

# Chapter 13

## ssRNA Phages: Life Cycle, Structure and Applications



Kaspars Tars

**Abstract** ssRNA phages belonging to the family Leviviridae are among the tiniest viruses, infecting various Gram-negative bacteria by adsorption to their pilus structures. Due to their simplicity, they have been intensively studied as models for understanding various problems in molecular biology and virology. Several of the studied ssRNA characteristics, such as coat protein–RNA interactions and the ability to readily form virus-like particles in recombinant expression systems, have fueled many practical applications such as RNA labeling and tracking systems and vaccine development. In this chapter, we review the life cycle, structure and applications of these small yet fascinating viruses.

### 13.1 Introduction

Bacteriophages belonging to the family *Leviviridae* are among the simplest known viruses, exhibiting positive-sense single-stranded RNA (ssRNA) genomes of just 3.5–4.5 kilobases, typically encoding only 4 proteins. Morphologically, *Leviviridae* particles present a spherical or isometric shape with a diameter of 30 nm.

Due to their simplicity, ssRNA phages have been used as models to study various processes in molecular biology and virology, including translation repression, RNA-protein interactions and virus evolution. Remarkably, the ssRNA phage MS2 was the first form of life for which a complete sequence of the genome was obtained, in 1976 in Walter Fiers's lab (Fiers et al. 1976). Four decades later, MS2 was the first form of life for which a 3D structure of the genome was established (Koning et al. 2016; Dai et al. 2017). In the meanwhile, ssRNA phages and their components have been found to be useful in many practical applications such as MS2 tagging, armoured RNA technology, drug delivery, nanoreactor construction and vaccine development, as discussed later in this chapter.

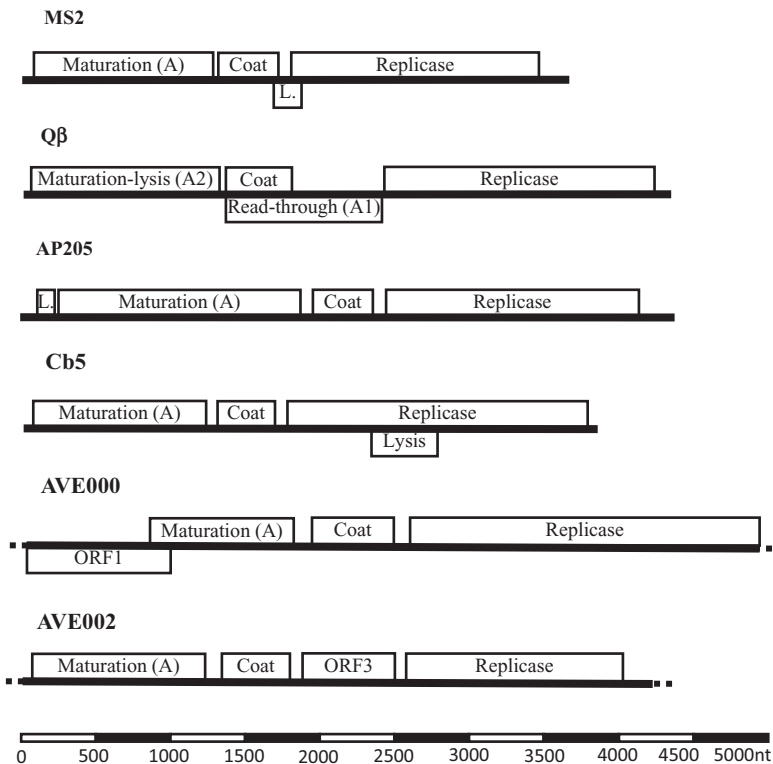
---

K. Tars (✉)  
Biomedical Research and Study Center, Riga, Latvia  
e-mail: [kaspars@biomed.lu.lv](mailto:kaspars@biomed.lu.lv)

### 13.2 Organization of the Genome, Classification and Relationship to Other Viruses

The genomes of all known ssRNA phages encode a so-called maturation, or “A” protein (AP, also known as A2 in genus *Allolevivirus*), a coat protein (CP) and a replicase subunit (RP) (Fig. 13.1). AP is a minor structural protein necessary for attachment to bacterial pilus structures, and CP is the major building block of the phage capsid. RP is the catalytic subunit of RNA-dependent RNA polymerase (RdRp). Additionally, many studied ssRNA phages encode a separate lysis protein (LP).

Family *Leviviridae* is further divided into two genera – *Levivirus* and *Allolevivirus*. In *Leviviridae* genomes, AP is encoded at the 5′ end, followed by CP and RP, always in the same order. In contrast, the position of LP in the genome may vary. Typically, the LP gene overlaps partially (Atkins et al. 1979; Olsthoorn et al. 1995) or completely (Kazaks et al. 2011; Rumnieks and Tars 2012) with the CP and/or RP genes,



**Fig. 13.1** Genomes of ssRNA phages MS2, Qβ, AP205, CB5, AVE000 and AVE002. Genes are shown as boxes. L. stands for lysis protein. In the incomplete sequences of phages AVE001 and AVE002, ORF1 and ORF3 of genes encoding two putative proteins are shown

albeit in different reading frames, except in the phage AP205, in which LP is encoded by a very short ORF at the very 5' end of the genome (Klovins et al. 2002). Phages belonging to genus *Allolevivirus* lack LP entirely and utilize a bifunctional A2 protein to carry out cell lysis (Karnik and Billeter 1983; Winter and Gold 1983; Bernhardt et al. 2001). Another hallmark of alloleviviruses is a gene encoding the read-through minor coat protein A1, whose function is somewhat unclear, although it has been shown that it is necessary for infection (Hofstetter et al. 1974).

The *Levivirus-Allolevivirus* division was assigned many decades ago when most studies on ssRNA phages were limited to those phages infecting *Escherichia coli* bacteria. Many other ssRNA phages have since been discovered and sequenced, and they do not seem to be particularly similar to any *Levivirus* or *Allolevivirus* representatives. Furthermore, *Levivirus* and *Allolevivirus* members are more closely related to each other than to many of the newly sequenced phages. Therefore, the classical *Levivirus-Allolevivirus* division is clearly outdated, and a new classification based solely on the sequence similarities of proteins should be introduced.

For quite some time, the sequences of relatively few ssRNA phages were known, but this situation changed dramatically in 2016, when two metagenome studies revealed more than 200 novel but somewhat incomplete genomes (Krishnamurthy et al. 2016; Shi et al. 2016). Additionally, 31 partial sequences including full-length CP gene were extracted from the NCBI nucleotide (nt) and environmental nucleotide (env\_nt) sequence databases (Lieknina et al. 2019). Although the corresponding phages themselves cannot be reconstructed from these partial sequences and the respective hosts remain unknown, these metagenome sequences revealed new data about the diversity of genomes and their encoded proteins. The sequences were so diverse that in many cases reliable assignment to the *Leviviridae* family was possible due only to the presence of the RP gene, the central part of which is remarkably conserved among all ssRNA phages. In contrast, the sequences of other proteins did not show any similarities to previously characterized members in many cases; for example, the presence of CP ORF could be detected only by assuming its approximate length and placement between the AP and RP genes. Some of the metagenome sequences (AVE002, AVE003, ESO001 and EMM000) seemed to contain two ORFs between the AP and RP genes. The first ORF after AP apparently encodes CP, as the expression of this gene from phages AVE002, ESO001 and EMM000 yielded virus-like particles (VLPs) that were morphologically similar to those of other phages (Lieknina et al. 2019). The expression of the second ORF (designated "ORF3", since it is the third ORF in the genome after AP and CP) of phages found in *E. coli* yielded either an insoluble product (AVE002, ESO001) or no product at all (AVE003, EMM000) (Lieknina et al. 2019). No cell lysis was observed in either case, suggesting that ORF3 is unlikely to be a candidate for encoding LP. It might still be the case that ORF3 is responsible for achieving lysis in host bacteria but is unable to do so in *E. coli*. However, the expression of LPs from other phages from different hosts has so far always led to cell lysis even when expressed in *E. coli*. Therefore, the identity and function of ORF3 remain unknown. Alternatively, ORF3 may be an artifact produced by sequencing errors. Thus, the ORF3 gene could actually be a 5' portion of the RP gene or a 3' extension of the CP gene, similar to the

A1 read-through protein in alloviviruses. ORF3 could also be a separate protein performing the same function as A1 in alloviviruses. However, this is rather unlikely since the “error model” would require multiple sequencing errors, and ORF3 does not display any homology to known proteins, including RP or A1, of any phage. Additionally, at least in AVE002, there are clearly defined SD sequences prior to the ORF3 and RP genes, further suggesting that ORF3 encodes a separate protein.

One of the metagenome sequences (AVE000) exhibited other unusual traits (Krishnamurthy et al. 2016). First, the incomplete genome was 4950 nucleotides long; therefore, the actual size of the full-length genome could be over 5000 bases. Second, the sequence included the another putative ORF with unknown function at the 5′ end, partially overlapping with AP in a different reading frame. The larger-than-usual size of the AVE000 genome might require larger particles, for which there is some indirect evidence, as discussed further in the VLP section.

*Leviviridae* members are not evolutionarily closely related to any other bacteriophages except for a very distant relationship to the only known dsRNA phages belonging to family *Cystoviridae*, since it is believed that all RNA viruses share common ancestry. The closest *Leviviridae* relatives appear to occur within genus *Mitovirus*, belonging to the family *Narnaviridae*, which consists of ssRNA viruses that infect the mitochondria of fungi (Hillman and Cai 2013). Mitoviruses do not have a protein capsid, and their genomes encode only one protein, RNA-dependent RNA polymerase (RdRp), which displays some sequence similarities to the RP of *Leviviridae* phages. The RdRp of the other *Narnaviridae* genus *Narnavirus* (infecting the cytoplasm of fungi) is even more distantly related to RP. Ourmia-like viruses, which infect plants, fungi and invertebrates, also harbor an RdRp that is distantly related to RP (Shi et al. 2016). Based on these similarities it has been suggested to create a separate ‘Narna-Levi’ clade consisting of *Leviviridae*, *Narnaviridae* and Ourmia-like viruses (Shi et al. 2016). However, the coat proteins of Ourmia-like viruses are clearly different from those of *Leviviridae* phages and, judging from similarities to other families, present a jelly roll  $\beta$ -barrel topology typical of many, if not most, viruses (Rastgou et al. 2009). Therefore, it appears that some of the mentioned viruses might carry RdRp genes that have been “borrowed” from each other via horizontal gene transfer. RP is the only protein from the *Leviviridae* family that clearly presents relatives within other viruses, while CP, AP, A1 and LP do not seem to be related to proteins from any other viruses or, indeed, to any other proteins in general.

## 13.3 Infectious Cycle

### 13.3.1 Adsorption, Genome Ejection and Penetration

ssRNA phages infect various Gram-negative bacteria by attachment to the sides of their pilus structures. In the case of *Escherichia coli*, most commonly conjugative plasmid-encoded F pili are used (Crawford and Gesteland 1964), but ssRNA phages infecting other bacteria have been reported to attach to completely different genome-encoded pili, such as polar pili in *Pseudomonas* (Bradley 1966) or swarmer cell-specific pili in *Caulobacter* (Schmidt 1966). Some phages display specificity to several hosts carrying certain conjugative pilus-encoding plasmids. For example, the bacteriophage PRR1 is specific to a variety of bacteria, provided that they harbour the incompatibility group IncP plasmid, which produces so-called P pili on the cell surface that serve as receptors for the phage (Olsen and Thomas 1973). Likewise, phage M is able to infect various bacteria carrying the IncM plasmid (Coetzee et al. 1983). However, most of the available knowledge about the initial steps of infection comes from studies of the closely related F-pili-specific phages MS2, R17 and f2. The physiological function of F-pili is to bind and bring donor and recipient bacteria close enough together for the subsequent exchange of genetic material. This is accomplished by retraction, which is essentially the shortening of F-pili, composed of many identical F-pilin protein monomers, via depolymerization. After retraction, the F plasmid enters bacteria through the conjugative pore formed by plasmid-encoded proteins. Apparently, ssRNA phages are able to repurpose this plasmid-transfer machinery to deliver their genomes inside the target cells.

