MetSV, a novel archaeal lytic virus targeting *Methanosarcina* strains

Katrin Weidenbach[1], Lisa Nickel[1], Horst Neve[2], Omer S. Alkhnbashi[3], Sven Künzel[4], Anne Kupczok[1], Thorsten Bauersachs[5], Liam Cassidy[6], Andreas Tholey[6], Rolf Backofen[3, 7] Ruth A. Schmitz[1]*

[1]Christian-Albrechts-University, Institute for General Microbiology, Am Botanischen Garten 1-9, 24118 Kiel, Germany

[2]Max Rubner-Institute, Department of Microbiology and Biotechnology, Hermann-Weigmann-Str. 1, 24103 Kiel, Germany

[3]Bioinformatics Group, Department of Computer Science, University of Freiburg, Georges-Köhler-Allee 106, 79110, Freiburg, Germany

[4]MPI for Evolutionary Biology, August-Thienemann-Str. 2, 24306 Plön, Germany

[5]Christian-Albrechts-University, Institute of Geosciences, Department of Organic Geochemistry, Ludewig-Meyn-Straße 10, 24118 Kiel, Germany

[6]Christian-Albrechts-University, Institute for Experimental Medicine; Department of Systematic Proteomics and Bioanalytic, Niemannsweg 11, 24105 Kiel, Germany

[7]Centre for Biological Signalling Studies (BIOSS), Cluster of Excellence, University of Freiburg, Germany

Running title: MetSV, a novel *Methanosarcina* virus

Key words: Methanosarcina mazei strain Gö1, virus, methanoarchaea, CRISPR/Cas

*Corresponding author: Ruth A. Schmitz, Institut für Allgemeine Mikrobiologie, Universität Kiel, Am Botanischen Garten 1-9, 24118 Kiel, Germany

Phone: + 49 (0431) 804334

Fax: + 49 (0431) 8802194.

Electronic mail address: rschmitz@ifam.uni-kiel.de

Abstract

A novel archaeal lytic virus targeting species of the genus *Methanosarcina* was isolated using *Methanosarcina mazei* strain Gö1 as host. Due to its spherical morphology the virus was designated *Methanosarcina* spherical virus (MetSV). Molecular analysis demonstrated that MetSV contains double stranded linear DNA with a genome size of 10,567 bp containing 22 open reading frames (ORFs) all oriented in the same direction. Functions were predicted for some of these ORFs, i. e. like DNA polymerase, ATPase, DNA-binding protein, as well as envelope (structural) protein. MetSV-derived spacers in CRISPR loci were detected in several published *Methanosarcina* draft genomes using bioinformatic tools, revealing the potential PAM motif (TTA/T). Transcription and expression of several predicted viral ORFs were validated by RT-PCR, PAGE analysis and LC-MS based proteomics. Analysis of core lipids by APCI mass spectrometry showed that MetSV and *M. mazei* both contain archaeol and glycerol dialkyl glycerol tetraether without cyclopentane moiety (GDGT-0). The MetSV host range is limited to *Methanosarcina* strains growing as single cells (*M. mazei*, *M. bakeri* and *M. soligelidi*). In contrast, strains growing as sarcina-like aggregates were apparently protected from infection. Heterogeneity related to morphology phases in *M. mazei* cultures allowed acquisition of resistance to MetSV after challenge by growing as sarcina-like aggregates. CRISPR/Cas mediated resistance was excluded since neither of the two CRISPR arrays showed MetSV-derived spacer acquisition. Based on these findings, we propose that changing the morphology from single cells to sarcina-like aggregates upon rearrangement of the envelope structure prevents infection and subsequent lysis by MetSV.

Importance

Methanoarchaea are among the most abundant organisms on the planet since they are present in high numbers in major anaerobic environments. They convert various carbon sources e.g. acetate, methylamines or methanol to methane and carbon dioxide, thus they have a

significant impact on the emission of major greenhouse gases. Today very little is known about viruses specifically infecting methanoarchaea, which most probably impact the abundance of methanoarchaea in microbial consortia. Here we characterize the first identified *Methanosarcina* infecting virus (MetSV) and show a mechanism for acquiring resistance against MetSV. Based on our results we propose that growth as sarcina-like aggregates prevents infection and subsequent lysis. These findings allow new insights into virus-host relationship in methanogenic community structures, their dynamics and their phase heterogeneity. Moreover, the availability of a specific virus provides new possibilities to deepen our knowledge on defence mechanisms of potential hosts and offers tools for genetic manipulation.

Introduction

Since the first archaeal viruses have been described in the 1970s, more than 100 have been reported (recently reviewed in (1)). Besides viruses infecting crenarchaea or halophilic archaea only very few viruses, proviruses and virus-like particles infecting methanoarchaea have been discovered today (reviewed in (2, 3)). Two of them, ψM1 and ψM2, have been described infecting *Methanothermobacter marburgensis* (formerly known as *Methanobacterium thermoautotrophicum).* Both viruses consist of a polyhedral head and a tail (4-6) and contain linear dsDNA with a genome size of approximately 27 kb and 26.3 kb, respectively. Thereby ψM2 is a deletion derivative of ψM1 that originated by homologous recombination under laboratory conditions. Furthermore, a defective provirus derivative with an approximately 70 % identity to the ψM2 genome has been identified in the genome of *Methanothermobacter wolfii* (7). Two morphologically similar viruses, i.e. ΦF1 and ΦF3, have been isolated from an anaerobic digester. Those viruses also consist of a polyhedral head

and a tail, but they differ in genome size, genome organisation and host specificity. Virus ΦF1 comprises a linear dsDNA with a size of approximately 85 +/- 5 kb and a broad host range. Virus ΦF3 contains dsDNA organized in a circular or terminally redundant linear genome with a size of approximately 36 +/- 2 kb and is only able to infect *M. thermoformicicum*. Sequence analysis revealed no genome similarities between ΦF1 and ΦF3 (8).

A few virus-like-particles (VLPs) have been reported which infect methanoarchaea. The first VLPs associated with *Methanococcus voltae* A3 were described by Wood *et al*. They detected extrachromosomal closed circular DNA and particles which appeared like viruses. Interestingly, an integrated chromosomal copy was verified, however neither infectivity nor inducibility of the VLPs could be demonstrated (9). Sequence analyses revealed an integrated copy of the A3 VLP in the genome of *M. voltae* A3 inside a provirus (MVV). It appears that the A3 VLP was integrated into the pre-integrated proviral genome, which might have led to the inactivation of the provirus (10). When growing *Methanofollis aquaemaris* sp. nov. in late exponential phase, VLP release by budding has been demonstrated, indicating the presence of a provirus, which was named MMV (11). Further VLPs have been detected, which are able to infect the archaeal genus *Methanosaeta* in the deep and anoxic sediments of the volcanic Lake Pavin (France) (12) and in *Methanosaeta* dominated methanogenic reactors (13). In addition, several proviruses were identified by genome sequence analyses in strains of *Methanobrevibacter ruminantium*; *Methanobrevibacter smithii* and *Methanococcus jannaschii* (14-17).

Here, we report the isolation and characterization of a new lytic archaeal virus infecting *Methanosarcina* strains, which now allows new insights into resistance formation in *Methanosarcina*. Besides it offers enormous potential for analysing the function of

CRISPR/Cas systems in *Methanosarcina* strains and the development of additional genetic tools.

Results

Isolation and morphological characterisation of virus MetSV

The MetSV was enriched and isolated from anaerobic sewage sludge collected in the digestion tower of a municipal the waste water treatment plant (Göttingen, Gemany), by serially challenging *M. mazei* DSM 3647 growing in liquid cultures with solutions of filtrated sludge. Transmission electron microscopy of negatively stained virus preparations revealed a spherical shape with a diameter of 56.6 ± 1.9 nm and a blackberry like envelope of the virus (Fig. 1). Disrupted virus particles (i.e., ghost particles without DNA) were typically filled with the dark staining solution. These damaged particles exhibited a unique thick structure of the coat envelope. The MetSV virus exhibits an overall morphotype similar to the previously reported spherical virus PSV infecting *Pyrobaculum* or *Thermoproteus* (18) however has a significantly smaller particle diameter than PSV (with approx. 100 nm diameter) and no helical nucleoparticles were detected. Due to the overall morphological appearence of the virus it was designated as Methanosarcina Spherical Virus (MetSV).

The linear genome of MetSV

DNA from MetSV was isolated as described in section Material and methods. Subsequent treatment with different restriction endonucleases demonstrated a linear, double stranded (ds) DNA with an approximate length of 11 kb (data not shown). The DNA was sequenced using a high throughput shot gun sequencing approach (using the 454-sequencing technology) and in

parallel by sub-cloning followed by Sanger sequencing. The size of the MetSV genome was determined to be 10,567 bp, which was in agreement with the endonuclease restriction analyses. Both ends of the linear genome were flanked by 59 nt long inverted terminal repeats (ITR). The G+C content was determined to be 38.46 %, which is in a similar range of the G+C content of the host DNA (41.5 %) (19). The MetSV genome contains 22 ORFs encoding potential proteins ranging from 3,630 Da to 36,700 Da. Thereof 11 long ORFs encoding proteins with a predicted size of approximately > 90 amino acids (aa) (Fig. 2) and 11 small ORFs encoding proteins with a size ≤ 90 aa. Promotor elements typical for *Methanosarcina* (TATA and BRE boxes) were identified for several ORFs (Table 1). All predicted ORFs were located on the same DNA strand resulting in a consistent direction of transcription (Fig. 2). From the 22 ORFs, 15 showed very low or no similarities to sequences in public databases, while only four ORFs showed homologies to proteins with an assigned function. ORF7 showed homology to a DNA polymerase B (37 % local identity (67 aa/179 aa)) and ORF13 to an ATPase (30 % local identity (64 aa/213 aa)) on the amino acid level. The latter contained a P-loop NTPase (P-loop containing nucleoside triphosphate hydrolase) domain and a HerA domain, as predicted using NCBI's conserved domains (20). Those ATPases are known to be involved in viral genome packing and segregation (21). For ORF6, a ribbon-helix-helix-domain was predicted using Phyre2 (22), which is often a part of DNA binding proteins like transcriptional regulators. ORF22 contains an Env-gp36 domain, which is a part of an envelope forming protein in the Lentivirus Feline Immunodeficiency Virus (FIV) (23), which potential suggest that ORF22 represents the structural coating protein. 3 ORFs showed similarities to hypothetical proteins (ORF3; ORF9 and ORF16). Furthermore second structure predictions were performed using the Phyre2 software. Thereby 7 ORFs with higher amounts of beta sheets were detected, however no jelly roll structures were found, which are characteristic for major capsid proteins of icosahedral viruses (see Table 1).

