

# Chapter 10

## Biology, Biodiversity and Application of Thermophilic Viruses

Kristine Uldahl and Xu Peng

**Abstract** Where there is life, there are viruses. Thermophilic viruses are essential in regulating the structure and composition of microbial communities in the terrestrial and marine hot environments. Moreover, they constitute an important source of novel enzymes of high biotechnological and industrial potential. This chapter focuses on the application potential of thermophilic viruses. Central to this topic are the many novel genetic and morphological features discovered in thermophilic viruses. In order to fully appreciate the huge potential these viruses have in basic research and biotechnology, we will introduce recent results derived from studies of structural adaptations, functional genomics, metagenomics, virus–host interactions and virus life cycle, all areas with remarkable characteristics. In the end, a section will cover more specific application areas including thermophilic viruses as nano-building blocks.

**Keywords** Thermophilic viruses • Genomics • Life cycle • Biotechnology • Application

### 10.1 Introduction

Since the first discovery of microorganisms in the boiling hot springs of Yellowstone National Park by Thomas Brock in 1966, many thermophilic organisms including bacteria, Archaea and fungi have been isolated from different hot environments.

---

K. Uldahl • X. Peng (✉)

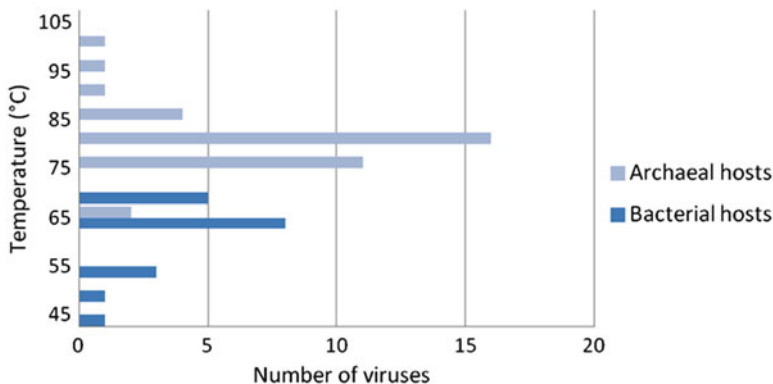
Archaea Centre, Department of Biology, University of Copenhagen,  
Ole Maaløes Vej 5, DK-2200 KBH N, Denmark  
e-mail: kristine.uldahl@bio.ku.dk; peng@bio.ku.dk

Where there is life, there are viruses. Viruses infecting thermophilic microorganisms are abundant in terrestrial and marine hot environments and have been isolated from a range of hot environments all over the world, including hydrothermal vents, acidic and neutral hot springs and self-heating organic material such as compost.

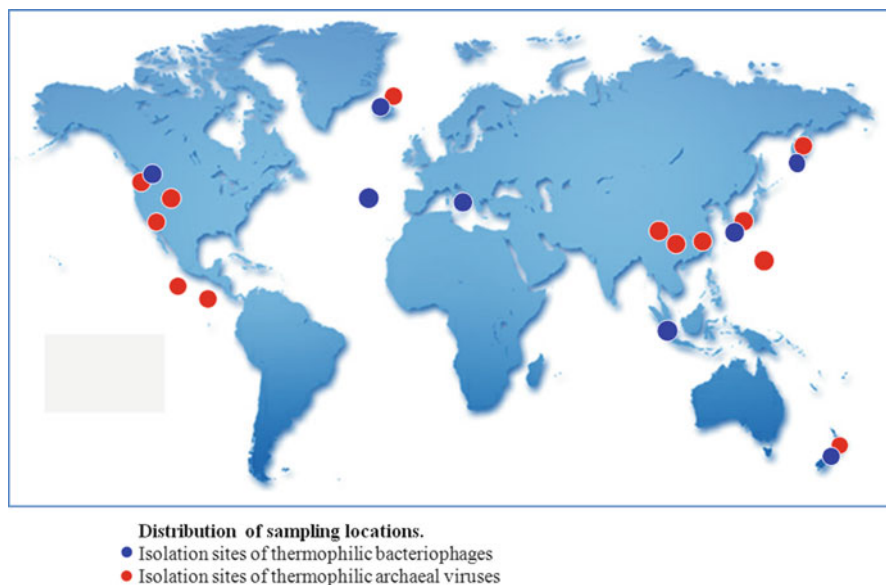
Viruses constitute a major component of the biosphere and play a significant role in nutrient and energy cycling of carbon, nitrogen and phosphorus (Krupovic et al. 2011). Moreover, viruses have important impacts on the evolution of their hosts. For example, viruses can influence the genetic diversity of prokaryotes by horizontal gene transfer among different species. As killers of bacteria and Archaea, viruses play an essential role in regulating the structure and composition of microbial communities (Weinbauer and Rassoulzadegan 2004). Importantly, thermophilic viruses can serve as model systems for studying life at high temperatures, and they constitute a valuable source of novel enzymes of high biotechnological and industrial potential.

Viruses are genetic elements that require host resources for their replication. They have single- or double-stranded RNA or DNA genomes ranging in size from a few thousand to a million bases. In order to thrive and multiply, viruses have to transmit from one host to another. When they exit the cell, they have to survive in the natural environment of the host (Madigan and Brock 2009). Thus, viruses have an extracellular form that enables them to exist outside the host for extended periods of time. Thermophilic viruses have adapted to hot environments same as their hosts. Thermophiles refer to organisms that grow at temperatures above 45°C, and if the temperatures are above 80°C, they are called hyperthermophiles. Above 65°C only prokaryotic life forms are found, and Archaea are the more successful domain (Lengeler et al. 1999). Archaea are adapted to conditions of chronic energy stress and have adaptations in membrane composition and metabolic pathways that provide a competitive advantage in a range of extreme environments (Valentine 2007). Although both Archaea and bacteria exist in hot environments, there is a trend that more Archaea and fewer bacteria are found when the temperature increases. This is also reflected by the temperature range of host cells for thermophilic bacteriophages and archaeal viruses. As illustrated by Fig. 10.1, the optimal growth temperatures for hosts of bacteriophages range from 45 to 75°C, while the hosts of archaeal viruses have optimal growth temperatures from 65 to 100°C. This implies that thermophilic bacteriophages and especially thermophilic archaeal viruses must employ some special strategies to cope with the detrimental high temperatures when they are outside of their host. It is worth noting that the number of thermophilic viruses isolated so far is still limited, and the distribution illustrated by Fig. 10.1 may change when more viruses are obtained. Moreover, as seen in Fig. 10.2, the locations where thermophilic viruses have been sampled are few due to the natural obstacles of hot environments. Thus, it is only with the development of advanced equipment that isolation and ecological studies of thermophilic and especially hyperthermophilic viruses have been possible.

This chapter focuses on the application potential of thermophilic viruses. Central to this topic are the many novel genetic and morphological features discovered in



**Fig. 10.1** The optimal growth temperature of thermophilic viral hosts. The graph shows an overview of the number of archaeal viruses and bacteriophages that infect hosts which have an optimal growth temperature exceeding 45°C



**Fig. 10.2** Distribution of sampling locations

thermophilic viruses. In order to fully appreciate the huge potential these viruses have in basic research and biotechnology, we will introduce recent results derived from studies of structural adaptations, functional genomics, metagenomics, virus–host interactions and virus life cycle, all areas with remarkable characteristics. At the end, a section will cover more specific application areas including thermophilic viruses as nanobuilding blocks.

## 10.2 Virion Morphotypes of Thermophilic Viruses

Thermophilic viruses are much less studied compared to the mesophilic viruses. In approximately 5,600 known viruses, only a few have been isolated from thermophiles (Lin et al. 2011). Tables 10.1 and 10.2 present 57 isolated thermophilic viruses, and the majority belong to archaeal genera. Although the number of known thermophilic viruses is limited, the morphological diversity is exceptional and includes unique morphotypes (Prangishvili and Garrett 2005), providing us with unique forms and complex features for possible future applications.

The thermophilic archaeal viruses are extremely diverse and can be categorised into nine distinct families including *Fuselloviridae*, *Bicaudaviridae*, *Ampullaviridae*, *Clavaviridae*, *Lipothrixviridae*, *Rudiviridae*, *Globulaviridae*, *Myoviridae* and *Siphoviridae* and four taxa of uncertain affiliation, represented by five novel viruses: STIV1, STIV2, PAV1, STSV1 (Pina et al. 2011) and TPV1 (Gorlas et al. 2012). Enrichment cultures from hot environments exhibit a range of morphotypes including exceptional forms not previously observed for any dsDNA virus. The newly described morphotypes of archaeal viruses include spindle, droplet and bottle shapes (Arnold et al. 2000a; Mochizuki et al. 2010). These morphotypes are not rare in hot environments, and electron microscopy studies from terrestrial hot springs suggest that spindles, filaments, rods and spheres predominate (Garrett et al. 2010). The thermophilic archaeal viruses have been isolated from several archaeal genera including *Sulfolobus*, *Acidianus*, *Thermoproteus*, *Aeropyrum*, *Stygiolobus*, *Pyrobaculum*, *Methanobacterium*, *Pyrococcus* and *Thermococcus* (Pina et al. 2011). *Pyrococcus* and *Pyrobaculum* have optimal growth temperatures of 100°C demonstrating the extreme temperatures the thermophilic archaeal viruses have to adapt to.

Most of the isolated thermophilic bacteriophages exhibit a head–tail morphology similar to that of a T bacteriophage and belong to the families of *Myoviridae* and *Siphoviridae*. Interestingly, in an extensive survey done by the Promega Corporation in which 115 *Thermus* bacteriophages were isolated, only 45% of the bacteriophages were tailed. This was surprising because tailless bacteriophages of all types comprise only 4% of about 5,150 bacteriophages observed in the electron microscope. Of tailless thermophilic bacteriophages, three have been classified, two belong to the family of *Tectiviridae* (Yu et al. 2006) and one belongs to the family of *Inoviridae*. The thermophilic bacteriophages have been isolated from six bacterial genera including *Bacillus*, *Geobacillus*, *Thermus*, *Meiothermus*, *Rhodothermus* and *Thermomonospora* (Lin et al. 2011).

It is noteworthy that out of the seven described virion morphotypes, the bacterial viruses are represented by three, and the archaeal viruses are represented in all seven morphotype categories.

### 10.2.1 Spindle-Shaped Viruses

Viruses with a spindle- or lemon-shaped morphotype have been observed in both terrestrial and marine hot environments. Viruses with spindle-shape morphology, single or two tailed, are common in and exclusive to the domain of Archaea

**Table 10.1** Morphology, taxonomical classification and characteristics of thermophilic archaeal viruses

Morphology	Family/genus	Virus	Host	Temp. host (°C)	Genome size	Lytic	Isolation site	References
Bacilliform	<i>Clavaviridae</i> <sup>a</sup>	APBV1	<i>Aeropyrum</i>	90	5.2	-	Japan, coastal hot spring	Mochizuki et al. (2010)
Bottle	<i>Ampullaviridae</i>	ABV	<i>Acidianus</i>	75	24	-	Italy, acidic hot spring	Haring et al. (2005)
Droplet	<i>Guttaviridae</i>	SNDV	<i>Sulfolobus</i>	80	20	-	New Zealand	Arnold et al. (2000a)
Head-tail	<i>Siphoviridae</i>	ψM1	<i>Methanobacterium</i>	65	30.4	-	From culture	Meile et al. (1989)
		ψM2		65	26.1	-		Pfister et al. (1998)
Linear	<i>Lipothrixviridae</i>	AFV1	<i>Acidianus</i>	75	21	-	Yellowstone hot spring	Bettstetter et al. (2003)
		AFV2		75	31.7	-	Italy, acidic hot spring	Prangishvili et al. (2006)
		AFV3		75	40.4	-		Vestergaard et al. (2008a)
		AFV6		75	39.5	-	Italy	
		AFV7		75	36.8	-		
		AFV8		75	38.1	-		
		AFV9		75	41.1	-	Russia, acidic hot spring	Bize et al. (2008)
		SIFV	<i>Sulfolobus</i>	80	40.8	-	Iceland, solfataric field	Arnold et al. (2000b)
		TTV1	<i>Thermoproteus</i>	85	15.9	+	Iceland, hot mud hole	Janekovic et al. (1983)
		TTV2		85	-	-	Iceland	
		TTV3		85	-	-		
	<i>Rudiviridae</i>	SIRV1	<i>Sulfolobus</i>	80	32.3	+	Iceland, solfataric field	Zillig et al. (1994)
		SIRV2		80	35.4	+		Prangishvili et al. (1999)
		SIRV	<i>Stygiolobus</i>	80-90	28	-	Azores, hot spring	Vestergaard et al. (2008b)
		ARV1	<i>Acidianus</i>	75	24.6	-	Italy, acidic hot spring	Vestergaard et al. (2005)
Spherical	<i>Globulaviridae</i>	PSV1	<i>Pyrobaculum</i>	100	28.3	-	Yellowstone hot spring	Haring et al. (2004)
		TTSV1	<i>Thermoproteus</i>	85	21.6	-	Indonesia, hot spring	Ahn et al. (2006)
	<i>Unclassified</i>	STIV	<i>Sulfolobus</i>	80	16.6	+	Yellowstone hot spring	Rice et al. (2004)
		STIV2		80	17.6	+	Iceland, hot spring	Happonen et al. (2010)

