frontiers | Frontiers in Microbiology

TYPE Original Research
PUBLISHED 05 August 2025
DOI 10.3389/fmicb.2025.1602937

OPEN ACCESS

EDITED BY
Marleen van Wolferen,
University of Freiburg, Germany

REVIEWED BY
Changyi Zhang,
University of Illinois at Urbana-Champaign, United States
Catherine Badel,
Université de Strasbourg, France

*CORRESPONDENCE
Eveline Peeters
✉ Eveline.Peeters@vub.be

RECEIVED 30 March 2025
ACCEPTED 11 July 2025
PUBLISHED 05 August 2025

CITATION
Xu Y, Peeters AI, Bervoets I, De Mey M, Baes R and Peeters E (2025) Effects of genomic location on ectopic integration and gene expression of a reporter gene cassette in *Sulfolobus acidocaldarius*. *Front. Microbiol.* 16:1602937. doi: 10.3389/fmicb.2025.1602937

Effects of genomic location on ectopic integration and gene expression of a reporter gene cassette in *Sulfolobus acidocaldarius*

Yifei Xu[1], Andries Ivo Peeters[2], Indra Bervoets[1], Marjan De Mey[2], Rani Baes[1] and Eveline Peeters[1]*

[1]Research Group of Microbiology, Department of Bioengineering Sciences, Vrije Universiteit Brussel, Brussels, Belgium, [2]Centre for Synthetic Biology, Department of Biotechnology, University of Ghent, Ghent, Belgium

In eukaryotes and bacteria, it is well-established that the genomic location of ectopic gene integration influences the expression level due to replication-associated gene dosage effects as well as effects mediated by chromatin organization. In contrast, in archaea, the impact of genomic location on gene expression remained unexplored. Here, we investigated this impact in the model archaeon *Sulfolobus acidocaldarius*, a crenarchaeal species that has a chromatin architecture with mixed eukaryotic-like and bacterial-like features. We aimed to integrate a standardized β-galactosidase (*lacS*) reporter cassette into diverse loci in the genome of *S. acidocaldarius* SK-1 for a comparative analysis. Nine integration mutant strains were successfully obtained, for which qRT-PCR analysis and *lacS* reporter gene assays revealed significant variation in transcriptional and translational expression of the reporter, respectively, demonstrating that genomic location strongly influences gene expression in *S. acidocaldarius*. However, variability in transcription levels and its regulation was shown to be primarily driven by transcriptional activity of neighboring genes, due to the high coding density in the *S. acidocaldarius* genome as well as a lack of insulator elements. In conclusion, this study not only provides insights into genome context effects, but also provides inspiration for the future design of genomic knock-in constructions in *S. acidocaldarius*.

KEYWORDS

archaea, Crenarchaeota, chromatin organization, reporter gene assays, transcription, genetic tools

1 Introduction

Within the Crenarchaeota, species belonging to the Sulfolobales are considered important models for fundamental studies of molecular biological processes in archaea (Lee et al., 2022), also providing evolutionary insights. In addition, they show promise as hosts for biotechnological applications due to their thermoacidophilic lifestyle—growing optimally at a high temperature between 75 and 80°C and a low pH between 2 and 3—and their metabolic characteristics, for example chemoorganotrophic metabolism (Crosby et al., 2019). Because of its genetic stability, *Sulfolobus acidocaldarius* is regarded as a chassis species for genetic studies, as well as for genetic engineering for biotechnological purposes (Chen et al., 2005; Wagner et al., 2012; Crosby et al., 2019).

Over the past two decades, a suite of genetic tools has been developed for *S. acidocaldarius* (Albers and Driessen, 2008; Wagner et al., 2012; Lee et al., 2022). These tools include

electroporation procedures to introduce exogenous DNA, shuttle plasmid vectors derived from native *Saccharolobus* viruses and plasmids, as well as genome engineering approaches based on homologous recombination. Indeed, the native homologous recombination machinery of *S. acidocaldarius* was shown to be sufficiently efficient to enable recombineering (Kurosawa and Grogan, 2005; Wagner et al., 2009; Suzuki and Kurosawa, 2017). Pyrimidine auxotrophy is commonly used as a selection strategy, with the *pyrEF* operon serving as a selection marker. Selection is performed with uracil and counterselection with 5-fluoroorotic acid (5-FOA), thereby enabling deletion of genomic segments or integration of exogenous DNA with a "pop-in pop-out" approach. This results in the removal of the selection marker, allowing the same marker to be used for multiple subsequent alterations (Wagner et al., 2012).

Using the "pop-in pop-out" method, a repertoire of markerless gene deletion mutants of *S. acidocaldarius* has been constructed, facilitating the study of gene functions (Wagner et al., 2012; Lee et al., 2022). In contrast, the use of this approach for the ectopic integration of a heterologous gene cassette, generating a "knock-in" mutant of *S. acidocaldarius*, has been more scarcely reported (Lee et al., 2022). Nevertheless, in the light of developing *S. acidocaldarius* as a biotechnological host, it is highly relevant to introduce novel genetic traits, such as heterologous enzyme expression, directly into the genome rather than relying on plasmid-based transformations, which might suffer stability issues. In the few reported cases of ectopic gene cassette integration into *S. acidocaldarius*, the genomic location for integration was primarily selected with the aim of minimizing impact on the functioning of genes overlapping or adjacent to the integration site (Zeldes et al., 2019; Lee et al., 2022). For example, the *upsE* locus in *S. acidocaldarius*, encoding UV-inducible pili, was chosen as a target site for ectopic integration of the glucose transporter operon from *Saccharolobus solfataricus*, as *upsE* was previously successfully deleted (Wagner et al., 2012). In another example, a cassette encoding a sulfur oxidation operon was integrated into the *Saci_1149* locus, concomitantly deleting a gene encoding an enzyme in the 3-hydroxypropionic acid/4-hydroxybutyrate (3HP-4HB) pathway, which was shown to be inactive in *S. acidocaldarius* (Zeldes et al., 2019). Finally, a gene cassette encoding a β-xylosidase from *S. solfataricus* was integrated into the *pyrEF* locus of *S. acidocaldarius*, which was already disrupted in the host strain (Lee et al., 2022). Recently, a CRISPR-COPIES pipeline was used to identify and target 8 chromosomal integration sites in the related species *Sulfolobus islandicus*, with the heterologous expression of glycerol dibiphytanyl glycerol tetraether ring synthase B serving as a proof-of-concept (Boob et al., 2025).

In the above case studies, with the exception of the recent study in *S. islandicus* (Boob et al., 2025), it has not been considered whether or how the genomic location of integration affects gene expression of the integrated gene or operon due to genetic context effects. To our knowledge, this is an understudied question in Sulfolobales and in archaea in general. In contrast, in eukaryotes it is well-established that overall transcriptional activity varies between chromosomal regions, thereby correlating with gene density and with chromatin organization into discrete compartments (Lieberman-Aiden et al., 2009; Bonev and Cavalli, 2016; Rowley and Corces, 2018). In bacteria, despite their simpler prokaryotic genome structure, gene expression is also subject to position-dependent effects (Cooke et al., 2019). This has been shown in *Escherichia coli* in various studies (Beckwith et al., 1966; Sousa et al., 1997; Block et al., 2012; Bryant et al., 2014; Brambilla and Sclavi, 2015; Scholz et al., 2022), such as by transposing a β-galactosidase (*lacZ*) reporter gene cassette to different chromosomal locations (Sousa et al., 1997). It was found that gene expression increases with an increasing proximity to the origin of replication, an effect linked to replication-associated gene dosage (Beckwith et al., 1966). Additional local effects, such as differences in genome accessibility to RNA polymerase, chromatin structure and organization, nucleoid-associated proteins (NAPs) and DNA topology were also observed (Block et al., 2012; Bryant et al., 2014; Brambilla and Sclavi, 2015; Scholz et al., 2022). For instance, also in *E. coli*, a GFP reporter cassette exhibited gene expression differences of up to 300-fold depending on chromosomal location (Bryant et al., 2014).

Intriguingly, while archaea share a similar gene organization to bacteria, typified by operonic structures and high coding density, their chromatin is structured using a combination of bacterial and eukaryotic mechanisms. In *S. acidocaldarius*, the chromosome is three-dimensionally organized into a higher-order structure reminiscent of eukaryotic chromatin, forming two chromosomal compartments: compartment A is characterized by a higher transcriptional activity, containing essential genes and replication origins and compartment B is characterized by a lower transcriptional activity and is enriched in non-essential genes and transposons (Takemata et al., 2019; Takemata and Bell, 2021; Pilatowski-Herzing et al., 2025). Unlike most archaea, Crenarchaea, including *S. acidocaldarius,* lack condensin, a protein involved in chromatin organization (Pilatowski-Herzing et al., 2025). Instead, Sulfolobales harbor an SMC protein named coalescin (ClsN), which is enriched in the B compartment and associated with gene silencing (Takemata et al., 2019). On a smaller scale, the chromatin of *S. acidocaldarius* and other Sulfolobales species is organized by a diverse repertoire of small (7–12 kDa) NAPs, including Cren7, Sul7, Alba, Sul10a and Sul12a, which contribute to chromatin structuring through diverse DNA interactions, including DNA kinking, compaction, bridging, looping and supercoiling (Grote et al., 1986; Lurz et al., 1986; Choli et al., 1988; Baumann et al., 1994; López-García et al., 1998; Bell et al., 2002; Guo et al., 2008; Driessen et al., 2016; Kalichuk et al., 2016; Zhang et al., 2020; Lemmens et al., 2022).

In this study, we address the hypothesis that the genomic location of ectopic integration affects gene expression in *S. acidocaldarius*, similarly as in eukaryotes and bacteria. Specifically, we aim to explore how a location within each of the higher-order A or B chromatin compartments in *S. acidocaldarius* affects gene expression, given that they show significant differences in global transcriptional activity in the native transcriptome (Takemata et al., 2019). This will be achieved by integrating a constitutive β-galactosidase (*lacS*) gene reporter cassette into different genomic locations which are selected based on differences in chromatin compartment and proximity to a replication origin, as well as the expression levels and predicted essentiality of neighboring genes.

2 Materials and methods

2.1 Microbial strains and cultivation conditions

The uracil-auxotrophic strain *S. acidocaldarius* SK-1 (Suzuki and Kurosawa, 2016) and all derived mutant strains were cultivated in

liquid in basic Brock medium (Brock et al., 1972) supplemented with 0.1% (w/v) N-Z-Amine, 0.2% (w/v) sucrose and 20 µg mL^{-1} uracil, acidified to pH 3.0 with sulfuric acid. Liquid *S. acidocaldarius* cultures were incubated at 75°C while shaking. Growth was monitored by measuring optical density at 600 nm (OD_{600}). Cells were cultivated until reaching an OD_{600} of between 0.4 and 0.5 (exponential growth phase) or OD_{600} of 1.0 (stationary growth phase). Nutrient starvation was performed based on Haurat et al. (2017). In this case, cells were cultivated until reaching OD_{600} 0.4 and collected by centrifugation at 4,000 g for 10 min at room temperature. The cell pellet was resuspended in fresh prewarmed Brock medium supplemented with 20 µg mL^{-1} uracil lacking N-Z-Amine and sucrose and incubated at 75°C for 4 h while shaking.