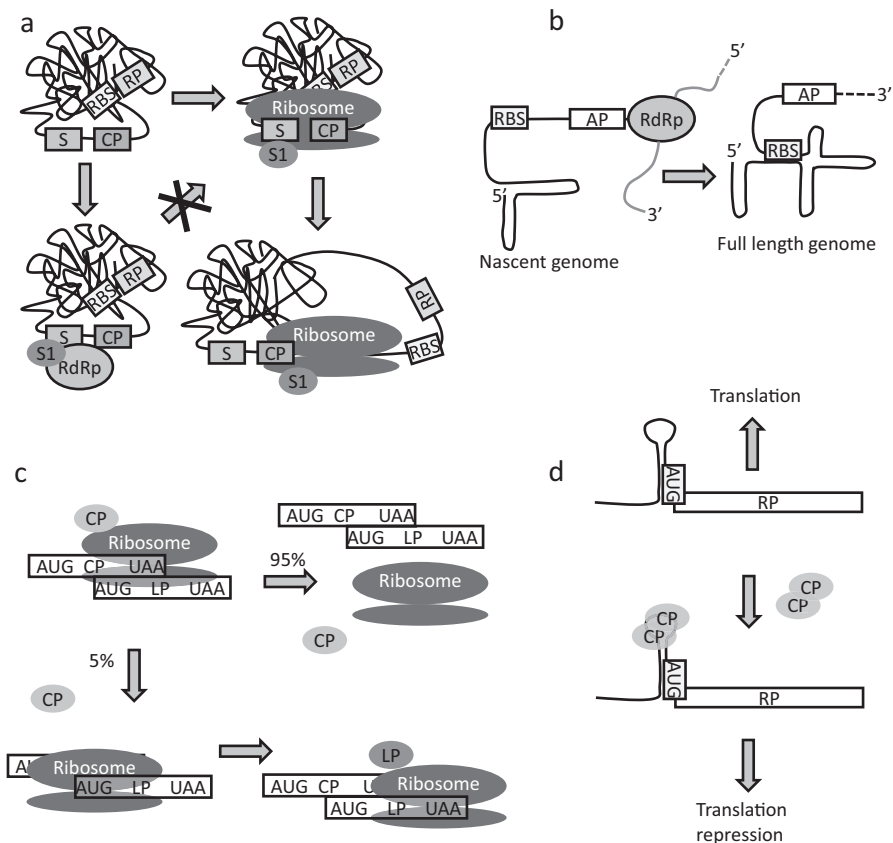
After attachment to pili on bacterial cell surface, AP is cleaved in two parts (Krahn et al. 1972), is ejected from particles together with the genome and enters bacteria via a poorly understood mechanism. A complex of only the AP and the genomic RNA is infectious to the cell (Shiba and Miyake 1975), suggesting that CP only protects viral RNA from environment and plays no other role in infectivity process. Although phage MS2 particles attach to both isolated and cell-bound F-pili (Valentine and Strand 1965), this attachment results in AP cleavage and genome ejection only if pili are attached to viable *E. coli* cells (Danziger and Paranchych 1970). Additionally, there is no evidence that the adsorption of the phage particles to F-pili actively triggers their retraction. However, it is known that F-pili can undergo retraction and elongation spontaneously (Clarke et al. 2008). Therefore, it is reasonable to assume that after adsorption, the occasional retraction of F-pili leads phage particles to be transported close to cell surface, triggering AP cleavage and genome ejection. The exact mechanism of further transport across the cell membrane remains unknown, but it can be speculated that AP, together with RNA, uses a conjugative pore intended for the transport of F-plasmid DNA.

### 13.3.2 *Synthesis of Proteins, Replication and Assembly*

Once the genomic RNA enters the cytoplasm of the host cell, it can act directly as mRNA for protein synthesis. The amounts of proteins to be translated are tightly regulated by a variety of mechanisms, including the accessibility of ribosome binding sites, the translational coupling of genes, the formation of RNA secondary structure elements, protein-RNA interactions and read-through of translation stop codons. The known expression regulation mechanisms of the MS2 phage are shown in Fig. 13.2. In the folded genome of the bacteriophage MS2, only the CP gene is initially accessible to the ribosome (Van Duin and Tsareva 2006). During the translation of the CP gene, the RNA secondary structure is partially disrupted, allowing access to the ribosome binding site of RP. Once the RP gene has been translated, the replication of the genome may begin. RP is the catalytic subunit of RNA-dependent RNA polymerase (RdRp), but the fully functional RdRp holoenzyme contains 3 additional proteins hijacked from the cellular translation machinery – the elongation factors EF-Tu and EF-Ts (Blumenthal et al. 1972) and ribosomal protein S1 (Wahba et al. 1974). Structural and functional studies have revealed that cellular components participate in the template binding, recognition and stabilization of the holoenzyme complex. By far the best-studied RdRp of ssRNA phages is that of phage Q $\beta$  due to its stability. The function, mechanism and structure of Q $\beta$  replicase have been studied in great detail – see (Tomita 2014) for a review. Q $\beta$  RdRp is among the fastest known RNA polymerases and can generate up to  $10^{10}$  copies of some artificially selected templates in 10 min (Chetverina and Chetverin 1993). For this reason, it has been proposed that Q $\beta$  RdRp could be used for the “room-temperature PCR” amplification of RNA, but template specificity requirements have prevented this type of application so far (Ugarov and Chetverin 2008). The RdRps of *Leviviridae* phages are remarkably specific for both the (+) and (–) strands of their own genomic RNAs and generally do not recognize other arbitrary templates. Additionally, it has been shown that MS2 RdRp does not recognize Q $\beta$  RNA or *vice versa* (Haruna and Spiegelman 1965). The specificity factors include sequences at the ends and in the middle of the genome, along with a high degree of secondary structure and long-distance base-pairing interactions (see (Rumnieks and Tars 2018) for a review).

In the folded RNA genome of phage MS2, AP translation is prevented by the formation of an extensive secondary structure around the ribosome binding site (Groeneveld et al. 1995). However, at the beginning of replication, the 5' end of the genome adopts other temporary conformations in which RBS is exposed to ribosomes, allowing the translation of the AP gene (van Meerten et al. 2001). This mechanism ensures that on average only one copy of AP is generated during each replication cycle, which makes sense since only one AP molecule must be present in the mature virion.

Although the initiation codon of the CP gene is accessible in the folded genome, the CP gene lacks a strong RBS sequence. Instead, the so-called S site is located prior to the CP gene. The S site shows affinity for the S1 protein, which is present



**Fig. 13.2** Regulation of gene expression in phage MS2. (a) In the folded genome, only the CP gene is available for translation. The CP gene is preceded by the S site, displaying affinity for the S1 protein, which is present in both the ribosome and RdRp. If RdRp is bound to the S site, translation from all genes is disabled. If the ribosome binds to the S site, upon translation, parts of the genome become unwound, exposing the RBS of the RP gene to make it accessible for translation. (b) In a full-length genome, AP RBS is normally hidden in the RNA secondary structure and not accessible to the ribosome. However, at the beginning of replication, when 5' end of the genome has been just synthesized by RdRp an alternative secondary structure is formed in the nascent genome, and AP RBS is temporarily accessible to the ribosome once per replication cycle. (c) The LP gene lacks its own RBS and can therefore be accessed only via the backsliding action of the ribosome. In <5% of cases, after the termination of CP synthesis, the ribosome slides backward and initiates LP translation. (d) The transcription of the RP gene is repressed by the TR-CP interaction around the initiation codon. The formation of the TR loop itself does not prevent translation, but the binding of the CP dimer to TR abolishes the binding of ribosomes

both in ribosomes and RdRp. Consequently, ribosomes and RdRp must compete for binding to the same RNA sequence (Van Duin and Tsareva 2006). Therefore, when RdRp is bound to the S site, the translation of all genes is prevented. This is important in the early stages of infection, when the production of more copies of the genome is more urgent than the translation of few existing copies.

As discussed later in the chapter in more detail, the translation of the RP gene is repressed not only by hiding its RBS in the folded genome but also by the binding of CP dimers to the sequence around the RP initiation codon (Gralla et al. 1974; Weber 1976). The synthesis of RP is thereby shut down in the late stages of infection when enough copies of the genome have accumulated, and further activity of RP is therefore not required.

After the synthesis of genomic RNA, AP and CP, the assembly of viral particles may begin. Due to the existence of extensive RNA secondary and tertiary structure, the genome must adopt a globular shape immediately after replication. The roughly spherical structure of the genome is further stabilized by the binding of AP, which interacts with several stem-loops of RNA (Dai, et al. 2017). Thus, the genome with bound AP seems to act as a nucleation center for virus assembly. After that, CP dimers attach to specific stem-loop regions in the genome, as discussed in the section on virion structure. It has been assumed that in the presence of cognate genomic RNA, no other cellular RNAs are packaged in the virion. Furthermore, co-infection experiments with MS2 and Q $\beta$  demonstrated that genomes of each phage are packaged only in their cognate capsids even when present in the same cell (Ling, et al. 1970). However, the recent analysis of phage Q $\beta$  virions by cryo-EM revealed that only approximately 20% of the examined particles contained the A2 protein and contained an RNA with a defined structure (Cui, et al. 2017). Although it can not be excluded that most particles were damaged during sample preparation, this seems to suggest that in at least some ssRNA phages, assembly results in a substantial amount of defective, noninfectious particles, possibly filled with cellular RNAs in a similar fashion to their recombinant VLPs, as discussed in Sect. 13.5.1.

### 13.3.3 Diverse Lysis Strategies in Leviviridae Phages

The last step in the viral life cycle of lytic phages is the lysis of the host cell. dsDNA phages typically make use of several genes to accomplish cell lysis, including genes encoding holins, lysins, spanins and additional proteins involved in the regulation of host lysis. In contrast, *Leviviridae* phages utilize only a single gene for the same purpose (see (Chamakura and Young 2019) for a review). Surprisingly, various clearly related ssRNA phages use at least three very different strategies to accomplish bacterial cell lysis. The LP of MS2 and other representatives of the *Levivirus* genus are inserted into the cellular membrane, creating channels that are permeable to ions (Goessens et al. 1988). This disrupts the electrostatic potential of the cellular membrane and activates cellular autolysins.



The target of the multifunctional A2 protein of phage Q $\beta$  and presumably other *Allolevivirus* members is the MurA enzyme, which catalyzes the first cytoplasmic step in proteoglycan wall synthesis, the transfer of enolpyruvate from phosphoenol pyruvate to UDP-N-acetylglucosamine. A2 binds to MurA, blocking the entry of its active site and thereby preventing the synthesis of the cell wall (Karnik and Billeter 1983; Winter and Gold 1983; Bernhardt et al. 2001). Remarkably, the A2 protein is able to block the enzymatic activity of MurA both in its free form and when present within the virion. Therefore, the mature particles of phage Q $\beta$  seem to be able to leave the host cell without the aid of any other factor.

Similar to MS2, the LP of phage M contains a single transmembrane helix. However, in the case of phage M, LP is able to interact with the MurJ protein, which is a flippase of lipid-linked precursors necessary for peptidoglycan synthesis. Apparently, the interaction of LP with MurJ locks the flippase in one of two conformations that is necessary for transport (Chamakura and Young 2019).

Since LP is potentially deadly to the host cell, it is important to avoid its overproduction in the early stages of infection. In MS2 and presumably other leviviruses, the translation initiation of the LP gene is coupled to the translation termination of the CP gene. The initiation codon of LP is located approximately 50 nucleotides upstream of the CP termination codon and lacks the preceding RBS sequence; therefore, it cannot be accessed by the ribosome directly. However, in approximately 5% of cases, after the termination of CP translation, the ribosome does not dissociate away from the template but slides backward and reinitiates translation at the LP start codon (Adhin and van Duin 1990). It is thereby ensured that LP accumulates in the cell slowly and only after the synthesis of CP.

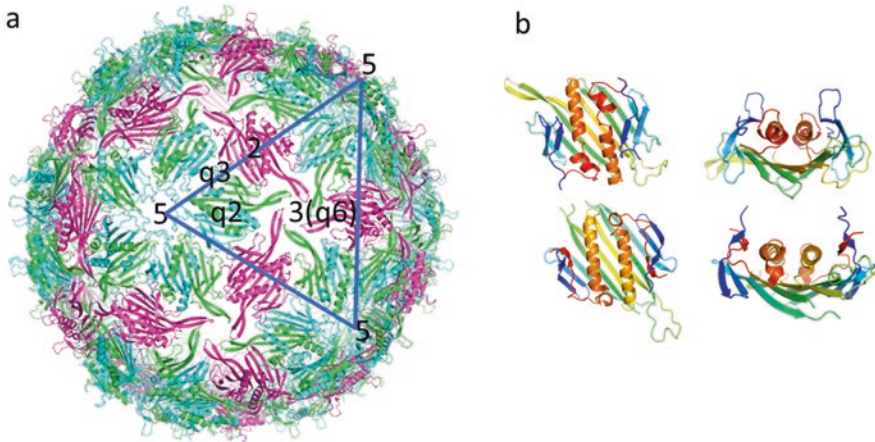
Since the LP genes of several other ssRNA phages are found at different locations in the genome, they have probably evolved independently; therefore, it might be possible that some of them use other yet to be discovered mechanisms of cell lysis.