155

156 *Screening for MetSV-derived spacers in methanosarcinal CRISPR arrays*

157 *M. mazei* type strain DSM 3647 contains two CRISPR/Cas-systems, which belong to the
158 subtypes IB and IIIC according to the nomenclature from Makarova *et al.* and Vestergaard *et*
159 *al.* (24-26). Both CRISPR arrays were screened for MetSV-derived spacers. No matches were
160 detected in the type strain (*M. mazei* DSM 3647). Draft genomes of additional
161 *Methanosarcina* isolates recently reported by Youngblut *et al*. (27) and available genomes
162 from the NCBI database were also screened for specific MetSV derived spacers. No spacers
163 without a mismatch were identified. Only spacers with one and up to six mismatches were
164 identified in *Methanosarcina* genomes (see Table S1). Responsible for the selection of certain
165 protospacers on the viral genome is the protospacer adjacend motif (PAM), which is defined
166 by the respective CRISPR/Cas system. By analysing the protospacers in MetSV a potential
167 PAM motif TTA/T was predicted.

168

169 *Expression of viral genes during infection: RNA and protein patterns*

170 RNA was isolated at different time points after infection of *M. mazei* cultures with the
171 MetSV. Transcription of predicted MetSV genes ORF6, ORF7, ORF13, ORF15, ORF19 and
172 ORF22 was evaluated by RT-PCR using specific primer pairs (Table S2). Generally, 30 min
173 after MetSV infection transcripts of all tested MetSV ORFs were detected. Transcripts were
174 still detectable after 90 min and 210 min past infection as shown in Fig. 3.

175 Furthermore, we performed a stack gel separation coupled to LC-MS/MS analysis (28) of the
176 purified MetSV proteome. Out of two samples, we identified twelve predicted ORF gene
177 products with, on average, sequence coverage greater than 50% (Supplemental Table S3 and
178 S4). In addition to the twelve confidently identified protein species, single proteolytic

peptides from three further virus proteins were found (data not shown). Additionally, two prominent protein bands with a size of approx. 11 kDa and 16 kDa were separately analysed showing that the major components were gene products of ORF16 (16 kDa band) and ORF 15 (11 kDa band) (Table S5 and S6).

Core lipid analysis

M. mazei DSM 3647 and the purified virus fraction were analysed by APCI mass spectrometry for their di- and tetraether content. We found that both samples contained the same core lipids in similar distributions but in significantly different quantities. Archaeol was the dominant membrane lipid that constituted 1318.85 and 40.81 ng/g dry cell material in the archaeal host and MetSV, respectively (see Table 6 and Fig. 5). In addition traces of glycerol dialkyl glycerol tetraether without cyclopentane moiety (GDGT-0) were detected in *M. mazei* DSM 3647 (1.76 ng/g dry cell material) and MetSV (0.07 ng/g dry cell material). These results indicate that MetSV contains an internal membrane and incorporates lipids obtained from the host. To exclude the possibility of contamination of the virus fraction by archaeal lipids during sample processing, a control experiment with filtered supernatant of a mature *M. mazei* culture was performed, in which no core lipids were detected.

Host range of MetSV

Different *Methanosarcina* strains available at the DSMZ were tested for sensitivity to MetSV. Actively growing cultures were incubated with MetSV and the optical turbidity was monitored at 600 nm. Most of the tested strains growing as single cells were lysed by MetSV within 4 h (*M. mazei* strains DSM 3647/DSM 2244/DSM 4556/DSM 6300/DSM 7222 and M .*bakeri* strain DSM 1311). The only exception was *M. acetivorans* DSM 2834TC2A growing

as single cells, which however was not lysed by MetSV (summarized in Table 3). In contrast, all strains typically growing as sarcina-like aggregates were not affected by MetSV (strains: *M. mazei* DSM 9195*; M. bakeri* DSM 800; *M. soligelidi* DSM 26065; *Methanosarcina spp*. DSM 11855 and *M. siliciliae* DSM 3028). Control strains that included other methanogenic archaea such as *Methanosphaera stadtmanaea, Methanosaeta thermophila* or the gram negative bacterium *E. coli* were not sensitive to MetSV (data not shown), strongly indicating a narrow host range of MetSV.

Challenging M. mazei *with MetSV and establishment of viral resistance*

M. mazei DSM 3647 was grown to a turbidity of approx. 0.2 – 0.3 at 600 nm, then MetSV containing lysate was added. Growth immediately stopped and cell lysis was observed after approximately 4 h of challenging and approximately 12 h later the majority of the cells was lysed (Fig. 6A). However, after several weeks of incubation at 37 °C the infected cultures started to grow again and cells tended to form large aggregates. These cells grew as typical sarcina-like aggregates and were not anymore affected by adding MetSV at various growth stages. Representative growth analyses are shown in Fig. 6. Chromosomal DNA was isolated from four independent aggregated cultures and confirmed to represent *M. mazei* chromosomal DNA. To exclude an involvement of the CRISPR/Cas system during resistance formation, the CRISPR arrays were examined for newly acquired spacers against MetSV. Therefore the first 29 spacers/repeats of the CRISPR IB array and the first nine spacers/repeats of the CRISPR IIIC array were PCR-amplified and analysed by sequencing, demonstrating that a MetSV-derived spacer was neither integrated in array IB nor in IIIC. However, in several cases a rearrangement of spacers was observed and spacers from the 3´end of the array were moved to the front or directly after the leader as depicted in Fig. 6C.

227

228 *Effects on global transcription of the resistant* M. mazei *strain in the presence of MetSV*

229 Transcriptome analysis were performed using whole genomic DNA microarrays and cDNAs
230 generated from total RNAs, which were isolated from uninfected cells and cells growing as
231 aggregates 200 h past infection (p.i.) with MetSV. In three independent biological analyses,
232 transcripts of 2334 ORFs were detectable; thereof 219 ORFs were identified with 3-fold
233 higher and 107 ORFs with 3-fold lower transcript abundance in the aggregated cells 200h p.i.
234 with MetSV compared to the non-infected cultures (Table 4 and Table S7). Several ORFs,
235 which have homologies to proteins involved in envelope formation and 35 ORFs annotated as
236 transporters had significantly elevated transcript levels (>3–fold), for example, a potential
237 operon *MM_3095-MM_3098* encoding two S-layer proteins and two ABC-transporter genes
238 (>10-20-fold). 107 ORFs had lower transcription levels in the aggregated cells 200 h p.i.,
239 compared to uninfected cells. 33 of these ORFs are coding for ribosomal proteins indicating
240 general slowdown of growth. Transcription levels of several ORFs were verified by qRT-PCR
241 as listed in Table 6. For the ORF *MM_0490* with the highest change in the transcript amounts
242 (annotated as phycocyanin alpha-subunit phycocyanobilin lyase) and *MM_3096* (encoding a
243 putative S-layer protein) the significantly increased transcript levels were verified by qRT-
244 PCR (442.74 +/- 43.46 and 10.18 +/- 1.14).

245

246 **Discussion**

247 Since the first discovery of methanoarchaea viruses in 1989, Methanosarcina Spherical Virus
248 (MetSV) is the only virus which was isolated and characterized for methanoarchaea for some
249 time. Besides the spherical virus morphology, the internal membrane and the presence of a
250 potential ATPase, which might be involved in genome packaging into procapsids, we did not

observe a protein which was similar to the known major capsid proteins, therefore we could not classify MetSV in one of the known virus families. In this respect it is tempting to speculate that MetSV belongs to a new virus family. In contrast, all other isolated methanoarchaea infecting viruses possess heads and tails and belong to the families of either *Myo*-or *Siphoviruses* (reviewed in (3)).

Genome organisation and viral replication

The genome of MetSV is double stranded, linear and has a size of 10,567 bp. This is in line with all other known archaeal viruses, all containing DNA (reviewed in (3)) and showing genome sizes from 5 – 144 kbp (reviewed in (29)). There are inverted terminal repeats (ITR) with a length of 59 nt on both ends of the MetSV genome. Quite a number of linear archaeal viruses contain ITRs including the *Haloarcula hispania* infecting viruses SH1 and HHIV-2, *Sulfolobus* virus SIRV1, the haloviruses His1 and His2 and the *Pyrobaculum* infecting PFV1 (2, 30-34). The presence of ITRs strongly suggests that MetSV also replicates via protein-primed DNA-replication like the bacteriophages PRD1 and Ø29 or the archaeoviruses His1, His2 and SH1, respectively (30, 33, 35, 36). A further indication for a protein-primed DNA replication is the predicted type B - polymerase, encoded at one end of the MetSV genome. Type B - polymerases are able to use proteins attached to the 5´ends of linear dsDNA to prime DNA-replication (reviewed in (29)). Functions were assigned for two other ORFs. First, ORF13 encoding an ATPase which contains a HerA (HAS-barrel/ATPase domain) and a P-loop NTPase domain, indicating that this ATPase is involved in DNA packing into the capsid, it has been shown for phage PDR1 ATPase (reviewed in (3)). Second ORF22 gene product contains an Env-gp36 domain, which is a part of a lentivirus surface glycoprotein (feline immunodeficiency retrovirus), thus we speculate that the ORF22 gene product might be a part of the MetSV envelope. In addition we predicted for several ORFs both archaeal promoter elements, the TATA- and BRE-box, which show high similarity to the *M. mazei*

promoter. This indicate that MetSV is using the transcription machinery of the host, as it has been reported for e.g. SH1 (37). In the experimental set up, however, it was not ascertainable that MetSV ORFs show different transcription patterns at varying time points after infection. Thus we could not distinguish between early and late viral genes as reported for other viruses (38), (39-42). Besides the 11 large ORFs, 11 small ORFs (sORFs) (≤ 90 aa) were predicted (Table 1). These sORFs encode small proteins ranging from 3.6 to 10 kDa. Some of them (i.e. ORF 11 and ORF 17) contain up to 43 % hydrophobic aa. Small proteins are described for viruses infecting eukaryotic cells like Influenza virus, Human Immunodeficiency virus (HIV), Fibropapillomavirus, poxvirus or paramyxovirus. Based on the high ratio of hydrophobic aa, most of the small proteins have transmembrane domains and are able to form so called viroporins or they function as regulators for expression or activity of larger cellular transmembrane proteins by binding to their membrane-spanning domains (reviewed in (43)). The identified MetSV sORFs might also be involved in envelope formation or modulating other large protein functions. The shortest ribosomally produced peptides from a phage genome have been recently reported to be involved in decision between lysogenic and lytic cycle of tempertate phages by functioning as a “viral quorum sensing” (44).

Host range

Generally the host range of methanogenic viruses can vary from rather narrow (e.g., ΦF3, which only infects *M. thermoautotrophicum)* to broad host range (e.g., virus ΦF1 infecting several different *M. thermoautotrophicum* strains) (8). MetSV was able to infect a variety of different *Methanosarcina* strains (see Table 2), but neither strains of *Methanosphaera stadtmanae* nor *Methanosaeta thermophile*. Consequently, MetSV represents the first virus specifically infecting the genus *Methanosarcina*. This is in line with the identification of

different MetSV derived spacers with one and up to six mismatches in their sequence exclusively in *Methanosarcina* (draft) genomes. Functional spacers can include mismatches unless the seed region is not mutated (45). Challenge assays clearly showed that the lytic virus MetSV can only infect *Methanosarcina* strains growing as single cells, and not as sarcina-like aggregates. Interestingly, *M. acetivorans*; that also grows as single cells under laboratory conditions was not infected by MetSV. The reason for this phenomenon might be that MetSV is sensitive to high salt concentrations since *M. acetivorans* was grown in high salt media. Additional evidence for this is provided by the observation that MetSV was also not infectious for *M. mazei* DSM 3647 when the strain was grown in a medium supplemented with 500 mM sodium chloride (data not shown). Further spacer analyses in different *Methanosarcina* draft genomes shows that MetSV is a ubiquitously distributed virus and strains with MetSV derived spacers might have an active CRISPR/Cas system in contrast to *M. mazei* DSM 3647 (46).

