1724–1739 *Nucleic Acids Research, 2023, Vol. 51, No. 4* *Published online 2 February 2023*
https://doi.org/10.1093/nar/gkad011

A clade of RHH proteins ubiquitous in Sulfolobales and their viruses regulates cell cycle progression

Li Xuyang, Lozano-Madueño Cristina, Martínez-Alvarez Laura* and Peng Xu*

Department of Biology, University of Copenhagen, Ole Maaløes Vej 5, 2200 Copenhagen N, Denmark

Received November 17, 2022; Revised December 30, 2022; Editorial Decision December 30, 2022; Accepted January 29, 2023

ABSTRACT

Cell cycle regulation is crucial for all living organisms and is often targeted by viruses to facilitate their own propagation, yet cell cycle progression control is largely underexplored in archaea. In this work, we reveal a cell cycle regulator (aCcr1) carrying a ribbon-helix-helix (RHH) domain and ubiquitous in the Thermoproteota of the order Sulfolobales and their viruses. Overexpression of several aCcr1 members including gp21 of rudivirus SIRV2 and its host homolog SiL_0190 of *Saccharolobus islandicus* LAL14/1 results in impairment of cell division, evidenced by growth retardation, cell enlargement and an increase in cellular DNA content. Additionally, both gp21 and SiL_0190 can bind to the motif AGTATTA conserved in the promoter of several genes involved in cell division, DNA replication and cellular metabolism thereby repressing or inducing their transcription. Our results suggest that aCcr1 silences cell division and drives progression to the S-phase in Sulfolobales, a function exploited by viruses to facilitate viral propagation.

INTRODUCTION

Viruses are dependent on host cellular machinery for propagation and therefore have evolved mechanisms to subvert host resources for their own benefit. One of the strategies is to inhibit cell division in order to keep host cells in a state favorable for DNA replication. Examples include many eukaryotic viruses (1), a couple of bacteriophages (2–4) and one family of archaeal viruses (5). While eukaryotic and bacterial viruses have been shown to inhibit cell division through direct protein-protein interaction with cell division machineries or cell cycle regulators, how the archaeal virus STSV2 inhibits cell division is not known (5).

The cell cycle of archaea of the order Sulfolobales (previously Crenarchaeaota, now phylum Thermoproteota following to the recent taxonomic classification of the archaea (6)) is reminiscent of that in eukaryotes displaying a gap 1 (G1) phase, a DNA synthesis phase (S phase), a long gap 2 (G2) phase and a rapid cell division (M and D) phase (7,8). In addition, most members of the Thermoproteota employ eukaryotic ESCRT (endosomal sorting complexes required for transport)-III and Vps4 proteins for cell division while other phyla, including the Halobacteriota, Methanobacteriota and Thermoplasmatota (previously grouped as phylum Euryarchaeota) use bacterial-like FtsZ for division (9,10). ESCRT-III and Vps4 coding genes have been identified in other archaeal genomes including those of Asgardarchaeota (Caspi and Dekker, 2018), suggestive of similar cell division mechanisms in these organisms.

Several transcription factors are known to activate or inhibit transcription of genes needed for entering into, or exit from, the cell division phase in eukaryotes (11–13) while much less is known about transcription factors controlling cell division in bacteria and archaea. A transcription factor conserved in many alphaproteobacteria, CtrA, controls cell division by binding to promoters of genes involved in cell cycle progression, as well as genes important for other cellular processes (14). In archaea, the only known transcription factor controlling cell division is a ribbon-helix-helix (RHH) protein, CdrS, which activates transcription of cell division genes (15,16).

RHH proteins are extremely widespread in Archaea and Bacteria, but absent in Eukarya. RHH domain containing proteins are DNA-binding and usually dimers (17). A previous study found RHH-containing proteins and glycosyltransferases to be extensively shared among crenarchaeal viruses, an outstanding feature given that viral families of the archaea domain share very few connector genes and hence appear to have distinct evolutionary histories (18). The distribution pattern of viral RHH transcription regulators suggested an important function during viral infection, and we set up to investigate the role of these proteins using rudivirus SIRV2, infecting *Saccharolobus islandicus* LAL14/1, as our system of study. We demonstrate in this work that an RHH protein highly conserved in archaeal viruses and their host genomes inhibits cell division.

*To whom correspondence should be addressed. Email: peng@bio.ku.dk
Correspondence may also be addressed to Laura Martinez-Alvarez. Email: laura.martinez@bio.ku.dk

MATERIALS AND METHODS

Strains and growth conditions

Cultures of *S. islandicus* LAL14/1 △ arrays (19) were grown at 78°C in SCVU medium (basic salts medium (20) supplemented with 0.2% sucrose, 0.2% casamino acids, vitamin mixture and 0.002% uracil) with agitation. After plasmid complementation of the △*pyrEF* mutation, uracil was removed from the medium. Induction of protein expression and knockdown cassettes was done by growing cells in ACV medium, which has the same composition as SCV but substitutes sucrose for 0.2% D-arabinose. *Escherichia coli* strains DH5α and Rossetta (DE3) were used for plasmid cloning and protein expression, respectively, and grown in LB medium supplemented with the appropriate antibiotics. Cellular strains used in this work are listed in Supplementary Table S5.

Virus propagation and infection of cell cultures

S. islandicus LAL14/1 cultures were grown until the optical density (OD) at 600 nm is 0.2, then infected with SIRV2 and incubated for around 48 h until the OD_{600} decreased. Cell were removed by centrifugation at 5000 × g for 10 min and the virions in the supernatant were recovered by addition of 10% polyethylene glycol 6000 (PEG6000) and 1 M NaCl, followed by an overnight incubation at 4°C. Virus was then collected by centrifugation at 11 000 × g for 30 min and resuspended in 10 mM Tris-acetate pH 6. Remaining cell debris was removed by centrifugation at 11 000 × g for 5 min. Virus titer was determined by plaque assay as described previously (21). Briefly, 100 μl of 10-fold serial dilutions of the virus were added to 2 ml fresh *S. islandicus* LAL14/1 △array cells (OD_{600} around 0.2) and incubated at 78°C for 30 min. Then, the infected cells were mixed with an equal volume of 0.4% Gelzan and spread onto the prewarmed 0.7% Gelzan/SCV plates. Plates were incubated at 78°C for 48 h and plaques formed were counted to determine the virus titer.

Plasmids, oligonucleotides and DNA sequencing

A list of the plasmids and oligonucleotides used in this study is provided in Supplementary Tables S5–S9. Oligonucleotides and gBlocks were synthesized by Integrated DNA Technologies, Inc. All constructs were verified through Sanger sequencing, using the services of Eurofins Genomics.

Plasmid construction

Plasmids were constructed using the vectors pGE2 and pEXA2 as backbones. The plasmid pEXA2 is a shuttle vector, it contains a *pyrEF* cassette as a selective marker in *Sulfolobus* (22) and pGE2 is a derivative of pEXA2 (19). Vector pGE2-MCS was constructed by inserting multiple cloning site 1 (MCS1) synthesized as a gBlock into pGE2 upstream of the pEXA_Sy primer binding-site using overlapping-end PCR.

SIRV2 gp21 was amplified from SIRV2 genomic DNA (NC_004086.1), the DNA fragment obtained was inserted between the NdeI and NotI restriction sites of pGE2, resulting in Sis/pGE2-SIRV2gp21 (Supplementary Table S6). The PCR products were later purified with the PCR purification kit (GeneJET K0702, Thermo Scientific).

The viral homologs of the protein SIRV2gp21 were selected from lipothrixvirus (SIFV0029 NP_445694.1), turrivirus (STIV-A61 YP_024996.1), bicaudavirus (SMV3gp63 YP_009226286.1) and usarudivirus (SIRV9gp19 YP_009362571.1). In addition, the archaeal factors SiL_0190 and SIL_RS14755 were amplified from the host *S. islandicus* LAL14/1 △arrays (19), whereas gBlocks were synthesized for the viral factors SIFV0029, STIV-A61, SMV3gp63 and SIRV9gp19 (Supplementary Table S6).

Construction of a gp21-knockout viral strain

Plasmids carrying CRISPR spacers and homologous sequences for recombination were constructed using protocols described earlier (23,24). Briefly, spacer fragments were generated by annealing of the corresponding complementary oligonucleotides and inserted in the LguI restriction sites of pGE2-MCS to yield plasmids carrying an artificial mini-CRISPR array; then, donor DNA fragments for homologous recombination were obtained by extension overlap PCR and inserted between the NheI and AvrII restriction sites (Supplementary Table S10). The resulting plasmid Sis/pGE2-d-SIRV2 gp21 was transformed by electroporation into *S. islandicus* LAL14/1 △arrays as described previously (25).

Protein expression and purification

The genes SIRV2 gp21 and SiL_0190 were PCR-amplified from the SIRV2 and *S.islandicus* LAL14/1 genomes, respectively, and inserted into the expression plasmid pET30a (Novagen). Recombinant SIRV2 gp21 and SiL_0190 with C-terminal His-tag were expressed using *Escherichia coli* RosettaTM (DE3) cells (Novagen). Protein expression was induced during logarithmic growth through the addition of 0.5 mM isopropyl-thiogalactopyranoside (IPTG) followed by an overnight incubation at 16°C. Cells were collected by centrifugation and the cell pellet was resuspended in lysis buffer (50 mM HEPES pH 7.4, 300 mM NaCl, 20 mM imidazole). Cells were lysed by sonication and French press and the cell extract was clarified by centrifugation at 13 000 × g for 30 min at 4°C. The supernatant was loaded onto a 5 ml His-trap column (GE Healthcare) for protein purification and eluted with a linear imidazole (20–200 mM imidazole) gradient in lysis buffer. Fractions were pooled and concentrated to 1 ml and purified further by size exclusion chromatography (SEC) using a Superdex 200 column using the chromatography buffer (50 mM HEPES pH 7.4, 300 mM NaCl).

RNA sequencing

Samples collected at 0, 6, 12 and 24 h post-induction with arabinose from duplicate cultures of *S. islandicus* LAL14/1 △ arrays harboring an empty pGE2 plasmid (control) or a plasmid expressing SIRV2 gp21. Total RNA was extracted using TRI reagent (Sigma-Aldrich) and RNA quality was

assessed using an RNA 6000 Nano kit in a 2100 Bioanalyzer (Agilent Technologies Denmark ApS). RNA was submitted to Novogene Co., Ltd for library preparation and NovaSeq PE150 sequencing. The BBTools package (BBMap – Bushnell B. – sourceforge.net/projects/bbmap) was used for read quality control: bbduk was used for quality trimming with the recommended settings, rRNA removal was done using bbsplit, leaving between 9.26 to 12.1 million reads for each library and optical duplicates were removed using the clumpify function. Post-QC reads were aligned to the reference genomes of *S. islandicus* LAL14/1 and SIRV2gp21 using bowtie2 (26) and reads for annotated regions in each library were counted with FeatureCounts v2.0.0 (27). Differential gene expression analysis was done using the R packages DESeq2 (28) with a Walden test and EdgeR (29) and genes with a $P \leq 0.05$ were considered as differentially regulated (Supplementary Table 2). Functional annotation of the open-reading frames in the *S. islandicus* LAL14/1 genome was achieved by retrieving the COG annotation of the genome from IMG/M (30), assigning KO terms with the KOfamKOALA web-server (31) and matching open-reading frames to the Conserved Domain Database (CDD) (downloaded on October 2019) (32) using DELTA-BLAST (33) (with the options -evalue 0.01, -max_hsps 1 and -max_target_seqs 1). To evaluate gene conservation within the Sulfolobales, OrthoFinder v2.5.4 (34) was used to determine orthologous genes and obtain a count-matrix of occurrences in 15 representative Sulfolobales genomes (Supplementary Table S4). The results of the above analyses are deposited in the Supplementary Table S3.

