Supplementary information for

TrmB1, a novel nucleoid-associated protein orchestrates dynamic chromatin organization and cell cycle progression in archaea

Qi Gan[1], Haodun Li[1], Qihong Huang*[1], Yunfeng Yang[1], Fei Sun[4], Xu Feng[1], Jinfeng Ni[1], Zhenfeng Zhang[2, 3], Qunxin She[1], Yulong Shen[1]*

[1]CRISPR and Archaea Biology Research Center, State Key Laboratory of Microbial Technology, Microbial Technology Institute, Shandong University, 266237, Qingdao, China.

[2]State Key Laboratory of Microbial Resources, Institute of Microbiology, Chinese Academy of Sciences (CAS), Beijing, China.

[3]College of Life Sciences, University of Chinese Academy of Sciences, Beijing, China.

[4]School of Life Science and Technology, Southeast University, Nanjing, China.

*Correspondence: yulgshen@sdu.edu.cn; huangqihong@sdu.edu.cn

Extended Data Figures Legends

Extended Data Fig. 1 SisTrmB1 is conserved in Sulfolobales.

a, Phylogenetic tree of SisTrmB1 homologs based on a Clustal Omega alignment. The tree was generated by Maximum Likelihood method with the LG+G model by MEGA 11. The scale bar refers to the phylogenetic distance. **b,** Genome organization of *sistrmB1*-like genes analyzed by SyntTax [1] in selected species of Sulfolobales. Predicted gene functions are indicated based on annotations in NCBI databases. Colors refer to orthologous genes.

Extended Data Fig. 2 Detection of the DNA binding activity of the mutants of SisTrmB1 at the DNA binding residues.

Structure of SisTrmB1 dimer predicted by AlphaFold3. The hot pink and wheat represent SisTrmB1 monomers, respectively. **b**, Structure of the AbfR2 (PDB:6CMV) of *S. acidocaldarius* [2]. **c**, Characterization of DNA binding capacity of SisTrmB1 and mutants by EMSA. **d,** Quantitative analysis of the results of different SisTrmB1 protein mutants. The quantification was performed using ImageJ, the Michaelis-Menten equation was used to fit the curve and derive the binding constant. At least two technical replications were performed for each set of EMSA assays. **e,** Structure of SisTrmB1-dimer-dsDNA complex predicted by AlphaFold3. R80 and R90 are shown as sticks.

Extended Data Fig. 3 SisTrmB1 bridges DNA.

a, The principle and schematic diagram of DNA bridging assay. Magnetic streptavidin beads were incubated with the 1,000 bp biotin-labeled bait DNA before SisTrmB1 was added. FAM-labeled dsDNA (500 bp) and the pBR322 plasmids with three topological states were chosen as the bridging substrates. **b,** The bridging ability of SisTrmB1 and its DNA-binding deficient mutants. An equal amount of substrate to the reaction was added. C, control without SisTrmB1. Each reaction was performed with at least three technical replicates. **c,** SisTrmB1 bridges linearized, supercoiled, and nicked plasmids. S, supercoiled; N, nicked; L, linearized. M: DNA marker. Each reaction was performed with at least three technical replicates. **d,** Schematic of proposed SisTrmB1-DNA

interaction. The SisTrmB1 dimer binds one dsDNA or binds and bridges two distant dsDNA. The orange ovals indicate dimeric SisTrmB1. Red and blue colors indicate dsDNA.

Extended Data Fig. 4 Growth curves of strains overexpressing DNA binding deficient mutants of SisTrmB1.

a, Western blotting verification of the expression of His-tagged SisTrmB1 mutants using anti-6×His-tag antibody. **b** and **c,** Growth curves of the SisTrmB1 mutant overexpression strains in STV (**b**) or ATV (**c**) medium. At least three biological replicates and two technical replicates were performed for each curve.

