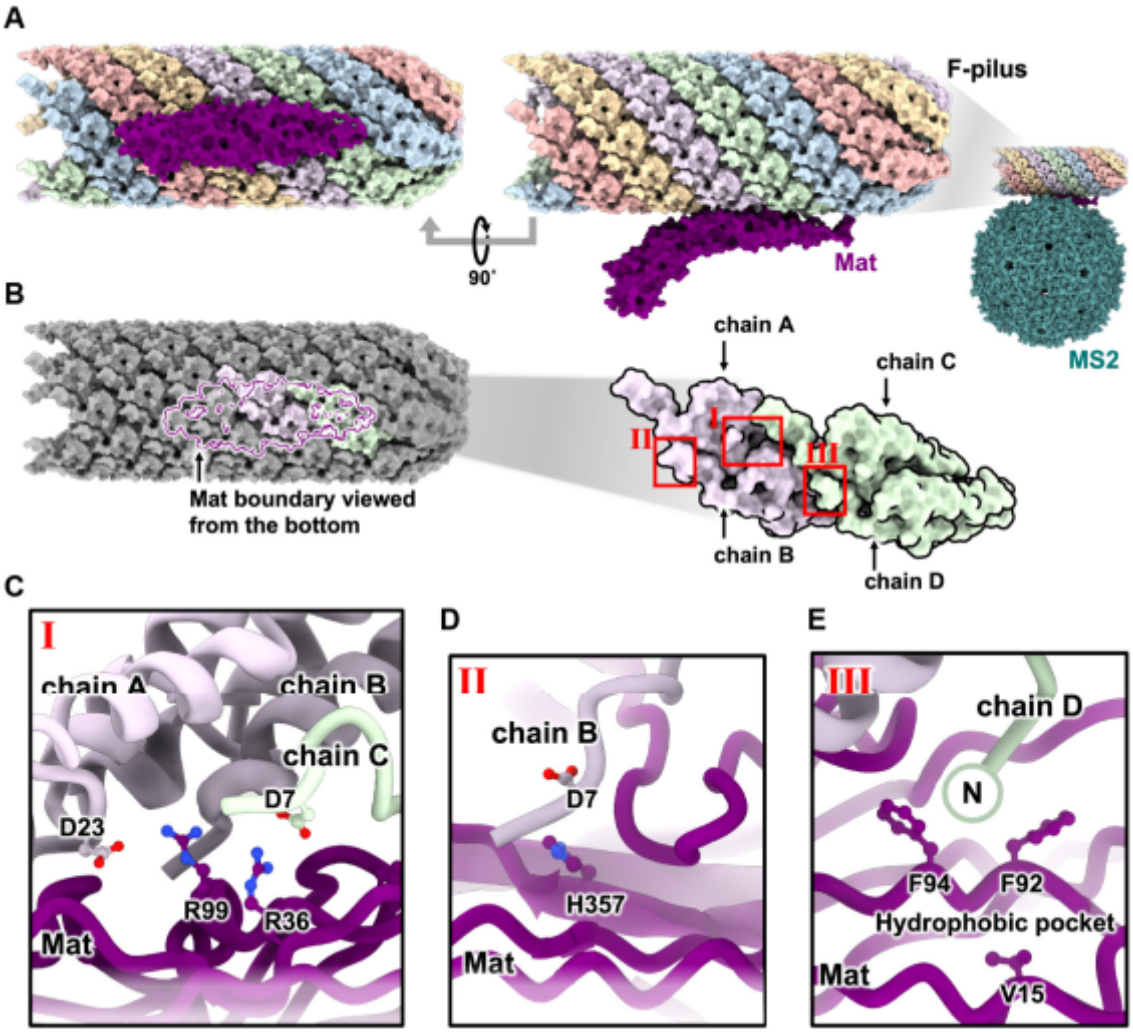


Figure 3. The interaction of MS2 and the F-pilus [31]. (**A**) The interaction of Mat^{MS2} (purple) and F-pilus, represented with five different colors for each helical strand. (**B**) Zoom-in view of panel A focusing on the binding site where Mat^{MS2}, shown as a purple boundary, is interacting with four pilin subunits. (**C**–**E**) Zoom-in views from panel B, denoted by roman numerals from I to III, rotated 90 to illustrate the reported contacts between Mat^{MS2} and pilin subunits.

4. The Genome Replication of ssRNA Phages

The ssRNA phage rep gene encodes the -subunit of the replicase (Rep), an RNA-dependent RNA polymerase, which contains the catalytic site responsible for replication. The replicase holoenzyme consists of the virally encoded -subunit as well as three essential host translation proteins. They are the elongation factor thermal stable (EF-Ts), the elongation factor thermal unstable (EF-Tu), and the S1 protein of the ribosome (Figure 4A) [39]. These proteins play crucial roles in efficient translation during normal cellular function. EF-Tu is a G-protein that functions in translocating amino-acylated tRNAs, EF-Ts facilitates the exchange of GDP for GTP for EF-Tu, and the S1 protein promotes translation initiation by stabilizing the mRNA [40,41].

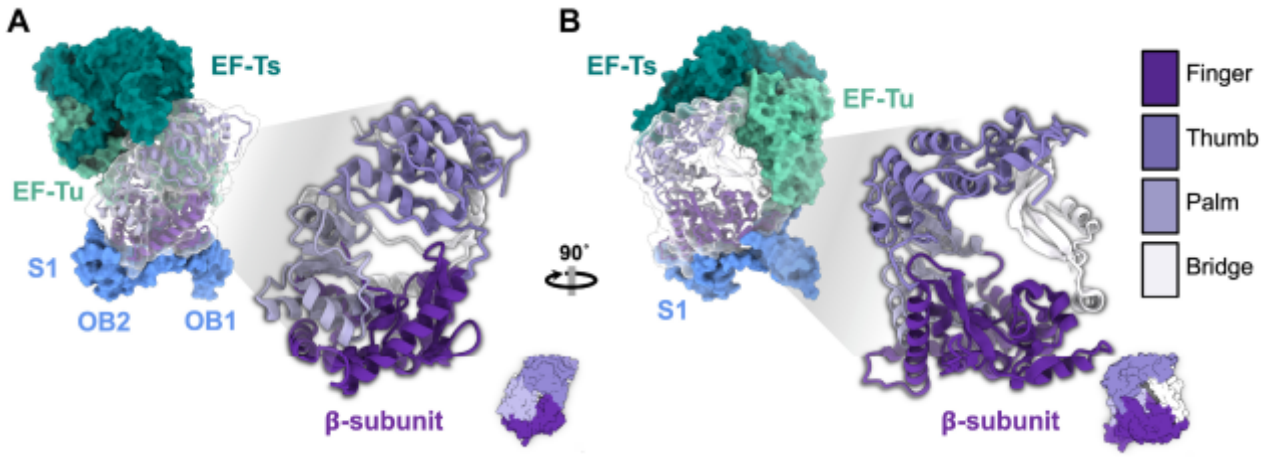


Figure 4. The Rep^{Q} holoenzyme complex (PDB 4R71). (**A**) The -subunit of the replicase (violet) forms a holoenzyme complex by hijacking the host factors: EF-Ts (dark green), EF-Tu (light green), and the ribosomal protein: S1 (blue). S1 is only shown with OB1 and OB2 domains, which interact with the -subunit. (**B**) The 90 rotation of the model from panel A. The ribbon model of -subunit which contains a four-domain architecture: finger, thumb, palm, and bridge (illustrated using different shades of purple/violet and white).

-subunit domain architecture: The structural and enzymatic investigations of ss-RNA replicase have primarily focused on the canonical Rep^{Q} due to its stability. Rep^{Q} has been shown to be expressed and purified successfully by fusing EF-Tu, EF-Ts, and Rep into a single chain connected by flexible linkers [42]. Structural studies reveal that Rep^{Q} contains three distinct domains: palm, thumb, and fingers. Additionally, Rep^{Q} possesses a bridge region that connects the finger and thumb domains and is important in complex formation (Figure 4B) [43,44]. The palm domain of Rep^{Q} contains the catalytic site, which coordinates two divalent cationic metal ions [45], typically calcium or magnesium, via three aspartic acid carboxylates for RNA polymerization. The two metal ions help coordinate the growing RNA strand and the next nucleotide.

