Extremophiles (1998) 2:131–140

REVIEW

Wolfram Zillig · Hans Peter Arnold · Ingelore Holz
David Prangishvili · Anja Schweier · Kenneth Stedman
Qunxin She · Hien Phan · Roger Garrett
Jakob K. Kristjansson

Genetic elements in the extremely thermophilic archaeon *Sulfolobus*

Received: January 22, 1998 / Accepted: February 16, 1998

Abstract This minireview summarizes what is known about genetic elements in the archaeal crenarchaeotal genus *Sulfolobus*, including recent work on viruses, cryptic plasmids, a novel type of virus satellite plasmids or satellite viruses, and conjugative plasmids (CPs), mostly from our laboratory. It does not discuss IS elements and transposons.

Key words Archaea · Crenarchaeota · *Sulfolobus* · Virus · Conjugative plasmid · Satellite virus · Vectors

Introduction

In the introduction to a previous review of our knowledge of viruses and plasmids in thermophilic and hyperthermophilic Archaea, we have called such elements "fragments of life" (Zillig et al. 1996). Rather than fragments formed by breakdown, however, they should be regarded as assemblies of modules from the pool of matter and function that also gave rise to organismic life. Studying their diversity thus should help in understanding early, perhaps even prebiotic, evolution. Viruses and plasmids have also provided model genomes for studying gene expression and its controls, as exemplified by the recognition of the archaeal transcription promoter in the *Sulfolobus* virus SSV1 (Reiter et al. 1988a). Moreover, they have been sources of enzymes, particularly of DNA and RNA polymerases and other tools for biotechnology (e.g., T7 RNA polymerase). Finally, they have been used as a basis in the development of multiple cloning and expression vectors, e.g., from phages λ and M13.

In view of this potential it appears astonishing that to date only our research group has undertaken a systematic search for viruses and plasmids in Crenarchaeota, although a fair number of euryarchaeotal genetic elements has been described and utilized by several laboratories (for review, see Zillig et al. 1988 and Reiter et al. 1988b). Genetic systems for Crenarchaeota are badly needed to develop the molecular genetics of this archaeal kingdom as well as to express the genes of its extremely thermophilic members that are refractory to expression in *Escherichia coli*, including correct folding and modification. After some initial work on viruses of *Thermoproteus* (Janekovic et al. 1983), we have decided to restrict ourselves to looking for genetic elements in the also extremely thermophilic but, in contrast to other representatives of this kingdom, aerobic and heterotrophic genus *Sulfolobus*. Many, although certainly not most, *Sulfolobus* strains can be grown on Gelrite plates (Zillig et al. 1994), both as single colonies allowing the isolation of pure strains and as lawns facilitating plaque tests, both prerequisites for genetic work.

Screening for genetic elements

Samples were taken from hydrothermal situations in solfataric fields in Iceland, Japan, New Zealand, and Italy, usually from sources with temperatures between 75° and 95°C and pH values between 1 and 5. The samples were adjusted to pH 5 and stored anaerobically to increase the viability of *Sulfolobus* cells. Optimally, the samples were plated within a few days either directly or after enrichment in liquid culture on Gelrite gels with 0.2% tryptone as sole carbon source. Total DNA and covalently closed circular (cccDNA) (Birnboim and Doly 1979) were prepared from

Communicated by W.G. Grant

W. Zillig (✉) · H.P. Arnold · I. Holz · D. Prangishvili · A. Schweier · K. Stedman
Max Planck Institut für Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, Germany
Tel. +49-89-85782223; Fax +49-89-85782728
e-mail: zillig@biochem.mpg.de

Q. She · H. Phan · R. Garrett
Institute of Molecular Biology, Copenhagen University, Solvgade 83 H DK-1307, Copenhagen K, Denmark

J.K. Kristjansson
University of Iceland, Keldnaholt, IS 112, Reykjavik, Iceland

cultures inoculated with single colonies. According to restriction fragment patterns of the chromosomal DNAs, these isolates fall into four "geographic" groups. Isolates from Pisciarelli, Italy, were all of the "*solfataricus*" type, clearly distinguished from the well-defined "*islandicus*" type comprising all strains from Iceland and from the more heterogeneous *japonicus* and *neozealandicus* strains. DNA-DNA cross-hybridization and 16S rRNA sequences (H.P. Klenk, personal communiation) showed, however, that all these isolates resemble the type strain of *S. solfataricus*, P1, DSM 1616 (Zillig et al. 1980), to an extent questioning their separation from the latter as distinct species (Grogan et al. 1990; Zillig et al. 1994).

Samples and cultures yielding conspicuous DNA restriction fragment patterns were spotted onto lawns of potential virus hosts to screen for plaque-forming elements. The supernatants of positive cultures were inspected for virus-like particles with an electron microscope. Hosts were selected by comparing virus production in liquid cultures, and plaque tests were developed with suitable hosts. Plasmids were identified by inspection of DNA restriction fragment patterns. Conjugative plasmids were identified by their capacity to be transferred directly and efficiently from cell to cell, and not via diffusible particles in supernatants as shown by the absence of inhibition zones in spot tests (Schleper et al. 1995). In this way, more than 30 novel viruses and more than 25 plasmids were found in *Sulfolobus* isolates, which are all closely related to *S. solfataricus* by DNA-DNA cross-hybridization and 16S rRNA sequence. All novel viruses were found in carrier states. We screened extensively for free viruses in polyethylene glycol (PEG) precipitates of supernatants of samples. Although we saw virus-like particles of various shapes in electron microscopic analysis, in a few cases attached to cells of *Thermoproteus* (e.g., Zillig et al. 1994), we were unable to isolate viable viruses from these specimens. Nine of the plasmids were cryptic, meaning that their genetic potential is unknown. Five were closely related to each other and appear to be satellites of fuselloviruses, and 15 were conjugative.

The viruses

Although the majority of the bacteriophages and also the viruses of Euryarchaeota are of the classical head-and-tail type, all crenarchaeotal viruses known to date belong to three or four novel virus families, which were created to account for their unique features (Table 1).

