Methylation deficiency of chromatin proteins is a non-mutational and epigenetic-like trait in evolved lines of the archaeon *Sulfolobus solfataricus*

Archaea are a distinct and deeply rooted lineage that harbor eukaryotic-like mechanisms, including several that manage chromosome function. In previous work, the thermoacidophilic crenarchaeon, *Sulfolobus solfataricus*, was subjected to adaptive laboratory evolution to produce three strains, called SARC, with a new heritable trait of super acid resistance. These strains acquired heritable conserved transcriptomes, yet one strain contained no mutations. Homologous recombination without allele replacement at SARC acid resistance genes caused changes in both phenotype and expression of the targeted gene.

As recombination displaces chromatin proteins, their involvement was predicted in the SARC trait. Native chromatin proteins are basic and highly abundant and undergo post-translational modification through lysine monomethylation. In this work, their modification states were investigated. In all SARC lines, two chromatin proteins, Cren7 and Sso7d, were consistently undermethylated, whereas other chromatin proteins were unaltered. This pattern was inheritable in the absence of selection and independent of transient exposure to acid stress.

The bulk of Sso7d was undermethylated at three contiguous N-terminal lysine residues but not at central or C-terminal regions. The N-terminal region formed a solvent-exposed patch located on the opposite side of the binding domain associated with the DNA minor groove. By analogy to eukaryotic histones, this patch could interact with other chromosomal proteins and be modulated by differential post-translational modification.

Previous work established an epigenetic-like mechanism of adaptation and inheritance in *S. solfataricus*. The identification of heritable epigenetic marks in this work further supports the occurrence of an epigenetic process in archaea.

The key structural feature of eukaryotic chromatin is the nucleosome composed of DNA wrapped around histone proteins identified by the histone fold domain. These “canonical” histones have broad distribution in higher organisms and are thought to play critical roles in gene regulation. Divergent “noncanonical” histone variants are also commonly found in eukaryotes but are implicated in a broader range of functions including DNA repair, meiotic recombination, and chromosome segregation (1). Noncanonical histone fold proteins are present in most members of the archaeal phylum Euryarchaeota and rarely in some members of the phylum Crenarchaeota (2–4). However, they are absent from many other crenarchaeotal orders, including the order *Sulfolobales*. Instead, the chromatin proteins of the *Sulfolobales* are small highly basic DNA-binding proteins much like those found in bacteria. However, archaeal chromatin exhibits a patchy phylogenetic distribution of critical features. Although Euryarchaeota have histones (2, 3), they are not post-translationally modified (5). In contrast, the chromatin proteins (not histones) in Crenarchaeota are modified, and some extensively (6–9). The chromatin proteins of *Sulfolobus* species in particular have received considerable study.

In the thermoacidophile, *Sulfolobus solfataricus*, there are seven reported chromatin proteins: Sso10b, Sso10b2, Sso7d1, Sso7d2, Cren7, Sso10a, and Sso10a2 (8–19). Sso10b (also known as Alba1) and Sso10b2 (Alba2) are paralogous and are present in other archaea. *In vivo* association of Alba1 with the sirtuin deacetylase, Sir2, implicated a contribution to chromatin structure and gene regulation (6), as did the apparent oligomerization of Sso10b paralogs (20). Sso10b members belonging to pfam01918 are also abundant in higher organisms. Sso7d and its paralogs belong to the Sso7c or Sul7d family and are abundant (nearly 5% of total protein) 7-kDa monomeric DNA-binding proteins distributed only among related *Sulfolobus* species (8, 18).

In *S. solfataricus* strain 98/2 (SULA) (NZ_CP011057.2), Sso7d is present in two copies called Sso7d1 and Sso7d2, whereas in the related *S. solfataricus* strain P2 (NC_002754.1), it is present in three copies. Based on its ability to induce DNA curvature, negative supercoiling, and compaction *in vitro*, the Sso7d ortholog, Sul7d, might promote DNA packaging *in vivo* (21). Homology between the Sul7d fold and the eukaryotic chromatin-binding (chromo)domain, a mediator of chromatin protein interactions (22), combined with its ability to disaggregate protein complexes (23) and to undergo lysine-specific monomethylation (8), implicates Sul7d...
Heritable chromatin hypomethylation links archaeal phenomics

in chromatin function. Cren7 is another S. solfataricus chromatin protein conserved across the Crenarchaeota, with a mass of 6.6 kDa, comprising nearly 1% of total protein and undergoing lysine monomethylation that likely plays an important structural role (9). It binds preferentially to dsDNA over ssDNA, although without apparent sequence preference, significantly increases the melting temperature of DNA, and constrains negative supercoil formation (9, 19). Despite having no amino acid sequence similarities, both Cren7 and Sso7d share similar tertiary protein structures and DNA-binding and -bending characteristics. Both proteins undergo lysine methylation on surface-exposed lysine residues to a different extent, although the effect of this post-translational modification remains a mystery (9, 10, 25). Co-crystals of both proteins bound to small dsDNA sequences determined that these proteins bind DNA in the minor groove, producing sharp kinks of 50–60° (19, 26, 27). There are two additional chromatin proteins in S. solfataricus, including Sso10a (28) and Sso10a2, reported solely through its N-terminal strand of HMG1 implicating post-translational modifications and recruitment (34). Post-translational modifications of the tertiary protein structures and DNA-binding and -bending characteristics. Both proteins undergo lysine methylation on surface-exposed lysine residues to a different extent, although the effect of this post-translational modification remains a mystery (9, 10, 25). Co-crystals of both proteins bound to small dsDNA sequences determined that these proteins bind DNA in the minor groove, producing sharp kinks of 50–60° (19, 26, 27). There are two additional chromatin proteins in S. solfataricus, including Sso10a (28) and Sso10a2, reported solely through its structural analysis (PDB² code 4HW0), whose roles in chromosome function are unknown.

In eukaryotic organisms, nonhistone minor groove-binding DNA-binding proteins have regulatory effects modulated by lysine methylation. High-mobility group (HMG) proteins represent a class of proteins that, similarly to Cren7 and Sso7d, bind to the minor groove of DNA with no sequence specificity and bend DNA by sharp kinking (29). HMG proteins play important roles in remodeling chromatin and modulating gene expression in eukaryotic organisms through their distortion of DNA structure by direct bending or protein–protein–protein complex with histones and transcription factors (30). Characterization of their architectural role on DNA found that mean bend angles were 67° and 78° for HMGB1 and -2, respectively (31). Like histones and crenarchaeal chromatin proteins, HMG proteins are subjected to post-translational modification, and their biological activity is highly regulated by their post-translational modification (PTM) state, including lysine methylation (32). For example, the male sex–determining factor (SRY) and lymphoid enhancer-binding factor 1 (LEF-1) are two examples of HMG architectural proteins that interact exclusively with the minor groove of DNA (33), induce sharp kinks on DNA (31), and form multiprotein–DNA (nucleoprotein) complexes (29). PTMs have been established as altering multiprotein–DNA complexes and have been shown to be targets for specific interactions and recruitment (34). Post-translational modifications of the N-terminal strand of HMG1 implicate the reversible modulation of modifications in HMG1–protein and HMG1–DNA interactions (35, 36). Gene regulation by tumor suppression protein p53, another minor groove binder, is highly modulated by its PTM state. In particular, lysine methylation regulates protein–protein interactions and stability (37, 38).