(continued)

**Table 10.1** (continued)

Morphology	Family/genus	Virus	Host	Temp. host (°C)	Genome size	Lytic	Isolation site	References
Spindle	<i>Fuselloviridae</i>	SSV1	<i>Sulfolobus</i>	80	15.4	-	Japan	Palm et al. (1991)
		SSV2		80	14.7	-	Iceland, hot spring	Stedman et al. (2003)
		SSV4		80	15.1	-	Iceland, hot spring	Peng (2008)
		SSV5		80	15.3	-	Iceland, hot spring	Redder et al. (2009)
		SSV6		80	15.6	-	Iceland, hot spring	
		SSV7		80	17.6	-	Iceland, hot spring	
		SSV/k1		80	17.3	-	Russia, acidic hot spring	Wiedenheft et al. (2004)
		SSVrh		80	16.4	-	Yellowstone hot spring	
		ASV1	<i>Acidianus</i>	75	24.1	-	USA	Redder et al. (2009)
		ATV		75	62.7	+	Italy, acidic hot spring	Prangishvili et al. (2006)
<i>Bicaudaviridae</i>	STSV1	<i>Sulfolobus</i>	80	75.2	-	China, acidic hot spring	Xiang et al. (2005)	
	<i>Unclassified</i>	PAV1	<i>Pyrococcus</i>	95–100	18	-	Deep sea hot vent	Geslin et al. (2003)
		TPV1	<i>Thermococcus</i>	85	21.5	-	Deep sea hot vent	Gorlas et al. (2012)

The temperature of the host refers to the optimum growth temperature of the Bacterium from which the virus was isolated from

Genome size shows the viral genome size in kb. A lytic life cycle state is identified with a (+) and a lysogenic life cycle without apparent lysis of host cells during any stage is identified with a (-)

<sup>a</sup>The approval of the family '*Clavaviridae*' is pending at the ITV

**Table 10.2** Morphology, taxonomical classification and characteristics of thermophilic bacteriophages

Morphology	Family/genus	Virus	Host	Temp. host (°C)	Genome size	Lytic	Isolation site	References
Linear	<i>Inoviridae</i>	PH75	<i>Thermus</i>	70–75	6.5	–	New Zealand, hot spring	Pederson et al. (2001)
Head-tail	<i>Myoviridae</i>	φTMA	<i>Thermus</i>	70–75	151.5	+	Japan, hot spring	Tamakoshi et al. (2011)
		φYS40		65	152.4	+	Japan, hot spring soil	Matsushita and Yanase (2009)
		TS2126		70–75	90		Iceland, hot tap water	Blondal et al. (2005b)
		RM378	<i>Rhodothermus</i>	65	130		Iceland	Blondal et al. (2003)
		MMP17	<i>Meiothermus</i>	50–65	33.5–39.5	+	China, hot spring	Lin et al. (2010)
		D6E	<i>Geobacillus</i>	65	34.7	+	Deep sea hot vent	Wang and Zhang (2010)
		GBSV1		65			China, offshore hot spring	Liu et al. (2009)
		BV1		65	35	+ <sup>a</sup>	China, inshore hot spring	Liu et al. (2010)
	<i>Siphoviridae</i>	GVE2		65–70	40.9	+	Deep sea hot vent	Liu and Zhang (2008)
		TSP4	<i>Thermus</i>	70–75	80	+	China, hot spring	Lin et al. (2010)
		P23-45		70–75	84.2	+	Russia, acidic hot spring	Minakhin et al. (2008)
	<i>Unclassified</i>	P74-26		70–75	83.3	+		
		BVW1	<i>Bacillus</i>	65–70	18	+	Deep sea hot vents	Liu et al. (2006)
		Tb1	<i>Thermomonospora</i>	45	43	+	Compost, self-heated organic material	Lawrence et al. (1986)
		Tf2		55	35	+		
		Tf3		55	45	+		
		Tf4		55	36	+		
Head, no tail	<i>Tectiviridae</i>	φIN93	<i>Thermus</i>	65	19.6	+	Japan, hot spring soil	Matsushita and Yanase (2009)
		P23-77		70–75	17	+	New Zealand, hot spring	Jaatinen et al. (2008)

The temperature of the host refers to the optimal growth temperature of the bacterium from which the virus was isolated from

Genome size shows the viral genome size in kb. A lytic life cycle state is identified with a (+) and a lysogenic life cycle without apparent lysis of host cells during any stage is identified with a (–)

<sup>a</sup>Evidence suggests the bacteriophage is lytic

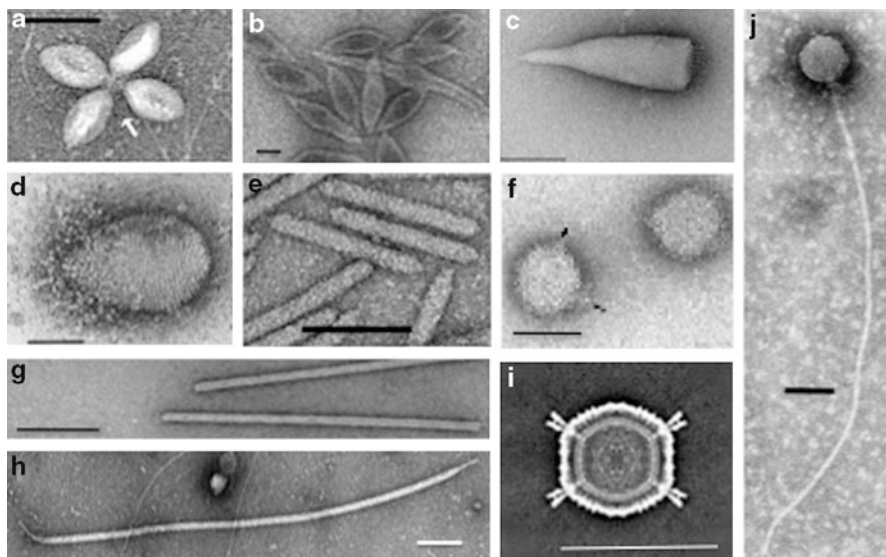
(Garrett et al. 2010). Thermophilic archaeal viruses with spindle shape have been divided into two families *Fuselloviridae* and *Bicaudaviridae*, and three viruses await classification.

*The Fuselloviridae Family.* All the known fuselloviruses infect and propagate in the hyperthermophilic genera *Sulfolobus* and *Acidianus*, including eight *Sulfolobus* spindle-shaped viruses (SSV1, 2, 4, 5, 6, 7, k1 and rh) and one *Acidianus* spindle-shaped virus (ASV1) (Table 10.1). The spindle-shaped virions are approximately 55–60 nm wide and 80–100 nm long (Fig. 10.3a). The exceptions are the viruses SSV6 and ASV1, whose virions tend to be pleomorphic (Redder et al. 2009). At one of the pointed ends of the spindle shape, a short tail of thin fibres is attached, which appear to be extremely sticky and readily attach to cellular fragments. The thin terminal fibres can also attach to the same type of fibres in other virions, which can produce rosette-like aggregates (Redder et al. 2009).

*The Bicaudaviridae Family.* The *Bicaudaviridae* family has only one member, the *Acidianus* two-tailed virus (ATV). The virions are exceptional in that they are extruded from host cells as tailless spindle-shaped particles, which then develop long tails at each pointed end. The tailless and the two-tailed virions have different virion dimensions. The tailless virions exhibit an average length of 243 nm and a maximum width of 119 nm, while the two-tailed virions show an average length of 744 nm and a maximum width of 85 nm (Fig. 10.3b) (Prangishvili et al. 2006). The tail development happens at temperatures above 75°C and independently of host cells or any energy sources (Haring et al. 2005). The only other known examples of extracellular viral morphogenesis comprise initial steps of infection or final steps in particle assembly and budding, and these are triggered on the cell surface of the host. The tails consist of tubes that terminate in an anchor-like structure. One function of the elongated, flexible tails might be to enhance the probability of virion adsorption to a new host cell (Prangishvili et al. 2006).

Three spindle-shaped viruses remain to be classified, *Pyrococcus abyssi* virus 1 (PAV1), *Thermococcus prieurii* virus 1 (TPV1) and *Sulfolobus tengchongensis* virus 1 (STSV1). PAV1 virions are approximately 120 nm long and 80 nm wide with a 15 nm long tail terminating in fibres. It contains a dsDNA of 18 kb (Geslin et al. 2003). STSV1 produces spindle-shaped virions (230 nm × 107 nm) with a single tail of variable length (Xiang et al. 2005). Similar to ATV, the spindle size of STSV1 seems inversely proportional to the length of the tail, that is, the longer the tail of a specific virion, the smaller the spindle is. This might reflect a reorganisation of structural components on the virions. The recently characterised TPV1 has a lemon-shaped virion approximately 140 nm long and 80 nm wide with a 15 nm long tail terminating in fibres, similar to the structure observed for PAV1. TPV1 contains a dsDNA of 21.5 kb (Gorlas et al. 2012).





**Fig. 10.3** Electron micrographs of (a) SSV6 (Redder et al. 2009); (b) ATV (Prangishvili et al. 2006); (c) ABV (Haring et al. 2005); (d) SNDV (Arnold et al. 2000a); (e) APBV1 (Mochizuki et al. 2010); (f) PSV (Haring et al. 2004); (g) SIRV2 (Prangishvili et al. 1999); (h) AFV3 (Vestergaard et al. 2008a); (i) STIV2 (Happonen et al. 2010) and (j) P23-45 (Minakhin et al. 2008). The white arrow in (a) indicates the interaction between tail fibers of four virions. I shows the central section of a cryo-EM picture of STIV2. Scale bars, 50 nm in D; 100 nm in (a), (b), (c), (f), (i) and (j); 200 nm in (e), (g) and (h)

### 10.2.2 Bottle-Shaped, Droplet-Shaped and Bacilliform Viruses

The virions of three archaeal viruses, ABV, APBV1 and SNDV, were all isolated from extremely hot terrestrial environments, and all of them display morphological features that are so unique that each of the viruses has been assigned to a new family.

*The Ampullaviridae Family.* The virion of *Acidianus* bottle-shaped virus (ABV) is structurally one of the most complex virions in the viral world and infects the hyperthermophilic *Acidianus* genus. The complex form resembles a bottle approximately 230 nm long and 4–75 nm wide (Fig. 10.3c). The lipid-containing envelope encases a cone-shaped core formed by a torroidally supercoiled nucleoprotein filament. At the broad end of the bottle shape, a disc is present to which 20 ( $\pm 2$ ) short, thick filaments are attached (Haring et al. 2005). Their function is still unknown. The virion seems to absorb to the host cell through its narrow end, suggesting that the tip of the virion could be involved in viral adsorption and channelling viral DNA into host cells.

*The Guttaviridae Family.* The *Sulfolobus neozelandicus* droplet-shaped virus (SNDV) exhibits a complex droplet morphology. The droplet is approximately 90 nm wide

and 180 nm long and carries multiple long, thin fibres that are attached at its pointed end (Fig. 10.3d). The surface appears to be helically ribbed and has been compared to resemble a beehive (Arnold et al. 2000a). The circular dsDNA genome has not been sequenced, and the virus does not presently exist in culture collections.

*The Clavaviridae Family.* The *Aeropyrum pernix* bacilliform virus 1 (APBV1) has a short stiff bacillus form. The virions are 140 nm long and 20 nm wide (Fig. 10.3e). One end is pointed; the other is rounded. The circular dsDNA genome of 5.2 kb is the smallest genome of known prokaryotic dsDNA viruses (Mochizuki et al. 2010). The approval of the family *Clavaviridae* is pending at the ITV.

### 10.2.3 Linear Viruses

Linear viruses represent the main virion morphotype in extreme geothermal environments. Linear viruses of bacteria and Eukarya carry either ssDNA or ssRNA. Interestingly, linear archaeal viruses isolated from hot environments carry dsDNA. The linear archaeal viruses have been classified into two new families: the stiff, rodlike *Rudiviridae* and the flexible, filamentous *Lipothrixviridae*. One thermophilic filamentous bacteriophage is known, which belongs to the *Inoviridae* family.