The *Fuselloviridae* comprise spindle ("lemon") -shaped viruses (Fig. 1a) with a core consisting of cccDNA, about 15 kbp in size, and a basic DNA-binding protein wrapped into an envelope containing at least two hydrophobic coat proteins and probably host lipids (Martin et al. 1984; Zillig et al. 1988; Reiter et al. 1988b). The virions carry short tails with tail fibers on one end and absorb during infection to the membrane of the host. The sequence of the genome of SSV1 from the Japanese lysogen *S. shibatae* is completely known (Palm et al. 1991) and that of the icelandic SSV3 has been partially determined (Q. She, R. Garrett, and K. Stedman, unpublished). Together with the strains carrying SSV3 and the partially characterized SSV2, 32 of 395 novel *Sulfolobus* isolates, i.e., 8% (many of them from southwest Iceland), were shown to harbor SSVs (I. Holz, W. Zillig, D. Prangishvili, H.P. Arnold, and K. Stedman, unpublished results). Thus, fuselloviruses appear widespread and frequent. The DNAs of the novel SSVs show high similarity to each other and lower similarity to SSV1 in DNA-DNA cross-hybridization according to Southern analysis. All SSVs studied in this respect appear to be temperate, and their proviruses seem to be integrated into the host chromosome similar to that of SSV1, which is inserted into an arginyl tRNA gene in *S. shibatae* (Reiter et al. 1989; Reiter and Palm 1990). The production of SSV1 (Martin et al. 1984) and some of the novel Icelandic SSVs (unpublished) is induced by UV irradiation. The cells are neither lysed nor killed and eventually return to the lysogenic state. In contrast to lysogeny of *E. coli* carrying the integrated λ prophage, where only the repressor gene is expressed and all others are inducible, in SSV1 lysogens all structural virus

Table 1. Viruses of Crenarchaeota

Virus	Shape	Size (length/ width, nm)	Type	dsDNA size (kbp)	Sequence known	Proteins	Coat	Host	Special features	Family
SSV1	Spindle	90/60	ccc	15.5	Total	>3	+	*S. shibatae*	–	*Fuselloviridae*
SSV2	Spindle	80/55	ccc	15	On way	?	+	*S. islandicus*	Helper of SSVx	″
SSV3	Spindle	80/55	ccc	15	Partial	?	+	*S. islandicus*	–	″
SSVx	Spindle	60/40	ccc	5.7	Total	?	+	*S. islandicus*	Satellite of SSV2	Unassigned
TTV1	Short filament	400/40	Linear	16	85%	4	+	*T. tenax*	–	*Lipothrixviridae*
TTV2	Flex. filament	1200/20	Linear	16	–	≥1	+	*T. tenax*	–	″
TTV3	Flex. filament	2500/30	Linear	27	–	≥1	+	*T. tenax*	–	″
DAFV	Flex. filament	2200/27	Linear	40	–	≥3	+	*A. ambivalens*	–	″
SIFV	Flex. filament	1950/24	Linear	42	94%	≥6	+	*S. islandicus*	Blocked DNA termini	″
TTV4	Stiff rod	500/30	Linear	17	–	2	+*	*T. tenax*		Unassigned
SIRV1[#]	Stiff rod	780/23	Linear	33	Total	≥5	−	*S. islandicus*	Covalently closed DNA termini	*Rudiviridae*
SIRV2	Stiff rod	900/23	Linear	36	–	≥5	−	*S. islandicus*		″
SNDV	Bearded droplet	145/75	ccc	20	On way	≥3	+	*Sulfolobus* sp. (New Zealand)	Modified DNA	Unassigned

Flex., flexible; *protein only; [#] sequenced variant; ccc, covalently closed circular; *S.*, *Sulfolobus*; *T.*, *Thermoproteus*; *A.*, *Acidianus*.

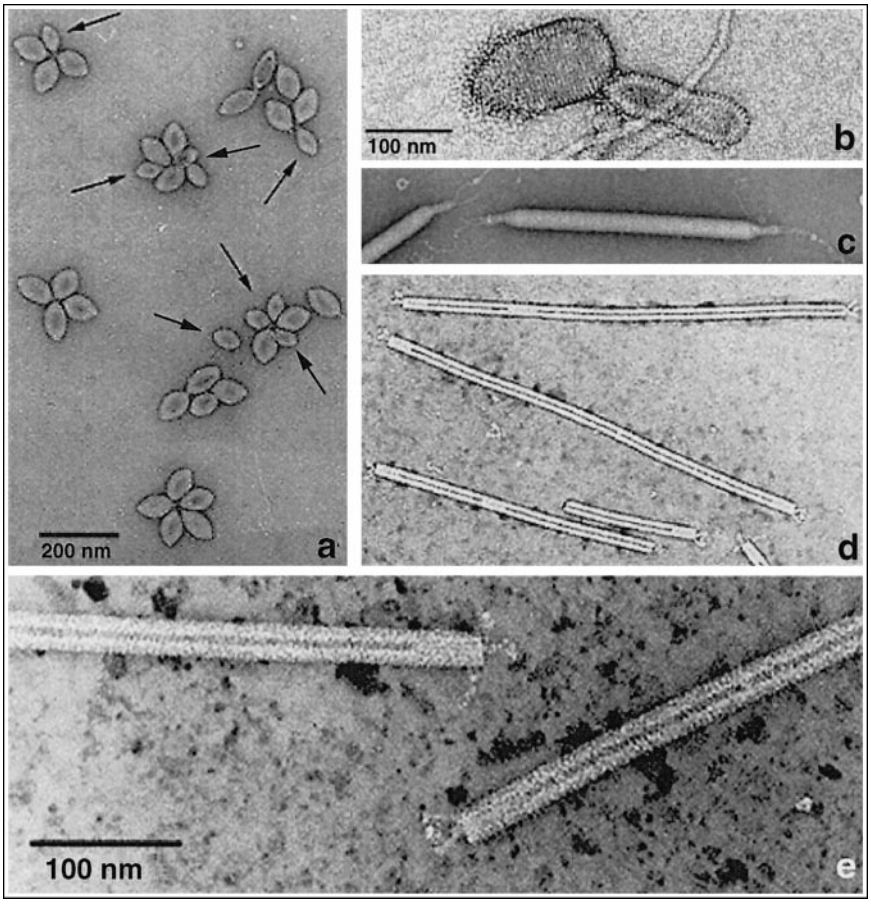

Fig. 1a–e. Electron photomicrographs of a purified preparation of the fusellovirus SSV2 (**a**) shows "rosettes" containing large SSV2 and small SSVx particles (*arrows*); the rudivirus SIRV2 (**d**, **e**); the unassigned viruses TTV4 (**c**); and SNDV (floating genus *Guttavirus*; **b**), all negatively stained. All except **b** and **e**, same magnification

genes are expressed (Reiter et al. 1987b). UV induction inactivates a host repressor of the short transcription unit, which is situated between two large "back-to-back" transcription units that in turn are switched on. The result is the replication of virus DNA, thus allowing the assembly of virions (Singer 1993). The UV induction and the integration of the provirus resemble those in lysogens carrying lambdoid prophages. The production of SSV2 is constitutive and appears enhanced when its host REY15/4 enters the stationary growth phase (Arnold 1998).

