ARTICLE

https://doi.org/10.1038/s41467-019-09479-1

A TetR-family transcription factor regulates fatty acid metabolism in the archaeal model organism *Sulfolobus acidocaldarius*

Kun Wang¹, David Sybers², Hassan Ramadan Maklad², Liesbeth Lemmens², Charlotte Lewyllie²,⁵, Xiao Xiao Zhou³, Frank Schult³, Christopher Bräsen³, Bettina Siebers³, Karin Valegård⁴, Ann-Christin Lindås¹ & Eveline Peeters²

Fatty acid metabolism and its regulation are known to play important roles in bacteria and eukaryotes. By contrast, although certain archaea appear to metabolize fatty acids, the regulation of the underlying pathways in these organisms remains unclear. Here, we show that a TetR-family transcriptional regulator (FadRSa) is involved in regulation of fatty acid metabolism in the crenarchaeon *Sulfolobus acidocaldarius*. Functional and structural analyses show that FadRSa binds to DNA at semi-palindromic recognition sites in two distinct stoichiometric binding modes depending on the operator sequence. Genome-wide transcriptomic and chromatin immunoprecipitation analyses demonstrate that the protein binds to only four genomic sites, acting as a repressor of a 30-kb gene cluster comprising 23 open reading frames encoding lipases and β-oxidation enzymes. Fatty acyl-CoA molecules cause dissociation of FadRSa binding by inducing conformational changes in the protein. Our results indicate that, despite its similarity in overall structure to bacterial TetR-family FadR regulators, FadRSa displays a different acyl-CoA binding mode and a distinct regulatory mechanism.

¹Department of Molecular Biosciences, The Wenner-Gren Institute, Stockholm University, Svante Arrhenius v. 20C, SE-10691 Stockholm, Sweden. ²Research Group of Microbiology, Department of Bioengineering Sciences, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussels, Belgium. ³Molekulare Enzymtechnologie und Biochemie, Biofilm Centre, ZWU, Fakultät für Chemie, Universität Duisburg-Essen, Universitätsstr. 2, 45117 Essen, Germany. ⁴Molecular Biophysics, Department of Cell and Molecular Biology, Uppsala University, Uppsala, Sweden. ⁵Present address: Laboratory of Cell Genetics, Department of Biology, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussels, Belgium. Correspondence and requests for materials should be addressed to A.-C.L. (email: Ann.Christin.Lindas@su.se) or to E.P. (email: Eveline.Peeters@vub.be)
he phylogenetic classification of archaea as a domain of life distinct from bacteria is supported by the nature of their membrane lipids having isoprenoid-based hydrocarbon chains instead of fatty acids. Despite the absence in archaenal membrane lipids, small amounts of fatty acids and derivatives have been detected in archaenal cells. The role of fatty acids for archaenal cellular physiology is not yet clear and a controversial issue of debate, although an involvement in the acylation or stabilization of membrane-bound energy-conversion proteins such as rhodopsin or cytochromes has been postulated. Many archaenal genomes have extensive sets of typical bacterial-like genes encoding fatty acid synthase type II (FAS-II) complex and β-oxidation enzymes. An outstanding question is whether these fatty acid metabolism genes perform anabolic or catabolic reactions, or both. Given the absence of genes encoding acyl-carrier protein (ACP) or ACP synthase, it has been postulated that a β-oxidation pathway might operate in the reverse direction in conjunction with acetyl-CoA C-acetyltransferase enzymes. These are abundantly encoded in archaenal genomes, sometimes in the direct neighborhood of β-oxidation genes.

Despite the abundance of fatty acid metabolism genes in many genomes, nothing is known about how their expression is regulated in archaea. In contrast, this is well characterized in bacteria, in which a tight regulation of the synthesis and degradation of fatty acids involves multiple transcription regulators that act in response to intracellular fatty acid-related metabolic signals. In Gram-negative bacteria a GntR-family regulator FadR has a dual role by coordinately repressing β-oxidation genes while activating FAS-II genes in response to acyl-CoA molecules, whereas a TetR-family malonyl-CoA-dependent regulator FabR controls the ratio between mono-unsaturated and saturated fatty acids. Gram-positive bacteria such as Bacillus subtilis use an identically named transcription factor FadR that belongs to the TetR family for the acyl-CoA dependent regulation of β-oxidation degradations and a DeoR family member FabP that regulates biosynthesis of saturated fatty acids and phospholipids. The mechanism of action of the bacterial acyl-CoA responsive TetR-like regulator has been unraveled by analysis of apo, ligand-bound, and DNA-bound crystal structures.

