



# Cas1 and Fen1 Display Equivalent Functions During Archaeal DNA Repair

Julia Wörtz<sup>1</sup>, Victoria Smith<sup>2</sup>, Jörg Fallmann<sup>3</sup>, Sabine König<sup>4,5</sup>, Tharani Thuraisingam<sup>6</sup>, Paul Walther<sup>6</sup>, Henning Urlaub<sup>4,5</sup>, Peter F. Stadler<sup>3,7,8,9,10,11,12,13</sup>, Thorsten Allers<sup>2</sup>, Frank Hille<sup>1</sup> and Anita Marchfelder<sup>1\*</sup>

<sup>1</sup> Biology II, Ulm University, Ulm, Germany, <sup>2</sup> School of Life Sciences, University of Nottingham, Nottingham, United Kingdom, <sup>3</sup> Department of Computer Science, Bioinformatics Group, Interdisciplinary Center for Bioinformatics, University of Leipzig, Leipzig, Germany, <sup>4</sup> Bioanalytical Mass Spectrometry Group, Max Planck Institute for Multidisciplinary Sciences, Göttingen, Germany, <sup>5</sup> Institute of Clinical Chemistry, University Medical Center Göttingen, Göttingen, Germany, <sup>6</sup> Central Facility for Electron Microscopy, Ulm University, Ulm, Germany, <sup>7</sup> German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany, <sup>8</sup> Competence Center for Scalable Data Services and Solutions, Leipzig Research Center for Civilization Diseases, University Leipzig, Leipzig, Germany, <sup>9</sup> Facultad de Ciencias, Universidad Nacional de Colombia, Bogotá, Colombia, <sup>10</sup> Institute for Theoretical Chemistry, University of Vienna, Vienna, Austria, <sup>11</sup> Center for RNA in Technology and Health, University of Copenhagen, Copenhagen, Denmark, <sup>12</sup> Santa Fe Institute, Santa Fe, NM, United States, <sup>13</sup> Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany

## OPEN ACCESS

### Edited by:

Sonja-Verena Albers,  
University of Freiburg, Germany

### Reviewed by:

Xu Peng,  
University of Copenhagen, Denmark  
Christopher Cooper,  
Peak Proteins, United Kingdom

### \*Correspondence:

Anita Marchfelder  
anita.marchfelder@uni-ulm.de

### Specialty section:

This article was submitted to  
Biology of Archaea,  
a section of the journal  
Frontiers in Microbiology

Received: 25 November 2021

Accepted: 21 February 2022

Published: 15 April 2022

### Citation:

Wörtz J, Smith V, Fallmann J,  
König S, Thuraisingam T, Walther P,  
Urlaub H, Stadler PF, Allers T, Hille F  
and Marchfelder A (2022) Cas1  
and Fen1 Display Equivalent  
Functions During Archaeal DNA  
Repair. *Front. Microbiol.* 13:822304.  
doi: 10.3389/fmicb.2022.822304

CRISPR-Cas constitutes an adaptive prokaryotic defence system against invasive nucleic acids like viruses and plasmids. Beyond their role in immunity, CRISPR-Cas systems have been shown to closely interact with components of cellular DNA repair pathways, either by regulating their expression or via direct protein-protein contact and enzymatic activity. The integrase Cas1 is usually involved in the adaptation phase of CRISPR-Cas immunity but an additional role in cellular DNA repair pathways has been proposed previously. Here, we analysed the capacity of an archaeal Cas1 from *Haloferax volcanii* to act upon DNA damage induced by oxidative stress and found that a deletion of the *cas1* gene led to reduced survival rates following stress induction. In addition, our results indicate that Cas1 is directly involved in DNA repair as the enzymatically active site of the protein is crucial for growth under oxidative conditions. Based on biochemical assays, we propose a mechanism by which Cas1 plays a similar function to DNA repair protein Fen1 by cleaving branched intermediate structures. The present study broadens our understanding of the functional link between CRISPR-Cas immunity and DNA repair by demonstrating that Cas1 and Fen1 display equivalent roles during archaeal DNA damage repair.

**Keywords:** CRISPR-Cas, Cas1, DNA repair, Fen1, archaea, *Haloferax volcanii*

## INTRODUCTION

Clustered regularly interspaced short palindromic repeats (CRISPR) and their associated genes (Cas) comprise a prokaryotic immune system, which defends bacteria and archaea from predatory mobile genetic elements, including viruses and plasmids (Hille et al., 2018; Nussenzweig and Marraffini, 2020). Its unique adaptive nature is enabled by proteins Cas1 and Cas2, which form

a complex that integrates short pieces of invading DNA, called spacers, into the chromosomal CRISPR array (Barrangou et al., 2007; Nuñez et al., 2014). Transcription of the CRISPR array generates a long precursor CRISPR-RNA (crRNA), which is processed by Cas proteins or cellular RNases into short mature crRNAs, each harbouring the sequence of a previously acquired spacer (Brouns et al., 2008). Cas effector nucleases bind and utilise crRNAs to sequence-specifically target and cleave complementary sequences, causing the degradation of invading DNA, thus protecting the cell (Barrangou et al., 2007).

Besides their role in prokaryotic immunity, the involvement of CRISPR-Cas systems in cellular functions beyond anti-viral defence has been subject of multiple studies (Faure et al., 2019a,b). One such function is the contribution to the repair of chromosomal DNA damage. In fact, CRISPR-Cas systems were first hypothesised to represent a novel DNA repair system due to similarity of enzymatic domains in Cas proteins with those found in DNA repair proteins like DNA helicases and nucleases such as RecB (Makarova et al., 2002, 2006). While this assumption has now been revised, a functional link between components of DNA repair pathways and CRISPR-Cas systems has been reported previously; during spacer acquisition, the DNA repair complexes RecBCD and AddAB (in Gram-negative and Gram-positive bacteria, respectively) are able to degrade invading DNA, thereby providing Cas1-Cas2 with DNA fragments for integration (Marraffini and Sontheimer, 2008). This mechanism biases the uptake of new spacers toward foreign genomes, thus avoiding auto-immunity (Levy et al., 2015; Modell et al., 2017). Moreover, the integration of a new spacer generates temporarily single-stranded regions within the CRISPR array. Sealing of those gaps by repair proteins like DNA polymerases and ligases are crucial to retain chromosome integrity (Ivančić-Baće et al., 2015).

There are also indications that Cas proteins may actively participate in DNA damage repair. Deletion of *cas1* in *Escherichia coli* yielded a phenotype that was sensitive to DNA damage and showed impaired chromosome segregation. The mechanism behind these observations likely relies on the ability of Cas1 to cleave branched DNA substrates that usually occur during DNA repair and recombination (Babu et al., 2011; Rollie et al., 2015). Moreover, Cas1 has been shown to associate with various repair proteins in *E. coli*, further strengthening the assumption that components of the *E. coli* CRISPR-Cas system are involved in DNA repair (Babu et al., 2011).

Another example illustrating the interplay of the CRISPR immune system and DNA repair was revealed in *Sulfolobus solfataricus*. Here, the CRISPR-associated protein Csa3a controls the expression of major CRISPR adaptation genes (Liu et al., 2015), is a key player in the activation of DNA repair genes (Liu et al., 2017), and regulates the DNA damage response (DDR) (Liu et al., 2020). The reason for the synergistic activation of both CRISPR-Cas and DNA repair pathways by Csa3a is probably due to the frequent acquisition of spacers from its own genome (roughly 7%) (Liu et al., 2017). The simultaneous activation of DDR genes thus reduces auto-immunity effects caused by self-targeting spacers (Liu et al., 2020).

The recognition of similar DNA substrates by Cas and repair nucleases might explain the mutually exclusive prevalence of

type II-A CRISPR-Cas systems and NHEJ (non-homologous end joining) genes. Csn2, an essential adaptation protein of type II-A systems (Heler et al., 2015; Wei et al., 2015), likely inhibits NHEJ repair by binding DNA at double-stranded breaks (DSBs) (Arslan et al., 2013), outcompeting the repair machinery and forcing bacteria to select against either of those systems (Bernheim et al., 2017).

In this study, we describe the interplay between CRISPR-Cas and DNA repair in the archaeon *Haloferax volcanii* by showing that Cas1 and repair protein Fen1 display equivalent roles during DNA repair. Fen1 is a highly conserved component of the cellular DNA metabolism found in Archaea and Eukarya that targets and removes 5' flap structures commonly associated with Okazaki fragments or base excision repair (Pan et al., 2011; Balakrishnan and Bambara, 2013; Grasso and Tell, 2014). *Haloferax volcanii* has a single *fen1* gene located on the main chromosome (Lestini et al., 2010) while *cas1* is encoded on a megaplasmid within a type I-B CRISPR-Cas system (Maier et al., 2019). The *cas* gene cassette is flanked by two CRISPR loci, with a third CRISPR locus present on the main chromosome (Maier et al., 2015).