Two very similar viruses, SIRV1 (Götz 1994; Zillig et al. 1994) and SIRV2 (Arnold 1998; D. Prangishvili, H.P. Arnold, U. Ziese, and W. Zillig, manuscript in preparation), represent the stiff rod-shaped, fragile *Rudiviridae* (Fig. 1d,e), which contain linear, double-stranded DNA that forms a tubelike double helix with one DNA-binding protein (see Fig. 3a). One turn of the SIRV2 superhelix measures 4.3 nm and comprises 16.5 turns of B DNA. The DNA is thus compressed by a factor of 12.7. The tube is plugged on both ends and each plug carries three tail fibers (Fig. 1e). A coat is absent. Besides the DNA-binding protein, a protein apparently forming the plugs and three or four minor proteins probably involved in forming the tail fibers have been identified (Arnold 1997). At both ends, the complementary strands of the DNA are linked by phosphodiester bonds creating hairpins (H. Blum, personal communication). The viral DNA is thus insensitive to the 5′-specific λ exonuclease but sensitive to the 3′-specific BAL31 exonuclease, which is known to initiate exonucleolytic breakdown at hairpins (Arnold 1997; Rohozinski et al. 1989). The length of the virus particles correlates with the size of the DNA. Although SIRV2 does not infect the icelandic *Sulfolobus* isolate REN2H1, wild-type SIRV1, which is entirely stable in its original carrier KVM10H3, can break this barrier to infection with a plaque-forming efficiency of 10^{-5}. The resulting virus was a mixture of variants, half of which proved stable upon further transfer while the others remained variable but yielded further stable variants upon repeated transfer. Although it is clear that infectious variants are generated (or present) with low frequency, the mode of their formation remains unclear. The sequence diversity in corresponding open reading frames (ORFs) of different variants is, however, up to 10% at the level of nucleotide sequence (D. Prangishvili, H.P. Arnold, U. Ziese, and W. Zilling, manuscript in preparation).

A novel *Sulfolobus* virus, SIFV (Arnold 1998; H.P. Arnold, I. Holz, U. Ziese, and W. Zillig, manuscript in preparation; Fig. 2d), the already described DAFV from *Acidianus* (*Desulfurolobus ambivalens*; Zillig et al. 1994) (Fig. 2c) and the *Thermoproteus tenax* viruses TTV1 (Fig. 2a), TTV2 (Fig. 2b), TTV3 (Janekovic et al. 1983; Rettenberger 1990), and possibly TTV4 (Rettenberger 1990) (Fig. 1c) belong to the *Lipothrixviridae*. Their virions are normally flexible filaments, ranging in length from 0.4 to more than 2 μm and in width from 20 to 40 nm, and contain linear double-stranded DNA. Both termini of SIFV DNA

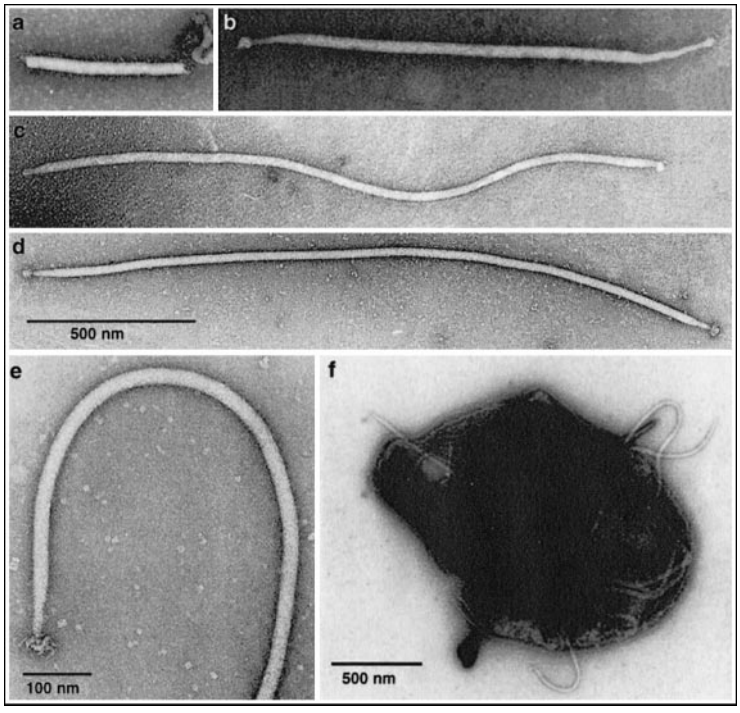

Fig. 2a–f. Electron photomicrographs of *Lipothrixviridae*. TTV1 (**a**), TTV2 (**b**), DAFV (**c**), SIFV (**d**), magnified moplike terminal structure of SIFV (**e**) and DAFV released from host cell (**f**). All negatively stained. **a**–**d**, same magnification

are blocked against digestion by 5′- and 3′-specific exonucleases. TTV1 is short and less flexible; TTV4 is a short, quite stiff, rod. All lipothrixviruses have envelopes. In TTV1 and SIFV these envelopes contain hydrophobic proteins and host lipids, the latter in compositions differing from those in the host membranes. TTV4 has a protein coat devoid of lipids. The core structure differs in different lipothrixiviruses. In TTV1 it is a superhelix made up of the DNA and two different DNA-binding proteins resembling the DNA-protein superhelix in SIRV2 (Fig. 3a). In SIFV, the DNA-protein core looks like a zipper with two rows of doughnut-shaped subunits shifted and tilted relative to each other, the latter at an angle of 30° (Fig. 3b,c). The DNA runs around and along the outer contours of the zipper as a superhelix with a period of 5 nm comprising 11 turns of B DNA, such that the length of the DNA is compressed by a factor of 7.5. This structure results in a noncircular cross section of the virion, 24 nm wide and 16 nm thick. It is unclear whether the structural differences between the slender wormlike members of the family and the short rods and the different nature of the TTV4 coat suffice to split this family in two.

Another unique virus, SNDV (Fig. 1b), has been found in a novel *Sulfolobus* isolate from New Zealand, STH3/3. It differs in many features from all other *Sulfolobus* viruses. In view of its singularity, a new family has, however, not been created. To list it, we thus propose to assign it to a "floating genus" *Guttavirus*, because it has the shape of a droplet with a dense beard of thin filaments on its pointed end (Zillig et al. 1996; Arnold 1998) (Fig. 1b). Its cccDNA is 20 kbp in size and is extensively modified in an unknown way so that many restriction enzymes cleave it either partially or not at all. The chromosomal DNA of the host is not modified. SNDV can be grown in a laboratory host, STH 1/3 from New Zealand, but a plaque test has so far not been developed. Electron photomicrographs indicate that the coat has a helical structure.

All these viruses were isolated from novel strains that harbor them in carrier states. Curing accompanied by the appearance of free virus has been observed for SIRV, SIFV, and SNDV in one-step growth experiments. In none of the carrier states is the virus–host relation completely understood. In contrast, the true lysogeny of *Sulfolobus* harboring SSVs resembles the lysogeny of bacteria.