To better understand the biotype of hyperthermoacidophily, adaptive laboratory evolution was used over a 3-year period to test whether such organisms harbor additional thermoacido-philic capacity (39–41). Three independent and genetically marked cell lines derived from a single type species were subjected to high-temperature serial passage, whereas culture acidity was gradually increased. A 178-fold increase in thermoacidopily was achieved after 29 increments of shifted culture pH, resulting in growth at pH 0.8 and 80 °C. These acid-adapted strains were named super acid–resistant Crenarchaeota (SARC). Surprisingly, SARC trait heritability could not be explained by mutation because one strain (SARC-I) had none. For the other two cell lines, mutation rates were 8.3-fold lower than a control cell line that was passaged without selection. In addition, these two SARC lines, transposition rates were 23-fold lower than in the control line, indicating an impact of the evolutionary process on both mutation and transposition. Transcriptomic analysis of the three SARC lineages revealed a conserved, heritable expression pattern that likely contributed to acid resistance. Finally, the role of chromatin proteins in the SARC gene expression pattern and its heritability was tested using homologous recombination without allele replacement to replace target regions with DNA lacking preexisting chromatin protein. Using this method, gene expression was altered in a locus-specific manner. When genes belonging to the SARC transcriptome were targeted, acid resistance decreased, whereas no phenotypic changes were apparent in genes not belonging to the SARC transcriptome. These data support an epigenetic-like model where archaeal chromatin proteins contribute to heritable gene expression states.

Due to the genetic stability of the SARC cell lines (39–41) and lack of dependence on stress physiology to elicit the evolved traits, studies were conducted on chromatin proteins from cells in balanced growth prepared using instrumented bioreactors. Purified chromatin proteins were evaluated to assess the impact of adaptive evolution on chromatin protein PTM. A heritable change that was conserved in all three evolved SARC lineages was identified in chromatin protein methylation state. This alteration may contribute to the evolved SARC trait of hyperthermoacidophily in a manner analogous to altered histone post-translational modification. These changes in PTMs were also not a response to acid stress, as the same modification state was apparent after cycling at pH 3. Mixed reaction-monitoring MS identified N terminus–specific methylation differences between the parental and acid-evolved Sso7d protein as well as a novel modification site occurring at the DNA–contact interface of the protein. A protein crystal model for the implications of the changes is discussed.

## Results

### Native chromatin purification and identification

To evaluate possible changes in the chromosomal organizational proteins during their adaptation to extreme acid resistance, native S. solfataricus chromatin proteins were examined from WT and evolved derivatives adapted to extreme acid (39–41). Chromatin proteins were purified from three parental cell lines (SULA, SULG, and SULM) and their SARC derivatives named SARC-C, SARC-I, and SARC-O, respectively (Table 1). Protein purification was as described previously (9) using cation-exchange and heparin-affinity chromatography. As reported previously (8–19), the S. solfataricus proteome, including S. solfa-

---

²The abbreviations used are: PDB, Protein Data Bank; HMG, high-mobility group; PTM, post-translational modification; SARC, super acid–resistant Crenarchaeota; ESI, electrospray ionization; MRM, multiple-reaction monitoring.
three independent trials of acid adaptation were performed with various parental strains of genetically similar *S. solfataricus*: the SULA lineage is WT, SULG has a deletion from SULA_0784–0827, and SULM has a deletion from SULA_0784–0827 and SULA_1707–1708.

### Table 1

*S. solfataricus* strains

| Lineage   | Growth pH | Growth rate at pH 1.0 | Strain name (SUL/SARC-X) |
|-----------|-----------|-----------------------|--------------------------|
| Parental  | 3.0       | No growth             | A (98/2 (39))            |
| SARC (terminal) | 1.0     | 0.0563 ± 0.005         | C                       |

### Table 2

*S. solfataricus* 98/2 chromatin proteins and predicted masses

| Protein   | ORF no.  | Average molecular weight minus N-terminal Met |
|-----------|----------|---------------------------------------------|
| Cren7     | SULA_1986| 6531.70                                      |
| Sso7D-1   | SULA_0352| 7148.35                                      |
| Sso7D-2   | SULA_2718| 7147.37                                      |
| Alba-1    | SULA_1972| 10454.25                                     |
| Alba-2    | SULA_1974| 10107.68                                     |
| Sso10a    | SULA_0268| 11111.17                                     |
| Sso10a-2  | SULA_0634| 15243.74                                     |
| SSB       | SULA_0168| 16006.74                                     |

* The N-terminal methionine residue is cleaved from each protein in this list according to MS and MS2 data.

### Table 3

*S. solfataricus* strain 98/2 (42), contains eight small molecular mass chromatin proteins that range from 6.6 to 16.1 kDa (Table 2). Of these eight proteins, Sso7d exists in two copies (Sso7d1 and Sso7d2) encoded by Ssol_0360 and Ssol_2697, respectively, that vary in mass by 1 Da due to mutation of E14Q. Mass overlap with isotopic envelopes precluded the ability to distinguish between their intact protein masses by LC-MS. For this reason, the protein values that were the sum of both proteins were referred to herein as Sso7d.

Purified native chromatin protein extracts yielded three individual *A*$_{280}$ peaks by heparin-affinity chromatography (Fig. 1A). These peak fractions contained between one and three abundant proteins as indicated by SDS-PAGE analysis. The number and apparent masses for the proteins for the parental and SARC cell lines appeared identical when comparing respective heparin chromatography fractions. All of the abundant proteins had low apparent masses consistent with those of archaeal chromatin proteins (Fig. 1B), and their identities were determined by LC-MS/MS analysis of in-gel and in-solution tryptic digests. Peptides were identified for seven of the eight reported *S. solfataricus* chromatinins, including SSBS, Sso10a, Sso7d-1, Sso7d-2, Cren7, Alba-1, and Alba-2. These MS analyses demonstrate that all three heparin fractions contained Sso7d protein. In contrast, Cren7 was present in the first and third peaks, whereas SSB, Alba-1, Alba-2, and Sso10a occurred only in the last heparin fractions. Also, the changes in relative abundance of these proteins was determined by densitometry analysis using ImageJ as compared with a recombinant Cren7 control based on its known protein mass (Fig. S1). The abundance of Cren7 increased 2–3-fold in SARC-C as compared with SULA, whereas the abundance for Sso7d decreased nearly 2-fold (Fig. S1D) and the abundance of the Alba proteins (Sso10b and Sso10b2) remained unchanged.