For cultivation of *S. acidocaldarius* on solid medium, Brock medium was supplemented with 1.5 mM $CaCl_2$, 5 mM $MgCl_2$ and 0.6% Gelrite as a solidifying agent (Wagner et al., 2012). For selection of marker-containing integration mutants ("pop-in" step), the medium was solely supplemented with 0.1% (w/v) N-Z-Amine, 0.2% (w/v) sucrose, while for selection of markerless knock-in mutants ("pop-out" step), the medium was supplemented with 0.1% (w/v) N-Z-Amine, 0.2% (w/v) sucrose, 20 µg mL^{-1} uracil and 200 µg mL^{-1} 5-FOA. Solid medium was prewarmed to 75°C before performing plating (Wagner et al., 2012).

E. coli strain DH5α was used for cloning and for propagation of plasmid constructs and was cultivated in Lysogeny Broth (LB) medium supplemented with 50 µg mL^{-1} ampicillin, if required. For selection of transformants from Golden Gate cloning, LB medium was used in which NaCl was omitted and 5% (w/v) sucrose was added. All strains used in this work are provided in Supplementary Table S1.

2.2 DNA manipulations and molecular cloning

Gibson assembly (Gibson et al., 2009) was used to construct a Golden Gate destination vector named pYX2304 (Figure 1). To this end, the backbone of suicide plasmid vector pSVA431 (Wagner et al., 2012) was amplified using oligonucleotides YX123 and YX124 (Supplementary Table S2) and the *sacB* gene was amplified from the pRN1-carrier_2_paqci plasmid vector using primers YX125 and YX126. Polymerase Chain Reaction (PCR) amplification was performed with KAPA HiFi DNA polymerase (Roche) using 10 ng plasmid DNA and 20 pmol of each primer in a 50-µL reaction with following PCR conditions: 3 min at 95°C, 35 cycles of 20 s at 98°C, 15 s at 60°C and 1 min/kb at 72°C, followed by a final step of 5 min at 72°C. PCR amplification was verified by agarose gel electrophoresis and products were purified using the Wizard® SV Gel and PCR Clean-Up System (Promega). Next, Gibson assembly was performed by incubating 100 ng of amplified pSVA431 backbone with an equimolar amount of the *sacB* gene fragment in NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs) at 50°C for 1 h, followed by heat-shock transformation into chemically competent *E. coli* DH5α. Transformants were screened by colony PCR using primers YX127 and YX128 and the sequence of the obtained construct was verified by Sanger sequencing (Eurofins Genomics).

The pYX2304 destination vector was subsequently used to construct all suicide knock-in plasmid vectors with Golden Gate assembly (Engler et al., 2008). For each of the targeted genomic

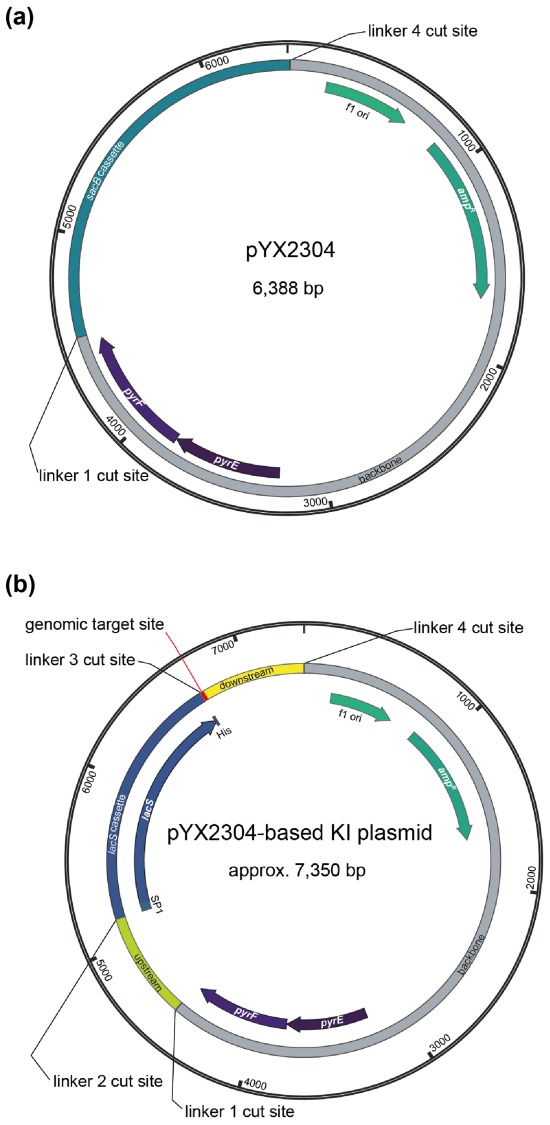

FIGURE 1
Plasmid maps of the Golden Gate destination vector pYX2304 **(a)** and the derived suicide knock-in (KI) plasmid vector **(b)** with indication of the different type IIS restriction sites with the following cut sites: linker 1 = 5′-AGGA-3′, linker 2 = 5′-TGAT-3′, linker 3 = 5′-GATG-3′ and linker 4 = 5′-GCAG-3′. The pYX2304 vector was designed to facilitate one-step assembly of PCR-amplified genomic flanking regions (~650 bp upstream and downstream of the target site) together with a central *lacS* reporter cassette. Fragments were joined using PaqCI-mediated Golden Gate cloning, enabling scarless and directional assembly based on the defined overhangs (refer to Supplementary Figure S2). The *lacS* cassette was amplified from a modified plasmid (pYX2301-UPDD) in which the native promoter was replaced with a synthetic constitutive promoter (P_{Ec10_Sal}) via Gibson assembly.

locations, primers (Supplementary Table S2) were designed to enable amplification of fragments of approximately 650 bp up- and downstream of the target site, respectively. These PCR reactions were performed using KAPA HiFi DNA polymerase (Roche) with 10 ng genomic DNA (gDNA) of *S. acidocaldarius* SK-1, which was extracted from liquid culture employing a QuickPick SML gDNA kit with magnetic bead purification (BioNobile). The *lacS* reporter cassette was PCR-amplified using KAPA HiFi DNA polymerase (Roche) with

primers YX133 and YX134 and pYX2301-UPDD plasmid DNA as a template. The latter was derived from pSVA431 by replacing the P*malE* promoter driving *lacS* expression with a synthetic constitutive promoter named P_{Ec10_Sa1}. This was realized via Gibson assembly, with the 56-bp P_{Ec10_Sa1} promoter sequence being introduced by primer YX091 (Supplementary Table S2). Golden Gate assembly was performed in 20-µL reactions consisting of 2 µL T4 ligase buffer (ThermoFisher), 1 µL T4 ligase (ThermoFisher), 1 µL PaqCI (Bioké), 0.3 µL PaqCI activator (Bioké), 100 ng pYX2304 plasmid DNA and equimolar amounts of the fragments. Reaction mixtures were subjected to following conditions: 50 cycles of 3 min at 37°C, 4 min at 16°C, followed by a final step of 5 min at 37°C and 5 min at 60°C. Subsequently, reaction mixes were heat-shock transformed into chemically competent *E. coli* DH5α and transformants were screened by colony PCR using customized primers (Supplementary Table S2). Finally, the sequence of the obtained construct was verified by Sanger sequencing (Eurofins Genomics). An overview of all plasmids constructed and used in this work is provided in Supplementary Table S3.

2.3 Construction of *Sulfolobus acidocaldarius* knock-in mutant strains

Construction of *S. acidocaldarius* markerless knock-in mutant strains was performed using a classical "pop-in pop-out" strategy as described (Wagner et al., 2012; Suzuki and Kurosawa, 2016) (Supplementary Figure S1). Competent *S. acidocaldarius* SK-1 cells were prepared by harvesting cells from a 100-mL culture at an OD_{600} of 0.2 through centrifugation at 6574 *g* for 20 min, followed by two washing steps in 20 mM sucrose and a final resuspension of the cells in this 20 mM sucrose solution, reaching a final concentration of 2×10^{10} cells mL^{-1}. Genetic transformation was performed with a Gene Pulser II electroporator (Bio-Rad) at 1.5 kV, 25 mF and 600 W with Gene Pulser 1-mm cuvettes (Bio-Rad) (Wagner et al., 2012). After 5 days of incubation at 75°C on solid Brock medium lacking uracil, transformant colonies were screened for β-galactosidase activity by spraying with 5 mg mL^{-1} 5-bromo-4-chloro-3-indolyl-beta-D-galacto-pyranoside (X-Gal) (ThermoFisher), followed by an incubation at 75°C for 30 min. Positive "blue" integrants were further confirmed by PCR analysis and were cultivated in liquid culture without addition of uracil. 50 µL of liquid culture was subsequently plated on Gelrite plates with uracil (20 µg mL^{-1}) and 5-FOA (200 µg mL^{-1}) for the second selection of "pop-out" of the vector backbone by homologous recombination and incubated for 5 days at 75°C. Transformant colonies were screened for β-galactosidase activity as described previously and successful integrants were confirmed through PCR analysis.

2.4 Quantitative reverse transcriptase PCR

Culture samples of 4 mL were centrifuged for 15 min at 4000 *g* and pellets were stabilized using an equal volume of RNAprotect (Qiagen) and subjected to RNA extraction using an SV Total RNA Isolation System kit (Promega), followed by removal of residual genomic DNA using a Turbo DNase kit (Ambion Life Technologies). Next, cDNA was prepared from 1 µg RNA using a Go-Script Reverse Transcriptase kit (Promega). Three biological replicates were included in this experiment, except for the stress condition, which was conducted without replication.

Quantitative reverse transcriptase PCR (qRT-PCR) was performed to determine relative transcriptional gene expression levels of *lacS* employing *tbp* (*saci_1336*) as a reference gene. qRT-PCR primers were designed with Primer3 software (Untergasser et al., 2012) (Supplementary Table S2) and tested for efficiency using *S. acidocaldarius* gDNA as a template. 20-µL qRT-PCR reactions were performed in an iCycler qPCR device (Bio-Rad) with each reaction mixture containing 1 µL 10-fold diluted cDNA, 1.6 pmol of each primer and GoTaq qPCR master mix (Promega). The following PCR conditions were used: 3 min at 95°C and 40 cycles of 10 s at 95°C and 30 s at 55°C. Quantification cycle (C_T) values were determined with iQ5 software (Bio-Rad). For each gene and biological replicate, three technical replicates were performed, for which the average C_T value was determined. Relative expression levels were normalized with respect to the reference gene and ratios were calculated using the $\Delta\Delta C_T$ method relative to the strain displaying the lowest expression level. Graphpad Prism 9 was employed to perform a two tailed, one sample *T* test for statistical analysis.