The exact 3D structures of the phage LPs other than the Q $\beta$  A2 protein are unknown. LPs from different distantly related phages share little, if any, sequence homology, with the only universally common motif being a putative transmembrane helix, or sometimes two transmembrane helices, as in the case of CB5 (Kazaks et al. 2011). In most cases, the LP gene partially or fully overlaps with the RP gene in a different reading frame. This seems to leave little room for the evolution of LPs since they must coevolve with RP, which is the most conservative of the ssRNA phage proteins. Nevertheless, LPs are actually very efficient and are even able to lyse bacteria other than the hosts of their phages. For example, the LPs of AP205 (host, *Acinetobacter* genospecies 16) and CB5 (host, *Caulobacter crescentus*) are both able to lyse *E. coli* bacteria (Klovins et al. 2002; Kazaks et al. 2011).

## 13.4 Structure of the Virion

### 13.4.1 Icosahedral Component – Coat Protein

Mature ssRNA phage virions are composed of icosahedral  $T = 3$  protein shells composed of 178 copies of CP and a single copy of AP (Dent et al. 2013; Koning et al. 2016; Dai et al. 2017). Prior to advances in cryo-EM that made asymmetric reconstructions possible, essentially the only available structural information was for the protein capsid, which was initially thought to consist of 180 chemically identical CP subunits arranged in a perfectly icosahedral shell (Valegård et al. 1990) (Fig. 13.3). Because of  $T = 3$  symmetry, CP exists in three slightly different conformations, A, B and C. There are very tight interactions among the two monomers of CP; therefore, the capsid can be regarded as being composed of 89 dimers. There are two types of dimers in the capsid – one AB dimer (composed of CPs in the A and B conformations) and another CC dimer (composed of two dimers in the C conformation). The structure of CP in the capsid was first solved for MS2 (Valegård et al. 1990; Golmohammadi et al. 1993) and related phages fr (Liljas et al. 1994), Q $\beta$  (Golmohammadi et al. 1996), GA (Tars et al. 1997), PP7 (Tars et al. 2000), PRR1 (Persson et al. 2008) and CB5 (Plevka et al. 2009) and seemed to be quite conserved with respect to secondary structure elements. In the CP monomer, the N-terminal hairpin is always followed by a 5-stranded beta-sheet and two C-terminal helices. In dimers, two beta sheets are joined together, forming a single 10-stranded beta-sheet.



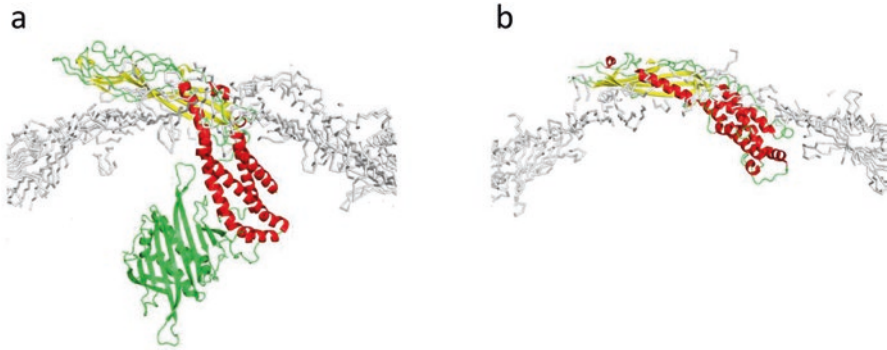
**Fig. 13.3** Structure of ssRNA phage capsids. **(a)** Structure of MS2 capsid. One facet of icosahedral particles is shown as a triangle. Approximate positions of icosahedral five-fold, three-fold and two-fold axes and quasi-symmetric  $q_3$  (relating three CP dimers),  $q_6$  (relating 6 CP dimers, coinciding with the icosahedral three-fold axis) and  $q_2$  (relating CP monomers within the AB dimer) axes are shown as well. **(b)** Structure of CP dimers in phages MS2 and AP205. Note the different placement of the terminal strand in the two cases

Deviation from this order is observed in the CP of phage AP205 (Shishovs et al. 2016), in which the N-terminal beta strand has “travelled” to the C-terminus, a circular permutation made possible by the close proximity of the C- and N-termini of the two monomers in the CP dimer. This permutation effectively exposes the C- and N-termini on the surface of AP205, whereas they are not well exposed and cluster together around quasi-threefold axes in other phages.

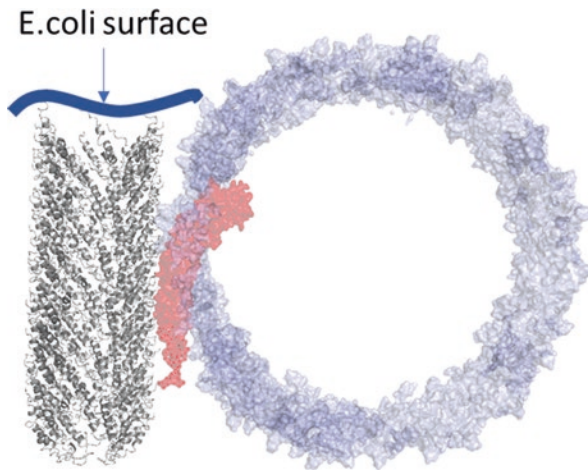
The exact interactions between CP dimers in capsids are surprisingly variable among even relatively closely related ssRNA phages. In the capsids of the studied *Levivirus* phages, CP dimers are held together merely by noncovalent protein-protein interactions. In some phages, like Q $\beta$  (Golmohammadi et al. 1996) and PP7 (Tars et al. 2000) covalent disulphide bonds link CP dimers around fivefold and quasi-sixfold axes, thereby significantly stabilizing the particles. In other cases, metal ions stabilize the structure of capsids in a similar way to that observed for many plant viruses (Tars et al. 2003). The capsid of bacteriophage CB5 exhibits calcium ions located in quasi-three-fold axes, which aid in holding 3 coat protein dimers together (Plevka et al. 2009). In some ssRNA phages, protein-RNA interactions also help to hold the dimers together; for example, CB5, in addition to being stabilized by metal ions, is further stabilized by RNA bases inserted between the CP dimers (Plevka et al. 2009).

### 13.4.2 Asymmetric Parts – AP, A1 and the Genome

Although it was initially believed that AP must traverse the prominent capsid pores around fivefold or quasi-sixfold axes, later cryo-EM studies revealed that one CC dimer in the capsid is replaced by a single copy of AP (Dent et al. 2013; Koning et al. 2016; Dai et al. 2017). The structure of AP is not similar to that of the CP dimer, yet AP is able to form relatively tight interactions with four different CP dimers. The introduction of AP in the capsid leads to some deviations from the icosahedral architecture – not only is a single CP dimer missing, but the neighboring CP dimers are also displaced from their perfectly icosahedral positions (Gorzelnik et al. 2016). Structurally, AP forms a single domain but can be described as being composed of two regions (Dai et al. 2017; Rumnieks and Tars 2017) – one that is predominantly helical, facing the genomic RNA, and a second containing beta strands ( $\beta$  part), facing the capsid exterior (Fig. 13.4) and interacting with the pilus receptor. While the helical parts are somewhat conserved among the related MS2 and Q $\beta$  phages, the  $\beta$  parts are very different. To some extent, this is reflected in their physiological functions – while the  $\beta$  part of the AP of MS2 interacts only with F pili, the  $\beta$  part of the A2 protein of Q $\beta$  also interacts with the MurA protein to promote cell lysis. While the structures of MS2 AP and Q $\beta$  A2 are quite different, their overall shapes are quite similar, forming a fairly extended, bent structure with their  $\beta$  part, extending tangentially from the virion surface. Recently, a 5 Å-resolution cryo-EM structure was reported for MS2 in complex with the F- pilus receptor (Meng et al. 2019) (Fig. 13.5). The  $\beta$  part of AP was found to be involved in



**Fig. 13.4** Placement of maturation protein in ssRNA phages Q $\beta$  (a) and MS2 (b). A section of phage particles around the maturation protein is shown. The particle shell is represented as a C $\alpha$  trace of CP molecules. Helices of the maturation protein are shown in red, strands are shown in yellow, and loops are shown in green. Note that the helical part faces the interior of the particles, while the  $\beta$  part is exposed on the surface. In Q $\beta$ , a single CP dimer is located in proximity to the the maturation protein



**Fig. 13.5** Interaction of phage MS2 with the F-pili receptor. A section of the phage is shown, revealing the position of AP. AP and CP are shown as semitransparent surface models in red and blue, respectively. F-pili are shown as a cartoon model. The location of the bacterial surface with respect to the orientation of the F-pili is indicated with a thick line

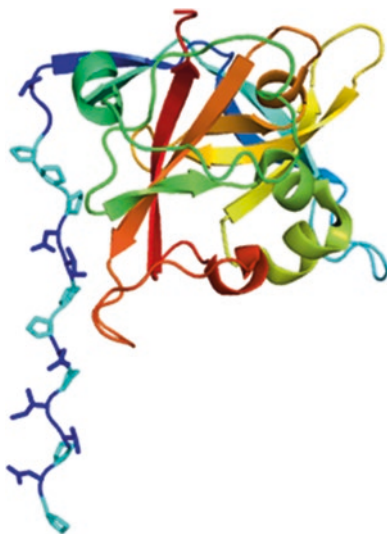
extensive hydrophobic and charged interactions with at least 6 pilin subunits. No significant changes in AP or genome structure were observed compared to unbound MS2 particles, confirming the previous observation that binding to F-pili *per se* does not induce AP cleavage or genome ejection. The interaction of phages with F-pili seems to be rather flexible, as three classes of MS2 particles with slight differences in their orientations with respect to the F-pili have been observed. The

particles bind to the pili so that the tip of AP bends away from the cell surface. Considering the relative orientation towards the cell and the hook-like structure of AP, it has been speculated that upon arrival at the cell surface, the particle could be mechanically opened like a coke can using AP as a pop tab.

In addition to a normal CP, *Allolevivirus* particles exhibit 3-10 copies of the A1 protein, which is an extended version of CP produced by the ribosome through an occasional read-through mechanism involving a leaky termination codon in which translation continues for an additional 600 nucleotides (Weiner and Weber 1971). The exact function of the A1 protein is unknown, although it has been shown that it is required for infection (Hofstetter, et al. 1974). Structurally, the A1 extension presents a roughly globular shape with a mixed  $\alpha/\beta$  architecture that is not observed in any other protein (Rumnieks and Tars 2011) (Fig. 13.6). The N-terminal part of the extension forms an unusually long polyproline type II helix. In cryo-EM reconstructions of Q $\beta$ , there are no traces of the A1 extension (Gorzelnik et al. 2016), suggesting that copies of A1 might be flexibly attached and/or randomly distributed in individual phage particles.