### Intact protein analysis

To determine why Cren7 and Sso7d proteins appeared in multiple heparin peak fractions, post-translational differences were evaluated by LC-MS analysis of intact protein from individual heparin peak fractions of parental SULA and acid-adapted SARC-C cell lines. All Cren7 protein was monomethylated, whereas all Sso7d was unacetylated. Both had varying degrees of methylation. In heparin peak 1, Cren7 purified from parental cells (Fig. 2A) appeared as a mixture of four isoforms ranging from 1 to 4 methylations. In heparin peak 3 (Fig. 2B), Cren7 varied little in intact mass composition compared with peak 1, occurring as four isoforms with 1–4 methylations. Acid-adapted Cren7 (Fig. 2, C and D) occurred as four isoforms with 0–3 methylations in both heparin peaks 1 and 3. Parental Sso7d protein from heparin peak 1 (Fig. 3A) occurred as five versions ranging from 1 to 5 methyl additions. Sso7d in heparin peaks 2 and 3 (Fig. 3, B and C) occurred as six isoforms with 0–5 methylations. Acid-adapted Sso7d in heparin peaks 1, 2, and 3 (Fig. 3, D–F) all occurred as six isoforms ranging from 0 to 5 methylations. Although there was no significant difference in the number and abundance of modifications among the same protein from different heparin peaks, both Cren7 and Sso7d from SARC cell lines had a significantly greater population of protein in the unmethylated form. Sso10b (Alba1), Sso10b2 (Alba2), and Sso10a protein identified in heparin peak 3 did not show any variation in modification state between the parental and acid-adapted cell lines (Figs. S2, S3, and S4, respectively).

To exclude any potential effect of a transient cellular response to culture pH on the PTM state of the chromatin proteins, SARC-C was grown at pH 3 and cycled for 30 generations prior to purification of its native chromatin proteins and their subsequent analysis. Fourier-transform ion cyclotron resonance-MS analysis of intact protein from purified chromatin fractions from SULA and SARC-C cultivated at pH 3 and SARC-C cultivated at pH 1 indicated that the chromatin proteins from these samples exhibited different degrees of methylation. In the WT SULA proteome, Cren7 protein in its least post-translationally modified form contained one acetyl modification, as indicated by a +42 Da mass difference from its predicted unmodified mass. Up to four monomethylation modifications occurred on Cren7 based on peak mass shifts of +14 Da. Protein isoforms contained one, two, or three methyl groups as +14 Da mass-shifted forms (Fig. S5).

Significant differences were observed in the relative abundance of methylated Cren7 protein isoforms for SULA grown at pH 3 as compared with SARC-C grown at pH 1 and pH 3 with an overall reduction in all methylated isoforms in the evolved cell lines (Fig. 4). Isoforms varied in methylation content from one to five, as indicated by the presence of +14 Da mass shifts (Figs. S5). The percentage distribution of Cren7 methylations in SULA was 17.4, 39.1, 35.9, and 7.6% for zero, one, two, and three methylations, respectively. For SARC-C cultured at pH 1, the distribution was 50.4, 32.4, 14.4, and 2.9% for zero, one, two, and three methylations. These variations were not a stress.
**Heritable chromatin hypomethylation links archaeal phenomics**

*Figure 1. FPLC separation and SDS-PAGE analysis of SULA *S. solfataricus* chromatin proteins.* A, chromatogram of WT whole-cell native chromatin protein extract fractionated by heparin affinity. B, SDS-PAGE analysis of purified native chromatin extracts. Eluted peak fractions from heparin resins were pooled and analyzed for protein identity. Gel lanes are labeled according to respective heparin peak fractions loaded. *mAU,* milliabsorbance units.

*Figure 2. Intact mass analysis of Cren7 from individual heparin peaks.* Mass differences of isoforms of Cren7 from SULA grown at pH 3 from heparin peak 1 (*A*) and peak 3 (*B*) and Cren7 from SARC-C grown at pH 1 from heparin peak 1 (*C*) and peak 3 (*D*).
response to acid, as SARC strains cultivated at pH 3 retained a similar methylation distribution as when cultivated at pH 1. For SARC-C cultured at pH 3, the distribution was 40.1, 36, 20, and 3.9% for zero, one, two, and three methylations. Importantly, the shift in protein isoform content was highly conserved between evolved cell lines. These independently evolved lines are biological triplicates (40), indicating that occurrence of the same overall pattern of protein undermethylation was a common response to the adaptive laboratory evolution process for the trait of increased acid resistance.
Sso7d exhibited the same pattern of undermethylation in SARC-C as compared with SULA. In the WT SULA proteome, Sso7d existed as an unmodified protein with five monomethylation modifications. Isoforms varied in methylation content from one to five as indicated by the presence of 14 Da mass shifts (Fig. S6). SULA Sso7d protein occurred primarily with one or two methyl groups with a much smaller fraction of its protein in the unmodified form relative to SARC-C (Fig. 5). The percentage distribution of Sso7d methylations in SULA was 17.1, 27.6, 29.5, 17.1, and 5.7% for zero, one, two, three, and four methylations, respectively. For SARC-C cultured at pH 1, the distribution was 40.3, 31.1, 17.6, 8.1, and 2.9% for zero, one, two, three, and four methylations. For SARC-C cultured at pH 3, the distribution was 35.7, 31.7, 20.5, 9.9, 3.7% for zero, one, two, three, and four methylations.

In contrast to the findings for Cren7 and Sso7d, the pattern of modification of other chromatin proteins, Alba-1, Alba-2, and Sso10a, was invariant. No differences in PTM state of the other chromatin proteins were evident, comparing protein from the parental and evolved isolates. This indicated that the changes in PTM state were specific to methylation and specific to Cren7 and Sso7d rather than affecting the entire suite of *S. solfataricus* chromatin proteins or other types of PTMs.

### Chromatin proteins of independent SARC lineages

Three independent SARC lineages were produced using adaptive laboratory evolution (39–41). All of them exhibited nearly identical and substantive increases in thermoacidophily. Similar MS experimental studies were conducted to test whether the pattern of undermethylation also occurred in the other terminal SARC lineages as compared with their parental cell lines. Intact proteins from individual heparin peaks from all parental (SULA, SULG, and SULM) and pH 1–adapted SARC cell lines (SARC-C, SARC-I, and SARC-O) were analyzed by ESI-TOF MS. Importantly, the pattern of undermethylation of Cren7 and Sso7d was conserved in all three individually adapted cell lines, SARC-C, SARC-I, and SARC-O.

To determine the total fractions of Cren7 and Sso7d that existed in each methylated isoform for each strain, all heparin peaks were pooled for each respective strain, and the overall population of protein at each methylation state was determined based on raw mass peak area values in all mass spectra. Separately, each heparin fraction was pooled between strains and analyzed similarly to look at variation between pooled and individual fractions. Pooled fraction relative abundance values agreed with the individual peaks, and all indicated that Cren7 and Sso7d from SARC cell lines were consistently undermethylated regardless of the lineage of the adapted cell line examined (Figs. 4 and 5). This indicated that the conserved pattern of undermethylation occurred three independent times during the evolution of the SARC cell lines to extreme acid resistance.

The averaged percentage distribution of Cren7 methylations for parental strains was 12.2, 37.6, 35.7, 11.4, and 3% for zero, one, two, three, and four methylations, respectively. For SARC strains (SARC-C, SARC-I, and SARC-O) cultured at pH 1 and pH 3 (p = 0.02 and p = 0.01) compared with the averaged parental strains. In addition, the population of Cren7 with two methyl groups is significantly smaller in SARC strains at pH 1 (p = 0.05) and pH 3 (p = 0.05). Other methylated Cren7 populations were generally smaller in SARC strains but did not pass a significance test (p \( \leq 0.05 \)).