*The Lipothrixviridae Family.* Eleven members belonging to this family have been isolated, including seven *Acidianus* filamentous viruses (AFV1, 2, 3, 6, 7, 8 and 9), one *Sulfolobus islandicus* filamentous virus (SIFV) and three *Thermoproteus tenax* viruses (TTV1-3) (Table 10.1). Lipothrixvirus filaments are surrounded by envelopes containing lipids obtained from the host. The filament is approximately 24 nm wide, ranging in length from 900 nm (AFV1) to 2,000 nm (AFV3 and SIFV) (Fig. 10.3h). Lipothrixviruses carry identical terminal structures at both ends, indicating both ends are able to bind to cell receptors. Members of this family show considerable diversity in their terminal structures. These structures can represent claws (AFV1), T-bars (AFV9), mop-like structures (SIFV), three (AFV1) or six (SIFV) short fibres or tips resembling bottle brushes (AFV2) (Pina et al. 2011). The lipothrixvirus SIFV has a termini structure that ends in a mop-like structure to which six tail fibres are attached. This structure unfolds like a spider's legs, before attaching to receptors on the host cell membrane (Arnold et al. 2000b). Apparently, members of the *Lipothrixviridae* family evolved different structures for cellular attachment, providing great flexibility in host recognition.

*The Rudiviridae Family.* This family contains two *S. islandicus* rudiviruses (SIRV1 and 2), one *Acidianus* rod-shaped virus (ARV1) and one *Stygiolobus* rod-shaped virus (SRV) (Table 10.1). The rod-shaped virions are nonenveloped, approximately 23 nm wide and 610–900 nm long (Fig. 10.3g). The length of the linear virion matches the length of the genomic dsDNA, a phenomenon also characteristic of bacterial inoviruses, which carry a ssDNA genome. The virion body is a tubelike superhelix formed by linear dsDNA and multiple copies of a highly glycosylated, DNA-binding capsid

protein (Vestergaard et al. 2008b). A plug is located at each end of the virion. It is approximately 50 nm wide and carries three short terminal fibres.

*The Inoviridae Family.* The only thermophilic member of this family, PH75, infects the bacterium *Thermus thermophilus* (Pederson et al. 2001). The filamentous virions are 6 nm wide and 1,000–2,000 nm long. The terminal structures of inoviruses are not identical. One end functions as attachment site to host cells. The virion attaches to the target cell via protein binding to a pilus. Pilus retraction then pulls the virion to the host's internal membrane. The virions comprise a helical capsid surrounding a core of circular ssDNA. Unlike other bacterial viruses, PH75 does not lyse host cells but exit by extrusion.

#### 10.2.4 Spherical Viruses

Two spherical virus families have been isolated from geothermal environments, *Globulaviridae* that infects Archaea and *Tectiviridae* that infects bacteria. Two spherical hyperthermophilic archaeal viruses, STIV1 and STIV2, still need to be assigned to a family.

*The Globulaviridae Family.* The family currently comprises two viral species, *Pyrobaculum* spherical virus (PSV) and *T. tenax* spherical virus 1 (TTSV1) (Fig. 10.3f; Table 10.1). PSV infects the anaerobic hyperthermophilic archaea, *Pyrobaculum*, which has a temperature optimum of a 100°C. The two viruses are highly similar in morphological and genomic properties. The spherical virions consist of a lipid envelope, which encases a helical nucleocapsid containing a linear dsDNA genome with inverted terminal repeats. The sphere of PSV has a diameter of approximately 100 nm (Pina et al. 2011).

*STIV Viruses.* The hyperthermophilic *Sulfolobus* turreted icosahedral viruses, STIV1 and STIV2 (Fig. 10.3i), remain taxonomically unclassified. The two viruses share a similar icosahedral structure and appear to be closely related. The virions have non-tailed icosahedral structures with an internal lipid monolayer, which encases the circular dsDNA (Fulton et al. 2009). They share an architectural similarity with viruses of the bacteriophage family, *Tectiviridae*. Image reconstruction of the STIV virion revealed a unique virus architecture including complex, turret-like projections extending from each of the vertices, which may be involved in host recognition.

*The Tectiviridae Family.* Two thermophilic bacteriophages have been assigned to this family, P23-77 and  $\phi$ IN93 (Table 10.2). The virions of tectiviruses consist of an icosahedral protein capsid with an internal membrane vesicle that encloses the linear dsDNA genome. The internal membrane contains lipids organised as a bilayer underneath the protein capsid, which can give a layered shell appearance of the spherical virion (Jaatinen et al. 2008). There is no indication of any tail structure. Both thermophilic tectiviruses have small genomes of approximately 17–20 kb, possessing only the basic functions needed for survival in hot environments.

### 10.2.5 Head–Tail Viruses

Several thermophilic viruses with head–tail morphotypes have been isolated from hot environments, most of which infect bacteria. Only two thermophilic archaeal viruses with head–tail morphology have been isolated so far; they infect the thermophilic *Methanobacterium*. The isolated head–tail viruses fall into the two families *Myoviridae* and *Siphoviridae*, and five still remain to be classified. The head has icosahedral symmetry and the tail structure differs in length and construct.

*The Myoviridae Family.* Eight thermophilic viruses belong to this family, with three infecting *Thermus*, one infecting *Rhodothermus* and *Meiothermus* each, and three infecting *Geobacillus* (Table 10.2). The hexagonal heads are approximately 60 nm wide, and the contractile tails vary in length between 80 and 150 nm. In T4, a mesophilic myovirus, the contractile tail ends with a complex baseplate with six long fibres radiating from it. These fibres are used for cell attachment (Comeau et al. 2007). A similar structure has been observed in the thermophilic myovirus  $\phi$ TMA. Like a typical myovirus, such as T4, the tail tube of  $\phi$ TMA protrudes from the bottom of the baseplate when the tail sheath contracts. Moreover, electron microscopy study showed lipid vesicles that appeared to be bound to the bottom of the baseplate of the  $\phi$ TMA page particles (Tamakoshi et al. 2011). Bacteriophage T4 is a very complex virus, more than 40 different proteins form the mature virion and the dsDNA genome comprises approximately 172 kb. Among the isolated thermophilic myoviruses,  $\phi$ TMA,  $\phi$ YS40 and RM378 have genome sizes ranging from 130 to 152 kb; the rest have genomes smaller than 100 kb. It has been hypothesised that the large genomes of  $\phi$ TMA,  $\phi$ YS40 and RM378 may encode a complexity similar to that of T4.

*The Siphoviridae Family.* Four thermophilic bacteriophages and two thermophilic archaeal viruses belong to this family (Table 10.2). The hexagonal heads are approximately 60 nm wide, and the tails have a length of approximately 150 nm and a width of about 8 nm. The tails are noncontractile with a helical structure; at the end of the tail, short terminal fibres are attached. The virions have nonenveloped dsDNA genomes, which have undergone extensive genetic exchange (Hendrix et al. 1999). The thermophilic siphovirus, TSP4, infecting *Thermus* sp. isolated from Southwest of China, has an extremely long and flexible tail of 785 nm in length and 10 nm in width. It has been found that the end of the tail frequently absorbs to cell debris, indicating attachment structures. Despite their exceptional habitats being separated by thousands of kilometres, the morphological characteristics of TSP4 showed high similarity with the thermophilic siphoviruses P23-45 (Fig. 10.3j) and P47-26, which have been isolated from hot springs in the Far East of Russia (Lin et al. 2010).

Five viruses with head–tail morphology remain to be classified, BVW1, Tb1 and Tf2-4. *Bacillus* virus W1 (BVW1) has a long tail of 300 nm in length and 15 nm in width, and the hexagonal head has a diameter of 70 nm. It contains a dsDNA of 18 kb (Liu et al. 2006). Tb1 has been isolated from *Thermonospora alba* and Tf2-4 have been isolated from *Thermonospora fusca*. All four *Thermonospora* bacteriophages possess polyhedral heads and long tails. The tail lengths for Tb1 and

Tf3 are approximately 260–280 nm and for Tf2 and Tf4 approximately 122–125 nm. Tail flexibility has been observed for the longer-tailed bacteriophages; however, flexibility is not apparent for the shorter-tailed types. The four bacteriophages contain dsDNA of 35–45 kb (Lawrence et al. 1986).

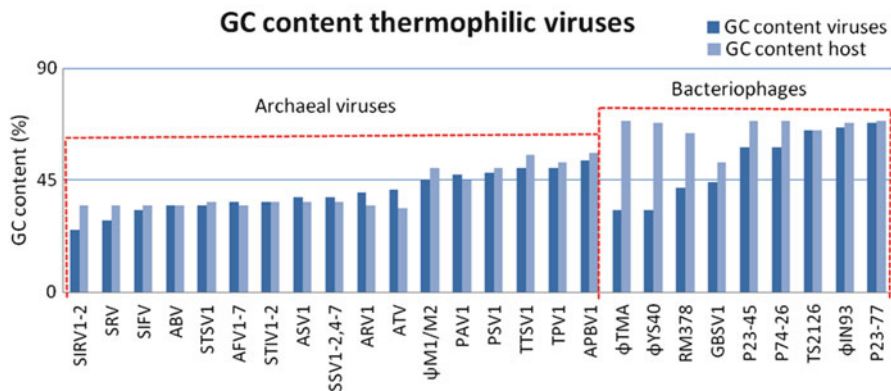
### 10.3 The -Omics Era and Gene Regulation of Thermophilic Viruses

Genomics and functional genomics have been instrumental in the past two decades in the field of biology. Although the number of thermophilic virus isolates is limited, the genomic and functional genomic data originated from them have provided insights into the diversity, ecology and gene regulation of these exceptional organisms.

#### 10.3.1 Genomic Properties

All thermophilic viruses isolated so far package circular or linear dsDNA genomes, with the exception of bacteriophage PH75 which packages circular ssDNA. The genomes of thermophilic viruses vary considerably in size. The smallest genomes belong to the archaeal virus APBV1 (5.2 kb) and the bacteriophage PH75 (6.5 kb). The largest genomes are found among thermophilic bacteriophages  $\phi$ YS40 and  $\phi$ TMA, which range from 151 to 153 kb. The largest archaeal viral genome is found in STSV1 (75.2 kb). Genome size in prokaryotes is often a good predictor of metabolic complexity; it has been hypothesised that the large genomes of  $\phi$ TMA and  $\phi$ YS40 may encode a complexity similar to that of the bacteriophage T4.

The genomes of most isolated thermophilic viruses have been sequenced, providing a wealth of information about the genetic diversity of these viruses. However, the genes of thermophilic viruses generally yield few significant matches to sequences in public sequence databases, and most predicted gene products lack recognisable functions and homologs in extant databases. Identified functions in archaeal viruses are confined to a few proteins. For example, in STIV1 the structure of A197 reveals a GT-A fold that is common to many members of the glycosyltransferase superfamily, suggesting a glycosyltransferase activity for A197. Viruses commonly decorate their proteins with sugars as means of regulating interactions with their hosts. While viruses can utilise their host's glycosylation machinery, it is clear that some also encode their own proteins for specialised glycosylation needs, for example, some lytic bacteriophages glycosylate their DNA to protect it from host restriction enzymes (Larson et al. 2006). In SIRV1 and SIRV2, Holliday junction cleaving enzymes have been characterised. Holliday junction resolving enzymes are ubiquitous and found in all living cells. They are essential for DNA recombination, recombination-related DNA repair and recombination-dependent DNA replication (Birkenbihl et al. 2001). Moreover, SIRV1 and SIRV2 both encode a dUTPase



**Fig. 10.4** Graph comparing the GC content of thermophilic archaeal viruses and bacteriophages with their respective hosts. More than half of the bacteriophages have a GC content of 45% or higher. Only one third of the archaeal viruses have a GC content of 45% or higher

which is involved in nucleotide metabolism (Prangishvili et al. 1998). Similarly, another thermophilic rudivirus, ARV1, encodes a thymidylate synthase (Vestergaard et al. 2005). These enzymes function in adjacent steps of the *de novo* synthesis pathway of thymidine nucleotides, and either enzyme can help to maintain a low dUTP–dTTP ratio, thereby, minimising misincorporation of uracil into DNA (Chen et al. 2002). This is important at high temperatures when dCTP deaminates more rapidly to yield dUTP. Thus both enzymes are most likely important for efficient replication and for the stability of the hyperthermophilic rudiviruses. In SSV1 an integrase has been identified, encoding a site-specific recombination system, suitable for further investigations into the regulatory mechanisms in the SSV1 genome (Muskhelishvili et al. 1992). Other functional categories characterised in archaeal viruses including polymerases, ligases and nucleases will be discussed in the application section. Furthermore, structural genomics has produced insights into the functions of some previously unknown viral genes. This was reviewed recently (Krupovic et al. 2012) and will not be discussed in this chapter.