Interactions with host proteins in the holoenzyme: Within the replication complex, it has been determined that the OB1 and OB2 domains of the S1 protein interact with the finger domain (Figure 4A), but the OB3-OB6 domains are flexible and cannot be resolved [46]. In the Q holoenzyme, EF-Tu domain 2 has been identified as interacting with the finger domain of the -subunit, while EF-Tu domain 3 and EF-Ts interact with the thumb domain of the -subunit [43,44]. It is notable that the EF-Tu and EF-Ts proteins are known to form a complex that has been observed outside of the Q holoenzyme [40]. These observations indicate the complex nature of the interactions within the replication complex, highlighting the intricate interplay between the various components involved in RNA polymerization.

The HMM search on ssRNA phage sequences (~15,000 sequences) from the 2020 literature identified two distinct orders of ssRNA phage Rep proteins based on 70% pairwise amino acid identity [3]. In 2022, a larger collection of ssRNA sequences (~65,000 sequences) also showed consistency with two major lineages as seen in the phylogeny generated (Figure 5A) [4]. Although the catalytic core of Rep is conserved, the host–factor interacting region displays a relative diversity (Figure 5B), suggesting that they have evolved to bind to different host factors from their respective hosts.

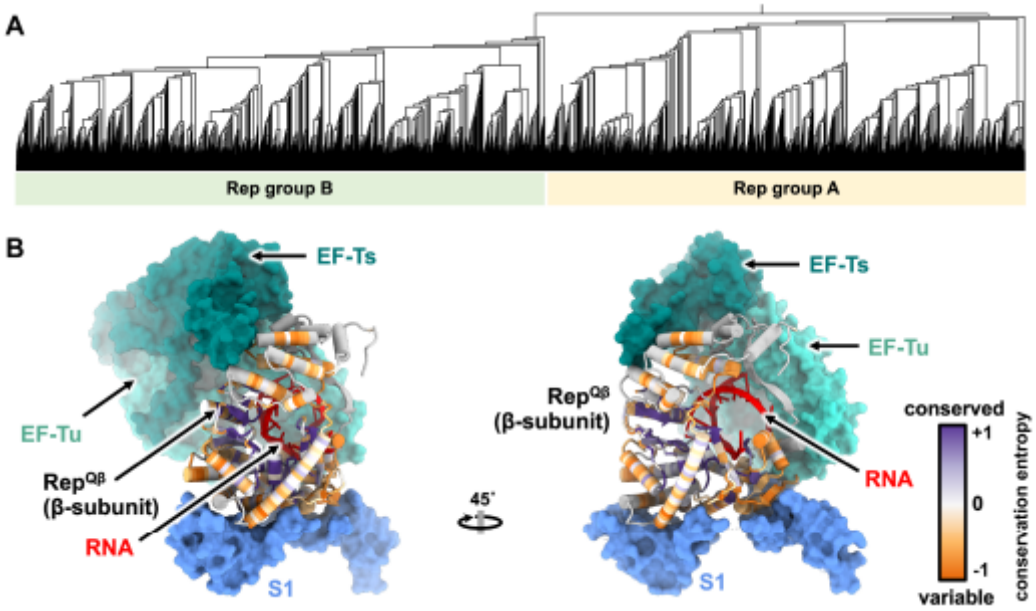


Figure 5. The phylogenetic tree and conservation among ssRNA phage replicases. (**A**) The phylogenetic tree of 15,452 ssRNA phage sequences with 90% average nucleotide identity from all available ssRNA phage sequences shows two distinct clades of ssRNA phages based on two distinct Rep groups [3]. The phylogenetic tree generated is based on the supplementary data in [4]. (**B**) The holoenzyme replicase complex of Q (PDB 4R71), bound to RNA (PDB 3BSN). The -subunit of the replicase is colored by conservation from MUSCLE alignment of 149 sequences from pre-2018 ssRNA phages. The conservation plot was created using ChimeraX software based on conservation entropy value (AL2CO) calculated by ChimeraX [47,48].

Initiation: Replication is initiated at the 3^0UTR of the positive-sense genome. It has been described that in Q the OB3 domain of the S1 protein is able to recognize and bind to two internal regions of RNA in the Q genome, 1247–1346, known as the S-site, and 2545–2867, known as the M-site [49]. This allows the complex to bind the RNA and position the CCA-3^0 sequence in the active site. While A is the terminal residue, replication starts at the penultimate C. When the terminal residue is mutated from A to G, U, or deleted, there is a significant reduction in replication [50]. From these studies, the idea emerged that the non-template A functions in stabilizing the initiation complex through Pi-Pi interactions with the GTP and the penultimate C [45].

Elongation: The bases enter through the NTP channel and hydrogen bond with the current nucleotide on the template strand, forming an RNA duplex. As each base is added, the duplex is driven towards the EF-Tu subunit of the complex. Once the ninth nucleotide is added to the growing strand, the non-template 3^0 adenosine Pi stacks onto a C-terminal asparagine of the -subunit [45]. This wedge region, formed by the C-terminus and EF-Tu, begins to destabilize the duplex and is the basis for its separation. Upon addition of the tenth nucleotide, the non-template A "flips" around the wedge through its interactions with domains 2 and 3 of EF-Tu and begins the separation of the template and growing RNA strands. At nucleotide 14, the template RNA begins to leave the replication complex through the exit tunnel formed by the -subunit-EF-Tu interface.

Termination: Once replication reaches the terminal 5^0 end of the genome, termination commences. The final template base, C, is added to the growing negative sense RNA, and then this complex shifts, allowing for the NTP binding site to open. It has been shown that this binding pocket is too large for CTP and UTP to bind in and too small for GTP, but it is the correct size for ATP, which is preferentially incorporated in the 3^0 growing strand [51]. In addition to the binding pocket size, the adenine base can Pi stack onto the 5^0 template strand G stabilizing the RNA duplex. This is the proposed mechanism of non-template A addition. Furthermore, it is thought that the S1 protein aids in termination of replication

through the binding of the growing RNA via the OB3 domain [52]. This prevents annealing of the template and growing strands upon release, but it has not currently been shown where the S1 protein binds on the negative sense RNA [53].

5. The Genome Packaging and Viral Assembly of ssRNA Phages

The genome packaging of ssRNA phages differs from dsDNA and dsRNA phages, in which ssRNA phages lack motors/effector proteins that help package the genome into the capsid [54,55]. ssRNA phages do not have a highly pressurized capsid. Instead, the gRNA of ssRNA phages is known to form a high level of secondary structures comprised of many RNA stem-loops, upon which the Mat and CPs will assemble to form a mature virion (Figure 6) [29,56]. The very first RNA stem-loop discovered was shown to bind with high affinity to CP dimers and was called "an operator". This is the RNA stem-loop that contains the start codon of the rep gene and is shown to play a role in suppressing the expression of rep. The operator stem-loop also acts as a strong encapsidation site for the coat protein to bind. These operators are shown to be conserved among known ssRNA phages [57–59].

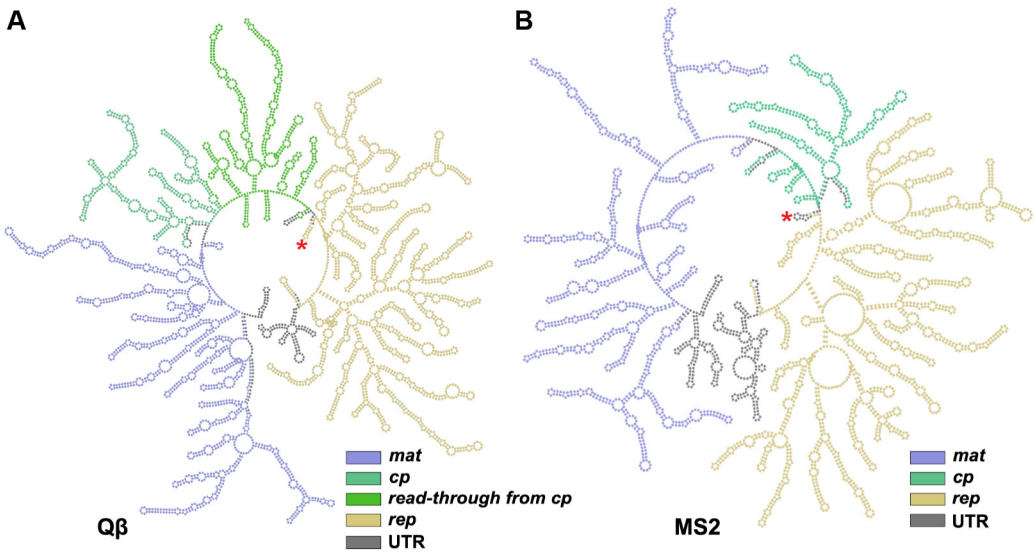


Figure 6. The secondary structures of gRNA from (**A**) Q [56] and (**B**) MS2 [20,29]. The genomes were colored based on open-reading frames encoded. The lys^{MS2} gene was not annotated because it overlaps both cp and rep. The lys^{Q} is its own mat. The untranslated regions (UTR) are colored grey. The operator, indicated by red asterisk, folds into the RNA stem containing 3^0 end of the preceding UTR and 5^0 end of the rep gene.