Electrophoresis mobility shift assay (EMSA)

DNA probes were obtained either as complementary oligonucleotides annealed into dsDNA or by PCR amplification of 200 bp fragments corresponding to the upstream region of the indicated genes of *S. islandicus* LAL14/1 (Supplementary Tables S7 and S8). DNA probes were radiolabeled with [γ-32P]-ATP (PerkinElmer) using the T4-polynucleotide kinase (New England Biolabs). For EMSA assays, 11.5 nM radiolabeled probe and 100 nM protein were incubated in 20 μl binding buffer (10 mM Tris–HCl, pH 8.0, 10 mM HEPES, 1 mM EDTA, 1 mM dithiothreitol, 50 mM KCl, 50 μg/ml bovine serum albumin) with increasing concentrations of cold competitor DNA (ratios 1:1, 1:10, 1:20 and 1:50) for 30 min at 48°C (35). A PCR fragment containing the T7 promoter was used as cold competitor DNA. Reactions were then mixed with the 6× loading buffer (0.02% bromophenol blue, 40% glycerol, v/v) and samples were loaded into a non-denaturing 12% acrylamide gel in 0.5× TBE buffer (65 mM Tris pH 7.6, 22.5 mM boric acid and 1.25 mM EDTA). Results were recorded by phosphorimaging with a Typhoon FLA-7000 scanner. The dissociation constant (K_D) of the binding of SiL_0190 and gp21 to the *cdvA* promoter was calculated by mixing 15 ng of labelled promoter (5.7 nM) with 1500 ng of cold *T7* promoter with the indicated protein concentrations. The data was fitted to the Hill equation using GraphPad Prism version 9.4.1.

Prediction of binding-motifs

Enriched motifs present in the 80 bp upstream region from the predicted start codon of the SiL_0190 or SIRV2 gp21-bound promoters, or the corresponding sequence of the homolog genes of other members of the Sulfolobales (*Sulfolobus acidocaldarius* DSM 639, *Sulfurisphaera tokodaii*, *Stygiolobus azoricus* DSM 6296, *Metallosphaera hakonensis* DSM 7519 and *Acidianus manzaensis*), were predicted by word-based analysis ($k = 6$, 7 and 8) using the RSAT pipeline (36,37).

Fluorescence microscopy and flow cytometry

Protein expression in triplicate cultures of the *S. islandicus* strains harboring constructs for the expression of different RHH-domain factors or pGE2 as a control were induced at OD_{600} 0.05 by the addition of D-arabinose to a concentration of 0.2%. The growth of the cultures was subsequently monitored by measuring absorbance at 600 nm and a 5 ml aliquot was collected at the indicated time points for microscopy and flow cytometry analyses as described previously (38,39). Briefly, for fluorescence microscopy 1.5 ml samples at $OD_{600} = 0.2$ or the equivalent number of cells at each collection point were centrifuged and fixed with 4% paraformaldehyde, followed by a series of washes with PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4). Next, each sample was spotted on a poly-L-lysine slide (Sigma-Aldrich), dried and stained with 10 μM 4',6-diamino-2-phenylindole (DAPI) (Invitrogen). After another round of PBS wash, the slides were covered with mounting medium (78% glycerol, 1× PBS, 0.2% polyvinylpyrrolidone) and stored at 4°C until analysis. The slides were analyzed with the Nikon Eclipse Ti inverted microscope at 100× magnification and images were collected using the software NIS-Element (RRID: SCR_014329). Relative cell size was determined by measuring cell diameter using ImageJ 1.53t (40) for 50 cells at the indicated timepoints after induction and normalizing the diameter of cells expressing viral factors to the diameter of control cells (harboring empty plasmid). For flow cytometry analyses, a volume of 300 μl of culture at $OD_{600} = 0.08$ or the equivalent number of cells were fixed in ethanol (final concentration of 70%) and incubated overnight. The fixed cells were centrifuged at 2,800 rpm for 20 min and resuspended in 1 ml wash buffer (10 mM Tris pH 7.5, 10 mM $MgCl_2$) and collected again by centrifugation. Finally, the cell pellets were resuspend in 140 μl of staining solution (the washing buffer plus 40 μg/ml ethidium bromide (Thermo-Scientific) and 100 μg/ml mithramycin A (Apollo Scientific), and stained for at least 20 min on ice. Stained cells were analyzed in an Apogee A10 Bryte (Apogee Flow Systems) and a total of 35 000 events were collected for each sample. Two parameters were measured: FSC (forward scattered light) and SSC (side scattered light). The results were plotted and analyzed with the FlowJo™ v10.8.1 and Flowing Software 2. The ratio of normal cells vs. enlarged cells was calculated for each factor by defining a region of interest containing the bulk of the control population as 'normal cells'. Cells to the right of the normal population where considered enlarged.

RNA extraction and cDNA synthesis

Samples (corresponding to ∼1 × 10^9 cells) were collected by centrifugation and stored in 1 ml TRI reagent (Sigma Aldrich) at −80°C. Total RNA was extracted using the TRI reagent following the instructions by the manufacturer. RNA pellets were washed and dissolved in nuclease-free water; RNA integrity verified by gel electrophoresis and concentration was measured using a Nanodrop 1000 spectrophotometer.

RNA samples were first treated at 37°C for 30 min with RNase-free DNase I (Invitrogen) in 50 μl reactions containing about 5μg RNA and 1unit DNase I. The treatment was stopped by adding 2 μl EDTA to a final concentration of 5 mM, followed by incubation at 75°C for 10 min. A 3 μl-treated RNA aliquote was reverse transcribed in 10 μl reaction using RevertAid RT reverse transcription kit (Thermo Scientific) according to the instructions of the manufacturer. As a control for possible contamination from genomic or viral DNA, another 3 μl RNA was processed in the same way except that no reverse transcriptase was added.

Quantitative real-time PCR

The reactions were performed in 10 μl mixtures containing 5 μl PowerUP SYBR Green mastermix (Thermo Scientific), 0.5 mM primers and about 1 ng cDNA. Separate reactions were prepared for detection of reference genes. Oligonucleotide pairs used for qPCR analysis are listed in Supplementary Table S9. The mixtures were prepared in 96-well microtiter PCR plates (Bio-Rad Laboratories), sealed with an adhesive cover and amplified on the Quant Studio 3 (Applied Biosystems) with the following program: denaturing at 95°C for 30 s, 40 cycles of 95°C for 5 s and 56°C for 30 s. Additionally, a non-template control (NTC) was used to check the possibility of unspecific amplification products and contamination. TBP, a housekeeping gene encoding the TATA-binding protein was used as the reference. PCR efficiencies and mean C_q values from the gene-expression analyses are deposited in Supplementary Table S11.

Sequences, identification of homologous proteins and phylogenetic analysis

A manually curated list of genomes from 200 archaeal viruses deposited in NCBI (downloaded on February 2020) and including proviral genomes of thaumarchaeal viruses (41) was used to construct a database containing 11 120 proteins hereby referred to as archaeaVir. A list of genomic accession numbers and identifiers is provided in Supplementary Table S1. Functional annotation of the proteins in the archaeaVir database was assigned by using the program hmmsearch of the HMMER suite (42) to match the sequences to the VOG (43) and pfam (44) databases, and by using DELTA-BLAST (33) to match sequences to the Conserved Domains Database (CDD) (32). Proteins annotated to have an RHH-domain were retrieved and used as queries for the identification of homologous proteins in the archaeaVir database by using PSI-BLAST (45) with a relaxed *e*-value of 0.1 to identify divergent homologs. If two or more queries had overlapping matches, they were merged into the same cluster of homologous proteins. Cellular homologs of gp21 were initially identified through a protein BLAST search to the genome of *S. islandicus* LAL14/1 (evalue 0.1), which retrieved three hits: SIL_RS00945, SIL_RS14755 and SIL_RS07870. A PSI-BLAST iteration of the search added one more hit corresponding to the conserved gene SIL_RS00595. A set of proteins from 15 members of the Sulfolobales order was then used to perform a protein BLAST search (*e*-value = 0.001) to identify the corresponding homolog in every genome, keeping the match with the highest score or the match that has the same genomic context as in the query. Synteny at the genomic level was evaluated using SyntTax (46). Sequences of the proteins containing more than two RHH domains, commonly found in homologs in the Bicaudaviridae viral family, were split into their individual domains for a better comparison. To construct a phylogenetic tree of the SIRV2 gp21 orthologs, viral and cellular sequences were aligned using MUSCLE (47), followed by maximum likelihood phylogenetic analysis with FastTree 2 (v2.1.3) (48) using an LG + CAT model and a Gamma20-based likelihood (option -gamma). The resulting tree was visualized with FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/) and is available as Supplementary Data 1.

Model of SIRV2 gp21 and SiL_0190

The structures of SiL_0190, SIRV2 gp21 and other members of the family were predicted using AlphaFold2 (49). Alignment of the predicted structures to a DNA molecule was done by superimposing the models to the crystal structure of the ArtG transcription regulator of Staphylococcus aureus bound to its target DNA (PDB: 3GXQ (50)).

RESULTS

Homologs of the SIRV2 gp21 RHH-domain transcription regulator are widespread in crenarchaeal viruses

To identify the RHH-containing proteins in archaeal viruses, we annotated 153 archaeal viral genomes deposited in NCBI and retrieved all proteins harboring a RHH domain. This search recovered 194 genes from 90 genomes, mainly from crenarchaeal viruses and with only three highly divergent occurrences in euryarchaeal haloviruses (Supplementary Table S1), similar to the taxonomic pattern observed by Iranzo et al. (18). Classification of the hits revealed that 115 of the 194 queries, including SIRV2 gp21 (also known as ORF59b), belonged to the same cluster and that they are present in 87 viral genomes (Figure 1A). The abundance and broad taxonomic distribution of the members of this cluster contrasts markedly with that of other viral RHH proteins, as the three largest clusters of the remaining queries contain only 19, 10 and 9 members each, including the previously studied rudiviral transcription factor SvtR (35), the homologs of SIRV4gp10, and the fuselloviral factor E73 (51). In addition, these clusters show a narrow distribution restricted to 1–3 viral families (Figure 1A).

The clade of SIRV2 gp21 homologs bears high similarity to a conserved cellular transcription regulator of Sulfolobales

The distribution pattern of RHH regulators in archaeal viruses suggests they were originally acquired from their

A

RHH proteins in archaeal viruses

Protein cluster	Frequency	Viral families
SIRV2gp21	115	10
SvtR	19	3
SIRV4gp10	10	2
E73	9	1
AFV3gp07	8	2
AvtR	8	1
ATSV1gp34	5	2
SSV11 ORF60	5	1
AFV9gp24	2	1
Other	13	4