The averaged percentage distribution of Sso7d methylations for parental strains (SULA, SULG, and SULM) was 8.9, 18.3, 28.2, 26.7, 13.5, and 3.4% for zero, one, two, three, four, and five methylations, respectively. For SARC strains (SARC-C, SARC-I, and SARC-O) cultured at pH 1 and pH 3 (p = 0.02 and p = 0.01) compared with the averaged parental strains. In addition, the population of Sso7d with two methyl groups is significantly smaller in SARC strains at pH 1 (p = 0.05) and pH 3 (p = 0.05). Other methylated Sso7d populations were generally smaller in SARC strains but did not pass a significance test (p \( \leq 0.05 \)).
SARC-I, and SARC-O) cultured at pH 1, the distribution was 30.5, 28.5, 21.5, 13.2, 5.1, and 1.1%. For SARC cultured at pH 3, the distribution was 29, 30.1, 24, 12.4, 4.1, and 1% (Fig. 5 and Fig. S6). The population of unmethylated Sso7D is significantly larger in SARC strains at both pH 1 (p = 0.03) and pH 3 (p = 0.02) compared with the parental strain average. All other methylated populations of Sso7D were smaller in SARC strains compared with parental strains but did not pass a significance test (p ≤ 0.05).

These data indicated that the PTM state of these chromatin proteins was both heritable and conserved in all three SARC lineages. Importantly, the pattern of methylation was also not a stress response to extreme culture acidity, because the same PTM state of these proteins was observed using SARC cells cultivated at pH 1 or pH 3. For both Cren7 and Sso7D, it is likely that lysine methylations were lost from multiple populations of methylated protein. For Cren7, much came from the population with two methyl groups.

Site-specific post-translational modification analysis

As reported previously (9, 10, 12, 27, 43), many of the lysine but not arginine residues in the native protein sequences of both Cren7 and Sso7D exhibited PTM. To determine whether there were any site-specific differences in the post-translational modification states of the individual chromatin proteins, methylated residues were analyzed by LC-MS/MS analysis of in-solution protein digests. Protein digests identified four methylated lysines, Lys-11, -16, -42, and -48, in Cren7 and five methylated residues from Sso7D, Lys-7, -40, -53, -61, and -63 (Figs. S7–S12). To account for modifications that might be difficult to identify using MS/MS analysis, LC-MRM-MS was conducted to obtain higher sensitivity and quantitative measurements of abundance of each modified residue.

Cren7 protein proved to be difficult to analyze using multiple-reaction monitoring (MRM), despite testing many different enzymes for digests. Because the level of methylation is lower than Sso7D, there was more extensive trypsin cleavage, producing fragments too small to detect in the mass analyzer. As a result, MRM analysis was only successful when repeated using Sso7D protein from two SARC lineages to test reproducibility between the independent isolates. MRM corroborated the methylated residues identified using MS/MS and identified a new residue, Lys-5 (Fig. S13).

Four of the 12 Cren7 lysines, Lys-11, -16, -42, and -48, were identified as methylated in this work. This varies slightly from previously reported results (44). Lys-11, -16, and -42 were validated with high confidence in this work, whereas other residues (Lys-5 and Lys-31) could not be reliably identified. The novel methylated site, Lys-48, has not been described previously. Slight variation between the abundance of the individual methylations observed in these data and those from previously published results may arise from differences in strains or culturing conditions. Importantly, the four identified sites are both solvent-facing and clustered on the same part of Cren7, which may allow them to regulate protein–protein interactions.

With combined MS/MS and MRM data, 6 of the 14 lysine residues in Sso7D were identified as having methyl additions: Lys-5, -7, -40, -53, -61, and -63. These also vary from previously published methylated residues in both abundance and location (Fig. 6). Previous reports claimed that modifications occurred only on surface-exposed lysine residues, specifically Lys-4, -6, -60, -62, and -63 (10, 12), resulting in no apparent difference in DNA-binding affinity despite methylation (25). Note that this work’s amino acid numbering is greater by one amino acid when compared with previous studies, most likely due to an unconsidered cleavage of N-terminal methionine. The new data support most of the previously identified residues and indicate that all methylated residues were solvent-exposed. However, two novel sites, Lys-40 and Lys-53, were identified, and the site Lys-64 could not be reliably identified in Sso7D from the strains tested. MRM data analysis found that a number of these Sso7D sites had decreased abundance in SARC lines. For example, the novel Lys-40 modification was nearly absent in all SARC cells. Another difference between WT and SARC Sso7D was the abundance of methylation on the N-terminal lysines Lys-5 and Lys-7. WT cells were significantly more methylated on Lys-5 (30-fold greater) and Lys-7 (60-fold greater) as compared with SARC cells (Fig. 6). All other modified peptides at the C terminus had similar abundance values, indicating that the change in Sso7D in the acid-evolved cell lines was mainly N terminus–specific.

Protein–DNA interaction model

The 2.0 Å resolution X-ray crystal structure of Sso7D from Sulfolobus acidocaldarius in complex with the DNA 8-mer GTAATTAC (PDB accession code 1BNZ) (27) was analyzed to interpret the possible impact of lysine methylation on DNA binding. This complex was chosen because S. acidocaldarius Sso7D shares 100% amino acid sequence identity with Sso7D from S. solfataricus. Three lysine residues, 5, 7, and 40, were of special interest because they were found to be differentially methylated by MS and are shown in Fig. 7. The β-strand containing Lys-5 and Lys-7 (shown in orange in Fig. 7) also features interspersed acidic residues (Glu-12 and Glu-14) that compose a “spine” of interacting positive and negative charges that are
likely important for Sso7D thermostability. Therefore, the majority of the differentially methylated lysine residues in Sso7D do not reside at the DNA-binding surface.

Instead, all of the modified lysine residues are solvent-exposed; for example, Lys-7 is close enough to the phosphodiester backbone to make a contact with only a minor shift in side-chain position. Lys-9 and Lys-22 make more distant (~4.0 Å), primarily electrostatic contacts with the DNA backbone but were not identified as being aberrantly methylated by MS.

Discussion

Data presented here show that two abundant chromatin proteins from evolved cell lines of *S. solfataricus* adapted to extreme acid resistance are undermethylated relative to parental cell line controls. This pattern of undermethylation was evident in each of the three independent cell lines all evolved to achieve the same degree of extreme acid resistance. This indicates that undermethylation is a conserved feature accompanying the altered extremophile trait. In addition, like all of the other evolved SARC phenotypes, undermethylation was found to be heritable and not a physiologic stress response to exogenous acid. After cycling the acid-evolved cell lines without selection, the same pattern of undermethylation was present. The change in post-translational modification state was also chromatin protein–specific; whereas Cren7 and Sso7d were altered, Alba-1, Alba-2, and Sso10 were not, indicating that methylation, not acetylation, was affected in these cells.