The operator-like stem-loops are found throughout the genome of ssRNA phages and are believed to act as "packaging signals" in ssRNA viruses [60]. Chang et al. 2022 modeled the entire gRNA of Q and identified a total of 77 RNA stem-loops (Figure 6A). Out of these, 59 stem-loops were identified to interact with the capsid within 5 Å distance (Figure 7A–E) [56]. Among these 59 RNA stem-loops in Q, 32 were identified as "operator-like" RNA stem-loops. These 33 stem-loops (32 operator-like and 1 actual translational operator, number 34) interact with the coat protein dimer in the same manner based on their model, and were proposed to be directly involved in the viral assembly (Figure 7D,E, red numbers). The remaining 26 RNA stem-loops were identified as non-operator-like stem-loops (Figure 7E, black numbers). Stem-loops 57 and 59 are those that interact with the internalized CP dimer and Mat^{Q}, respectively (Figure 7D,E). In contrast, in the high-resolution cryo-EM structure of MS2 solved by Dai et al. 2017 (Figure 7F–J) [20], 14 RNA stem-loops, out of 71 in the MS2 gRNA, interacted with the mature virion capsid and were resolved to high resolutions, thanks to their high affinity for the CP shell (Figure 7I, black-outline circles). Chang et al. later on built a complete model of the MS2 gRNA [29],

which allows us to identify five additional RNA stem-loops shown to be "operator-like" (Figure 7I, dotted-outline circles). This increases the total to 40 RNA stem-loops defined as stem-loops within less than 5 Å to the inner surface of the MS2 capsid (Figure 7I,J). Among those, 19 stem-loops were identified as operator-like stem-loops (Figure 7J, red numbers), while 21 stem-loops were identified as non-operator-like stem-loops (Figure 7J, black numbers). The actual MS2 translational operator is stem-loop 18, while stem-loop 40 interacts with Mat^{MS2} (Figure 7I,J). These results revealed the gRNA packing preference in mature virions. With the Mat defined as the "north pole", in both MS2 and Q, the 5^0 end of the genome resides towards the southern hemisphere of the capsid, while the 3^0 end resides towards the northern hemisphere (Figure 7D,I). Notably, there is a crack in the Q mature capsid around where the Mat points [21], which suggests that this region of the capsid is the last to assemble around the RNA.

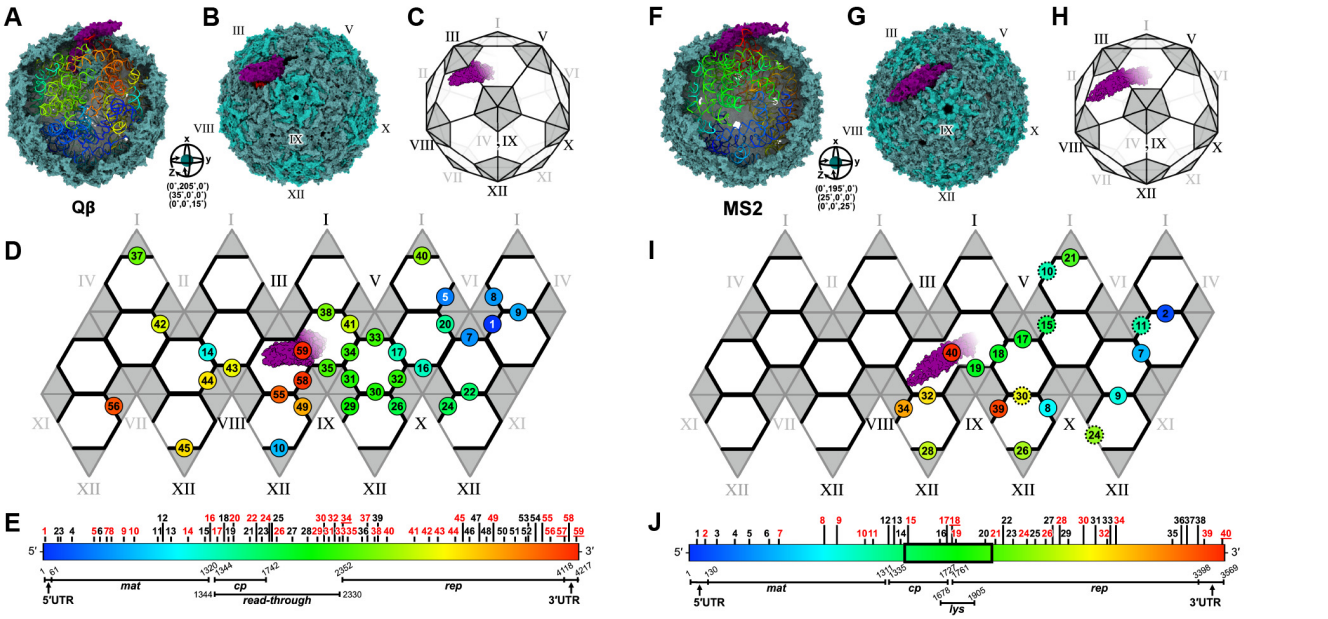
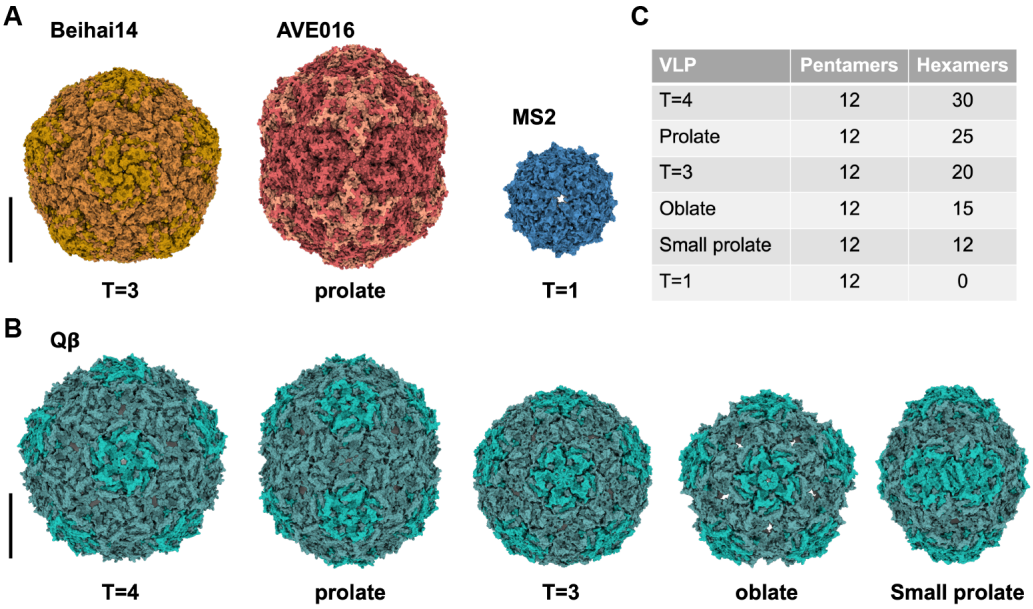


Figure 7. The genome packaging and CP/RNA stem–loop interactions of Q (**A**–**E**) and MS2 (**F**–**J**). (**A**) The atomic model of Q (PDB 7LHD) with the capsid (green) partially removed to reveal gRNA (rainbow). (**B**) The model from panel A illustrated at a different rotational angle around the XYZ axes, as indicated in the inset. (**C**) A cartoon representation of the model in panel (**B**), showing the pentameric faces (gray) and hexametric faces (white) of the Q capsid. Each pentameric face was labelled with roman numerals from I to XII to establish their locations on the capsid. The roman numerals, shown in gray, represent the faces that point inwards to the paper, while the ones shown in black represent those that point outwards. (**D**) The planar representation of the icosahedral surface from panel (**C**) shows the CP dimers (black line) that interact with "operator-like" RNA stem-loops. The RNA stem-loops are shown with numbers corresponding to those labelled in panel (**E**) [56]. (**E**) The gRNA sequence of Q shown in rainbow to match those in panels (**A**,**D**). The red numbers indicate operator-like RNA stem-loops plotted in panel (**D**), while the black numbers are non-operator-like RNA stem-loops. (**F**) The atomic model of MS2 (combined PDB from 6NM5, 5MSF and RNA modeled using conformation 1 from Chang et al. [29]) represented as in panel A. (**G**) The atomic model from panel (**F**) rotated around the XYZ axes, as indicated in the inset. (**H**) The cartoon representation of panel (**F**), as shown in panel (**C**). (**I**) The planar representation of the icosahedral surface from panel (**H**) showing the MS2 gRNA stem-loops interacting with capsid, represented in the same way as in panel (**D**). The solid black-outline circles are the stem-loops identified by Dai et al. [20], while the black dotted-outline circles were the additional stem-loops identified in gRNA Conformation 1 by Chang et al. [29]. (**J**) The gRNA sequence of MS2 is shown in a rainbow to match those in panel (**F**,**I**). The region marked with black box is the region reported to be a flexible region by Dai et al. [20].