In this work, we focus on characterizing the transcriptional regulation of genes encoding fatty acid metabolism functions in the thermoacidophilic crenarchaeon Sulfolobus acidocaldarius, which is genetically tractable and considered to be a major archaenal model organism. S. acidocaldarius has an extensive gene cluster, comprising genes Saci_1103 until Saci_1126, encoding homologs of the three β-oxidation enzymes acyl-CoA dehydrogenase, enoyl-CoA hydratase, and hydroxyacyl-CoA dehydrogenase. Also, genes encoding members of the thiolase superfamily presumably catalyzing the last step of the β-oxidation cycle, i.e., ketoacyl-CoA thiolases as well as acetyl-CoA acetyltransferases were identified within the cluster. In addition, genes encoding lipid degradation functions are present in this genomic region. Concerning these latter functions, Saci_1105 and Saci_1116 code for enzymes that were experimentally shown to display esterase activity. The Saci_1103-Saci_1126 gene cluster also comprises a gene, Saci_1107, encoding a predicted TetR-like transcription factor for which we hypothesized that it might be involved in regulating the expression of the gene cluster. We aim at performing structural, biochemical, genetic and genomic analyses of this regulator, named FadRsa, thereby unveiling the function and mode of action of an acyl-CoA-responsive transcriptional regulator in an archaenal microorganism.

Results
FadRsa structure. S. acidocaldarius harbors a 30-kb gene cluster consisting of genes encoding enzymes involved in lipid and fatty acid metabolism and a putative regulator (Saci_1107, Fig. 1a). As a first step towards functional characterization of this regulator, we performed a crystallographic analysis of the protein encoded by Saci_1107 (Figure 1b and Table 1). Size exclusion chromatography (SEC) indicated that the purified recombinant protein behaves as a homogenous population of 44-kDa sized dimers (Supplementary Figure 1). The asymmetric unit of the 2.4-Å resolution crystal structure also contains a homodimer with an exclusive alpha-helical structure. Each subunit displays two functional domains: an N-terminal helix-turn-helix (HTH) DNA-binding domain (α1–α3) and a C-terminal domain (α4–α9) of which α8 and α9 stabilize dimerization. The overall Ω-shape structure of the dimer validates its classification as a TetR family member.

Although BLAST analyses initially did not reveal which bacterial regulators could be considered as potential functional homologs for the protein encoded by Saci_1107, a superposition revealed structural similarity with the previously characterized TetR-family FadR transcription regulators in Bacillus sp., FadRb (RMSD = 4.23 Å) and FadRc (RMSD = 5.88 Å), and Thermus thermophilus, FadRd (RMSD = 11.85 Å) (Fig. 1c). Conservation is significantly higher for the N-terminal than for the C-terminal domains (Supplementary Table 1) as also confirmed by a structure-based sequence alignment (Fig. 1d). This structural similarity led us to propose to name this protein FadRsa accordingly.

Upon solving the FadRsa crystal structure, one of the subunits (subunit B) was found to have additional unassigned electron density in the C-terminal domain. This could be explained by fitting it with an acyl-CoA molecule (Fig. 1b), which was likely derived from Escherichia coli during heterologous overexpression. The best fit was obtained with heptanoyl-CoA. Given the low intracellular abundance of odd-chained acyl-CoA molecules it is possible that a mixture of even-chained short-chain acyl-CoA molecules was present in the ligand binding pockets of different protein molecules packed in the crystal. The unintended cocRYystallization of acyl-CoA with FadRsa (Fig. 1b) suggests that it is a specific ligand of the protein. This further supports the hypothesis that the regulatory role of this transcription factor is connected to acyl-CoA metabolism.

Genome-wide DNA-interaction map of FadRsa. As a next step toward unraveling FadRsa function, we employed chromatin immunoprecipitation (ChIP) in combination with next-generation sequencing (ChIP-seq). A total of 14 significant and reproducible in vivo-associated genomic loci were identified (Fig. 2a and Supplementary Table 2). The two highest enrichment scores were observed within the Saci_1103-Saci_1126 gene cluster. More specifically, both high-enrichment binding regions were located within the intergenic region of the divergently organized operon encoding the fadRsa gene itself and a putative esterase-encoding gene (peaks 1 and 2). Within the Saci_1103-Saci_1126 gene cluster, two additional low-enrichment binding regions were observed within the coding sequence of gene Saci_1115 and in the intergenic region separating a divergently encoded β-oxidation operon and a putative transcription factor gene, respectively (peaks 3 and 4). Targeted chromatin immunoprecipitation quantitative polymerase chain reaction (ChIP-qPCR) validated the observed enrichments (Fig. 2b), which were not observed anymore upon deleting fadRsa (Supplementary Figure 2). All sequences enriched in the ChIP-seq analysis were subjected to a computational binding motif prediction, yielding a 16-base pair
(bp) motif with dyad symmetry that is conserved in 13 of 14 binding regions (Fig. 2c and Supplementary Table 2). Besides the four experimentally identified binding sites, an in silico screening revealed three additional putative binding sites in the gene cluster, of which one is located within the open reading frame (ORF) of Saci_1106 (Supplementary Figure 3, Supplementary Table 3). Possibly they were not captured in the ChIP-seq analysis but are functional in other conditions.