Although ssRNA phages nominally present a single-stranded RNA genome, this is correct only in the sense that a single strand of RNA is indeed packaged in each virion. However, more than 70% of the genome is involved in short- and long-distance base-pair interactions forming stem-loops, pseudoknots and other regions that essentially consist of dsRNA (Skripkin et al. 1990). Furthermore, the genome has a well-defined 3D structure that is more or less identical in all virions. The genome structure was first visualized by cryo-EM in phage MS2, initially at medium resolution (Koning et al. 2016) and later at high resolution (Dai et al. 2017), making it possible to observe most of the genome, although only some regions, particularly those forming interactions with the protein shell, were visible at a near-atomic resolution. Later, similar asymmetric cryo-EM structures were reported for phage Q $\beta$

**Fig. 13.6** Structure of the Q $\beta$  A1 extension. The structure is shown as a cartoon model, rainbow colored from the N- to C-termini. In the N-terminal polyproline helix, the side-chains are shown as stick models with prolines colored in cyan and other residues in blue

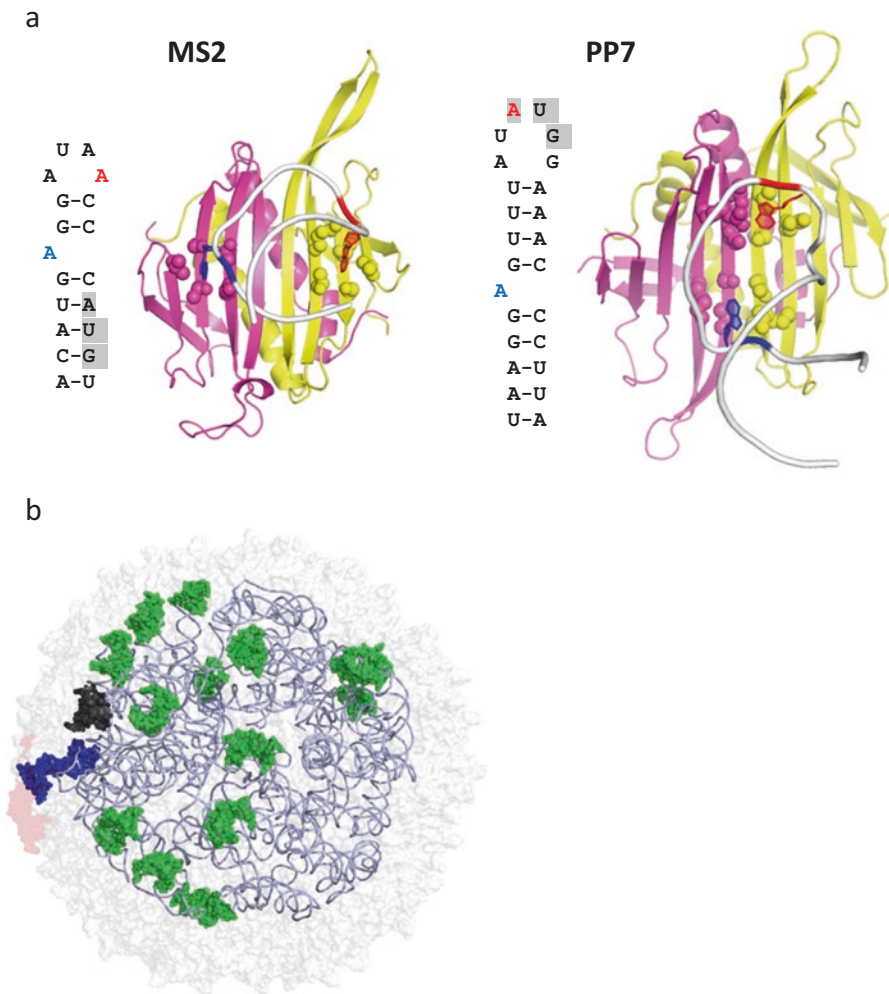


(Gorzelnik et al. 2016; Cui et al. 2017). Interestingly, unlike MS2, one copy of an isolated CP dimer was found inside the particle (Cui et al. 2017) (Fig. 13.4a), located near the A2 protein and bound to genomic RNA. Therefore, the Q $\beta$  virion is actually composed of 180 CP subunits, although a single CP dimer is not a part of the protein shell.

### 13.4.3 RNA-Coat Protein Interactions

For quite some time, the only available structural information about protein-RNA interactions in ssRNA phage particles came from the studies of CP dimers in complex with a 19 nucleotide-long stem-loop fragment known as TR (translation repression) from the genome region located around the replicase start codon in bacteriophage MS2 (Valegård et al. 1994). The main physiological role of this interaction seems to be the repression of the translation of the replicase gene in late stages of infection, when the presence of the replicase is no longer needed (Gralla et al. 1974; Weber 1976). Additionally, the same interaction provides a nucleation site for capsid assembly and contributes to the packaging of correct RNA inside the virions. The CP-TR interaction in phage MS2 has been studied in great detail both structurally and functionally and remains one of the best understood RNA-protein interactions in general (see (Rumnieks and Tars 2018) for a review), which has fuelled several practical applications, as discussed in later sections. Similar interactions with the same biological purpose exist in some (but not all) other ssRNA phages, including PP7 (Chao et al. 2008), PRR1 (Persson et al. 2008) and Q $\beta$  (Rumnieks and Tars 2014). Surprisingly, while the TR-binding sites of the MS2, PRR1 and Q $\beta$  CPs are clearly similar, the PP7-binding site is quite different. In both MS2 and PP7, the main specificity determinants of CP-RNA interactions are two adenine bases, one of which is located in the loop, while the other forms a bulge in the stem (Fig. 13.7a). The two adenines occupy symmetrical binding pockets in the CP dimer. However, while in MS2 both pockets is formed by residues belonging only to one monomer, in PP7, pockets are found at the interface of monomers. Therefore, the TR-binding sites of the two phages are located at completely different sites of CP, suggesting that the CP-TR interaction might have evolved in two independent ways in ssRNA phages. However, it is now clear that the CP-TR interaction plays a rather minor role in the life cycle of phage, as TR-deficient mutants are viable and only marginally less fit than wt phages (Peabody 1997a; Licis et al. 2000). Furthermore, some other phages, such as AP205 and CB5, do not seem to present this specific interaction at all.

In addition to the TR-CP interaction, numerous other fragments of ssRNA are involved in CP-RNA interactions. In most cases, these are other stem-loop structures from the genome with a similar appearance. In a cross-linking study, more than 50 stem loops were identified as being bound to CP (Rolfsson et al. 2016), many of which were visualized in later asymmetric cryo-EM reconstructions of MS2 (Dai et al. 2017). However, only 15 stem loops were visible in a cryo-EM



**Fig. 13.7** Protein – RNA interactions in ssRNA phages. (a) Interactions of the TR stem-loop with MS2 and PP7 CP dimers. Two adenine bases participating in the most important sequence-specific interactions are shown with blue and red stick models. CP monomers are shown with yellow and magenta cartoon models with the residues forming the binding pockets for adenine bases shown as sphere models. Note the markedly different placement of the two adenine-binding pockets in MS2 and PP7. The secondary structure of the TR stem loops is shown as well, with both important adenylates indicated in the same color as in the stick models. The replicase gene initiation codon is highlighted in gray. (b) Asymmetric cryo-EM reconstruction of phage MS2. CP and AP are shown as semitransparent light gray and red surface models, respectively. High-resolution CP-binding genome stem loops are shown as green sphere models. TR (same as in panel a for MS2) is shown in black and AP interacting stem-loop in blue. A lower-resolution genome model is shown as a coil

reconstruction at an atomic resolution (Fig. 13.7b), suggesting that the remainder are partially disordered, indicating weaker binding to CP. One of the 15 stem-loop interactions was identified as a previously known CP-TR complex located close to but not in direct contact with AP. In addition to CP-RNA interactions, AP also interacts with segments of RNA, notably with a 24 nucleotide-long stem-loop at the 3' end of the genome. The distribution of RNA in the virion is somewhat uneven, with the densest portion occupying roughly one-half of the capsid volume, where AP and most of the high-resolution stem-loop-CP complexes are found. The cryo-EM reconstruction of phage Q $\beta$  shows a subatomic 4.7 Å resolution; therefore, the details of its genome structure are not as well resolved as in the case of MS2, and atomic models have not been built for parts of the genome. As discussed previously, unlike MS2, a single isolated CP dimer was found to be located inside of the Q $\beta$  particle, bound to genomic RNA. Although it is impossible to determine with certainty at which genomic RNA segment the isolated CP dimer is bound due to the limited resolution, it might be the TR sequence, since the correlation coefficient with the known crystal structure of the Q $\beta$  CP-TR complex was 0.91 (Cui et al. 2017).

### 13.5 Practical Applications of ssRNA Phages and Their Components

Over the past decades, surprisingly many applications of ssRNA phages and their components that are useful both in fundamental research and product development have been found. Broadly speaking, all of these applications can be grouped into three main classes, relying on the use of intact ssRNA phages, their VLPs or CP-RNA interactions.

ssRNA phages themselves are used mainly in environmental and disinfection studies, serving as markers for the tracking of viral and microbial sources or as surrogate models for the control of viral contamination in various samples. The applications of intact ssRNA phages are not further discussed in this chapter. CP-RNA interactions have contributed significantly to the development of various imaging applications. The VLPs of ssRNA phages are being used in vaccine and drug delivery development. Some applications, such as armoured RNA technology and MS2 display, combine the use of VLPs and CP-RNA interactions. In the following section, a concise summary of the best-known applications involving ssRNA phage VLPs and CP-RNA interactions is given.



### 13.5.1 VLPs of ssRNA Phages

Several applications of ssRNA phages rely on the production of recombinant VLPs from the phages. In most cases, the expression of the CP gene alone in bacteria or yeast is sufficient to produce soluble CP, and CP dimers then spontaneously assemble into VLPs in host cells. Recombinant VLPs of numerous ssRNA phages have been produced in this way (Kastelein et al. 1983; Kozlovskaya et al. 1986; Peabody 1990; Kozlovskaya et al. 1993). Recently, VLPs of 80 previously unknown ssRNA phages were produced using CP sequences from metagenome sequencing data (Liekniina et al. 2019). In most cases, recombinant VLPs are morphologically indistinguishable from the respective phages when imaged via conventional negative-staining EM. Furthermore, in two known cases involving phages MS2 (Golmohammadi et al. 1993; Valegård et al. 1997) and CB5 (Plevka et al. 2009), when the icosahedrally averaged crystal structures of both VLPs and the respective phages were determined, the capsid structures were found to be virtually identical. However, in some cases, recombinant VLPs are somewhat heterogeneous as previously observed in a cryo-EM reconstruction of bacteriophage AP205 VLPs and may include a mixture of  $T = 3$ ,  $T = 1$  and somewhat irregular particles (Shishovs et al. 2016). Mutant VLPs of phages MS2 and PP7 have been shown to exhibit  $T = 1$  (Asensio et al. 2016) and  $T = 4$  (de Martin Garrido, et al. 2019; Zhao, et al. 2019) symmetries. During the characterization of 80 VLPs from metagenome data, several VLPs were found to display considerable deviations from classical  $T = 3$  particles. In two cases, only smaller  $T = 1$  particles could be observed, while in the case of AVE000, somewhat larger heterogeneous particles were observed, some of which may exhibit  $T = 4$  symmetry. In the case of AVE016, elongated  $T = 3$   $Q = 4$  VLPs were produced. Since VLPs originate from metagenome sequences, the actual phages are not available. Therefore, it is difficult to judge whether the observed shapes and sizes of larger-than  $T = 3$  particles are merely a consequence of the artificial production system or the respective phages themselves indeed exhibit  $T = 4$  or  $T = 3$   $Q = 4$  symmetry. Intriguingly, as discussed in the previous section, AVE000 presents a very long genome, possibly of more than 5000 nucleotides. Therefore, it might be that larger  $T = 4$  particles are actually present not only in artificial VLPs but also in the phage itself, to enable the packaging of larger genomes. In contrast, the smaller  $T = 1$  particles could not possibly be present in native phages since there would not be enough space for their genome.

Although ssRNA phage VLPs obviously do not contain a genome, they are packaged with various RNAs that are acquired during assembly following expression in host cells (Pickett and Peabody 1993). This is largely due to the nonspecific interactions of CP with any RNA, in which positively charged lysine and arginine residues interact with negatively charged RNA phosphate groups, and some stacking interactions among aromatic residues and RNA bases seem to contribute to unspecific binding as well. Nevertheless, it has been shown that in addition to other cellular RNAs, the recombinant VLPs of ssRNA phages contain substantial amounts of CP mRNA, a property that can be used in several applications, as discussed below.

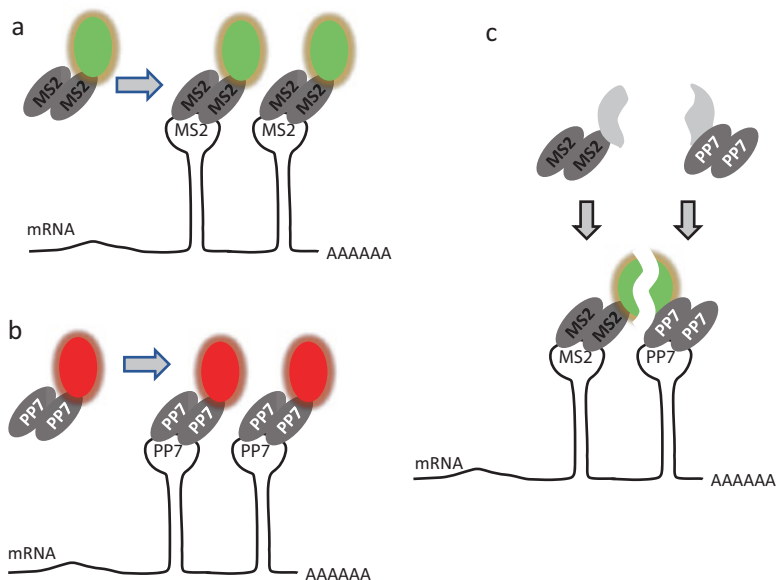
### 13.5.2 *MS2 Tagging: Identification of Protein-RNA Interactions, RNA Labeling and Tracking*

MS2 tagging refers to a variety of techniques relying on specific CP-TR interactions in phage MS2. For this purpose, the formation of VLPs is not desired; therefore, assembly-deficient MS2 CP mutants are used. One MS2-tagging technique referred to as MS2 BioTRAP is a method for the identification of protein-RNA interactions (Bardwell and Wickens 1990; Tsai et al. 2011). In RNA molecules of interest, several copies of TR are introduced. The RNA molecules are coexpressed with the MS2 coat protein modified with an HB tag sequence, which enables biotinylation *in vivo*. As a result, the RNA molecule of interest is decorated with biotinylated copies of MS2 CP. Any proteins bound to the RNA molecule of interest can be extracted together with the RNA using streptavidin beads. The identity of the bound proteins can be revealed by mass spectrometry or other suitable techniques. Using similar approach called MS2-TRAP it is possible to identify also RNA-RNA interactions, such as binding of miRNAs to their target mRNAs (Yoon et al. 2012; Yoon and Gorospe 2016).