Plasmids

Table 2 lists nonconjugative plasmids in novel *Sulfolobus* isolates. The first section shows so-called cryptic plasmids of various sizes of which only size and geographic origin are known (Zillig et al. 1996). Strain TIK4/2, which harbors pTIK4, has the capacity to efficiently outcompete other strains. It is a swarmer and shows chemotaxis (D. Prangishvili and W. Zillig, unpublished results). It has not been shown, however, whether these features depend on the plasmid. One or other of the elements of about 5–7 kbp may end up in the following section (see next paragraph) when their sequences become known.

The second section of Table 2 lists five plasmids, four of which are related according to strong partial sequence similarity (pRN1 and pRN2: Keeling et al. 1996 and EMBL Data Bank accession number SIU93082; SSVx: H.P.

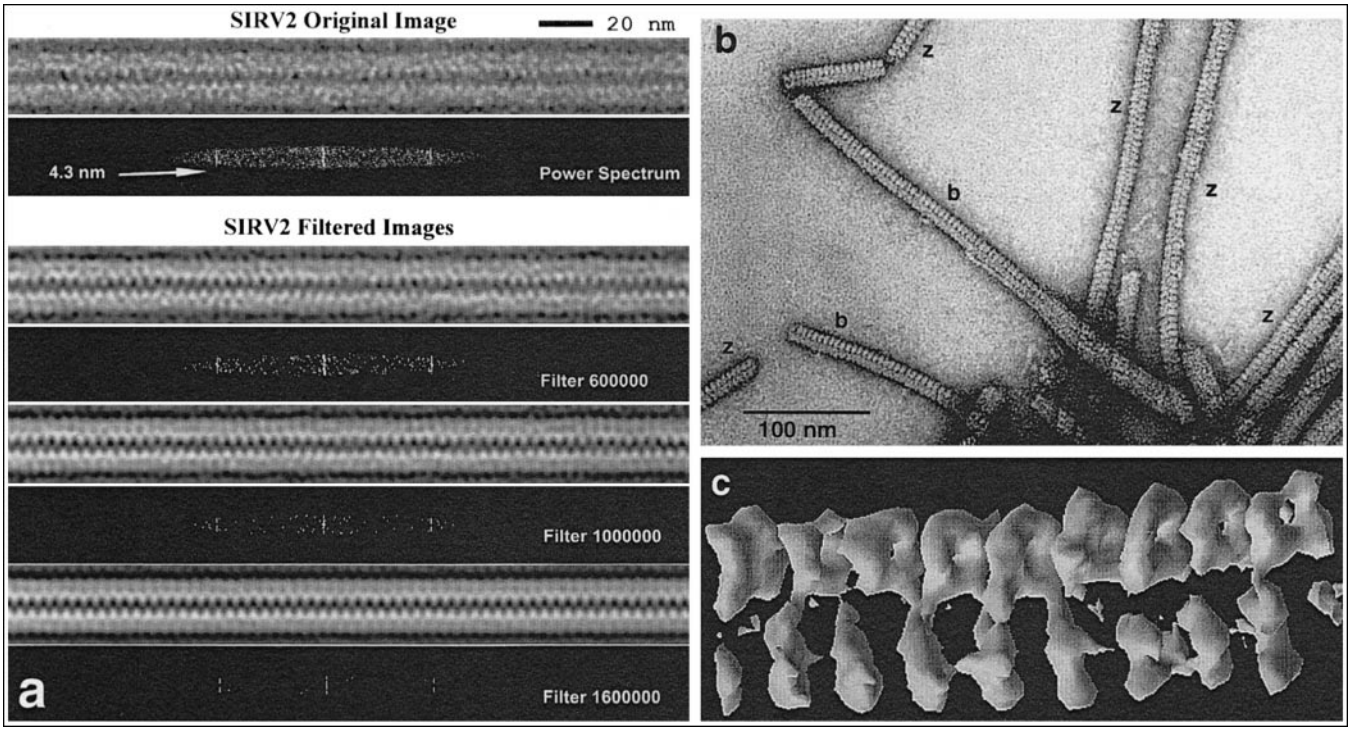

Fig. 3a–c. Structures of SIFV and SIRV. **a** Power spectrum of SIRV. **b** "Stripped" SIFV particles showing the brushlike view (*b*) and the zipper-like view (*z*), negatively stained. **c** Reconstruction of the SIFV core (brush-like view) after electron tomography. (**a** and **c**, courtesy of Ulrike Ziese)

Table 2. Nonconjugative plasmids of *Sulfolobales*

Designation	Host	DNA size (kb)	Copy number	Function
pSTH4	*S. neozealandicus*	4.7	High	Cryptic
pTAV4	*S. neozealandicus*	6.2	High	Cryptic
pWHI1	*S. neozealandicus*	6.8	Low	Cryptic
pWHI2	*S. neozealandicus*	15.4	High	Cryptic
pTIK4	*S. neozealandicus*	14.3	High	? (see text)
pSTH3	*S. neozealandicus*	35.0	High	Cryptic
pIT3	*S. solfataricus*	4.9	~10	Cryptic
pRN1	*S. islandicus*	5.5	20	Cryptic
pRN2	*S. islandicus*	6.9	35	Cryptic
pHE7	*S. islandicus*	7.5	15	Cryptic
pDL10	*Desulfurolobus ambivalens*	7.0	High	Widespread, H_2S aut.
SSVx	*S. islandicus*	5.7	High	SSV2 satellite

The host designations are invalid (see text).
aut., autotrophy; *S.*, *Sulfolobus*.

Arnold, Q. She, R. Garrett, and W. Zillig, manuscript in preparation; and pDL10: A. Kletzin, personal communication) and one, pHEN2, which has not been sequenced yet, that appears to be related according to significant although weak DNA-DNA cross-hybridization (Zillig et al. 1994). Plasmids pRN1 and pRN2 were both found in one isolate, REN1H1 from Reykjanes, southwest Iceland, in high copy number (Zillig et al. 1994). The plasmid (or virus?) SSVx was found as a satellite DNA in a preparation of the fusellovirus SSV2, also from Reykjanes, which was purified by buoyant density gradient centrifugation (Arnold 1998) (Fig. 1a). The particles containing the viral and the satellite DNAs differ in size corresponding to the difference in the size of the DNAs (see Fig. 1a). The SSVx DNA shows only limited homology to the viral DNA (see following) but high sequence similarity to the other plasmids, especially pRN2 (Fig. 4).