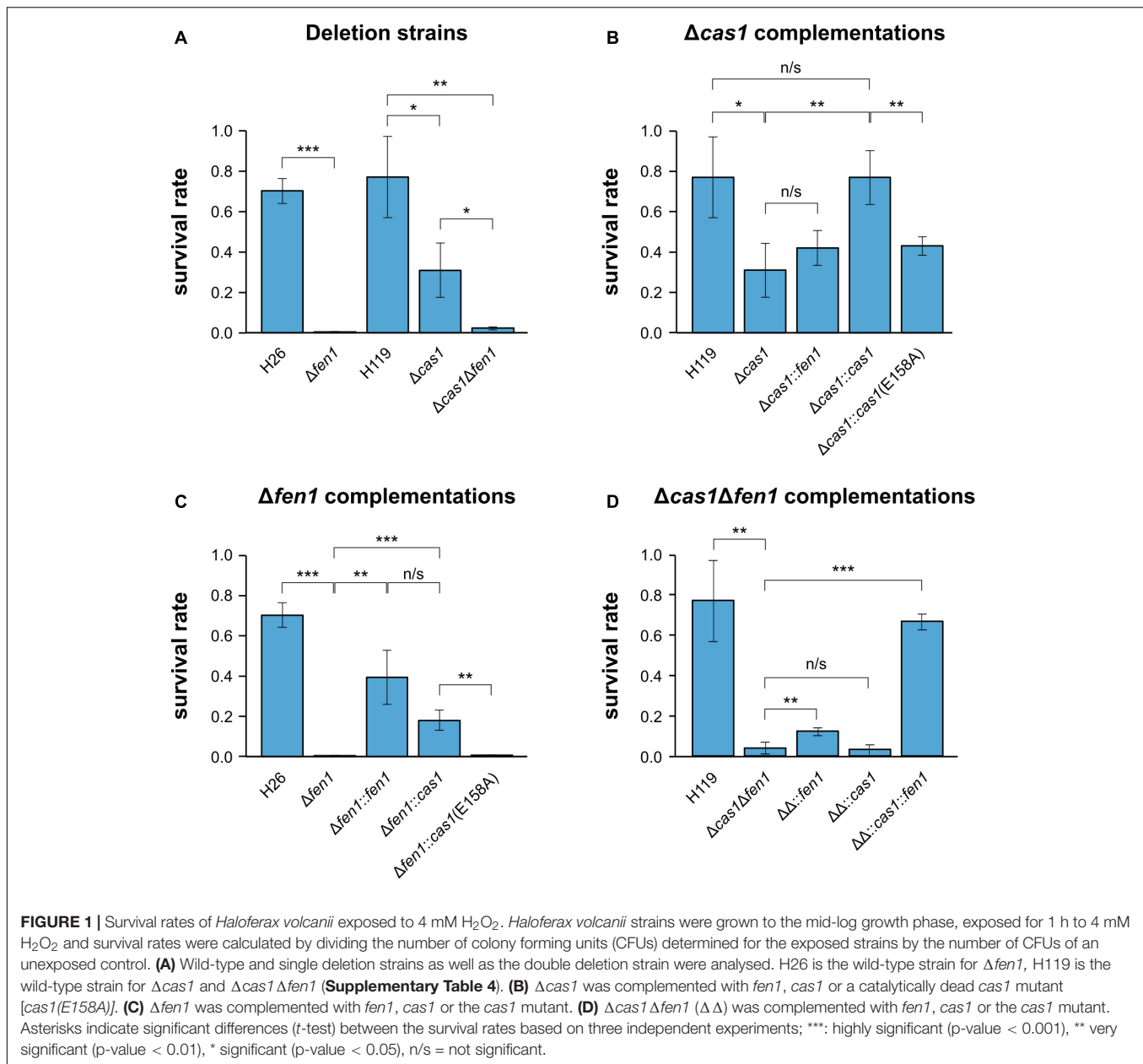
We found that deletions of either *cas1* or *fen1* caused similar phenotypes that exhibited sensitivity to oxidative stress and DNA damage. Double deletions caused cell elongation as well as increased cell size and when exposed to ultraviolet light (UV) stress survival rates were greatly reduced. We hypothesise that, given the similar substrate specificity and enzymatic activities of Cas1 and Fen1, both proteins are able to process 5' flap secondary DNA structures that occur during DNA repair and DNA replication.

## RESULTS

### Cas1 and Fen1 Are Required for DNA Repair

In order to investigate whether there is a functional link between components of the CRISPR-Cas system and the DNA repair machinery, we induced DNA damage in the *H. volcanii* wild-type strains and mutant daughter strains, harbouring a deletion of either *cas1*, *fen1* or both genes, respectively, and evaluated their fitness. To evaluate DNA repair, all strains were exposed to oxidative stress using H<sub>2</sub>O<sub>2</sub> and survival rates were measured. Both single deletion strains displayed a significant reduction in survival compared to the wild-type, with  $\Delta fen1$  showing a higher sensitivity than  $\Delta cas1$  (Figure 1A). The  $\Delta cas1 \Delta fen1$  double mutant displayed a growth defect comparable to the  $\Delta fen1$  single mutant. These results indicate that both Cas1 and Fen1 are involved in the cellular oxidative stress response.

To exclude any polar effects of the gene deletions, we complemented the  $\Delta cas1$  strain with a plasmid-born copy of the gene and were able to detect growth rescue in the complemented strain (Figure 1B). Moreover, when mutating the active site of Cas1 in the complementing gene (E158A, Supplementary Figure 1) (Babu et al., 2011; Kim et al., 2013; Rollie et al., 2015), growth rescue was impaired, further strengthening our



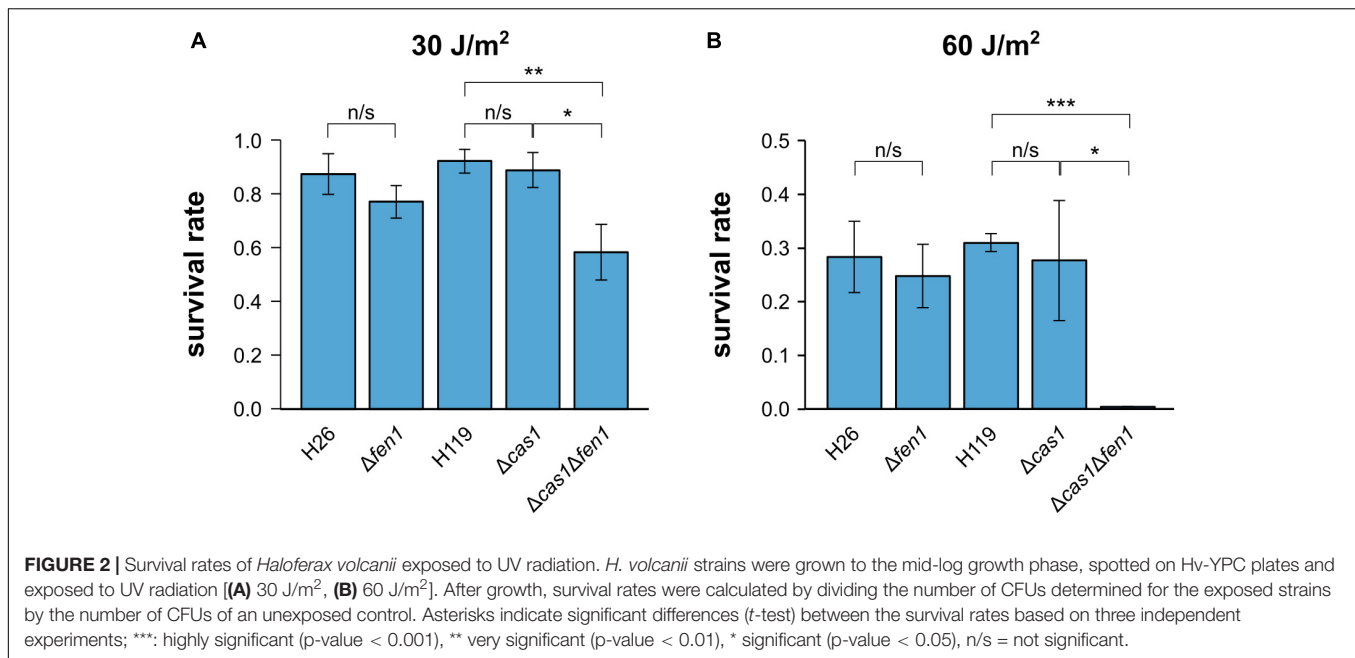
assumption that the enzymatic activity of Cas1 is directly involved in DNA repair (Figure 1B).

We also supplemented the *Δcas1* strain with a plasmid-borne copy of *fen1* but were not able to observe increased rescue, indicating that the DNA damage induced by H<sub>2</sub>O<sub>2</sub> cannot be compensated by higher levels of Fen1 (Figure 1B). Previous reports showed that Cas1 has not only a 5' flap processing activity but additional activities (Han et al., 2009; Wiedenheft et al., 2009; Beloglazova et al., 2015), which might explain why Fen1 cannot completely rescue a missing Cas1. Interestingly, when we overexpressed Cas1 in the *Δfen1* background, survival rates increased compared to both the uncomplemented *Δfen1* strain and the *Δfen1* strain harbouring a plasmid expressed active-site mutant of *cas1* (Figure 1C), clearly demonstrating the capacity

of Cas1 to mitigate DNA damage. It is noteworthy that the Cas1 protein is only present in low levels in wild-type cells under standard conditions (Jevtić et al., 2019), which could explain why the overexpression of Cas1 showed a significant increase in cell survival, as opposed to Fen1 in the *Δcas1* strain.

In the *Δcas1Δfen1* background, complementation with only *cas1* did not significantly increase the survival rate, while complementation with only *fen1* showed a slight but statistically significant growth rescue (p-value: 1.46E-02) (Figure 1D). These data are consistent with the results of the single deletion strains, where *Δfen1* showed an elevated sensitivity toward oxidative stress (p-value: 3.47E-06) compared to *Δcas1* (p-value: 1.68E-02).

To further evaluate the ability of Cas1 and Fen1 to repair DNA damage, we exposed each single deletion strain and the double



mutant to UV radiation (Figure 2). We observed a significant decrease in cell survival at both 30 and 60 J m<sup>-2</sup>, whereby only 1% of  $\Delta cas1\Delta fen1$  cells survived at 60 J m<sup>-2</sup> (Figure 2B). Interestingly, the single mutants did not display an increased sensitivity to UV radiation, indicating that UV-induced damage can be efficiently repaired in the presence of either enzyme alone.

### Absence of Cas1 and Fen1 Cause Increased Cell Size and Cell Elongation

Our results indicate that both Cas1 and Fen1 play vital roles during the DNA damage response. To gain further insights into their roles in cellular growth, we compared morphological features of the wild-type and the  $\Delta cas1\Delta fen1$  deletion strain.

Analysis of the strains via light microscopy revealed that the double mutant displayed an inflated phenotype under normal growth conditions, while the single deletion strains ( $\Delta cas1$  and  $\Delta fen1$ , respectively) lacked any obvious change in phenotype (Figure 3A), indicating that the deletion of both genes has effects on cellular pathways other than DNA damage response. The altered morphology of the double mutant was confirmed by scanning electron microscopy (Figure 3B). Here, the cells showed a variety of elongated and inflated shapes, indicating defects in the growth cycle and division of the double mutant strain. This assumption was further strengthened by analysing cell size and DNA content using flow cytometry (Figure 3C and Supplementary Figure 2). Here, we observed a shift of the peak absorbance of the DNA content in the  $\Delta cas1\Delta fen1$  strain compared to the wild-type, indicating fewer genome copies per cell in the mutant. Considering the enlarged phenotype of the mutant, these results suggest that the mutant fails to properly replicate its genome, leading to impaired cell division, resulting in the cell inflating. It can therefore be hypothesised that Cas1 and Fen1 may play a role in DNA-related pathways, such as DNA

replication, to maintain genome integrity not only during stress responses but under normal growth conditions as well.