Another conceptually similar MS2 tagging technology can be used for the tracking mRNA in living cells (Fig. 13.8). At the 3' end of an mRNA of interest, multiple copies of TR are inserted. The MS2 coat protein is fluorescently labeled by, for example, fusion to green fluorescence protein (GFP). As a result, upon the binding of labeled MS2 CP, the RNA of interest also becomes fluorescently labeled. Using confocal microscopy, it is further possible to track the path of the tagged RNA of interest in the cell (Bertrand et al. 1998). Similar technology has been developed for phage PP7 (Larson et al. 2011), which also displays a specific interaction of its CP with TR. Several adaptations of the described method exist, some of which make use of both MS2 and PP7 TR sequences located close together in the target RNA. In this case, MS2 and PP7 CPs are labeled with different fragments of an engineered GFP or its variants that are unable to emit fluorescent signals by themselves (Wu et al. 2014; Park et al. 2020). However, upon binding to their respective TRs, the MS2 and PP7 CPs bring the two fragments of GFP together, resulting in fluorescence. Compared to standalone MS2 or PP7 techniques, this significantly reduces the background from unbound CP-GFP molecules. In another adaptation of the PP7 and MS2 RNA labeling method, the two CPs are labeled with different fluorescent proteins (Hocine et al. 2013). Thus, two different RNAs harboring MS2 and PP7 TRs can be simultaneously tracked in a living cell.

### 13.5.3 *Design of Riboswitches*

Riboswitches are regulatory components of mRNA that are able to alter gene expression upon binding to small molecule effectors. The activity of a riboswitch may lead to various changes in mRNA, such as accessibility of the ribosome



**Fig. 13.8** Use of CP-RNA interactions in mRNA visualization. In the 3'-end of target mRNA, several phage-specific TR sequences are inserted. CP dimers of phages MS2 (a) and PP7 (b) are modified by genetic fusion with fluorescent proteins, enabling detection of labeled mRNAs of interest. By combining different specificities of TR interactions in PP7 and MS2 and two different fluorescence proteins, it is possible to track two different mRNAs simultaneously. (c) To reduce background fluorescence resulting from proteins, unbound to target mRNAs, a combined MS2-PP7 approach has been introduced. Both MS2 and PP7 TRs are introduced in the mRNA sequence next to each other. MS2 and PP7 CP dimers are modified by the attachment of GFP segments, which are unable to emit fluorescence signals by themselves. The two halves of GFP are brought together and begin to emit light when the CP dimers bind to their respective TRs

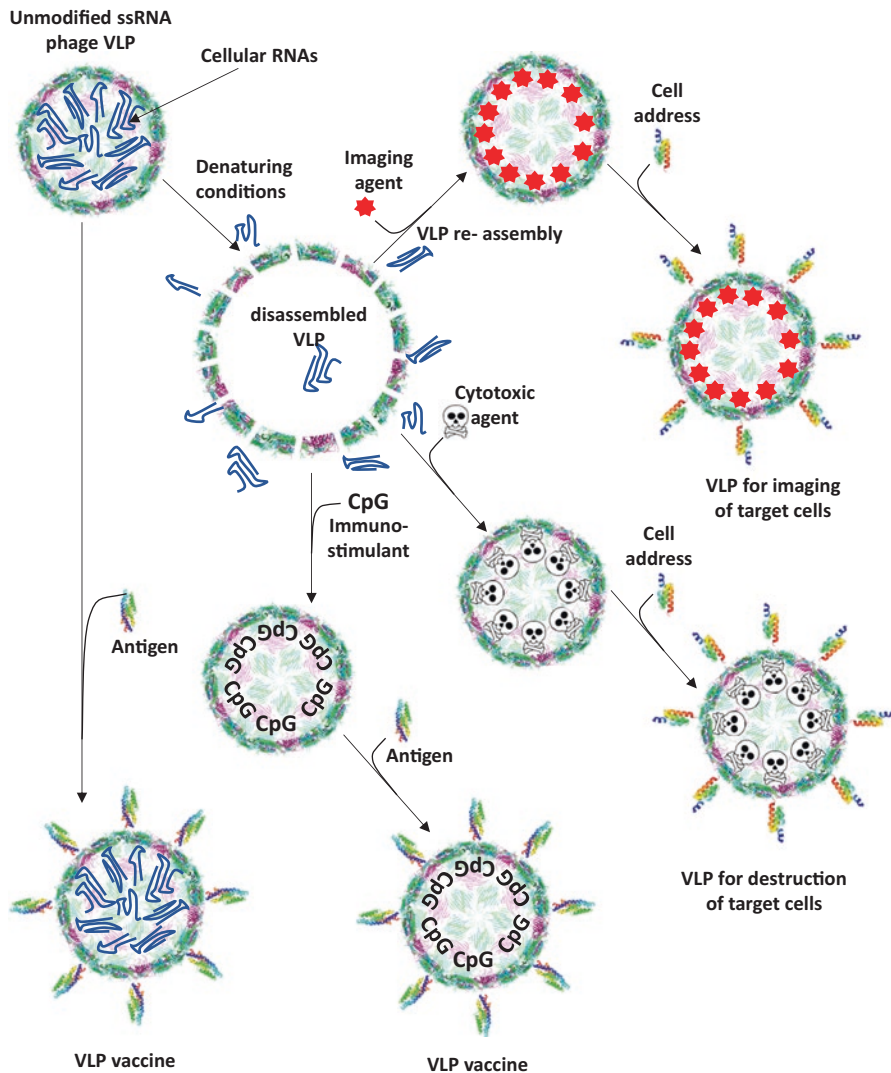
binding site, the formation of transcription terminator hairpins, self-cleavage by induced ribozyme activity or modifications of splice sites. In general, the binding of effector molecules alters the secondary structure of mRNA. This property has been used in selection, where the MS2 TR is inserted in close proximity to the riboswitch sequence (Wu et al. 2019). The experiment can be designed in an “On” or “Off” configuration, so that the RNA structure after the binding of the effector molecule becomes compatible or incompatible, respectively, with the formation of a TR hairpin. The presence or absence of a TR hairpin can in turn be detected by treatment with MS2 CP-GFP.

### 13.5.4 Armored RNA Technology

To detect pathogenic RNA viruses by RT-qPCR in environmental, food or clinical samples, reliable process quality control RNAs are necessary. However, RNA is particularly vulnerable to degradation; therefore, special precautions must be taken to preserve intact control RNAs. One strategy is to use an “armored” cage around RNA molecules to prevent the access of RNases. The natural ability of ssRNA phage capsids to package their genomes has been utilized for the protection of control RNAs. The RNA of interest can be genetically fused with the TR operator and coexpressed with the CP gene of the corresponding phage. As a result, the control RNA becomes encapsulated in VLPs and effectively protected from attack by RNases. This technology has been well developed using phage MS2 VLPs (Pasloske et al. 1998) in which RNA sequences from the genomes of pathogenic viruses, including HCV, HIV, SARS, West Nile virus and many others, have been packaged, and a wide variety of armored RNAs are commercially available for routine testing. Phage Q $\beta$  VLPs have been used for the same purpose, and it has been shown that the resulting armored RNAs are significantly more stable compared to those packaged in MS2 VLPs (Yao et al. 2019). Although compared to RNA, the degradation of DNA is of significantly less concern, a similar “armored DNA” technique has been developed. In this case, packaging is performed by re-assembling phage MS2 VLPs *in vitro* in the presence of dsDNA fragments (Zhang et al. 2015).

### 13.5.5 Targeted Delivery

The VLPs of ssRNA phages can be used as nanocontainers for the targeted delivery of various diagnostic or therapeutic agents (see (Pumpens et al. 2016) for a review) (Fig. 13.9). The surface of VLPs can be decorated with an “address”, capable of recognizing a particular cell type. Examples of the addresses used for the targeting of ssRNA phage VLPs include cancer cell-targeting proteins(Elsohly et al. 2017) or peptides(Carrico et al. 2008), glycans(Rhee et al. 2012), DNA aptamers(Cohen and Bergkvist 2013), antibodies(ElSohly et al. 2015) and the Z-domain of the *S. aureus* A protein (Zhao et al. 2019). The surface of VLPs can be further modified by the attachment of cell-penetrating peptides(Wei et al. 2009) or PEG chains(Kovacs et al. 2007). Various cargoes can be packaged inside VLPs, including siRNAs(Galaway and Stockley 2013), miRNAs(Wang et al. 2016), dyes(Aanei et al. 2018), small molecule drugs(Finbloom et al. 2018), toxins(Wu et al. 1995), quantum dots(Ashley et al. 2011), nanoparticles(Freivalds et al. 2014), metal ions(Kolesanova et al. 2019), radionuclides (Aanei et al. 2016) and other agents. If an attached address is used in the same VLPs, the cargo can be delivered only to target cells, enabling their selective labeling, activity modulation or destruction. In pioneering studies, it was shown that MS2 VLPs packaged with the ricin A chain and decorated with transferrin on their exterior were able to selectively and



**Fig. 13.9** Applications of ssRNA phage VLPs in vaccine development and drug delivery. The surface of VLPs can be modified, resulting in antigen presentation for vaccine or attachment of address for targeted nanocontainer delivery. VLPs can be disassembled and reassembled to pack immunostimulatory CpGs for vaccine development or cytotoxic or imaging agents for destruction or visualization of target cells

efficiently kill leukemia cells with exposed transferrin receptors (Wu et al. 1995). Since those studies, many other ssRNA phage VLP-derived targeted nanocontainers with various loads have been developed, most of which are aimed at treating various forms of cancer.

However, even though in many cases, the VLPs of ssRNA phages have indeed been able to deliver various cargoes to target cells, their use as delivery agents in immunocompetent animal models or humans faces several challenges. Most importantly, VLPs are very potent immunogens, which is a property that is extremely useful in their application in vaccine development but represents a hurdle for drug delivery. If VLP-based nanocontainers are injected into animals, a massive immune response will clear them from the circulation, especially after repeated injections. To some extent, this effect can be minimized by the PEGylation of VLPs (Kovacs et al. 2007). Furthermore, for targeted delivery, the surface of VLPs must be decorated with a suitable address that is able to recognize particular cell types. The address itself may provoke an immune response, further complicating VLP-based delivery applications.

### 13.5.6 Nanoreactors

VLPs can be packaged with enzymes, efficiently transforming them into nanoscale reactors performing certain enzymatic reactions or even providing whole metabolic pathways. Similar systems exist in nature. For example, many bacteria make use of so-called microcompartments, which are quasi-icosahedral protein shells containing encapsulated cores of enzymes (see (Kerfeld et al. 2018) for a review). Microcompartments protect the rest of the cell from potentially toxic reaction products and provide selectivity filters to substrates, products and co-factors. They also increase the effectiveness of metabolic pathways by sequestering, locally concentrating and protecting enzymes from the action of proteases, thereby drawing considerable interest in the field of biotechnology. However, bacterial microcompartments contain several different structural proteins, complicating their engineering. The icosahedral architecture of VLPs is somewhat similar to that of microcompartments, but VLPs are considerably simpler in the case of ssRNA phages composed of only one type of protein. Therefore, substantial efforts have been devoted to the engineering of ssRNA VLPs for use as nanoreactors. Several enzymes have been successfully encapsulated in Q $\beta$  VLPs by coexpression with Q $\beta$  CP, including peptidase E and firefly luciferase (Fiedler et al. 2010). Packaging was promoted by RNA linker sequence, consisting of Q $\beta$  TR and RNA aptamer, able to bind arginine-rich peptide, attached to enzymes. Alternatively, enzymes such as alkaline phosphatase can be modified by attachment to negatively charged oligomers such as the DNA analogue of TR or acidic peptides and packaged *in vitro* in MS2 VLPs (Glasgow et al. 2012). Furthermore, by changing the charges of residues lining the pores around the fivefold and quasi-sixfold axes of VLPs, it is possible to alter the rate of the enzymatic reaction (Glasgow et al. 2015), demonstrating that it is possible to regulate the transport of substrates and products through pores to some extent. In another development, two enzymes from the indigo biosynthesis pathway, pyridoxal phosphate (PLP)-dependent tryptophanase TnaA and nicotinamide adenine dinucleotide phosphate (NADPH)-dependent monooxygenase FMO, were

covalently attached to the inner surface of MS2 VLPs using the spy-catcher technique described in Sect. 13.5.7. (Giessen and Silver 2016). The spy tag was genetically inserted in the inner loop of MS2 CP, while the spy-catcher tag was added to two enzymes, TnaA and FMO. The coexpression of all components led to the efficient production of nanoreactors that were 60% more efficient in producing indigo inside the bacterial host compared to naked enzymes. Furthermore, the purified nanoreactors retained 95% of their activity after incubation for 7 days at 25 °C, while isolated enzymes retained only 5% of their activity after similar treatment.