One surprise when studying thermophilic viral genomes has been the variance in their genomic GC content. The GC content of nucleic acids is known to be correlated with the stability of their double helix. GC pairs are more stable than AT pairs because they have an additional hydrogen bond. Thus higher GC base pairing leads to higher thermal stability of the DNA (Galtier and Lobry 1997). This can be exploited by organisms living in thermal environments. It has been suggested that high GC content may be a selective response to high temperature. However, a study comparing GC content and optimal growth temperature for numerous prokaryotes failed to demonstrate the predicted correlation (Hurst and Merchant 2001). Looking at GC content of thermophilic viruses emphasises this finding. Approximately 35% of the thermophilic archaeal viruses and 55% of the thermophilic bacteriophages have a GC content of 45% or higher (Fig. 10.4). This was surprising given the relatively higher optimal growth temperatures of archaeal hosts than bacterial hosts (Fig. 10.2).



One extreme example is the archaeal rudiviruses SIRV1 and SIRV2 which have a GC content of 25% while the hyperthermophilic host, *Sulfolobus*, has an average GC content of about 37%. Apparently, factors other than GC base pairing, such as DNA-binding protein(s) in virions, play major roles in protecting viruses against heat.

### 10.3.2 Metagenomics

The classification of viruses has traditionally been based on morphological characteristics as demonstrated by the previous section. The system can only be used for viruses that are abundant enough for both microscopic study and genomic sequencing, which significantly biases our view of diversity towards the culturable fraction of the virus community (Pride and Schoenfeld 2008). One factor limiting the discovery of thermophilic viruses has been the reliance on culture-dependent methods for virus isolation. Hot environments are characterised by extreme conditions. For example, deep sea hydrothermal vents have strong physicochemical gradients such as temperatures from 2°C to more than 350°C and high hydrostatic pressure, lack of solar energy and prevalence of chemosynthesis (Gorlas et al. 2012). No viral-cultivation study can fully mimic the temperature and pressure extremes that characterise deep sea hydrothermal vents; this imposes a selective pressure preventing the propagation of viruses unable to adapt to laboratory conditions. Studies have shown that only 1–15% of microbial organisms are culturable under laboratory conditions, demonstrating the amount of genomes still to be discovered (Singh et al. 2009). The introduction of viral metagenomics has revolutionised the field of environmental virology by allowing the exploration of viral communities in a variety of environments, including extreme environments. Viral metagenomics involves viral particle purification followed by library construction and sequencing (Rosario and Breitbart 2011). Viral metagenomic studies of hot environments have provided information concerning viral biogeography, diversity and community structure and new viral types and contributed to the discovery of a potential archaeal RNA virus.

There are approximately 26 published metagenomic studies investigating viral communities, among which, five investigate hot environments. From these studies, it is clear that environmental viral communities are different from those observed in culture. Analysis of viral metagenomes has shown that about 30% of these sequences have detectable similarity to those in GenBank and about half of these are most similar to other known viruses (Rosario and Breitbart 2011). The high percentage of unknown sequences demonstrates the vast novelty of genetic information to be obtained from viruses. Despite the lack of sequence identification from viral metagenomes, studies have used this technique to catalogue viruses in environmental samples based on the identifiable sequences and to investigate community composition through statistical analyses.

A metagenomic analysis of samples from Yellowstone hot springs investigated viral biogeography at a local and global scale (Schoenfeld et al. 2008). Two pools

were investigated, Bear Paw (74°C) and Octopus (93°C). The two metagenomic libraries, one from each pool, showed a relative close similarity (nearly 25%). This was surprising, given that microbial populations are temperature-dependent and the surface temperature of the two hot springs differ by 19°C. Alignment of the metagenomes to whole-genome sequences of six cultivated thermophilic viruses revealed striking conservation of certain sequences. Comparison with the genome of PSV showed median identities of 60 and 51% to Bear Paw and Octopus, respectively. PSV was isolated from a hot spring with notably different geochemistry and more than 30 km away from both hot springs. The high median identities between PSV and the two metagenome libraries illustrate the global conservation of sequences in hot springs with different physical and biogeochemical properties (Schoenfeld et al. 2008). This correlates with the findings that groups of highly similar *Sulfolobus* viruses and *Thermus* bacteriophages have been isolated from hot springs on different continents (Wiedenheft et al. 2004; Yu et al. 2006).

In general, viral metagenomics has revealed an enormous diversity and abundance of viruses in all environments. Estimates of viral diversity suggest the existence of several thousand unidentified virus types (Weinbauer and Rassoulzadegan 2004). With metagenomics, it is possible to assemble complete genomes of unidentified virus-like particles obtained from environmental viromes, thus identifying novel viruses. A metagenomic study investigating an enriched environmental sample from a hot spring in Yellowstone national park yielded two novel viral genomes, HAV1 and HAV2. Neither viral genome shows any clear similarity to other known archaeal viruses; only HAV2 shows morphological similarities with the two-tailed spindle-shaped ATV virus and limited sequence similarity between two genes. HAV1 has a linear 23 kb genome and HAV2 has a circular 18 kb genome; both have dsDNA. They yielded few significant matches in public sequence databases, reinforcing the notion of wide genetic diversity in archaeal viruses. The study attempted to identify the hosts of the two viruses by isolating archaeal strains from the environmental sample. However, none of the isolated strains were infected (Garrett et al. 2010). Based on the morphological and limited sequence, similarities to ATV, HAV1 and HAV2 were inferred as archaeal viruses. It can be difficult to identify potential archaeal hosts, especially because no reliable protocol procedures for transfection of viral DNA into potential hosts have been developed. Metagenomic analyses can be a helpful tool for identifying new viruses, although information concerning hosts and virus life cycle can be difficult to confirm.

Identification of viral hosts is crucial in understanding the role of viruses in an ecosystem. It would aid in determining how viruses impact the microbial diversity of a given ecosystem, as they are the only predators in hot environments above 60°C. A study tried to identify specific hosts for a viral assemblage obtained from a hydrothermal vent in the Northeast Pacific. The viral metagenome was compared with a comprehensive database of spacers derived from the clustered regularly interspaced short palindromic repeat (CRISPR) adaptive immune system (Anderson et al. 2011). The CRISPR system is an antiviral defence mechanism found in both Archaea and bacteria (Marraffini and Sontheimer 2010). Cells incorporate genomic sequences as CRISPR spacers from invading viruses and plasmids into their CRISPR



loci. When invading viruses or plasmids have a match to a pre-existing CRISPR spacer sequence in the host genome, these elements are recognised as pathogenic invaders. Thus, CRISPR spacer sequences provide an adaptive, heritable record of past infections and express small CRISPR RNAs that guide targeting of invasive nucleic acids, rendering the cell immune to these viruses (Marraffini and Sontheimer 2010). The CRISPR loci act effectively as libraries of previous viral infections which can be compared with environmental viromes, thus potentially identifying virus–host relationships. A study by Anderson et al. (2011) showed that CRISPR spacers from vent isolates and from thermophiles in general have a higher percentage of matches to the vent virome than to other marine or terrestrial hot spring viromes. Spacers derived from strains belonging to 23 different taxonomic groups matched sequences from the hot vent virome; most notably, a wide range of both archaeal and bacterial taxonomic groups has CRISPR spacers matching the marine vent viromes. Interestingly, a high percentage of hits to spacers from mesophilic hosts suggested that the marine vent virome was comprised of viruses that have the potential to infect diverse taxonomic groups of multiple thermal regimes in both the bacterial and the archaeal domains (Anderson et al. 2011).

Metagenomics has also resulted in findings that suggest the existence of novel positive-strand RNA viruses that probably replicate in hyperthermophilic archaeal hosts and are highly divergent from RNA viruses that infect eukaryotes and bacteria. To date, all archaeal viruses possess dsDNA genomes, except one that possess ssDNA. Finding a RNA virus infecting an archaeal host could enhance our knowledge of RNA viruses and shed light on the origin of the enormous diversity of RNA viruses infecting eukaryotes (Bolduc et al. 2012). Further studies need to identify the potential hosts of the RNA viruses, possibly through sequence comparison between the RNA viral genomes and CRISPR spacers of cellular genomes from the same environment.

Advances in the field of metagenomics will provide more information concerning important aspects of viral ecology. Moreover, functional viral metagenomics are being developed to discover novel viral enzymes that can be used for diagnostic and biotechnological purposes (Schoenfeld et al. 2010).

### ***10.3.3 Functional Genomics and Gene Regulation***

A virus is totally dependent on its host cell for viral reproduction. In order to multiply, a virus must take control of its host cell's molecular machineries. Perhaps the most important mechanism for achieving this control is a viral code for strong positive signals to promote viral gene expression and other signals to repress expression of some cellular genes. When a cell is infected, much is going on inside the cell at the molecular level, such as transcription of the incoming viral genes to form viral mRNAs, and their translation to produce early viral proteins, including the enzymes necessary to replicate viral DNA (Collier and Oxford 2006). Transcriptomics is a useful tool for investigating the subtle but definite changes in

gene expression of both the host and the virus in the period of a virus infection. The transcriptome reflects the genes that are expressed at any given time. Therefore, it can provide insights into which genes are being up- or downregulated during an infection cycle. Moreover, viral transcriptomics can be used to compare transcription patterns of viruses under a wide range of conditions (Walther et al. 2011).

Microarray analysis of an infected culture of *Sulfolobus solfataricus* with STIV1 has revealed insights into the timing and extent of virus transcription, as well as differential regulation of host genes. The infection cycle of STIV1 is lytic with almost all cells in a culture being killed by the virus. Transcription of viral genes was first detected at 8 h post infection (hpi), and at 16 hpi, most viral genes were expressed. The majority of the viral genes are transcribed between 16 and 24 hpi. Around 24 hpi, a shift takes place from virus replication to preparation for lysis, with general cell lysis detected at 32 hpi. Although the expression starts at different timepoints for different genes, little temporal control was observed. During the infection, 177 host genes were determined to be differentially expressed, with 124 genes upregulated and 53 genes downregulated. The upregulated genes were dominated by genes associated with DNA replication and repair and those of unknown function, suggesting that STIV1 uses host proteins to aid the replication of its own DNA (Ortmann et al. 2008). An important upregulated gene concerns an ESCRT (endosomal-sorting complex required for transport) III homolog. The ESCRT III complex serves in eukaryotes as a protein-sorting machinery and functions as a pair of molecular scissors that cuts cell endosomal membranes (Wollert et al. 2009). The ESCRT III homolog has recently been reported essential for the cell division in *Sulfolobales* (Ettema and Bernander 2009); the upregulation may suggest its involvement in the release of STIV1 virions. The downregulated genes, mostly detected at 32 hpi, were associated with energy production and metabolism (Ortmann et al. 2008). Interestingly, the majority (94%) of the host genes showed no differential regulation between infected and uninfected cells, suggesting that the virus life cycle is tailored to avoid host stress response.

An infection study of SSV1 with *S. solfataricus* as a host investigated the transcriptome fluctuations of a virus with a lysogenic life cycle. Upon infection, the SSV1 genome is rapidly and site specifically integrated into a tRNA gene of the host paralleled by a short slowdown of growth. Even after infection with an excess of virus particles, the host recovers well and often grows even better than before, as if the presence of the viral genome gives some advantage (Schleper et al. 1992). Upon UV irradiation of the host cells, a strong replication of the viral DNA is induced and large amounts of particles (up to 100 per cell) are released into the culture medium. This occurs without apparent lysis of the host cells. However, at this stage, the cell growth is significantly retarded. The study found that the first viral transcripts already can be found after 1 hpi, while most viral genes are active at 8.5 hpi. The viral genes are clustered at 9 operons expressed into 10 transcripts, comprising both regulatory and structural genes. The regulatory genes are the first to be transcribed, and the genes coding for the coat protein of the virus are produced at a later stage. Similar to those of eukaryotic viruses and bacteriophages, the transcripts of SSV1 can be categorised according to their time of appearance

and putative functional roles, showing that SSV1 exhibits a tight chronological transcriptional regulation (Frols et al. 2007). With respect to host gene regulation, transcriptomes were compared between UV irradiated and nonirradiated cells for both infected and uninfected cultures, and the response to UV irradiation was then compared between infected and uninfected cells. Only small differences in genome-wide analysis were detected between infected and uninfected cells, possibly due to the indirect comparison.