B

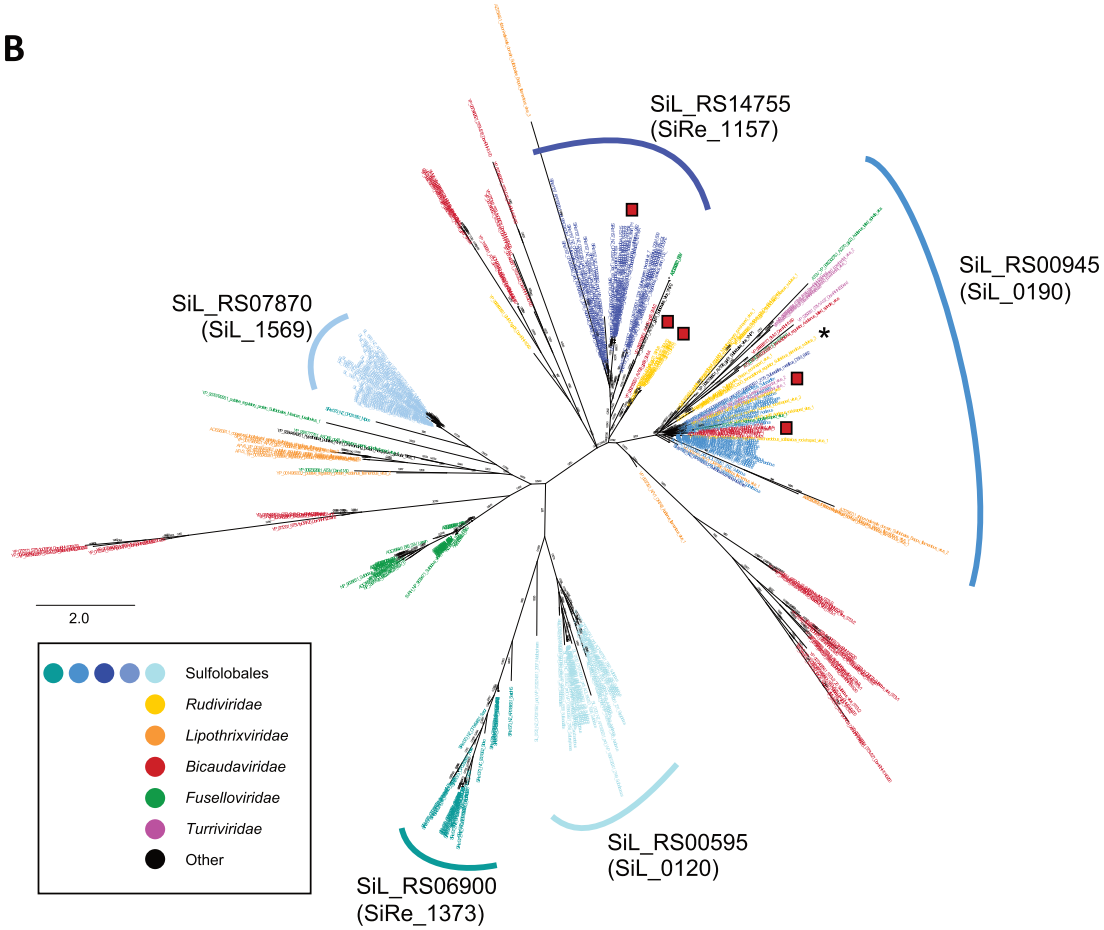

Figure 1. RHH-domain proteins in archaeal viruses. (**A**) Classification of RHH proteins in archaeal viruses. **(B)** Phylogenetic analysis of the SIRV2 gp21 homologs in the Sulfolobales and their viruses. Color-coding is the same as in the inset panel with cellular homologs colored in shades of blue and viral homologs colored according to their viral taxonomy. The asterisk indicates the position of SIRV2 gp21 and the red squares indicate other transcription factors used in this work. Branches were labeled following the gene locus name in *S. islandicus* LAL14/1 (SIL- nomenclature) with the protein name in parenthesis corresponding to the old protein annotation (now outdated). In cases where no SiL- protein name is assigned, the name of the homologous protein in *S. islandicus* REY15 (SiRe- nomenclature) is used.

hosts (18). A BLAST search to investigate the presence of cellular homologs of SIRV2 gp21 in the genome of *S. islandicus* LAL14/1 revealed the presence of five small genes with RHH-domains with the gene identifiers: SIL_RS00945 (also known as SiL_0190), SIL_RS07870 (SiL_1569), SIL_RS00595 (SiL_0120), SIL_RS14755 (no SiL_ annotation) and a homolog of SiRe_1373 (gene 1373 of *S. islandicus* REY15A, the closest relative of LAL14/1 (Supplementary Figure 2C)) not annotated in the genome of LAL14/1 but located in the region between 1 286 819 to 1 286 682 bp. SiL-format protein names are not available for all LAL14/1 genes because these correspond to the previous RefSeq genome annotation where some small genes had not been accurately predicted. The gene identifier (SIL-format) assigned following RefSeq prokaryotic reannotation of genomes in 2017 (52) is used when no protein name is available. The five gp21-homologs in LAL14/1correspond to highly conserved proteins in the Sulfolobales and possess identical genomic neighborhoods (Supplementary Figure S1A–E). Additionally, they appear to be essential to the host as revealed by a previous transposon mutagenesis and sequencing analysis of essential genes in *S. islandicus* (53). A phylogenetic analysis of all viral and cellular homologs of SIRV2 gp21 splits them into several branches, five of which are represented each by one of the LAL14/1 homologs while four are composed uniquely of viral proteins. In addition, there are several interspersed viral homologs (Figure 1B). Interestingly, proteins from diverse viral families show high similarity to the clades of SiL_0190 and SIL_RS14755, with SIRV2 gp21 clustering with the former. While protein identity ranges between 100% and 75.9% among the cellular homologs of SiL_0190, viral members of the clade share between 86.2% and 26.1% identity to SiL_0190 (Supplementary Figure 1). We reasoned that such a high conservation within viruses that have little in common at the genomic level or mode of infection points out to a key function in their interaction with their crenarchaeal hosts.

SIRV2 gp21 regulates transcription of cellular genes involved in cell division and metabolism

SIRV2 gp21 is a 59-amino-acid protein (6.9 kDa) containing a ribbon-helix-helix (RHH) motif indicative of transcription regulator activity (17). RHH-transcription regulators consist of two RHH motifs intertwined in which the N-terminal portion of each RHH monomer folds as a β-sheet and mediates site-specific interactions with the DNA (16). A model of the dimeric structure of SIRV2 gp21 obtained using AlphaFold is shown in Figure 2A. To investigate the role of SIRV2 gp21 during infection of its host *S. islandicus*, we first attempted to take advantage of the recently developed genetic engineering system for SIRV2 (21,24). However, our efforts to construct a viral strain encoding a His-tagged variant of gp21 or a knock-out mutant lacking gp21 were unsuccessful (Supplementary Figure 4A, 4C). Considering that gp21 is a conserved gene in all members of the *Rudiviridae* and regarded as part of their core genome (24), our results suggest that gp21 may be an essential gene for the virus. Initial efforts to identify the binding sequence of gp21 within the viral genome using recombinant protein expressed in, and purified from, *E. coli* showed nonspecific binding to DNA at increased protein concentrations (data not shown). Therefore, we cloned gp21 with a C-terminal His-tag into a plasmid vector for expression in *S. islandicus*. The tag was incorporated to the C-terminus which was shown to be unessential for the replication of SIRV2 (24). Induction of gp21 expression triggered strong growth retardation (Figure 2B), leading us to hypothesize that the target genes of this protein are located within the host genome. Chromatin immunoprotein sequencing (ChIP-seq) analysis to determine the binding sites of gp21 was, however, hindered by the low yield of the protein. Thus, we performed transcriptome profiling by RNA sequencing (RNA-seq) to compare gene expression remodeling after expression of gp21 (Supplementary Table S2).

The principal component analysis of the data shows that the control cultures (cells carrying the empty plasmid) group apart from cells expressing gp21 (Figure 2C). By comparing control versus gp21-expressing cultures (i) before induction (at 0 hours post-induction) and (ii) after induction, we observed 269 and 155 genes, respectively, that are differentially expressed in the gp21-expressing samples with a $\log_2$ fold-change over 1.4 (Supplementary Table S3). We reasoned that the high number of differentially expressed genes could be the result of cascade effects initiated by the gp21-mediated changes and that if members of the SiL_0190 clade are so highly conserved, their target genes would also be conserved within the Sulfolobales, thus we used conservation as criterion to further filter the results. This allowed us to identify 103 genes differentially expressed, of which 8 showed strong modulation in the gp21-expressing cultures even before induction with arabinose, probably due to leaky expression of gp21 from the arabinose promoter (22) (Supplementary Table S3, Figure 2D). The majority of the differentially expressed genes belonged to the COG (Clusters of Orthologous Proteins) modules of Metabolism (47.6% of the genes) and Cellular processes and signaling (17.5%), followed by genes involved in Information storage and processing (9.7%). A more detailed analysis of the functional categories revealed an overrepresentation of genes involved in Cell cycle control, division and chromosome partitioning (11-fold) and Replication, recombination and repair (2-fold) when compared to the distribution of COG terms in the genome of *S. islandicus* LAL14/1 (Figure 2D). Among the metabolism-related categories, genes belonging to the Energy production and conversion (20% of the genes), and Inorganic transport and metabolism (7.7% of the genes) were preferentially regulated in our samples (Figure 2D). Genes from the categories of Transcription, and Translation, ribosomal structure and biogenesis were depleted (5- and 6-fold, respectively) from our dataset. We were particularly interested in genes *sil_1169-sil_1171* (*cdvA, escrt-III-1, vps4*), *sil_1195* (*escrt-III-2*), *sil_1935–1936* (*segAB*), *sil_1121* (*topG2*), *sil_1206* (*cdt1*), *sil_1226* (*gins23*) and *sil_1733* (*cdc6 1*). Expression of the above genes was downregulated 2–6 times (1.5 - 2.5 log2-fold) in the gp21-expressing cultures, with the exception of *sil_1206*, which was upregulated 2.8 times (about 1.7 $\log_2$-fold) (Figure 2E, Supplementary Table S3). The first four are key components of the *S. islandicus* cell division machinery (CdvABC) driving cytokinesis

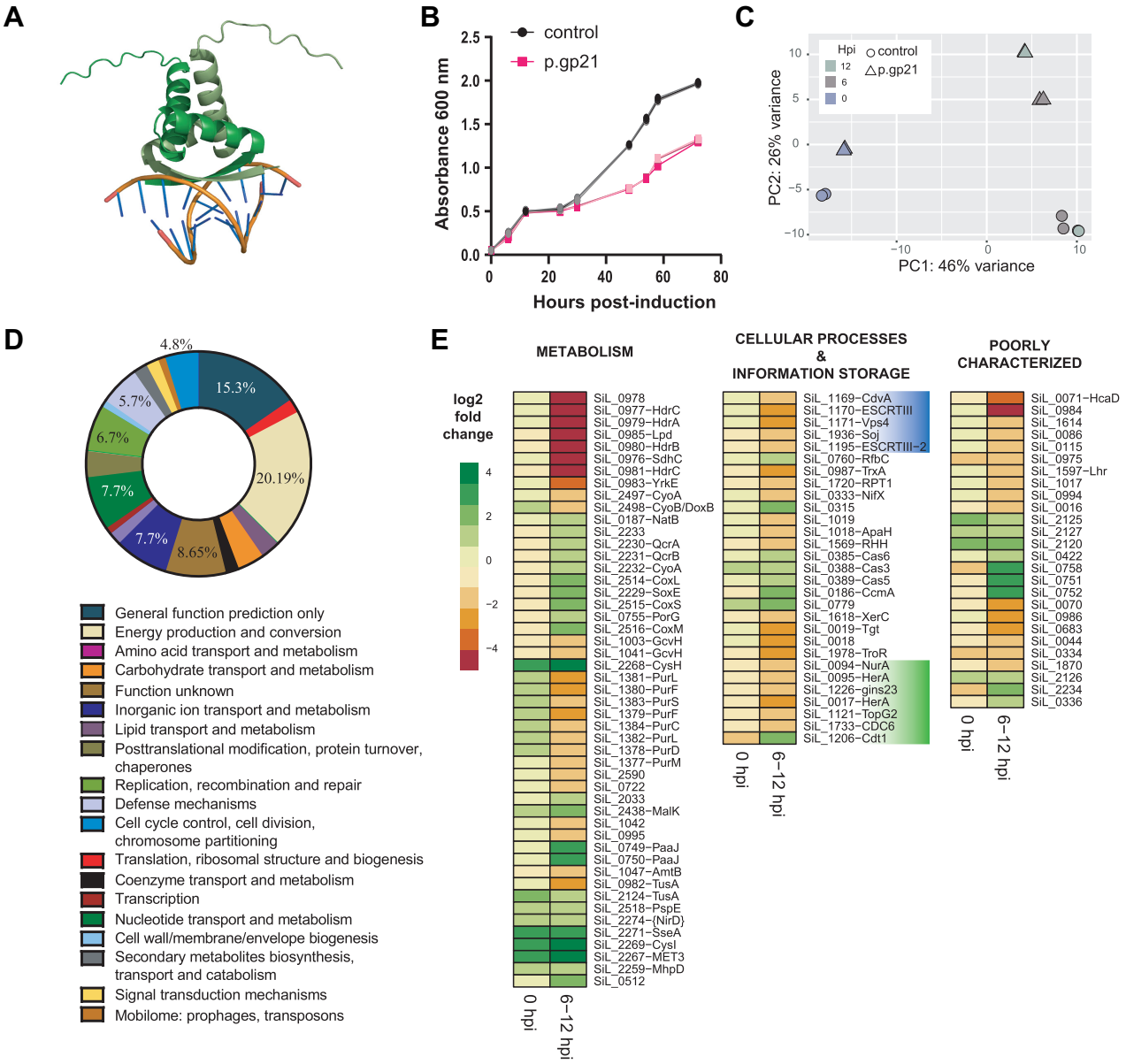

Figure 2. Expression of gp21 alters the transcription profile of genes involved in cell division and DNA replication. (**A**) Model of the SIRV2 gp21 dimer. SIRV2 gp21 was modeled with AlphaFold2 and superimposed to the structure of the RHH-domain ArtA regulator with its target DNA (PDB: 3GXQ). Each chain of the gp21 dimer is indicated in different shades of green. The per-residue confidence score of the model (pLDDT) is high (>90) for the first 40 residues of each monomer and decreases sharply at the C-terminal region corresponding to the unstructured region of the protein. (**B**) Overexpression of gp21 elicits growth retardation. Growth in ACV medium of *S. islandicus* LAL14/1 harboring an empty plasmid (control) or a plasmid expressing gp21 under the control of the inducible arabinose promoter (p.gp21). Growth was monitored by measuring absorbance at 600 nm. (**C**) Principal component analysis of RNA-seq data from *S.islandicus* LAL14/1 cultures taken at 0, 6 or 12 h post-induction (hpi) with arabinose. Control cultures harbor an empty plasmid, while p.gp21 indicates cells containing a plasmid for the expression of SIRV2 gp21. (**D**) Clusters of Orthologous Proteins (COG) functional annotation of the genes selected as having a significant differential transcription in cultures expressing SIRV2 gp21 according to their COG-Category (bottom) classification. (**E**) Heatmap of the transcriptional profile of selected differentially expressed genes in p.gp21 vs. control cultures at 0 hpi with arabinose (basal gp21 expression) and at 6 and 12 hpi, colored by their $\log_2$ fold change and clustered by their COG-Module annotation. Genes belonging to the category of Cell cycle control, cell division and chromosome partitioning are shaded in blue. Genes of the category of replication, recombination and repair are shaded in green.