### 13.5.7 Vaccine Development

Probably the most promising medical application of ssRNA phage VLPs lies in their use in vaccine development (Fig. 13.9). Similar to many other VLPs, the VLPs of ssRNA phages act as very potent immunogens. There are several reasons that VLPs are very immunogenic, the most important of which is their ability to cross-link B-cell receptors, greatly enhancing B-cell proliferation (Bachmann and Zinkernagel 1996). This property can be used to fuse weak antigens to the surface of VLPs to enhance their immune response. Importantly, the coupling of antigens to VLPs may overcome B-cell tolerance to self-antigens (Bachmann et al. 1993) – a property that can be used to create antibodies against certain undesired proteins of organisms; successful examples include but are not limited to the reduction of angiotensin levels in case of high blood pressure (Tissot et al. 2008) and the reduction of proprotein convertase subtilisin/kexin type 9 (PCSK9) levels to reduce low-density lipoprotein cholesterol (Crossey et al. 2015). Although the VLPs of many other viruses can be used in a similar fashion, ssRNA phage VLPs present certain advantages. First, they are exceptionally easy to produce in *Escherichia coli* or yeast systems. Second, they are packaged with unspecific cellular RNAs, which greatly increases their immune response by activating TLR3 and TLR7 toll-like receptors. If necessary, their RNA content can be exchanged for something else – for example, immunostimulatory agents such as CpG oligonucleotides, to achieve an enhanced TLR9 response (Bachmann et al. 2003). Third, antigens can be easily coupled to the VLPs of ssRNA phages by either genetic fusion or chemical coupling. Owing to these advantages, various ssRNA phage VLP scaffold vaccine candidates for a wide variety of diseases have been reported in well over 100 scientific articles, in many cases including proof-of-principle studies in animals or even clinical trials (see (Pumpens, et al. 2016) for a review). In general, all ssRNA phage VLP vaccine candidates can be divided into two broad groups. The first group is represented by candidate prophylactic vaccines against infectious diseases, such as influenza (Tissot et al. 2010; Jegerlehner et al. 2013), West Nile fever (Spohn et al. 2010), Lyme disease (Marcinkiewicz et al. 2018), cervical cancer caused by human papilloma virus (Tumban et al. 2012; Zhai et al. 2019) or malaria (Ord et al. 2014). The second group comprises therapeutic vaccines targeted against self-antigens, cancer antigens or even small molecules, as exemplified by angiotensin (Tissot et al. 2008),

tumour-specific carbohydrates(Yin et al. 2015, 2016) and nicotine(Maurer et al. 2005), respectively. In both cases, the corresponding antigens must be attached to VLPs. In the case of protein or peptide antigens, this can be achieved either by genetic fusion or chemical coupling, while for nonprotein antigens, chemical coupling is the only option. Regarding genetic fusion, tolerability to insertions depends heavily on the choice of the particular VLP. In the VLPs of phages MS2, PP7 and Q $\beta$ , the N- and C- termini of three adjacent dimers are closely clustered together around quasi-threefold axes(Shishovs et al. 2016); therefore, longer insertions in the N- and C- termini are usually not tolerated due to steric incompatibility. The so-called AB loop located on the VLP surface has also been utilized for short insertions. It is possible to genetically fuse the N- and C- termini of two copies of CP, resulting in a covalent CP dimer in which genetic manipulations such as insertions can be performed in only one monomer, thereby reducing the density of insertions twofold. This way, MS2 VLPs tolerate insertions in AB loop up to 10aa in length (Peabody 1997b; Peabody et al. 2008). Genetic fusions with longer peptides can be achieved much more easily in the case of bacteriophage AP205 VLPs (Tissot et al. 2010) because, compared to MS2 or Q $\beta$ , AP205 CP has both its C- and N-termini exposed on the VLP surface (Shishovs et al. 2016), as discussed in the previous section on VLP structure.

Tolerability to longer insertions in AP205 has enabled the development of another coupling technique known as the spy-catcher approach. This technique relies on the use of the engineered collagen adhesion domain CnaB2 from *Streptococcus pyogenes*(Zakeri et al. 2012). In the CnaB2, a covalent isopeptide bond is autocatalytically formed between the Lys31 and Asp117 residues. It is possible to split the CnaB2 protein into two parts, one peptide with only 13 residues containing Lys31 and another containing the remaining 116 residues of the protein, including Asp117. When mixed together, the two parts spontaneously form covalent isopeptide bonds, similar to native proteins(Li et al. 2014). The 13 residue-long spy tag peptide can be added to the antigen of interest, while the longer spy-catcher sequence can be genetically fused to the C-terminus of AP205 CP, which leaves the VLPs relatively intact. Then, the two components can be mixed together, and covalent bonds between them form autocatalytically within a few minutes(Brune et al. 2016). Using this technique, a variety of antigens have been coupled to AP205 VLPs targeting malaria, HPV, cancer antigens and allergy-associated self-antigens (Brune et al. 2016; Thrane et al. 2016; Janitzek et al. 2019).

Although the genetic coupling technique is superior in the sense that only one round of protein expression and purification has to be performed, it fails frequently due to the formation of an insoluble product or failure to assemble into VLPs. For this reason, VLPs and antigens of choice are very often produced and purified separately and then linked together by chemical coupling. In many cases, amine-to-sulfhydryl cross-linkers, such as succinimidyl 6-((beta-maleimidopropionamido)hexanoate are used (Marcinkiewicz et al. 2018). Succinimide moiety reacts with the free amino groups of lysines or N-termini on the VLP surface, while maleimide reacts with the free sulfhydryl groups of the antigen. If a protein antigen does not contain free surface-exposed cysteines, they can be introduced by genetic



engineering, or existing surface lysines can be modified with the SATA (N-succinimidyl S-acetylthioacetate) reagent, which attaches free sulfhydryls to amines (Bachmann et al. 2018). Other chemical coupling methods can also be used, such as click chemistry (Polonskaya et al. 2017).

If chemical coupling is used, the antigens do not need to be proteins or peptides; for example, it has been shown that high levels of high-affinity anti-nicotine antibodies can be produced by immunization with nicotine, chemically coupled to the surface of bacteriophage Q $\beta$ , potentially acting as a cure for nicotine addiction (Maurer et al. 2005).

### 13.5.8 MS2 Display

MS2 display is a development that is somewhat similar to classical phage display techniques. As discussed previously, short peptide sequences can be inserted into the AB surface loop of the MS2 coat protein without compromising the stability of VLPs. Therefore, a library of VLPs can be created with randomized 6-10-residue-long sequences in AB loops (Peabody et al. 2008). Furthermore, the obtained VLPs can be “fished” out with a target object displaying affinity to the particular peptide within the AB loop. Crucially, the VLPs of ssRNA phages are known to package their own mRNA. Therefore, after affinity selection, mRNA from the “fished” modified VLPs can be easily extracted, reverse transcribed and sequenced, revealing the sequences of the peptides in AB loops. MS2 display can be used to screen for peptides with high affinity to monoclonal antibodies (Chackerian et al. 2011). In this case, the affinity-selected VLPs can be further directly used as vaccine candidates. Using the monoclonal antibody 5A8, displaying high activity in *Plasmodium falciparum* growth inhibition assays, Ord et al. were able to create a VLP vaccine candidate for malaria (Ord et al. 2014). Using a similar approach involving monoclonal antibodies against virulence factor AIP4 from *Staphylococcus aureus*, the same research group was able to create a VLP vaccine candidate for *Staphylococcus aureus* infection (O’Rourke, et al. 2015).

## References

- Aanei IL, ElSohly AM, Farkas ME, Netirojjanakul C, Regan M, Murphy ST, O’Neil JP, Seo Y, Francis MB (2016) Biodistribution of antibody-MS2 viral capsid conjugates in breast cancer models. *Mol Pharm* 13:3764–3772
- Aanei IL, Huynh T, Seo Y, Francis MB (2018) Vascular cell adhesion molecule-targeted MS2 viral capsids for the detection of early-stage atherosclerotic plaques. *Bioconjug Chem* 29:2526–2530
- Adhin MR, van Duin J (1990) Scanning model for translational reinitiation in eubacteria. *J Mol Biol* 213:811–818

- Asensio MA, Morella NM, Jakobson CM, Hartman EC, Glasgow JE, Sankaran B, Zwart PH, Tullman-Ereck D (2016) 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* 16:5944–5950
- Ashley CE, Carnes EC, Phillips GK, Durfee PN, Buley MD, Lino CA, Padilla DP, Phillips B, Carter MB, Willman CL et al (2011) Cell-specific delivery of diverse cargos by bacteriophage MS2 virus-like particles. *ACS Nano* 5:5729–5745
- Atkins JF, Steitz JA, Anderson CW, Model P (1979) Binding of mammalian ribosomes to MS2 phage RNA reveals an overlapping gene encoding a lysis function. *Cell* 18:247–256
- Bachmann MF, Zinkernagel RM (1996) The influence of virus structure on antibody responses and virus serotype formation. *Immunol Today* 17:553–558
- Bachmann MF, Rohrer UH, Kundig TM, Burki K, Hengartner H, Zinkernagel RM (1993) The influence of antigen organization on B cell responsiveness. *Science* 262:1448–1451
- Bachmann MF, Storni T, Maurer P, Tissot A, Schwarz K, Meijerink E, Lipowsky G, Pumpens P, Cielens I, Renhofa R (2003) Packaging of immunostimulatory cpG into virus-like particles: method of preparation and use. In: Office EP (ed) WIPO patent: WO2004084940. <https://patentscope.wipo.int/search/en/detail.jsf?docId=WO2004084940>
- Bachmann MF, Zeltins A, Kalnins G, Balke I, Fischer N, Rostaher A, Tars K, Favrot C (2018) Vaccination against IL-31 for the treatment of atopic dermatitis in dogs. *J Allergy Clin Immunol* 142:279–281
- Bardwell VJ, Wickens M (1990) Purification of RNA and RNA-protein complexes by an R17 coat protein affinity method. *Nucleic Acids Res* 18:6587–6594
- Bernhardt TG, Wang IN, Struck DK, Young R (2001) A protein antibiotic in the phage Qbeta virion: diversity in lysis targets. *Science* 292:2326–2329
- Bertrand E, Chartrand P, Schaefer M, Shenoy SM, Singer RH, Long RM (1998) Localization of ASH1 mRNA particles in living yeast. *Mol Cell* 2:437–445
- Blumenthal T, Landers TA, Weber K (1972) Bacteriophage Qb replicase contains the protein biosynthesis elongation factors EF Tu and EF Ts. *Proc Natl Acad Sci USA* 69:1313–1317
- Bradley DE (1966) The structure and infective process of a *Pseudomonas aeruginosa* bacteriophage containing ribonucleic acid. *J Gen Microbiol* 45:83–96
- Brune KD, Leneghan DB, Brian IJ, Ishizuka AS, Bachmann MF, Draper SJ, Biswas S, Howarth M (2016) Plug-and-display: decoration of virus-like particles via isopeptide bonds for modular immunization. *Sci Rep* 6:19234
- Carrico ZM, Romanini DW, Mehl RA, Francis MB (2008) Oxidative coupling of peptides to a virus capsid containing unnatural amino acids. *Chem Commun*:1205–1207
- Chackerian B, Caldeira JD, Peabody J, Peabody DS (2011) Peptide epitope identification by affinity selection on bacteriophage MS2 virus-like particles. *J Mol Biol* 409:225–237
- Chamakura K, Young R (2019) Phage single-gene lysis: finding the weak spot in the bacterial cell wall. *J Biol Chem* 294:3350–3358
- Chao JA, Patskovsky Y, Almo SC, Singer RH (2008) Structural basis for the coevolution of a viral RNA-protein complex. *Nat Struct Mol Biol* 15:103–105
- Chetverina HV, Chetverin AB (1993) Cloning of RNA molecules in vitro. *Nucleic Acids Res* 21:2349–2353
- Clarke M, Maddera L, Harris RL, Silverman PM (2008) F-pili dynamics by live-cell imaging. *Proc Natl Acad Sci USA* 105:17978–17981
- Coetzee JN, Bradley DE, Hedges RW, Fleming J, Lecatsas G (1983) Bacteriophage M: an incompatibility group M plasmid-specific phage. *J Gen Microbiol* 129:2271–2276
- Cohen BA, Bergkvist M (2013) Targeted in vitro photodynamic therapy via aptamer-labeled, porphyrin-loaded virus capsids. *J Photochem Photobiol B* 121:67–74
- Crawford EM, Gesteland RF (1964) The adsorption of bacteriophage R17. *Virology* 22:165–167
- Crossey E, Amar MJA, Sampson M, Peabody J, Schiller JT, Chackerian B, Remaley AT (2015) A cholesterol-lowering VLP vaccine that targets PCSK9. *Vaccine* 33:5747–5755