Three ORFs in SSVx DNA, 51, 890, and 76, comprising about 50% of the genome, are homologous to ORFs 52, 980, and 81, in the same order, in pRN2. ORF52 shows homology to *cop* genes encoding copy number control proteins in bacterial plasmids, and ORF980 encodes a putative replication protein. The function of ORF81 is unknown. Two ORFs in SSVx immediately downstream of ORF76

show 67% and 31% sequence identity to ORFs a153 and b251, respectively, in the genome of SSV1 (SSV2 has not yet been sequenced). Homologous ORFs are absent in pRN1 and pRN2, suggesting that SSVx is a recombinant of a pRN-like plasmid with a short sequence of an SSV, possibly involved in packaging the DNA into the core which, in the course of morphopoesis of SSV1 (Reiter et al. 1987a), is then wrapped into the hydrophobic virus coat. Accordingly it was shown by transfection experiments that the spread of SSVx in a *S. solfataricus* culture depends on the presence of SSV2 as helper. The plasmid pDL10 was found in 10 of 12

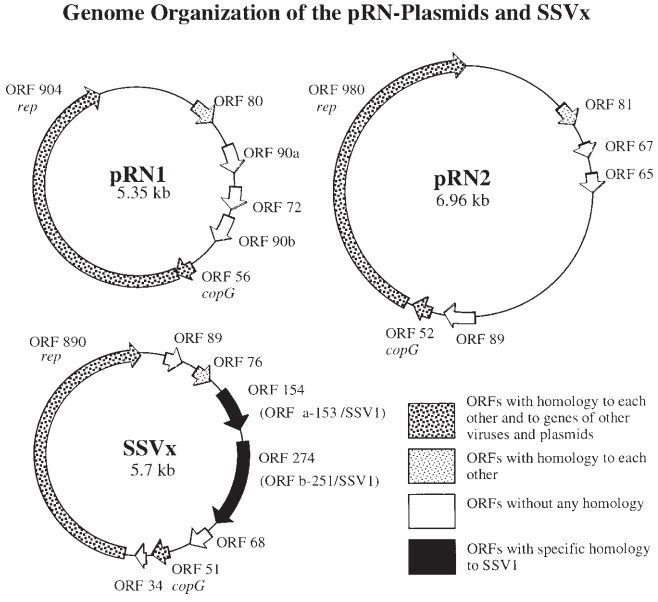

Fig. 4. Schematic presentation of maps showing ORFs in the genomes of the plasmids pRN1 and pRN2 and the satellite SSVx of the fusellovirus SSV2

Acidianus (formerly *Desulfurolobus*) isolates from various locations throughout Iceland, suggesting the capacity to spread. Like SSVx, its sequence shows higher similarity to pRN2 than to pRN1 (A. Kletzin, personal communication). It is amplified in the sulfur-reducing growth mode of its autotrophic hosts. On the basis of (a) the requirement of SSV2 for the spread of SSVx, (b) the independent isolation of an SSV and of homologs of the satellite from the same location, and (c) the wide radiation of pDL10 in Iceland, we suggest that SSVx and possibly pDL10 can generally be packaged into and transported by fuselloviruses. The latter would then be helpers similar to bacteriophage P2, which packages the genome of the parasitic phage P4. Phage P4 shows only very limited homology to phage P2 (Christie and Calendar 1990). The specificity and the extent of the interaction of these satellites with their helper viruses have yet to be elucidated.

Table 3 lists conjugative plasmids (CPs) found in *Sulfolobus* isolates of which two, pNOB8 and pNOB8-33, are from Japan (Schleper et al. 1995) and the others from Iceland (Prangishvili et al., in manuscript). Except for the incomplete pING3, these self-spreading elements are efficiently and directly transferred from donor to recipient cells and not via diffusible particles through the medium. In contrast to some bacterial conjugative systems, pili are apparently not involved in the formation of pairs between donor and recipient, which occurs within 40 min after mixing. The contact involves large areas of the cell surfaces. Transcipients, i.e., recipients carrying the plasmids after transfer, contain multiple additional, probably plasmid-encoded, proteins in their membranes (M. John, unpublished data from this laboratory). Immediately after transfer, the plasmids are efficiently replicated to high copy number in the transcipients. In the course of about 5 h, expression of the plasmid genome establishes the transfer apparatus for the next round of conjugation. A copy number control apparently encoded by the plasmids leads to a

Table 3. Conjugative plasmids of *Sulfolobus*

Designation	Size (kbp)	Host		Remarks	Relationship
		*S.s.*P1	*S.i.*HVE10/4		
pARN3/2	26.1	+	+		Closely Related
pARN4/2	26.5	+	+		
pHVE14/5	36.5	+	−		
pKEF9/1	29.8	+	−		
pING1	24.6	+	+		Variants in subfamily
pING4	25.1	+	+		
pING5	25.1	+	+		
pING6	25.1	+	+		
pING3	6.0	(+)	?	Nonconjugative	
pHVE12/4	27.3	+	−		
pSOG2/4 wt.	28.4	+	+		Variants in subfamily
clone 1	28.7	+	+		
pNOB8	42.6	+	?		del. variant
pNOB8-33	33.0	+	?		

del., deletion; *S.s.*, *Sulfolobus solfataricus*; *S.i.*, *Sulfolobus islandicus*.

progressive reduction of the copy number on growth of the transcipients. In its native host NOB8H1, pNOB8 is stably maintained at low copy number, but in the "foreign" recipient *S. solfataricus* strain P1 the plasmid is lost by curing although it encodes Par functions, apparently because these work only in proper hosts.

Total sequences were determined for the large, apparently complete CP pNOB8 from Japan (She et al., in manuscript) and the smaller, probably incomplete pING1 from Iceland (K. Stedman, I. Holz, Q. She, R. Garrett, and W. Zillig, manuscript in preparation) (see schematic comparison in Fig. 5). About half the sequence of pING1 is very similar to about a third of that of pNOB8. The arrangement of corresponding highly similar ORFs is nearly identical in both genomes. ORF630b in pNOB8 shows significant although weak sequence similarity to ATPases of the TrbE family involved in bacterial and interdomain conjugation. ORF715 of pNOB8, which is homologous to ORF630b, shows weak similarity, especially in motifs, to ATPases of the TraG family that are also involved in conjugative transfer in bacterial systems. This is the only evidence for limited homology between archaeal and bacterial conjugation systems. A variant pNOB8-33, in which about 8 kbp of the pNOB8 genome situated between two exactly repeated 85-bp sequences have been deleted, lacks two ORFs homologous to the ParB and ParA proteins, respectively, both involved in partitioning and thus maintenance of bacterial plasmids and chromosomes. It has however maintained another ORF homologous to ParB and a transposase showing high sequence identity to transposase genes in the *Sulfolobus* chromosome (She et al. 1998). In pING1, the entire region between the TraG homolog and the other end of the region deleted in pNOB8-33, including the second ParB homolog, is missing. Accordingly, pING1 is not stably maintained but is lost by curing on continuous growth of *S. solfataricus* P1 transcipients. Another variant in the pING subfamily of CPs, pING4, appears to differ from pING1 by a simple rearrangement. The sequence of a much smaller plasmid, pING3, is identical to part of the pING4 sequence from which it appears to have looped out by a recombination process. Plasmid pING3 contains one ORF, 320, which does not find a homolog in pNOB8. It is incomplete in being unable to effect its own transfer but is subject to being transferred by the transfer apparatus of a complete CP from the same subfamily. It should thus contain a specific origin of replication allowing replicative transfer, *ori* T. Like the complete plasmids of this subfamily, pING3 is subject to curing. As curing should be facilitated by low copy number, its single ORF is a candidate for a copy number control gene as well (K. Stedman, I. Holz, Q. She, R. Garrett, and W. Zillig, manuscript in preparation).