Extended Data Fig. 5 SisTrmB1 overexpression compromises the cell cycle of *Sa. islandicus*.

a, Diagrammatic overview of the experimental design and sampling strategy for cell cycle synchronization of SisTrmB1 overexpression strain. The arabinose was added 2 h before release for the induction of SisTrmB1. **b,** Flow cytometry profiles of SisTrmB1 overexpression strain and the control strain after release. Cells containing one and two chromosomes (1C and 2C) are indicated. Each set of experiments was conducted at least twice.

Extended Data Fig. 6 Examples showing AT-rich DNA binding preference of SisTrmB1 revealed by ChIP-Seq.

a, AT content (top) and ChIP-seq (bottom) of pSeSD and pSeSD-SisTrmB1 cells. AT content below the genomic average (64.7%) is plotted in reverse. **b** and **c,** SisTrmB1 binds to both intergenic and protein-coding sequences. AT content (top) and positions of annotated genes (bottom, 23S rRNA and 16S rRNA in green, tRNA gene in yellow).

Extended Data Fig. 7 RNA-Seq analysis of SisTrmB1 overexpression strain.

a, Volcano plot for differential gene expression in pSeSD and pSeSD-SisTrmB1. Both strains were grown in ATV medium for 6 h induction with arabinose. The Y-axis (-log10, padj) represents the statistical significance of the fold change, and the x-axis represents the log2 fold change in gene expression. Genes exhibiting ≥2-fold up- and down-regulation are highlighted in red and blue, respectively. **b,** Cluster heatmap of the gene expression difference in pSeSD and pSeSD-SisTrmB1.

Genes are clustered with log2 (FPKM + 1) and their expression levels are indicated by different colors with red representing the highest and blue indicating the lowest.

Extended Data Fig. 8 Hi-C contact maps and Pearson correlation matrices of synchronized samples and SisTrmB1 overexpression cells.

a, Pearson correlation matrices of Hi-C Data for E233S (CK), SC-1h, and SC-5h at 10-kb bin. The compartment index below was generated from the PC1 by PCA. The A and B compartments are indicated by red and blue bars, respectively. The data were combined from two biological replicates, respectively. **b**, Hi-C contact heatmaps of cells carrying pSeSD and pSeSD-SisTrmB1. The contact matrices represent interaction frequencies for pairs of 5-kb bins. **c**, Pearson correlation matrices of Hi-C data for strains carrying pSeSD and pSeSD-SisTrmB1 at 10-kb bin. The compartment index below was generated by PC1 and the A and B compartments are indicated by red and blue bars, respectively.

Extended Data Fig. 9 Correlation between SisTrmB1 enrichment and transcript abundance of SisTrmB1 overexpression cells.

a and **b,** Scatterplots comparing SisTrmB1 enrichment and transcripts (RPKM) of control (pSeSD) cells (**a**) and the overexpression cells (**b**). Spearman rank correlation coefficient (*r*) and corresponding *P* value are shown. The data were analyzed at 1-kb resolution. **c** and **d,** Comparison of transcript abundance (Log10 RPKM) in A compartment (**c**) and B compartment (**d**) of the SisTrmB1 overexpression cells and the control cells. The significant differences were calculated using the two-tailed Student's t-test with a significance threshold of $p < 0.05$ (****$P < 0.0001$).