Resistance formation

M. mazei resistance against MetSV infection was detected after a change of morphology phenotype from single cells to aggregates. It is known that in some cases bacteria change their surface layer proteins in a reversible manner by different expressions patterns resulting in a phenotypically heterogenic population. This phase variation increases the overall fitness and reduces the sensitivity to phage infection (reviewed by (47)). One example is the different expression in the phases Bvg^+ or Bvg^- of BvgAS, a two component system from *Bordetella*, which is involved in the regulation of the expression of surface proteins. In Bvg^+ phase cells (BvgAS is active), pertactin is expressed, which is the recognition site for phage BPP-1. Consequently in Bvg^- phase cells (BvgAS is inactive), the pertactin expression is supressed resulting in an increased protection against BPP-1 infection (48-50). Transcriptome analysis

of MetSV induced aggregated *M. mazei* cells versus non challenged cells revealed several differently transcribed genes that are potentially involved in altering the cell surface structure (see Table 3). Those ORFs encode S-layer proteins or ABC-transporters, which are known to be involved in segregation of S-layer proteins (reviewed in (51)). The ORF with highest transcript levels in the aggregated cells *MM_0490*, annotated as phycocyanin alpha-subunit phycocyanobilin lyase, is almost conserved in the Methanosarcinales. Using Phyre2 prediction a possible function in cell adhesion was predicted. Furthermore, Saunders *et al.* (52) reported a PBS-lyase heat repeat domain containing protein from *M. bakeri*, which had homologies to secreted proteins from *M. burtonii* involved in envelope formation. Considering those findings and the highly elevated transcript levels of *MM_0490* in the aggregated cells after MetSV infection, it is tempting to speculate, that this protein is crucially involved in envelope formation and the formation of sarcina-like aggregations. Consequently, we propose resistance against MetSV is mediated by phase variation of the morphology phenotype from single cells to aggregates caused by differences in envelope structures through different expression of surface proteins and not by CRISPR-derived immunity.

Material and methods

Strains and plasmids

Strains and plasmids used in this study are listed in Table 6.

Growth conditions

All archaeal strains were grown under strictly anaerobic conditions. The growth media were generally supplemented with 100 µg/ml ampicillin to prevent bacterial contamination. *M. mazei* isolates (DSM 3647, 2244, 4556, 6300, 9195), *M. soligelidi* (DSM 26065) and *M. bakeri* (DSM 800 and 1311) strains were grown in minimal medium under a N_2/CO_2 (80:20)

350 atmosphere at 37 °C without shaking, as described previously (53, 54). The medium was
351 supplemented with 150 mM methanol (MeOH), 40 mM acetate, 2 mM cysteine and 1 mM
352 sodium sulfide. In case of *Methanosarcina spp*. (DSM 11855), minimal medium was
353 additionally supplemented with 5 % (v/v) rumen fluid (clarified) and the strain was incubated
354 at 55 °C. *M. acetivorans* (DSM 2834) was grown in high salt medium as described by Sowers
355 and Schreier (55). *M. siliciliae* (DSM 3028) was grown in DSMZ medium no. 324
356 (*Methanolobus* II medium. Growth was monitored by determining the turbidity of the cultures
357 at 600 nm ($O.D._{600}$).
358
359
360 *Virus enrichment and isolation*

361 Digester sludge was filtered (0.2 µm, polyethersulfon membrane, Sarstedt, Nümbrecht,
362 Germany) under anaerobic conditions and the filtrate was added to 5 ml of an exponentially
363 growing *M. mazei* DSM 3647 cultures. After cell-lysis of the culture, the virus containing
364 supernatant was again anaerobically filtrated and 10 % virus filtrate (final concentration) was
365 added to a fresh *M. mazei* DSM 3647 culture. After 10 passage,s only one type of virus was
366 detectable by transmission electron microscopy.

367 *Transmission electron microscopy*

368 One liter of *M. mazei* culture was grown to a turbidity of ~ 0.2 at 600 nm and supplemented
369 with MetSV containing filtrate to 1 % final concentration and further incubated at 37 °C.
370 After cell lysis (24 h), viruses were harvested using a modified protocol as described by
371 Colombet *et al.* (56). Briefly, after adding 10 % PEG 8000 and 0.6 % NaCl (Roth, Karlsruhe;
372 Germany), the cultures were incubated for approximately 24 h in the dark at 4 °C. The white
373 layer, which contains the viruses was transferred to a sterile centrifugation tube and harvested
374 by centrifugation at 8,000 x g for 20 min at 4 °C. The pellet was re-suspended in 300 µl SM-

375 buffer (0.1 M NaCl; 8 mM $MgSO_4x7H_2O$; 50 mM Tris-HCl; 0.005 % glycerol (wt/vol); pH
376 7.0). Then, 1 M KCl (final concentration) was added and the mixture was incubated for
377 20 min on ice. The mixture was centrifuged at 12,000 x g for 10 min at 4 °C. The viruses
378 containing supernatant was transferred to a new tube and stored at -80 °C. After negative
379 staining with 2 % (wt/vol) uranyl acetate (Serva, Heidelberg, Germany), grid samples were
380 analysed with a transmission electron microscope (Tecnai10, FEI Thermo Fisher Scientific,
381 Eindhoven, the Netherlands) equipped with a Megaview G2 CCD-camera (Emsis, Münster,
382 Germany) as described in detail previously (57).

383

384 *MetSV infection*

385 *M. mazei* cultures (50 ml each) were grown in closed bottles to a turbidity of ~0.2 at 600 nm;
386 then 0.5 ml filtrated (0.2 µm) supernatant of a MetSV-lysed culture was added subsequently.
387 The cultures were further incubated at 37 °C until the cells were completely lysed (8-12 h) or
388 aggregation-formation was obtained (within two weeks).

389

390 *Isolation of viral DNA*

391 Lysed *M. mazei* cultures were filtered (0.2 µm) and the filtrate was centrifuged at 30,000 x g
392 for six h at 4 °C. The pellet was re-suspended overnight in 200 µl nuclease free water (Roth,
393 Karlsruhe, Germany) at 4 °C. DNaseI (Life Technologies, Darmstadt, Germany) treatment
394 was performed for 30 min at 37 °C to destroy contaminating *M. mazei* DNA. After DNaseI
395 treatment and separation of DNAseI, DNA was isolated using the QiaAmp Mini Elute Virus
396 Spin Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. The
397 obtained elution fraction was treated with RNaseA (Life Technologies, Darmstadt, Germany)
398 for 30 min at 37 °C, followed by a proteinase K (Life Technologies, Darmstadt, Germany)

399 treatment for 30 min at 37 °C. To inactivate and remove proteins, ethanol precipitation was
400 performed.

401

402 *Virus genome: Sequencing, assembly and annotation*

403 Viral DNA was isolated. Afterwards, the DNA were processed followed the Rapid Library
404 Preparation Method Manual (Roche, Mannheim, Germany). Multiplex Identifier (MID)
405 Adaptors for Rapid Library (Roche, Mannheim, Germany) were ligated to the DNA
406 fragments. The DNA fragments were cleaned and quantified using the Agilent 2100
407 Bioanalyzer High Sensitivity DNA Analysis kit (Agilent Technologies, Waldbronn,
408 Germany). The individual samples were combined to a library pool. The sample of the final
409 library was run on an Agilent 2100 Bioanalyzer (Agilent Technologies, Waldbronn,
410 Germany) prior to emulsion PCR and sequencing as recommended by Roche. The library was
411 subsequently sequenced on a 454 GS-FLX using Titanium sequencing chemistry. Contigs
412 were reconstructed using Mira v4.0 (58) with the job options “genome,denovo,accurate”. In
413 addition, a shot gun cloning library was generated for closing the gaps. The viral DNA was
414 restricted with HindIII (NEB, Frankfurt, Germany) and cloned into pCR-TOPO® 2.1-vector
415 (Life technologies, Darmstadt; Germany). The contigs were aligned to the shot gun sequences
416 using the CodonCode aligner (CodonCode Corporation, Centerville, USA). ORFs were
417 predicted using ORF finder (National Center for Biotechnology Information (NCBI,
418 (http://www.ncbi.nlm.nih.gov/)). The sequence was deposited at NCBI with the accession
419 MF186604. Phyre2 was used to predict protein domains by homology detections (22).

420

421 *Sequencing CRISPR arrays of* M. mazei *DSM 3647*

CRISPR array IB (spacers 1-29) was amplified using the primers C1for/C1rev2. For the CRISPR array IIIC (spacers 1-9) the primers CRISPR_IIIC_for/CRISPR_IIICsp_r were used (listed in Table S2). The obtained PCR products were TOPO-TA cloned into pCR-TOPO® 2.1-vector (Life Technologies, Darmstadt, Germany) as described by the manufacturer's instructions. Plasmid DNA was generally transformed into *E. coli* DH5α according to the method of Inoue *et al*. (59). The insert of the resulting plasmids was sequenced by the method of Sanger (at the Institute of Clinical Molecular Biology (IKMB), Kiel, Germany).

Bioinformatic analysis of Methanosarcina *draft genomes*

All *Methanosarcina mazei* draft genomes in this study were downloaded from Youngblut *et al*. (27) and NCBI in 2015 (for accession numbers see Table S1). CRISPR loci were predicted using an in-house tool on all draft genomes. CRISPR loci orientations were predicted using the CRISPRstrand tool (60). The consensus direct repeat for each CRISPR loci is the most frequent direct repeat in this CRISPR array. All CRISPR spacers were compiled into a single database and compared against the MetSV sequence using FASTA tool version 36.3.6 (61) with optimizing the number of allowed hits for the MetSV. To estimate the significance of a hit between the MetSV sequence and the CRISPR spacer with a certain number of mismatches, we used a shuffling approach, which allowed us to distinguish between a random match (false positive) and truth match between spacer sequences and MetSV. Thus, for each spacer sequence in this dataset a di-or tri-nucleotide shuffle was generated. According to the literature, the number of mismatches can be very high in some instances (62). Therefore, we used a modified version of IntaRNA (63) to predict RNA-DNA duplexes. Protospacer adjacent motif (PAM) plays a significant role in CRISPR-mediated immunity. In our study,

we determine the consensus PAM by extracting five nucleotides upstream of the putative protospacers from the virus genome and then aligning them.