Site-specific changes in post-translational modification state were identified that suggested the changes in methylation may play a role in protein–protein interactions. Six of the 14 lysine residues in Sso7d and 4 of the 12 in Cren7 were capable of methylation according to MS/MS and MRM-MS. The N terminus–specific changes that occurred in Sso7d affected two surface-exposed lysines involved in salt bridge formation. It has also been demonstrated that salt bridges that form between two protein monomers may contribute to the stabilization of protein–protein complexes (45). Methylated lysine residues on histones in eukaryotes are also known to form salt bridges in complex with other proteins (46). Similar to histones, HMG proteins undergo methylation, and their biological function is highly regulated by the location and type of modification (reviewed in Ref. 32). These PTMs can also induce or repress their potential interactions with other proteins in forming macromolecular complexes ultimately affecting the expression of certain genes at and downstream from binding sites (47, 48, 50–52). It is therefore interesting to speculate that methylation of Sso7d may serve a similar role in forming or altering protein–protein complexes on DNA and, as a result, alter gene expression. Ongoing ChIP-Seq experiments seek to identify binding locations of these proteins, their different patterns between parental and SARC lineages, and how binding changes correlate with gene expression values.

The mechanism of methylation pattern inheritance and the existence of reader/writer proteins requires investigation in archaea. These mechanisms may not be homologous to eukaryotic histone systems, as *Sulfolobus* appears to lack the Royal family of histone Lys-Met recognition domains (chromo, tudor, phd, etc.) and the SET domain that methylates histone tails (53). However, *Sulfolobus* does have various proteins that contain the methylation domains from the Dot1 histone-body methylase and has METTL10 and METTL21A, which are generic protein lysine methyltransferases (54). The aKMT lysine methyltransferase known to methylate Cren7 and other proteins in *vitro* (44) uses a Methyltransf_25 domain.

In previous studies, one of the SARC cell lines (SARC-I) acquired a heritable transcriptome and phenotype without undergoing genomic mutation, whereas the other lines suffered mutation but not within conserved genes (41). This implicated a non-Mendelian process, such as epigenetics, in trait inheritance. It was also found that homologous recombination without allele replacement at SARC transcriptome genes perturbed expression of those genes and reduced the acid resistance phenotype, without changing DNA sequence. This effect was locus-specific and did not occur when non-SARC transcriptome genes were targeted. As recombination machinery displaces chromatin proteins, these results supported a model in which chromatin proteins contributed to the regulation and heritability of traits. An alternative epigenetic mechanism could function through DNA methylation, whose role is still unknown in archaea (55). However, in the work presented here, the conserved, heritable hypomethylation of Cren7 and Sso7D for all three cell lines, and for unmutated SARC-I in particular, suggests that chromatin protein PTM state is a likely factor in this process. For example, because the affected chromatin proteins bind to the minor groove of DNA with no sequence specificity, their interactions with other DNA-binding proteins may confer selective genome interaction. The undermodification of Cren7 and Sso7D may perturb protein–protein interactions and thereby influence transcription at these loci. This may explain the transcriptomic differences and conservation that are apparent between SARC and WT cell lines while providing a mechanistic basis for the archaean epigenetic-like system.

Payne et al. (41) described an epigenetic-like system where heritable expression patterns and phenotypes can be acquired without mutation yet can be perturbed by homologous recombination at relevant genes. This work identifies epigenetic marks that could contribute to chromatin protein–mediated regulation. Together, these works provide the basis for an epigenetic system in archaea.
Materials and methods

Archaeal strains and cultivation

Parental *S. solfataricus* and SARC cell lines were cultivated as described previously (39–41) in basal salts medium (56) as modified by Brock. Complex medium was supplemented with 0.2% (w/v) tryptone and adjusted to the desired pH using sulfuric acid. Cultures (50 ml) were grown with aeration by agitation in screw-capped flasks at 80 °C to mid-exponential phase and subcultured into 2.0-liter flasks containing 500 ml of basal salts medium with 0.2% tryptone. Cultures were maintained aerobically at 80 °C with agitation (200 rpm), and cellular growth was monitored spectrophotometrically at 540 nm. For determination of PTM state during various growth states, cultures (50 ml) were grown with aeration by agitation at 75 °C to mid-exponential phase and subcultured into Applikon 3-liter bioreactors containing 2.0 liters of basal salts medium with 0.2% tryptone. Cultures were maintained aerobically at 75 °C with agitation (200 rpm) and supplemented with 1 volume of air per volume of medium (1 v/vm). Cells were harvested at mid-exponential phase (~0.5 OD) or early stationary phase (~1.0 OD), as indicated, by centrifugation at 5,000 g for 10 min. Cell pellets were stored at −20 °C until used for purification of proteins.

Chromatin isolation and purification

Native chromatin proteins used for intact mass analysis and MRM methylation site quantitation were purified using chromatographic methods. Cell pellets were thawed and resuspended in 30 mM sodium phosphate, 0.1 mM EDTA, 1 mM DTT, pH 6.6, buffer and lysed by intermittent sonication of 30-s duration for 30 min on ice. Cellular debris and insoluble proteins were pelleted by centrifugation at 13,000 × g for 15 min. Soluble protein was decanted and precipitated with 70% (w/v) ammonium sulfate final concentration. Precipitated proteins were resuspended in 30 mM sodium phosphate, 0.1 mM EDTA, 1 mM DTT, pH 6.6, buffer and injected into an AKTA purifier FPLC. Samples were loaded onto 1 ml of hand-packed Macro High-S Support (Bio-Rad) columns at 0.25 ml/min. Proteins were eluted from the column by a 25-ml linear gradient from 0 to 0.75 M potassium chloride. Peak fractions were pooled and dialyzed overnight into 30 mM sodium phosphate, 0.1 mM EDTA, 1 mM DTT, pH 6.6, buffer. Samples were fractionated identically using High-S a second time with a shallow gradient of 45 ml. Peak fractions were pooled, dialyzed, and purified using HiTrap heparin-affinity (GE Healthcare) 1-ml columns. Following heparin-affinity chromatography, peak fractions were either saved as individual peaks for compositional analysis or pooled for quantitative analysis of post-translation modification state of the purified chromatin proteins. Composition of the heparin fractions was analyzed by SDS-PAGE on 4–20% TGX (Bio Rad) pre-cast gels. All protein samples were dialyzed into 50 mM ammonium bicarbonate, pH 8.0, for MS analysis.

Native chromatin proteins used to identify methylated sites using LC-MS/MS could be quickly enriched in large quantities using immunoprecipitation. Polyclonal antibodies (Pierce) prepared in rabbits using recombinant Cren7 or Sso7D were incubated with Protein A–conjugated Sepharose beads (Invitrogen) for 1 h and washed three times with PBS. Cell mass equivalent to 2 × 10^10 cells was resuspended in lysis buffer (50 mM HEPES-KOH, pH 7.5, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate) and lysed using a cup-horn sonicator. Cell lysate was clarified at 15,000 × g for 20 min, and the supernatant was incubated for 1 h with the antibody bound to beads at room temperature with agitation. Beads were washed five times with PBS, and the captured protein was eluted using 50 mM Tris-HCl, pH 8, 10 mM EDTA, 1% SDS. Protein was digested with Trypsin Gold (Promega) and purified using an S-Trap minicolumn (Protifi) following the manufacturer’s protocol. Samples were dried using a SpeedVac, stored at −80 °C, and then reconstituted in 50 μl of 5% formic acid in water prior to LC injection. Antibody specificity for Cren7 or Sso7D was validated using MS.