(54). The operon of *sil_1935-36* code for the components of the segrosom SegAB, involved in chromosome organization prior to partitioning (55). TopG2 is a reverse gyrase presumably involved in DNA replication and repair in the crenarchaea (56). Cdt1 and Cdc6-1 recognize the origins of chromosome replication, and Gins23 is part of the MCM complex, the motor driving unwinding of DNA at the replication fork (57,58). Among the genes assigned to the Metabolism category, components of the heterodisulfide reductase complex (Hdr) were strongly downregulated, as well as genes involved in purine biosynthesis (*sil_1377–1384*) (4 and 2 $\log_2$-fold, respectively). On the other hand, genes coding for cytochrome subunits (*sil_2229–2233*) were upregulated (1.5–2.3 $\log_2$-fold), similarly to several genes in the Inorganic ion transport and metabolism (e.g. *sil_2267–2271*) (3.3–4.4 $\log_2$-fold). These results led us to speculate that gp21 is involved in the regulation of genes required for cell division and genome replication.

Overexpression of members of the sil_0190 clade triggers growth arrest and cell enlargement

We reasoned that inhibition of cell division genes in gp21 overexpressing cells could be related to the growth arrest phenotype observed in these cultures. We performed fluorescence microscopy and flow cytometry analyses of cells at different time-points after induction of plasmid-borne gp21 or an empty plasmid as a control. Surprisingly, we observed a remarkable increase in cell size noticeable after 24 hours post-induction (hpi) (Figure 3A), when 66.7 ± 2.3% of the cells showed cell enlargement equivalent to a 35-fold increase in the cell volume (Figure 3B). Cell enlargement continued with incubation time and cells at 72 hpi had expanded their diameter up to 5-fold, corresponding to an increase in volume of ∼150 times (Figure 3B). Additionally, we observed an increment in the number of genome equivalents within the population of large cells. Whereas the monoploid *Sulfolobus* oscillates between having one chromosome copy during the G1 phase (1C) and two copies after genome replication (2C), DNA content in large cells was significantly increased by 12 hpi when cells mostly contain over two chromosome equivalents (2C) and peaked after 30 hpi, with the majority of the cells containing between 4 and 10 genome equivalents (Figure 3A). Yet, we observed a fraction of the population showing the characteristics of normal cell growth, with respect to size and genome content, that increased from 33% at 24 hpi to 45% at 72 hpi (Figure 3A). This fraction of cells probably explains the continued, albeit slower, growth observed for the gp21-expressing samples (Figure 2A).

The enlarged-cell phenotype with cells containing more than two genome equivalents has been previously observed in *Saccharolobus* and *Sulfolobus* species after silencing of cell division genes, in cells overexpressing dominant-negative mutants of the Cdv complex and in cells where cell division has been inhibited as a result of virus infection (5,54,59,60), suggesting that overexpression of SIRV2 gp21 results in inhibition of cell division, but not DNA replication. Moreover, the high conservation of members of the clade among different families of the Sulfolobales and their viruses suggests that the homologs perform an equivalent function. To explore the functional equivalence of the members of this branch, we transformed a plasmid with an inducible-expression cassette containing RHH-domain proteins of the SiL_0190 and SIL_RS14755 phylogenetic branches, namely the cellular SIL_RS14755 representative of its branch and the members of the SiL_0190 branch including cellular SiL_0190, the viral SIFV gp29 from a lipothrixvirus, the turriviral STIV gp35, SMV3 gp63 from a bicaudavirus and the rudiviral SIRV9 gp19, or an empty vector as a control. All these proteins are predicted by AlphaFold2 to fold into an RHH motif and to dimerize, with the exception of SIFV gp29, which already contains two RHH motifs in the same peptide chain and folds as a monomer (Supplementary Figure 3). We subsequently monitored changes in cell growth curve, size and genome content occurring after induction of target gene expression (Figure 3C). All the constructs transformed with the exception of the vector harboring SIL_RS14755, showed a 10-fold reduction in their transformation efficiency when compared to the empty vector and a small colony phenotype suggestive of slower growth, and induced growth retardation after induction of protein expression (Figure 3D), similar to what was observed with SIRV2 gp21. We were unable to grow the strains harboring a plasmid for SIFV gp29 expression in liquid culture, implying that leaky expression of these factors already causes aberrant growth. Overexpression of SiL_0190 triggered similar changes as we observed for gp21: 63.8 ± 5.2% of the cells increased their size by 24 hpi, decreasing slightly to around 40% with prolonged incubation, while the diameter of this fraction of the population increased up to around 6-times by 72 hpi equivalent to a raise in cell volume of 150-fold (Figure 3E). Moreover, the enlarged cells contained over 6 genome equivalents after 72 h of incubation (Figure 3C). The viral homologs STIV gp35 and SMV3 gp63 also triggered cell enlargement and increased DNA content, albeit in a milder way and with the strongest phenotype seen at 24 hpi, when 57.7 ± 0.7% and 43.4 ± 1.5% of the cells had increased their diameter by 4- and 3-fold, respectively (Figure 3E). Contrastingly, the fraction of the population corresponding to large cells decreased during incubation, stabilizing at 20–25%. Whereas some STIV gp35-expressing cells contained around 6 genome equivalents, cells harboring SMV3 gp63 showed a moderate doubling of their DNA content equivalent to 3–4 chromosome copies (Figure 3E). The milder phenotype caused by the turriviral and bicaudaviral fractions could reflect differences between the native hosts of these viruses and *S. islandicus* LAL14/1. On the other hand, overexpression of SIL_RS14755 did not affect cell size or genome content, indicative of a different function of the members of this clade. Overall, our results support a conserved role of the SiL_0190 clade in the regulation of cell division in the Sulfolobales.

To further investigate if SIRV2 gp21 and SiL_0190 were functionally equivalent, we attempted to substitute gp21 in the genome of SIRV2 with SiL_0190 while keeping intact the native regulatory regions of the former. We employed a CRISPR-based genome editing strategy where a strain harboring a plasmid containing a spacer targeting gp21 and a donor DNA fragment consisting of SiL_0190 flanked by the adjacent sequences to gp21 in the viral genome was used to

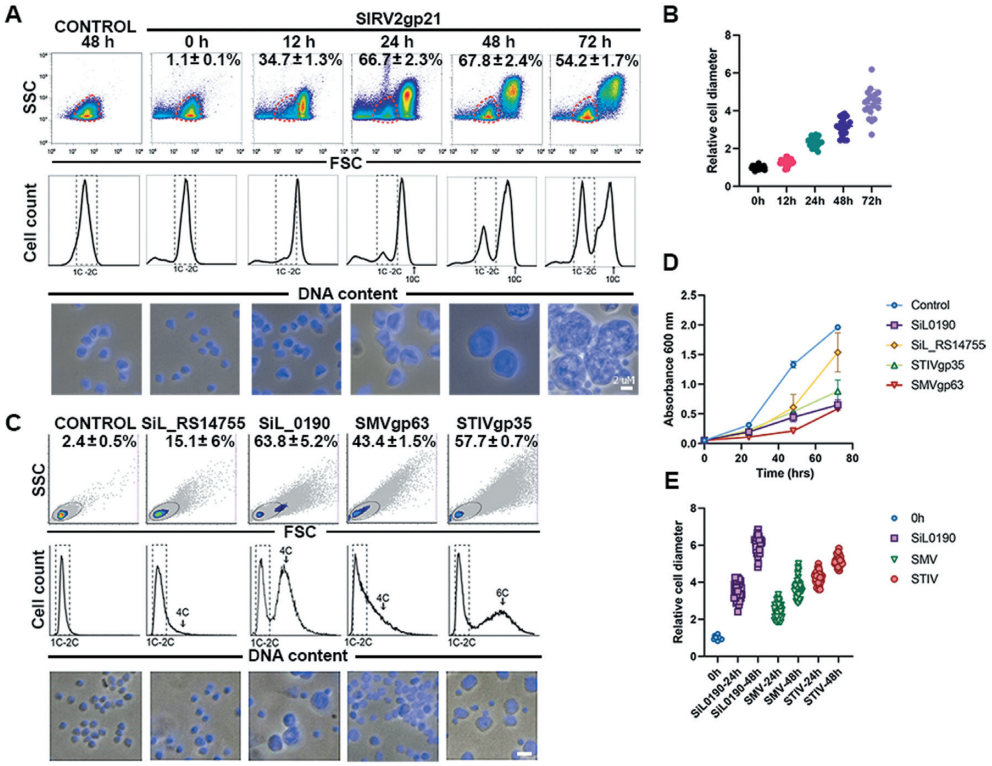

Figure 3. Members of the SiL_0190 clade induce cell enlargement and increased DNA content. (**A**) Effects of expression of SIRV2 gp21 on cell size (top panel) and DNA content (middle panel) at different time-points after induction measured by flow cytometry and microscopy (bottom panel). Bar = 2 um. The control corresponds to cells harboring an empty plasmid at 72 hpi. The percentage of large cells is indicated in the graphs of cell size (top panel). The region of interest marked as a red, dotted line in the top panels indicates normal cells. The dotted square in the middle panel indicates cells with normal DNA content corresponding to 1–2 chromosome copies. (**B**) Average cell diameter of 25 gp21-expressing cells. Bars correspond to the standard deviation of the data. (**C**) Distribution of cell size and DNA content (top and middle panels) and microscopy images (bottom panel) of cultures expressing viral and cellular homologs of the SiL_0190 clade, an empty plasmid (control) or SiL_RS14755 at 24 hpi. (**D**) Growth curves after induction with arabinose of *S. islandicus* strains harboring an empty plasmid (control), or plasmids expressing SIL_RS14755, SiL_0190, STIV gp25 or SMV gp63. The graph shows the average and the standard deviation of three biological replicates. (**E**) Average cell diameter of 50 random cells from the cultures in panel C at 24 or 48 h post-induction of protein expression.

facilitate recombination after infection (21). We were unable to obtain the recombinant virus after several attempts, suggesting that SiL_0190 cannot complement the function of SIRV2 gp21 during infection (Supplementary Figure 4B).