Protein analysis by SDS-PAGE and Mass Spectrometry

Filtered MetSV-lysed *M. mazei* DSM 2647 cultures were centrifuged at 30,000 x g for six h at 4 °C. The resulting pellet was re suspended in de-ionized H_2O over night at 4 °C. Protein extracts were generated by adding SDS-PAGE loading buffer (125 mM Tris-HCl (pH6.8); 140 mM SDS; 0.3 mM bromophenol blue; 10 % β-mercatoethanol [v/v]) and incubating the re-suspended pellet at 100 °C for 10 min. Samples were analysed by 12.5 % SDS-PAGE according to Laemmli (64). Prior to silver staining, the gels were washed with solution 1 (20 ml ethanol, 4 ml glacial acetic acid, 16 ml H_2O) over night and two times for 20 min with solution 2 (39 ml ethanol, 91 ml H_2O) at room temperature. Silver staining was performed as described by Rabilloud *et al*. (65).

Additionally, a stacking gel based SDS-PAGE followed by LC-MS analysis (28) was employed for the analysis of the proteome of two MetSV samples. Briefly, 10 µg aliquots of the protein extract, either without or with washing of the filtered extract, sample #1 and #2 respectively, were suspended in SDS-PAGE loading buffer, heated at 95 °C for 5 min and loaded onto 12% SDS-PAGE. The samples were migrated into the gel via constant voltage (40 V) for 15 min, followed by 80 V for a further 5 min to allow the proteins to enter the resolving gel, then a Coomassie Brilliant Blue staining was performed.

In-gel digestion

For in-gel digestion, protein bands, and protein containing regions (ca. 25 mm^2 of the stacking gel cleaned SDS-PAGE, were excised and destained with 30 % acetonitrile (ACN)/70 % ammonium bicarbonate (50 mM, pH: 7.4). Destained gel slices were dehydrated by addition of 100 % ACN. 70 µl of 10 mM 1,4-D/L-dithiothreitol were added for disulfide reduction (60 min at 56 °C). The supernatant was removed and 70 µl of 55 mM iodoacetamide were added for cysteine alkylation incubation in the dark for 60 min at room temperature. Gel slices were dehydrated with 100 % ACN. For enzymatic digestion of the protein bands, 1.5 µg of the enzyme (trypsin or GluC, sequencing grade, Promega, Madison, USA), or 0.3 µg for the stacking gel protein containing regions, were added in 10 µl 50 mM ammonium bicarbonate, pH: 7.4. After 10 min incubation, additional 90 µl ammonium bicarbonate buffer were added and the digestion was continued for 16 h at 37 °C. Supernatants were transferred to new tubes and the peptides were extracted from the gel pieces with increasing concentrations of acetonitrile (60 % ACN/70 % ammonium bicarbonate and 100 % ACN), pooled with the supernatants and dried by applying vacuum in a SpeedVac (Eppendorf, Concentrator plus, Wesseling-Berzdorf, Germany). For LC-MS/MS analysis, samples were re-suspended in 20 µl 3 % acetonitrile, 0.05 % formic acid.

LC-MS/MS analysis of peptides

For LC-MS/MS analysis, 5 µl of each peptide sample were used. Peptide separation was performed on a Dionex Ultimate 3000 UHPLC system (Acclaim® PepMap 100 analytical column, 2 µm, 100 Å, 75 µm×500 mm) coupled online to a Thermo Q Exactive Plus Orbitrap mass spectrometer. For chromatographic separation of peptides from the protein bands, 0.05 % formic acid as buffer A and 80 % ACN/0.04 % FA as buffer B were used for the following gradient: 0 min, 5 % B; 5 min, 5 % B; 95 min, 70 % B; 100 min, 95 % B; 110 min, 95 % B; 110.1 min, 5 % B; 120 min, 5 % B. For chromatographic separation of peptides from the stack gel approach, the same eluents were used across the following gradient: 0 min, 4 %

B; 2 min 4 % B; 82 min, 20 % B; 122 min, 40 % B; 130 min, 90 % B; 140 min, 90 % B; 141 min, 4 % B; 150 min, 4 % B. The flow-rate was set to 300 nl/min with a column oven temperature to 30 °C for protein band peptides, and 45 °C for the peptides from the stack gel approach. For MS/MS of the protein band peptides, the 10 most intensive ions were fragmented by using HCD at normalized collision energy of 25 (full scan MS: 300-2000 *m/z*, resolution 70,000). For MS/MS of the stack gel peptides, the 15 most intensive ions were fragmented by HCD at a normalized collision energy of 27.5 (full scan MS: 350-1300 *m/z*, resolution 70,000).For automatic data interpretation, a SequestHT database search was performed on a Proteome Discoverer (version 2.2.0.388) against FASTA databases containing the predicted MetSV ORFs, the proteome of *Methanosarcina mazei* Gö1 (UniProt, accessed 2017.06.26) and common impurities (cRAP). The result filter settings were as follows: peptide confidence - high; peptide rank – 1; min. 1 % of base peak; precursor mass tolerance – 7 ppm; fragment mass tolerance - 0.02 Da, enzyme specificity - trypsin or GluC; max. missed cleavage sites – 2; length - 6 to 144 amino acids; static modifications – carbamidomethyl (Cys); dynamic modifications – oxidation (Met), acetylation (protein N-terminus). To be classified as identified, proteins required at least two high confidence peptides, and a minimum of one unique (proteotypic) peptide.

Lipid extraction

Acid hydrolysis was carried out on freeze-dried biomass of *M. mazei* DSM 3647 and purified cell material of MetSV to release di- and tetraether lipids. In addition, a virus-free culture of *M. mazei* was harvested during its late stationary growth phase and purified as described above to test whether archaeal lipids were successfully retained during the clean-up procedure. All organic tissue (between 3.1 and 20.1 mg) was reacted with 7.5 ml of 5 % hydrochloric acid (HCl) in MeOH for 4 h at 80 °C and then allowed to cool to room

518 temperature. The supernatant was separated from the biomass by centrifugation (4,696 x g; 10
519 min) and then transferred to an 80 ml centrifuge tube. Subsequently, the cell material was
520 extracted three times with 5 ml of a solvent mixture of dichloromethane (DCM): MeOH (3:1;
521 v/v). High performance liquid chromatography (HPLC) grade water (~10 ml) was added to
522 the combined supernatants, resulting in a phase separation. The lipid-containing organic
523 bottom layer was transferred to a round bottom flask and the remaining water layer was
524 washed twice each with 11 ml of DCM. The combined lipid extract was then reduced under
525 vacuum, transferred to a pre-weighted vial and dried under a gentle stream of nitrogen. All
526 lipid extracts were stored at -20 °C and re-dissolved in DCM:MeOH (9.1, v/v) at a
527 concentration of 1 mg/ml prior to analysis.

528

529 *Lipid analysis*

530 Archaeal core lipids were analyzed following the analytical procedure originally described by
531 Hopmans *et al.* (66) and afterwards modified by Liu *et al.* (67). Target compounds were
532 eluted using an Alliance 2695 (Waters, UK) HPLC system interfaced to a ZQ (Micromass,
533 UK) single quadruple mass spectrometer. The HPLC was fitted with a Grace Prevail Cyano
534 HPLC column (3 µm, 150 × 2.1 mm i.d.) and a guard column of the same material.
535 Separation was achieved at a constant temperature of 30 °C with a flow-rate of 0.2 ml/min
536 and following gradient: 90 % A (*n*-hexane)/10 % B (*n*-hexane/propan-2-ol, 90:10, v/v) held
537 isocratically for 5 min, followed by a linear gradient to 81 % A/19 % B in 20 min, then a
538 linear gradient to 100 % B in 35 min, which was held for 5 min. The column was re-
539 equilibrated with 90 % A/10 % B at 0.2 ml/min for 15 min before the next injection.

540 The mass spectrometer was equipped with an atmospheric pressure chemical
541 ionization interface (APCI) source operated in positive ion mode. MS settings were as

follows: source 150 °C, vaporizer 500 °C, corona 2 µA, cone voltage 40 V, extractor 3 V, RF Lens 0.1 V, de-solvation gas (N_2) 8 l/min. Detection of archaeal core lipids was achieved by single ion recording of their protonated molecular ions $[M+H]^+$ (dwell time 234 ms) as reported by Hopmans *et al*. (66) and Liu *et al.* (67). Absolute concentrations of the target compounds were determined after peak integration in the QuanLynx application manager and comparison of area counts with those of an external standard curve obtained by serial dilutions of a synthetic tetraether diol with 77 carbon atoms (68). Similar response factors were assumed for archaeal lipids and the external standard.

RNA Isolation

M. mazei total RNA from different growth phases was isolated by Isol-RNA Lysis Reagent extraction as reported by Nickel *et al.* (46) (5'PRIME GmbH, Hilden, Germany) followed by DNaseI treatment (Life Technologies, Darmstadt; Germany). Quantity and purity of RNA was determined as described previously (69).

Reverse transcriptase (RT-)PCR

One Step RT-PCR Kit (Qiagen, Hilden, Germany) was used for RT-PCR according to manufacturer´s instructions. Samples of 0.5 µg total RNA from *M. mazei* DSM 3647 (wild type) isolated at different time points after infection with MetSV (0, 60, 90 and 210 min) and primers listed in Table S2 were used. The respective PCR fragments were evaluated on a 1 % agarose gel.

Quantitative (q)RT -PCR

qRT-PCR was performed as described by Veit *et al*. (70) using Quatitect® SYBR Green RT PCR Kit (Qiagen, Hilden, Germany) and the ViiA 7 Cycler (Applied Biosystems by Life

567 Technologies, Darmstadt; Germany). Primer pairs used for quantitative RT-PCR reactions are
568 listed in Table S2. Fold change in transcript abundance for each ORF was determined by
569 comparison with the Ct (threshold cycle) of transcripts of three control *M. mazei* genes (i.e.
570 *MM1621*, *MM2181*, *MM1215*) (70, 71). The fold change in the abundance of a transcript was
571 calculated using the following formula *fold change* = $2^{-\Delta\Delta Ct}$ (72).
572
573 *Microarray chip generation*
574 Oligonucleotides (length: 70 nt) for the 3371 predicted ORFs and additional 42 new ORFs
575 from *M. mazei* identified during RNAseq analyses (73) were generated (Eurofins, Ebersberg,
576 Germany). All oligonucleotides were bioinformatically analysed before synthesis to exclude
577 unspecific hybridisation by MWG, Eurofins Genomics (Ebersberg, Germany). The
578 oligonucleotides were resolved in dimethylsulfoxid to a final concentration of 0.3 ng/nl and
579 then spotted onto aminosilane-coated Type-7* microarray slides (GE Healthcare Europe
580 GmbH, Freiburg, Germany) with a Lucidea Array Spotter Generation V (GE Healthcare
581 Europe GmbH, Freiburg, Germany). Each oligonucleotide was spotted in duplicate on the
582 slides and fixed by UV-light. The microarray chips were stored in a vacuum desiccator at
583 20 °C.
584
585 *RNA labelling, hybridization and microarray analysis*
586 The cDNA generation and simultaneous labelling with fluorescent Cy-3 and Cy-5 dyes was
587 performed by using the SuperScript III Reverse Transcriptase Kit (Life Technologies,
588 Darmstadt; Germany). The Lucidea SlidePro hybridisation chamber was used for
589 hybridization to microarray slides. Hybridisation, scanning, data normalization and evaluation
590 was performed using the GenePix Pro software version 6.0 (Axon Instruments, Union City,
591 USA) as described (69, 71). Three independent biological replicates including a dye swap