2.5 Western blotting

Western blotting was performed based on the His-tagged LacS. To this end, 4 mL of *S. acidocaldarius* cells, cultivated until an OD_{600} of 0.4, were pelleted by centrifugation at 4000 *g* for 15 min at 4°C. The pellet was resuspended in 200 µL extraction buffer (50 mM Tris–HCl, 50 mM NaCl, 15 mM $MgCl_2$, 1 mM DTT) to which 0.1% Triton X-100 was added. Cells were incubated at 4°C for 30 min while rotating for lysis. Total protein concentration was determined using Bradford Assay solution (TCI Chemicals), using BSA for the generation of the standard curve.

Normalized protein extracts (30 µg) were mixed with LDS (ThermoFisher) and subjected to sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) using 4–12% NuPAGE™ Bis-Tris Mini Protein gels (ThermoFisher) and NuPAGE™ MES SDS Running Buffer (ThermoFisher). After SDS-PAGE, proteins were electroblotted onto a polyvinylidene difluoride (PVDF) membrane (Bio-Rad) using a Trans-Blot Turbo transfer system (Bio-Rad) operated at 1.3 A, up to 25 V for 7 min. The membrane was then incubated in Phosphate Buffer Saline (PBS) buffer with 5% (v/v) milk and 0.1% (v/v) Tween20 for 1 h at room temperature followed by an overnight incubation at 4°C with anti-His-tag mouse antibody (Proteintech, Cat No. 66005-1-Ig) diluted at 1:5000 in the same buffer. Excess primary antibody was washed away with PBS buffer containing 0.1% Tween20, followed by incubation with a secondary goat anti-mouse HRP-conjugated antibody (Proteintech, Cat No. SA00001-1) diluted at 1:5000 in the same buffer. After 1 h incubation at room temperature, unbound antibodies were washed away with PBS buffer containing 0.1% Tween20, twice, and PBS buffer for final wash. Visualization was performed with a Pierce™ ECL Western Blotting Substrate kit (ThermoFisher) and an ImageQuant800 imager (Cytiva). Band intensities were quantified using ImageJ (Schneider et al., 2012) as integrated density values (area × mean gray value). Values are reported in arbitrary units.

2.6 β-galactosidase reporter gene assays

β-galactosidase assays were performed as described (Van der Kolk et al., 2020) with some modifications. The σ-nitrophenyl-β-D-galactopyranoside (ONPG) conversion rate was measured spectrophotometrically at 410 nm in a microplate reader (Infinite M Nano+, TECAN) every 5 min, while being incubated at 42°C for 4 h. The slope of the conversion curves was taken as the value for β-galactosidase activity, which is a proxy for LacS expression. For each sample, the assay was performed in triplicate. The slopes were determined by fitting a Monod-type model with smoothing to the conversion data:

$$OD_{410}(t) = \frac{sl * dt * m}{sl * dt + m}$$

with

$$dt = \frac{(t - t_{off}) + \sqrt{(t - t_{off})^2 + d^2}}{2}$$

where sl = slope, m = maximum value, t = time, t_{off} = offset and d = period around the offset that the transition takes place. The slope was determined by fitting the model to ranges of datapoints from t_0-t_{max} to t_0-t_5. Local minima of the fit error of these fits were then searched and from those fits, the slope of the fit resulting in the slope with the lowest standard error was selected. When the resulting fit was not deemed adequate, the average value of the fitted d parameter of the other technical and/or biological replicates was used as fixed value for a subsequent fitting. The geometric mean of the slopes was calculated, together with the geometric standard error. If the fit error or error from the technical replicates was larger than the geometric standard error, that one was used instead.

3 Results

3.1 Establishment of an efficient suicide vector cloning system for construction of knock-in mutant strains

In this study, we used a knock-in suicide plasmid vector that is based on the pSVA431 plasmid vector harboring a *S. solfataricus pyrEF* gene cassette and a pGEM-T Easy cloning vector containing an ampicillin resistance cassette and f1 origin of replication for *E. coli* (Wagner et al., 2012). To facilitate the construction of knock-in suicide vectors, we have chosen to employ a one-step Golden Gate assembly approach. To this end, pSVA431 was converted into a Golden Gate destination vector named pYX2304 (Figure 1a). This vector combines the pSVA431 backbone with a *sacB* gene cassette flanked by type IIS restriction sites. The *sacB* gene encodes levansucrase, an enzyme that converts sucrose into levans, which accumulates in the *E. coli* periplasm causing toxicity (Gay et al., 1985). During cloning, it can thus be used as a negative selection marker in presence of sucrose, enabling the selection of transformants harboring a vector in which the *sacB* cassette is excised and replaced by the fused *lacS* cassette and target up- and downstream regions (Supplementary Figure S2).

The designed knock-in suicide plasmid vector (Figure 1b) can be used to create markerless knock-in mutant strains with a single-crossover recombination approach, employing the *pyrEF* cassette as a selection marker (Wagner et al., 2012) and the uracil auxotrophic *S. acidocaldarius* SK-1 (Suzuki and Kurosawa, 2016) as a host strain (Supplementary Figure S1). *S. acidocaldarius* SK-1 is derived from *S. acidocaldarius* M31, which harbors a 31-bp deletion in the *pyrEF* region (Reilly and Grogan, 2001), and was further engineered by deleting the *suaI* endonuclease-encoding gene, thereby inactivating its restriction-modification system (Suzuki and Kurosawa, 2016). As compared to *S. acidocaldarius* MW001, SK-1 offers the advantage that methylation of the plasmid DNA is not required, thereby simplifying the genetic transformation procedure. Otherwise, given its uracil auxotrophy, the use of SK-1 as a host strain enables a classical "pop-in pop-out" scheme in which, in a first selection step directly following transformation, cultivation on solid medium lacking uracil results in selection of integrant mutant strains, with the entire plasmid vector being integrated into the genome ("pop-in" step) (Wagner et al., 2012). Successfully obtained integrant mutant strains are then subjected to a second selection step, which entails a counterselection because of the presence of uracil and 5-FOA, resulting in the selection of strains that have undergone a second single-crossover recombination event ("pop-out" step), leading either to the wild-type genomic sequence or to the desired knock-in mutant sequence (Supplementary Figure S1).

The *lacS* reporter cassette consisted of a *lacS* gene from *S. solfataricus*, previously validated as a reporter system in *S. acidocaldarius* (Berkner et al., 2007), fused to a C-terminal 6xHis-tag and under control of a synthetic *Sulfolobus* promoter named P_{Ec10_Sa1}, which is based on the native *Saci_2137* promoter, driving expression of a putative aminotransferase (Liu et al., 2014). This in-house developed promoter has a small size of 56 bp (Supplementary Table S1), thereby contributing to a minimal vector size, as well as a constitutive expression profile of moderate strength and a non-native sequence, which helps to avoid unwanted homologous recombination with the genome.

3.2 Efficiencies in genome integration and in obtaining markerless knock-in mutant strains

We selected 11 target positions for insertion within the *S. acidocaldarius* SK-1 genome, based on (i) variations in the relative distance of the site to the nearest origin of replication, (ii) transcriptional expression levels of nearby or overlapping genes under physiological conditions (Baes et al., 2023) and (iii) predicted essentiality assessed based on homology with *S. islandicus* genes (Zhang et al., 2018) (Figure 2; Table 1). Additionally, the target sites were chosen to be evenly distributed between the two chromosomal compartments A and B and for this reason, target sites were selected in two chromosomal segments: between oriC1 and oriC2 and between oriC1 and oriC3 (Figure 2; Table 1). In all cases with one exception (*cmp*), the reporter gene cassette was specifically targeted to intergenic regions to minimize unintended effects on the expression of adjacent genes (Figure 3; Supplementary Figure S3). Intergenic regions include those located between genes transcribed in the same direction, whether part of an operon (e.g., *ccc1*) or not (e.g., *sdhC*), as well as between convergently (e.g., *cmp*) or divergently transcribed genes