Members of the SiL_0190 clade of transcription regulators bind to a sequence motif AGTATTA

To determine the mechanism by which SiL_0190 homologs inhibit cell division in Sulfolobales, we investigated possible binding of SiL_0190 and SIRV2 gp21 to the promoter region of the differentially expressed genes identified in the RNA-seq experiment and to viral promoters, including the promoter of gp21. We performed competition assays where a recombinant protein was mixed with 200 bp dsDNA fragments corresponding to the labeled promoters of the target genes and different concentrations of competitor DNA. The results of PAGE-EMSA reveal three patterns: (i) no DNA binding, as evidenced by a lack of retardation in DNA migration after electrophoresis (e.g. SIRV2 gp01 promoter, Figure 4A, B, top row), (ii) unspecific, low affinity binding, quickly displaced by the addition of competitor DNA in a ratio of 1:10 (e.g. the promoters of the bacteriophage T7 polymerase, SIRV2 *gp15* or *sil_0985*, Figure 4A, B, top row) and (iii) specific binding evidenced by retardation of DNA migration at high concentrations of competitor DNA up to 1:50 (Figure 4A, B). SIRV2 gp21 and SiL_0190 bind fifteen and eleven DNA probes, respectively, in a specific manner. Promoters bound by both proteins include the regulons of the cell division machinery components *sil_1169* (*cdvA*) and *sil_1195* (*escrt-III-2*); the regulatory sequences of genes involved in DNA metabolism *sil_1121* (*topG2*), *sil_1936 (segA*) and the promoters of *sil_0086* (uncharacterized AAA-type ATPase), *sil_0336* (DUF3311), *sil_0779* (hypothetical protein), *sil_1153* (*nrdA*, ribonucleotide reductase) and *sil_2516* (*coxM*, glyceraldehyde oxidoreductase subunit). In addition, gp21, but not SiL_0190, specifically recognizes the promoters of *sil_1206* (*cdt1*), sil_0095 (HerA-like helicase) and *sil_0979* (hypothetical protein), plus the viral promoters of the middle-late genes *gp17*, *gp20-gp21* (Figure 4A, B, bottom row). These slight differences in target preference suggest that gp21 may have additional functions than the host factor that explain why SiL_0190 was unable to substitute for gp21 (Supplementary Figure 4B).

We determined the dissociation constant (K_D) of the interaction between SiL_0190 and gp21 to the *cdvA* promoter (*sil_1169*), 15.1 and 21.5 nM, respectively (Figure 4C), showing that the affinity to DNA of the two proteins

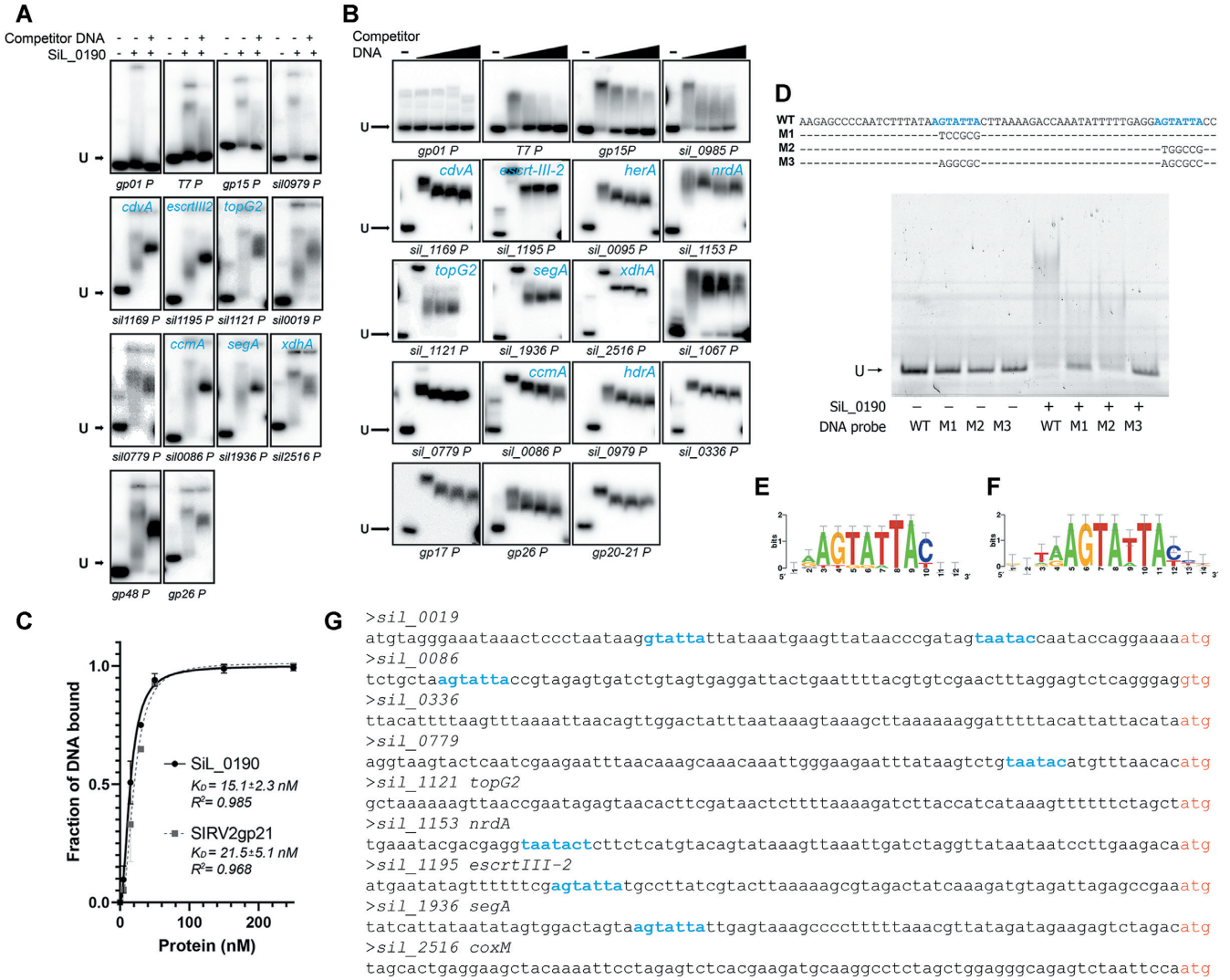

Figure 4. SiL_0190 binds to the motif in the promoters of cell division and DNA replication genes. (**A**) Representative results of mobility shift assays for SiL_0190. A total of 11.5 nM labeled probe was incubated with 100 nM protein with or without competitor DNA (ratio of 1:10). The first lane of every experiment, marked with a '-', corresponds to the labeled probe without protein nor competitor DNA. The name of the probe is indicated below each experiment. Arrows point to the unbound (U) probe. (**B**) Representative results of mobility shift assays for SIRV2 gp21. Assays were conducted as described in panel A with additional concentrations of cold competitor DNA (ratios 1:1, 1:10, 1:20 and 1:50). (**C**) Apparent binding affinities (K_D) of SiL_0190 and SIRV2 gp21. Curves represent the average of shifted *cdvA* promoter as a function of protein concentration. Average data from three independent experiments, bars indicate the standard deviation. (**D**) Top: Upstream region of the *cdvA* gene of *S. islandicus* LAL14/1 and variants with mutations of the AGTATTA motif (highlighted in blue). Bottom: EMSA assay results for the cdvA promoter variants in panel D for SiL_0190. (**E, F**) Enriched motifs identified in the SiL_0190- or the SIRV2gp21-bound promoters (E, F, respectively). (**G**) Occurrences of the DNA-binding motif AGTATTA in the target genes of SiL_0190.

is very similar. Binding to the promoter of *cdvA* is also highly specific, as the dissociation constant of SiL_0190 remains the same in the presence of excess competitor DNA (Supplementary Figure 5). To determine the binding site of SiL_0190 we used 100 bp sequences of the target promoters to predict enriched motifs with the RSAT oligo-analysis tool (van Helden *et al.*, 1998; Nguyen *et al.*, 2018). The analysis identified two motifs enriched in the SiL_0190-bound promoters with the seed sequence GAGTCT and AGTATTA (Supplementary Figure 6A–D). Mutagenesis along the promoter of *cdvA* abolished binding only when the motif AGTATTA was absent, suggesting this oligonucleotide represents the SiL_0190 binding site (Supplementary Figure 6C, D). This motif is present twice in the *cdvA*-regulatory region, between positions −10 to −2 and −40 to −46 (Figure 4D), hence we specifically mutated each copy and evaluated its contribution to the binding of both SiL_0190 and SIRV2 gp21 (Figure 4D-E). We observed impaired binding when only one copy was present and mutation of both copies was necessary to completely abrogate the binding (Figure 4E, Supplementary Figure 6A–D), demonstrating that AGTATTA is the DNA-binding motif of SiL_0190 and gp21. We additionally tested binding to the promoter of *sil_1936* that has a single occurrence of the motif, observing abrogation of DNA binding capacity after its mutation (Supplementary Figure 6E, F). The

SiL_0190-binding motif is present, however, in only five of eleven bound targets (Figure 4A, F), thus it remains to be determined which factors mediate recognition of the other promoters. On the other hand, we were able to identify the binding-motif in the 150 bp region upstream from the predicted start codon of 240 genes of *S. islandicus*, of which 33 were differentially expressed in the RNAseq experiment (Table 1).

We performed quantitative RT-qPCR to investigate whether SiL_0190 regulates expression of genes that were up- or down- regulated by SIRV2 gp21 (Figure 5A). RNA was extracted from cells overexpressing Sil_0190 and the subsequent RT-qPCR demonstrated gene regulation highly similar to what was observed for gp21-overexpressing cells (Figure 5A). Among the tested genes, *cdvA*, *escrtIII-2* and *segA* were downregulated −2.5, −2.8 and −1.1 $\log_2$ fold and the ribonucleotide reductase *nrdA* genes were upregulated 3.8 $\log_2$ fold by SiL_0190 overexpression.

To investigate if the other members of the clade share the same DNA-binding motif, we predicted enriched oligonucleotides in the promoters of the homologs of the SiL_0190 target genes in four different members of Sulfolobales (Supplementary Figure 8). A motif with the seed sequence GTATTA was identified in three of the four species and occurrences of the seed were identified in the analyzed promoters of the four organisms, in particular the promoters of *cdvA*, *escrtIII-2* and *nrdA* consistently contain the motif in all cases analyzed (Supplementary Figure 8).

Earlier transcriptome analyses showed that SIRV2 gp21 expression was detectable at 30 min post-infection and reached maximum at 1 h (61,62), and the host gp21-targets identified here followed the same regulation pattern (repression or activation) (Supplementary Figure 9). For example, *cdvA* was found to be downregulated around 2.2 $\log_2$ fold at 2 h post SIRV2 infection and the ribonucleotide reductase genes were shown to be upregulated 2.1 $\log_2$ fold at the same time point. This strongly suggests that these changes were due to the activity of gp21 and that the gene regulation following gp21 overexpression observed in this study is biologically relevant.

SiL_0190 is a cell cycle regulator that drives the transition between the D- and S- phases of the cell cycle

Previous work on *S. acidocaldarius* identified *saci_0942* (the homolog of *siL_0190* in this organism) as a putative regulator of the transition to the stationary phase (63). Additionally, Lundgren *et al.* (63) identified a subset of genes sharing a similar transcription pattern and harboring the consensus element TGTATTAT, which they named as ccr-2 (cell cycle regulon 2) box. This motif resembles the SiL_0190 binding motif identified in this work (AGTATTA) and the latter is present in the promoters of the homologs of the ccr-2 box-containing genes in *S. islandicus*. We therefore name the gp21/SiL_0190 clade as aCcr1 (archaeal cell cycle regulator 1), the first master regulator of the cell cycle characterized in Sulfolobales and rename the ccr-2 box as aCcr1-box.

It was reported that induction of aCcr1 *saci_0942* occurred during the G1/S transition (63). Interestingly, the homologs of the genes regulated by aCcr1 overexpression in *S. islandicus* LAL14/1 exhibited a consensus expression pattern in the synchronized *S. acidocaldarius* (63). That is, genes downregulated by aCcr1 shown in this work (*cdvA, escrtIII-2, segA, topR2*) were induced in an earlier stage (D- or D/G1) than *aCcr1* whereas those upregulated by aCcr1 (*saci_1107, saci_2209*, see Table 1) were induced at a later stage in the cell cycle (S-phase) (Figure 5B, Table 1). This strongly suggests that aCcr1 homologs regulate cell cycle progression in Sulfolobales by inhibiting the transcription of cell division genes and other D/G1-phase genes, and by inducing transcription of S/G2-phase genes (Figure 5C).