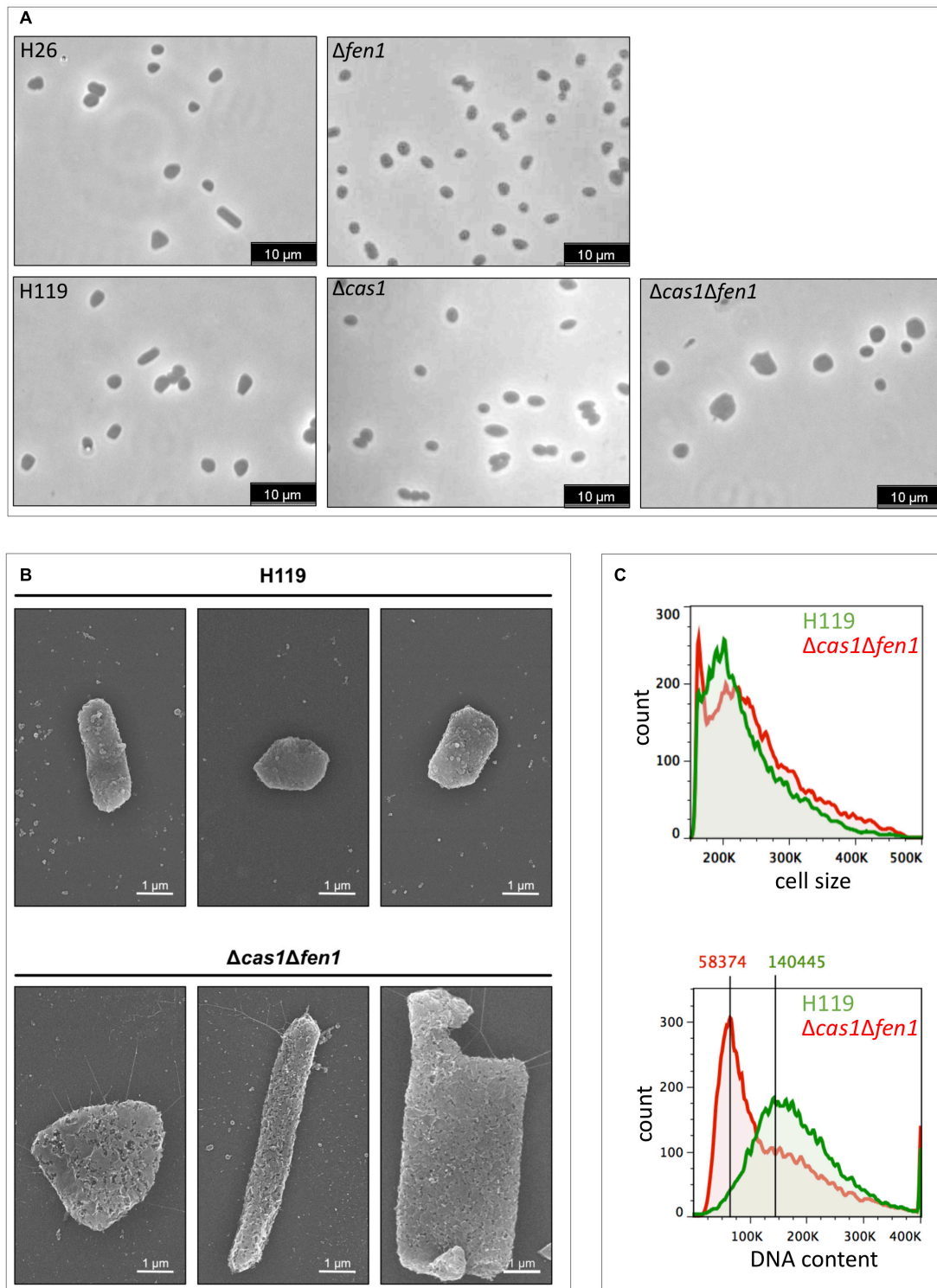
To assess whether the inflated phenotype of the  $\Delta cas1\Delta fen1$  mutant affects its ability to deal with osmotic stress, we monitored the growth of *H. volcanii* in media supplemented with different salt concentrations<sup>1</sup> via OD<sub>650nm</sub> measurements (Figure 4). Under standard conditions (18% salt), we observed a similar growth rate of the wild-type and double deletion strain (Figure 4A).

Interestingly, in both low and high salt conditions (15% and 23% salt, respectively), the mutant displayed a shorter lag phase compared to the wild-type, indicating an increased adaptation rate to osmotic stress in the absence of Cas1 and Fen1 (Figures 4B,C). However, when grown at 23% salt, the optical density of the double mutant culture declined shortly after reaching the stationary phase at roughly 80 h post-inoculation, while the wild-type was still in the logarithmic growth phase at a similar timepoint.

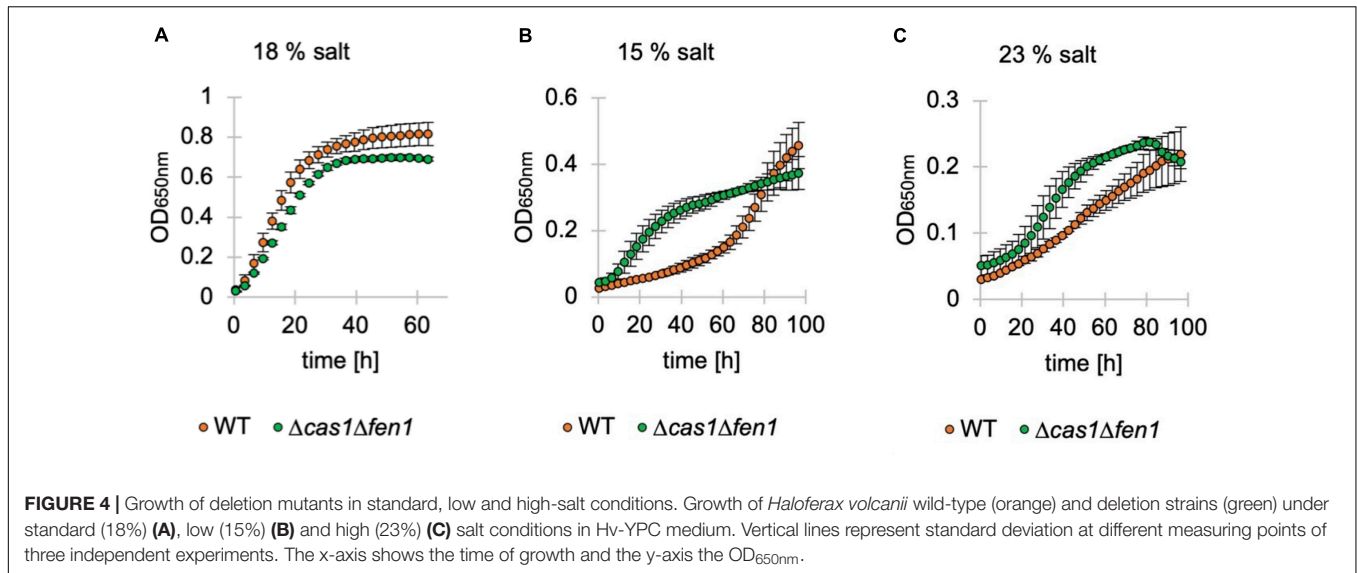
### Cas1 Interacts With DNA Repair Proteins

To exclude the possibility that the deletion of *cas1* and *fen1* has an effect on the expression level of other DNA repair proteins, thus indirectly causing impaired fitness during the cellular DNA damage response, we investigated gene expression patterns in both the wild-type and  $\Delta cas1\Delta fen1$  deletion strains using RNA sequencing. When comparing up- and down-regulated genes (Tables 1, 2 and Supplementary Tables 1, 2), only minor changes were observed in the gene expression levels of various functional groups that are not involved in DNA damage repair or DNA metabolism. However, the absence of Cas1 and Fen1 caused the

<sup>1</sup>Standard growth conditions (18% salt water) correspond to 15% NaCl, low salt (15% salt water) corresponds to 10.8% NaCl and high salt (23% salt water) corresponds to 19.2% NaCl.



**FIGURE 3 |** Cell morphology and DNA content of wild-type and the  $\Delta cas1\Delta fen1$  deletion strain. **(A)** *Haloferax volcanii* wild-type and deletion strains were grown to mid-log growth phase and cell morphology was analysed by light microscopy. **(B)** Scanning electron microscopic pictures of wild-type strain H119 (upper panel) and  $\Delta cas1\Delta fen1$  (lower panel) grown to stationary phase. **(C)** Flow cytometry of wild-type H119 (green) and  $\Delta cas1\Delta fen1$  deletion strains (red) grown to mid-log growth phase. The FS (forward light scatter) channel was used to evaluate cell size (upper panel) and the FL1 (fluorescence 1) channel was used to evaluate the DNA content of acridine orange-stained cells (lower panel). Modal population values for DNA content highlights a decreased DNA content in the  $\Delta cas1\Delta fen1$  mutant (41.5% of wild-type value).



up-regulation of altogether 15 putative proviral genes (Dyall-Smith et al., 2021) in the genome of *H. volcanii* (Table 1 and Supplementary Table 1), which might be relevant for the immunity-related functions of Cas1.

Considering the minor changes in observed expression patterns, we conclude that the deletion of *cas1* and *fen1* does not strongly impact cellular pathways on a transcriptional level.

We further analysed direct protein-protein interactions of Cas1 by conducting pull-down analyses. Cas1 co-purified with

a variety of proteins involved in replication and DNA repair (Table 3), including several helicases and the DNA mismatch

**TABLE 1 |** RNA expression profile of the  $\Delta cas1 \Delta fen1$  deletion mutant compared to the wild-type.

Gene_ID	Annotation/Gene name	logFC
<b>Provirus regions</b>		
HVO_0369	hypothetical protein; ProVir2 prediction	2.18
HVO_1434	hypothetical protein; ProVir5 prediction	2.16
HVO_A0218	oxidoreductase; ProVir4 prediction	2.01
HVO_0276A	homolog to HGPV1-ORF9; ProVir6 prediction	2.00
<b>Iron metabolism</b>		
HVO_1721	2Fe-2S iron-sulfur cluster binding domain-containing protein, <i>ferA3</i>	2.04
<b>Transposases</b>		
HVO_2817	transposase (ISH51)	2.06
HVO_A0258	ISH4-type transposase homolog	2.06
<b>Other</b>		
HVO_B0028	xylose dehydrogenase (NAD/NADP dependent), <i>xacA</i>	2.32
HVO_B0342	oxidoreductase (Luciferase family protein), <i>mer2</i>	2.19
HVO_0641	cob(II)yrinic acid a,c-diamide reductase, <i>bluB</i>	2.10
HVO_B0343	hydrolytic enzyme LpID, <i>lpID</i>	2.07
HVO_0694	purine phosphoribosyltransferase, <i>gptA</i>	2.05
HVO_1205	taxis cluster protein CheD, <i>cheD</i>	2.00

Shown are up-regulated genes ( $\log_{2}FC \geq 2.00$ ) ( $\log_{2}FC$ : log<sub>2</sub> fold change.). Six genes for hypothetical proteins are also up-regulated with  $\log_{2}FC \geq 2.00$  (Supplementary Table 1), for a complete list of up-regulated genes see Supplementary Table 1. Provirus prediction according to Dyall-Smith et al. (2021).