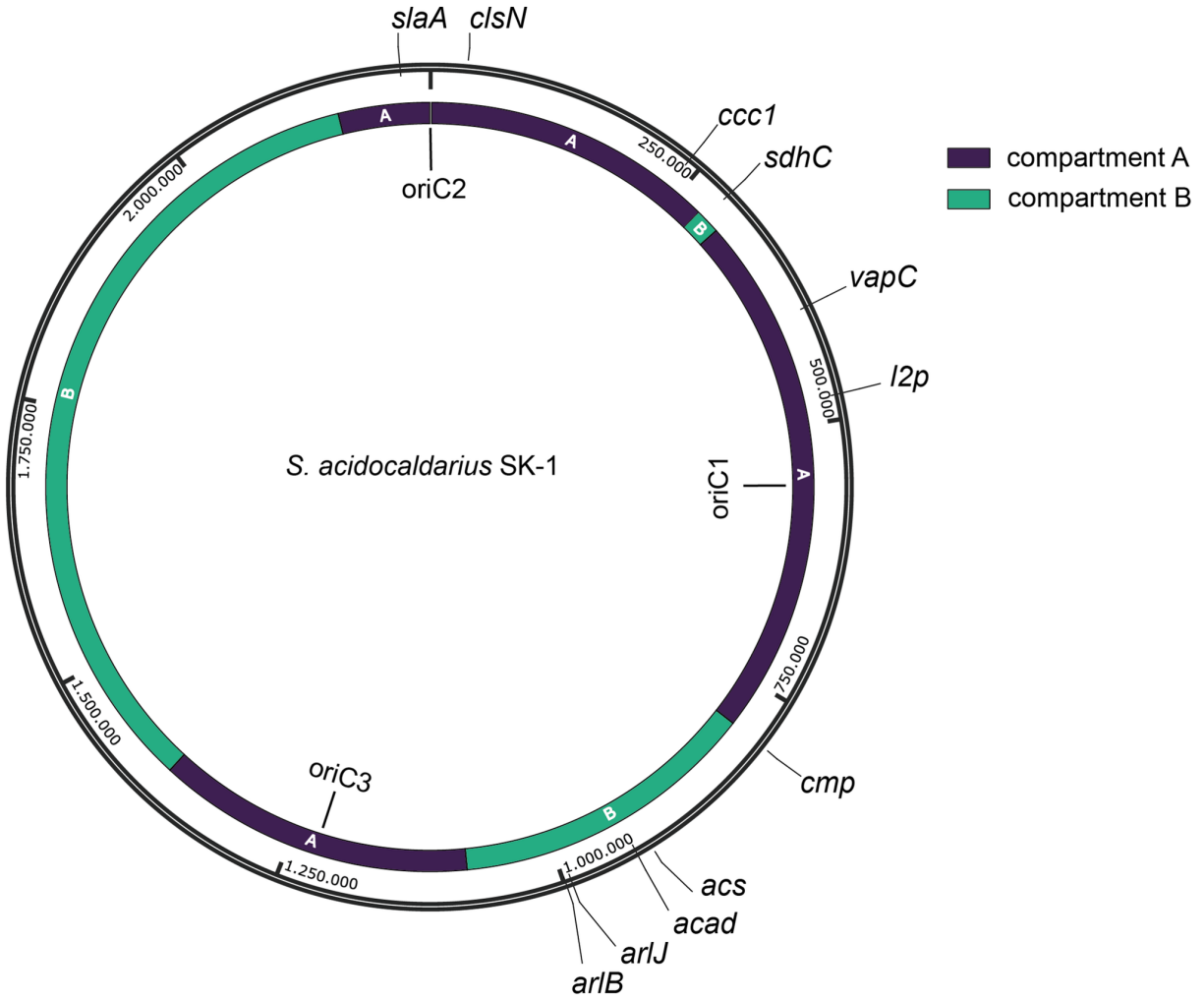

FIGURE 2
Map of the *S. acidocaldarius* SK-1 genome sequence with indication of genomic position, replication origins (oriC1, oriC2 and oriC3) (Duggin et al., 2008) and target sites (with naming according to Table 1). Chromosome compartments A and B (Takemata et al., 2019) are color-indicated. Assignment of compartments was performed on a gene-by-gene basis, which was possible given the high similarity between the genome sequences of *S. acidocaldarius* SK-1 and DSM639 strains.

(e.g., *acad*). The insertion positions and corresponding knock-in strains were named based on the nearest gene (Table 1).

All genetic constructs were transformed into *S. acidocaldarius* SK-1 followed by a selection in uracil-free medium, yielding high numbers of colony-forming units (CFUs). Upon X-gal treatment, enabling colorimetric detection of *lacS*$^+$ colonies, it appeared that only a small fraction of CFUs represented "pop-in" *lacS*$^+$ integrants (Figure 4a). The nature of the "background" *lacS*$^-$ CFUs is unknown. On average, the fraction of "pop-in" *lacS*$^+$ CFUs was higher for locations in compartment A as compared to compartment B (an average of 0.56% *versus* 0.22%) (Figure 4a). Transformations that yielded a higher "pop-in" *lacS*$^+$ integrant fraction, such as for *slaA* and *ccc1* knock-in constructs, were possibly characterized by more a favorable genomic recombination as compared to transformations with knock-in constructs like *acad* or *arlJ*.

Purified "pop-in" integrant strains were subsequently cultivated on 5-FOA- and uracil-containing plates to select for "pop-out" *lacS*$^+$ integrant strains, which still contained the *lacS* cassette, but with the remainder of the vector removed by a crossover event (Figure 1b). For all knock-in constructions, with the exception of *sdhC* and *cmp*, "pop-out" *lacS*$^+$ integrant strains were easily obtained and the correct integration of the *lacS* cassette was verified by colony PCR (Figure 4b). In contrast, for the construction of the *acs* and *cmp* knock-in strains, colony PCR revealed partial or full presence of the wildtype genotype and despite multiple selection attempts, we failed in obtaining these knock-in strains. For the *cmp* knock-in strain this could be possibly attributed to disruption of expression of the *cmp* gene itself, which is a highly essential gene (Table 1).

3.3 Transcriptional gene expression analysis of *lacS* in the different knock-in strains

A qRT-PCR approach was used to evaluate transcriptional expression levels of the *lacS* reporter gene in the different knock-in strains (Figure 5). Although all strains were cultivated in identical

TABLE 1 Overview of the 11 insertion positions, with indication of name, nearest origin of replication, distance to nearest origin of replication, chromosomal position, chromosomal compartment, as well as gene number, annotation, expression level (Exp. level) and predicted essentiality of the nearest gene (Zhang et al., 2018).

Name	Nearest oriC	Distance to oriC (bp)	Chromosomal position	Chromosomal compartment	Gene number	Exp. level[a]	Predicted essentiality[b]
clsN	oriC2	32,139	32,281	A	*Saci_0046*	1.955	3
ccc1	oriC2	234,430	235,243	A	*Saci_0278*	55.012	13
sdhC	oriC1	278,118	278,353	B	*Saci_0325*	129.5	20
vapC	oriC1	162,379	394,047	A	*Saci_0467*	1.47	23
l2p	oriC1	83,191	473,288	A	*Saci_0594*	53.09	2
cmp	oriC1	231,169	789,249	A	*Saci_0985*	11.012	0
acs	oriC3	311,117	914,836	B	*Saci_1111*	53.76	18
acad	oriC3	297,343	928,402	B	*Saci_1123*	56.12	43
arlJ	oriC3	238,551	987,911	B	*Saci_1172*	18.05	N. D.
arlB	oriC3	230,907	993,346	B	*Saci_1178*	35.18	11
slaA	oriC2	26,044	2,199,897	A	*Saci_2355*	11.554	7

All information in this table is based on alignment with the *S. acidocaldarius* DSM639 genome sequence (Chen et al., 2005).
[a]Expression level is expressed in CPM value (Baes et al., 2023). N. D. indicates not determined.
[b]Predicted essentiality is expressed on a scale between 0 and 48 with the former the highest essentiality.

conditions and despite *lacS* being under control of the same constitutive promoter, large variations in transcriptional levels were observed (Figure 5a). Relative transcription is calculated with respect to the knock-in strain displaying the lowest *lacS* transcription, *arlJ*.

Transcriptional levels were observed to be highest in the *slaA* knock-in strain, 55-fold higher than in the *arlJ* knock-in strain. This could be attributed to the target site position, immediately downstream of the *slaA* stop codon, making it highly probable that *slaA* transcription, which is known to be constitutively expressed at high level (Baes et al., 2023), runs through to the *lacS* cassette. The same applies for other mutant strains that display significantly higher transcription levels with respect to the *arlJ* knock-in strain (Figure 5a). For the *vapC* knock-in strain, in which *lacS* is transcribed at a 6.2-fold higher level, the *lacS* cassette was integrated downstream of the *vapC* gene preceding a putative terminator site (Figure 5a; Supplementary Figure S3). For the *ccc1* knock-in strain, in which *lacS* is transcribed 5.2-fold higher than in the *arlJ* knock-in strain, the target site is located in between two operonic genes (Figure 5a; Supplementary Figure S3).

Only in the *clsN* and *acad* knock-in strains, the target sites are located in regions where no read-through transcription from adjacent genomically encoded transcription units is expected, namely in an intergenic region in between the two divergently oriented putative promoters (Figure 3; Supplementary Figure S3). For these strains, transcriptional expression of *lacS* is expected to be driven solely by the P_{Ec10_Sa1} promoter and indeed, both strains have very similar transcriptional levels with respect to the *arlJ* knock-in strain, namely 1.88-fold and 1.86-fold for *clsN* and *acad*, respectively (Figure 5a).

Although a trend was observed of compartment A-targeted knock-in strains having higher transcriptional expression levels than compartment B-targeted knock-in strains, this cannot be linked to a global effect, but rather to the expression levels of adjacent genes, which are not insulated from the *lacS* cassette. To a certain extent, a correlation was observed between previously detected expression levels of the adjacent gene (Baes et al., 2023) expected to cause read-through transcription and the relative *lacS* transcriptional level (Supplementary Table S4). Given that for the two "insulated" knock-in strains *clsN* and *acad*, which both have similar transcriptional expression levels and of which the *clsN* target site is located in compartment A and the *acad* target site in compartment B, a global gene silencing effect was not apparent.

3.4 Transcriptional regulation of *lacS* in the different knock-in strains

Transcriptional expression of *lacS* was also monitored in response to stress conditions, either nutritional starvation (Figure 5b) or stationary phase growth (Figure 5c). Although the P_{Ec10_Sa1} promoter is expected to drive transcription of *lacS* in a constitutive manner, differential transcription was observed in several knock-in strains, either positively or negatively. Trends are similar for both nutritional starvation and stationary growth phase. In the *slaA* and *vapC* knock-in strains, a transcriptional downregulation of *lacS* is observed, while in the *sdhC* and *arlB* knock-in strain, an upregulation is observed (Figures 5b,c). Upon comparing the regulatory effects on *lacS* expression with the differential transcriptional patterns of nearby genes observed in previous transcriptomic datasets under nutritional starvation and stationary phase conditions (Supplementary Tables S5, S6) (Bischof et al., 2018; Takemata et al., 2019), indications of read-through effects were found in some but not all "non-insulated" knock-in strains. For example, in the *arlB* knock-in strain, the cassette is inserted directly downstream of the *arlB* ORF (*saci_1178*) (Figure 3; Supplementary Figure S3) and an upregulation of *lacS* is observed (Figures 5b,c), consistent with the known upregulation of *arlB* under nutritional starvation and stationary phase conditions (Supplementary Tables S5, S6). A similarly clear correlation was found for the *ccc1* and *sdhC* knock-in strains. In contrast, for the *slaA*, *vapC*, *l2p*, and *arlJ* knock-in strains differential expression of *lacS* was not entirely correlated with differential expression of adjacent genes expected to cause read-through expression in both transcriptomic datasets (Figures 5b,c, Supplementary Tables S5, S6).