DISCUSSION

The pioneer work by Lundgren *et al.* first described cell-cycle specific transcription of over 160 genes in the thermoproteon *S. acidocaldarius* (63), pointing to transcriptional regulation as a mechanism for the control of cell cycle progression in Sulfolobales. However, no cell-cycle regulatory factor has been described so far for Thermoproteota. In this work, we identify and characterize aCcr1, a conserved, essential protein in Sulfolobales that functions as a regulator of cell cycle progression. aCcr1 downregulates the transcription of important components of the Cdv (cell division) machinery CdvA and CdvB2, SegA, a component of the chromosome segregation system (64) and the reverse gyrase TopR2 that regulates DNA supercoiling in crenarchaeota (65). It additionally upregulates the transcription of the ribonucleotide reductase NrdA that converts ribonucleotides into deoxyribonucleotides, an essential step for DNA synthesis, and CoxM, a component of the CO dehydrogenase important for CO oxidation in Sulfolobales (66). aCcr1 binds to the motif AGTATTA (*ccr2*-box, Figure 4) that is present in the promoter region of several of its target genes. aCcr1 is highly conserved in Sulfolobales, with amino acid identities between 81% and 100% (Supplementary Figure 2A) similarly to its target genes, which all have homologs in all Sulfolobales and whose promoters also contain the motif AGTATTA (Supplementary Figure 7).

Our knowledge about cell cycle control in archaea is still very limited. The cell cycle transcription regulator CdrS (cell division regulator short) present in the Halobacteriota phylum, was recently identified and characterized (15,16). CdrS harbors an RHH domain that binds to promoters, and activates the transcription of 18 genes, including four involved in cell division (*ftsZ1, ftsZ2*), DNA replication and partitioning (*sepF* and *smc*) and several genes involved in metabolism (15,16). While only two cell cycle regulators have been described in the archaea domain (CdrS and aCcr1), it is interesting to note that both contain the RHH structural motif, whereas the majority of the transcription factors encoded in archaeal genomes harbor helix-turn-helix (HTH) motifs (67,68). On the other hand, RHH transcription factors are the most abundant in archaeal viruses (18). The common use of RHH transcription factors despite the fact that cell division machineries between the Halobacteriota and the Thermoproteota are completely different (the tubulin FtsZ system and the ESCRT-based machinery, respectively (reviewed by (9,10)), suggests a common theme in cell cycle control in the Archaea domain. RHH-domain proteins have so far been described solely in bacteria and archaea (17). Intriguingly, while transcriptional control ap-

Table 1. *S. islandicus* LAL14/1 promoters containing the ccr-2 box and transcription profile of their homologs in *S. acidocaldarius*

SisLAL14 gene name	Saci gene name	Annotation	ccr-2 box	Transcription pattern Saci*	Cell cycle phase	RNAseq profile +
SiL_0086#	Saci_0157	ATPase, predicted component of phage defense system	Y	cyclic		down
SiL_0094-SiL_0095	Saci_0148	NurA-like nuclease I HerA-like helicase	Y	N		down
SiL_0286	Saci_1107	AcrR TetR family transcriptional regulator	Y	cyclic	G1/S	up
SiL_0287	Saci_2219	Sterol carrier protein	Y	cyclic	S	up
SiL_0345	Saci_2245	Predicted metal-dependent hydrolase of the TIM-barrel fold	Y	N		up
SiL_0416	n.a.	PaaD Metal-sulfur cluster biosynthetic enzyme	Y	n.a.		up
SiL_0486	n.a.	NADH-FMN oxidoreductase RutF, flavin reductase (DIM6/NTAB) family	Y	n.a.		up
SiL_0572	Saci_2042	HyuB N-methylhydantoinase B/acetone carboxylase, alpha subunit	Y	gradual		down
SiL_0593	n.a.	Metal-dependent hydrolase of the beta-lactamase superfamily II	Y	n.a.		up
SiL_0612	n.a.	CRISPR-associated protein, Cas1 family	Y	n.a.		up
SiL_0619–0620	n.a.	VapBC type II toxin-antitoxin	Y	n.a.		up
SiL_0631–0632	n.a.	RHH/CopG DNA binding protein| PIN domain containing protein	Y	n.a.		up
SiL_0728	n.a.	Hypothetical	Y	n.a.		up
SiL_0841	n.a.	SAM-dependent methyltransferase	Y	n.a.		down
SiL_0934	Saci_1846	Thermopsin-like protease	Y	N		down
SiL_0980	Saci_0329	HdrB, heterodisulfide reductase, subunit B	Y	N		down
SiL_1018	Saci_0884	ApaH Serine/threonine protein phosphatase PP2A family	Y	N		down
SiL_1067	n.a.	Hypothetical	Y	n.a.		up
SiL_1153#	Saci_1353	NrdA, ribonuclease reductase	Y	cyclic	G1/S	up
SiL_1169#	Saci_1374	cdvA	Y	cyclic	D	down
SiL_1195#	Saci_1416	ESCRT-III-2	Y	cyclic		down
SiL_1597	Saci_0814	Lhr-like helicase	Y	N		down
SiL_1614	Saci_2356	Hypothetical	Y	N		down
SiL_1813	Saci_0773	SpoVK, AAA ATPase	Y	N		down
SiL_1935–1936#	Saci_0203–0204	SegAB	Y	cyclic	D/G1	down
SiL_2120	n.a.	DNA-binding beta-propeller fold protein YncE	Y	n.a.		up
SiL_2124–2127	n.a.	TusA-related sulfurtransferase, hypothetical, hypothetical, Predicted peroxiredoxin	Y	n.a.		up
SiL_2184	n.a.	Alpha-D-xyloside xylohydrolase	Y	n.a.		down
SiL_2234	Saci_1754	Hypothetical	Y	N		up
SiL_2267–69	Saci_2201–2203	MET3 sulfate adenylyltransferase, CysH phosphoadenylylsulfate reductase (thioredoxin), CysI sulfite reductase (NADPH) hemoprotein beta-component	Y	cyclic		up
SiL_2284	Saci_1122	CaiC Acyl-CoA synthetase (AMP-forming)/AMP-acid ligase II	Y	Y		up
SiL_2338	n.a.	Hypothetical	Y	n.a.		up
SiL_2438	Saci_1166	MalK	Y	gradual		up
SiL_0019#	n.a.	Queuine tRNA-ribosyltransferase related protein	N	n.a.		down
SiL_0336#	Saci_1744	Protein of unknown function (DUF3311)	N	N		up
SiL_0779#	Saci_0236	Hypothetical	N	N		up
SiL_1121#	topG2	topR2	N	cyclic	G1/S	down
SiL_2516#	Saci_2269	CoxM, Aerobic-type carbon monoxide dehydrogenase, middle subunit CoxM/CutM homolog	N	gradual		up

* Data from Lundgren *et al.* (2007).
n.a., not applicable; Y, yes; N, no.
Binding to SiL_0190 in EMSA assays.
+ Transcription pattern in GP21-expressing cultures (Supplementary Table 3); up, upregulated; down, downregulated.

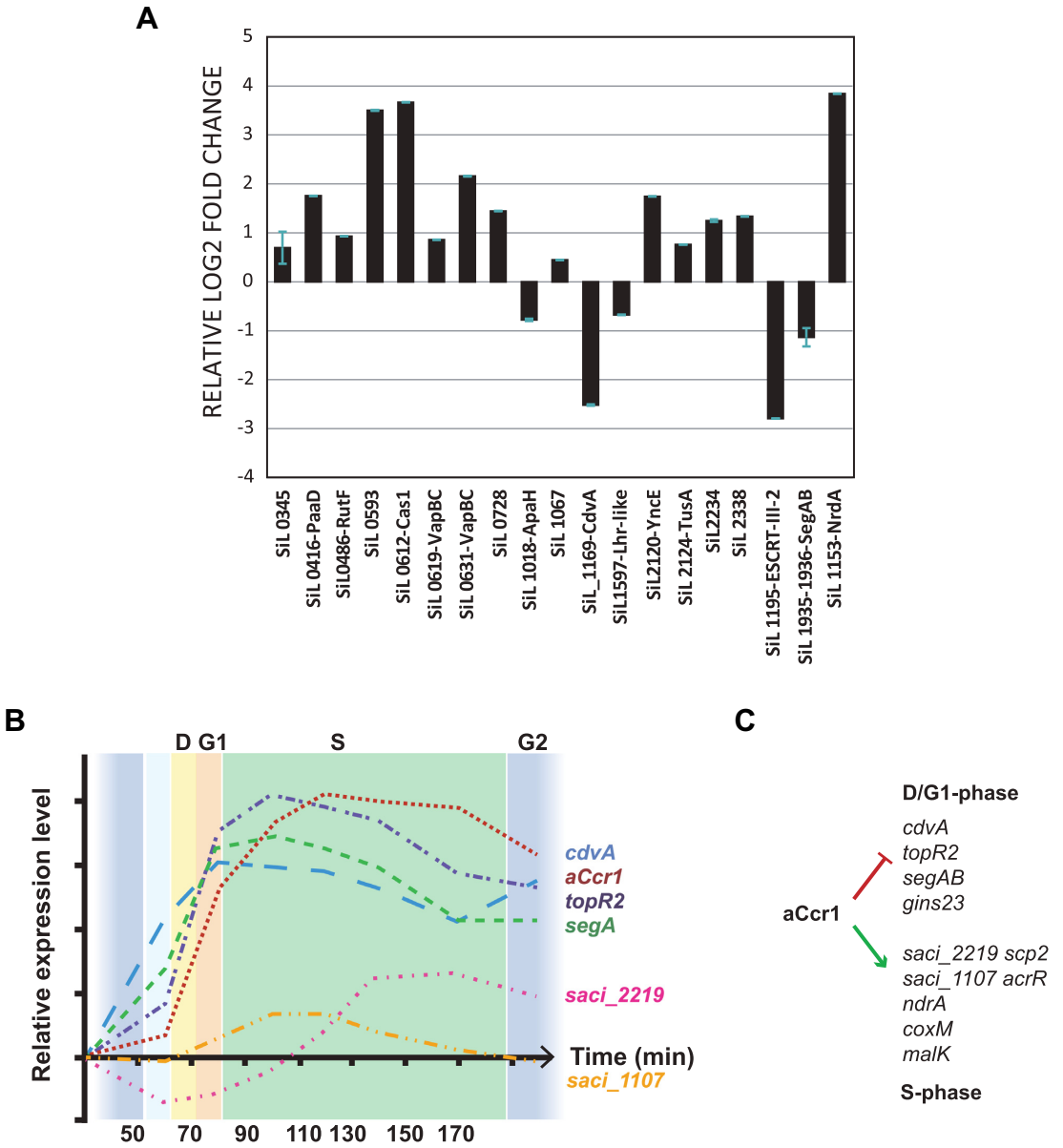

Figure 5. Expression of aCcr1 target genes after the induction of aCcr1 and model of the aCcr1-mediated regulation of cell cycle progression in the Sulfolobales. (**A**) *S. islandicus* □arrays cultures harboring a plasmid for the inducible expression of SiL_0190 were sampled and analyzed by RT-qPCR at 6 h post-induction with arabinose. TBP was used as the normalization reference. The graph shows the average and standard deviation of three biological replicates. (**B**) Expression profiles of cyclic genes in *S. acidocaldarius* as determined by Lundgren *et al.* (63). The homologs in *S. islandicus* LAL14/1 of the genes shown in the graph contain the aCcr1-box in their promoter region and are recognized by SiL_0190 and SIRV2 gp21. (**C**) Model of the effect of aCcr1 expression in *S. acidocaldarius* over the transcription of its target genes. Induction of aCcr1 correlates with the decrease or plateau of genes induced earlier during the D- and G1 phases (e.g. *cdvA, segA, topR2, gins23*) and with the induction of S-phase genes (e.g. *scp2, acrR*).

pears as a shared mechanism of cell cycle regulation for both bacteria and eukaryotes, to our knowledge no transcription factor of the RHH family has been implicated in cell cycle control in bacteria. Additional work is required to elucidate the role of RHH-domain transcription factors in regulating cell cycle progression in the Archaea domain.

There is only one previous report of an archaeal virus family disrupting the cell cycle of its host: infection with viruses of the *Bicaudaviridae* family STSV2 and SMV1 arrest cells in the S-phase and triggers transcriptional repression of components of the ESCRT cell division machinery, resulting in cell enlargement and an increase in the number of chromosomes per cell (5). Interestingly, whereas some bicaudaviruses, namely ATV, ATV2 and SMV3, encode homologs of aCcr1, STSV2 contains several divergent members of the gp21 family that do not belong to the aCcr1 clade, hence it remains to be determined the mechanism used by STSV2 to regulate the cell cycle. Whereas STSV2-infection causes marked cell enlargement, this is not observed during SIRV2-infection. This is due to SIRV2 being a lytic virus having and infection cycle that lasts around 8 hours (69).