**TABLE 2 |** RNA expression profile of the  $\Delta cas1 \Delta fen1$  deletion mutant compared to the wild-type.

Gene_ID	Annotation/Gene name	logFC
<b>Deleted genes</b>		
HVO_2873	flap endonuclease, <i>fen1</i>	-8.13
HVO_A0211	Cas1 protein, <i>cas1</i>	-7.63
<b>Transposases</b>		
HVO_A0279	Transposase (ISH18)	-4.41
HVO_2051	Transposase (ISH51)	-2.23
<b>Transcription regulators</b>		
HVO_2507	Asn family transcription regulator, <i>trh7</i>	-3.17
HVO_2522	Asn family transcriptional regulator, <i>trh8</i>	-2.97
<b>tRNA metabolism</b>		
HVO_1092	ribonuclease P protein component 2, <i>mp2</i>	-2.68
HVO_3052	tRNA Gly	-2.35
<b>Iron metabolism</b>		
HVO_A0541	ABC-type transport system periplasmic substrate-binding protein (probable substrate iron-III)	-2.45
HVO_B0044	<i>iucA</i> iron transport protein A, <i>iucA</i>	-2.43
HVO_2588	isocitrate dehydrogenase, <i>iucD</i>	-2.28
<b>Other</b>		
HVO_1228	halocyanin domain protein (membrane), <i>hcpE</i>	-2.92
HVO_2508	carbamoyl-phosphate synthase small subunit, <i>carA</i>	-2.56
HVO_2361	carbamoyl-phosphate synthase large subunit, <i>carB</i>	-2.04
HVO_B0045	diaminobutyrate decarboxylase, <i>bdb</i>	-2.47
HVO_B0046	diaminobutyrate pyruvate aminotransferase, <i>dat</i>	-2.26
HVO_2606	PQQ repeat-containing protein	-2.72
HVO_2607	PQQ repeat-containing protein	-2.29

Shown are down-regulated genes ( $\log_{2}FC \leq -2.00$ ) ( $\log_{2}FC$ : log<sub>2</sub> fold change.). Nine genes for hypothetical proteins are also down-regulated with  $\log_{2}FC \leq -2.00$  (Supplementary Table 2), for a complete table of down-regulated genes see Supplementary Table 2.

**TABLE 3** | Proteins that were co-purified with Cas1.

Gene ID	Annotation/Gene name	Peptides/Unique spectra counts
<b>Replication and repair</b>		
HVO_0393	UvrABC system protein A, <i>uvrA</i>	73
HVO_0029	UvrABC system protein B, <i>uvrB</i>	32
HVO_0415	repair helicase UvrD, <i>uvrD</i>	38
HVO_0349	DNA-directed RNA polymerase subunit A, <i>rpoA1</i>	56
HVO_0347	DNA-directed RNA polymerase subunit B, <i>rpoB2</i>	44
HVO_0858	DNA-directed DNA polymerase B (intein-containing), <i>polB1</i>	36
HVO_2380	AAA-type ATPase (CDC48 subfamily), <i>cdc48a</i>	66
HVO_0854	DNA double-strand break repair ATPase Rad50, <i>rad50</i>	56
HVO_B0118	Smc-like protein Sph2; homolog of Rad50, <i>sph2</i>	30
HVO_0552	DNA mismatch repair protein MutS, <i>muts1b</i>	49
HVO_0014	ATP-dependent DNA helicase Hel308a, <i>hel308a</i>	44
HVO_1018	Hef-associated 3 exonuclease, <i>recJ3</i>	35
HVO_2889	DHH/RecJ family phosphoesterase RecJ4, <i>recJ4</i>	33
HVO_0220	ATP-dependent DNA helicase MCM, <i>mcm</i>	34
<b>Cas protein</b>		
HVO_A0206	Cas8b, <i>cas8</i>	58
<b>Ribonucleases</b>		
HVO_0874	zinc-dependent nuclease CPSF1, <i>cpsf1</i>	44
HVO_2724	ribonuclease J, <i>mJ</i>	35
<b>Sensing kinases</b>		
HVO_1811	sensor box histidine kinase	32
HVO_B0154	receiver/sensor box histidine kinase	41

Two independent pull-down analyses using a FLAG-Cas1 fusion protein resulted in the co-purification of interacting proteins. Corresponding peptides of the co-purifying proteins were sequenced by mass spectrometry (MS). Only proteins that are functionally interesting, e.g., involved in DNA repair, CRISPR-Cas functions, and having unique spectra/peptide counts  $\geq 30$  are listed. The complete table of co-purified proteins with a spectra/peptide count  $\geq 30$  is found in **Supplementary Table 3**. The complete set of MS data including all co-purified proteins (including those with a spectra/peptide count  $< 30$ ) has been deposited in PRIDE. Column peptides/unique spectra counts: number of peptides and unique spectra counts identified.

repair protein MutS, and was associated with three members of the UvrABC system that acts in nucleotide excision repair. For a final confirmation of these interactions reverse pull-down experiments have to be carried out where the proteins identified by the Cas1 pull-down are FLAG-tagged and the presence of Cas1 in the elution fraction is determined after FLAG-purification.

Other proteins that co-purified with Cas1 included RNases, sensing kinases, and proteins associated with amino acid synthesis, vitamin and glucose metabolism (**Table 3** and **Supplementary Table 3**). Interestingly, only five peptides of Fen1 were detected during the MS analysis (data not shown, the complete set of co-purified proteins has been deposited in

PRIDE), suggesting that the interaction between Cas1 and Fen1 is only weak or does not occur at all. This indicates that their functional link might not be based on physical interactions within the cell.

## Cas1 Has 5' Flap Endonuclease Activity

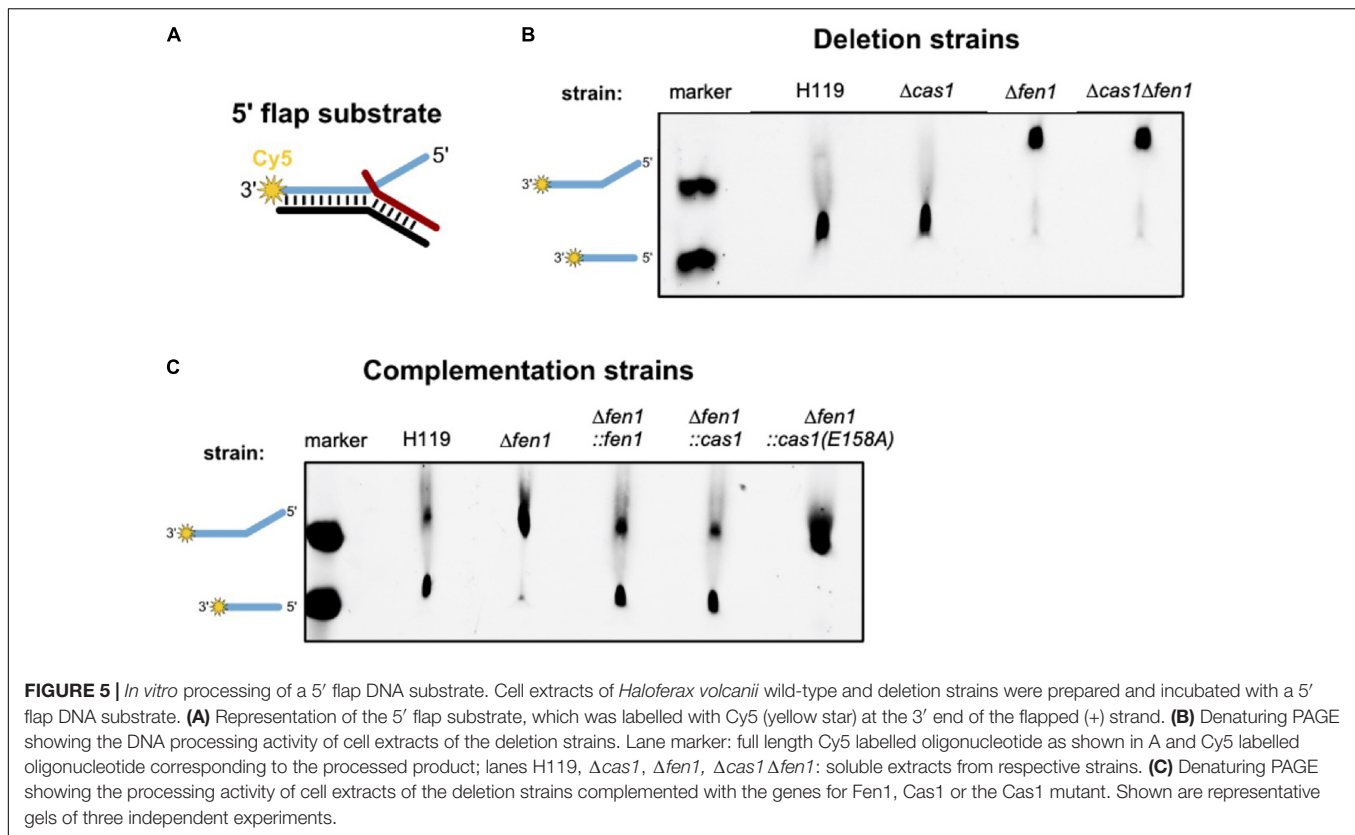
Previous studies have revealed that Cas1 has DNA processing activities and is able to cleave branched DNA substrates such as Holliday junctions and 5' flap structures (Babu et al., 2011; Rollie et al., 2015). We hypothesised that the type I-B Cas1 homolog of *H. volcanii* might possess similar activities on branched DNA substrates. In the context of DNA repair, processing of such structures by Cas1 might compensate for a deletion of the flap endonuclease Fen1. To test this, we attempted to express and purify both nucleases recombinantly in *E. coli* but failed to achieve adequate yield and purity. Therefore, we prepared cell extracts of *H. volcanii* wild-type as well as *cas1* and *fen1* deletion mutants and compared their ability to process a double flap DNA substrate (**Figure 5**).

We observed processing of the branched substrate by cell extracts of the wild-type and  $\Delta cas1$  strain, whereas cell extracts of the  $\Delta fen1$  strain yielded no DNA cleavage product, suggesting that only Fen1 is able to process the DNA double flap substrate (**Figure 5B**). It is noteworthy that Cas1 levels are considerably low under normal growth conditions (Jevtić et al., 2019), potentially too low to observe substrate cleavage in our assay. Therefore, we repeated the cleavage assay using cell extracts of a  $\Delta fen1$  strain that constitutively expressed an *in trans* copy of *cas1* and were able to detect processing of the double flap substrate (**Figure 5C**). Moreover, this effect was not observable in the  $\Delta fen1$  strain expressing a nuclease-defective variant of Cas1 (E158A), which strongly indicates that both Fen1 and Cas1 are able to recognise and process a DNA flap.