Very few studies have investigated the change in gene expression profiles when a bacteriophage infects a hyperthermophilic bacterium. A study analysing the interaction between GVE2, a deep sea thermophilic bacteriophage, and its host *Geobacillus* sp. E263 showed differential regulation of host genes. A comparison of protein/gene expression profiles of GVE2-infected and noninfected *Geobacillus* sp. E263 bacteria revealed that among the 20 differentially expressed host genes, 13 were upregulated and 7 were downregulated in response to GVE2 infection. Based on homology searches in GenBank, 19 of the 20 proteins involved in GVE2 infection shared homology with proteins of known function. The 19 proteins have diverse metabolic functions and can be grouped into three different categories based on cellular function, suggesting a coordinated response to virus infection (Wei and Zhang 2010).

One of the current limitations of transcriptomic methods is that they require millions of cells as a starting material. Therefore, a future challenge will be the determination of transcriptome maps for single cells, which will open an avenue for investigating the regulatory heterogeneity in a microbial population upon infection (Sorek and Cossart 2010). This may lead to the discovery of advantageous genes/defence mechanisms in some cells of a given population. Furthermore, single cell transcriptomics will also allow studies of the transcriptome of individual cells from unculturable species, a major advantage with the high amount of unculturable thermophilic viruses and their hosts. Another field that could advance our understanding of thermophilic viruses' impact on their environments is metatranscriptomics. The analyses of the ecology in hyperthermophilic environments are generally less complex than that of aquatic or soil systems, making it easier to deal with big dataset covering many organisms (Walther et al. 2011). Comparing metatranscriptomes from different hot environments could elucidate the impact of viruses, helping us understand the roles viruses play in ecological and geochemical processes in hot environments.

## 10.4 Virus–Host Interactions

The number of studies characterising virus life cycles in hot environments is limited, and the interactions have rarely been characterised in depth, beyond the basic level. According to studies on isolated archaeal virus–host systems, the dominant type of the viral life cycle differs from that of bacterial systems (Abedon 2009). The majority of known thermophilic archaeal viruses establish a chronic infection in which virions are continuously produced, and the host cells remain alive although the

growth rate is often retarded. Adaptation to energy stress conditions is hypothesised to be the crucial factor that distinguishes Archaea from Bacteria (Valentine 2007) and may be what has driven and even favoured such virus–host relationships. The only hyperthermophilic bacteriophage that exhibits a chronic mode of infection is the filamentous bacteriophage of the *Inoviridae* family, PH75, which does not lyse host cells even when virions are actively produced (Pederson et al. 2001). This infection mode is rare for bacteriophages but appears to be common among thermophilic archaeal viruses and is often referred to as the carrier state. However, the carrier state can be interrupted and transformed into a lytic state by stress factors such as UV irradiation. The hallmark of the lytic life cycle is the lysis of the host cells when virions are released, thus killing the host. The lytic cycle is the mode of reproduction for all hyperthermophilic bacteriophages with the exception of PH75. A few archaeal viruses are reported to be purely lytic (TTV1, STIV and SIRV). Determining the nature of the infection cycle is not always easy. For example, the rudivirus SIRV2 was thought to establish a carrier state infection, mainly based on the lack of a decrease in OD in infected cultures. However, the virus lyses the cells. The exceptional egress mechanism involved in the release of virions from the cell results in lysed cells in form of empty spheres. The cells appear intact and OD measurements lack the sophistication to identify that the measured cells are no longer ‘living’. Thus, other thermophilic archaeal viruses could turn out to be lytic upon further investigations.

#### 10.4.1 Carrier State Infection

AFV1 is similar to most thermophilic archaeal viruses. Thus, it does not kill its host during reproductive infection. A study investigating the virus–host interactions between AFV1 and its host showed that the growth rate of host cells, *Acidianus hospitalis*, is nearly completely blocked in the initial stage of infection. The cell growth recovers slowly after initial infection, and from 2 hpi, a generation time of 20 h is observed, in contrast to a generation time of 11 h for uninfected cells. Mature virus particles start to be released from infected cells 4–5 hpi. After several successive dilutions of virus-carrying cultures and prolonged incubation, the virus was still present in host cells, indicating a stable carrier state of host–virus relationship. Under certain growth conditions, a balance between production of virions and multiplication of host cells can be observed. DNA bands originating from the viral AFV1 genome could not be distinguished in the restriction digestion pattern of total cellular DNA at any stage of infection. This indicates a low intracellular copy number of virus DNA, probably less than 10 copies per cell (Bettstetter et al. 2003). Another thermophilic virus, which reproduces by combining a lysogenic life cycle with various levels of viral particle production in a carrier state mode, is SSV2. A comparative study of SSV2 physiology in the natural host *S. islandicus* versus the foreign host *S. solfataricus* provided evidence of differently regulated SSV2 life cycles in the two hosts. An initial infection of the foreign host *S. solfataricus* retards

cell growth. However, when SSV2 settles in the foreign host *S. solfataricus*, host growth is no longer inhibited. This is in strict contrast to what is observed in the initial infection stage. Neither does the SSV2 genome exceed over a few copies per cell throughout extended cultivation. This suggests that a series of interactions between SSV2 and its foreign host leads to a coexistence harmony between them. Further studies may reveal whether this conversion of host response is a general scheme for the host defence of *S. solfataricus* to a viral infection. In the natural host *S. islandicus*, SSV2 replication is characterised by a physiological induction. Viral genome copy number increases up to 50-fold within 4 hpi after the sudden halt of cell growth at OD<sub>600</sub> of about 1.3 which corresponds to a late exponential growth phase. Growth inhibition of the host correlates with the virus replication induction. Interestingly, the inhibition effect is reversible. When released into a fresh medium, the SSV2-induced *S. islandicus* culture resumes their exponential growth and hosts regain control over SSV2 replication (Contursi et al. 2006).

### 10.4.2 Lytic Life Cycle

Virus–host interactions in hyperthermophilic bacteriophages are not well studied. Therefore, the best examples of lytic thermophilic virus–host interactions are from the domain of Archaea. As mentioned earlier, only few thermophilic archaeal viruses are lytic. However, the life cycles of the lytic viruses, STIV1 and SIRV2, have revealed a unique mechanism of virion release (Prangishvili and Quax 2011). The viruses differ significantly in their morphological and genomic properties but exploit the same mechanism for the release of mature virions from the host cell. A study done by Bize et al. (2009) investigated the life cycle of SIRV2. SIRV2 infects *S. islandicus* and virions are assembled in the cytoplasm of the host cell; approximately 8–10 hpi, they start to be released through well-defined apertures in the cell envelope. Formation of these openings is preceded and facilitated by the generation of virus-induced cellular structures of pyramidal shape, virus-associated pyramids (VAPs), located at the cell envelope and pointing outwards. The VAPs perforate the membrane and S-layer and opens up to release the preassembled virions in the cytoplasm to the surrounding environment. After virion release, the cell envelope remains as a stable empty shell. The timing of VAP disruption and virus release must be strictly controlled by virus-encoded functions, such that cell lysis does not occur until the virions have been assembled, as for any lytic virus (Bize et al. 2009). The extent of modifications caused by SIRV2 on the host cells results in a radically transformed cell, which functions as a complex viral factory. It has been suggested that the lysogenic life cycle provides an intracellular refuge for virus populations in hot environments. However, viruses with a purely lytic life cycle such as SIRV2 demonstrate that virus particles can survive in extreme ecosystems long enough to encounter new host cells. The SIRV2 virions are well adapted to hot environments, being almost as stable at 80°C as bacteriophages infecting mesophilic bacteria are at 37°C (De Paepe and Taddei 2006).

## 10.5 Applications

The previous sections have covered thermophilic viral characteristics, including novel features not previously observed. Many of these features may have application potential in basic research or biotechnology. However, application development in thermophilic viruses has so far concentrated on a few identified enzymes and an emerging interest to use them as nanobuilding blocks. Virus-derived vectors and enzymes have been essential research tools since the first days of molecular biology, and compared to mesophilic viruses, an obvious advantage of thermophilic viruses is their thermostability. The next section will summarise thermophilic viral enzymes which are already trademarked or have the potential for biotechnological/industrial exploitation. Moreover, the potential of thermophilic viruses as nanobuilding blocks will be addressed.

### 10.5.1 *Application of Thermophilic Viral Enzymes in Biotechnology and Basic Research*

Many research applications of enzymes are centred on nucleic acid metabolism. Especially DNA polymerases have been used in molecular biology techniques such as whole-genome amplification, PCR and DNA sequencing (Tabor and Richardson 1987; Saiki et al. 1988; Zhang et al. 1992). Viral polymerases are highly diverse in terms of primary amino acid sequence and biochemical activities (Schoenfeld et al. 2010) and are functionally distinct from their cellular counterparts. Many applications of DNA polymerases depend on thermostability up to 95°C, making hyperthermophilic viruses ideal subjects for discovering viral DNA polymerases with biotechnological potential. Recently, a thermostable DNA polymerase from a viral metagenome was found to be a potent RT-PCR enzyme (Moser et al. 2012). Out of 21,198 Sanger sequence reads derived from a viral metagenome library constructed from Octopus hot spring (93°C) in Yellowstone National Park, hundreds of potential *pol* genes were identified and 59 complete *pol* genes were tested for polymerase activity. Among these, 3173 Pol demonstrated both high thermostability and innate reverse transcriptase (RT) activity (Table 10.3). As the first reported virus-derived thermostable RT Pol, 3173 Pol also exhibited high sensitivity and high specificity comparable to two-enzyme RT-PCR systems. Obviously, 3173 Pol fulfils the requirements for a facile single-enzyme RT-PCR reagent (Moser et al. 2012). The 3173 Pol-based PyroScript RT-PCR master mix (Lucigen) provides demonstrated advantages over the commonly used two-enzymes RT-PCR systems. Among other documented performance problems, the two-enzyme RT-PCR systems require an initial low temperature reverse transcription that reduces specificity, increases reaction time and impairs synthesis through complex secondary structures. In contrast, the high thermostability of 3173 Pol should allow a 'hot start' (Moser et al. 2012).

**Table 10.3** Current and future uses of thermophilic viruses

Enzyme/nanoparticles	Viral source	Current and emerging applications	References
ssDNA/RNA ligase 1	TS2126	Production of ssDNA templates for rolling-circle replication or rolling-circle transcription experiment	EpiCentre (2012)
		Production of ssDNA templates for RNA polymerase and RNA polymerase inhibitor assays	Blondal et al. (2005b)
Polynucleotide ligase 1	RM378	Phosphorylation of nucleic acids with modified phosphates for subsequent modifications and/or labeling	Blondal et al. (2005a) and Prokazyne (2012)
Nonspecific nuclease	GBSV1	RNA sequencing Removal of nucleic acids during protein purification Antiviral agents	Song and Zhang (2008)
Protein-primed DNA polymerase	ABV	PCR Genome amplification	Peng et al. (2007)
3173 DNA polymerase	Metagenome	RT-PCR Genome amplification DNA sequencing	Moser et al. (2012)
Nanoparticles	SSV1	Nanocompartments	Reiter et al. (1987) and Steinmetz et al. (2008a)
	SIRV2	Imaging Drug delivery	

PCR has been used for about two decades as a powerful tool in molecular biology. However, the size of DNA product that can be amplified by PCR is still limited. With the best PCR enzyme to date, Phusion, a maximum of 20 kb can be reached. Although the bacteriophage phi29 DNA polymerase presents the highest processivity described so far for a DNA polymerase, and up to 70 kb fragments can be generated (Blanco et al. 1989; Rodriguez et al. 2005), the enzyme is not thermostable and therefore cannot be used in PCR. Biochemical and structural studies have shown that the high processivity of Phi29 DNA Pol is dependent on one of the two insertions in the protein sequence which are a characteristic of protein-primed DNA Pol families (Rodriguez et al. 2005). Recently, a putative protein-primed DNA polymerase gene was identified from the genome of the hyperthermophilic archaeal virus ABV which also contains the two insertions in the protein sequence (Table 10.3) (Peng et al. 2007). This feature may render the ABV Pol a high processivity, which, in combination with the high thermostability, makes the enzyme a promising candidate for PCR amplification of large DNA fragments.