According to DNA-DNA cross-hybridization and restriction fragment patterns, the CPs listed in Table 3 have been assigned to subfamilies. The members of three of these, the pNOB, the pING, and the pSOG subfamilies, are variants related to each other by recombination events. In pNOB8-33, this is a deletion between two perfectly repeated 85-bp sequences by homologous recombination. pING1, pING4, pING6, and the incomplete pING3 (see foregoing) are also recombinants of each other. In contrast, the genomes of the original pSOG2/4 and a cloned stable derivative formed or selected by repeated transfer of this plasmid into a foreign host, *S. solfataricus* strain P1, each contain nonhomologous in addition to shared sequences. The plasmids pARN3/2, pARN4/2, and pHVE14/5 clearly differ in restriction patterns but are close relatives according to Southern cross-hybridization (Prangishvili et al., in manuscript). The cloned plasmids are usually stable in con-

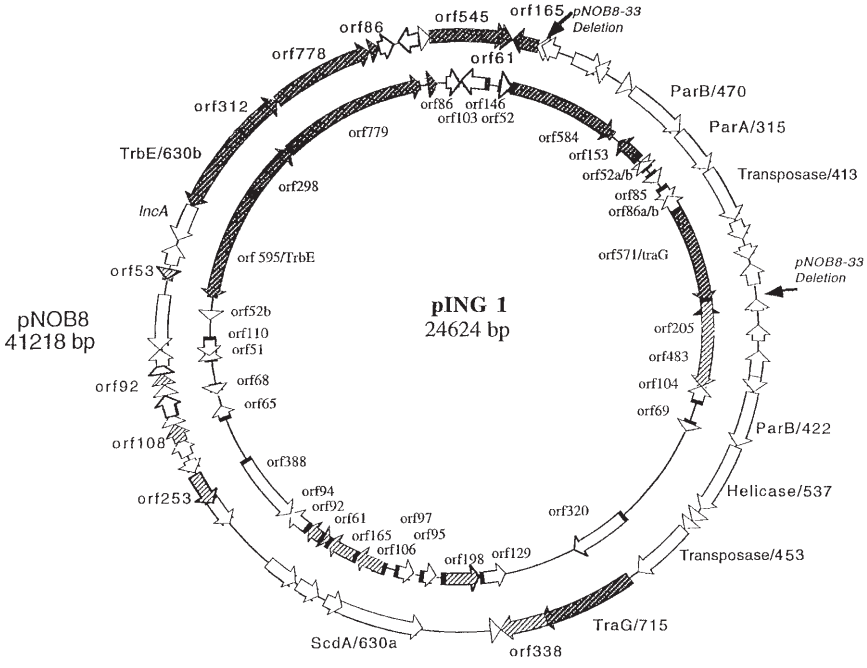

Fig. 5. Concentrically drawn maps of ORFs in the CPs pNOB8 and pING1 show the similarity of the organization of the conserved portions of the genomes and the presence of a portion in the pNOB8 genome, which is absent in pING1. *Dark shading*, more than 70% sequence similarity; *medium shading*, more than 50% similarity; *light shading*, more than 25% similarity.
Note added in proof: The open reading frames TraG/715 and ORF338 in pNOB8 as well as orf571/traG and ORF483 in pING1 are one continuous open reading frame in each plasmid

jugative transfer from donors to recipients of the same host strain but may form variants when the recipient is changed or when introduced by electroporation or during growth in transcipients. Plasmids of the four subfamilies effect incompatibility to superconjugation by other members of the same subfamily but are compatible with members of different subfamilies.

Utilization of virus and plasmid genomes for the construction of cloning and expression vectors

The first successful transformation of *S. solfataricus* utilized the CP pNOB8 as the vector, the *S. solfataricus lacS* gene (Cubellis et al. 1990) as the marker, and a β-galactosidase-negative *S. solfataricus* mutant, PH1, with an insertion sequence (IS) element disrupting the *lacS* gene (Schleper et al. 1994), as the recipient (Elferink et al. 1996). Shuttle vectors have been constructed from a small fragment of the genome of SSV1 containing the origin of replication (Cannio et al. 1996a), from the plasmid pGT5 of the euryarchaeote *Pyrococcus furiosus* (Aagard et al. 1996), and from the *Sulfolobus* plasmid pRN1 (R. Cannio, P. Contursi, M. Rossi, and S. Bartolucci, personal communication). The *Sulfolobus* ADH gene (Cannio et al. 1996b) and a thermoadapted hygromycin B phosphotransferase gene from *E. coli* were used to monitor transformation and to select for transformants.

Both viruses and CPs are self-spreading elements with near 100% spreading efficiency. Self-spreading shuttle vectors combining the genomes of these elements with those of *E. coli* vectors without interfering with the functions of both do not require markers to follow transformation after selective passage through *E. coli*. They are thus self-selecting. Additional advantages of SSV1-based vectors are the stable integration of the virus genome into the *Sulfolobus* chromosome and the possibility to control the copy number by UV induction. The major difficulty is to find locations in the genome that do not destroy virus functions. The first functional shuttle vector of this type has been constructed by statistically introducing single cuts into the virus genome with a multiply cutting restriction enzyme, followed by ligating an *E. coli* vector into the single cuts and selecting virus genomes carrying the insert by passage through *E. coli*, and finally isolating the working shuttle vector by passage through *Sulfolobus* (Figs. 6 and 7) (K. Stedman and W. Zillig, manuscript in preparation).

The analogous construction of a CP-based shuttle vector is hindered by the strong inhibitory effects of these multicopy plasmids on growth and colony-forming efficiency of transcipients. In its native host NOB8H1, however, stable maintenance of the complete CP pNOB8 is guaranteed by the Par functions of the plasmid (see above) although the copy number is low. Curing of NOB8H1 should therefore produce a useful recipient.