## DISCUSSION

Thus far it has only been demonstrated in bacteria that components of the prokaryotic immune system CRISPR-Cas may be involved in other cellular pathways, such as DNA repair. In this study, we demonstrate a functional link between CRISPR-Cas systems and DNA metabolism in Archaea, by showing that Cas1 and the replication enzyme Fen1 play equivalent roles during the DNA damage response in *H. volcanii*.

Fen1 is found in Archaea and Eukarya and cleaves flapped DNA structures that occur during DNA replication and repair (Pan et al., 2011; Balakrishnan and Bambara, 2013). Accordingly, deletion of the homologous gene in *H. volcanii* causes reduced survival rates following DNA-damaging oxidative stress. Deletion of *cas1* had similarly adverse effects on cell survival under these conditions, indicating that the CRISPR-Cas immune protein is involved in DNA repair as well. Additionally, we found that the enzymatic activity of the Cas1 nuclease is directly responsible for DNA repair, as a mutation in the active site of the protein also resulted in reduced survival after oxidative stress. The mechanistic basis for repairing DNA damage is likely linked to



the ability of the nuclease Cas1 to cleave flapped DNA substrates in a similar manner to Fen1.

Pull-down experiments showed that Cas1 seems to physically interact with a variety of DNA binding and processing enzymes, some of which are involved in DNA repair. Whether these interactions affect the role of Cas1 in DNA repair remains to be revealed, however, it is tempting to speculate that Cas1 is part of a multi-protein repair complex in *H. volcanii*. The formation of such a complex might determine the role of Cas1 in either DNA repair or CRISPR-Cas immunity, which would be advantageous for the cell. During viral defence, Cas1 forms a complex with Cas2 and integrates short pieces of DNA into the CRISPR array (Nuñez et al., 2014). Recruiting the Cas1-Cas2 complex to DNA lesions during repair could trigger spacer acquisition from the organism's own genome and consequently lead to auto-immunity. The interaction of Cas1 with DNA repair enzymes would prevent this problem, since complex formation with Cas2 is constrained and thus prevents spacer acquisition. Further experiments are required to reveal the factors that influence the role of Cas1 in either damage repair or immunity. In our pull-down experiments Cas2 could not be identified as interaction partner of Cas1. Cas2 is not detectable in proteome analyses of cells grown under standard conditions (Jevtić et al., 2019; Schulze et al., 2020) which might explain this observation.

In addition to its role in DNA repair, Cas1 might also play a role in DNA replication. In *E. coli*, it has been proposed that Cas1 is involved in genome segregation, based on the detection of abnormally elongated cells of a *cas1* knockout

strain after mitomycin C treatment (Babu et al., 2011). In *H. volcanii*, a single deletion of either *fen1* or *cas1* did not cause any notable morphological changes in cell shape or size, compared to the wild-type under unstressed conditions. However, a simultaneous deletion of both genes led to an inflated phenotype and reduced DNA content per cell, with the latter being indicative of fewer genome copies per cell. Similar morphological alterations have been observed before, where UV-induced DNA damage in *H. volcanii* results in increased cell size and cell elongation (Delmas et al., 2013). These results suggest that both Cas1 and Fen1 might process flapped DNA substrates during replication and thus maintain genome integrity under conditions of genomic stress.

## CONCLUSION

This study expands our knowledge of the role of CRISPR-Cas systems beyond immunity, by demonstrating that, similar to its bacterial homologs, the archaeal Cas1 in *H. volcanii* may participate in cellular DNA repair.

## MATERIALS AND METHODS

Strains, plasmids and oligonucleotides used are listed in **Supplementary Tables 4–6**. *E. coli* strain DH5a (Invitrogen, Thermo Fischer Scientific, Waltham, MA, United States) was used for plasmid cloning and grown aerobically at 37°C in 2YT medium (Miller, 1972).



## Growth Curves

*Haloferax volcanii* strains were grown in Hv-YPC medium (Dyall-Smith, 2008) at 45°C with an optimal concentration of 18% salt water (2.46 M NaCl, 88 mM MgCl<sub>2</sub>, 85 mM MgSO<sub>4</sub>, 56 mM KCl, 12 mM Tris-HCl, pH 7.5). For low and high salt stress, salt concentration was adjusted to 15% and 23%, respectively, without modifying the ion ratios. While standard conditions (18% salt water) correspond to 15% NaCl, low salt (15% salt water) corresponds to 10.8% NaCl and high salt (23% salt water) corresponds to 19.2% NaCl. Growth of cultures was monitored in 96 well microtiter plates using an Epoch2 NS Microplate Spectrophotometer (BioTek Instruments, Bad Friedrichshall, Germany). Strains were precultured in Hv-YPC medium to OD<sub>650nm</sub> of 0.4–0.7, after dilution to an OD<sub>650nm</sub> of 0.05 they were transferred to microtiter plates. Cultures were incubated aerobically with orbital shaking at 45°C and OD<sub>650nm</sub> was measured every 30 min. Outer wells were filled with salt water as evaporation barriers (Liao et al., 2016). The growth curves represent the average of at least three biological replicates.

## Plasmid Cloning

pBlue-NdeI-Cas1-EcoRV was generated by ligating a linearised pBluescriptII KS (cleaved with EcoRV) with the insert that contained *cas1* and the restriction sites NdeI and EcoRV. The insert was amplified from *H. volcanii* genomic DNA using primers Cas1up-FLAGC and Cas1do-FLAGC, primers contained the restriction sites NdeI and EcoRV.

pTA927-Cas1E158A-CFLAG was generated using a site-directed mutagenesis kit (Agilent, Santa Clara, CA, United States) according to the manufacturer's instructions with primers Cas1E158A fw and Cas1E158A rev and pTA927-Cas1-CFLAG as a template.

Plasmid pTA927-FLAGcontrol was cloned as follows. Using primers FLAG NdeI fw and FLAG EcoRI rev with pTA927-NZ-NFLAG as template, the FLAG sequence (ATG-3xFLAG-TGA) was amplified. The primers contained restriction sites for *ApaI* (5') and *EcoRI* (3'), after digestion of the resulting polymerase chain reaction (PCR) product with *ApaI* and *EcoRI* ligation with pTA927 (digested with *ApaI* and *EcoRI*) resulted in pTA927-FLAGcontrol.

Plasmid pTA231.p.fdx was obtained by ligation of pTA231 (digested with *NotI* and *EcoRI*) with a p.fdx-t.syn fragment excised with *NotI* and *EcoRI* from pMA-T-FdxSyn (synthesised by GeneArt, Thermofisher Scientific), resulting in pTA231-p.fdx.

In order to generate pTA231-p.fdx-NFLAG, a PCR with primers 231NFLAG fw and 231NFLAG rev and pTA927-Cas1-NFLAG as a template was performed resulting in PCR fragment ATG-3xFLAG, primers contained a *NdeI* restriction site (5' terminus) and a *SnaBI* as well as a *XbaI* restriction site (3' terminus). The purified PCR product was digested with *NdeI* and *SnaBI* and ligated to pTA231-p.fdx (digested with *NdeI* and *XbaI*), to generate pTA231-p.fdx-NFLAG.

Plasmid pTA230-p.fdx-NFLAG was generated by digesting pTA231-p.fdx-NFLAG with *NotI* and *EcoRI* to obtain the p.fdx-NFLAG-t.syn fragment (p.fdx-ATG-3x FLAG-t.syn), which was subsequently ligated with the pTA230 vector (linearised with *NotI* and *EcoRI*).

Plasmids pTA230-p.fdx-Cas1-NFLAG and pTA230-p.fdx-Fen1-NFLAG were obtained as follows. *fen1* and *cas1* genes were amplified from *H. volcanii* genomic DNA using primers Fen1 *SnaBI* fw/Fen1 *XbaI* rev and Cas1 *EcoRV* fw/Cas1 *XbaI* rev, respectively. The resulting PCR products were ligated with pBluescriptII KS (*EcoRV* digested), resulting in pBlue-*SnaBI*-Fen1-*XbaI* and pBlue-*EcoRV*-Cas1-*XbaI*. These plasmids were then digested with *SnaBI/XbaI* or *EcoRV/XbaI* to excise the *fen1* and *cas1* gene fragments, respectively, which were subsequently ligated with pTA230-p.fdx-NFLAG (linearised with *SnaBI* and *XbaI*) yielding pTA230-p.fdx-Fen1-NFLAG and pTA230-p.fdx-Cas1-NFLAG.

For generation of pTA230-p.fdx-Cas1E158A-NFLAG, a site-directed mutagenesis kit (Agilent, Santa Clara, CA, United States) was used according to the manufacturer's instructions with primers Cas1E158A fw and Cas1E158A rev and pTA230-p.fdx-Cas1-NFLAG as a template.

To obtain the pTA962-p.fdx-Fen1-NFLAG-Cas1 overexpression plasmid, the *cas1* insert was generated with PCR using primers 5'XmaJI Cas1 + Start fw/3'BglII Cas1 + Stopp rev and plasmid pTA230-p.fdx-Cas1-NFLAG as template. The resulting PCR product was cleaved with *XmaJI/BglII* and then ligated with pTA962 (cleaved with *XmaJI/BglII*). To generate p.fdx-Fen1-NFLAG, the *fen1* insert was obtained with PCR using primers 5'ApaI p.fdx fw/3'XmaJI Fen1 + Stopp rev and pTA230-p.fdx-Fen1-NFLAG as template. The resulting PCR product was cloned into pTA962 via *ApaI/XmaJI* restriction sites.

The *cas1* deletion plasmid pTA131-UPDO(*cas1*) was obtained by amplifying the *cas1* gene with 480 bp up- and 587 bp downstream sequences using primers Cas1KOU/P/Cas1KODO. The resulting fragment was ligated with pTA131 (linearised with *EcoRV*). The resulting plasmid pTA131-UP-Cas1-DO was used as template for an inverted PCR with 5' phosphorylated primers iCas1 fw/iCas1 rev, the resulting PCR product was ligated to obtain pTA131-UPDO(*cas1*). This plasmid contains only the up- and downstream regions of *cas1*.