Supplementary Tables

Supplementary Table S1. Strains and plasmids used in this study

Strain or vector	Description	References or sources
Escherichia coli DH5α	Plasmid amplification	Laboratory strain
E. coli BL21(DE3) codon plus-RIL	Protein expression	Laboratory strain
Sa. islandicus REY15A (E233S)	Δ*pyrEF*Δ*lacS*	Deng et al.[3]
pSeSD	A *Saccharolobus-E. coli* shuttle vector carrying an expression cassette controlled under a synthetic strong promoter $P_{araS\text{-}SD}$	Peng et al.[4]
pSeSD-SisTrmB	pSeSD carrying wild-type SisTrmB1 encoding sequence	This study
pSeSD-SisTrmB1-C-His	Expression of SisTrmB1 with His-tag at the C-terminal	This study
pSeSD-SisTrmB1-R80A-C-His	pSeSD carrying SisTrmB1 encoding sequence with R80A substitution	This study
pSeSD-SisTrmB1-K82A-C-His	pSeSD carrying SisTrmB1 encoding sequence with K82A substitution	This study
pSeSD-SisTrmB1-K87A-C-His	pSeSD carrying SisTrmB1 encoding sequence with K87A substitution	This study
pSeSD-SisTrmB1-R90A-C-His	pSeSD carrying SisTrmB1 encoding sequence with R90A substitution	This study
pSeSD-SisTrmB1-K82A/K87A-C-His	pSeSD carrying SisTrmB1 encoding sequence with K82A/K87A substitutions	This study
pSeSD-SisTrmB1-R80A/K82A-C-His	pSeSD carrying SisTrmB1 encoding sequence with R80A/K82A substitutions	This study
pSeSD-SisTrmB1-K87A/R90A-C-His	pSeSD carrying SisTrmB1 encoding sequence with K87A/R90A substitutions	This study
pSeSD-SisTrmB1-R80A/R90A-C-His	pSeSD carrying SisTrmB1 encoding sequence with R80A/R90A substitutions	This study
pSeSD-SisTrmB1-R80A/K82A/K87A/R90A-C-His	pSeSD carrying SisTrmB1 encoding sequence with R80A/K82A/K87A/R90A substitutions	This study
pET22b-SisTrmB1-K82A-C-His	pET22b carrying SisTrmB1 encoding sequence with K82A substitution	This study
pET22b-SisTrmB1-K87A-C-His	pET22b carrying SisTrmB1 encoding sequence with K87A substitution	This study
pET22b-SisTrmB1-R90A-C-His	pET22b carrying SisTrmB1 encoding sequence with R90A substitution	This study
pET22b-SisTrmB1-K82A/K87A-C-His	pET22b carrying SisTrmB1 encoding sequence with K82A/K87A substitutions	This study
pET22b-SisTrmB1-R80A/K82A-C-His	pET22b carrying SisTrmB1 encoding sequence with R80A/K82A substitutions	This study
pET22b-SisTrmB1-K87A/R90A-C-His	pET22b carrying SisTrmB1 encoding sequence with K87A/R90A substitutions	This study
pET22b-SisTrmB1- R80A/R90A-C-His	pET22b carrying SisTrmB1 encoding sequence with R80A/R90A substitutions	This study
pET22b-SisTrmB1-R80A/K82A/K87A/R90A-C-His	pET22b carrying SisTrmB1 encoding sequence with R80A/K82A/K87A/R90A substitutions	This study