Preparation of recombinant Sso7d and Cren7

The *S. solfataricus* Sso7d ORF (SULA_0352) and Cren7 ORF (SULA_1986) were amplified from genomic DNA using Phusion High-Fidelity DNA polymerase (New England Biolabs) and oligonucleotide primers Sso7d-F and Sso7d-R and primers Cren7-F and Cren7-R, respectively. The PCR amplicons were digested with Ncol and Xhol and ligated into pET28b (Novagen), creating plasmids pPB1311 (Sso7d) and pPB1109 (Cren7).

*Escherichia coli* strain BL21 Rosetta (Novagen) was transformed with pPB1311 and pPB1109 for expression of recombinant C-terminal hexahistidine-tagged Sso7d and Cren7, respectively. For preparation of recombinant Sso7d and Cren7, cells were grown to an optical density at 600 nm of 0.3, and expression was induced by the addition of 1 mM isopropyl-β-D-thiogalactopyranoside followed by a 3-h additional incubation. Cells were harvested by centrifugation, resuspended in 20 mM sodium phosphate, 0.5 M NaCl, 20 mM imidazole, 1 mM β-mercaptoethanol, pH 7.4 (buffer A), and lysed by sonication followed by centrifugation. The soluble extract was heated at 80 °C for 20 min and centrifuged at 10,000 × g for 20 min to remove denatured proteins. The supernatant was applied to a 1.0-ml HisTrap HP column (GE Healthcare). The column was equilibrated with 10 column volumes of buffer A, and supernatant was then applied to the column followed by a wash with 10 column volumes of buffer A. Bound protein was eluted using a linear gradient from 0 to 0.5 M imidazole over 20 ml. Purity of protein was analyzed by SDS-PAGE after staining with Coomassie Blue G250. Protein was concentrated and dialyzed into PBS using Centricron YM-3 (Amicon) filters prior to storage of aliquots at −80 °C.

LC-MS

Heparin-purified chromatin extracts were separated by reverse-phase HPLC using a Waters nanoAcquity 1.8-μm HSS T3 C18 column (75-μm inner diameter × 250-mm length) on a Waters UPLC system. Separations were completed using a flow rate of 500 μl/min with mobile phase A consisting of 0.1% formic acid in water and mobile phase B as 0.1% formic acid in acetonitrile. Initial conditions were held at 5% B for 2 min followed by an increase to 80% B over 12 min. Conditions were held at 80% B for 1 min before returning to 5% B for 8 min for column equilibration. Samples were introduced into the mass
Heritable chromatin hypomethylation links archaeal phenomics

spectrometer using the nanospray source. ESI was performed at 3 kV capillary voltage, 80 °C source temperature, and nebulizer gas flow at 4.8 bar. Mass spectra were acquired on a Waters G2S Synapt TOF (Waters) in the positive-ion mode from \( m/z \) 410 to 3,500. All ESI \( m/z \) spectra of intact proteins were deconvoluted to molecular mass using the MaxEnt 1 algorithm in MassLynx (version 4.1) software.

**Fourier-transform ion cyclotron resonance MS**

Approximately 0.5 µg of protein sample was loaded onto a ProSwift RP-10R column (1.0 mm × 5 cm) (Thermo Scientific). A gradient was developed over 90 min at a flow rate of 20 µl/min using an Agilent 1100 LC (Agilent Technologies), with solvent A consisting of 0.1% formic acid in \( H_2O \) and solvent B consisting of 0.1% formic acid in acetonitrile. The gradient consisted of initial conditions of 0% B held for 1 min, followed by linear increases to 80% B at 75 min and 100% B at 78 min. 100% B was held for 1 min, and the system was returned to initial conditions at 80 min and held for 10 min. Mass spectra were acquired on a Bruker Solarix 7.05-tesla Fourier-transform ion cyclotron resonance mass spectrometer (Bruker Daltonics). Profiled spectra were acquired in the positive mode from \( m/z \) 150.59 to 3,000 with an estimated resolving power of 200,000 at \( m/z \) of 400. Prior to acquisition, the mass spectrometer was calibrated using a 0.1 mg/ml NaTFA solution. Spectra were deconvoluted, and intact mass was determined using Compass DataAnalysis version 4.0 (Bruker Daltonics). Theoretical isotope distributions were calculated using the IDCalc software, which is based on the method described by Kubinyi (57).

**LC-MS/MS**

The trypic digest samples purified using immunoprecipitation were subjected to nano-LC-MS/MS analysis on a Sciex 6600 Triple-TOF with Spray source and Calibrant Delivery System coupled to an Ultimate 3000 nano-LC system (Dionex Corp.). The peptides were separated on a C18 Pep Map column (1-mm inner diameter × 150-mm length, 3-µm particle size, 100-Å pore; Dionex) by applying an acetonitrile gradient (acetonitrile plus 0.1% formic acid, 60-min gradient of 0–40%, then a 2-min gradient of 30–40%, and then held at 40% for 5 min) at a flow rate of 50 µl/min, and introduced into the mass spectrometer using the nanospray source. All MS methods for the 6600 Triple-TOF used the information-dependent acquisition mode. The 6600 Triple-TOF was operated with the following parameters: nanospray voltage, positive mode; survey scan range, 400–1,250 \( m/z \); MS/MS scan range, 100–1,500 \( m/z \); 48 MS/MS scans/cycle; rolling collision energy. The MS/MS spectra were searched against the S. solfataricus proteome from UniProt using ProteinMetrics Byonic and Byologic (version 3.2). Database search criteria were as follows: enzyme, trypsin; missed cleavages, 2; mass, monoisotopic; fixed modification, carbamidomethyl (C); variable modification, acetylation (K) and methylation (K); precursor tolerance, 20 ppm; product ion tolerance, 60 ppm.

**LC-MRM-MS**

An MRM assay was developed to confirm methylation of specific Lys residues in these proteins. Protein sequences were imported into a Skyline file with Lys residues containing a variable methylation modification. A theoretical digest with LysC was used to generate a list of peptides for each protein, with LysC not being able to cleave at methylated Lys residues. From this list of peptides, peptides containing up to three consecutive methylated Lys residues were accepted for the MRM assay. Through inspection of the peptide sequences, a minimum of three transitions for each possible peptide were selected, and a scouting LC-MRM analysis of LysC-digested samples was performed. For peptides with a positive signal, additional transitions (minimum of 5/peptide) were added to improve specificity and sensitivity.

LC-MRM experiments were performed on a Dionex U3000 HPLC and AB-Sciex 4000 QTrap mass spectrometer. Chromatographic separations were carried out on a Dionex Acclaim C18 Pepmap column (1.0-mm inner diameter × 150-mm length; Thermo Fisher Scientific). Chromatography solvents were 0.1% formic acid (Fluka) in water (Solvent A) and 0.1% formic acid in acetonitrile (Honeywell). The column oven was maintained at 60 °C. Samples were injected in 0% B and separated by a linear gradient of 0–40% B over 40 min. MRM analysis was performed in positive mode. Transitions used a collision energy that was optimized for this particular QTrap MS instrument from a large library of peptides. Raw data files were imported into Skyline for analysis. Acceptance of MRM results required (i) co-elution of all transitions, (ii) quality peaks that were clearly demarcated from the background and a signal greater than 3 times the local noise, and (iii) consistent ranking of transition peak intensities among replicate samples.