We predict the existence of several regulatory factors whose successive activation drives the transcription of different sets of genes during cell cycle progression. A similar mechanism has been described in the alphaproteobacteria, where a circuit of antagonistic regulators act together to control transcription across the different phases of the cell cycle (reviewed in (70)). In eukaryotes, transcriptional control of the cell cycle is linked to the activity of cyclins and cyclin dependent kinases, which, among other changes, regulate the activity of the three RNA polymerases (71)).

Regulation of cell growth, cytokinesis and DNA replication and segregation is exquisitely and robustly coupled in the three domains of life. These processes are linked together by several layers of regulation that involve different mechanisms and control strategies, and we envisage a similarly complex picture for the archaea domain. For example, post-translational control of the cell division machinery by the proteasome was reported in *S. acidocaldarius*, where the 20S proteasome degrades CdvB and triggers cell division by allowing the constriction of the CdvB1:CdvB2 division ring (72). Post-translational mechanisms could be involved in regulation of aCcr1, since plasmid-driven overexpression of aCcr1 homologs in this work constantly resulted in low protein yield, despite strong transcription from the arabinose promoter.

The widespread presence of aCcr1 homologs in crenarchaeal viruses is striking, as it encompasses six different families of viruses that have diverse hosts and infection modes, including both lytic and lysogenic viruses (Figure 1). This highlights the importance of this protein for successful viral infection. The pressure to maintain sequence and functional conservation of aCcr1 homologs is remarkable, as exemplified by the high identity (over 40%) at the amino acid level of the viral aCcr1 homologs and the fact that no other protein is shared among so many archaeal viral taxons. We hypothesize that SIRV2 and other viruses benefit from regulating cell cycle by driving the host cell into a state where components important for viral replication are present. For example, the ribonucleotide reductase NrdA (Sil_1153), one of the highly activated genes following SIRV2 infection, has been proposed to be important for viral replication by converting ribonucleotides into deoxyribonucleotides for DNA replication and repair (61).

There are multiple reports of eukaryotic viruses regulating cell cycle during infection through different mechanisms (reviewed by (1)) and a few bacterial viruses that mostly disrupt proper functioning of the cell division protein FtsZ (3,73–76). The only report of an archaeal virus family disrupting the cell cycle of its host is the *Bicaudaviridae* family, including STSV2 and SMV1. Infection by these viruses arrests cells in the S-phase and triggers transcriptional repression of components of the ESCRT cell division machinery, resulting in cell enlargement and an increase in the number of chromosomes per cell (5). Interestingly, some bicaudaviruses, namely ATV, ATV2 and SMV3, encode homologs of aCcr1 while STSV2 contains divergent homologs clustering in other branches of the phylogenetic tree of viral RHH proteins (Figure 1), thus, not belonging to the aCcr1 clade. Hence it remains to be determined if the RHH-domain factors in STSV2 are involved in cell cycle control.

In this work, we started by investigating widespread viral transcription regulators of which some, including SIRV2 gp21 characterized here, are homologous and functionally similar to the global cell cycle regulator aCcr1 of the Sulfolobales. We have functionally characterized the clade of aCcr1 as a transcription regulator that drives progression into the S-phase of the cell cycle. Our preliminary data suggests that other branches of the family may perform novel, different functions in Sulfolobales, and future work will shed light on the role of these cellular and viral clades of the family.

This work is published together with the study by Yang *et al.* (77) describing SiRe_0197, homolog of SiL_0190, as aCcr1 of *S. islandicus* REY15A. They identified and characterized aCcr1 of *S. islandicus* REY15A by investigating transcription regulators with cyclic transcriptional patterns and obtained similar results to ours. Additionally, they provide ChIP-seq data confirming that GTATTA is the seed of the binding motif and that *cdvA* is an in vivo target of aCcr1. Together, both studies provide insight into cell cycle control in the archaea and about viral strategies to subvert host cellular processes.

DATA AVAILABILITY

The RNAseq data underlying this article is deposited in the GEO repository under the series record GSE21874.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We acknowledge de Biocomputing Core Facility of the Department of Biology, University of Copenhagen for access to resources used for bioinformatic analyses.
Author contribution: X.L., L.M.-A. and X.P. conceived the experiments. X.L., C.L.-M. and L.M.-A. performed the experiments. X.L., L.M.-A. and X.P. wrote the paper. All authors revised and approved the manuscript.

FUNDING

Independent Research Fund Denmark/Natural Sciences [DFF-0135-00402]; Novo Nordisk Foundation/Hallas-Moeller Ascending Investigator Grant [NNF17OC0031154 to X.P.]. Funding for open access charge: DFF; Novo Nordisk.
Conflict of interest statement. None declared.

This paper is linked to: doi:10.1093/nar/gkad006.

REFERENCES

1. Fan,Y., Sanyal,S. and Bruzzone,R. (2018) Breaking bad: how viruses subvert the cell cycle. *Front. Cell Infect. Microbiol.*, **8**, 396.
2. Haeusser,D.P., Hoashi,M., Weaver,A., Brown,N., Pan,J., Sawitzke,J.A., Thomason,L.C., Court,D.L. and Margolin,W. (2014) The kil peptide of bacteriophage λ blocks Escherichia coli cytokinesis via ZipA-dependent inhibition of FtsZ assembly. *PLoS Genet.*, **10**, e1004217.

3. Kiro,R., Molshanski-Mor,S., Yosef,I., Milam,S.L., Erickson,H.P. and Qimron,U. (2013) Gene product 0.4 increases bacteriophage T7 competitiveness by inhibiting host cell division. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 19549–19554.
4. Stewart,C.R., Deery,W.J., Egan,E.S.K., Myles,B. and Petti,A.A. (2013) The product of SPO1 gene 56 inhibits host cell division during infection of Bacillus subtilis by bacteriophage SPO1. *Virology*, **447**, 249–253.
5. Liu,J., Cvirkaite-Krupovic,V., Baquero,D.P., Yang,Y., Zhang,Q., Shen,Y. and Krupovic,M. (2021) Virus-induced cell gigantism and asymmetric cell division in archaea. *Proc. Natl. Acad. Sci. U.S.A.*, **118**, e2022578118.
6. Rinke,C., Chuvochina,M., Mussig,A.J., Chaumeil,P.-A., Davín,A.A., Waite,D.W., Whitman,W.B., Parks,D.H. and Hugenholtz,P. (2021) A standardized archaeal taxonomy for the genome taxonomy database. *Nat. Microbiol.*, **6**, 946–959.
7. Lundgren,M. and Bernander,R. (2005) Archaeal cell cycle progress. *Curr. Opin. Microbiol.*, **8**, 662–668.
8. Lundgren,M., Malandrin,L., Eriksson,S., Huber,H. and Bernander,R. (2008) Cell cycle characteristics of crenarchaeota: unity among diversity. *J. Bacteriol.*, **190**, 5362–5367.
9. Caspi,Y. and Dekker,C. (2018) Dividing the archaeal way: the ancient cdv cell-division machinery. *Front. Microbiol.*, **9**, 174.
10. Ithurbide,S., Gribaldo,S., Albers,S.-V. and Pende,N. (2022) Spotlight on FtsZ-based cell division in archaea. *Trends Microbiol.*, **30**, 665–678.
11. Bähler,J. (2005) Cell-cycle control of gene expression in budding and fission yeast. *Annu. Rev. Genet.*, **39**, 69–94.
12. Galderisi,U., Jori,F.P. and Giordano,A. (2003) Cell cycle regulation and neural differentiation. *Oncogene*, **22**, 5208–5219.
13. Haase,S.B. and Wittenberg,C. (2014) Topology and control of the cell-cycle-regulated transcriptional circuitry. *Genetics*, **196**, 65–90.
14. Francis,N., Poncin,K., Fioravanti,A., Vassen,V., Willemart,K., Ong,T.A.P., Rappez,L., Letesson,J.-J., Biondi,E.G. and De Bolle,X. (2017) CtrA controls cell division and outer membrane composition of the pathogen *Brucella abortus*. *Mol. Microbiol.*, **103**, 780–797.
15. Darnell,C.L., Zheng,J., Wilson,S., Bertoli,R.M., Bisson-Filho,A.W., Garner,E.C. and Schmid,A.K. (2020) The ribbon-helix-helix domain protein CdrS regulates the tubulin homolog ftsZ2 to control cell division in archaea. *Mbio*, **11**, e01007-20.
16. Liao,Y., Vogel,V., Hauber,S., Bartel,J., Alkhnbashi,O.S., Maaß,S., Schwarz,T.S., Backofen,R., Becher,D., Duggin,I.G. *et al.* (2021) CdrS is a global transcriptional regulator influencing cell division in haloferax volcanii. *Mbio*, **12**, e0141621.
17. Schreiter,E.R. and Drennan,C.L. (2007) Ribbon-helix-helix transcription factors: variations on a theme. *Nat. Rev. Microbiol.*, **5**, 710–720.
18. Iranzo,J., Koonin,E.V., Prangishvili,D. and Krupovic,M. (2016) Bipartite network analysis of the archaeal virosphere: evolutionary connections between viruses and capsidless mobile elements. *J. Virol.*, **90**, 11043–11055.
19. He,F., Vestergaard,G., Peng,W., She,Q. and Peng,X. (2016) CRISPR-cas type I-A cascade complex couples viral infection surveillance to host transcriptional regulation in the dependence of Csa3b. *Nucleic Acids Res.*, **45**, 1902–1913.
20. Zillig,W., Kletzin,A., Schleper,C., Holz,I., Janekovic,D., Hain,J., Lanzendörfer,M. and Kristjansson,J.K. (1993) Screening for sulfolobales, their plasmids and their viruses in icelandic solfataras. *Syst. Appl. Microbiol.*, **16**, 609–628.
21. Alfastsen,L., Peng,X. and Bhoobalan-Chitty,Y. (2021) Genome editing in archaeal viruses and endogenous viral protein purification. *STAR Protoc.*, **2**, 100791.
22. Peng,N., Deng,L., Mei,Y., Jiang,D., Hu,Y., Awayez,M., Liang,Y. and She,Q. (2012) A synthetic arabinose-inducible promoter confers high levels of recombinant protein expression in hyperthermophilic archaeon Sulfolobus islandicus. *Appl. Environ. Microbiol.*, **78**, 5630–5637.
23. Li,Y., Pan,S., Zhang,Y., Ren,M., Feng,M., Peng,N., Chen,L., Liang,Y.X. and She,Q. (2016) Harnessing type I and type III CRISPR-cas systems for genome editing. *Nucleic Acids Res.*, **44**, e34.
24. Mayo-Muñoz,D., He,F., Jørgensen,J.B., Madsen,P.K., Bhoobalan-Chitty,Y. and Peng,X. (2018) Anti-CRISPR-based and CRISPR-based genome editing of sulfolobus islandicus rod-shaped virus 2. *Viruses*, **10**, E695.
25. Deng,L., Zhu,H., Chen,Z., Liang,Y.X. and She,Q. (2009) Unmarked gene deletion and host-vector system for the hyperthermophilic crenarchaeon Sulfolobus islandicus. *Extremophiles*, **13**, 735–746.
26. Langmead,B. and Salzberg,S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Methods*, **9**, 357–359.
27. Liao,Y., Smyth,G.K. and Shi,W. (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, **30**, 923–930.
28. Love,M.I., Huber,W. and Anders,S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550.
29. Robinson,M.D., McCarthy,D.J. and Smyth,G.K. (2010) edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, **26**, 139–140.
30. Chen,I.-M.A., Markowitz,V.M., Chu,K., Palaniappan,K., Szeto,E., Pillay,M., Ratner,A., Huang,J., Andersen,E., Huntemann,M. *et al.* (2017) IMG/M: integrated genome and metagenome comparative data analysis system. *Nucleic Acids Res.*, **45**, D507–D516.
31. Aramaki,T., Blanc-Mathieu,R., Endo,H., Ohkubo,K., Kanehisa,M., Goto,S. and Ogata,H. (2020) KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics*, **36**, 2251–2252.
32. Marchler-Bauer,A., Bo,Y., Han,L., He,J., Lanczycki,C.J., Lu,S., Chitsaz,F., Derbyshire,M.K., Geer,R.C., Gonzales,N.R. *et al.* (2017) CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res.*, **45**, D200–D203.
33. Boratyn,G.M., Schäffer,A.A., Agarwala,R., Altschul,S.F., Lipman,D.J. and Madden,T.L. (2012) Domain enhanced lookup time accelerated BLAST. *Biol. Direct*, **7**, 12.
34. Emms,D.M. and Kelly,S. (2019) OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.*, **20**, 238.
35. Guillière,F., Peixeiro,N., Kessler,A., Raynal,B., Desnoues,N., Keller,J., Delepierre,M., Prangishvili,D., Sezonov,G. and Guijarro,J.I. (2009) Structure, function, and targets of the transcriptional regulator SvtR from the hyperthermophilic archaeal virus SIRV1. *J. Biol. Chem.*, **284**, 22222–22237.
36. van Helden,J., André,B. and Collado-Vides,J. (1998) Extracting regulatory sites from the upstream region of yeast genes by computational analysis of oligonucleotide frequencies. *J. Mol. Biol.*, **281**, 827–842.
37. Nguyen,N.T.T., Contreras-Moreira,B., Castro-Mondragon,J.A., Santana-Garcia,W., Ossio,R., Robles-Espinoza,C.D., Bahin,M., Collombet,S., Vincens,P., Thieffry,D. *et al.* (2018) RSAT 2018: regulatory sequence analysis tools 20th anniversary. *Nucleic Acids Res.*, **46**, W209–W214.
38. Bernander,R. and Poplawski,A. (1997) Cell cycle characteristics of thermophilic archaea. *J. Bacteriol.*, **179**, 4963–4969.
39. Martínez-Alvarez,L., Deng,L. and Peng,X. (2017) Formation of a viral replication focus in *Sulfolobus* cells infected by the rudivirus SIRV2. *J. Virol.*, **91**, e00486-17.
40. Schneider,C.A., Rasband,W.S. and Eliceiri,K.W. (2012) NIH image to ImageJ: 25 years of image analysis. *Nat. Methods*, **9**, 671–675.
41. Krupovic,M., Makarova,K.S., Wolf,Y.I., Medvedeva,S., Prangishvili,D., Forterre,P. and Koonin,E.V. (2019) Integrated mobile genetic elements in thaumarchaeota. *Environ. Microbiol.*, **21**, 2056–2078.
42. Eddy,S.R. (2011) Accelerated profile HMM searches. *PLoS Comput. Biol.*, **7**, e1002195.
43. Grazziotin,A.L., Koonin,E.V. and Kristensen,D.M. (2017) Prokaryotic virus orthologous groups (pVOGs): a resource for comparative genomics and protein family annotation. *Nucleic Acids Res.*, **45**, D491–D498.
44. El-Gebali,S., Mistry,J., Bateman,A., Eddy,S.R., Luciani,A., Potter,S.C., Qureshi,M., Richardson,L.J., Salazar,G.A., Smart,A. *et al.* (2019) The Pfam protein families database in 2019. *Nucleic Acids Res.*, **47**, D427–D432.
45. Altschul,S.F., Madden,T.L., Schäffer,A.A., Zhang,J., Zhang,Z., Miller,W. and Lipman,D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, **25**, 3389–3402.
46. Oberto,J. (2013) SyntTax: a web server linking synteny to prokaryotic taxonomy. *BMC Bioinformatics*, **14**, 4.
47. Edgar,R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.*, **32**, 1792–1797.