592 experiment were performed. Differences in transcript levels of selected ORFs were
593 additionally verified by qRT-PCR as detailed above.
594

595 **Acknowledgements**

596 We thank Daniela Hallack, Christopher Noack and Heinke Buhtz for technical support. As
597 well as Alena Wiegandt for MS analysis of the 11 and 16 kD protein bands. Armin
598 Ehrenreich (Department of Microbiology, Technical University of Munich, Germany) is
599 acknowledged for spotting microarrays and Thierry Benvegnu for providing an aliquot of the
600 tetraether diol standard. This study was supported by the DFG (FOR1680, SCHM1052/12-1
601 and 2).
602

603 **Tables**

604 **Table 1:** Predicted ORFs of MetSV

predicted ORF	Genome position	**BRE**/TATA box (start nt BRE/start nt TATA box)	number of aa/predicted protein size	Second structure prediction (Phyre2)	Similarities (using DELTA BLAST)	Query cover/ E value/ Identity (DELTA BLAST)	Domain prediction	Function prediction	Proteome analysis*
ORF1	295-465	**GAA**TTTTAAAT (-28/-25)	56 aa/ 6160 Da	47 % disordered 78 % α helix 7 % β strand	pyruvate formate-lyase [*Desulfuromonas* sp. TF]	89% 6.7 42%			**(+/-)**
ORF2	608-733	**GAA**-N2-TATAGTTATA (-38/-33)	41 aa/ 4510 Da	27 % disordered 80 % α helix 0 % β strand	no similarities	----			+
ORF3	733-876	**AAA**-N2-TAATGAAAAA (-35/-30)	47 aa/ 5170 Da	23 % disordered 49 % α helix 17 % β strand	PREDICTED: protein piccolo, partial [*Myotis brandtii*]	81% 0.078 41%	DUF4451, domain with unknown function (NCBI)		
ORF4	903-1025	not detected	41 aa/ 4510 Da	29 % disordered 29 % α helix 32 % β strand	no similarities	----			+
ORF5	1167-1445	not detected	92 aa/ 10120 Da	14 % disordered 14 % α helix 57 % β strand	no similarities	----			+
ORF6	1417-2151	**CAA**-N-AATATA (-25/-21)	244/ 26840 Da	25 % disordered 84 % α helix 0 % β strand	uncharacterized protein Dsimw501_GD16	55% 5.8 18%	Ribbon-helix-helix-domain	pot. DNA-binding protein	**(+/-)**

					851 [*Drosophila simulans*]		(Phyre2)		
ORF7	2313-3392	**GTA**-N4-AAAAGTAAA (-50/-43)	359 aa/ 39490 Da	9 % disordered 32 % α helix 22 % β strand	type B DNA polymerase [*Bacillus cereus*]	72% 2e-16 30%	Conserved Protein Domain Family DNA_pol_B_2 (NCBI)	DNA polymerase	**(+/-)**
ORF8	3728-3967	**GAA**-N3-TAAAATAT (-45/-39)	79 aa/ 8690 Da	35 % disordered 38 % α helix 16 % β strand	hypothetical protein AHOG_08260 [*Actinoalloteichus hoggarensis*]	75% 2.9 35%			**(+/-)**
ORF9	3999-4187	not detected	62 aa/ 6820 Da	10 % disordered 90 % α helix 0 % β strand	hypothetical protein [*Bacillus* sp. FJAT-22058]	66% 0.068 44%	Conserved Protein Domain Family Urocanase (NCBI)		+
ORF10	4344-4697	**GAC**-N5-TTTAAAGTATATA (-35/-27)	117 aa/ 12870 Da	35 % disordered 68 % α helix 2 % β strand	hypothetical protein A2020_10980 [*Lentisphaerae bacterium* GWF2_45_14]	32% 0.34 45%	Conserved Protein Domain Family HTH/ Conserved Protein Domain Family Zn-ribbon_8 (NCBI)		+
ORF11	4679-4822	not detected	47 aa/ 5170 Da	15 % disordered 0 % α helix	hypothetical protein	85% 1.6			**(+/-)**

				89 % β strand	SBOR_1092 [*Sclerotinia borealis* F-4128]	38%			
ORF12	4819-4944	not detected	41 aa/ 4510 Da	32 % disordered 7 % α helix 51 % β strand	no similarities	----			
ORF13	5263-5871	**GAG**-N-TTTTAAAAATTAT (-31/-27)	202 aa/ 22220 Da	8 % disordered 51 % α helix 18 % β strand	ATPase [*Halodesulfurarchaeum formicicum*]	84% 1e-10 30%		ATPase	**(+/-)**
ORF14	6007-6108	not detected	33 aa/ 3630 Da	67 % disordered 0 % α helix 27 % β strand	no similarities	----			
ORF15	6020-6328	**GAC**-N4-TTTTTTTAATT (-52/-45)	102 aa/ 11220 Da	24 % disordered 36 % α helix 27 % β strand	MCE family protein [*Mycobacterium obuense*]	98% 1.3 32%			+
ORF16	6403-6843	**CAA**CACATTAAA (-24/-18)	146 aa/ 16060 Da	25 % disordered 24 % α helix 45 % β strand	hypothetical protein APR63_09490 [*Desulfuromonas* sp. SDB]	88% 3e-06 34%			+
ORF17	7024-7191	**GAC**-N3-AATAAT (-24/-18)	55 aa/ 6050 Da	11 % disordered 91 % α helix 0 % β strand	hypothetical protein SAMN05444173_2079 [*Opitutus* sp. GAS368]	73% 0.89 36%			**(+/-)**
ORF18	7202-7522	**GAC**-N-ATTTAAA (-31/-27)	106 aa/ 11660 Da	50 % disordered 52 % α helix 14 % β strand	hypothetical protein [*Pseudoalteromonas* sp. S3431]	61% 2.3 35%			+
ORF19	7519-	not detected	90 aa/ 9900	16 % disordered	no similarities	----			+

	7791		Da	16 % α helix 49 % β strand					
ORF20	7804-8547	**GAG**TATTA (-29/-26)	247 aa/ 27170 Da	11 % disordered 13 % α helix 54 % β strand	hypothetical protein [*Bacillus* sp. JCM 19047]	22% 3.5 40%			+
ORF21	8560-9393	**GAC**TATTT (-42/-39)	277 aa/ 30470 Da	46 % disordered 34 % α helix 11 % β strand	no similarities				+
ORF22	9409-10413	**GAA**AAATT AA (-25/-22)	334 aa/ 36740 Da	9 % disordered 16 % α helix 48 % β strand	Beta-galactosidase [uncultured *Ruminococcus* sp.]	32% 3.9 27%	an Env-gp36 domain (NCBI)	putative envelope protein	+

* "+" representss a sequence coverage about 50 %; "(+/-)" a sequence coverage between 30- 49 %

610

611 **Table 2:** Host specificity and morphologically phenotypes of different *Methanosarcina*
612 strains

strain/DSM-No.	**cell morphology**	**isolated from environment**	**Effects of MetSV challenge**
M. mazei Gö1* (3A)	single cells	laboratory derived strain	lysis
M. mazei DSM 3647	single cells	anaerobic sewage digester	lysis
M. mazei DSM 2244	single cells	cow dung	lysis
M. mazei DSM 4556	single cells	alkaline mud, oil exploration drilling site	lysis
M. mazei DSM 6300	single cells	biomethanation granules	lysis
M. mazei DSM 7222	single cells	anaerobic sewage digester	lysis
M. mazei DSM 9195	sarcina- like aggregates	paddy field soil	no lysis
M. bakeri DSM 800	sarcina- like aggregates	anaerobic sewage digester	no lysis
M. bakeri DSM 1311	single cells	anaerobic sediment	lysis
M. acetivorans DSM 2834TC2A	single cells	marine mud	no lysis
M. soligelidi DSM26065	sarcina- like aggregates	permafrost-affected soil	no lysis
Methanosarcina spp. DSM 11855	sarcina- like aggregates	ovine rumen	no lysis
M. siliciliae DSM 3028	sarcina- like aggregates	lake sediment	no lysis

613

614

615 **Table 3:** Microarray analyses of aggregated cells after MetSV infection vs. uninfected cells

Function	Number of genes
Higher transcript levels	
Transporter	35
Transcription regulation	11
Metabolism	58
Stress	13
RNA/DNA modification	2
Conserved/hypothetical proteins	95
aa biosynthesis/ RNA processing	5
Lower transcript levels	
Ribosomal proteins	33
Envelope/transporter	9
Conserved/hypothetical proteins	17
Metabolism	21
DNA/RNA synthesis	5
Transcription regulation	3
aa biosynthesis	19

616

617

618 **Table 4:** Core lipids of *M. mazei* DSM 3647 and MetSV measured by APCI mass

619 spectrometry

Sample	archaeol (ng/g dry biomass)	GDGT-0 (ng/g dry biomass)	archaeol	GDGT-0
M. mazei DSM 3647	1318.85	1.76	99.87 %	0.13 %
MetSV	40.06	0.07	99.83 %	0.17 %

620

621

622 **Table 5:** Transcription levels of MetSV infected aggregates vs. uninfected cells determined

623 by qRT-PCR and microarray analyses.