## Generation of Knockout Mutants

Deletion strains were generated using the pop-in pop-out method as described previously (Bitan-Banin et al., 2003). Transformations were performed according to the PEG600 protocol as described in the HaloHandbook (Dyall-Smith, 2008). Strain  $\Delta cas1$  was generated as described (Klein, 2011), in short: For the generation of the  $\Delta cas1$  knockout strain, the H119 wild-type strain was transformed with the *cas1* deletion plasmid pTA131-UPDO(*cas1*) and plated onto Hv-Ca medium (Dyall-Smith et al., 2021) supplemented with L-tryptophan (40  $\mu$ g/ml) to select for pop-in candidates which were subsequently identified by PCR using the primers Cas1 KO UP/Cas1 KO DO. Positive clones were then grown in Hv-Ca medium supplemented with L-tryptophan (40  $\mu$ g/ml) and subsequently plated onto Hv-Ca agar plates with L-tryptophan (40  $\mu$ g/ml), 5-FOA (5-fluoroorotic acid) (10  $\mu$ g/ml) and 10  $\mu$ g/ml uracil to select for pop-out clones. Knockout candidates were identified by PCR using the same primers as mentioned above and the additional primer pair Cas1 *EcoRV* fw/Cas1 *ApaI* rev. Homozygous deletion strains were finally confirmed via Southern Blot analysis (Sambrook and Russell, 2001; Klein, 2011).

In order to establish a  $\Delta cas1 \Delta fen1$  knockout strain, the  $\Delta cas1$  strain was transformed with *fen1* deletion plasmid pCN6 (Meslet-Cladière et al., 2007). Transformants were selected by plating on Hv-Ca medium supplemented with tryptophan (40  $\mu\text{g/ml}$ ). Pop-in candidates were subsequently plated onto Hv-Ca agar plates supplemented with L-tryptophan (40  $\mu\text{g/ml}$ ), 5-FOA (10  $\mu\text{g/ml}$ ) and 10  $\mu\text{g/ml}$  uracil. Pop-in and pop-out clones were identified by PCR using the primers Fen1 HVO UP/Fen1 HVO DO and Fen1 *Sna*BI fw/Fen1 *Xba*I rev. Finally, homozygous pop-out candidates were verified by Southern Blot analysis (Sambrook and Russell, 2001) (**Supplementary Figure 3**). Genomic DNAs of candidate clones K1, K4, K17, K20, K22 and of wild-type H119 were extracted via spooling method (Dyall-Smith, 2008). Concentrations of genomic DNAs were determined with a NanoPhotometer N60, the DNA was stored at 4°C. Genomic DNA was digested with *Sal*I and a Southern blot analysis was made as described in Sambrook and Russell (2001). Radioactively labeled PCR products were used as hybridisation probes to detect DNA fragments on the southern blot membrane. PCR products amplified with Fen1 Sonde fw/Fen1 Sonde rev and Fen1 (HVO) UP/Fen1 (HVO) DO, respectively, using H119 gDNA as a template were used as probes. For radiolabeling, 50  $\mu\text{Ci}$  [ $\alpha$ - $^{32}\text{P}$ ]-dCTP and the DECAprime<sup>TM</sup> II DNA labeling kit (Life Technologies/Merck, Germany) were used. Two separate probes were generated, one that binds upstream of the *fen1* coding sequence, and another one that binds directly in the *fen1* gene. After hybridisation the membrane was exposed to an x-ray film (**Supplementary Figure 3**).

## H<sub>2</sub>O<sub>2</sub> and UV Radiation Survival Assays

Wild-type and deletion strains were inoculated in 4 ml Hv-YPC medium, strains that carried a plasmid were inoculated in 4 ml selective medium with the appropriate supplements and incubated over night at 45°C until culture density reached an OD<sub>650nm</sub> of 0.3–0.5. To determine survival rates after H<sub>2</sub>O<sub>2</sub> exposure, cultures that reached the adequate OD<sub>650nm</sub> were aliquoted into two 2 ml reaction tubes by pipetting 490  $\mu\text{l}$  culture into each tube. One tube was supplemented with 10  $\mu\text{l}$  of a 200 mM H<sub>2</sub>O<sub>2</sub> stock solution (final concentration 4 mM), and 10  $\mu\text{l}$  of 18% saltwater was added to the other tube (control). H<sub>2</sub>O<sub>2</sub> stock solutions were freshly prepared for each experiment by diluting 30% H<sub>2</sub>O<sub>2</sub> with 18% saltwater. Samples were incubated at 45°C and 450 rpm for 1h. Subsequently, samples were serially diluted (1:10) in 18% saltwater and 20  $\mu\text{l}$  of the 10<sup>-3</sup> to 10<sup>-6</sup> dilutions were spotted onto prewarmed Hv-YPC agar plates in duplicate. After the liquid was absorbed by the agar, plates were incubated at 45°C for 3 days until colonies could be counted. Survival rates were determined by dividing the number of colonies of the 4 mM H<sub>2</sub>O<sub>2</sub> approach by the number of colonies of the control. P-values were determined by *t*-tests (two sample assuming equal variances).

To compare the sensitivity of the wild-type and deletions strains toward UV radiation, strains were incubated as described above until an OD<sub>650nm</sub> of 0.3–0.5. Next, cultures were serially diluted (1:10) in 18% saltwater and 20  $\mu\text{l}$  of the dilutions 10<sup>-2</sup> to 10<sup>-6</sup> were spotted onto Hv-YPC agar plates in duplicate. After the liquid was absorbed by the agar, plates were exposed

to UV radiation (either 30 or 60 J m<sup>-2</sup>) in a Stratalink<sup>®</sup> UV Crosslinker (Stratagene, Agilent, Santa Clara, CA, United States). Control plates were not exposed to UV radiation. To prevent photoreactivation, all plates were incubated in the dark at 45°C for at least three days. Colonies were counted and survival rates were calculated by dividing the number of colonies on irradiated plates by the number of colonies on the control plates. All experiments were performed in triplicate.

## Light Microscopy

Liquid cultures were inoculated in 4 ml Hv-YPC medium and incubated at 45°C overnight. After cultures reached an OD<sub>650nm</sub> of 0.3–0.5 (exponential phase) and 0.9–1.2 (stationary phase), respectively, 1 ml of each culture was transferred into a reaction tube and centrifuged at 1,500 x g for 10 min. The supernatant was discarded, and the obtained pellets were resuspended in 50  $\mu\text{l}$  (derived from exponential culture) and 150  $\mu\text{l}$  (derived from stationary culture) Hv-YPC medium, respectively, and applied onto agarose-covered microscope slides. Agarose slides were prepared as follows: 0.1 g agarose was solved in 4 ml bidistilled water by heating the suspension in a microwave until the agarose was completely dissolved. Meanwhile, 6 ml 18% saltwater was prewarmed in a water bath and added to the hot agarose solution. With the help of a brush, the agarose solution was applied onto the slides. 2  $\mu\text{l}$  of the resuspended pellets were pipetted onto the hardened agarose slides and covered with a coverslip. The edges of the cover glasses were sealed with clear nail polish and microscopic analysis was done at 100 x magnification with oil immersion on a Leica DM5500 B light microscope.

## Scanning Electron Microscopy

For scanning electron microscopy H119 and  $\Delta cas1 \Delta fen1$  cells were grown to stationary phase (OD<sub>650nm</sub>: 1.2) in Hv-YPC, adsorbed on silicon platelets and fixed for 1 h with a final concentration of 2.5% glutaraldehyde in 0.1 M phosphate buffer with 1% saccharose. Afterward, samples were fixed with 2% of osmium tetroxide in phosphate buffer for 20 min, dehydrated in a graded series of propanol and then critical point dried using carbon dioxide in a CPD BalTec030 critical point dryer. Finally, samples were mounted on specimen stubs and rotary coated by electron beam evaporation with about 2 nm of platinum. Samples were imaged in a Hitachi S-5200 scanning electron microscope at an accelerating voltage of 10 kV using the secondary electron signal.

## Co-immunoprecipitation and Mass Spectrometry

The FLAG-Cas1 fusion protein and as a control the FLAG protein were expressed in *Haloflex* cells (strain  $\Delta cas1$ ). 500 ml Hv-Ca medium supplemented with 40  $\mu\text{g/ml}$  L-tryptophan were inoculated from a plate and grown until OD<sub>650nm</sub> of 0.7. After incubation for 1 h cultures were harvested by centrifugation at 10,000 x g for 30 min at 4°C. The resulting pellets were washed with enriched PBS [2.5 M NaCl, 150 mM MgCl<sub>2</sub>, 1xPBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na<sub>2</sub>PO<sub>4</sub>, 2 mM KHPO<sub>4</sub>, pH 7.4)] and stored at -80°C. Pellets were thawed on ice

and resuspended in 10 ml lysis buffer (50 mM Tris-HCl, pH 7.4, 1 mM EDTA, 10 mM MgCl<sub>2</sub>, 1 mM CaCl<sub>2</sub>) containing 150 μl of proteinase inhibitor (Sigma-Aldrich/Merck, Darmstadt, Germany) and 100 μl RQ1 RNase-free DNase (Promega, Walldorf, Germany). After 30 min incubation on ice, cells were disrupted by ultrasonic treatment and subsequently centrifuged at 100,000 x g at 4°C for 30 min. The supernatants (S100) were transferred into fresh falcons and RNase A was added at a final concentration of 400 μg/ml. After incubation for 30 min at 37°C, NaCl was added to a final concentration of 150 mM. A total of one FLAG control purification and two FLAG-Cas1 purifications were carried out.

For affinity purification, 400 μl of anti-FLAG M2 affinity gel (Sigma-Aldrich/Merck, Darmstadt, Germany) was washed 10 times with 10 ml of ice-cold washing buffer (50 mM Tris/HCl, pH 7.4, 150 mM NaCl) before the lysate was added. After incubation overnight at 4°C, anti-FLAG M2 affinity gel was washed 10 times with 10 ml of ice-cold washing buffer. Proteins were eluted from the affinity gel using washing buffer supplemented with 150 ng/l 3xFLAG peptide and fractions were collected in 2 ml reaction tubes and finally precipitated using acetone (99%). Protein concentrations were determined using ROTI<sup>®</sup>Nanoquant according to the manufacturer's protocol, extracts were snap frozen in liquid nitrogen and stored at -80°C. The FLAG elution fraction was concentrated using an Amicon Ultra - 4 10,000 NMWL device (centrifugation for 15 min at 7,500 xg and 4°C). 50 mM Na<sub>3</sub>PO<sub>4</sub>/150 mM NaCl buffer was added to the retentate to a final volume of 200 μl. The sample was then loaded onto a Superdex<sup>™</sup> 75 Increase 10/300 gel filtration column. 1 ml fractions containing Cas1 and eluting with the complex size of approximately 670 kDa were collected and used for mass spectrometry analyses.