- Cui Z, Gorzelnik KV, Chang JY, Langlais C, Jakana J, Young R, Zhang J (2017) 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* 114:11697–11702
- Dai X, Li Z, Lai M, Shu S, Du Y, Zhou ZH, Sun R (2017) In situ structures of the genome and genome-delivery apparatus in a single-stranded RNA virus. *Nature* 541:112–116
- Danziger RE, Paranchych W (1970) Stages in phage R17 infection III Energy requirements for the F-pili mediated eclipse of viral infectivity. *Virology* 40:554–564
- de Martin GN, Crone MA, Ramlaul K, Simpson PA, Freemont PS, Aylett CHS (2019) Bacteriophage MS2 displays unreported capsid variability assembling T = 4 and mixed capsids. *Mol Microbiol Dent KC*, Thompson R, Barker AM, Hiscox JA, Barr JN, Stockley PG, Ranson NA (2013) The asymmetric structure of an icosahedral virus bound to its receptor suggests a mechanism for genome release. *Structure* 21:1225–1234
- ElSohly AM, Netirojjanakul C, Aanei IL, Jager A, Bendall SC, Farkas ME, Nolan GP, Francis MB (2015) Synthetically modified viral capsids as versatile carriers for use in antibody-based cell targeting. *Bioconjug Chem* 26:1590–1596
- Elsohly AM, MacDonald JI, Hentzen NB, Aanei IL, El Muslemany KM, Francis MB (2017) Ortho-Methoxyphenols as convenient oxidative bioconjugation reagents with application to site-selective Heterobifunctional cross-linkers. *J Am Chem Soc* 139:3767–3773
- Fiedler JD, Brown SD, Lau JL, Finn MG (2010) RNA-directed packaging of enzymes within virus-like particles. *Angew Chem Int Ed Engl* 49:9648–9651
- Fiers W, Contreras R, Duerinck F, Haegeman G, Iserentant D, Merregaert J, Min Jou W, Molemans F, Raeymaekers A, Vandenberghe A et al (1976) Complete nucleotide sequence of bacteriophage MS2 RNA: primary and secondary structure of the replicase gene. *Nature* 260:500–507
- Finbloom JA, Aanei IL, Bernard JM, Klass SH, Elledge SK, Han K, Ozawa T, Nicolaides TP, Berger MS, Francis MB (2018) Evaluation of three morphologically distinct virus-like particles as nanocarriers for convection-enhanced drug delivery to Glioblastoma. *Nanomaterials*:8
- Freivalds J, Kotelovica S, Voronkova T, Ose V, Tars K, Kazaks A (2014) Yeast-expressed bacteriophage-like particles for the packaging of nanomaterials. *Mol Biotechnol* 56:102–110
- Galaway FA, Stockley PG (2013) MS2 Viruslike particles: a robust, semisynthetic targeted drug delivery platform. *Mol Pharm* 10:59–68
- Giessen TW, Silver PA (2016) A catalytic Nanoreactor based on in vivo encapsulation of multiple Enzymes in an engineered protein Nanocompartment. *Chembiochem* 17:1931–1935
- Glasgow JE, Capelhart SL, Francis MB, Tullman-Ercek D (2012) Osmolyte-mediated encapsulation of proteins inside MS2 viral capsids. *ACS Nano* 6:8658–8664
- Glasgow JE, Asensio MA, Jakobson CM, Francis MB, Tullman-Ercek D (2015) Influence of electrostatics on small molecule flux through a protein Nanoreactor. *ACS Synth Biol* 4:1011–1019
- Goessens WHF, Driessen AJM, Wilschut J, van Duin J (1988) A synthetic peptide corresponding to the C-terminal 25 residues of phage MS2 coded lysis protein dissipates the protonmotive force in *Escherichia coli* membrane vesicles by generating hydrophilic pores. *EMBO J* 7:867–873
- Golmohammadi R, Valegård K, Fridborg K, Liljas L (1993) The refined structure of bacteriophage MS2 at 2.8 Å resolution. *J Mol Biol* 234:620–639
- Golmohammadi R, Fridborg K, Bundule M, Valegård K, Liljas L (MS2 co-authors) (1996) The crystal structure of bacteriophage Qb at 3.5 Å resolution. *Structure* 4:543–554
- Gorzelnik KV, Cui Z, Reed CA, Jakana J, Young R, Zhang J (2016) Asymmetric cryo-EM structure of the canonical Allovivivirus Qbeta reveals a single maturation protein and the genomic ssRNA in situ. *Proc Natl Acad Sci USA* 113:11519–11524
- Gralla J, Steitz JA, Crothers DM (1974) Direct physical evidence for secondary structure in an isolated fragment of R17 bacteriophage mRNA. *Nature* 248:204–208
- Groeneveld H, Thimon K, van Duin J (MS2 co-authors) (1995) Translational control of maturation-protein synthesis in phage MS2: a role for the kinetics of RNA folding. *RNA* 1:79–88
- Haruna I, Spiegelman S (1965) Specific template requirements of RNA replicases. *Proc Natl Acad Sci USA* 54:579–587

- Hillman BI, Cai G (2013) The family narnaviridae: simplest of RNA viruses. *Adv Virus Res* 86:149–176
- Hocine S, Raymond P, Zenklusen D, Chao JA, Singer RH (2013) Single-molecule analysis of gene expression using two-color RNA labeling in live yeast. *Nat Methods* 10:119–121
- Hofstetter H, Monstein H, Weissmann C. (Qb co-authors). (1974) The readthrough protein A1 is essential for the formation of viable Qb particles. *Biochim Biophys Acta* 374:238–251
- Janitzek CM, Peabody J, Thrane S, Carlsen PHR, Theander TG, Salanti A, Chackerian B, Nielsen MA, Sander AF (2019) A proof-of-concept study for the design of a VLP-based combinatorial HPV and placental malaria vaccine. *Sci Rep*:9
- Jegerlehner A, Zabel F, Langer A, Dietmeier K, Jennings GT, Saudan P, Bachmann MF (2013) Bacterially produced recombinant influenza vaccines based on virus-like particles. *PLoS One* 8:e78947
- Karnik S, Billeter M. (Qb co-authors). (1983) The lysis function of RNA bacteriophage Qb is mediated by the maturation (A2) protein. *EMBO J* 2:1521–1526
- Kastelein RA, Berkhout B, Overbeek GP, van Duin J (1983) Effect of the sequences upstream from the ribosome-binding site on the yield of protein from the cloned gene for phage MS2 coat protein. *Gene* 23:245–254
- Kazaks A, Voronkova T, Rumnieks J, Dishlers A, Tars K (2011) Genome structure of caulobacter phage phiCb5. *J Virol* 85:4628–4631
- Kerfeld CA, Aussignargues C, Zarzycki J, Cai F, Sutter M (2018) Bacterial microcompartments. *Nat Rev Microbiol* 16:277–290
- Klovins J, Overbeek GP, van den Worm SH, Ackermann HW, van Duin J (2002) Nucleotide sequence of a ssRNA phage from *Acinetobacter*: kinship to coliphages. *J Gen Virol* 83:1523–1533
- Kolesanova EF, Melnikova MV, Bolshakova TN, Rybalkina EY, Sivov IG (2019) Bacteriophage MS2 as a tool for targeted delivery in solid tumor chemotherapy. *Acta Nat* 11:98–101
- Koning RI, Gomez-Blanco J, Akopjana I, Vargas J, Kazaks A, Tars K, Carazo JM, Koster AJ (2016) Asymmetric cryo-EM reconstruction of phage MS2 reveals genome structure in situ. *Nat Commun* 7:12524
- Kovacs EW, Hooker JM, Romanini DW, Holder PG, Berry KE, Francis MB (2007) Dual-surface-modified bacteriophage MS2 as an ideal scaffold for a viral capsid-based drug delivery system. *Bioconjug Chem* 18:1140–1147
- Kozlovskaya TM, Cielens I, Dreilina D, Dislers A, Baumanis V, Ose V, Pumpens P (1993) Recombinant RNA phage Q beta capsid particles synthesized and self-assembled in *Escherichia coli*. *Gene* 137:133–137
- Kozlovskaya TM, Pumpens P, Dreilina DE, Tsimanis AJ, Ose VP, Tsibinogin VV, Gren EJ (1986) Formation of capsid-like structures as a result of expression of coat protein gene of RNA phage fr. *Dokl. Akad. Nauk SSSR* 287:452–455
- Krahn PM, O'Callaghan RJ, Paranchych W (1972) Stages in phage R17 infection. VI Injection of A protein and RNA into the host cell. *Virology* 47:628–637
- Krishnamurthy SR, Janowski AB, Zhao G, Barouch D, Wang D (2016) Hyperexpansion of RNA bacteriophage diversity. *PLoS Biol* 14:e1002409
- Larson DR, Zenklusen D, Wu B, Chao JA, Singer RH (2011) Real-time observation of transcription initiation and elongation on an endogenous yeast gene. *Science* 332:475–478
- Li L, Fierer JO, Rapoport TA, Howarth M (2014) Structural analysis and optimization of the covalent association between SpyCatcher and a peptide tag. *J Mol Biol* 426:309–317
- Licis N, Balklava Z, van Duin J (2000) Forced retroevolution of an RNA bacteriophage. *Virology* 271:298–306
- Liekina I, Kalnins G, Akopjana I, Bogans J, Sisovs M, Jansons J, Rumnieks J, Tars K (2019) Production and characterization of novel ssRNA bacteriophage virus-like particles from metagenomic sequencing data. *J Nanobiotechnol* 17:61
- Liljas L, Fridborg K, Valegård K, Bundule M, Pumpens P (1994) Crystal structure of bacteriophage fr capsids at 3.5 Å resolution. *J Mol Biol* 244:279–290