Besides the near-icosahedral T = 3 mature virion and perfect icosahedral T = 3 VLPs, it has been previously shown that ssRNA phage capsid proteins are capable of assembling into non-canonical T = 4 [19,61–63], T = 1 [64], elongated T = 3 Q = 4 (prolate) [19], unusually large T = 3 [19], and even tubular capsids [65] when overexpressed or mutated (Figure 8A and Appendix A; Table A1). Chang et al. demonstrated that Q was able to assemble into small populations of non-canonical capsid forms through wild-type infection [56]. The various non-canonical forms of Q reported include T = 4, prolate, oblate and small prolate capsids (Figure 8B). All these forms of Q contain 12 pentamers of CPs, but the number of hexamers changes in correlation with the size of capsid (Figure 8C). This suggests that the manipulation of the CP hexameric units incorporated into the particles might influence the size of Q capsid. In conjunction with this, the size of RNA might also influence the capsid assembly. Indeed, Chang et al. 2022 showed that when only cp/read-through were overexpressed, the population of oblate and small prolate particles increased [56].



VLP	Pentamers	Hexamers
T=4	12	30
Prolate	12	25
T=3	12	20
Oblate	12	15
Small prolate	12	12
T=1	12	0

Figure 8. Examples of different forms of VLPs observed in ssRNA phages. (**A**) The VLPs observed from overexpression and/or mutagenesis of CPs. The Beihai14 (PDB 6YFD) and AVE016 (PDB 6YFB) are non-culturable ssRNA phages with their CPs recombinantly expressed and forming a large T = 3 and prolate particles, respectively. The S37P mutation of CP^{MS2} resulted in T = 1 particle (PDB 4ZOR) (**B**) The VLPs of Q, which were observed through wild-type infection (PDB 5VLY, 7LGE, 7LGF, 7LGG, and 7LGH). (**C**) The table illustrates the number of pentamers and hexamers for each VLP form. Scale bars in this figure denote 100 Å.

The current proposed model of gRNA packaging and assembly in ssRNA phages may be referred to as co-replicational assembly [56]. This model describes that the packaging starts during the replication of negative-sense to positive-sense RNA. The 5^0 end of the nascent positive-sense RNA folds into its secondary structure, exposing the operator-like RNA stem loops which act as a packaging signal, providing a nucleation site for CP dimers (referred to as C/C dimer). Upon binding to these packaging signals, the CP dimers change their conformation from C/C to A/B, which promotes pentamer formation. As replication progresses, CP recruitment continues, leading to RNA collapse. This forms the intermediate state, leading to the folding of tertiary structures of the RNA. Upon replication termination, the 3^0 end of the RNA becomes exposed and recruits the Mat, resulting in a mature virion.

6. The Lysis of the Host by ssRNA Phages

Double-stranded DNA (dsDNA) phages usually lyse their host cells through multi-gene lysis systems, which are typically composed of endolysins and holins. Endolysins

are muralytic enzymes that target the peptidoglycan cell wall synthesis pathway, while holins are cytoplasmic membrane-spanning proteins that create "holes" on the membrane, allowing endolysins to access and degrade the peptidoglycan cell wall [66]. Unlike dsDNA phages, ssRNA phage host lysis is induced through a single-gene Lys [67]. Lys is encoded by the lys gene, and, unlike their other structural and functional genes, is widely distributed throughout the genome of ssRNA phages. Typically, it is embedded out of frame in other genes (Figure 1A).

The Lys proteins of known ssRNA phages have been classified into two groups: (1) non-peptidoglycan (PG)-targeting and (2) PG-targeting Lys proteins (Figure 9A). The mechanism of non-PG-targeting Lys remains inscrutable. Unlike the latter group, as the name suggests, they do not target enzymes related to peptidoglycan biosynthesis. The canonical Lys^{MS2} is shown to have four domains [68]. Through comparative analysis of sequences, several candidates were identified as non-PG-targeting [68]. Interestingly, one of the culturable ssRNA phages targeting P. aeruginosa, LeviOr01, does not have a Lys. Although the lys candidate has been proposed, the start codon found in the annotation is an arbitrary 5^0-TTA-3^0 (which codes for leucine). Thus, it is likely that the annotation is not actually a Lys. LeviOr01 was shown to form clear plaques on P. aeruginosa PcyII-10. The sub-population of PcyII-10, after being passed for multiple rounds, can still produce LeviOr01. Hence, it is proposed by the authors that this ssRNA phage is capable of inducing a carrier state in PcyII-10 [10].

A

Phages	Lys target	Size
MS2	non-PG target	75
Qβ	MurA	420
PP7	MurJ	55
LeviOr01	Unidentified	Unidentified
AP205	non-PG target	35
ϕCb5	non-PG target	136
PRR1	non-PG target	54
M	MurJ	37
C-1	non-PG target	65
Hgal1	non-PG target	65

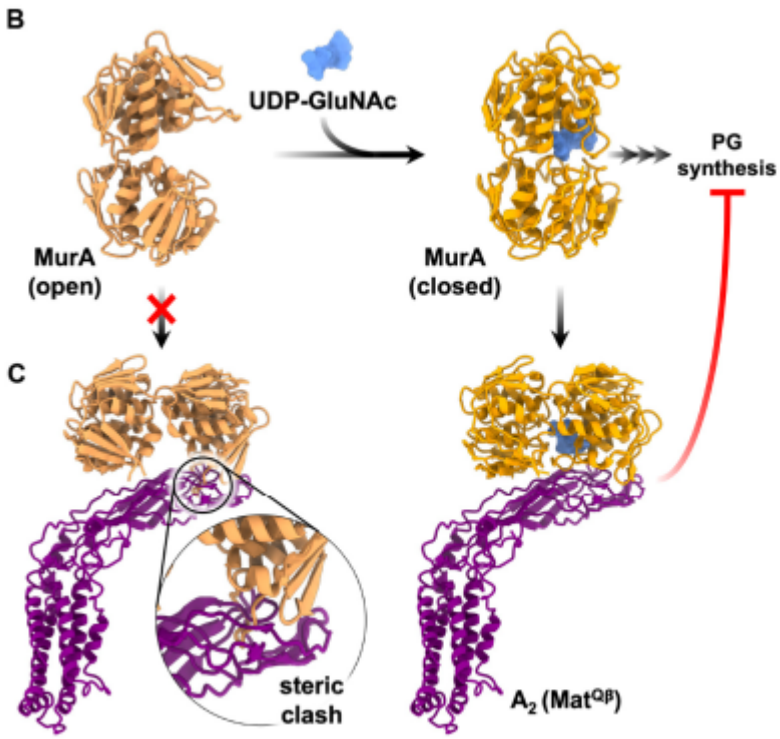


Figure 9. Cell lysis induced by Lys of ssRNA phages with the structural mechanism only revealed for Q. (**A**) Single lysis proteins from previously known ssRNA phages, with the name, target and size in amino acid length listed. (**B**) The MurA in its open conformation (PDB 1EJD) without its substrate, UDP-GluNAc. Upon the substrate binding, MurA changes to a closed conformation (PDB 3VCY). The released product will be further used by downstream enzymes for peptidoglycan biogenesis. (**C**) The open conformation of MurA was unable to bind to A_2 (PDB 5VM7) due to the steric clashes between the loops from MurA and the tip of Mat -region. In the closed conformation, the MurA/UDP-GluNAc complex can bind to A_2, inhibiting MurA function.