Supplementary Table S2. Oligonucleotides used in this study

Primers	Sequence[a,b] (5'-3')
sistrmB1-NdeI-F	GTGGAATTCATATGATGTCGGAAACCCAATT
sistrmB1-SalI-R	AGGCGTCGATTACAATGGCTTAAACTCCTTT
sistrmB1-SalI-C-His-R	AGGCGTCGACAATGGCTTAAACTCCTTT
sistrmB1-R80A/K82A/K87A/R90A-SOE-F	ATTAGTAATG**GCA**ATA**GCG**GAGCCAGGGAAC**GCG**GCTGGA**GCA**CCAAGATATTTA
sistrmB1-R80A/K82A/K87A/R90A-SOE-R	ATAAATATCTTGG**TGC**TCCAGC**CGC**GTTCCCTGGCTC**CGC**TAT**TGC**CATTACTAAT
sistrmB1-R80A/R90A-SOE-F	ATTAGTAATG**GCA**ATAAAGGAGCCAGGGAACAAGGCTGGA**GCA**CCAAGATATTTA
sistrmB1-R80A/R90A-SOE-R	ATAAATATCTTGG**TGC**TCCAGCCTTGTTCCCTGGCTCCTTTAT**TGC**CATTACTAAT
sistrmB1-K82A-SOE-F	ATTAGTAATGAGAATA**GCG**GAGCCAGGGAACAAGG
sistrmB1-K82A-SOE-R	CCTTGTTCCCTGGCTC**CGC**TATTCTCATTACTAAT
sistrmB1-K87A-SOE-F	GGAGCCAGGGAAC**GCG**GCTGGAAGACCAAGATA
sistrmB1-K87A-SOE-R	TATCTTGGTCTTCCAGC**CGC**GTTCCCTGGCTCC
sistrmB1-R80A-SOE-F	GATGGGATTAGTAATG**GCA**ATAAAGGAGCCAGGGA
sistrmB1-R80A-SOE-R	TCCCTGGCTCCTTTAT**TGC**CATTACTAATCCCATC
sistrmB1-R90A-SOE-F	AGGGAACAAGGCTGGA**GCA**CCAAGATATTTATAC
sistrmB1-R90A-SOE-R	GTATAAATATCTTGG**TGC**TCCAGCCTTGTTCCCT
qPCR- *sistrmB1*-F	CCTTGTTCCCTGGCTCCTTT
qPCR- *sistrmB1*-R	TGAAAGGCGATGCTAGAGGT
qPCR-16S-F	CGCAAGACTGAAACTTAAAGGA
qPCR-16S-R	AGTCAGGCAAGGTCGTTAG
Biotin-dbr-F	ATCTGTAATTGCCTTAAGCCCAACTTTA
dbr-1000bp-R	ATTGCCTCAACCAAAAGGAAACTCCCCA
FAM-AT-F	TTAATATTATTTAATATTTTATAAAATTTTTATTAATTTAATTTTTAATTATTTT
AT-R	AAAATAATTAAAAATTAAATTAATAAAAATTTTATAAAATATTAAATAATATTAA
FAM-GC-F	GCGGCCGGGCCCGGGCCCCGCCGGGCCCCGCCGGGCCCCGCCGGGCCCCGCGCCG
GC-R	CGGCGCGGGGCCCGGCGGGGCCCGGCGGGGCCCGGCGGGGCCCGGGCCCGGCCGC

[a]The underlined denote sites of the restriction enzymes. [b]The mutated codons are indicated in boldface

Supplementary Table S3. Summary of mRNA levels of functional genes in the SisTrmB1 overexpression strain based on comparative transcriptomic analysis

Gene	Log2FoldChange (SisTrmB1_OE/CTRL)
Cell division, genome maintenance, and chromosome segregation	
CdvA (*sire_1173*)	-0.6208
ESCRT-III-1 (*sire_1200*)	-3.1125
ESCRT-III-2 (*sire_1388*)	-1.5550
ESCRT-III-3 (*sire_1550*)	-5.0631
ESCRT-III (*sire_1174*)	-0.4939
Vps4 (*sire_1175*)	-0.0121
SegA (*sire_1962*)	-2.7431
SegB (*sire_1961*)	-1.2527
SMC	
ClsN (*sire_0055*)	1.1353
NAPs	
Cren7 (*sire_1111*)	-1.4812
Alba1 (*sire_1125*)	-1.2553
Alba2 (*sire_1123*)	0.1953
Sis7d (*sire_0668*)	4.2118
Sis7d (*sire_2648*)	-0.4518
Sul12a (*sire_2004*)	-5.53816
DNA replication and transcription	
Dpo2N (*sire_0614*)	2.4491
Dpo2C (*sire_0615*)	2.7801
Orc1-2 (*sire_1231*)	1.8347
Orc1-3 (*sire_0002*)	-1.3104
Orc1-1 (*sire_1740*)	-1.0076
TFB1 (*sire_1555*)	-0.2978
TFB3 (*sire_1717*)	1.2723

1. Deng L, Zhu H, Chen Z, Liang YX, She Q. Unmarked gene deletion and host-vector system for the hyperthermophilic crenarchaeon Sulfolobus islandicus. *Extremophiles* **13**, 735-746 (2009).