- Ling CM, Hung PP, Overby LR. (Q-Beta co-authors). (1970) Independent assembly of Q $\beta$  and MS2 phages in doubly infected *Escherichia coli*. *Virology* 40:920–929
- Marcinkiewicz AL, Lieknina I, Kotelovica S, Yang X, Kraiczy P, Pal U, Lin YP, Tars K (2018) Eliminating factor H-binding activity of *Borrelia burgdorferi* CspZ combined with virus-like particle conjugation enhances its efficacy as a Lyme disease vaccine. *Front Immunol* 9:181
- Maurer P, Jennings GT, Willers J, Rohner F, Lindman Y, Roubicek K, Renner WA, Muller P, Bachmann MF (2005) A therapeutic vaccine for nicotine dependence: preclinical efficacy, and phase I safety and immunogenicity. *Eur J Immunol* 35:2031–2040
- Meng R, Jiang M, Cui Z, Chang JY, Yang K, Jakana J, Yu X, Wang Z, Hu B, Zhang J (2019) Structural basis for the adsorption of a single-stranded RNA bacteriophage. *Nat Commun* 10:3130
- Olsen RH, Thomas DD (1973) Characteristics and purification of PRR1, an RNA phage specific for the broad host range *Pseudomonas* R1822 drug resistance plasmid. *J Virol* 12:1560
- Olsthoorn RCL, Garde G, Dayhuff T, Atkins JF, van Duin J (MS2 co-authors) (1995) Nucleotide sequences of a single-stranded RNA phage from *Pseudomonas aeruginosa*: kinship to coliphages and conservation of regulatory RNA structures. *Virology* 206:611–625
- Ord RL, Caldeira JC, Rodriguez M, Noe A, Chackerian B, Peabody DS, Gutierrez G, Lobo CA (2014) A malaria vaccine candidate based on an epitope of the *Plasmodium falciparum* RH5 protein. *Malar J* 13:326
- O'Rourke JP, Daly SM, Hall PR (2015) Development of a Mimotope vaccine targeting the *Staphylococcus aureus* quorum sensing pathway (vol 9, e111198, 2014). *PLoS One*:10
- Park SY, Moon HC, Park HY (2020 Jan) Live-cell imaging of single mRNA dynamics using split superfolder green fluorescent proteins with minimal background. *RNA* 26(1):101–109
- Pasloske BL, Walkerpeach CR, Obermoeller RD, Winkler M, DuBois DB (1998) Armored RNA technology for production of ribonuclease-resistant viral RNA controls and standards. *J Clin Microbiol* 36:3590–3594
- Peabody DS (1990) Translational repression by bacteriophage MS2 coat protein expressed from a plasmid. Development of a Mimotope vaccine targeting the *Staphylococcus aureus* quorum sensing pathway. *J Biol Chem* 265:5684–5689
- Peabody DS (1997a) Role of the coat protein-RNA interaction in the life cycle of bacteriophage MS2. *Mol GenGenet* 254:358–364
- Peabody DS (1997b) Subunit fusion confers tolerance to peptide insertions in a virus coat protein. *Arch Biochem Biophys* 347:85–92
- Peabody DS, Manifold-Wheeler B, Medford A, Jordan SK, do Carmo Caldeira J, Chackerian B (2008) Immunogenic display of diverse peptides on virus-like particles of RNA phage MS2. *J Mol Biol* 380:252–263
- Persson M, Tars K, Liljas L (2008) The capsid of the small RNA phage PRR1 is stabilized by metal ions. *J Mol Biol* 383:914–922
- Pickett GG, Peabody DS (1993) Encapsulation of heterologous RNAs by bacteriophage MS2 coat protein. *Nucl Acids Res* 21:4621–4626
- Plevka P, Kazaks A, Voronkova T, Kotelovica S, Dishlers A, Liljas L, Tars K (2009) The structure of bacteriophage phiCb5 reveals a role of the RNA genome and metal ions in particle stability and assembly. *J Mol Biol* 391:635–647
- Polonskaya Z, Deng SL, Sarkar A, Kain L, Comellas-Aragones M, McKay CS, Kaczanowska K, Holt M, McBride R, Palomo V et al (2017) T cells control the generation of nanomolar-affinity anti-glycan antibodies. *J Clin Invest* 127:1491–1504
- Pumpens P, Renhofa R, Dishlers A, Kozlovska T, Ose V, Pushko P, Tars K, Grens E, Bachmann MF (2016) The true story and advantages of RNA phage capsids as Nanotools. *Intervirology* 59:74–110
- Rastgou M, Habibi MK, Izadpanah K, Masenga V, Milne RG, Wolf YI, Koonin EV, Turina M (2009) Molecular characterization of the plant virus genus *Ourmiavirus* and evidence of interkingdom reassortment of viral genome segments as its possible route of origin. *J Gen Virol* 90:2525–2535

- Rhee JK, Baksh M, Nycholat C, Paulson JC, Kitagishi H, Finn MG (2012) Glycan-targeted virus-like nanoparticles for photodynamic therapy. *Biomacromolecules* 13:2333–2338
- Rolfsson O, Middleton S, Manfield IW, White SJ, Fan B, Vaughan R, Ranson NA, Dykeman E, Twarock R, Ford J et al (2016) Direct evidence for packaging signal-mediated assembly of bacteriophage MS2. *J Mol Biol* 428:431–448
- Rumnieks J, Tars K (2011) Crystal structure of the read-through domain from bacteriophage Qbeta A1 protein. *Protein Sci* 20:1707–1712
- Rumnieks J, Tars K (2012) Diversity of pili-specific bacteriophages: genome sequence of IncM plasmid-dependent RNA phage M. *BMC Microbiol* 12:277
- Rumnieks J, Tars K (2014) Crystal structure of the bacteriophage qbeta coat protein in complex with the RNA operator of the replicase gene. *J Mol Biol* 426:1039–1049
- Rumnieks J, Tars K (2017) Crystal structure of the maturation protein from bacteriophage Qbeta. *J Mol Biol* 429:688–696
- Rumnieks J, Tars K (2018) Protein-RNA interactions in the single-stranded RNA bacteriophages. *Subcell Biochem* 88:281–303
- Schmidt JM (1966) Observations on the adsorption of Caulobacter bacteriophages containing ribonucleic acid. *J Gen Microbiol* 45:347–353
- Shi M, Lin XD, Tian JH, Chen LJ, Chen X, Li CX, Qin XC, Li J, Cao JP, Eden JS et al (2016, Dec 22) Redefining the invertebrate RNA virosphere. *Nature* 540(7634):539–543
- Shiba T, Miyake T (1975) New type of infectious complex of *E. coli* RNA phage. *Nature* 254:157–158
- Shishovs M, Rumnieks J, Diebold C, Jaudzems K, Andreas LB, Stanek J, Kazaks A, Kotelovica S, Akopjana I, Pintacuda G et al (2016) Structure of AP205 coat protein reveals circular permutation in ssRNA bacteriophages. *J Mol Biol* 428:4267–4279
- Skripkin EA, Adhin MR, de Smit MH, van Duin J (1990) Secondary structure of bacteriophage MS2. Conservation and biological significance. *J Mol Biol* 211:447–463
- Spohn G, Jennings GT, Martina BE, Keller I, Beck M, Pumpens P, Osterhaus AD, Bachmann MF (2010) A VLP-based vaccine targeting domain III of the West Nile virus E protein protects from lethal infection in mice. *Virology* 7:146
- Tars K, Bundule M, Fridborg K, Liljas L (1997) The crystal structure of bacteriophage GA and a comparison of bacteriophages belonging to the major groups of *Escherichia coli* leviviruses. *J Mol Biol* 271:759–773
- Tars K, Fridborg K, Bundule M, Liljas L (2000) Crystal structure of phage PP7 from *Pseudomonas aeruginosa* at 3.7 Å resolution. *Virology* 272:331–337
- Tars K, Zeltins A, Liljas L (2003) The three-dimensional structure of cocksfoot mottle virus at 2.7 Å resolution. *Virology* 310:287–297
- Thrane S, Janitzek CM, Matondo S, Resende M, Gustavsson T, de Jongh WA, Clemmensen S, Roeffen W, van de Vegte-Bolmer M, van Gemert GJ et al (2016) Bacterial superglue enables easy development of efficient virus-like particle based vaccines. *J Nanobiotechnol* 14:30
- Tissot AC, Maurer P, Nussberger J, Sabat R, Pfister T, Ignatenko S, Volk HD, Stocker H, Muller P, Jennings GT et al (2008) Effect of immunisation against angiotensin II with CYT006-AngQb on ambulatory blood pressure: a double-blind, randomised, placebo-controlled phase IIa study. *Lancet* 371:821–827
- Tissot AC, Renhofa R, Schmitz N, Cielens I, Meijerink E, Ose V, Jennings GT, Saudan P, Pumpens P, Bachmann MF (2010) Versatile virus-like particle carrier for epitope based vaccines. *PLoS One* 5:e9809
- Tomita K (2014) Structures and functions of Qbeta replicase: translation factors beyond protein synthesis. *Int J Mol Sci* 15:15552–15570
- Tsai BP, Wang X, Huang L, Waterman ML (2011) Quantitative profiling of in vivo-assembled RNA-protein complexes using a novel integrated proteomic approach. *Mol Cell Proteomics* 10(M110):007385

- Tumban E, Peabody J, Tyler M, Peabody DS, Chackerian B (2012) VLPs displaying a single L2 epitope induce broadly cross-neutralizing antibodies against human papillomavirus. *PLoS One* 7:e49751
- Ugarov VI, Chetverin AB (2008) Functional circularity of legitimate Qbeta replicase templates. *J Mol Biol* 379:414–427
- Valegård K, Liljas L, Fridborg K, Unge T (1990) The three-dimensional structure of the bacterial virus MS2. *Nature* 345:36–41
- Valegård K, Murray JB, Stockley PG, Stonehouse NJ, Liljas L (1994) Crystal structure of an RNA bacteriophage coat protein-operator complex. *Nature* 371:623–626
- Valegård K, Murray JB, Stonehouse NJ, van den Worm S, Stockley PG, Liljas L (1997) The three-dimensional structures of two complexes between recombinant MS2 capsids and RNA operator fragments reveal sequence-specific protein-RNA interactions. *J Mol Biol* 270:724–738
- Valentine RC, Strand M (1965) Complexes of F-pili and RNA bacteriophage. *Science* 148:511–513
- Van Duin J, Tsareva N (2006) Single-stranded RNA Phages. In: Calendar R (ed) *The bacteriophages*. Oxford University Press, New York, pp 175–196
- van Meerten D, Girard G, van Duin J (2001) Translational control by delayed RNA folding: identification of the kinetic trap. *RNA* 7:483–494
- Wahba AJ, Miller MJ, Niveleau A, Landers TA, Carmichael GG, Weber K, Hawley DA, Slobin LI (1974) Subunit I of Qbeta replicase and 30 S ribosomal protein S1 of *Escherichia coli* Evidence for the identity of the two proteins. *J Biol Chem* 249:3314–3316
- Wang GJ, Jia TT, Xu XX, Chang L, Zhang R, Fu Y, Li YL, Yang X, Zhang K, Lin GG et al (2016) Novel miR-122 delivery system based on MS2 virus like particle surface displaying cell-penetrating peptide TAT for hepatocellular carcinoma. *Oncotarget* 7:59402–59416
- Weber H (1976) The binding site for coat protein on bacteriophage Qb RNA. *Biochim Biophys Acta* 418:175–183
- Wei B, Wei YX, Zhang K, Wang J, Xu RH, Zhan S, Lin GG, Wang W, Liu M, Wang JN et al (2009) Development of an antisense RNA delivery system using conjugates of the MS2 bacteriophage capsids and HIV-1 TAT cell penetrating peptide. *Biomed Pharmacother* 63:313–318
- Weiner AM, Weber K (1971) Natural read-through at the UGA termination signal of Q-beta coat protein cistron. *Nature New Biol* 234:206–209
- Winter RB, Gold L. (Qb co-authors). (1983) Overproduction of bacteriophage Qb maturation (A2) protein leads to cell lysis. *Cell* 33:877–885
- Wu M, Brown WL, Stockley PG (1995) Cell-specific delivery of bacteriophage-Encapsidated ricin-a chain. *Bioconjug Chem* 6:587–595
- Wu B, Chen J, Singer RH (2014) Background free imaging of single mRNAs in live cells using split fluorescent proteins. *Sci Rep* 4:3615
- Wu MJ, Andreasson JOL, Kladowang W, Greenleaf W, Das R (2019) Automated Design of Diverse Stand-Alone Riboswitches. *ACS Synth Biol* 8:1838–1846
- Yao L, Li F, Qu M, Guo Y, Jiang Y, Wang L, Zhai Y (2019) Development and evaluation of a novel Armored RNA technology using bacteriophage Qbeta. *Food Environ Virol* 11:383–392
- Yin Z, Chowdhury S, McKay C, Baniel C, Wright WS, Bentley P, Kaczanowska K, Gildersleeve JC, Finn MG, BenMohamed L et al (2015) Significant impact of Immunogen design on the diversity of antibodies generated by carbohydrate-based anticancer vaccine. *ACS Chem Biol* 10:2364–2372
- Yin Z, Dulaney S, McKay CS, Baniel C, Kaczanowska K, Ramadan S, Finn MG, Huang X (2016) Chemical synthesis of GM2 Glycans, Bioconjugation with Bacteriophage Qbeta, and the induction of anticancer antibodies. *ChemBiochem* 17:174–180
- Yoon JH, Gorospe M (2016) Identification of mRNA-interacting factors by MS2-TRAP (MS2-tagged RNA affinity purification). *Methods Mol Biol* 1421:15–22
- Yoon JH, Srikantan S, Gorospe M (2012) MS2-TRAP (MS2-tagged RNA affinity purification): tagging RNA to identify associated miRNAs. *Methods* 58:81–87

- Zakeri B, Fierer JO, Celik E, Chittock EC, Schwarz-Linek U, Moy VT, Howarth M (2012) Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. *Proc Natl Acad Sci USA* 109:E690–E697
- Zhai L, Yadav R, Kunda NK, Anderson D, Bruckner E, Miller EK, Basu R, Muttill P, Tumban E (2019) Oral immunization with bacteriophage MS2-L2 VLPs protects against oral and genital infection with multiple HPV types associated with head & neck cancers and cervical cancer. *Antivir Res* 166:56–65
- Zhang L, Sun Y, Chang L, Jia T, Wang G, Zhang R, Zhang K, Li J (2015) A novel method to produce armored double-stranded DNA by encapsulation of MS2 viral capsids. *Appl Microbiol Biotechnol* 99:7047–7057
- Zhao L, Kopylov M, Potter CS, Carragher B, Finn MG (2019) Engineering the PP7 virus capsid as a peptide display platform. *ACS Nano* 13:4443–4454