**Author contributions**—T. J., S. M., S. P., K. V. C., and P. B. designed experiments. T. J., R. G., S. M., S. P., E. O., and C. M. performed experiments. J. A., R. G., S. M., S. P., E. O., and C. M. performed experiments. T. J., R. G., S. M., S. P., E. O., and C. M. performed data analysis and interpretation. T. J., S. P., M. A. W., and P. B. wrote the paper.

**References**

1. Talbert, P. B., and Henikoff, S. (2010) Histone variants—ancient wrap artists of the epigenome. *Nat. Rev. Mol. Cell Biol.* 11, 264–275 CrossRef Medline
2. Marín-Ramírez, L., Levine, K. M., Morales, M., Zhang, S., Moreland, R. T., Baxevanis, A. D., and Landsman, D. (2011) The Histone database: an integrated resource for histones and histone fold-containing proteins. *Database (Oxford)* 2011, bar048 CrossRef Medline
3. Sandman, K., and Reeve, J. N. (2006) Archaeal histones and the origin of the histone fold. *Curr. Opin. Microbiol.* 9, 520–525 CrossRef Medline
4. Cubonová, L., Sandman, K., Hallam, S. J., Delong, E. F., and Reeve, J. N. (2005) Histones in crenarchaea. *J. Bacteriol.* 187, 5482–5485 CrossRef Medline
5. Forbes, A. J., Patrie, S. M., Taylor, G. K., Kim, Y. B., Jiang, L., and Kelleher, N. L. (2004) Targeted analysis and discovery of posttranslational modifications in proteins from methanogenic archaea by top-down MS. *Proc. Natl. Acad. Sci. U.S.A.* 101, 2678–2683 CrossRef Medline
6. 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 CrossRef Medline
7. Edmondson, S. P., Qiu, L., and Shriver, J. W. (1995) Solution structure of the DNA-binding protein Sac7d from the hyperthermophile *Sulfolobus acidocaldarius*. *Biochemistry* 34, 13289–13304 CrossRef Medline
Heritable chromatin hypomethylation links archaeal phenomics

8. Edmondson, S. P., and Shriver, J. W. (2001) DNA binding proteins Sac7d and Sso7d from Sulfolobus. Methods Enzymol. 334, 129–145 CrossRef Medline

9. Guo, L., Feng, Y., Zhang, Z., Yao, H., Luo, Y., Wang, J., and Huang, L. (2008) Biochemical and structural characterization of Cren7, a novel chromatin protein conserved among Crenarchaea. Nucleic Acids Res. 36, 1129–1137 CrossRef Medline

10. Baumann, H., Knapp, S., Lundback, T., Ladenstein, R., and Hard, T. (1994) Solution structure and DNA-binding properties of a thermostable protein from the archaean Sulfolobus solfataricus. Nat. Struct. Biol. 1, 808–819 CrossRef Medline

11. Biyani, K., Kahsai, M. A., Clark, A. T., Armstrong, T. L., Edmondson, S. P., and Shriver, J. W. (2005) Solution structure, stability, and nucleic acid binding of the hyperthermophile protein Sso10b2. Biochemistry 44, 14217–14230 CrossRef Medline

12. 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 Archaeabacterium, Sulfolobus solfataricus. Biochim. Biophys. Acta 950, 193–203 CrossRef Medline

13. Dijk, J., and Reinhardt, R. (1986) The structure of DNA-binding proteins from Eu- and Archaeabacteria. in Bacterial Chromatin (Gualerzi, C. O., and Pon, C. L., eds) pp. 184–218, Springer-Verlag, Berlin

14. Kahsai, M. A., Vogler, B., Clark, A. T., Edmondson, S. P., and Shriver, J. W. (2005) Solution structure, stability, and flexibility of Sso10a: a hyperthermophile coiled-coil DNA-binding protein. Biochemistry 44, 2822–2832 CrossRef Medline

15. Sandman, K., and Reeve, J. N. (2005) Archaeal chromatin proteins: different structures but common function? Curr. Opin. Microbiol. 8, 656–661 CrossRef Medline

16. Sun, F., and Huang, L. (2013) Sulfolobus chromatin proteins modulate strand displacement by DNA polymerase B1. Nucleic Acids Res. 41, 8182–8195 CrossRef Medline

17. Wardleworth, B. N., Russell, R. J., Bell, S. D., Taylor, G. L., and White, M. F. (2002) Structure of Alba: an archaeal chromatin protein modulated by acetylation. EMBO J. 21, 4654–4662 CrossRef Medline

18. White, M. F., and Bell, S. D. (2002) Holding it together: chromatin in the Archaea. Trends Genet. 18, 621–626 CrossRef Medline

19. Zhang, Z., Gong, Y., Guo, L., Jiang, T., and Huang, L. (2010) Structural insights into the interaction of the crenarchaeal chromatin protein Cren7 with DNA. Mol. Microbiol. 76, 749–759 CrossRef Medline

20. Jelinska, C., Conroy, M. J., Craven, C. J., Hounslow, A. M., Archer, S. J., Ball, L. J., Murzina, N. V., Broadhurst, R. W., Raine, A. R., and Sso7d from Sulfolobus solfataricus. Proc. Natl. Acad. Sci. U.S.A. 115, 12271–12276 CrossRef Medline

21. Napoli, A., Zivanovic, Y., Bocs, C., Bhuler, C., Rossi, M., Porterre, P., and Ciaramella, M. (2002) DNA bending, compaction and negative supercoiling by the architectural protein Sso7d of Sulfolobus solfataricus. Nucleic Acids Res. 30, 2656–2662 CrossRef Medline

22. Ball, L. J., Murzina, N. V., Broadhurst, R. W., Raine, A. R., Archer, S. J., Stott, F. J., Murzina, A. G., Singh, P. B., Domaile, P. J., and Laue, E. D. (1997) Structure of the chromatin binding (chromo) domain from mouse modifier protein 1. EMBO J. 16, 2473–2481 CrossRef Medline

23. Guagliardi, A., Cerchia, L., Moracci, M., and Rossi, M. (2000) The chromosomal protein Sso7d of the crenarchaeon Sulfolobus solfataricus rescues aggregated proteins in an ATP hydrolysis-dependent manner. J. Biol. Chem. 275, 31813–31818 CrossRef Medline

24. Scholert, J., Dixit, V., Hoang, V., Simbahan, J., Drozda, M., and Blum, P. (2004) Occurrence and characterization of mercury resistance in the hypertherophilic archaean Sulfolobus solfataricus by use of gene disruption. J. Bacteriol. 186, 427–437 CrossRef Medline

25. Lundback, T., Hansson, H., Knapp, S., Ladenstein, R., and Hard, T. (1998) Thermodynamic characterization of non-sequence-specific DNA-binding by the Sso7d protein from Sulfolobus solfataricus. J. Mol. Biol. 276, 775–786 CrossRef Medline