2. Peng N*, et al.* A synthetic arabinose-inducible promoter confers high levels of recombinant protein expression in hyperthermophilic archaeon Sulfolobus islandicus. *Appl Environ Microbiol* **78**, 5630-5637 (2012).

Supplementary Materials and Methods

Bioinformatic analysis

Homology searches were performed using Protein BLAST (National Center for Biotechnology Information) and SyntTax[1]. Clustal Omega version 1.2.4 was used to generate sequence alignment[5]. The evolutionary analyses were conducted with MEGA11 [6] by using the Maximum Likelihood method and Le_Gascuel_2008 model [7]. SWISS-MODEL [8] and AlphaFold 3.0 [9] were used for protein structure prediction.

Generation of the overexpression strains of SisTrmB1 mutants

SistrmB1 and mutant derivatives were engineered by PCR to have a C-terminal 6×His-tag sequence. The *sistrmB1* gene and its mutant fragments were cloned into pSeSD plasmids digested by *NdeI* and *SalI*. The plasmids were transformed into E233S by electroporation, generating strains named pSeSD-SisTrmB1-R80A-C-His, pSeSD-SisTrmB1-R90A-C-His, pSeSD-SisTrmB1-K82A-C-His, pSeSD-SisTrmB1-K87A-C-His, pSeSD-SisTrmB1-R80A/K82A-C-His, pSeSD-SisTrmB1-K82A/K87A-C-His, pSeSD-SisTrmB1-K87A/R90A-C-His and pSeSD-SisTrmB1-R80A/R90A-C-His for short. Growth curves of SisTrmB1 mutant overexpression strains were obtained by inoculating cells in STV or ATV medium from an initial OD_{600} of 0.03-0.05 and subsequent culture and measurement. The samples were taken every 6 or 12 h to get the OD_{600} values. Cells containing pSeSD were used as the control.

Electrophoretic mobility shift assay (EMSA)

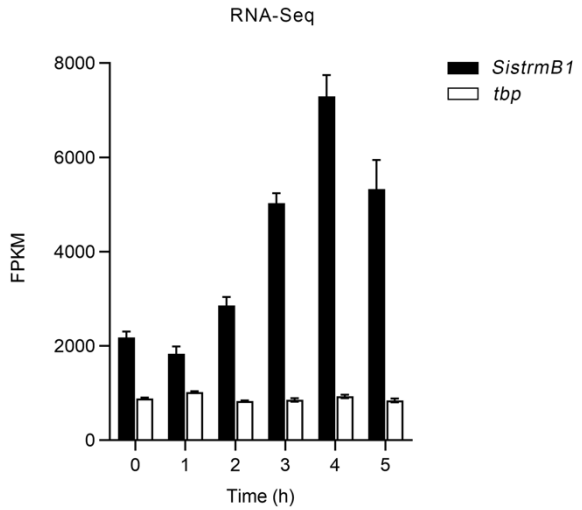
Circular supercoiled pBR322 plasmids were purified using a Plasmid Extraction Kit (Omega Bio-tek, Norcross, USA) following the provided protocol with two additional wash steps. linearized and nicked pBR322 plasmid DNAs were prepared by treatment with *BamHI* (Takara Biomedical Technology, Beijing) and Nb. *Bpu10I* (Thermo Fisher Scientific), respectively, and were purified by using a DNA Cycle-Pure Kit (Omega Bio-tek, Norcross, USA). For the reaction with plasmid DNA as substrate, different concentrations of SisTrmB1 (0, 0.05, 0.1, 0.2, 0.4, and 0.8 μM) were taken with 300 ng of plasmid pBR322 with different types (supercoiled, linearized, and nicked), and the reaction system was 50 mM Tris-HCl pH 6.8, 25 mM NaCl. The reaction was incubated at room temperature or 37°C for 30 min. After incubation, the results were detected by 1% agarose gel

electrophoresis with 0.5×TAE buffer at 110 V, 40 min, and visualized by EC3 Imaging System (Ultra-Violet Products Ltd, Cambridge UK).