MM ORF #	predicted function	qRT-PCR	micro array
MM_0002	Dipeptide ABC transporter, binding protein	1.68 ± 0.26	0.13 ± 0.08
MM_0276	Superfamily II DNA and RNA helicase	2.72 ± 0.61	9.11 ± 12.02
MM_0490	Phycocyanin alpha-subunit phycocyanobilin lyase	442.74 ± 43.46	113.01 96.7
MM_0807	Protein translation initiation factor 6 (IF-6)	0.48 ± 0.06	0.24 ± 0.25
MM_1009	Protein translocase, subunit SecE	0.47 ± 0.02	0.06 ± 0.04
MM_1601	Cobalamin biosynthesis protein CobN	265.87 ± 62.27	49.82 ± 57.2
MM_2021	Conserved protein	2.89 ± 1.33	12.07 ± 15.59
MM_2833	ABC transporter, ATP-binding protein	5.15 ± 0.47	74.89 ± 61.67
MM_3096	Hypothetical protein (S-layer domain protein)	10.18 ± 1.14	17.35 ± 21.08

624

625 **Table 6**: Strains and plasmids

Strain	Genotype or description	reference
Methanosarcina. mazei DSM 3647	type strain	DSMZ; ***No. 3647***
M. mazei DSM 2244		DSMZ; ***No. 2244***
M. mazei DSM 4556		DSMZ; ***No.4556***
M. mazei DSM 6300		DSMZ; ***No.6300***
M. mazei DSM 7222		DSMZ; ***No.7222***
M. mazei DSM 9195		DSMZ; ***No. 9195***
M. bakeri DSM 1311		DSMZ; ***No. 1311***
M. bakeri DSM 800	type strain	DSMZ; ***No. 800***
M. acetivorans DSM 2834	type strain	DSMZ; ***No. 2834***
M. soligelidi DSM 26065	type strain	DSMZ; ***No. 26065***
M. spp. DSM 11855		DSMZ; ***No.11855***
M. siliciliae DSM 3028	type strain	DSMZ; ***No. 3028***

Escherichia coli DH5a	general cloning strain	(74)
Virus	**description**	**reference**
Methanosarcina spherical virus (MetSV)		this study
Plasmid	**description**	**reference**
pCR®2.1-TOPO®	*plac, lacZα-ccB, KanR, amp^R, pUC ori*, cloning vector	Life technologies, (Darmstadt; Germany)

626

627

628

629 **Figure legends**

630

631 **Figure 1: Transmission electron micrograph of 5 intact and of 3 empty (negatively**
632 **stained) MetSV particles**

633

634 **Figure 2: Genome map of MetSV.** All predicted ORFs with a size > 90 aa are depicted.
635 Dark grey arrows indicate ORFs with a predicted function. All predicted ORFs are located on
636 the same strand and thus have the same direction of transcription. The nucleotide sequences of
637 the inverted terminal repeats (ITR) on both ends are depicted.

638

639 **Figure 3: Detection of MetSV transcripts after *M. mazei* infection**. 0, 30, 90 and 210 min
640 past infection with MetSV transcripts of the predicted ORFs 6, 7, 13, 19 and 22 were analysed
641 by RT-PCR using specific primer sets (see Table S2). The respective RT-PCR products were
642 analysed by agarose gel electrophoresis; Lane 1-4, ORF6; lane 5-8, ORF7; lane 9-12, ORF13;
643 lane 13-16, ORF19; lane 17-20, ORF15; lane 21-24, ORF22; $T_{0\ min}$ lanes 1, 5, 9, 13, 17 and
644 21; $T_{30\ min}$ 2, 6, 10, 14, 18 and 22; $T_{90\ min}$ lanes 3, 7, 11, 15, 19 and 23; $T_{210\ min}$ lanes 4, 8, 12,
645 16, 20 and 24; M, 1 kb marker (ThermoFisher Scientific, Darmstadt, Germany).

646 **Figure 4: Protein pattern of MetSV.** Total protein extracts MetSV were separated by SDS-
647 PAGE followed by silver staining.. Lane 1: MetSV protein extract after a second wash step of
648 the virus particles lane 2: MetSV protein extract without a second wash step; M: LMW
649 marker (GE Healthcare Europe GmbH, Freiburg, Germany).

650 **Figure 5: HPLC–APCI-MS chromatogram.** Distribution of archaeol, glycerol dialkyl
651 glycerol tetraether witout cyclopentane mojety (GDGT-0) in *M. mazei* DSM 3647 (A);
652 MetSV (B) and in the control sample (C) were illustrated.

653

654 **Figure 6: Challenging *M. mazei* with MetSV and establishment of virus resistance. (A)**
655 Representative growth curve of an early exponential growing *M. mazei* cultures (50 ml)
656 (turbidity at 600 nm ~ 0.2) challenged with 0.5 ml filtrated MetSV-lysed culture supernatant
657 (indicated by the arrow). After approximately 4 h most of the cells were lysed by MetSV. **(B)**
658 Approximately 200 h p.i. with MetSV the *M. mazei* (DSM 3467) culture started to form
659 significant aggregates, which were not susceptible for the virus**.** The CRISPR arrays IB of
660 four independent aggregated cultures (1-4) were cloned and sequenced, but no integration of a
661 MSV specific spacer could be detected. However, rearrangements of the existing spacers
662 (original numbering) were obtained **(C).**

663

1. **Snyder JC, Bolduc B, Young MJ.** 2015. 40 Years of archaeal virology: Expanding viral diversity. Virology **479-480:**369-378.
2. **Pina M, Bize A, Forterre P, Prangishvili D.** 2011. The archeoviruses. FEMS Microbiol Rev **35:**1035-1054.
3. **Pietila MK, Demina TA, Atanasova NS, Oksanen HM, Bamford DH.** 2014. Archaeal viruses and bacteriophages: comparisons and contrasts. Trends Microbiol **22:**334-344.
4. **Meile L, Jenal U, Studer D, Jordan M, Leisinger T.** 1989. Characterization of OM1, a virulent phage of Methanobacterium thermoautotrophicum Marburg. Arch Mikrobiol **152:**105-110.
5. **Jordan M, Meile L, Leisinger T.** 1989. Organization of Methanobacterium thermoautotrophicum bacteriophage psi M1 DNA. Mol Gen Genet **220:**161-164.
6. **Pfister P, Wasserfallen A, Stettler R, Leisinger T.** 1998. Molecular analysis of *Methanobacterium* phage psiM2. Mol Microbiol **30:**233-244.
7. **Luo Y, Pfister P, Leisinger T, Wasserfallen A.** 2001. The genome of archaeal prophage PsiM100 encodes the lytic enzyme responsible for autolysis of *Methanothermobacter wolfeii*. J Bacteriol **183:**5788-5792.
8. **Nölling J, Groffen A, De Vos WM.** 1993. Φ Fl and Φ F3, two novel virulent, archaeal phages infecting different thermophilic strains of the genus *Methanobacterium*. Journal of General Microbiology **139:**251 251-2516.
9. **Wood AG, Whitman WB, Konisky J.** 1989. Isolation and characterization of an archaebacterial viruslike particle from *Methanococcus voltae* A3. J Bacteriol **171:**93-98.
10. **Krupovic M, Bamford DH.** 2008. Archaeal proviruses TKV4 and MVV extend the PRD1-adenovirus lineage to the phylum Euryarchaeota. Virology **375:**292-300.
11. **Lai MC, Chen SC.** 2001. *Methanofollis aquaemaris* sp. nov., a methanogen isolated from an aquaculture fish pond. Int J Syst Evol Microbiol **51:**1873-1880.
12. **Borrel G, Colombet J, Robin A, Lehours AC, Prangishvili D, Sime-Ngando T.** 2012. Unexpected and novel putative viruses in the sediments of a deep-dark permanently anoxic freshwater habitat. ISME J **6:**2119-2127.
13. **Chien IC, Meschke JS, Gough HL, Ferguson JF.** 2013. Characterization of persistent virus-like particles in two acetate-fed methanogenic reactors. PLoS One **8:**e81040.
14. **Leahy SC, Kelly WJ, Altermann E, Ronimus RS, Yeoman CJ, Pacheco DM, Li D, Kong Z, McTavish S, Sang C, Lambie SC, Janssen PH, Dey D, Attwood GT.** 2010. The genome sequence of the rumen methanogen *Methanobrevibacter ruminantium* reveals new possibilities for controlling ruminant methane emissions. PLoS One **5:**e8926.
15. **Samuel BS, Hansen EE, Manchester JK, Coutinho PM, Henrissat B, Fulton R, Latreille P, Kim K, Wilson RK, Gordon JI.** 2007. Genomic and metabolic adaptations of *Methanobrevibacter smithii* to the human gut. Proc Natl Acad Sci U S A **104:**10643-10648.
16. **Hansen EE, Lozupone CA, Rey FE, Wu M, Guruge JL, Narra A, Goodfellow J, Zaneveld JR, McDonald DT, Goodrich JA, Heath AC, Knight R, Gordon JI.** 2011. Pan-genome of the dominant human gut-associated archaeon, *Methanobrevibacter smithii,* studied in twins. Proc Natl Acad Sci U S A **108 Suppl 1:**4599-4606.
17. **S. Makino NA, H. Koike and M. Suzuki.** 1999. Prophages inserted in archaebacterial genomes. Proc Japan Acad **75:**166-171.
18. **Haring M, Peng X, Brugger K, Rachel R, Stetter KO, Garrett RA, Prangishvili D.** 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:**233-242.
19. **Deppenmeier U, Johann A, Hartsch T, Merkl R, Schmitz RA, Martinez-Arias R, Henne A, Wiezer A, Baumer S, Jacobi C, Bruggemann H, Lienard T, Christmann A, Bomeke M, Steckel S, Bhattacharyya A, Lykidis A, Overbeek R, Klenk HP, Gunsalus RP, Fritz HJ, Gottschalk G.** 2002. The genome of *Methanosarcina mazei:* evidence for lateral gene transfer between bacteria and archaea. J Mol Microbiol Biotechnol **4:**453-461.

20. **Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, Bryant SH.** 2015. CDD: NCBI's conserved domain database. Nucleic Acids Res **43:**D222-226.
21. **Iyer LM, Makarova KS, Koonin EV, Aravind L.** 2004. Comparative genomics of the FtsK-HerA superfamily of pumping ATPases: implications for the origins of chromosome segregation, cell division and viral capsid packaging. Nucleic Acids Res **32:**5260-5279.
22. **Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ.** 2015. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc **10:**845-858.
23. **Olmstedt RA, Hirsch VM, Purcello RH, Johnson PR.** 1989. Nucleotide sequence analysis of feline immunodeficiency virus: Genome organization and relationship to other lentiviruses. Proc Natl Acad Sci **86:**8088-8092.
24. **Makarova KS, Haft DH, Barrangou R, Brouns SJ, Charpentier E, Horvath P, Moineau S, Mojica FJ, Wolf YI, Yakunin AF, van der Oost J, Koonin EV.** 2011. Evolution and classification of the CRISPR-Cas systems. Nat Rev Microbiol **9:**467-477.
25. **Makarova KS, Wolf YI, Alkhnbashi OS, Costa F, Shah SA, Saunders SJ, Barrangou R, Brouns SJ, Charpentier E, Haft DH, Horvath P, Moineau S, Mojica FJ, Terns RM, Terns MP, White MF, Yakunin AF, Garrett RA, van der Oost J, Backofen R, Koonin EV.** 2015. An updated evolutionary classification of CRISPR-Cas systems. Nat Rev Microbiol **13:**722-736.
26. **Vestergaard G, Garrett RA, Shah SA.** 2014. CRISPR adaptive immune systems of Archaea. RNA Biol **11:**156-167.
27. **Youngblut ND, Wirth JS, Henriksen JR, Smith M, Simon H, Metcalf WW, Whitaker RJ.** 2015. Genomic and phenotypic differentiation among *Methanosarcina mazei* populations from Columbia River sediment. ISME J **9:**2191-2205.
28. **Hung CW, Klein T, Cassidy L, Linke D, Lange S, Anders U, Bureik M, Heinzle E, Schneider K, Tholey A.** 2016. Comparative Proteome Analysis in Schizosaccharomyces pombe Identifies Metabolic Targets to Improve Protein Production and Secretion. Mol Cell Proteomics **15:**3090-3106.
29. **Wang H, Peng N, Shah SA, Huang L, She Q.** 2015. Archaeal extrachromosomal genetic elements. Microbiol Mol Biol Rev **79:**117-152.
30. **Bamford DH, Ravantti JJ, Ronnholm G, Laurinavicius S, Kukkaro P, Dyall-Smith M, Somerharju P, Kalkkinen N, Bamford JK.** 2005. Constituents of SH1, a novel lipid-containing virus infecting the halophilic euryarchaeon *Haloarcula hispanica*. J Virol **79:**9097-9107.
31. **Jaakkola ST, Penttinen RK, Vilen ST, Jalasvuori M, Ronnholm G, Bamford JK, Bamford DH, Oksanen HM.** 2012. Closely related archaeal *Haloarcula hispanica* icosahedral viruses HHIV-2 and SH1 have nonhomologous genes encoding host recognition functions. J Virol **86:**4734-4742.
32. **Blum H, Zillig W, Mallok S, Domdey H, Prangishvili D.** 2001. The genome of the archaeal virus SIRV1 has features in common with genomes of eukaryal viruses. Virology **281:**6-9.
33. **Bath C, Cukalac T, Porter K, Dyall-Smith ML.** 2006. His1 and His2 are distantly related, spindle-shaped haloviruses belonging to the novel virus group, Salterprovirus. Virology **350:**228-239.
34. **Rensen EI, Mochizuki T, Quemin E, Schouten S, Krupovic M, Prangishvili D.** 2016. A virus of hyperthermophilic archaea with a unique architecture among DNA viruses. Proc Natl Acad Sci U S A **113:**2478-2483.
35. **Savilahti H, Bamford DH.** 1993. Protein-primed DNA replication: role of inverted terminal repeats in the *Escherichia coli* bacteriophage PRD1 life cycle. J Virol **67:**4696-4703.
36. **Gonzalez-Huici V, Salas M, Hermoso JM.** 2000. Sequence requirements for protein-primed initiation and elongation of phage O29 DNA replication. J Biol Chem **275:**40547-40553.
37. **Porter K, Russ BE, Yang J, Dyall-Smith ML.** 2008. The transcription programme of the protein-primed halovirus SH1. Microbiology **154:**3599-3608.