DNA polymerases constitute just one group of useful enzymes; other viral enzymes with distinct and useful properties have recently been discovered in thermophilic viruses. Thermostable RNA ligases 1 from bacteriophages TS2126 and RM378 were recently isolated and characterised (Table 10.3) (Blondal et al. 2003, 2005a, b). RNA ligases 1 have the ability to ligate single-stranded nucleic acids by catalysing the ATP-dependent formation of phosphodiester bonds between 5'-phosphate and 3'-hydroxyl termini of single-stranded RNA or DNA. The biological role of bacteriophage RNA ligases has been primarily studied in T4, where the T4 RNA ligase 1, together with the bacteriophage encoded polynucleotide kinase, repairs cleaved tRNA molecules, thereby counteracting the defence mechanism of the bacterial host. The actual role of RNA ligases in TS2126 and RM378 bacteriophages has not been determined, although it is likely that these enzymes are a part of a repair machinery that responds to RNA degradation by the host. The T4 RNA ligase is a very important tool in molecular biology and is used in numerous protocols. Applications include RNA ligase-mediated rapid amplification of cDNA ends (RLM-RACE) which can be used in mapping the 5' and 3' ends of RNA molecules (Liu and Gorovsky 1993), ligation of oligonucleotide adaptors to cDNA or single-stranded primer extension products for PCR (Zhang et al. 1992) and various 5' nucleotide modifications of nucleic acids. For example, a rapid 5'-labelling method of single-stranded DNA/RNA was developed based on the utilisation of an adenylylated intermediate in the reaction of T4 RNA ligase. This method is useful for fluorescence-, isotope- or biotin-labelling of the 5' ends of both oligo- and polynucleotides (Kinoshita et al. 1997).

Investigations into the overall identity of the RNA ligase 1 sequences from TS2126 and RM378 showed low similarity to each other. The amino acid sequence of RNA ligase 1 from the TS2126 bacteriophage showed more similarity to T4 (18%) than RM378 ligase 1 (15%), indicating that the two proteins of thermophilic origin have evolved and adapted independently to the elevated thermal conditions (Blondal et al. 2005a). The characterisation of TS2126 RNA ligase 1 revealed that it is stable at 60–65°C for an extended time period but loses activity at higher temperatures, similar to the thermostability of RM378 RNA ligase 1. When comparing the activity of TS2126 and RM378 and a commercial T4 RNA ligase 1, the TS2126 ligase showed ~30 and 10 times higher specific activity, respectively. This striking difference in ligation efficiency was also observed in the ssDNA ligation experiments where TS2126 RNA ligase 1 was much more effective than RNA ligases from RM378 and T4. The TS2126 RNA ligase 1 exhibits extremely high activity, high ligation efficiency and moderate thermostability; thus, it can be used in the RLM-RACE assay at elevated temperatures (65°C). This may produce better results if the 5' donor end of mRNA molecules has a secondary structure that inhibits efficient ligations with T4 RNA ligase at 37°C (Blondal et al. 2005b). The ability to ligate ssDNA at elevated temperatures may also lead to a major advancement in specific applications in molecular biology. In fact, patents covering RNA ligases 1 from both TS2126 and RM378 have been sold to Epicentre® (an Illumina® company). The RNA ligase 1 from TS2126 is already commercially available under the registered trademark CircLigase™. The thermostable ATP-dependent ligase that



catalyses intramolecular ligation (i.e. circularisation) of ssDNA templates having a 5'-phosphate and a 3'-hydroxyl group can be used for very specific applications. In contrast to T4 DNA ligase and Ampligase® DNA ligase, which ligate dsDNA ends, CircLigase™ ligates ends of ssDNA. The enzyme is therefore useful for making circular ssDNA molecules from linear ssDNA. Circular ssDNA molecules can be used as substrates for rolling-circle replication or rolling-circle transcription. Linear ssDNA of >30 bases is circularised by the CircLigase™ enzyme. Under standard reaction conditions, virtually no linear concatamers or circular concatamers are produced. In addition to its activity on ssDNA, CircLigase™ also has activity in ligating a single-stranded nucleic acid having a 3'-hydroxyl ribonucleotide and a 5'-phosphorylated ribonucleotide or deoxyribonucleotide (EpiCentre 2012). This can be used to map transcription termination sites of prokaryotic RNAs.

Following the characterisation of RNA ligase 1, a polynucleotide kinase (PNK) was identified and characterised in RM378 (Table 10.3), elucidating a defence mechanism against the host similar to that observed in T4. RNA ligase 1 functions together with PNK 1; together, they repair cleaved tRNA molecules. Whereas RNA ligase 1 is essential for ligation of the cleaved tRNA molecules, PNK 1 modifies the tRNA fragments to make appropriate substrates for the ligation step (Silber et al. 1972; Lillehaug and Kleppe 1977). PNK 1 is a bifunctional nucleic acid processing enzyme with 5'-kinase and 3'-phosphatase activities, catalysing the restoration of 5'-phosphate and 3'-hydroxyl termini in cleaved ssRNA and ssDNA. It functions in two major ways: (1) to remove the 2':3'-cyclic phosphate from the 5' tRNA fragment and (2) to add a phosphate group to the 5'-hydroxyl group of the 3'-tRNA fragment using ATP as the phosphate donor (Wang et al. 2002). The RNA ligase 1 and the PNK 1 are thus part of the same system, acting in concert to repair RNA and DNA. A study showed that RM378 PNK 1 carries out the same or very similar processes as those of the T4 PNK 1 (Blondal et al. 2005a). Characterisation of the RM378 PNK 1 protein showed thermostability up to 60–65°C; this is consistent with the optimum temperature of the natural environment of its host, *Rhodothermus marinus*. The RM378 PNK 1 shares only its 5'-kinase domain with the PNK family but no apparent homology to the 3'-phosphatase domain in that family. Its similarity with the T4 5'-kinase is low, but they share the P-loop motif, which is characteristic of many phosphotransferase families. Because of its unique phosphatase domain, RM378 PNK 1 resembles the mammalian PNKs rather than other phage PNKs in terms of domain arrangement (Zhu et al. 2004; Blondal et al. 2005a). This is a very interesting observation, because eukaryotic PNKs function as repair enzymes on double-stranded DNA. The T4 PNK 1 is used for labelling the 5'-termini of nucleic acids, and the labelled products can be used as markers for gel electrophoresis, primers for DNA sequencing, primers for PCR and probes for hybridisation (Stahl et al. 1991; Hilario 2004). The RM378 PNK 1 may have advantages in some protocols due to its stability at higher temperatures. The company Prokazyne is selling the RM378 PNK 1 under the registered trademark ThermoPhage™ polynucleotide kinase, and their protocol mentions several application areas including labelling of nucleic acids using 32P- $\gamma$ -ATP for probes and DNA sequencing, phosphorylation of nucleic acids for subsequent ligation for cloning, phosphorylation of

oligonucleotides for ligase reaction like the ligase chain reaction and similar procedures and phosphorylation of nucleic acids with modified phosphates (i.e. thiol-phosphates) for subsequent modifications and/or labelling. ThermoPhage™ polynucleotide kinase is the only available thermostable enzyme of its kind (Prokazyme 2012) opening up for exploration of high-temperature applications with a PNK.

Another study characterised a novel nonspecific nuclease from the thermophilic bacteriophage GBSV1 (Table 10.3) (Song and Zhang 2008). Nucleases are defined as a group of enzymes which are capable of hydrolyzing the phosphodiester linkages of nucleic acids. According to the substrates they hydrolyze, nucleases are divided into two groups: sugar specific nucleases (deoxyribonucleases and ribonucleases) and sugar nonspecific nucleases. Sugar nonspecific nucleases are characterised by their ability to hydrolyze both DNA and RNA without exhibiting pronounced base preferences. Sugar nonspecific nucleases play very important roles in different aspects of basic genetic mechanisms, including DNA salvage, repair, recombination and degradation. Moreover, they are involved in nutrition, scavenging for nucleotides and phosphates for growth and metabolism (Hsia et al. 2005). They have been isolated from a wide variety of sources including fungi, bacteria and viruses. The majority of these enzymes are intracellular, but some have been reported to be extracellular in nature (Legerski et al. 1978; Rangarajan and Shankar 1999). The ability of sugar nonspecific endonucleases to recognise a wide variety of nucleic acid structures has led to considerable efforts to evaluate their role in different cellular processes as well as application as analytical tools to study nucleic acid structure. Applications also include rapid sequencing of RNA, the removal of nucleic acids during protein purification and the use as antiviral agents (Rangarajan and Shankar 1999; Song and Zhang 2008). Compared with mesophilic enzymes, thermostable nucleases may possess novel properties in structures and biological functions. The novel GBSV1 nonspecific nuclease purified from GBSV1 is the first nuclease isolated from a thermophilic virus. GBSV1 nonspecific nuclease is able to degrade a variety of nucleic acids, including RNA, ssDNA and dsDNA that is either circular or linear. It is active at temperatures ranging from 20 to 80°C with an optimal temperature of 60°C, which is higher than those of most reported nonspecific nucleases. GBSV1 nonspecific nuclease could be obtained in large quantity by expression in *E. coli*. This would facilitate its biotechnological applications (Song and Zhang 2008).

### ***10.5.2 Industrial Potential of Thermophilic Viruses***

Thermophilic viruses have adapted to extreme hot environments. Besides high temperature, hot environments are often accompanied by a set of harsh conditions, consisting of physical extremes (e.g. pressure or radiation) and geochemical extremes such as salinity and pH. All biomolecules within the virus particles, including proteins, nucleic acids and possibly lipids, must be adapted to these conditions. Therefore, thermophilic viruses offer a source of thermostable enzymes that display outstanding stability against high temperatures and can often cope with a set of harsh conditions. Driven by increasing industrial demands for biocatalysts that

can cope with industrial process conditions, considerable efforts have been devoted to the search for such enzymes. Biocatalysis uses natural catalysts, such as enzymes, to perform chemical transformations on organic compounds compared with organic synthesis; biocatalysts often have far better chemical precision, which can lead to more efficient production of single stereoisomers, fewer side reactions and a lower environmental burden. Approximately 3,000 different enzymes have been identified, and many of these have found their way into biotechnological and industrial applications (van den Burg 2003). However, there is a demand for enzymes which can withstand industrial reaction conditions. The reasons to exploit enzymes that are stable and active at elevated temperatures are obvious. At elevated temperatures, the solubility of many reaction components, in particular polymeric substrates, is significantly improved. Moreover, the risk of contamination, leading to undesired complications, is reduced at higher temperatures (van den Burg 2003). As a result, increasing attention is given to microorganisms that are able to thrive in hot environments. The exceptional diversity and novelty of thermophilic archaeal viral genomes indicates that novel enzymes may be encoded. One example is the putative provirus XQ2 entrapped in the genome of *S. solfataricus* P2 (She et al. 2001). The genome of the XQ2 is about 60 kb which encodes many hypothetical proteins as well as recognisable metabolic enzymes. The sizes of thermophilic viral genomes range from 5 to 153 kb, showing the coding capacity for metabolic genes. Recently, an EU funded collaborative project, HotZyme, was initiated aiming at a global investigation of biodiversity of the hot environments on Earth. Most importantly, novel hydrolases with improved performance in different industrial fields are expected to be discovered from thermophilic microorganisms and their viruses (HotZyme 2012).

### 10.5.3 Nanotechnological Potential of Thermophilic Viruses

A virus is a nanoscaled biomolecular unit composed of genes, protecting capsid proteins, and eventually envelopes. In recent years, naturally occurring bionanoparticles, including virus capsids and protein cages, have been studied and utilised as templates and building blocks for applications in biotechnology. Virus capsids and protein cages have received special attention because they are naturally self-assembled with atomic precision. Viral nanoparticles (VNPs) have several advantages such as their nanometre range size, their propensity to self-assemble into monodisperse nanoparticles of discrete size and shape, their high degree of symmetry and polyvalency, their relative ease of producing large quantities of material, and their exceptional stability, robustness and biocompatibility (Steinmetz et al. 2008b). The particles are composed of programmable units, which can be modified by either genetic modification or chemical bioconjugation methods. Due to these properties, viruses hold promise for development as amenable platforms for diverse applications in biotechnology, electronics and medicine (Wiedenheft et al. 2007). Naturally occurring protein cage architectures, such as the spherical cagelike architectures of ferritins, can mediate the deposition of 'hard' inorganic materials within the spatial confines of a 'soft' protein container. For example, Cowpea chlorotic mottle

virus (CCMV) is composed of 180 identical coat proteins that self-assemble to form a protein cage where RNA is encapsulated (Speir et al. 1995). CCMV undergoes reversible swelling when pH varies. Capsid swelling of CCMV at pH levels greater than 6.5 results in pore opening of the capsid protein assembly (Tama and Brooks 2002). Capsid protein swelling allows for the exchange of molecules, which in turn enables the release or entrapment of target molecules (Douglas and Young 1998). The development of VNPs for nanotechnological applications often requires that they withstand harsh conditions during either fabrication or use. Bioconjugation technique often requires the organic solvent dimethyl sulfoxide (DMSO) in concentrations of at least 20% by volume, and many applications rely on temperature stability. This has prompted efforts aimed at exploiting the thermostability and unique features found in hyperthermophiles. A recent study investigated *S. islandicus* rod-shaped virus 2 (SIRV2) as a candidate VNP (Table 10.3) (Steinmetz et al. 2008a). SIRV2 particles were found to be stable in two different solvent/water mixtures that are of relevance for bioconjugation and mineralisation reactions. Moreover, SIRV2 particles offer attachment sites allowing selective chemical modification. It was found that the major coat protein (CP) forms the virus body while the minor CP is located in the tail fibres at the end of the particles. Interestingly, amine reactivity showed that the minor CP could be selectively labelled, and this labelling reaction targeted the ends of the particles only. This suggests that various functional molecules can be installed at different positions in the VNP, offering multiple and various chemical attachment sites. SIRV2 remained intact and infectious in DMSO in concentrations up to 50% by volume for at least 6 days and is naturally stable at 80°C and pH 3. Overall, SIRV2 represents an extremely stable and structurally interesting VNP with the potential for novel biotechnological applications (Steinmetz et al. 2008a).