Protein chemical cross-linking

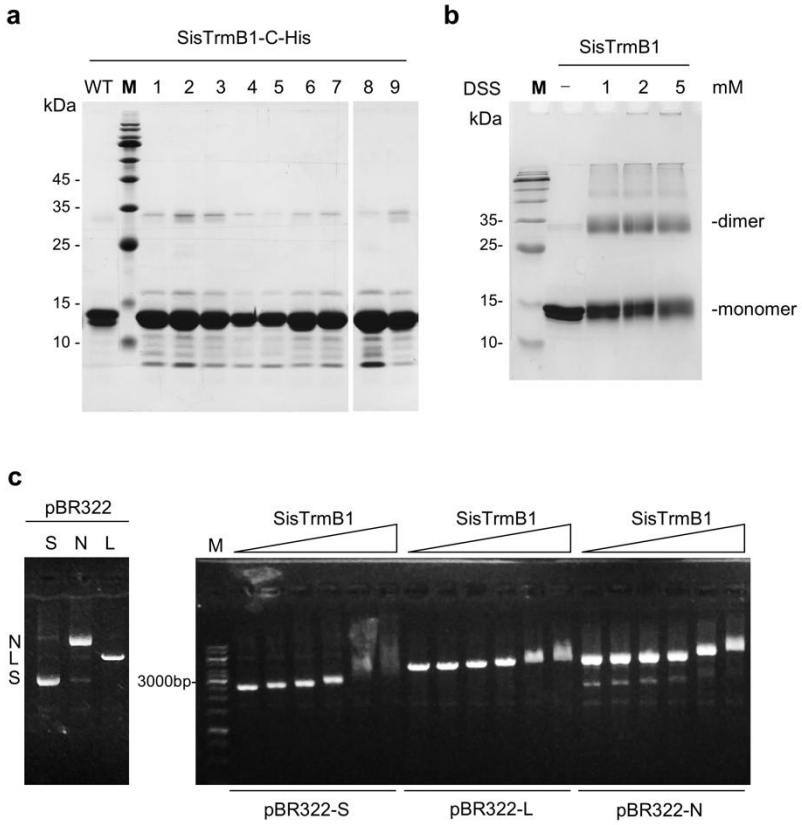
A quantity of 1 μg SisTrmB1 was cross-linked with DSS (Disuccinimidyl subera) crosslinker (Thermo Scientific) with a final concentration of 0, 1, 2, and 5 mM in solution. The mixtures were incubated at room temperature for 30 min. Then the samples were treated by adding 5×loading buffer (without 2-Mercaptoethanol) and boiling for 10 min before fractionating by 15 % sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE).

Supplementary Figures



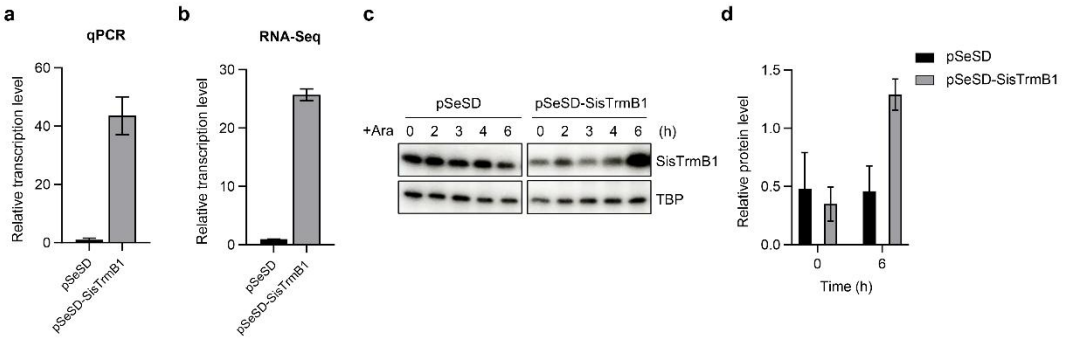
Supplementary Fig. 1 The transcription levels of *sistrmB1* during one cell cycle based on the RNA-Seq data[10].