38. **Kessler A, Brinkman AB, van der Oost J, Prangishvili D.** 2004. Transcription of the rod-shaped viruses SIRV1 and SIRV2 of the hyperthermophilic archaeon *sulfolobus*. J Bacteriol **186:**7745-7753.
39. **Maaty WS, Steffens JD, Heinemann J, Ortmann AC, Reeves BD, Biswas SK, Dratz EA, Grieco PA, Young MJ, Bothner B.** 2012. Global analysis of viral infection in an archaeal model system. Front Microbiol **3:**411.
40. **Frols S, Dyall-Smith M, Pfeifer F.** 2012. Biofilm formation by haloarchaea. Environ Microbiol **14:**3159-3174.
41. **Okutan E, Deng L, Mirlashari S, Uldahl K, Halim M, Liu C, Garrett RA, She Q, Peng X.** 2013. Novel insights into gene regulation of the rudivirus SIRV2 infecting *Sulfolobus* cells. RNA Biol **10:**875-885.
42. **Ortmann AC, Brumfield SK, Walther J, McInnerney K, Brouns SJ, van de Werken HJ, Bothner B, Douglas T, van de Oost J, Young MJ.** 2008. Transcriptome analysis of infection of the archaeon *Sulfolobus solfataricus* with Sulfolobus turreted icosahedral virus. J Virol **82:**4874-4883.
43. **DiMaio D.** 2014. Viral miniproteins. Annu Rev Microbiol **68:**21-43.
44. **Erez Z, Steinberger-Levy I, Shamir M, Doron S, Stokar-Avihail A, Peleg Y, Melamed S, Leavitt A, Savidor A, Albeck S, Amitai G, Sorek R.** 2017. Communication between viruses guides lysis-lysogeny decisions. Nature **541:**488-493.
45. **Semenova E, Jore MM, Datsenko KA, Semenova A, Westra ER, Wanner B, van der Oost J, Brouns SJ, Severinov K.** 2011. Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. Proc Natl Acad Sci U S A **108:**10098-10103.
46. **Nickel L, Weidenbach K, Jager D, Backofen R, Lange SJ, Heidrich N, Schmitz RA.** 2013. Two CRISPR-Cas systems in *Methanosarcina mazei* strain Go1 display common processing features despite belonging to different types I and III. RNA Biol **10:**779-791.
47. **Dy RL, Richter C, Salmond GP, Fineran PC.** 2014. Remarkable Mechanisms in Microbes to Resist Phage Infections. Annu Rev Virol **1:**307-331.
48. **Decker KB, James TD, Stibitz S, Hinton DM.** 2012. The Bordetella pertussis model of exquisite gene control by the global transcription factor BvgA. Microbiology **158:**1665-1676.
49. **Stibitz S, Aaronson W, Monack D, Falkow S.** 1989. Phase variation in Bordetella pertussis by frameshift mutation in a gene for a novel two-component system. Nature **338:**266-269.
50. **Liu M, Deora R, Doulatov SR, Gingery M, Eiserling FA, Preston A, Maskell DJ, Simons RW, Cotter PA, Parkhill J, Miller JF.** 2002. Reverse transcriptase-mediated tropism switching in Bordetella bacteriophage. Science **295:**2091-2094.
51. **Davidson AL, Dassa E, Orelle C, Chen J.** 2008. Structure, function, and evolution of bacterial ATP-binding cassette systems. Microbiol Mol Biol Rev **72:**317-364, table of contents.
52. **Saunders NF, Ng C, Raftery M, Guilhaus M, Goodchild A, Cavicchioli R.** 2006. Proteomic and computational analysis of secreted proteins with type I signal peptides from the Antarctic archaeon *Methanococcoides burtonii*. J Proteome Res **5:**2457-2464.
53. **Deppenmeier U, Blaut M, Mahlmann A, Gottschalk G.** 1990. Reduced coenzyme F420: heterodisulfide oxidoreductase, a proton- translocating redox system in methanogenic bacteria. Proc Natl Acad Sci U S A **87:**9449-9453.
54. **Ehlers C, Veit K, Gottschalk G, Schmitz RA.** 2002. Functional organization of a single *nif* cluster in the mesophilic archaeon *Methanosarcina mazei* strain Gö1. Archaea **1:**143-150.
55. **Sowers KRS, H. J. .** 1995. Archaea: A Laboratory Manual. Cold Spring Harbor Lab, Plainview, NY **2**.
56. **Colombet J, Robin A, Lavie L, Bettarel Y, Cauchie HM, Sime-Ngando T.** 2007. Virioplankton 'pegylation': use of PEG (polyethylene glycol) to concentrate and purify viruses in pelagic ecosystems. J Microbiol Methods **71:**212-219.

57. **Casey E, Mahony J, Neve H, Noben JP, Dal Bello F, van Sinderen D.** 2015. Novel phage group infecting *Lactobacillus delbrueckii* subsp. lactis, as revealed by genomic and proteomic analysis of bacteriophage Ldl1. Appl Environ Microbiol **81:**1319-1326.
58. **Chevreux B, Wetter T, Suhai S.** 1999. Genome Sequence Assembly Using Trace Signals and Additional Sequence Information. , abstr Computer Science and Biology: Proceedings of the German Conference on Bioinformatics (GCB) 99,
59. **Inoue H, Nojima H, Okayama H.** 1990. High efficiency transformation of *Escherichia coli* with plasmids. Gene **96:**23-28.
60. **Alkhnbashi OS, Costa F, Shah SA, Garrett RA, Saunders SJ, Backofen R.** 2014. CRISPRstrand: predicting repeat orientations to determine the crRNA-encoding strand at CRISPR loci. Bioinformatics **30:**i489-496.
61. **Pearson WR, Lipman DJ.** 1988. Improved tools for biological sequence comparison. Proc Natl Acad Sci U S A **85:**2444-2448.
62. **Manica A, Zebec Z, Steinkellner J, Schleper C.** 2013. Unexpectedly broad target recognition of the CRISPR-mediated virus defence system in the archaeon *Sulfolobus solfataricus*. Nucleic Acids Res **41:**10509-10517.
63. **Busch A, Richter AS, Backofen R.** 2008. IntaRNA: efficient prediction of bacterial sRNA targets incorporating target site accessibility and seed regions. Bioinformatics **24:**2849-2856.
64. **Laemmli UK.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. . Nature **227:**680–685.
65. **Rabilloud T, Carpentier G, Tarroux P.** 1988. Improvement and simplification of low-background silver staining of proteins by using sodium dithionite. Electrophoresis **9:**288-291.
66. **Hopmans EC, Schouten S, Pancost RD, van der Meer MT, Sinninghe Damste JS.** 2000. Analysis of intact tetraether lipids in archaeal cell material and sediments by high performance liquid chromatography/atmospheric pressure chemical ionization mass spectrometry. Rapid Commun Mass Spectrom **14:**585-589.
67. **Liu XL, Summons RE, Hinrichs KU.** 2012. Extending the known range of glycerol ether lipids in the environment: structural assignments based on tandem mass spectral fragmentation patterns. Rapid Commun Mass Spectrom **26:**2295-2302.
68. **Brard M, Laine C, Rethore G, Laurent I, Neveu C, Lemiegre L, Benvegnu T.** 2007. Synthesis of archaeal bipolar lipid analogues: a way to versatile drug/gene delivery systems. J Org Chem **72:**8267-8279.
69. **Weidenbach K, Gloer J, Ehlers C, Sandman K, Reeve JN, Schmitz RA.** 2008. Deletion of the archaeal histone in *Methanosarcina mazei* Go1 results in reduced growth and genomic transcription. Mol Microbiol **67:**662-671.
70. **Veit K, Ehlers C, Schmitz RA.** 2005. Effects of nitrogen and carbon sources on transcription of soluble methyltransferases in *Methanosarcina mazei* strain Go1. J Bacteriol **187:**6147-6154.
71. **Veit K, Ehlers C, Ehrenreich A, Salmon K, Hovey R, Gunsalus RP, Deppenmeier U, Schmitz RA.** 2006. Global transcriptional analysis of *Methanosarcina mazei* strain Go1 under different nitrogen availabilities. Mol Genet Genomics **276:**41-55.
72. **Talaat AM, Lyons R, Howard ST, Johnston SA.** 2004. The temporal expression profile of *Mycobacterium tuberculosis* infection in mice. Proc Natl Acad Sci U S A **101:**4602-4607.
73. **Jager D, Sharma CM, Thomsen J, Ehlers C, Vogel J, Schmitz RA.** 2009. Deep sequencing analysis of the *Methanosarcina mazei* Go1 transcriptome in response to nitrogen availability. Proc Natl Acad Sci U S A **106:**21878-21882.
74. **Hanahan D.** 1983. Studies on transformation of Escherichia coli with plasmids. J Mol Biol **166:**557-580.