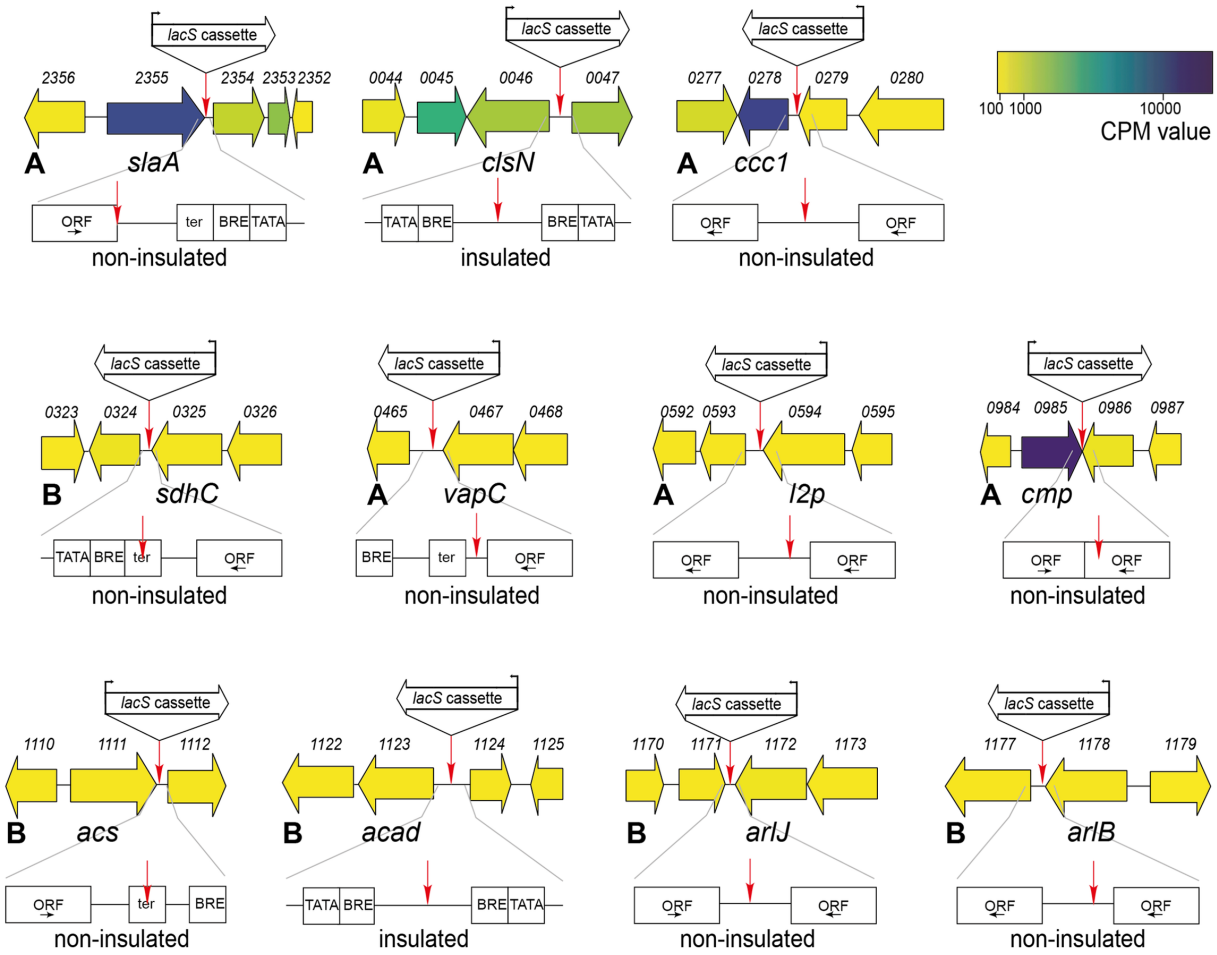

FIGURE 3
Genetic organization in the genomic environment of the 11 target sites with indication of the target site and the orientation of the *lacS* cassette. Numbers refer to gene locus tags, with xxxx referring to *Saci_xxxx*. Gene arrows are color-coded according to expression levels (CPM value) (Baes et al., 2023). For each knock-in strain, a schematic representation is given by the location of the target site with respect to the ORFs and transcriptional elements. ORFs are boxed with the direction of transcription indicated by an arrow. Putative transcriptional elements are indicated with TATA = TATA box, BRE = factor B recognition element and ter = terminator element. Please refer to Supplementary Figure S3 for more information on the sequence. For each knock-in strain, the corresponding chromatin compartment (A or B) is indicated, along with a prediction of whether the *lacS* cassette is insulated or not.

In almost all cases, there was a lack of correlation in differential expression observed for the gene of interest either under nutritional starvation or in stationary phase growth (Supplementary Tables S5, S6). Moreover, in the *arlJ* knock-in strain, in which the *lacS* cassette is integrated into the intergenic region between the converging 3′ ends of the *arlJ* (*saci_1172*) and *arnR1* (*saci_1171*) genes (Lassak et al., 2013) (Figure 3), it is unclear which of the two genes, if any, *lacS* expression would be expected to correlate with, especially since no obvious transcriptional terminator sequences were identified within this region (Supplementary Figure S3).

With exception of *lacS* expression in the *acad* knock-in strain under nutritional starvation, for which an upregulation was observed, little or no transcriptional regulation was observed for the "insulated" knock-in strains *clsN* and *acad* (Figures 5b,c). This insulation is further strengthened by the observation that several adjacent genes (*saci_0047* for *clsN* and *saci_1123* and *saci_1124* for *acad*) display transcriptional regulation under nutritional starvation and stationary phase conditions (Supplementary Tables S5, S6).

3.5 Translational level and activity analysis of LacS in the different knock-in strains

To assess the abundance and activity of LacS, which is expressed as a His-tagged protein, on the protein level, we performed anti-6xHis western blotting and ONPG-based LacS activity assays (Figure 6). Both approaches demonstrated a variability among the knock-in strains with compartment A strains showing a relatively higher average activity as compared to compartment B strains (an average of 2.412×10^{-4} versus 1.419×10^{-4}) (Figure 6b), with the *clsN* and *slaA* knock-in strains seemingly exhibiting the highest activity, together with the *vapC* knock-in strain.

An apparent higher translational activity in compartment A versus compartment B knock-in strains might be, at least partially, reflecting differences in transcriptional activities, which follow the same trend (Figure 5a). The observed differences in translational activity for the different strains were less pronounced as those observed at the transcriptional level. For instance, while compartment

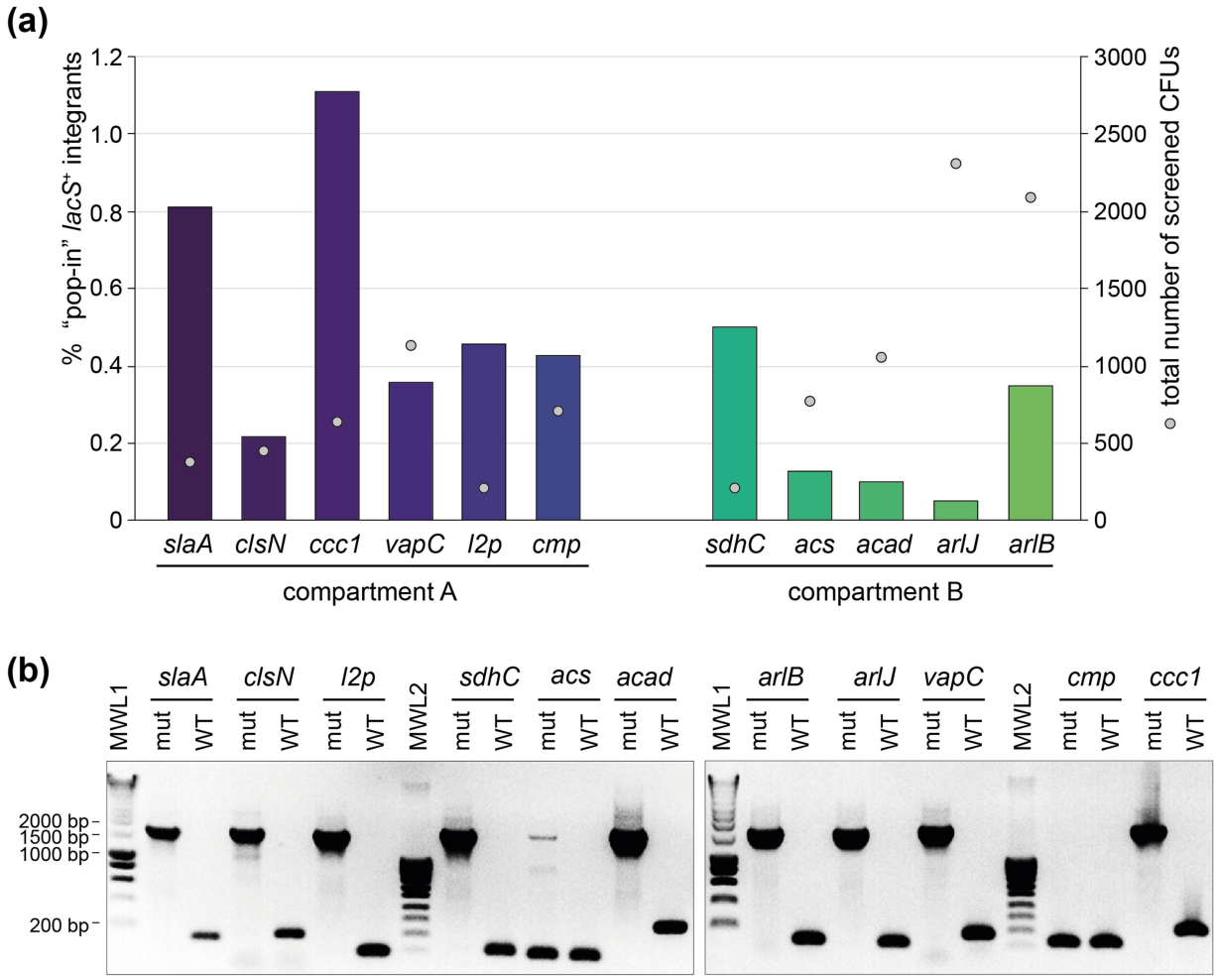

FIGURE 4
Construction of the *lacS*$^{+}$ knock-in mutant strains. **(a)** Percentage of "pop-in" *lacS*$^{+}$ integrant CFUs after the "pop-in" construction step, depicted in bars for each of the intended knock-in strains (Table 1), grouped according to their chromosome compartment. This percentage is based on X-gal colorimetric screening and expressed with respect to "background" CFUs. Sphere symbols indicate the total number of screened CFUs for each of the knock-in constructions. **(b)** Colony PCR of selected "pop-out" *lacS*$^{+}$ integrant strains to verify integration of the *lacS* cassette, leading to an amplicon of approximately 1,700 bp. The name of the knock-in strain under construction is indicated above each lane. For each construct, the sequences of the forward and reverse primers, which hybridize to the genome just upstream and downstream of the target integration site, are provided in Supplementary Table S2. WT, wild type; mut, "pop-out" *lacS*$^{+}$ integrant; MWL, molecular weight ladder.

A strains, such as the *clsN* and *slaA* knock-in strains, exhibited slightly higher protein activities as compared to compartment B strains, such as the *arlJ* knock-in strain, the variation in protein activity was relatively modest. Moreover, the correlation between transcriptional and translational levels was very limited with a R^2 of 0.1296 and a lack of statistical significance (p value = 0.3413) (Figure 7a). Interestingly, a somewhat larger correlation can be found for protein activity and distance to the closest origin of replication (R^2 = 0.432) than for transcriptional expression (R^2 = 0.3471), although with low statistical significance (p values of 0.0543 and 0.0951, respectively) (Figures 7b,c).