The second group of Lys proteins is the PG-targeting Lys, which targets enzymes involved in PG synthesis (Figure 9A). These PG-targeting Lys include Q, M, and PP7 [67]. Nonetheless, the structural information for ssRNA Lys proteins is only limited to Lys^{Q}, also known as A_2. A_2, which also functions as Mat^{Q}, promotes lysis by inhibiting MurA (UDP-N-acetylglucosamine enolpyruvyl transferase). This enzyme catalyzes the first committed step in the biosynthesis of PG. The binding of MurA to its substrate, UDP-N-acetylglucosamine (UDP-GlcNAc), causes a conformational change of MurA from an open to a closed conformation. The closed conformation of MurA allows A_2 to bind to the UDP-GlcNAc-MurA complex, thus inhibiting MurA's function (Figure 9B,C) [24,69]. Although it is lacking structural information, Lys^{M} and the recently identified Lys^{PP7} are shown to inhibit MurJ, an enzyme with lipid II flippase activity [70]. One of the key residues in E. coli MurJ that confers function of Lys^{M} and Lys^{PP7} is Q244. Studies have shown that the mutation of this residue to proline (Q224P) causes resistance to both Lys^{M} and Lys^{PP7}. This residue is located on transmembrane domain (TMD) 7, one of the 14 TMDs of MurJ [70,71]. However, it is noteworthy that Lys^{M} and Lys^{PP7} are different in structure and sequence. In fact, sequence analysis of Lys^{PP7} suggests that it might be a non-PG-targeting Lys [68,71].

7. Conclusions and Future Directions

Even with the many years of research on ssRNA phages, current knowledge of the structures and biology of ssRNA phages still revolves around ssRNA coliphages (Figure 2), except for the structures of the CPs (Appendix A). However, recent advancements in metagenomes and metatranscriptomes [2–5] have unveiled a myriad of ssRNA phages that exceeded previous expectations. The mature virions of the ssRNA coliphage, MS2 and Q, were shown to contain a single copy of Mat and a capsid that was composed of 178 copies of CPs encapsidating gRNA. Despite infecting the same host and utilizing the same receptor as MS2, Q notably exhibits structural variations, containing an internalized CP dimer in its mature virion [22]. Currently, with around ten ssRNA phages that can be cultured, their diversity and variation have already been demonstrated. It is unknown if the mature virions/morphology observed in ssRNA coliphages will be valid for all culturable ssRNA phages, such as PP7, LeviOr01, Cb5, and AP205, since they have also evolved to target different types of retractile pili. The mature structures of these phages are worth exploring to understand the biology of ssRNA phages, especially the Mat, which is challenging to purify. Structural determination using cryo-EM is a promising strategy to unravel the structures of these ssRNA mature virions.

With an increase in the number of ssRNA phage sequences, it is possible that the host range and retractile pilus receptors that these ssRNA phages target could be more diverse. Exploring the interaction of Mat/pilus unveils information necessary to understand the adsorption of these phages. With enough structural information, motif searches into metagenomic data could allow us to identify new receptors or new hosts for new ssRNA phages, particularly facilitated by the recent breakthrough in protein structure prediction tools such as AlphaFold2, RoseTTAFold, I-TASSER, and ESMFold [72–75]. It is worth noting that the entry process of these ssRNA phages remains unknown, particularly regarding the involvement of the Mat in coordinating gRNA delivery into the cell. Although it was shown in MS2 that pilus detachment is required for gRNA entry [37], it is unclear how gRNA translocates through the cell membrane/peptidoglycan into the host cell. This phenomenon could potentially be unraveled through cryo-electron tomography.

Although there are complete atomic models of gRNA for coliphage MS2 and Q and different forms of VLPs (Figures 2 and 8, and Appendix A), which suggests a potential genome packaging pathway for ssRNA phages, the proposed model has yet to be tested or visualized in real time in situ. Using combinations of time-resolved cryo-EM [76] and focus-ion beam (FIB)/cryo-ET [77] could potentially unravel the gRNA-packaging mechanism both in vitro and in situ, and observe any phage assembly intermediates during gRNA packing.

The model for gRNA replication by Rep for ssRNA phages is mainly derived from coliphage Q. It is reasonable to suggest that the gRNA replication carried out by replicase might function differently for different host targets. One explanation would be that the host factors involved in the process might differ for ssRNA phages targeting different hosts. It has been shown in Q that the host factor for Q replication (Hfq) modulates Rep^{Q} activity for negative-sense gRNA synthesis [78,79]. Recently, additional host factors for Rep^{MS2} have also been identified, including initiation factor-1 (IF-1) and IF-3. While IF-1 was shown to promote replication, IF-3 was shown to inhibit it [80]. The structural interactions of these additional host factors with the holoenzyme replicase/RNA complex are yet to be explored.

Lysis is the terminal step of the ssRNA infectious cycle taken to release new viral particles into the environment in order to infect a new host and repeat the cycle. Structural information on ssRNA Lys proteins is only limited to A_2 of Q, a PG-targeting Lys. Even more mysterious is the Lys of the non-PG-targeting group. There currently remains a lack of structural information and knowledge of the mechanisms of this group.

In summary, there are still unexplored structural ventures that could potentially help in understanding and capturing the life cycle of ssRNA phages in situ, especially for non-E. coli ssRNA phages. Additionally, it is important to explore metagenome sequences to identify hosts for non-culturable ssRNA phages, which could potentially become tangible with more structural information.

Author Contributions: J.T., Z.L. and J.Z. wrote the first draft of the review manuscript. Z.H. performed relevant data analysis. All authors contributed to the revision of this article and approved the submitted version. All authors have read and agreed to the published version of the manuscript.

Funding: The support was received from TAMU T3 and X-grants, the Welch Foundation grant A-1863, NSF grant MCB-1902392, and NIH grants R21AI156846, R01GM141659, as well as funding from the Center for Phage Technology, jointly sponsored by Texas AgriLife and Texas A&M University.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable; No new data were generated.

Conflicts of Interest: The authors declare no conflict of interest. J.Z. is a consultant of Lynntech Inc., College station, Texas. The funders had no role in the writing of the manuscript.

Appendix A

Table A1. List of PDB identifiers associated with ssRNA phages' structures. This list excludes PDB related to replicase structures.

Phages	PDB#	Method	Brief Description
Culturable ssRNA phages			
MS2	2MS2	X-ray	T = 3 VLP
	1BMS	X-ray	T = 3 VLP (P78N)
	1MST	X-ray	T = 3 VLP (E76D)
	1ZDI	X-ray	CP/RNA complex, WT RNA
	1ZDH	X-ray	CP/RNA complex, C-5 variant RNA (U ! C mutant)
	1AQ3	X-ray	CP(T59S)/RNA complex
	1AQ4	X-ray	CP(T45A)/RNA complex
	1MVA	X-ray	T = 3 VLP, CP(T45A)
	1MVB	X-ray	T = 3 VLP, CP(T59S)
	1ZDJ	X-ray	CP/RNA complex, loop RNA (5^0-GGAUCACC-3^0), shorter RNA operator derived from WT RNA operator
	1ZDK	X-ray	CP/RNA complex, clamp RNA, elongated stem of C-5 variant RNA operator

Table A1. Cont.