The lack of maintenance of CPs lacking Par functions is not necessarily a negative feature, because it opens the possibility to use these plasmids as "trojan horses" for the

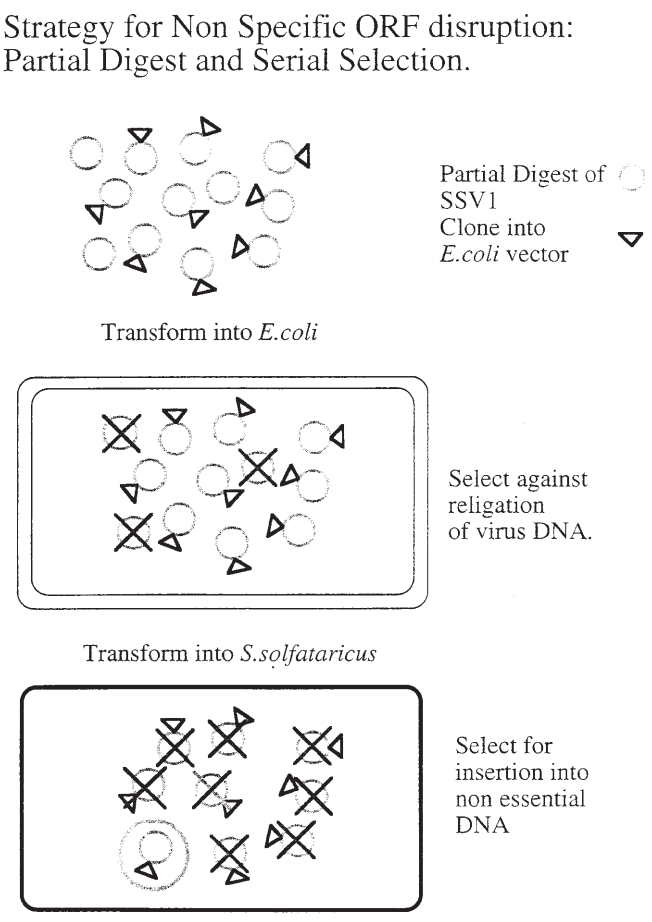

Fig. 6. Schematic presentation of the strategy of isolation of a self-selecting (self-spreading) shuttle vector containing a functioning SSV1 portion and a small *E. coli* vector portion

introduction of sequences that have the capacity to integrate into the *Sulfolobus* chromosome by homologous recombination, e.g., for specific inactivation ("knockout") or introduction of manipulated genes. The transcipients should lose the plasmids by curing but maintain the chromosomal alterations resulting from recombination with the introduced sequences. After conjugation of the *lacS* negative nonreverting *S. solfataricus* strain PH1 with pNOB8 carrying an intact *lacS* gene, we have observed the infrequent formation of large stable recombinants that carried an intact *lacS* gene in the chromosome and had been cured of the plasmid upon growth (Elferink et al. 1996).

Evolution

At first sight, the crenarchaeotal viruses appear unique, although fuselloviruses might be regarded as deformed head-and-tail phage. The organization of the genomes, the virus–host relationship including the lysogeny of *Sulfolobus* strains carrying SSV proviruses integrated into the host chromosomes, and other mainly negative controls, how-

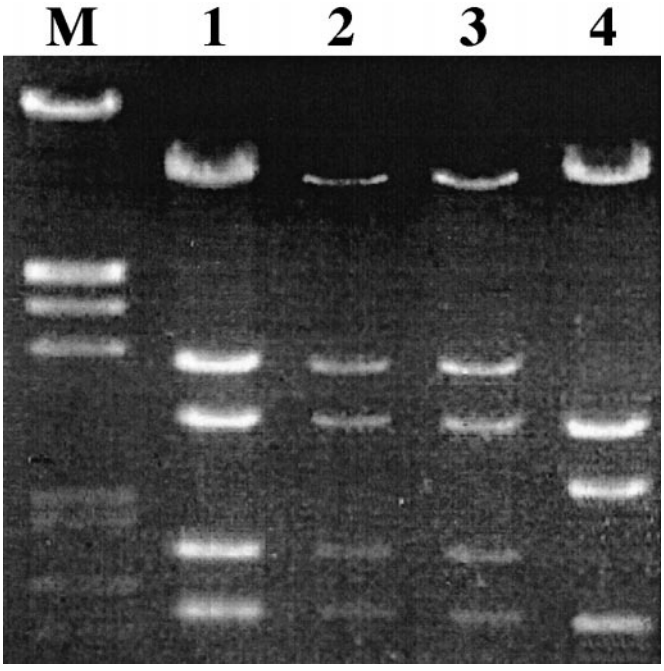

Fig. 7. A shuttle vector for *Sulfolobatus solfataricus* and *E. coli*. *Lane 1*: cccDNA isolated from a *S. solfataricus* selection. This DNA was electroporated into *E. coli*. *Lane 2*: cccDNA from a single clone from *E. coli* transformed with DNA from *Sulfolobus*. *Lane 3*: cccDNA from *S. solfataricus* transformed with DNA from a single *E. coli* clone. *Lane 4*: wild-type SSV1 DNA. *M*, size marker

ever, resemble corresponding features of lambdoid bacteriophages to an extent suggesting homology rather than convergence, although the fast evolution of viruses hinders the confirmation of this notion by sequence comparison. In view of the differences of the hosts, in lifestyle as well as in their incompatible transcription machineries, it is hard to imagine that these viruses spread from one domain into the other. They might rather have evolved in concert with their hosts from an ancestral virus of the common ancestor of the three domains. The DNA termini of rudiviruses resemble the DNA termini of some eukaryal viruses containing linear DNAs (Rohozinski et al. 1989). Archaeal viruses show modular organization resembling that in bacteriophages (Campbell and Botstein 1983; Campbell 1988). This is apparent in SSV1 where about half the genome harboring two inducible transcription units in opposite orientation contains DNA genes (in bacteriophages termed early genes) in which cysteine codons occur in normal frequency, whereas the other half including the genes for the structural proteins (in bacteriophages termed late genes) contains only one cysteine codon. Cysteine codons are not of the RNY type considered primeval (Shepherd 1981). Their absence in one of the modules might thus indicate its primeval nature. As discussed, the other module, which because of its cysteine content should even be younger, perhaps already existed in the common ancestral period of biotic (organismic) evolution. If so, viruses might even date back to prebiotic evolution. The definition of viruses as entities requiring hosts for propagation seems to exclude this possibility except when we assume that the primordial soup or surface was the primordial host.

Are plasmids, especially the apparently selfish CPs, also that old? The significant although weak sequence similarity between archaeal and bacterial CPs in two related ORFs encoding NTPases that are possibly involved in DNA transfer indicates that a primitive conjugation system might already have existed in the common ancestral state. Was it involved in the efficient gene transfer supposed to have linked individuals of common ancestral populations?