4 Discussion

In conclusion, our study indicates that while genomic location can influence gene expression levels in *S. acidocaldarius*, the evidence for a strong position-dependent effect independent of local genomic context is limited, especially at the transcriptional level. Instead, transcriptional variability can be largely attributed to the transcriptional activity of neighboring genes rather than to the genomic position itself for most of the chosen target positions. Indeed, these positions did not insulate the *lacS* cassette from surrounding transcriptional influences, leading to expression levels that were primarily dictated by local transcriptional context. This can be explained by the high coding density in the *S. acidocaldarius* genome, with 2,292 protein-encoding genes predicted for 2,225,959 bp (Chen et al., 2005). Therefore, while almost all target sites are in intergenic regions, only in the *clsN* and *acad* knock-in strains these target sites can be assumed not to be part of a transcriptional unit. Moreover, due to a lack of the availability of genetic elements for archaea, for example a transcriptional terminator, we could not integrate an insulator element into the reporter gene cassette as is typically done in bacteria (Scholz et al., 2022).

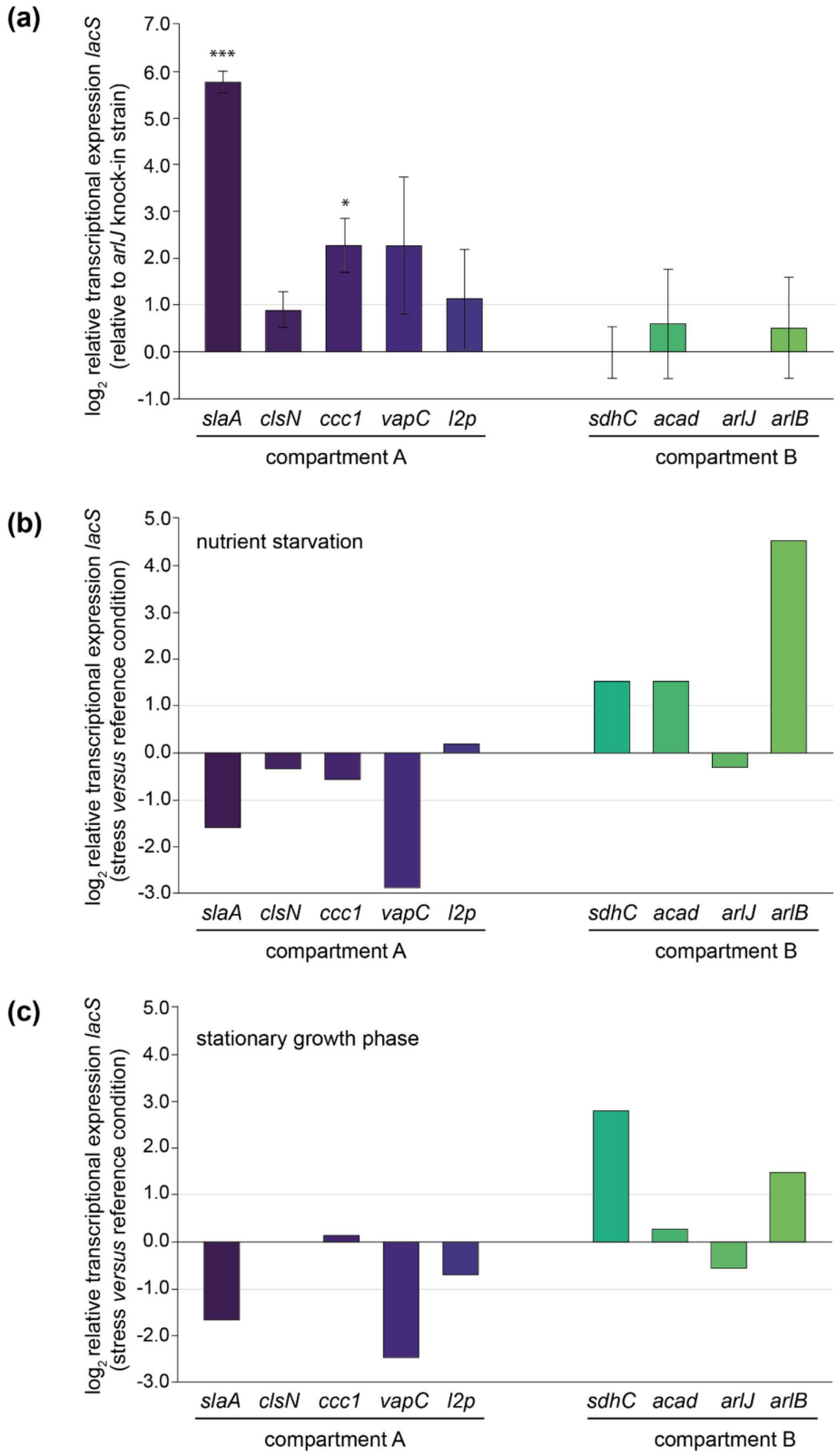

FIGURE 5
Relative transcriptional expression levels of *lacS* as monitored by qRT-PCR analysis. **(a)** Relative transcriptional *lacS* expression in each of the knock-in strains relative to the *arlJ* knock-in strain, grouped according to the chromosomal compartment. Experiments have been performed in biological triplicates with error bars representing standard deviations. Statistically significant differential expression was determined by a one-sample *T*-test (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$). **(b)** Relative transcriptional *lacS* expression in nutrient starvation conditions *versus* the reference growth condition in each of the knock-in strains. **(c)** Relative transcriptional *lacS* expression in stationary *versus* exponential growth phase.

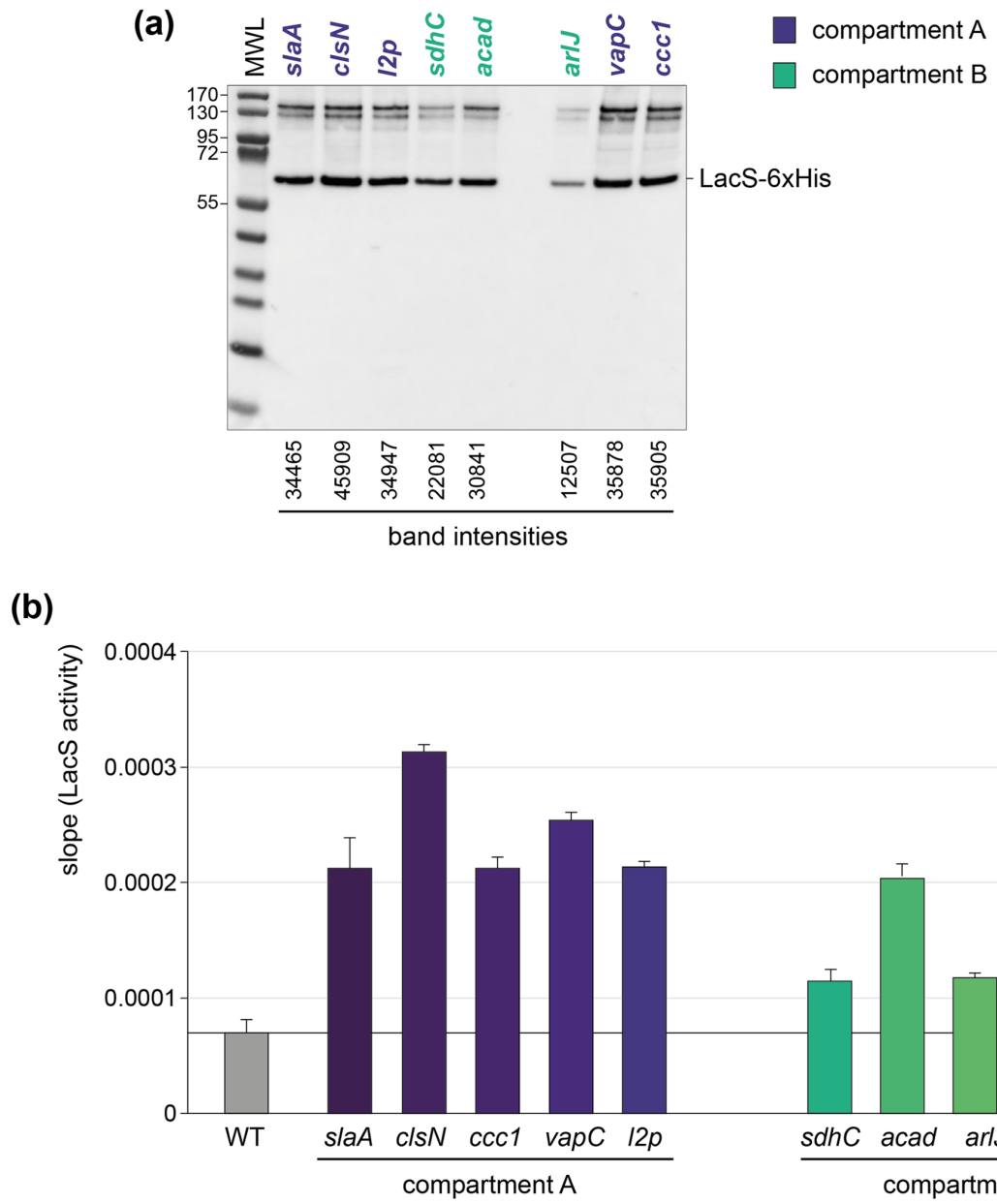

FIGURE 6
Translational and activity levels of LacS. **(a)** Western blot analysis employing anti-6xHis antibodies for the detection of LacS-6xHis in the different knock-in strains. The *arlB* knock-in strain was not included in this analysis because of a mutation in the His-tag-encoding sequence. MWL = molecular weight ladder. Protein molecular weights are indicated in kDa. For the bands corresponding to LacS-6xHis, intensities were determined with ImageJ. **(b)** LacS activities as measured with an ONPG assay for the different knock-in strains. More information on how raw LacS activity data are processed is provided in Supplementary Figure S4.

The strong influence of the activity of adjacent transcription units is not only apparent for constitutive expression levels, but also for transcription regulation. For most knock-in strains, similar regulatory trends were observed for two different stress conditions: nutritional starvation and stationary phase growth. Moreover, the regulatory effect corresponds to that of the adjacent gene in several cases. For example, this is the case for the *arlB* knock-in strain. The *arlB* gene encodes the main structural unit of the archaellum and is part of a larger operon that encodes additional structural and functional proteins of this motility structure (Lassak et al., 2012). This operon is known to be transcriptionally upregulated in response to stress conditions such as nutrient limitation, through a complex interplay of different transcription regulators and kinases (Lassak et al., 2013; Haurat et al., 2017; Li et al., 2017; Bischof et al., 2018). By incorporating the reporter gene cassette directly downstream of the *arlB* ORF, it also appears to be subject to this regulation. This was most clearly observed in response to nutritional starvation, but also in response to stationary phase growth, for which an upregulation of *arlB* was observed in a previous transcriptomic study (Takemata et al., 2019).