48. Price,M.N., Dehal,P.S. and Arkin,A.P. (2010) FastTree 2 – approximately maximum-likelihood trees for large alignments. *PLoS One*, **5**, e9490.
49. Jumper,J., Evans,R., Pritzel,A., Green,T., Figurnov,M., Ronneberger,O., Tunyasuvunakool,K., Bates,R., Žídek,A., Potapenko,A. *et al.* (2021) Highly accurate protein structure prediction with AlphaFold. *Nature*, **596**, 583–589.
50. Ni,L., Jensen,S.O., Ky Tonthat,N., Berg,T., Kwong,S.M., Guan,F.H.X., Brown,M.H., Skurray,R.A., Firth,N. and Schumacher,M.A. (2009) The Staphylococcus aureus pSK41 plasmid-encoded ArtA protein is a master regulator of plasmid transmission genes and contains a RHH motif used in alternate DNA-binding modes. *Nucleic Acids Res.*, **37**, 6970–6983.
51. Schlenker,C., Goel,A., Tripet,B.P., Menon,S., Willi,T., Dlakić,M., Young,M.J., Lawrence,C.M. and Copié,V. (2012) Structural studies of E73 from a hyperthermophilic archaeal virus identify the 'RH3' domain, an elaborated ribbon-helix-helix motif involved in DNA recognition. *Biochemistry*, **51**, 2899–2910.
52. Haft,D.H., DiCuccio,M., Badretdin,A., Brover,V., Chetvernin,V., O'Neill,K., Li,W., Chitsaz,F., Derbyshire,M.K., Gonzales,N.R. *et al.* (2018) RefSeq: an update on prokaryotic genome annotation and curation. *Nucleic Acids Res.*, **46**, D851–D860.
53. Zhang,C., Phillips,A.P.R., Wipfler,R.L., Olsen,G.J. and Whitaker,R.J. (2018) The essential genome of the crenarchaeal model *Sulfolobus islandicus*. *Nat. Commun.*, **9**, 4908.
54. Liu,J., Gao,R., Li,C., Ni,J., Yang,Z., Zhang,Q., Chen,H. and Shen,Y. (2017) Functional assignment of multiple ESCRT-III homologs in cell division and budding in Sulfolobus islandicus. *Mol. Microbiol.*, **105**, 540–553.
55. Yen,C.-Y., Lin,M.-G., Chen,B.-W., Ng,I.W., Read,N., Kabli,A.F., Wu,C.-T., Shen,Y.-Y., Chen,C.-H., Barillà,D. *et al.* (2021) Chromosome segregation in Archaea: segA- and SegB-DNA complex structures provide insights into segrosome assembly. *Nucleic Acids Res.*, **49**, 13150–13164.
56. Couturier,M., Gadelle,D., Forterre,P., Nadal,M. and Garnier,F. (2020) The reverse gyrase TopR1 is responsible for the homeostatic control of DNA supercoiling in the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Mol. Microbiol.*, **113**, 356–368.
57. Lang,S. and Huang,L. (2015) The Sulfolobus solfataricus GINS complex stimulates DNA binding and processive DNA unwinding by minichromosome maintenance helicase. *J. Bacteriol.*, **197**, 3409–3420.
58. Samson,R.Y., Xu,Y., Gadelha,C., Stone,T.A., Faqiri,J.N., Li,D., Qin,N., Pu,F., Liang,Y.X., She,Q. *et al.* (2013) Specificity and function of archaeal DNA replication initiator proteins. *Cell Rep.*, **3**, 485–496.
59. Samson,R.Y., Obita,T., Freund,S.M., Williams,R.L. and Bell,S.D. (2008) A role for the ESCRT system in cell division in archaea. *Science*, **322**, 1710–1713.
60. Yang,N. and Driessen,A.J.M. (2015) The saci_2123 gene of the hyperthermoacidophile Sulfolobus acidocaldarius encodes an ATP-binding cassette multidrug transporter. *Extremophiles*, **19**, 101–108.
61. Okutan,E., Deng,L., Mirlashari,S., Uldahl,K., Halim,M., Liu,C., Garrett,R.A., She,Q. and Peng,X. (2013) Novel insights into gene regulation of the rudivirus SIRV2 infecting Sulfolobus cells. *RNA Biol.*, **10**, 875–885.
62. Quax,T.E.F., Voet,M., Sismeiro,O., Dillies,M.-A., Jagla,B., Coppée,J.-Y., Sezonov,G., Forterre,P., van der Oost,J., Lavigne,R. *et al.* (2013) Massive activation of archaeal defense genes during viral infection. *J. Virol.*, **87**, 8419–8428.
63. Lundgren,M. and Bernander,R. (2007) Genome-wide transcription map of an archaeal cell cycle. *Proc. Natl. Acad. Sci. U.S.A.*, **104**, 2939–2944.
64. Kalliomaa-Sanford,A.K., Rodriguez-Castañeda,F.A., McLeod,B.N., Latorre-Roselló,V., Smith,J.H., Reimann,J., Albers,S.V. and Barillà,D. (2012) Chromosome segregation in Archaea mediated by a hybrid DNA partition machine. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 3754–3759.
65. Garnier,F., Couturier,M., Débat,H. and Nadal,M. (2021) Archaea: a gold mine for topoisomerase diversity. *Front. Microbiol.*, **12**, 661411.
66. Sokolova,T.G., Yakimov,M.M., Chernyh,N.A., Lun'kova,E.Yu., Kostrikina,N.A., Taranov,E.A., Lebedinskii,A.V. and Bonch-Osmolovskaya,E.A. (2017) Aerobic carbon monoxide oxidation in the course of growth of a hyperthermophilic archaeon, sulfolobus sp. ETSY. *Microbiology*, **86**, 539–548.
67. Flores-Bautista,E., Hernandez-Guerrero,R., Huerta-Saquero,A., Tenorio-Salgado,S., Rivera-Gomez,N., Romero,A., Ibarra,J.A. and Perez-Rueda,E. (2020) Deciphering the functional diversity of DNA-binding transcription factors in Bacteria and Archaea organisms. *PLoS One*, **15**, e0237135.
68. Lemmens,L., Maklad,H.R., Bervoets,I. and Peeters,E. (2019) Transcription regulators in Archaea: homologies and differences with bacterial regulators. *J. Mol. Biol.*, **431**, 4132–4146.
69. Bize,A., Karlsson,E.A., Ekefjärd,K., Quax,T.E.F., Pina,M., Prevost,M.-C., Forterre,P., Tenaillon,O., Bernander,R. and Prangishvili,D. (2009) A unique virus release mechanism in the Archaea. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 11306–11311.
70. Panis,G., Murray,S.R. and Viollier,P.H. (2015) Versatility of global transcriptional regulators in alpha-proteobacteria: from essential cell cycle control to ancillary functions. *FEMS Microbiol. Rev.*, **39**, 120–133.
71. Delgado-Román,I. and Muñoz-Centeno,M.C. (2021) Coupling between cell cycle progression and the nuclear RNA polymerases system. *Front. Mol. Biosci.*, **8**, 691636.
72. Tarrason Risa,G., Hurtig,F., Bray,S., Hafner,A.E., Harker-Kirschneck,L., Faull,P., Davis,C., Papatziamou,D., Mutavchiev,D.R., Fan,C. *et al.* (2020) The proteasome controls ESCRT-III-mediated cell division in an archaeon. *Science*, **369**, eaaz2532.
73. Bhambhani,A., Iadicicco,I., Lee,J., Ahmed,S., Belfatto,M., Held,D., Marconi,A., Parks,A., Stewart,C.R., Margolin,W. *et al.* (2020) Bacteriophage SP01 gene product 56 inhibits Bacillus subtilis cell division by interacting with FtsL and disrupting Pbp2B and FtsW recruitment. *J. Bacteriol.*, **203**, e00463-20.
74. Faubladier,M. and Bouché,J.P. (1994) Division inhibition gene dicF of Escherichia coli reveals a widespread group of prophage sequences in bacterial genomes. *J. Bacteriol.*, **176**, 1150–1156.
75. Hernández-Rocamora,V.M., Alfonso,C., Margolin,W., Zorrilla,S. and Rivas,G. (2015) Evidence that bacteriophage λ kil peptide inhibits bacterial cell division by disrupting FtsZ protofilaments and sequestering protein subunits. *J. Biol. Chem.*, **290**, 20325–20335.
76. Ragunathan,P.T. and Vanderpool,C.K. (2019) Cryptic-prophage-encoded small protein DicB protects Escherichia coli from phage infection by inhibiting inner membrane receptor proteins. *J. Bacteriol.*, **201**, e00475-19.
77. Yang,Y., Liu,J., Fu,X., Zhou,F., Zhang,S., Zhang,X., Huang,Q., Krupovic,M., She,Q., Ni,J. *et al.* (2022) A novel RHH family transcription factor aCcr1 and its viral homologs dictate cell cycle progression in archaea. bioRxiv doi: https://doi.org/10.1101/2022.07.07.499082, 22 November 2022, preprint: not peer reviewed.