Phages	PDB#	Method	Brief Description
	6MSF	X-ray	CP/RNA complex, F6 RNA, derived from WT and C-5 version
	5MSF	X-ray	CP/RNA complex, F5 RNA, derived from WT and C-5 version
	7MSF	X-ray	CP/RNA complex, F7 RNA, derived from WT and C-5 version
	1U1Y	X-ray	CP/RNA complex, F5 RNA with 2-aminopurine at -10 position
	2BU1	X-ray	CP/RNA complex, RNA operator with 5-bromouracil at -5 position
	2C4Q	X-ray	CP/RNA complex, RNA operator with pyridin-2-one at -5 position
	2C4Y	X-ray	CP/RNA complex, RNA operator with 2-thiouracil at -5 position
	2C4Z	X-ray	CP/RNA complex, RNA operator with 2-thiouracil at -5 and -6 position
	2C50	X-ray	CP/RNA complex, RNA operator with adenine at -5 position
	2C51	X-ray	CP/RNA complex, RNA operator with guanine at -5 position
	2BNY	X-ray	CP(N87A)/ MS2 RNA complex
	2BQ5	X-ray	CP(N87A/E89K)/MS2 RNA complex
	2BS0	X-ray	CP(N87A/E89K)/variant Q RNA complex
	2BS1	X-ray	CP(N87A/E89K)/Q RNA complex
	1ZSE	X-ray	CP(N87S)/Q RNA complex
	2B2D	X-ray	CP(N87S/E89K)/Q RNA complex
	2B2E	X-ray	CP(N87S/E89K)/MS2 RNA complex
	2B2G	X-ray	CP(N87S)/MS2 RNA complex
	2IZ8	X-ray	CP/RNA complex, C-5 RNA operator with cytosine at -7 position
	2IZ9	X-ray	CP/RNA complex, RNA operator with pyridin-4-one at -5 position
	2IZN	X-ray	CP/RNA complex, RNA operator with guanine at -10 position, and GC pairs above and below -10 position changed to CG
	2IZM	X-ray	CP/RNA complex, C-5 RNA operator with cytosine at -10 position
	2VTU	X-ray	Octahedral VLP (artifact) from covalent CP dimer fusion
	4ZOR	X-ray	T = 1 VLP (S37P)
	6RRS	Cryo-EM	T = 3 VLP
	6RRT	Cryo-EM	T = 4 VLP
Q	1QBE	X-ray	T = 3 VLP
	5KIP	Cryo-EM	T = 3 VLP
	5VLY	Cryo-EM	T = 3 VLP
	5VLZ	Cryo-EM	Mature virion, capsid and Mat (A2)
	5VM7	Cryo-EM	MurA-A2 complex
	7LGE	Cryo-EM	T = 4 VLP
	7LGF	Cryo-EM	Prolate VLP
	7LGG	Cryo-EM	Oblate VLP
	7LGH	Cryo-EM	Small prolate VLP
	7LHD	Cryo-EM	Complete atomic model of QT
	7TJD	X-ray	= 1 VLP
PP7	6N4V	Cryo-EM	T = 4 VLP
	2QUX	X-ray	CP/RNA complex
	2QUD	X-ray	CP dimer
	1DWN	X-ray	T = 3 VLP
AP205	5JZR	NMR	CP dimer
	5FS4	X-ray	CP dimer
	5LQP	Cryo-EM	T = 3 VLP
Cb5	2W4Z	X-ray	T = 3 VLP, from mature virions
	2W4Y	X-ray	T = 3 VLP, from overexpression of CP
PRR1	2VF9	X-ray	T = 3 VLP
	4ANG	X-ray	T = 3 VLP with RNA operator
fr	1FRS	X-ray	T = 3 VLP, WT
	1FR5	X-ray	T = 3 VLP, 4 residue deletion in FG loop
GA	1GAV	X-ray	T = 3 VLP

Table A1. Cont.

Phages	PDB#	Method	Brief Description
Nonculturable, pre-2018, ssRNA phages			
AVE002	6YF9	X-ray	T = 3 VLP
AVE015	6YFA	X-ray	T = 3 VLP
AVE016	6YFB	X-ray	Prolate VLP
AVE019	6YFC	X-ray	T = 3 VLP
Beihai14	6YFD	X-ray	T = 3 VLP, unusual CP
Beihai19	6YFE	X-ray	T = 3 VLP
Beihai21	6YFF	X-ray	T = 3 VLP
Beihai32	6YFG	X-ray	T = 3 VLP
EMS014	6YFH 6YFI	X-ray X-ray	T = 3 VLP CP dimer
ESE001	6YFJ	X-ray	T = 3 VLP
ESE007	6YFK	X-ray	T = 3 VLP
ESE020	6YFL	X-ray	T = 3 VLP
ESE021	6YFM	X-ray	T = 3 VLP
ESE058	6YFN	X-ray	T = 3 VLP
NT-214	6YFQ	X-ray	T = 3 VLP
NT-391	6YFR	X-ray	T = 3 VLP
Wenzhou1	6YFT	X-ray	T = 3 VLP
Wenzhou4	6YFU	X-ray	T = 3 VLP
PQ-465	6YFS	X-ray	T = 3 VLP
AC	6YF7	X-ray	T = 3 VLP
GQ-112	6YFP	X-ray	T = 3 VLP
GQ-907	6YFO	X-ray	T = 3 VLP

References

1. Loeb, T. Isolation of a bacteriophage specific for the F plus and Hfr mating types of Escherichia coli K-12. Science **1960**, 131, 932–933. [CrossRef] [PubMed]
2. Krishnamurthy, S.R.; Janowski, A.B.; Zhao, G.; Barouch, D.; Wang, D. Hyperexpansion of RNA Bacteriophage Diversity. PLoS Biol. **2016**, 14, e1002409. [CrossRef] [PubMed]
3. Callanan, J.; Stockdale, S.R.; Shkoporov, A.; Draper, L.A.; Ross, R.P.; Hill, C. Expansion of known ssRNA phage genomes: From tens to over a thousand. Sci. Adv. **2020**, 6, eaay5981. [CrossRef] [PubMed]
4. Neri, U.; Wolf, Y.I.; Roux, S.; Camargo, A.P.; Lee, B.; Kazlauskas, D.; Chen, I.M.; Ivanova, N.; Zeigler Allen, L.; Paez-Espino, D.; et al. Expansion of the global RNA virome reveals diverse clades of bacteriophages. Cell **2022**, 185, 4023–4037.e18. [CrossRef]
5. Wolf, Y.I.; Silas, S.; Wang, Y.; Wu, S.; Bocek, M.; Kazlauskas, D.; Krupovic, M.; Fire, A.; Dolja, V.V.; Koonin, E.V. Doubling of the known set of RNA viruses by metagenomic analysis of an aquatic virome. Nat. Microbiol. **2020**, 5, 1262–1270. [CrossRef]
6. Callanan, J.; Stockdale, S.R.; Adriaenssens, E.M.; Kuhn, J.H.; Rumnieks, J.; Pallen, M.J.; Shkoporov, A.N.; Draper, L.A.; Ross, R.P.; Hill, C. Leviviricetes: Expanding and restructuring the taxonomy of bacteria-infecting single-stranded RNA viruses. Microb. Genom. **2021**, 7, 000686. [CrossRef]
7. Epstein, S. National Academy of Sciences: Abstracts of Papers Presented at the Autumn Meeting, 29 October, La Jolla, California, 30 October-1 November 1961, Los Angeles. Science **1961**, 134, 1425–1437. [CrossRef]
8. Miyake, T.; Yanagisawa, K.; Watanabe, I. Alteration of a host character by infection with RNA phage Q-beta. Jpn. J. Microbiol. **1966**, 10, 141–148. [CrossRef]
9. Bradley, D.E. The Structure and Infective Process of a Pseudomonas Aeruginosa Bacteriophage Containing Ribonucleic Acid. Microbiology **1966**, 45, 83–96. [CrossRef]
10. Pourcel, C.; Midoux, C.; Vergnaud, G.; Latino, L. A carrier state is established in Pseudomonas aeruginosa by phage LeviOr01, a newly isolated ssRNA levivirus. J. Gen. Virol. **2017**, 98, 2181–2189. [CrossRef]