References

Aagaard C, Leviev I, Aravalli RN, Forterre P, Prieur D, Garrett RA (1996) General vectors for archaeal hyperthermophiles: strategies based on a mobile intron and a plasmid. FEMS Microbiol Rev 18:93–104

Arnold HP (1998) Isolierung und Beschreibung neuer Viren der crenarchaealen Gattung *Sulfolobus*. Dissertation der Biologischen Fakultät für Biologie der Ludwig-Maximilans-Universität, München

Birnboim HC, Doly J (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res 7:1513–1523

Campbell A (1988) Phage evolution and specification. In: Calendar R (ed) The bacteriophages, vol 1. Plenum, New York, pp 1–14

Campbell A, Botstein D (1983) Evolution of the lambdoid phages. In: Hendrix RW, et al (eds) Lambda II. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, pp 365–380

Cannio R, Contursi P, Rossi M, Bartolucci S (1996a) An autonomously replicating, transforming vector for *Sulfolobus solfataricus*: expression in *Sulfolobus* cells of a mesophilic drug resistance adapted to the higher temperature. Abstracts of the first international congress on extremophiles, Estoril, Portugal, p 35

Cannio R, Fiorentino G, Carpinelli P, Rossi M, Bartolucci S (1996b) Cloning and overexpression in *Escherichia coli* of the genes encoding NAD-dependent alcohol dehydrogenase from two *Sulfolobus* species. J Bacteriol 178:301–305

Christie GE, Calendar R (1990) Interactions between satellite bacteriophage P4 and its helpers. Annu Rev Genet 24:465–490

Cubellis MV, Rozzo C, Montecucchi P, Rossi M (1990) Isolation and sequencing of a new β-galactosidase-encoding archaebacterial gene. Gene 94:89–94

Elferink M, Schleper CH, Zillig W (1996) Transformation of the extremely thermophilic archaeon *Sulfolobus solfataricus* via self-spreading vector. FEMS Microbiol Lett 137:31–35

Götz D (1994) Charakterisierung des *Sulfolobus islandicus* – virus SIRV. Diplomarbeit der Fakultät für Biologie der Ludwig-Maximilians-Universität, München

Grogan D, Palm P, Zillig W (1990) Isolate B12, which harbours a virus-like element, represents a new species of the archaebacterial genus *Sulfolobus, Sulfolobus shibatae*, sp. nov. Arch Microbiol 154:594–599

Janekovic D, Wunderl S, Holz I, Zillig W, Gierl A, Neumann H (1983) TTV1, TTV2 and TTV3, a family of viruses of the extremely thermophilic, anaerobic, sulfur reducing archaebacterium *Thermoproteus tenax.* Mol Gen Genet 192:39–45

Keeling PJ, Klenk H-P, Singh RK, Feeley O, Schleper CH, Zillig W, Doolittle WF, Sensen CW (1996) Complete nucleotide sequence of the *Sulfolobus islandicus* multicopy plasmid pRN1. Plasmid 35:141–144

Martin A, Yeats S, Janekovic D, Reiter W-D, Aicher W, Zillig W (1984) SAV1, a temperature u.v.-inducible DNA virus-like particle from the archaebacterium *Sulfolobus acidocaldarius.* EMBO J 3(9):2165–2168

Neumann H, Schwass V, Eckerskorn C, Zillig W (1989) Identification and characterization of the genes encoding the structural proteins of the *Thermoproteus tenax* virus TTVI. Mol Gen Genet 217:105–110

Palm P, Schleper C, Grampp BW, Yeats S, McWilliam P, Reiter W-D, Zillig W (1991) Complete nucleotide sequence of the virus SSV1 of the archaebacterium *Sulfolobus shibatae*. Virology 185:242–250

Reiter W-D, Palm P (1990) Identification and characterization of a defective SSV1 genome integrated into a tRNA gene in the archaebacterium *Sulfolobus* sp. B12. Mol Gen Genet 221:65–71

Reiter W-D, Palm P, Henschen A, Lottspeich F, Zillig W, Grampp B (1987a) Identification and characterization of the genes encoding three structural proteins of the *Sulfolobus* virus-like particle SSV1. Mol Gen Genet 206:144–153

Reiter W-D, Palm P, Yeats S, Zillig W (1987b) Gene expression in Archaebacteria: physical mapping of constitutive and UV-inducible transcripts from the *Sulfolobus* virus-like particle SSV1. Mol Gen Genet 209:270–275

Reiter W-D, Palm P, Zillig W (1988a) Analysis of transcription in the archaebacterium *Sulfolobus* indicates that archaebacterial promoters are homologous to eukaryotic pol II promoters. Nucleic Acids Res 16:1–19

Reiter W-D, Zillig W, Palm P (1988b) Archaebacterial viruses. In: Marmarosch K, Ruply FA, Shatkin AJ (eds) Advances in virus research, vol 34. Academic Press, Orlando, pp 143–188

Reiter W-D, Palm P, Yeats S (1989) Transfer RNA genes frequently serve as integration sites for prokaryotic genetic elements. Nucleic Acids Res 17(5):1907–1914

Rettenberger M (1990) Das Virus TTV1 des extrem thermophilen Schwefel-Archaebakteriums *Thermoproteus tenax*: Zusammensetzung und Struktur. Dissertation der Fakultät für Biologie der Ludwig-Maximilians-Universität, München

Rohozinski J, Girton LE, van Ettern JL (1989) *Chlorella* viruses contain linear nonpermuted double-stranded DNA genomes with covalently closed hairpin ends. Virology 168:363–369

Schleper C, Röger R, Singer T, Zillig W (1994) An insertion element of the extremely thermophilic archaeon *Sulfolobus solfataricus* transposes into the endogenous β-galactosidase gene. Mol Gen Genet 243:91–96

Schleper C, Holz I, Janekovic D, Murphy J, Zillig W (1995) A multicopy plasmid of the extremely thermophilic archaeon *Sulfolobus* effects its transfer to recipients by mating. J Bacteriol 177:4417–4425

She Q, Phan H, Garrett RA, Albers S-V, Stedman KM, Zillig W (1998) Genetic profile of pNOB8 from *Sulfolobus*: the first conjugative plasmid from an archaeon. Extremophiles 2(4)

Shepherd JCW (1981) Periodic correlations in DNA sequences and evidence suggesting their evolutionary origin in a comma-less code. J Mol Evol 17:94–102

Singer T (1993) Kontrolle der Synthese des UV-induzierten Transkripts T_{ind} des Virus′ SSV1. Diplomarbeit der Fakultät für Biologie der Ludwig-Maximiliand-Universität, München

Zillig W, Stetter KO, Wunderl S, Schulz W, Priess H, Scholz I (1980) The *Sulfolobus*-"Caldariella" group: taxonomy on the basis of the structure of DNA-dependent RNA polymerases. Arch Microbiol 125:259–269

Zillig W, Reiter W-D, Palm P, Gropp F, Neumann H, Rettenberger M (1988) Viruses of Archaebacteria. In: Calendar R (ed) The Bacteriophages, vol 1. Plenum, New York, pp 517–568

Zillig W, Kletzin A, Schleper CH, Holz I, Janekovic D, Hain J, Lanzendörfer M, Kristjanson JK (1994) Screening for *Sulfolobales*, their plasmids and their viruses in Icelandic solfataras. Syst Appl Microbiol 16:609–628

Zillig W, Prangishvili D, Schleper CH, Elferink M, Holz I, Albers S, Janekovic D, Götz D (1996) Viruses, plasmids and other genetic elements of thermophilic and hyperthermophilic *Archaea*. FEMS Microbiol Rev 18:225–236