There was a lack of a clear correlation between transcriptional and translational levels. This suggests that translational efficiency and LacS activity may be influenced by factors beyond transcription alone, such

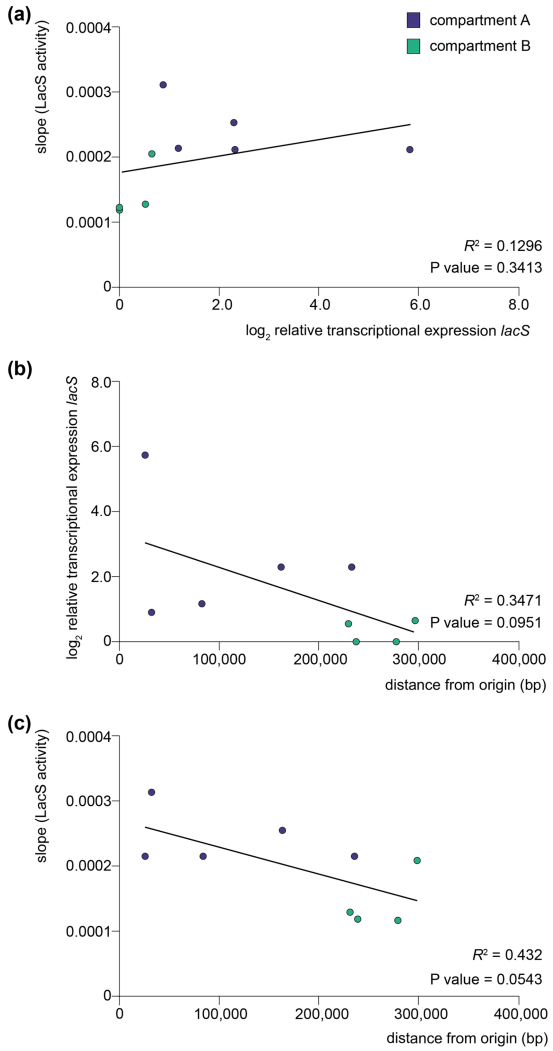

FIGURE 7
Correlation between transcription, translation and distance to origin of replication. **(a)** Scatter plot between translational levels as monitored by LacS activities and relative transcriptional expression levels of *lacS* as monitored by qRT-PCR analysis. **(b)** Scatter plot between relative transcriptional expression levels of *lacS* as monitored by qRT-PCR analysis and distance to the closest origin of replication. **(c)** Scatter plot between translational levels as monitored by LacS activities and distance to the closest origin of replication. For each plot, a linear regression line has been fitted and the coefficient of determination (R^2) and two-tailed *p* value is displayed. Data points are colored according to the corresponding compartment (A or B).

as mRNA stability and ribosome accessibility. This lack of correlation between transcription and translation was previously observed for dynamic effects in response to nutrient limitation (Bischof et al., 2018) and heat shock (Baes et al., 2023), indicating a prevalence of post-transcriptional and post-translational mechanisms in *S. acidocaldarius*. In a similar study in *S. islandicus* employing a *lacS* reporter gene cassette, significant positional effects on translation level were also observed (Boob et al., 2025).

Several of the ectopic integration sites explored in our study could be employed in the future for the construction of knock-in strains, whether it is for the generation of strains with altered phenotypic characteristics or for high-level production of recombinant protein. It should be noted that in our study, the *lacS* ORF was directly fused to a promoter, without including a 5′-untranslated region (5'-UTR). Although the presence of a Shine-Dalgarno (SD)-containing 5′-UTR sequence might be considered less relevant for gene expression, given the prevalent occurrence of leaderless transcripts in archaea (Wurtzel et al., 2010; Schmitt et al., 2020), it has been shown that the presence of a 5′-UTR in plasmid-based heterologous expression constructs in Sulfolobales can significantly impact the gene expression level by affecting both transcriptional and translational efficiency (Peng et al., 2009, Ao et al., 2013, Kuschmierz et al., 2024). More specifically, the integration of a 5'-UTR derived from the Alba-encoding gene *saci_1322* from *S. acidocaldarius* harboring a SD sequence in a heterologous protein expression plasmid led to a significantly higher protein production (Kuschmierz et al., 2024). Therefore, if the goal is to obtain a maximal yield of recombinant protein production in a stable *S. acidocaldarius* platform, it could be advised to integrate the *alba*-derived 5′-UTR into an expression cassette ectopically integrated into the genome target site just downstream of the *slaA* ORF.

While our study focused on identifying intergenic loci that enable stable gene integration and its heterologous expression, we acknowledge that not all sites may be completely neutral and that polar effects might be present for adjacent genes. Although for the knock-in strains that were successfully obtained in this study, the integration did not appear to impact strain viability or growth under standard conditions, a detailed phenotypic characterization was not performed. It remains possible that transcriptional or translational interference could alter expression of up- or downstream genes. Future studies involving transcriptomic or proteomic profiling could help assess these effects and further validate site neutrality.

Data availability statement

All data presented in this study can be found in either the Supplementary material or the online dataset 10.5281/zenodo.16025639.

Author contributions

YX: Conceptualization, Formal analysis, Investigation, Writing – original draft. AP: Formal analysis, Investigation, Writing – review & editing. IB: Supervision, Writing – review & editing. MM: Resources, Supervision, Writing – review & editing. RB: Formal analysis, Supervision, Writing – review & editing. EP: Conceptualization, Resources, Supervision, Visualization, Writing – original draft.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This research was funded by the Vrije Universiteit Brussel (Strategic Research Program SRP91), by Flanders Innovation and Entrepreneurship (VLAIO) (Moonshot project "TACBIO" [HBC.2020.2618]), by the Bijzonder Onderzoeksfonds (iBOF project "POSSIBL" [iBOF/21/092]) and by

the Research Foundation Flanders (FWO-Vlaanderen) [G012323N]. YX was funded by a PhD scholarship from the Chinese Scholarship Council.

Acknowledgments

We would like to thank Norio Kurosawa for the generous gift of the *S. acidocaldarius* SK-1 strain and Karl Jonckheere for technical assistance.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2025.1602937/full#supplementary-material

References

Albers, S. V., and Driessen, A. J. (2008). Conditions for gene disruption by homologous recombination of exogenous DNA into the *Sulfolobus solfataricus* genome. *Archaea* 2, 145–149. doi: 10.1155/2008/948014

Ao, X., Li, Y., Wang, F., Feng, M., Lin, Y., Zhao, S., et al. (2013). The Sulfolobus initiator element is an important contributor to promoter strength. *J. Bacteriol.* 195, 5216–522. doi: 10.1128/JB.00768-13

Baes, R., Grünberger, F., Pyr Dit Ruys, S., Couturier, M., De Keulenaer, S., Skevin, S., et al. (2023). Transcriptional and translational dynamics underlying heat shock response in the thermophilic crenarchaeon *Sulfolobus acidocaldarius*. *MBio* 14:e0359322. doi: 10.1128/mbio.03593-22

Baumann, H., Knapp, S., Lundbäck, T., Ladenstein, R., and Härd, T. (1994). Solution structure and DNA-binding properties of a thermostable protein from the archaeon *Sulfolobus solfataricus*. *Nature* 1, 638–653.

Beckwith, J. R., Signer, E. R., and Epstein, W. (1966). Transposition of the lac region of *E. coli*. *Cold Spring Harb. Symp. Quant. Biol.* 31, 393–401. doi: 10.1101/SQB.1966.031.01.051

Bell, S. D., Botting, C. H., Wardleworth, B. N., Jackson, S. P., and White, M. F. (2002). The interaction of Alba, a conserved archaeal chromatin protein, with Sir2 and its regulation by acetylation. *Science* 296, 148–151. doi: 10.1126/science.1070506

Berkner, S., Grogan, D., Albers, S.-V., and Lipps, G. (2007). Small multicopy, non-integrative shuttle vectors based on the plasmid pRN1 for *Sulfolobus acidocaldarius* and *Sulfolobus solfataricus*, model organisms of the (cren-)archaea. *Nucleic Acids Res.* 35:e88. doi: 10.1093/nar/gkm449

Bischof, L. F., Haurat, M. F., Hoffmann, L., Albersmeier, A., Wolf, J., Neu, A., et al. (2018). Early response of *Sulfolobus acidocaldarius* to nutrient limitation. *Front. Microbiol.* 9:3201. doi: 10.3389/fmicb.2018.03201

Block, D. H., Hussein, R., Liang, L. W., and Lim, H. N. (2012). Regulatory consequences of gene translocation in bacteria. *Nucleic Acids Res.* 40, 8979–8992. doi: 10.1093/nar/gks694

Bonev, B., and Cavalli, G. (2016). Organization and function of the 3D genome. *Nat. Rev. Genet.* 17, 661–678. doi: 10.1038/nrg.2016.112

Boob, A. G., Zhang, C., Pan, Y., Zaidi, A., Whitaker, R. J., and Zhao, H. (2025). Discovery, characterization, and application of chromosomal integration sites in the hyperthermoacidophilic crenarchaeon *Sulfolobus islandicus*. *bioRxiv*. doi: 10.1101/2025.03.16.643552

Brambilla, E., and Sclavi, B. (2015). Gene regulation by H-NS as a function of growth conditions depends on chromosomal position in *Escherichia coli*. *G3 (Bethesda, Md)* 5, 605–614. doi: 10.1534/g3.114.016139

Brock, T. D., Brock, K. M., Belly, R. T., and Weiss, R. L. (1972). *Sulfolobus*: a new genus of sulfur-oxidizing bacteria living at low pH and high temperature. *Arch. Mikrobiol.* 84, 54–68. doi: 10.1007/BF00408082

Bryant, J. A., Sellars, L. E., Busby, S. J., and Lee, D. J. (2014). Chromosome position effects on gene expression in *Escherichia coli* K-12. *Nucleic Acids Res.* 42, 11383–11392. doi: 10.1093/nar/gku828

Chen, L., Brügger, K., Skovgaard, M., Redder, P., She, Q., Torarinsson, E., et al. (2005). The genome of *Sulfolobus acidocaldarius*, a model organism of the Crenarchaeota. *J. Bacteriol.* 187, 4992–4999. doi: 10.1128/JB.187.14.4992-4999.2005

Choli, T., Henning, P., Wittmann-Liebold, B., and Reinhardt, R. (1988). Isolation, characterization and microsequence analysis of a small basic methylated DNA-binding protein from the Archaebacterium, *Sulfolobus solfataricus*. *Biochim. Biophys. Acta* 950, 193–203. doi: 10.1016/0167-4781(88)90011-5

Cooke, K., Browning, D. F., Lee, D. J., Blair, J. M. A., McNeill, H. E., Huber, D., et al. (2019). Position effects on promoter activity in *Escherichia coli* and their consequences for antibiotic-resistance determinants. *Biochem. Soc. Trans.* 47, 839–845. doi: 10.1042/BST20180503