11. Coffi, H. Lysotypie des Acinetobacter. Master's Thesis, Laval University, Québec, QC, Canada, 1995.
12. Schmidt, J.M.; Stanier, R.Y. Isolation and Characterization of Bacteriophages Active against Stalked Bacteria. J. Gen. Microbiol. **1965**, 39, 95–107. [CrossRef] [PubMed]
13. Olsen, R.H.; Thomas, D.D. Characteristics and purification of PRR1, an RNA phage specific for the broad host range Pseudomonas R1822 drug resistance plasmid. J. Virol. **1973**, 12, 1560–1567. [CrossRef] [PubMed]
14. Coetzee, J.N.; Bradley, D.E.; Hedges, R.W.; Fleming, J.; Lecatsas, G. Bacteriophage M: An incompatibility group M plasmid-specific phage. J. Gen. Microbiol. **1983**, 129, 2271–2276. [CrossRef] [PubMed]
15. Sirgel, F.A.; Coetzee, J.N.; Hedges, R.W.; Lecatsas, G. Phage C-1: An IncC group; plasmid-specific phage. J. Gen. Microbiol. **1981**, 122, 155–160. [CrossRef] [PubMed]
16. Nuttall, D.; Maker, D.; Colleran, E. A method for the direct isolation of IncH plasmid-dependent bacteriophages. Lett. Appl. Microbiol. **1987**, 5, 37–40. [CrossRef]
17. Carmichael, G.G.; Weber, K.; Niveleau, A.; Wahba, A.J. The host factor required for RNA phage Qbeta RNA replication in vitro. Intracellular location, quantitation, and purification by polyadenylate-cellulose chromatography. J. Biol. Chem. **1975**, 250, 3607–3612. [CrossRef]
18. Blumenthal, T.; Landers, T.A.; Weber, K. Bacteriophage Q replicase contains the protein biosynthesis elongation factors EF Tu and EF Ts. Proc. Natl. Acad. Sci. USA **1972**, 69, 1313–1317. [CrossRef]
19. Rumnieks, J.; Lieknina, I.; Kalnins, G.; Sisovs, M.; Akopjana, I.; Bogans, J.; Tars, K. Three-dimensional structure of 22 uncultured ssRNA bacteriophages: Flexibility of the coat protein fold and variations in particle shapes. Sci. Adv. **2020**, 6, eabc0023. [CrossRef] [PubMed]
20. Dai, X.; Li, Z.; Lai, M.; Shu, S.; Du, Y.; Zhou, Z.H.; Sun, R. In situ structures of the genome and genome-delivery apparatus in a single-stranded RNA virus. Nature **2017**, 541, 112–116. [CrossRef]
21. Gorzelnik, K.V.; Cui, Z.; Reed, C.A.; Jakana, J.; Young, R.; Zhang, J. Asymmetric cryo-EM structure of the canonical Allolevivirus Qbeta reveals a single maturation protein and the genomic ssRNA in situ. Proc. Natl. Acad. Sci. USA **2016**, 113, 11519–11524. [CrossRef]
22. Koning, R.I.; Gomez-Blanco, J.; Akopjana, I.; Vargas, J.; Kazaks, A.; Tars, K.; Carazo, J.M.; Koster, A.J. Asymmetric cryo-EM reconstruction of phage MS2 reveals genome structure in situ. Nat. Commun. **2016**, 7, 12524. [CrossRef]
23. Zhong, Q.; Carratala, A.; Nazarov, S.; Guerrero-Ferreira, R.C.; Piccinini, L.; Bachmann, V.; Leiman, P.G.; Kohn, T. Genetic, Structural, and Phenotypic Properties of MS2 Coliphage with Resistance to ClO(2) Disinfection. Environ. Sci. Technol. **2016**, 50, 13520–13528. [CrossRef] [PubMed]
24. Cui, Z.; Gorzelnik, K.V.; Chang, J.Y.; Langlais, C.; Jakana, J.; Young, R.; Zhang, J. Structures of Qbeta virions, virus-like particles, and the Qbeta-MurA complex reveal internal coat proteins and the mechanism of host lysis. Proc. Natl. Acad. Sci. USA **2017**, 114, 11697–11702. [CrossRef]
25. Steitz, J.A. Identification of the A protein as a structural component of bacteriophage R17. J. Mol. Biol. **1968**, 33, 923–936. [CrossRef]
26. Steitz, J.A. Isolation of the A protein from bacteriphage R17. J. Mol. Biol. **1968**, 33, 937–945. [CrossRef]
27. Reed, C.A.; Langlais, C.; Kuznetsov, V.; Young, R. Inhibitory mechanism of the Qbeta lysis protein A2. Mol. Microbiol. **2012**, 86, 836–844. [CrossRef]
28. Rumnieks, J.; Tars, K. Crystal Structure of the Maturation Protein from Bacteriophage Qbeta. J. Mol. Biol. **2017**, 429, 688–696. [CrossRef]
29. Chang, J.Y.; Cui, Z.; Yang, K.; Huang, J.; Minary, P.; Zhang, J. Hierarchical natural move Monte Carlo refines flexible RNA structures into cryo-EM densities. RNA **2020**, 26, 1755–1766. [CrossRef] [PubMed]
30. Manchak, J.; Anthony, K.G.; Frost, L.S. Mutational analysis of F-pilin reveals domains for pilus assembly, phage infection and DNA transfer. Mol. Microbiol. **2002**, 43, 195–205. [CrossRef] [PubMed]
31. Meng, R.; Jiang, M.; Cui, Z.; Chang, J.Y.; Yang, K.; Jakana, J.; Yu, X.; Wang, Z.; Hu, B.; Zhang, J. Structural basis for the adsorption of a single-stranded RNA bacteriophage. Nat. Commun. **2019**, 10, 3130. [CrossRef]
32. Bradley, D.E. Shortening of Pseudomonas aeruginosa pili after RNA-phage adsorption. J. Gen. Microbiol. **1972**, 72, 303–319. [CrossRef]
33. Bayer, M.; Eferl, R.; Zellnig, G.; Teferle, K.; Dijkstra, A.; Koraimann, G.; Hogenauer, G. Gene 19 of plasmid R1 is required for both efficient conjugative DNA transfer and bacteriophage R17 infection. J. Bacteriol. **1995**, 177, 4279–4288. [CrossRef]
34. Paranchych, W.; Ainsworth, S.K.; Dick, A.J.; Krahn, P.M. Stages in phage R17 infection. V. Phage eclipse and the role of F pili. Virology **1971**, 45, 615–628. [CrossRef] [PubMed]
35. Danziger, R.E.; Paranchych, W. Stages in phage R17 infection. 3. Energy requirements for the F-pili mediated eclipse of viral infectivity. Virology **1970**, 40, 554–564. [CrossRef] [PubMed]
36. Dent, K.C.; Thompson, R.; Barker, A.M.; Hiscox, J.A.; Barr, J.N.; Stockley, P.G.; Ranson, N.A. The asymmetric structure of an icosahedral virus bound to its receptor suggests a mechanism for genome release. Structure **2013**, 21, 1225–1234. [CrossRef] [PubMed]
37. Harb, L.; Chamakura, K.; Khara, P.; Christie, P.J.; Young, R.; Zeng, L. ssRNA phage penetration triggers detachment of the F-pilus. Proc. Natl. Acad. Sci. USA **2020**, 117, 25751–25758. [CrossRef] [PubMed]