Figure 1

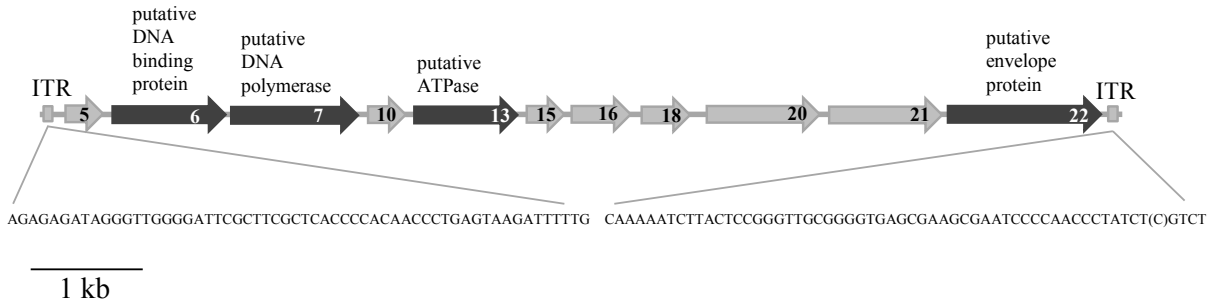

Figure 2

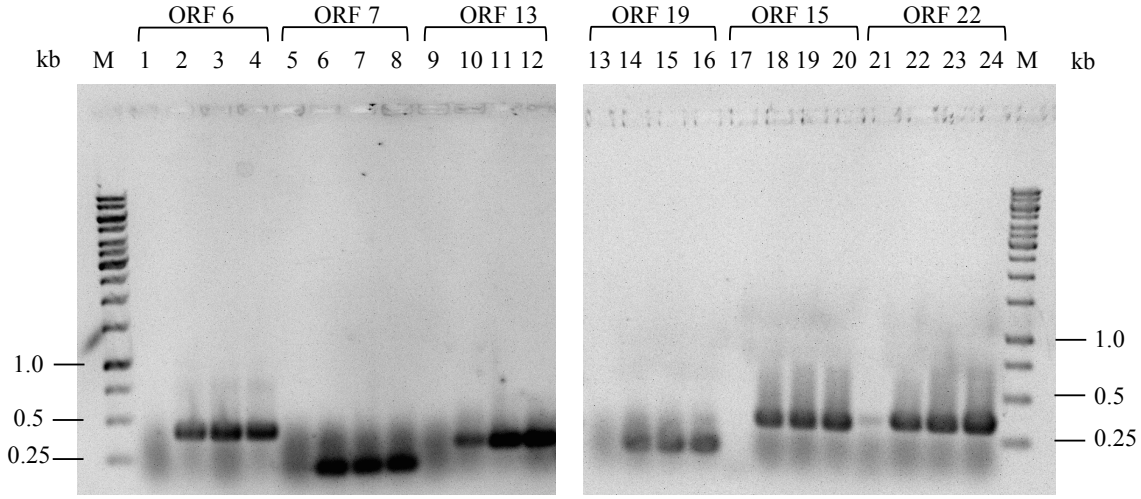

Figure 3

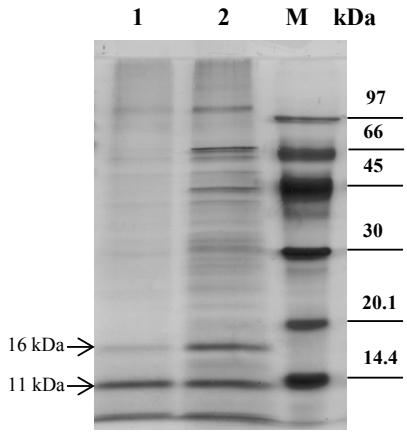

Figure 4

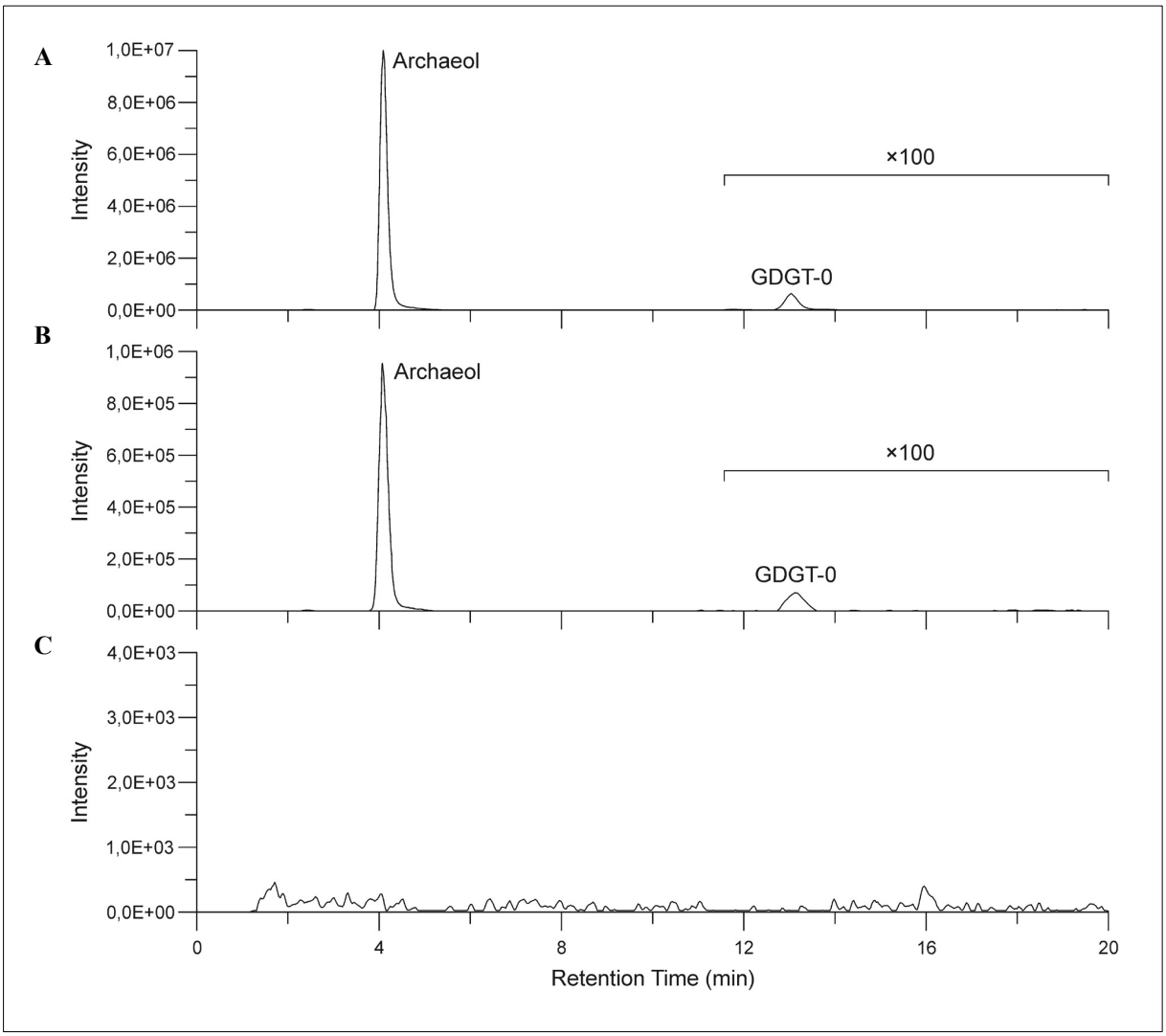

Figure 5

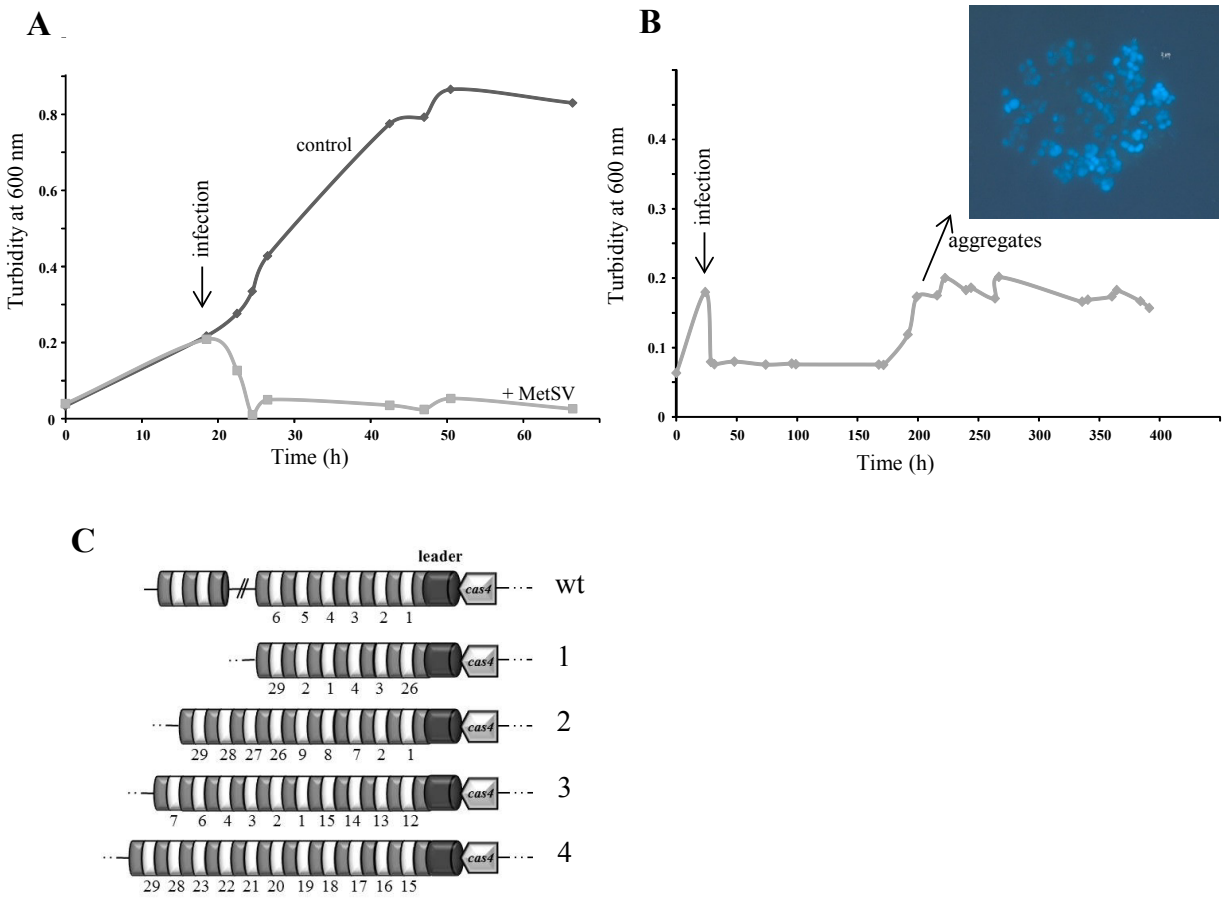

Figure 6