26. Feng, Y., Yao, H., and Wang, J. (2010) Crystal structure of the crenarchaeal conserved chromatin protein Cren7 and double-stranded DNA complex. Protein Sci. 19, 1253–1257 CrossRef Medline

27. Robinson, H., Gao, Y. G., McCrary, B. S., Edmondson, S. P., Shriver, J. W., and Wang, A. H. (1998) The hyperthermophile chromosomal protein Sac7d sharply kinks DNA. Nature 392, 202–205 CrossRef Medline

28. Teale, M. J., Kahsai, M., Singh, S. K., Edmondson, S. P., Gupta, R., Shriver, J. W., and Meehan, E. (2003) Cloning, expression, crystallization and preliminary X-ray analysis of the DNA-binding domain Sso10a from Sulfolobus solfataricus. Acta Crystallogr. D Biol. Crystallogr. 59, 1320–1322 CrossRef Medline

29. Thomas, J. O., and Travers, A. A. (2001) HMGI1 and 2, and related “architectural” DNA-binding proteins. Trends Biochem. Sci. 26, 167–174 CrossRef Medline

30. Tjian, R., and Maniatis, T. (1994) Transcriptional activation: a complex puzzle with few easy pieces. Cell 77, 5–8 CrossRef Medline

31. Zhang, J., McCauley, M. J., Maher L. J 3rd, Williams, M. C., and Israeloff, N. E. (2009) Mechanism of DNA flexibility enhancement by HMGB proteins. Nucleic Acids Res. 37, 1107–1114 CrossRef Medline

32. Zhang, Q., and Wang, Y. (2010) HMGI modifications and nuclear function. Biochim. Biophys. Acta 1799, 28–36 CrossRef Medline

33. Cleynen, I., and Vande Ven, W. J. (2008) The HMGA proteins: a myriad of functions (review). Int. J. Oncol. 32, 289–305 Medline

34. McCarthy, S., Vande Ven, W. J. (2005) Obligate heterodimerization of the archaeal Alba2 protein with Alba1 provides a structural “ DNA-binding proteins. Biochim. Biophys. Acta 1799, 28–36 CrossRef Medline

35. Dijk, J. and Reinhardt, R. (1986) The structure of DNA-binding proteins from Eu- and Archaeabacteria. in Bacterial Chromatin (Gualerzi, C. O., and Pon, C. L., eds) pp. 184–218, Springer-Verlag, Berlin

36. White, M. F., and Bell, S. D. (2002) Holding it together: chromatin in the Archaea. Trends Genet. 18, 621–626 CrossRef Medline

37. McCarthy, S., Gradigno, J., Johnson, T., Payne, S., Lipzen, A., Martin, J., Schackwitz, W., Moriyama, E., and Blum, P. (2015) Complete genome sequence of Sulfolobus solfataricus strain 98/2 and evolved derivatives. Genom. Annu. Ann. 3, 40549–15 Medline

38. Irvine, K., Sun, F., and Blum, P. (1995) Purification and characterization of a maltase from the extremely thermophilic crenarchaeote Sulfolobus solfataricus. J. Bacteriol. 177, 482–485 CrossRef Medline

39. Zhang, K., Zheng, S., Yang, J., Shen, Y., and Cheng, Z. (2013) Comprehensive profiling of protein lysine acetylation in Escherichia coli. J. Proteome Res. 12, 844–851 CrossRef Medline

40. Yu, D., Tsi, C. J., and Nussinov, R. (1997) Hydrogen bonds and salt bridges across protein-protein interfaces. Protein Eng. 10, 999–1012 CrossRef Medline

41. Lee, J., Thompson, J. R., Botuyan, M. V., and Mer, G. (2008) Distinct binding modes specify the recognition of methylated histones H3K4 and H4K20 by JMJD2A-tudor. Nat. Struct. Mol. Biol. 15, 109–111 CrossRef Medline

42. Bianchi, M. E., and Agresti, A. (2005) HMG proteins: dynamic players in gene regulation and differentiation. Curr. Opin. Genet. Dev. 15, 496–506 CrossRef Medline

43. Frasca, F., Rustighi, A., Malaguarnera, R., Altamura, S., Vigneri, P., Del Sal, G., Giancotti, V., Pezzino, V., Vigneri, R., and Manfioletti, G. (2006)

J. Biol. Chem. (2019) 294(19) 7821–7832 7831
HMGA1 inhibits the function of p53 family members in thyroid cancer cells. *Cancer Res.* 66, 2980–2989 CrossRef Medline

Hyun, K., Jeon, J., Park, K., and Kim, J. (2017) Writing, erasing and reading histone lysine methylations. *Exp. Mol. Med.* 49, e324 CrossRef Medline

Hamamoto, R., Saloura, V., and Nakamura, Y. (2015) Critical roles of non-histone protein lysine methylation in human tumorigenesis. *Nat. Rev. Cancer* 15, 110–124 CrossRef Medline

Couturier, M., and Lindás, A. C. (2018) The DNA methylome of the hyperthermooacidophilic crenarchaeon *Sulfolobus acidocaldarius*. *Front. Microbiol.* 9, 137 CrossRef Medline

Zinder, S. H., and Brock, T. D. (1978) Production of methane and carbon dioxide from methane thiol and dimethyl sulfide by anaerobic lake sediments. *Nature* 273, 226–228 CrossRef

Kubinyi, H. (1991) Calculation of isotope distributions in mass-spectrometry: a trivial solution for a nontrivial problem. *Anal. Chim. Acta* 247, 107–119 CrossRef

Heritable chromatin hypomethylation links archaeal phenomics

Couturier, M., and Lindás, A. C. (2018) The DNA methylome of the hyperthermooacidophilic crenarchaeon *Sulfolobus acidocaldarius*. *Front. Microbiol.* 9, 137 CrossRef Medline

Zinder, S. H., and Brock, T. D. (1978) Production of methane and carbon dioxide from methane thiol and dimethyl sulfide by anaerobic lake sediments. *Nature* 273, 226–228 CrossRef

Kubinyi, H. (1991) Calculation of isotope distributions in mass-spectrometry: a trivial solution for a nontrivial problem. *Anal. Chim. Acta* 247, 107–119 CrossRef

Hyun, K., Jeon, J., Park, K., and Kim, J. (2017) Writing, erasing and reading histone lysine methylations. *Exp. Mol. Med.* 49, e324 CrossRef Medline

Hamamoto, R., Saloura, V., and Nakamura, Y. (2015) Critical roles of non-histone protein lysine methylation in human tumorigenesis. *Nat. Rev. Cancer* 15, 110–124 CrossRef Medline

Couturier, M., and Lindás, A. C. (2018) The DNA methylome of the hyperthermooacidophilic crenarchaeon *Sulfolobus acidocaldarius*. *Front. Microbiol.* 9, 137 CrossRef Medline

Zinder, S. H., and Brock, T. D. (1978) Production of methane and carbon dioxide from methane thiol and dimethyl sulfide by anaerobic lake sediments. *Nature* 273, 226–228 CrossRef

Kubinyi, H. (1991) Calculation of isotope distributions in mass-spectrometry: a trivial solution for a nontrivial problem. *Anal. Chim. Acta* 247, 107–119 CrossRef