Crosby, J. R., Laemthong, T., Lewis, A. M., Straub, C. T., Adams, M. W., and Kelly, R. M. (2019). Extreme thermophiles as emerging metabolic engineering platforms. *Curr. Opin. Biotechnol.* 59, 55–64. doi: 10.1016/j.copbio.2019.02.006

Driessen, R. P. C., Lin, S. N., Waterreus, W. J., Van Der Meulen, A. L. H., Van Der Valk, R. A., Laurens, N., et al. (2016). Diverse architectural properties of Sso10a proteins: evidence for a role in chromatin compaction and organization. *Sci. Rep.* 6, 1–11. doi: 10.1038/srep29422

Duggin, I. G., McCallum, S. A., and Bell, S. D. (2008). Chromosome replication dynamics in the archaeon *Sulfolobus acidocaldarius*. *Proc. Natl. Acad. Sci. USA* 105, 16737–16742. doi: 10.1073/pnas.0806414105

Engler, C., Kandzia, R., and Marillonnet, S. (2008). A one pot, one step, precision cloning method with high throughput capability. *PLoS One* 3:e3647. doi: 10.1371/journal.pone.0003647

Gay, P., Le Coq, D., Steinmetz, M., Berkelman, T., and Kado, C. I. (1985). Positive selection procedure for entrapment of insertion sequence elements in gram-negative bacteria. *J. Bacteriol.* 164, 918–921. doi: 10.1128/jb.164.2.918-921.1985

Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A. 3rd, and Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* 6, 343–345. doi: 10.1038/nmeth.1318

Grote, M., Dijk, J., and Reinhardt, R. (1986). Ribosomal and DNA binding proteins of the thermoacidophilic archaebacterium *Sulfolobus acidocaldarius*. *Biochim. Biophys. Acta* 873, 405–413. doi: 10.1016/0167-4838(86)90090-7

Guo, L., Feng, Y., Zhang, Z., Yao, H., Luo, Y., Wang, J., et al. (2008). Biochemical and structural characterization of Cren7, a novel chromatin protein conserved among Crenarchaea. *Nucleic Acids Res.* 36, 1129–1137. doi: 10.1093/nar/gkm1128

Haurat, M. F., Figueiredo, A. S., Hoffmann, L., Li, L., Herr, K., A, J. W., et al. (2017). ArnS, a kinase involved in starvation-induced archaellum expression. *Mol. Microbiol.* 103, 181–194. doi: 10.1111/mmi.13550

Kalichuk, V., Béhar, G., Renodon-Cornière, A., Danovski, G., Obal, G., Barbet, J., et al. (2016). The archaeal "7 kDa DNA-binding" proteins: extended characterization of an old gifted family. *Sci. Rep.* 6, 1–10. doi: 10.1038/srep37274

Kurosawa, N., and Grogan, D. W. (2005). Homologous recombination of exogenous DNA with the *Sulfolobus acidocaldarius* genome: properties and uses. *FEMS Microbiol. Lett.* 253, 141–149. doi: 10.1016/j.femsle.2005.09.031

Kuschmierz, L., Wagner, A., Schmerling, C., Busche, T., Kalinowski, J., Bräsen, C., et al. (2024). 5′-untranslated region sequences enhance plasmid-based protein production in *Sulfolobus acidocaldarius*. *Front. Microbiol.* 15:1443342. doi: 10.3389/fmicb.2024.1443342

Lassak, K., Neiner, T., Ghosh, A., Klingl, A., Wirth, R., and Albers, S.-V. (2012). Molecular analysis of the crenarchaeal flagellum. *Mol. Microbiol.* 83, 110–124. doi: 10.1111/j.1365-2958.2011.07916.x

Lassak, K., Peeters, E., Wrobel, S., and Albers, S.-V. (2013). The one-component system ArnR: a membrane-bound activator of the crenarchaeal archaellum. *Mol. Microbiol.* 88, 125–139. doi: 10.1111/mmi.12173

Lee, A., Jin, H., and Cha, J. (2022). Engineering of *Sulfolobus acidocaldarius* for Hemicellulosic biomass utilization. *J. Microbiol. Biotechnol.* 32, 663–671. doi: 10.4014/jmb.2202.02016

Lemmens, L., Wang, K., Ruykens, E., Nguyen, V. T., Lindås, A. C., Willaert, R., et al. (2022). DNA-binding properties of a novel crenarchaeal chromatin-organizing protein in *Sulfolobus acidocaldarius*. *Biomol. Ther.* 12:524. doi: 10.3390/biom12040524

Li, L., Banerjee, A., Bischof, L. F., Maklad, H. R., Hoffmann, L., Henche, A. L., et al. (2017). Wing phosphorylation is a major functional determinant of the Lrs14-type biofilm and motility regulator AbfR1 in *Sulfolobus acidocaldarius*. *Mol. Microbiol.* 105, 777–793. doi: 10.1111/mmi.13735

Lieberman-Aiden, E., van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., et al. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326, 289–293. doi: 10.1126/science.1181369

Liu, H., Orell, A., Maes, D., van Wolferen, M., Lindås, A. C., Bernander, R., et al. (2014). BarR, an Lrp-type transcription factor in Sulfolobus acidocaldarius, regulates an aminotransferase gene in a β-alanine responsive manner. *Mol. Microbiol.* 92, 625–639. doi: 10.1111/mmi.12583

López-García, P., Knapp, S., Ladenstein, R., and Forterre, P. (1998). *In vitro* DNA binding of the archaeal protein Sso7d induces negative supercoiling at temperatures typical for thermophilic growth. *Nucleic Acids Res.* 26, 2322–2328. doi: 10.1093/nar/26.10.2322

Lurz, R., Grote, M., Dijk, J., Reinhardt, R., and Dobrinski, B. (1986). Electron microscopic study of DNA complexes with proteins from the Archaebacterium *Sulfolobus acidocaldarius*. *EMBO J.* 5, 3715–3721. doi: 10.1002/j.1460-2075.1986.tb04705.x

Peng, N., Xia, Q., Chen, Z., Liang, Y. X., and She, Q. (2009). An upstream activation element exerting differential transcriptional activation on an archaeal promoter. *Mol. Microbiol.* 74, 928–939. doi: 10.1111/j.1365-2958.2009.06908.x

Pilatowski-Herzing, E., Samson, R. Y., Takemata, N., Badel, C., Bohall, P. B., and Bell, S. D. (2025). Capturing chromosome conformation in Crenarchaea. *Mol. Microbiol.* 123, 101–108. doi: 10.1111/mmi.15245

Reilly, M. S., and Grogan, D. W. (2001). Characterization of intragenic recombination in a hyperthermophilic archaeon via conjugational DNA exchange. *J. Bacteriol.* 183, 2943–2946. doi: 10.1128/JB.183.9.2943-2946.2001

Rowley, M. J., and Corces, V. G. (2018). Organizational principles of 3D genome architecture. *Nat. Rev. Genet.* 19, 789–800. doi: 10.1038/s41576-018-0060-8

Schmitt, E., Coureux, P. D., Kazan, R., Bourgeois, G., Lazennec-Schurdevin, C., and Mechulam, Y. (2020). Recent advances in archaeal translation initiation. *Front. Microbiol.* 11:584152. doi: 10.3389/fmicb.2020.584152

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. doi: 10.1038/nmeth.2089

Scholz, S. A., Lindeboom, C. D., and Freddolino, L. (2022). Genetic context effects can override canonical cis regulatory elements in *Escherichia coli*. *Nucleic Acids Res.* 50, 10360–10375. doi: 10.1093/nar/gkac787

Sousa, C., de Lorenzo, V., and Cebolla, A. (1997). Modulation of gene expression through chromosomal positioning in *Escherichia coli*. *Microbiology* 143, 2071–2078. doi: 10.1099/00221287-143-6-2071

Suzuki, S., and Kurosawa, N. (2016). Disruption of the gene encoding restriction endonuclease SuaI and development of a host-vector system for the thermoacidophilic archaeon *Sulfolobus acidocaldarius*. *Extremophiles* 20, 139–148. doi: 10.1007/s00792-016-0807-0

Suzuki, S., and Kurosawa, N. (2017). Development of the multiple gene knockout system with one-step PCR in thermoacidophilic crenarchaeon *Sulfolobus acidocaldarius*. *Archaea* 2017:7459310. doi: 10.1155/2017/7459310

Takemata, N., and Bell, S. D. (2021). Multi-scale architecture of archaeal chromosomes. *Mol. Cell* 81, 473–487.e6. doi: 10.1016/j.molcel.2020.12.001

Takemata, N., Samson, R. Y., and Bell, S. D. (2019). Physical and functional compartmentalization of archaeal chromosomes. *Cell* 179, 165–179.e18. doi: 10.1016/j.cell.2019.08.036

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., et al. (2012). Primer3--new capabilities and interfaces. *Nucleic Acids Res.* 40:e115. doi: 10.1093/nar/gks596

Van der Kolk, N., Wagner, A., Wagner, M., Waßmer, B., Siebers, B., and Albers, S. V. (2020). Identification of XylR, the activator of arabinose/xylose inducible regulon in *Sulfolobus acidocaldarius* and its application for homologous protein expression. *Front. Microbiol.* 11:1066. doi: 10.3389/fmicb.2020.01066

Wagner, M., Berkner, S., Ajon, M., Driessen, A. J., Lipps, G., and Albers, S. V. (2009). Expanding and understanding the genetic toolbox of the hyperthermophilic genus *Sulfolobus*. *Biochem. Soc. Trans.* 37, 97–101. doi: 10.1042/BST0370097

Wagner, M., van Wolferen, M., Wagner, A., Lassak, K., Meyer, B. H., Reimann, J., et al. (2012). Versatile genetic tool box for the Crenarchaeote *Sulfolobus acidocaldarius*. *Front. Microbiol.* 3:214. doi: 10.3389/fmicb.2012.00214

Wurtzel, O., Sapra, R., Chen, F., Zhu, Y., Simmons, B. A., and Sorek, R. (2010). A single-base resolution map of an archaeal transcriptome. *Genome Res.* 20, 133–141. doi: 10.1101/gr.100396.109

Zeldes, B. M., Loder, A. J., Counts, J. A., Haque, M., Widney, K. A., Keller, L. M., et al. (2019). Determinants of Sulphur chemolithoautotrophy in the extremely thermoacidophilic Sulfolobales. *Environ. Microbiol.* 21, 3696–3710. doi: 10.1111/1462-2920.14712

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. doi: 10.1038/s41467-018-07379-4

Zhang, Z., Zhan, Z., Wang, B., Chen, Y., Chen, X., Wan, C., et al. (2020). Archaeal chromatin proteins Cren7 and Sul7d compact DNA by bending and bridging. *MBio* 11, 1–15. doi: 10.1128/mBio.00804-20