#### 10.5.4 Perspectives

Enzymes from thermophilic viruses offer the opportunity to greatly expand the reaction conditions of biocatalysis. Furthermore, unique features in morphology and virus–host interactions may be used in tailored applications in biomedicine and nanotechnology. However, the current repertoire of viral enzymes only hints at their overall potential. The most commonly used enzymes are derived from a surprisingly small number of cultivated viruses, which is remarkable considering the enormous morphological and genomic diversity of viruses revealed over the past decade (Schoenfeld et al. 2010). Developments in the cultivation and production of thermophilic viruses, and developments related to the cloning and expression of their genes in heterologous hosts, will help in the search of novel enzymes with biotechnological/industrial potential. In addition, thermophilic viruses are likely to be useful tools for studying host evolution and host biochemical pathways, providing us with information concerning life in hot environments. Thermophilic viruses have many unique features not observed in other studied viruses; there may be a potential to exploit these features in biomedicine.

Viral genomes are relatively simple compared with those of their hosts and contain a comparatively high proportion of genes coding for structural proteins (e.g. coat and tail) together with proteins involved in nucleic acid metabolism and lysis (Schoenfeld et al. 2010). The density of certain genes in viral genomes offers a great advantage when looking for novel enzymes. For example, a typical bacterial genome of about 2 Mb contains only a single *pol I* gene (coding for DNA polymerase I). By contrast, between 20 and 40 *pol* genes per 2 Mb were found in viral metagenomic sequences. Metagenomic approaches have been used in ecological investigations of hot environments. However, viral metagenomics can also be used to uncover novel biological features of viruses, ultimately, producing useful enzymes. Many challenges are associated with metagenomic-based enzyme discovery. An inherent difficulty in viral metagenomics is isolating adequate amounts of genomic material for large-insert library construction (Edwards and Rohwer 2005). Another challenge arises during sequence assembly. Very large scale sequencing of viral metagenomics should permit the assembly of large contiguous stretches of DNA and potentially entire genomes. However, the high degree of sequence polymorphism within viral populations has largely confounded attempts to assemble large contigs with high confidence (Schoenfeld et al. 2008). Assembly at high stringency tends to prevent misassembling noncontiguous parts of the genome, but such stringent assembly can lead to overestimation of unique viral types and can prevent discovery of genes, enzymes and genomes. The identification of viral genes from sequence assemblies is difficult owing to the high diversity of viral genes and the relatively low numbers of viral genomes in public sequence databases. Moreover, most viral coding sequences (even for a well-studied bacteriophage such as T4) have no similarity to any known genes. This is especially distinct for thermophilic viruses of the Archaea (Prangishvili and Garrett 2005), lowering the likelihood of finding genes by similarity. These problems are being addressed, and viral metagenomics offers a mean of exploring genetic diversity within the vast uncultivated portion of thermophilic viruses.

## 10.6 Conclusions

Thermophilic viruses show a huge potential as useful tools in a broad field of biotechnology. Novel genetic and morphological features, in combination with a huge unexplored diversity, promise new discoveries which are essential to the development of new applications both in industry and basic research.

## References

- Abedon ST (2009) Phage evolution and ecology. *Adv Appl Microbiol* 67:1–45  
Ahn DG, Kim SI et al (2006) TTSV1, a new virus-like particle isolated from the hyperthermophilic crenarchaeote *Thermoproteus tenax*. *Virology* 351(2):280–290

- Anderson RE, Brazelton WJ et al (2011) Using CRISPRs as a metagenomic tool to identify microbial hosts of a diffuse flow hydrothermal vent viral assemblage. *FEMS Microbiol Ecol* 77(1):120–133
- Arnold HP, Ziese U et al (2000a) SNDV, a novel virus of the extremely thermophilic and acidophilic archaeon *Sulfolobus*. *Virology* 272(2):409–416
- Arnold HP, Zillig W et al (2000b) A novel lipothrixvirus, SIFV, of the extremely thermophilic crenarchaeon *Sulfolobus*. *Virology* 267(2):252–266
- Bettstetter M, Peng X et al (2003) AFV1, a novel virus infecting hyperthermophilic archaea of the genus *Acidianus*. *Virology* 315(1):68–79
- Birkenbihl RP, Neef K et al (2001) Holliday junction resolving enzymes of archaeal viruses SIRV1 and SIRV2. *J Mol Biol* 309(5):1067–1076
- Bize A, Peng X et al (2008) Viruses in acidic geothermal environments of the Kamchatka Peninsula. *Res Microbiol* 159(5):358–366
- Bize A, Karlsson EA et al (2009) A unique virus release mechanism in the Archaea. *Proc Natl Acad Sci U S A* 106(27):11306–11311
- Blanco L, Bernad A et al (1989) Highly efficient DNA synthesis by the phage phi 29 DNA polymerase. Symmetrical mode of DNA replication. *J Biol Chem* 264(15):8935–8940
- Blondal T, Hjorleifsdottir SH et al (2003) Discovery and characterization of a thermostable bacteriophage RNA ligase homologous to T4 RNA ligase I. *Nucleic Acids Res* 31(24):7247–7254
- Blondal T, Hjorleifsdottir S et al (2005a) Characterization of a 5'-polynucleotide kinase/3'-phosphatase from bacteriophage RM378. *J Biol Chem* 280(7):5188–5194
- Blondal T, Thorisdottir A et al (2005b) Isolation and characterization of a thermostable RNA ligase I from a *Thermus scotoductus* bacteriophage TS2126 with good single-stranded DNA ligation properties. *Nucleic Acids Res* 33(1):135–142
- Bolduc B, Shaughnessy DP et al (2012) Identification of novel positive-strand RNA viruses by metagenomic analysis of archaea-dominated yellowstone hot springs. *J Virol* 86(10):5562–5573
- Chen R, Wang H et al (2002) Roles of uracil-DNA glycosylase and dUTPase in virus replication. *J Gen Virol* 83(Pt 10):2339–2345
- Collier LH, Oxford JS (2006) *Human virology: a text for students of medicine, dentistry, and microbiology*. Oxford University Press, Oxford/New York
- Comeau AM, Bertrand C et al (2007) Modular architecture of the T4 phage superfamily: a conserved core genome and a plastic periphery. *Virology* 362(2):384–396
- Contursi P, Jensen S et al (2006) Characterization of the *Sulfolobus* host-SSV2 virus interaction. *Extremophiles* 10(6):615–627
- De Paepe M, Taddei F (2006) Viruses' life history: towards a mechanistic basis of a trade-off between survival and reproduction among phages. *PLoS Biol* 4(7):e193
- Douglas T, Young M (1998) Host-guest encapsulation of materials by assembled virus protein cages. *Nature* 393(6681):152–155
- Edwards RA, Rohwer F (2005) Viral metagenomics. *Nat Rev Microbiol* 3(6):504–510
- EpiCentre (2012) *CircLigase™ ssDNA ligase*. <http://www.epibio.com/item.asp?id=445>. Retrieved 8 June 2012
- Ettema TJ, Bernander R (2009) Cell division and the ESCRT complex: a surprise from the archaea. *Commun Integr Biol* 2(2):86–88
- Frols S, Gordon PM et al (2007) Elucidating the transcription cycle of the UV-inducible hyperthermophilic archaeal virus SSV1 by DNA microarrays. *Virology* 365(1):48–59
- Fulton J, Bothner B et al (2009) Genetics, biochemistry and structure of the archaeal virus STIV. *Biochem Soc Trans* 37:114–117
- Galtier N, Lobry JR (1997) Relationships between genomic G+C content, RNA secondary structures, and optimal growth temperature in prokaryotes. *J Mol Evol* 44(6):632–636
- Garrett RA, Prangishvili D et al (2010) Metagenomic analyses of novel viruses and plasmids from a cultured environmental sample of hyperthermophilic neutrophiles. *Environ Microbiol* 12(11):2918–2930
- Geslin C, Le Romancer M et al (2003) PAV1, the first virus-like particle isolated from a hyperthermophilic euryarchaeote, “*Pyrococcus abyssi*”. *J Bacteriol* 185(13):3888–3894



- Gorlas A, Koonin EV et al (2012) TPV1, the first virus isolated from the hyperthermophilic genus *Thermococcus*. *Environ Microbiol* 14(2):503–516
- Happonen LJ, Redder P et al (2010) Familial relationships in hyperthermo- and acidophilic archaeal viruses. *J Virol* 84(9):4747–4754
- Haring M, Peng X et al (2004) Morphology and genome organization of the virus PSV of the hyperthermophilic archaeal genera *Pyrobaculum* and *Thermoproteus*: a novel virus family, the Globuloviridae. *Virology* 323(2):233–242
- Haring M, Rachel R et al (2005) Viral diversity in hot springs of Pozzuoli, Italy, and characterization of a unique archaeal virus, acidianus bottle-shaped virus, from a new family, the Ampullaviridae. *J Virol* 79(15):9904–9911
- Hendrix RW, Smith MCM et al (1999) Evolutionary relationships among diverse bacteriophages and prophages: all the world's a phage. *Proc Natl Acad Sci U S A* 96(5):2192–2197
- Hilario E (2004) End labeling procedures: an overview. *Mol Biotechnol* 28(1):77–80
- HotZyme (2012) The HotZyme project. <http://hotzyme.com/project/>. Retrieved 8 June 2012
- Hsia KC, Li CL et al (2005) Structural and functional insight into sugar-nonspecific nucleases in host defense. *Curr Opin Struct Biol* 15(1):126–134
- Hurst LD, Merchant AR (2001) High guanine-cytosine content is not an adaptation to high temperature: a comparative analysis amongst prokaryotes. *Proc R Soc Lond B Biol Sci* 268(1466):493–497
- Jaatinen ST, Happonen LJ et al (2008) Biochemical and structural characterisation of membrane-containing icosahedral dsDNA bacteriophages infecting thermophilic *Thermus thermophilus*. *Virology* 379(1):10–19
- Janekovic D, Wunderl S et al (1983) Ttv1, Ttv2 And Ttv3, a family of viruses of the extremely thermophilic, anaerobic, sulfur reducing archaeobacterium *Thermoproteus-tenax*. *Mol Gen Genet* 192(1–2):39–45
- Kinoshita Y, Nishigaki K et al (1997) Fluorescence-, isotope- or biotin-labeling of the 5' end of single-stranded DNA/RNA using T4 RNA ligase. *Nucleic Acids Res* 25(18):3747–3748
- Krupovic M, Prangishvili D et al (2011) Genomics of bacterial and archaeal viruses: dynamics within the prokaryotic virosphere. *Microbiol Mol Biol Rev* 75(4):610–635
- Krupovic M, Peixeiro N et al (2012) Archaeal tetrathionate hydrolase goes viral: secretion of a sulfur metabolism enzyme in the form of virus-like particles. *Appl Environ Microbiol* 78:5463–5465
- Larson ET, Reiter D et al (2006) Structure of A197 from *Sulfolobus turreted* icosahedral virus: a crenarchaeal viral glycosyltransferase exhibiting the GT-A fold. *J Virol* 80(15):7636–7644
- Lawrence HM, Merivuori H et al (1986) Preliminary characterization of bacteriophages infecting the thermophilic actinomycete thermomonospora. *Appl Environ Microbiol* 52(4):631–636
- Legerski RJ, Hodnett JL et al (1978) Extracellular nucleases of pseudomonas BAL 31. III. Use of the double-strand deoxyriboexonuclease activity as the basis of a convenient method for the mapping of fragments of DNA produced by cleavage with restriction enzymes. *Nucleic Acids Res* 5(5):1445–1464
- Lengeler JW, Drews G et al (1999) *Biology of the prokaryotes*. Thieme, Stuttgart/New York. Malden, MA, Distributed in the USA by Blackwell Science
- Lillehaug JR, Kleppe K (1977) Phosphorylation of tRNA by T4 polynucleotide kinase. *Nucleic Acids Res* 4(2):373–380
- Lin LB, Hong W et al (2010) Isolation and characterization of an extremely long tail *Thermus* bacteriophage from Tengchong hot springs in China. *J Basic Microbiol* 50(5):452–456
- Lin LB, Han J et al (2011) Isolation and characterization of a new bacteriophage MMP17 from *Meiothermus*. *Extremophiles* 15(2):253–258
- Liu X, Gorovsky MA (1993) Mapping the 5' and 3' ends of *Tetrahymena thermophila* mRNAs using RNA ligase mediated amplification of cDNA ends (RLM-RACE). *Nucleic Acids Res* 21(21):4954–4960
- Liu B, Zhang XB (2008) Deep-sea thermophilic *Geobacillus* bacteriophage GVE2 transcriptional profile and proteomic characterization of virions. *Appl Microbiol Biotechnol* 80(4):697–707