38. Schoulaker, R.; Engelberg-Kulka, H. Escherichia coli mutant temperature sensitive for group I RNA bacteriophages. J. Virol. **1978**, 25, 433–435. [CrossRef] [PubMed]
39. Gytz, H.; Mohr, D.; Seweryn, P.; Yoshimura, Y.; Kutlubaeva, Z.; Dolman, F.; Chelchessa, B.; Chetverin, A.B.; Mulder, F.A.; Brodersen, D.E.; et al. Structural basis for RNA-genome recognition during bacteriophage Qbeta replication. Nucleic Acids Res. **2015**, 43, 10893–10906. [CrossRef]
40. Kawashima, T.; Berthet-Colominas, C.; Wulff, M.; Cusack, S.; Leberman, R. The structure of the Escherichia coli EF-Tu.EF-Ts complex at 2.5 A resolution. Nature **1996**, 379, 511–518. [CrossRef]
41. Sengupta, J.; Agrawal, R.K.; Frank, J. Visualization of protein S1 within the 30S ribosomal subunit and its interaction with messenger RNA. Proc. Natl. Acad. Sci. USA **2001**, 98, 11991–11996. [CrossRef]
42. Kita, H.; Cho, J.; Matsuura, T.; Nakaishi, T.; Taniguchi, I.; Ichikawa, T.; Shima, Y.; Urabe, I.; Yomo, T. Functional Qbeta replicase genetically fusing essential subunits EF-Ts and EF-Tu with beta-subunit. J. Biosci. Bioeng. **2006**, 101, 421–426. [CrossRef]
43. Takeshita, D.; Tomita, K. Assembly of Qbeta viral RNA polymerase with host translational elongation factors EF-Tu and -Ts. Proc. Natl. Acad. Sci. USA **2010**, 107, 15733–15738. [CrossRef]
44. Kidmose, R.T.; Vasiliev, N.N.; Chetverin, A.B.; Andersen, G.R.; Knudsen, C.R. Structure of the Qbeta replicase, an RNA-dependent RNA polymerase consisting of viral and host proteins. Proc. Natl. Acad. Sci. USA **2010**, 107, 10884–10889. [CrossRef]
45. Takeshita, D.; Tomita, K. Molecular basis for RNA polymerization by Qbeta replicase. Nat. Struct. Mol. Biol. **2012**, 19, 229–237. [CrossRef]
46. Kutlubaeva, Z.S.; Chetverina, H.V.; Chetverin, A.B. The Contribution of Ribosomal Protein S1 to the Structure and Function of Qbeta Replicase. Acta Naturae **2017**, 9, 24–30. [CrossRef] [PubMed]
47. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Meng, E.C.; Couch, G.S.; Croll, T.I.; Morris, J.H.; Ferrin, T.E. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. Protein Sci. **2021**, 30, 70–82. [CrossRef]
48. Pei, J.; Grishin, N.V. AL2CO: Calculation of positional conservation in a protein sequence alignment. Bioinformatics **2001**, 17, 700–712. [CrossRef] [PubMed]
49. Miranda, G.; Schuppli, D.; Barrera, I.; Hausherr, C.; Sogo, J.M.; Weber, H. Recognition of bacteriophage Qbeta plus strand RNA as a template by Qbeta replicase: Role of RNA interactions mediated by ribosomal proteins S1 and host factor. J. Mol. Biol. **1997**, 267, 1089–1103. [CrossRef] [PubMed]
50. Takeshita, D.; Yamashita, S.; Tomita, K. Molecular insights into replication initiation by Qbeta replicase using ribosomal protein S1. Nucleic Acids Res. **2014**, 42, 10809–10822. [CrossRef]
51. Takeshita, D.; Yamashita, S.; Tomita, K. Mechanism for template-independent terminal adenylation activity of Qbeta replicase. Structure **2012**, 20, 1661–1669. [CrossRef]
52. Tomita, K. Structures and functions of Qbeta replicase: Translation factors beyond protein synthesis. Int. J. Mol. Sci. **2014**, 15, 15552–15570. [CrossRef] [PubMed]
53. Vasilyev, N.N.; Kutlubaeva, Z.S.; Ugarov, V.I.; Chetverina, H.V.; Chetverin, A.B. Ribosomal protein S1 functions as a termination factor in RNA synthesis by Qbeta phage replicase. Nat. Commun. **2013**, 4, 1781. [CrossRef]
54. Hohn, T. Packaging of Genomes in Bacteriophages: A Comparison of ssRNA Bacteriophages and dsDNA Bactgeriophages. Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. **1976**, 276, 143–150. [CrossRef]
55. de Haas, F.; Paatero, A.O.; Mindich, L.; Bamford, D.H.; Fuller, S.D. A symmetry mismatch at the site of RNA packaging in the polymerase complex of dsRNA bacteriophage phi6. J. Mol. Biol. **1999**, 294, 357–372. [CrossRef]
56. Chang, J.Y.; Gorzelnik, K.V.; Thongchol, J.; Zhang, J. Structural Assembly of Qbeta Virion and Its Diverse Forms of Virus-like Particles. Viruses **2022**, 14, 225. [CrossRef] [PubMed]
57. Fouts, D.E.; True, H.L.; Celander, D.W. Functional recognition of fragmented operator sites by R17/MS2 coat protein, a translational repressor. Nucleic Acids Res. **1997**, 25, 4464–4473. [CrossRef] [PubMed]
58. Lim, F.; Spingola, M.; Peabody, D.S. The RNA-binding site of bacteriophage Qbeta coat protein. J. Biol. Chem. **1996**, 271, 31839–31845. [CrossRef] [PubMed]
59. Persson, M.; Tars, K.; Liljas, L. PRR1 coat protein binding to its RNA translational operator. Acta Crystallogr. Sect. D Biol. Crystallogr. **2013**, 69, 367–372. [CrossRef]
60. Comas-Garcia, M. Packaging of Genomic RNA in Positive-Sense Single-Stranded RNA Viruses: A Complex Story. Viruses **2019**, 11, 253. [CrossRef]
61. de Martin Garrido, N.; Crone, M.A.; Ramlaul, K.; Simpson, P.A.; Freemont, P.S.; Aylett, C.H.S. Bacteriophage MS2 displays unreported capsid variability assembling T = 4 and mixed capsids. Mol. Microbiol. **2020**, 113, 143–152. [CrossRef]
62. Zhao, L.; Kopylov, M.; Potter, C.S.; Carragher, B.; Finn, M.G. Engineering the PP7 Virus Capsid as a Peptide Display Platform. ACS Nano **2019**, 13, 4443–4454. [CrossRef] [PubMed]
63. Lieknina, I.; Kalnins, G.; Akopjana, I.; Bogans, J.; Sisovs, M.; Jansons, J.; Rumnieks, J.; Tars, K. Production and characterization of novel ssRNA bacteriophage virus-like particles from metagenomic sequencing data. J. Nanobiotechnol. **2019**, 17, 61. [CrossRef]
64. Asensio, M.A.; Morella, N.M.; Jakobson, C.M.; Hartman, E.C.; Glasgow, J.E.; Sankaran, B.; Zwart, P.H.; Tullman-Ercek, D. A Selection for Assembly Reveals That a Single Amino Acid Mutant of the Bacteriophage MS2 Coat Protein Forms a Smaller Virus-like Particle. Nano Lett. **2016**, 16, 5944–5950. [CrossRef] [PubMed]
65. Cielens, I.; Ose, V.; Petrovskis, I.; Strelnikova, A.; Renhofa, R.; Kozlovska, T.; Pumpens, P. Mutilation of RNA phage Qbeta virus-like particles: From icosahedrons to rods. FEBS Lett. **2000**, 482, 261–264. [CrossRef]

66. Young, R. Phage lysis: Three steps, three choices, one outcome. J. Microbiol. **2014**, 52, 243–258. [CrossRef]
67. Chamakura, K.R.; Young, R. Single-gene lysis in the metagenomic era. Curr. Opin. Microbiol. **2020**, 56, 109–117. [CrossRef]
68. Chamakura, K.R.; Edwards, G.B.; Young, R. Mutational analysis of the MS2 lysis protein L. Microbiology **2017**, 163, 961–969. [CrossRef]
69. Reed, C.A.; Langlais, C.; Wang, I.N.; Young, R. A(2) expression and assembly regulates lysis in Qbeta infections. Microbiology **2013**, 159, 507–514. [CrossRef]
70. Chamakura, K.R.; Sham, L.T.; Davis, R.M.; Min, L.; Cho, H.; Ruiz, N.; Bernhardt, T.G.; Young, R. A viral protein antibiotic inhibits lipid II flippase activity. Nat. Microbiol. **2017**, 2, 1480–1484. [CrossRef]
71. Adler, B.A.; Chamakura, K.; Carion, H.; Krog, J.; Deutschbauer, A.M.; Young, R.; Mutalik, V.K.; Arkin, A.P. Multicopy suppressor screens reveal convergent evolution of single-gene lysis proteins. Nat. Chem. Biol. **2023**, 19, 759–766. [CrossRef] [PubMed]
72. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Zidek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. Nature **2021**, 596, 583–589. [CrossRef] [PubMed]
73. Baek, M.; DiMaio, F.; Anishchenko, I.; Dauparas, J.; Ovchinnikov, S.; Lee, G.R.; Wang, J.; Cong, Q.; Kinch, L.N.; Schaeffer, R.D.; et al. Accurate prediction of protein structures and interactions using a three-track neural network. Science **2021**, 373, 871–876. [CrossRef] [PubMed]
74. Zheng, W.; Zhang, C.; Li, Y.; Pearce, R.; Bell, E.W.; Zhang, Y. Folding non-homologous proteins by coupling deep-learning contact maps with I-TASSER assembly simulations. Cell Rep. Methods **2021**, 1, 100014. [CrossRef] [PubMed]
75. Lin, Z.; Akin, H.; Rao, R.; Hie, B.; Zhu, Z.; Lu, W.; Smetanin, N.; Verkuil, R.; Kabeli, O.; Shmueli, Y.; et al. Evolutionary-scale prediction of atomic-level protein structure with a language model. Science **2023**, 379, 1123–1130. [CrossRef] [PubMed]
76. Amann, S.J.; Keihsler, D.; Bodrug, T.; Brown, N.G.; Haselbach, D. Frozen in time: Analyzing molecular dynamics with time-resolved cryo-EM. Structure **2023**, 31, 4–19. [CrossRef]
77. Wagner, F.R.; Watanabe, R.; Schampers, R.; Singh, D.; Persoon, H.; Schaffer, M.; Fruhstorfer, P.; Plitzko, J.; Villa, E. Preparing samples from whole cells using focused-ion-beam milling for cryo-electron tomography. Nat. Protoc. **2020**, 15, 2041–2070. [CrossRef]
78. Franze de Fernandez, M.T.; Eoyang, L.; August, J.T. Factor fraction required for the synthesis of bacteriophage Qbeta-RNA. Nature **1968**, 219, 588–590. [CrossRef]
79. Franze de Fernandez, M.T.; Hayward, W.S.; August, J.T. Bacterial proteins required for replication of phage Q ribonucleic acid. Pruification and properties of host factor I, a ribonucleic acid-binding protein. J. Biol. Chem. **1972**, 247, 824–831. [CrossRef] [PubMed]
80. Wagner, A.; Weise, L.I.; Mutschler, H. In vitro characterisation of the MS2 RNA polymerase complex reveals host factors that modulate emesviral replicase activity. Commun. Biol. **2022**, 5, 264. [CrossRef]