SDS-PAGE-separated protein samples were processed as described by Shevchenko et al. (1996). The resulting peptides were loaded to nano HPLC coupled with Exploris Orbitrap Mass spectrometer (Thermo Fisher Scientific, Waltham United States). The peptides were separated with a linear gradient of 5–40% buffer B (80% acetonitrile and 0.1% formic acid) at flow rate of 300 nL/min over 48 min total gradient time. The MS data was acquired by scanning the precursors in mass range from 350 to 1400 m/z at a resolution of 60,000 at m/z 200. Top30 precursor ion were chosen for MS2 by using data-dependent acquisition (DDA) mode at a resolution of 15,000 at m/z 200 with auto IT. Data analysis and search was performed against Uniprot\_Haloferax\_volcanii database (July 2021, 3911 entries) using Maxquant Software (1.6.17.0) and results were annotated with Scaffold5 software. Search parameters were 0.1 FDR on peptide and protein level; Carbamidomethyl was set as fixed modification and oxidation of methionine and acetylation at the N-terminus of the protein as variable modifications.

To evaluate the data, first all proteins for which more than five peptides were found in the FLAG control were removed from the final tables (Table 3 and Supplementary Table 3). In addition, for proteins that were found in the FLAG only control (identified with four or less peptides) the number of peptides found in the FLAG only control was subtracted from those found in the FLAG-Cas1 samples. In addition, only proteins found

in both FLAG-Cas1 purifications were included in Table 3 and Supplementary Table 3.

## Flow Cytometry

DNA content and cell size of *H. volcanii* cells were determined via flow cytometry. Cultures were prepared in 5 ml Hv-YPC broth and grown at 45°C with 8 rpm rotation in two successive overnight dilutions until an OD<sub>650nm</sub> of ~0.4 was reached. Acridine orange solution was added to the cells at a final concentration of 1 μg/ml. Samples were analysed using an FC500 flow cytometer (Beckman Coulter; University of Nottingham Flow Cytometry facility) equipped with a 488 nm laser and 528/28 emission filter to measure acridine orange fluorescence. Samples were run on the lowest speed setting and at least 20,000 cells were acquired for each sample. Data was analysed using Flow Jo v7.6 (Tree Star Inc.). Cells were gated based on forward and side scatter and doublets were excluded by height/area analysis.

## Preparation of Cell Extracts and Nuclease Assays

For the preparation of total protein extracts, 150 ml cultures were grown in Hv-YPC medium (strains without plasmids) or Hv-Min medium (strains carrying a plasmid) until they reached an OD<sub>650nm</sub> of 0.6–0.8. Cells were harvested by centrifugation at 10,000 x g and 4°C for 30 min, cell pellets were washed once with enriched PBS [2.5 M NaCl, 150 mM MgCl<sub>2</sub>, 1xPBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na<sub>2</sub>PO<sub>4</sub>, 2 mM KHPO<sub>4</sub>, pH 7.4)] and stored at -80°C until further usage. Pellets were thawed on ice and resuspended in 2 ml Lysis buffer (100 mM Tris-HCl pH 7.5, 10 mM EDTA) and subsequently disrupted by ultrasonic treatment. To separate soluble and non-soluble lysate fractions, suspensions were centrifuged at 100,000 x g and 4°C for 30 min. The supernatants (S100) were transferred into 2 ml reaction tubes and total protein concentration was measured using ROTI<sup>®</sup>Nanoquant solution according to the manufacturer's instructions. Extracts were snap frozen in liquid nitrogen and stored at -80°C until further use.

For the *in vitro* 5' flap processing experiments, 20 μg of the S100 protein extracts were mixed with 150 ng double flap DNA substrate and 10x IVP buffer (500 mM Tris-HCl pH 7.4, 500 mM MgCl<sub>2</sub>, 50 mM KCl) and incubated for 2 h at 45°C. The longer flap strand was ordered as oligonucleotide labelled with Cy5 at the 3' end (Sigma-Aldrich, Taufkirchen, Germany). (Figure 5, as described in Craggs et al., 2014). Proteins were inactivated and DNA was denatured by heat treatment (90°C, 10 min) and reactions were loaded onto a denaturing 10% PAA gel. Cy5 signals were detected on a ChemiDoc using standard parameters for Cy5.

## RNA Sequencing

Three replicates of wild-type H119 and deletion strain  $\Delta cas1\Delta fen1$  were cultured in Hv-YPC medium and grown to OD<sub>650nm</sub> of 0.6–0.7. Total RNA was isolated using NucleoZOL<sup>™</sup> (Machery and Nagel, Düren, Germany) and RNA samples were sent to vertis Biotechnologie AG (Martinsried, Germany) for sequencing. Total RNA was treated with T4 Polynucleotide Kinase (NEB, Frankfurt,

Germany) and rRNA was depleted using an in-house protocol. Samples were fragmented using an ultrasonic treatment before cDNA library was prepared. For this, first-strand cDNA synthesis was performed using a 3'-adapter primer and the M-MLV reverse transcriptase. After cDNA purification, the 5' Illumina TruSeq sequencing adapter was ligated to the 3' end of the antisense cDNA and the sample amplified using TruSeq\_Sense\_primer and TruSeq\_Antisense\_primer to 10–20 ng/μl using a high-fidelity DNA polymerase. Finally, cDNA was purified using the Agencourt AMPure XP kit (Beckman Coulter Genomics), samples were pooled (equimolar) and the pool size was fractionated (200–550 bp) by preparative agarose gel electrophoresis before the samples were sequenced on an Illumina NextSeq 500 system using 1 × 75 bp read length.

## Bioinformatics Analyses

Quality control was conducted on raw reads and after each processing step using fastqc (v0.11.9) (<sup>2</sup>accessed 09/09/20) and MultiQC (v1.10.1) (Ewels et al., 2016). Adapters were trimmed using trimgalore (v0.6.3) (<sup>3</sup>accessed 09/09/20) and reads were mapped against the *H. volcanii* genome with STAR (v2.7.3a) (Dobin et al., 2013). Unique reads were extracted and counted using featureCounts (v2.0.1) (Liao et al., 2014) and DESeq2 (v1.30.0) (Love et al., 2014) was used for differential expression analysis. Command lines for the processing steps are available in the **Supplementary Material**. A modular workflow assembler which was used to generate workflows for this analysis is available at <https://github.com/jfallmann/MONSDA> (Fallmann et al., 2022).

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: <https://www.ncbi.nlm.nih.gov/>, PRJNA776127; <https://www.ebi.ac.uk/pride/archive/>, PXD029952 (Vizcaíno et al., 2016).

<sup>2</sup> <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>

<sup>3</sup> <https://github.com/FelixKrueger/TrimGalore>

## REFERENCES

- Arslan, Z., Wurm, R., Brener, O., Ellinger, P., Nagel-Steger, L., Oesterhelt, F., et al. (2013). Double-strand DNA end-binding and sliding of the toroidal CRISPR-associated protein Csn2. *Nucleic Acids Res.* 41, 6347–6359. doi: 10.1093/nar/gkt315
- Babu, M., Beloglazova, N., Flick, R., Graham, C., Skarina, T., Nocek, B., et al. (2011). 'A dual function of the CRISPR-Cas system in bacterial antiviral immunity and DNA repair'. *Mol. Microbiol.* 79, 484–502.
- Balakrishnan, L., and Bambara, R. A. (2013). 'Flap Endonuclease I'. *Annu. Rev. Biochem.* 82, 119–138.
- Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., et al. (2007). 'CRISPR provides acquired resistance against viruses in prokaryotes'. *Science* 315, 1709–1712.
- Beloglazova, N., Lemak, S., Flick, R., and Yakunin, A. F. (2015). "Analysis of nuclease activity of Cas1 proteins against complex DNA substrates," in *CRISPR: Methods and Protocols*, eds M. Lundgren, E. Charpentier, and P. C. Fineran (New York, NY: Humana Press-Springer). doi: 10.1007/978-1-4939-2687-9\_16

## AUTHOR CONTRIBUTIONS

JW, VS, SK, and TT did the experiments. JE, SK, PW, HU, PS, TA, FH, and AM performed data curation. AM conceptualised the project. FH and AM wrote the original draft. FH, JW, JE, SK, VS, TT, PW, HU, PS, TA, and AM reviewed and edited the draft. PW, HU, PS, TA, and AM provided the resources and funding. All authors contributed to the article and approved the submitted version.

## FUNDING

Work in the AM laboratory as well as in the HU laboratory was funded by the DFG priority programme "CRISPR-Cas functions beyond defence" SPP2141 (Ma1538/25-1 and UR225- 7/1). Data analysis was in part funded by the German Federal Ministry for Education and Research [BMBF 031A538B (de.NBI/RBC)] to PS. This work was also supported by the Biotechnology and Biological Sciences Research Council (BBSRC) [BB/M008770/1] to VS. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

## ACKNOWLEDGMENTS

We thank Juliane Benz, Lisa-Katharina Maier, and Carola Klein for help with generating plasmids pTA927Cas1-NFLAG, pTA927-Cas1-CFLAG, pTA231-p.fdx and strain Δ*cas1*, as well as Irma Merdian and Susanne Schmidt for expert technical assistance. Furthermore we thank Lisa-Katharina Maier for helpful discussions and for critically reading the manuscript.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2022.822304/full#supplementary-material>

- Bernheim, A., Calvo-Villamañán, A., Basier, C., Cui, L., Rocha, E. P. C., Touchon, M., et al. (2017). 'Inhibition of NHEJ repair by type II-A CRISPR-Cas systems in bacteria'. *Nat. Comm.* 8:2094. doi: 10.1038/s41467-017-02350-1
- Bitan-Banin, G., Ortenberg, R., and Mevarech, M. (2003). 'Development of a Gene Knockout System for the Halophilic Archaeon *Haloflex volcanii* by Use of the *pyrE* Gene'. *J. Bacteriol.* 185, 772–778. doi: 10.1128/JB.185.3.772-778.2003
- Brouns, S. J., Jore, M. M., Lundgren, M., Westra, E. R., Slijkhuys, R. J., Snijders, A. P., et al. (2008). 'Small CRISPR RNAs guide antiviral defense in prokaryotes'. *Science* 321, 960–964. doi: 10.1126/science.1159689
- Craggs, T. D., Hutton, R. D., Brenlla, A., White, M. F., and Penedo, J. C. (2014). 'Single-molecule characterization of Fen1 and Fen1/PCNA complexes acting on flap substrates'. *Nucleic Acids Res.* 42, 1857–1872. doi: 10.1093/nar/gkt1116
- Delmas, S. I., Duggin, G., and Allers, T. (2013). DNA damage induces nucleoid compaction via the Mre11-Rad50 complex in the archaeon *Haloflex volcanii*. *Mole. Microbiol.* 87, 168–179. doi: 10.1111/mmi.12091

- Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). 'STAR: ultrafast universal RNA-seq aligner'. *Bioinformatics* 29, 15–21. doi: 10.1093/bioinformatics/bts635
- Dyall-Smith, M. (2008). *The Halo handbook: protocols for haloarchaeal genetics*. Melbourne: Haloarchaeal Genetics Laboratory, 14.
- Dyall-Smith, M., Pfeiffer, F., Chiang, P. W., and Tang, S. L. (2021). The Novel Halovirus Hardycor1, and the Presence of Active (Induced) Proviruses in Four Haloarchaea. *Genes* 2021, 12. doi: 10.3390/genes12020149
- Ewels, P., Magnusson, M., Lundin, S., and Käller, M. (2016). 'MultiQC: summarize analysis results for multiple tools and samples in a single report'. *Bioinform.* 32, 3047–3048. doi: 10.1093/bioinformatics/btw354
- Fallmann, J., Goldmann, R., and Stadler, P. F. (2022). MONSDA, Modular Organizer of Nextflow and Snakemake Driven hts Data Analysis. (in preparation)
- Faure, G., Makarova, K. S., and Koonin, E. V. (2019a). 'CRISPR–Cas: Complex Functional Networks and Multiple Roles beyond Adaptive Immunity'. *J. Mole. Biol.* 431, 3–20. doi: 10.1016/j.jmb.2018.08.030
- Faure, G., Shmakov, S. A., Yan, W. X., Cheng, D. R., Scott, D. A., Peters, J. E., et al. (2019b). 'CRISPR–Cas in mobile genetic elements: counter-defence and beyond'. *Nat. Rev. Microb.* 17, 513–525. doi: 10.1038/s41579-019-0204-7
- Grasso, S., and Tell, G. (2014). 'Base excision repair in Archaea: Back to the future in DNA repair'. *DNA Repair* 21, 148–157. doi: 10.1016/j.dnarep.2014.05.006
- Han, D., Lehmann, K., and Krauss, G. (2009). 'SSO1450—a CAS1 protein from *Sulfolobus solfataricus* P2 with high affinity for RNA and DNA'. *FEBS Lett.* 583, 1928–1932. doi: 10.1016/j.febslet.2009.04.047
- Heler, R., Samai, P., Modell, J. W., Weiner, C., Goldberg, G. W., Bikard, D., et al. (2015). 'Cas9 specifies functional viral targets during CRISPR-Cas adaptation'. *Nature* 519, 199–202. doi: 10.1038/nature14245
- Hille, F., Richter, H., Wong, S. P., Bratovič, M., Ressel, S., and Charpentier, E. (2018). 'The Biology of CRISPR-Cas: Backward and Forward'. *Cell* 172, 1239–1259. doi: 10.1016/j.cell.2017.11.032
- Ivančić-Baće, I., Cass, S. D., Wearne, S. J., and Bolt, E. L. (2015). 'Different genome stability proteins underpin primed and naïve adaptation in *E. coli* CRISPR-Cas immunity'. *Nucleic Acids Res.* 43, 10821–10830. doi: 10.1093/nar/gkv1213
- Jevtić, Z., Stoll, B., Pfeiffer, F., Sharma, K., Urlaub, H., Marchfelder, A., et al. (2019). The Response of *Haloferax volcanii* to Salt and Temperature Stress: A Proteome Study by Label-Free Mass Spectrometry. *Proteomics* 19:1800491. doi: 10.1002/pmic.201800491
- Kim, T. Y., Shin, M., Huynh, L., Yen, T., and Kim, J. S. (2013). Crystal structure of Cas1 from *Archaeoglobus fulgidus* and characterization of its nucleolytic activity. *Biochem. Biophys. Res. Commun.* 441, 720–725. doi: 10.1016/j.bbrc.2013.10.122
- Klein, C. (2011). *Die Untersuchung des CRISPR/Cas Systems in Haloferax volcanii*. Germany: Ulm University.
- Lestini, R., Duan, Z., and Allers, T. (2010). 'The archaeal Xpf/Mus81/FANCM homolog Hef and the Holliday junction resolvase Hjc define alternative pathways that are essential for cell viability in *Haloferax volcanii*'. *DNA Repair* 9, 994–1002. doi: 10.16/j.dnarep.2010.06.012
- Levy, A., Goren, M. G., Yosef, I., Auster, O., Manor, M., Amitai, G., et al. (2015). 'CRISPR adaptation biases explain preference for acquisition of foreign DNA'. *Nature* 520, 505–510. doi: 10.1038/nature14302
- 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. doi: 10.1093/bioinformatics/btt656
- Liao, Y., Williams, T. J., Walsh, J. C., Ji, M., Poljak, A., Curmi, P. M. I., et al. (2016). 'Developing a genetic manipulation system for the Antarctic archaeon, *Halorubrum lacusprofundi*: investigating acetamidase gene function. *Sci. Rep.* 6:34639. doi: 10.1038/srep34639
- Liu, T., Li, Y., Wang, X., Ye, Q., Li, H., Liang, Y., et al. (2015). 'Transcriptional regulator-mediated activation of adaptation genes triggers CRISPR de novo spacer acquisition'. *Nucleic Acids Res.* 43, 1044–1055. doi: 10.1093/nar/gku1383
- Liu, T., Liu, Z., Ye, Q., Pan, S., Wang, X., Li, Y., et al. (2017). 'Coupling transcriptional activation of CRISPR-Cas system and DNA repair genes by Csa3a in *Sulfolobus islandicus*'. *Nucleic Acids Res.* 45, 8978–8992. doi: 10.1093/nar/gkx612
- Liu, Z., Sun, M., Liu, J., Liu, T., Ye, Q., Li, Y., et al. (2020). 'A CRISPR-associated factor Csa3a regulates DNA damage repair in *Crenarchaeon Sulfolobus islandicus*'. *Nucleic Acids Res.* 48, 9681–9693. doi: 10.1093/nar/gkaa694
- 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. doi: 10.1186/s13059-014-0550-8
- Maier, L.-K., Dyall-Smith, M., and Marchfelder, A. (2015). 'The Adaptive Immune System of *Haloferax volcanii*'. *Life* 5, 521–537. doi: 10.3390/life5010521
- Maier, L.-K., Stachler, A.-E., Brendel, J., Stoll, B., Fischer, S., Haas, K. A., et al. (2019). 'The nuts and bolts of the *Haloferax* CRISPR-Cas system I-B'. *RNA Biol.* 16, 469–480. doi: 10.1080/15476286.2018.1460994
- Makarova, K. S., Aravind, L., Grishin, N. V. I., Rogozin, B., and Koonin, E. V. (2002). 'A DNA repair system specific for thermophilic Archaea and bacteria predicted by genomic context analysis'. *Nucleic Acids Res.* 30, 482–496. doi: 10.1093/nar/30.2.482
- Makarova, K. S., Grishin, N. V., Shabalina, S. A., Wolf, Y. I., and Koonin, E. V. (2006). 'A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action'. *Biol. Direct.* 1:7. doi: 10.1186/1745-6150-1-7
- Marraffini, L. A., and Sontheimer, E. J. (2008). 'CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA'. *Science* 322, 1843–1845. doi: 10.1126/science.1165771
- Meslet-Cladière, L., Norais, C., Kuhn, J., Briffotiaux, J., Sloostra, J. W., Ferrari, E., et al. (2007). 'A novel proteomic approach identifies new interaction partners for proliferating cell nuclear antigen'. *J. Mol. Biol.* 372, 1137–1148. doi: 10.1016/j.jmb.2007.06.056
- Miller, J. H. (1972). *Experiments in Molecular Genetics*. Cold Spring Harbour, N. Y.: Cold Spring Harbour Laboratory Press.
- Modell, J. W., Jiang, W., and Marraffini, L. A. (2017). 'CRISPR–Cas systems exploit viral DNA injection to establish and maintain adaptive immunity'. *Nature* 544, 101–104. doi: 10.1038/nature21719
- Nuñez, J. K., Kranzusch, P. J., Noeske, J., Wright, A. V., Davies, C. W., and Doudna, J. A. (2014). 'Cas1-Cas2 complex formation mediates spacer acquisition during CRISPR-Cas adaptive immunity'. *Nat. Struct. Mol. Biol.* 21, 528–534. doi: 10.1038/nsmb.2820
- Nussenzweig, P. M., and Marraffini, L. A. (2020). 'Molecular Mechanisms of CRISPR-Cas Immunity in Bacteria'. *Annu. Rev. Genet.* 54, 93–120. doi: 10.1146/annurev-genet-022120-112523
- Pan, M., Kelman, L. M., and Kelman, Z. (2011). 'The archaeal PCNA proteins'. *Biochem. Soc. Trans.* 39, 20–24. doi: 10.1042/BST0390020
- Rollie, C., Schneider, S., Brinkmann, A. S., Bolt, E. L., and White, M. F. (2015). 'Intrinsic sequence specificity of the Cas1 integrase directs new spacer acquisition. *eLife* 4:e08716. doi: 10.7554/eLife.08716
- Sambrook, J., and Russell, D. (2001). *Molecular Cloning: A Laboratory Manual*. New York, NY: Cold Spring Harbour Press.
- Schulze, S., Adams, Z., Cerletti, M., De Castro, R., Ferreira-Cerca, S., Fufezan, C., et al. (2020). 'The Archaeal Proteome Project advances knowledge about archaeal cell biology through comprehensive proteomics'. *Nat. Comm.* 11:3145. doi: 10.1038/s41467-020-16784-7
- Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996). 'Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels'. *Anal. Chem.* 68, 850–858. doi: 10.1021/ac950914h
- Vizcaino, J. A., Csordas, A., del-Toro, N., Dianes, J. A., Griss, J., Lavidas, I., et al. (2016). '2016 update of the PRIDE database and its related tools'. *Nucleic Acids Res.* 44, D447–D456.
- Wei, Y., Terns, R. M., and Terns, M. P. (2015). 'Cas9 function and host genome sampling in Type II-A CRISPR-Cas adaptation'. *Genes Dev.* 29, 356–361. doi: 10.1101/gad.257550.114
- Wiedenheft, B., Zhou, K., Jinek, M., Coyle, S. M., Ma, W., and Doudna, J. A. (2009). 'Structural basis for DNase activity of a conserved protein implicated

in CRISPR-mediated genome defense'. *Structure* 17, 904–912. doi: 10.1016/j.str.2009.03.019

**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.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in

this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

*Copyright © 2022 Wörtz, Smith, Fallmann, König, Thuraingam, Walther, Urlaub, Stadler, Allers, Hille and Marchfelder. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*