- Liu B, Wu SJ et al (2006) Two novel bacteriophages of thermophilic bacteria isolated from deep-sea hydrothermal fields. *Curr Microbiol* 53(2):163–166
- Liu B, Zhou FF et al (2009) Genomic and proteomic characterization of a thermophilic *Geobacillus* bacteriophage GBSV1. *Res Microbiol* 160(2):166–171
- Liu B, Wu SJ et al (2010) Complete genome sequence and proteomic analysis of a thermophilic bacteriophage BV1. *Acta Oceanol Sin* 29(3):84–89
- Madigan MT, Brock TD (2009) Brock biology of microorganisms. Pearson/Benjamin Cummings, San Francisco
- Marraffini LA, Sontheimer EJ (2010) CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat Rev Genet* 11(3):181–190
- Matsushita I, Yanase H (2009) The genomic structure of thermus bacteriophage phi IN93. *J Biochem* 146(6):775–785
- Meile L, Jenal U et al (1989) Characterization of Psi-M1, a virulent phage of *Methanobacterium thermoautotrophicum* Marburg. *Arch Microbiol* 152(2):105–110
- Minakhin L, Goel M et al (2008) Genome comparison and proteomic characterization of *Thermus thermophilus* bacteriophages P23-45 and P74-26: siphoviruses with triplex-forming sequences and the longest known tails. *J Mol Biol* 378(2):468–480
- Mochizuki T, Yoshida T et al (2010) Diversity of viruses of the hyperthermophilic archaeal genus *Aeropyrum*, and isolation of the *Aeropyrum pernix* bacilliform virus 1, APBV1, the first representative of the family Clavaviridae. *Virology* 402(2):347–354
- Moser MJ, Difrancesco RA et al (2012) Thermostable DNA polymerase from a viral metagenome is a potent rt-PCR enzyme. *PLoS One* 7(6):e38371
- Muskhelishvili G, Palm P et al (1992) SSV1-encoded site-specific recombination system in *Sulfolobus shibatae*. *Mol Gen Genet* 237:334–342
- Ortmann AC, Brumfield SK et al (2008) Transcriptome analysis of infection of the archaeon *Sulfolobus solfataricus* with *Sulfolobus turreted* icosahedral virus. *J Virol* 82(10):4874–4883
- Palm P, Schleper C et al (1991) Complete nucleotide-sequence of the virus Ssv1 of the archaeobacterium *Sulfolobus shibatae*. *Virology* 185(1):242–250
- Pederson DM, Welsh LC et al (2001) The protein capsid of filamentous bacteriophage PH75 from *Thermus thermophilus*. *J Mol Biol* 309(2):401–421
- Peng X (2008) Evidence for the horizontal transfer of an integrase gene from a fusellovirus to a pRN-like plasmid within a single strain of *Sulfolobus* and the implications for plasmid survival. *Microbiol-Sgm* 154:383–391
- Peng X, Basta T et al (2007) Genome of the *Acidianus* bottle-shaped virus and insights into the replication and packaging mechanisms. *Virology* 364(1):237–243
- Pfister P, Wesserfallen A et al (1998) Molecular analysis of *Methanobacterium* phage Psi M2. *Mol Microbiol* 30(2):233–244
- Pina M, Bize A et al (2011) The archeoviruses. *FEMS Microbiol Rev* 35(6):1035–1054
- Prangishvili D, Garrett RA (2005) Viruses of hyperthermophilic Crenarchaea. *Trends Microbiol* 13(11):535–542
- Prangishvili D, Quax TEF (2011) Exceptional virion release mechanism: one more surprise from archaeal viruses. *Curr Opin Microbiol* 14(3):315–320
- Prangishvili D, Klenk HP et al (1998) Biochemical and phylogenetic characterization of the dUTPase from the archaeal virus SIRV. *J Biol Chem* 273(11):6024–6029
- Prangishvili D, Arnold HP et al (1999) A novel virus family, the Rudiviridae: structure, virus-host interactions and genome variability of the *Sulfolobus* viruses SIRV1 and SIRV2. *Genetics* 152(4):1387–1396
- Prangishvili D, Vestergaard G et al (2006) Structural and genomic properties of the hyperthermophilic archaeal virus ATV with an extracellular stage of the reproductive cycle. *J Mol Biol* 359(5):1203–1216
- Pride DT, Schoenfeld T (2008) Genome signature analysis of thermal virus metagenomes reveals Archaea and thermophilic signatures. *BMC Genomics* 9:420
- Prokazyme (2012) ThermoPhage™ Polynucleotide kinase. <http://www.prokazyme.com/ProductSheets/Thermophage-PNK-version-3.0.pdf>. Retrieved 8 June 2012



- Rangarajan S, Shankar V (1999) Extracellular nuclease from *Rhizopus stolonifer*: purification and characteristics of – single strand preferential – deoxyribonuclease activity. *Biochim Biophys Acta* 1473(2–3):293–304
- Redder P, Peng X et al (2009) Four newly isolated fuselloviruses from extreme geothermal environments reveal unusual morphologies and a possible interviral recombination mechanism. *Environ Microbiol* 11(11):2849–2862
- Reiter D, Palm P et al (1987) Identification and characterization of the genes encoding three structural proteins of the *Sulfolobus* virus-like particle SSV1. *Mol Gen Genet* 206:144–153
- Rice G, Tang L et al (2004) The structure of a thermophilic archaeal virus shows a double-stranded DNA viral capsid type that spans all domains of life. *Proc Natl Acad Sci U S A* 101(20):7716–7720
- Rodriguez I, Lazaro JM et al (2005) A specific subdomain in phi29 DNA polymerase confers both processivity and strand-displacement capacity. *Proc Natl Acad Sci U S A* 102(18):6407–6412
- Rosario K, Breitbart M (2011) Exploring the viral world through metagenomics. *Curr Opin Virol* 1(4):289–297
- Saiki RK, Gelfand DH et al (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239(4839):487–491
- Schleper C, Kubo K et al (1992) The particle SSV1 from the extremely thermophilic archaeon *Sulfolobus* is a virus: demonstration of infectivity and of transfection with viral DNA. *Proc Natl Acad Sci U S A* 89(16):7645–7649
- Schoenfeld T, Patterson M et al (2008) Assembly of viral metagenomes from yellowstone hot springs. *Appl Environ Microbiol* 74(13):4164–4174
- Schoenfeld T, Liles M et al (2010) Functional viral metagenomics and the next generation of molecular tools. *Trends Microbiol* 18(1):20–29
- She Q, Peng X et al (2001) Gene capture in archaeal chromosomes. *Nature* 409(6819):478
- Silber R, Malathi VG et al (1972) Purification and properties of bacteriophage T4-induced RNA ligase. *Proc Natl Acad Sci U S A* 69(10):3009–3013
- Singh J, Behal A et al (2009) Metagenomics: concept, methodology, ecological inference and recent advances. *Biotechnol J* 4(4):480–494
- Song Q, Zhang X (2008) Characterization of a novel non-specific nuclease from thermophilic bacteriophage GBSV1. *BMC Biotechnol* 8:43
- Sorek R, Cossart P (2010) Prokaryotic transcriptomics: a new view on regulation, physiology and pathogenicity. *Nat Rev Genet* 11(1):9–16
- Speir JA, Munshi S et al (1995) Structures of the native and swollen forms of cowpea chlorotic mottle virus determined by X-ray crystallography and cryo-electron microscopy. *Structure* 3(1):63–78
- Stahl CP, Zucker-Franklin D et al (1991) Effects of human interleukin-6 on megakaryocyte development and thrombocytopoiesis in primates. *Blood* 78(6):1467–1475
- Stedman KM, She QX et al (2003) Relationships between fuselloviruses infecting the extremely thermophilic archaeon *Sulfolobus*: SSV1 and SSV2. *Res Microbiol* 154(4):295–302
- Steinmetz NF, Bize A et al (2008a) Site-specific and spatially controlled addressability of a new viral nanobuilding block: *Sulfolobus islandicus* rod-shaped virus 2. *Adv Funct Mater* 18:3478–3486
- Steinmetz NF, Findlay KC et al (2008b) Layer-by-layer assembly of viral nanoparticles and polyelectrolytes: the film architecture is different for spheres versus rods. *Chembiochem* 9(10):1662–1670
- Tabor S, Richardson CC (1987) DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. *Proc Natl Acad Sci U S A* 84(14):4767–4771
- Tama F, Brooks CL III (2002) The mechanism and pathway of pH induced swelling in cowpea chlorotic mottle virus. *J Mol Biol* 318(3):733–747
- Tamakami M, Murakami A et al (2011) Genomic and proteomic characterization of the large Myoviridae bacteriophage varphiTMA of the extreme thermophile *Thermus thermophilus*. *Bacteriophage* 1(3):152–164
- Valentine DL (2007) Adaptations to energy stress dictate the ecology and evolution of the Archaea. *Nat Rev Microbiol* 5(4):316–323

- van den Burg B (2003) Extremophiles as a source for novel enzymes. *Curr Opin Microbiol* 6(3):213–218
- Vestergaard G, Haring M et al (2005) A novel ruidivirus, ARV1, of the hyperthermophilic archaeal genus *Acidianus*. *Virology* 336(1):83–92
- Vestergaard G, Aramayo R et al (2008a) Structure of the *Acidianus* filamentous virus 3 and comparative genomics of related archaeal lipothrixviruses. *J Virol* 82(1):371–381
- Vestergaard G, Shah SA et al (2008b) Stygiolobus rod-shaped virus and the interplay of crenarchaeal ruidiviruses with the CRISPR antiviral system. *J Bacteriol* 190(20):6837–6845
- Walther J, Sierocinski P et al (2011) Hot transcriptomics. *Archaea* 2010:897585
- Wang YQ, Zhang XB (2010) Genome analysis of deep-Sea thermophilic phage D6E. *Appl Environ Microbiol* 76(23):7861–7866
- Wang LK, Lima CD et al (2002) Structure and mechanism of T4 polynucleotide kinase: an RNA repair enzyme. *EMBO J* 21(14):3873–3880
- Wei D, Zhang X (2010) Proteomic analysis of interactions between a deep-sea thermophilic bacteriophage and its host at high temperature. *J Virol* 84(5):2365–2373
- Weinbauer MG, Rassoulzadegan F (2004) Are viruses driving microbial diversification and diversity? *Environ Microbiol* 6(1):1–11
- Wiedenheft B, Stedman K et al (2004) Comparative genomic analysis of hyperthermophilic archaeal Fuselloviridae viruses. *J Virol* 78(4):1954–1961
- Wiedenheft B, Flenniken ML et al (2007) Bioprospecting in high temperature environments; application of thermostable protein cages. *Soft Matter* 3:1091–1098
- Wollert T, Wunder C et al (2009) Membrane scission by the ESCRT-III complex. *Nature* 458(7235):172–177
- Xiang XY, Chen LM et al (2005) *Sulfolobus tengchongensis* spindle-shaped virus STSV1: Virus-host interactions and genomic features. *J Virol* 79(14):8677–8686
- Yu MX, Slater MR et al (2006) Isolation and characterization of *Thermus* bacteriophages. *Arch Virol* 151(4):663–679
- Zhang L, Cui X et al (1992) Whole genome amplification from a single cell: implications for genetic analysis. *Proc Natl Acad Sci U S A* 89(13):5847–5851
- Zhu H, Yin S et al (2004) Characterization of polynucleotide kinase/phosphatase enzymes from *Mycobacteriophages* omega and Cjw1 and vibriophage KVP40. *J Biol Chem* 279(25):26358–26369
- Zillig W, Kletzin A et al (1994) Screening for *sulfolobales*, their plasmids and their viruses in Icelandic solfataras. *Syst Appl Microbiol* 16(4):609–628