The samples of *Sa. islandicus* REY15A (E233S) were collected each hour after cell cycle synchronization for RNA-Seq. *tbp* was chosen as a control. The values were based on three biological repeats.



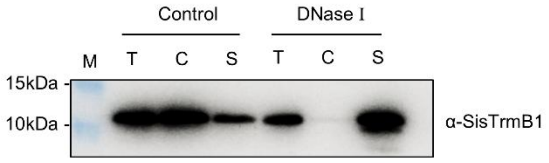
Supplementary Fig. 2 Biochemical properties of SisTrmB1.

a, Heterologous expression and purification of SisTrmB1 and its mutants in *E. coli*. The theoretical molecular mass of SisTrmB1 is 14.3 kDa. The proteins were all fused with 6×His tag at the C-terminus. WT: wild type SisTrmB1. Lanes 1-9, SisTrmB1 mutants with alanine substitution: R80A, K82A, K87A, R90A, R80A/K82A, K82A/K87A, K87A/R90A, R80A/R90A, and R80A/K82A/K87A/R90A. M: protein marker. **b,** Chemical cross-linking of SisTrmB1. Different concentrations of DSS (1, 2, and 5 mM) and 1 μg SisTrmB1 were added to the reaction mixture with the purified SisTrmB1 as control. The positions of the SisTrmB1 monomer and dimer are labeled. DSS: Disuccinimidyl subera. M: protein marker. **c,** SisTrmB1 binds to linearized, supercoiled, and nicked plasmids. The figure on the left shows the electrophoretic position of the three states of pBR322. S: supercoiled; N: nicked; L: linearized. M: DNA marker. Each reaction was performed with at least three technical replicates.



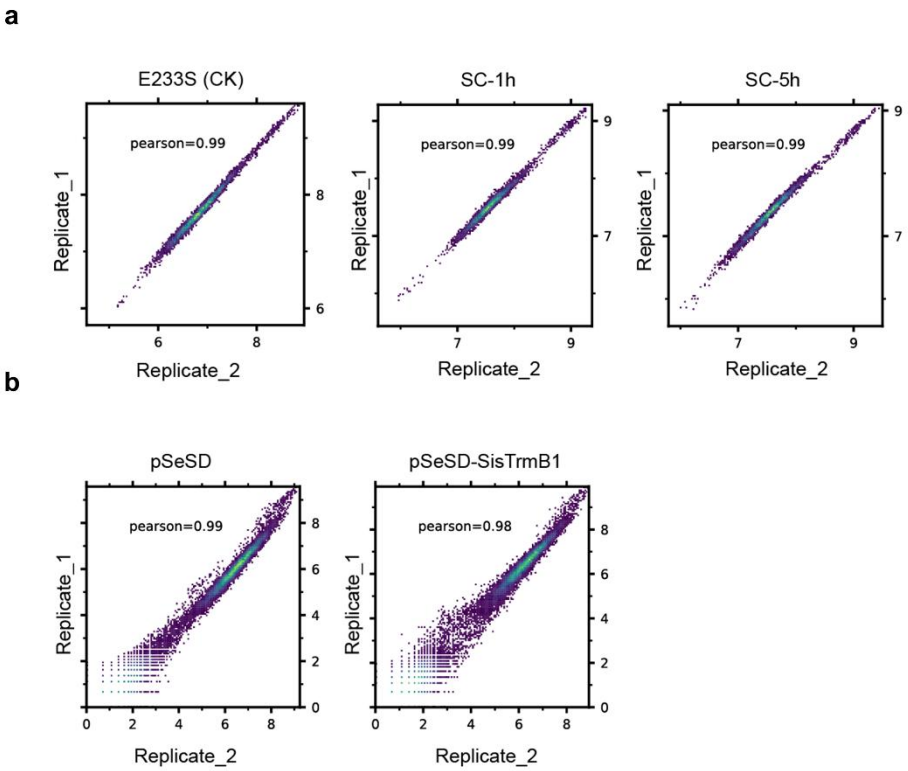
Supplementary Fig. 3 Expression of SisTrmB1 at mRNA and protein levels after arabinose induction.

a and **b,** Relative mRNA level of SistrmB1 detected by qRT-PCR (**a**) and from RNA-Seq data (**b**) in SisTrmB1 overexpression strain after 6 h induction of arabinose. 16s rRNA gene was chosen as the internal reference. The data was normalized by the FPKM value of SisTrmB1 in pSeSD. The RNA-Seq was performed by three biological replicates. **c,** Western blotting analysis of the protein levels of SisTrmB1 in the overexpression strain and control after arabinose induction. TBP (TATA-box binding protein) was used as a control. The experiments were performed at least three times, with representative images being shown. **d,** Quantitative analysis of the results in (c) by ImageJ. The relative protein levels of SisTrmB1 at 0 h and 6 h were shown.



Supplementary Fig. 4 SisTrmB1 binds to chromatin DNA.

Related to Fig. 4a. Total protein (T) of cells in exponential phase was separated into chromatin (C) and soluble (S) fractions, and the samples were separated by SDS-PAGE. SisTrmB1 was detected by Western blotting. For samples treated with DNase I, total protein sample was incubated with DNase I before fractioning.



Supplementary Fig. 5 The reproducibility of Hi-C data.

a, Correlation of the two biological replicates of Hi-C data for E233S (CK), SC-1h and SC-5h. **b,** Correlation of the two biological replicates of Hi-C data for pSeSD and pSeSD-SisTrmB1.

References

1. Oberto J. SyntTax: a web server linking synteny to prokaryotic taxonomy. *BMC Bioinformatics* **14**, 4 (2013).

2. Vogt MS, Volpel SL, Albers S-V, Essen L-O, Banerjee A. Crystal structure of an Lrs14-like archaeal biofilm regulator from Sulfolobus acidocaldarius. *Acta Crystallographica Section D* **74**, 1105-1114 (2018).

3. Deng L, Zhu H, Chen Z, Liang YX, She Q. Unmarked gene deletion and host-vector system for the hyperthermophilic crenarchaeon Sulfolobus islandicus. *Extremophiles* **13**, 735-746 (2009).

4. Peng N*, et al.* A synthetic arabinose-inducible promoter confers high levels of recombinant protein expression in hyperthermophilic archaeon Sulfolobus islandicus. *Appl Environ Microbiol* **78**, 5630-5637 (2012).

5. Sievers F*, et al.* Fast, scalable generation of high‐quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* **7**, 539 (2011).

6. Tamura K, Stecher G, Kumar S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Mol Biol Evol* **38**, 3022-3027 (2021).

7. Le SQ, Gascuel O. An improved general amino acid replacement matrix. *Mol Biol Evol* **25**, 1307-1320 (2008).

8. Guex N, Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18**, 2714-2723 (1997).

9. Abramson J*, et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493-500 (2024).

10. Yang Y*, et al.* A novel RHH family transcription factor aCcr1 and its viral homologs dictate cell cycle progression in archaea. *Nucleic Acids Res* **51**, 1707-1723 (2023).