Electrophoretic mobility shift assays (EMSAs) with DNA probes encompassing the centers of the binding regions verified that the observed ChIP-seq enrichment regions represent direct and specific FadRSa–DNA interactions (Fig. 2d, Supplementary Table 2 and Supplementary Figure 4). Densitometric analysis of EMSA autoradiographs performed with the high-enrichment targets revealed a formation of two electrophoretically distinct FadRSa–DNA complexes with high affinity and positive-binding interaction.
Mechanisms of DNA binding. To further unravel mechanisms of DNA binding by FadRSa, we determined the cocystal structure of the protein–DNA complex to a resolution of 3.29 Å using a duplex DNA containing the predicted FadRSa binding motif in the control region of the fadRSa gene itself, corresponding to ChIP-seq peak 2 (Table 1 and Fig. 3a, b). The asymmetric unit contained six FadRSa subunits, organized as three dimers, and two DNA duplex molecules thus representing a nonbiological assembly (Fig. 3a), although the protein–DNA molecular interactions within this structure are representative of the biologically relevant complexes (see below, DNA-binding stoichiometry of FadRSa).

In each FadRSa subunit in the cocystal structure, residues of the recognition helix α3 and the α2–α3 loop interact with the major groove of DNA with the establishment of an extensive number of contacts (Fig. 3b, Supplementary Note 1 and Supplementary Data 1). Base-specific contacts mainly consist of hydrophobic interactions between FadRSa residues Tyr47, Leu49, Tyr51, and Phe52 and methyl groups of thymines (Fig. 3c) similar as in other TetR-like regulators (Supplementary Figure 5), in addition to electrostatic interactions between Gly48 and the N7 group of guanines. The role of these residues for DNA binding was further investigated by performing site-directed alanine substitution and analyzing the mutant proteins in EMSA (Supplementary Figure 6a). FadRSaY47A, FadRSaY51A, and FadRSaY53A are all negatively affected in DNA-binding affinity and cooperativity. With FadRSaG48A, no DNA binding was observed at all demonstrating that Gly48 is a crucial residue (Supplementary Figure 6a). Besides protein–DNA contacts, a weak electrostatic protein–protein contact was also observed between Asn37 residues of FadRSa dimers bound on different sides of the DNA helix (Fig. 3d and Supplementary Note 1).

Table 1 Data collection and refinement statistics (values in parentheses are for outer resolution shell)

|                         | FadRSa | FadRSa-DNA | FadRSa:lauroyl-CoA |
|-------------------------|--------|------------|-------------------|
| **Data collection**     |        |            |                   |
| Space group             | P2₁    | P2₁       | P2₁               |
| a, b, c (Å)             | 42.0, 98.8, 54.8, 178.3, 46.0, 91.8, 53.5 |
| α, β, γ (°)             | 90.0, 106.4, 90.0, 90.0, 90.0, 112.1, 90.0 |
| Wavelength (Å)          | 0.97895 | 0.97625 | 0.97625 |
| Resolution (Å)          | 47.2-2.4 | 49.4-3.29 | 45.9-1.90 |
| Rmerge (%)              | 0.046 | 0.095 | 0.065 (0.727) |
| Redundancy              | 10.34 (1.67) | 12.59 (1.96) | 10.03 (1.07) |
| Completeness (%)        | 95.0 (72.9) | 99.5 (95.7) | 94.4 (66.0) |
| Refinement              | 2.2 | 6.8 | 3.3 |
| Resolution (Å)          | 47.2-2.4 | 49.4-3.29 | 45.9-1.90 |
| No. of reflections      | 72,292 | 276,981 | 100,877 |
| No. of atoms            | 3075 | 9345 | 4837 |
| Protein                 | 3075 | 9345 | 4837 |
| Ligand/ion              | 56 | 1783 | 122 |
| Water                   | 9 | 0 | 67 |
| Wilson B-factors (Å²)   | 53.5 | 97.4 | 34.0 |
| R.M.S. deviation (Å)    | 0.009 | 0.008 | 0.002 |
| Bond lengths (Å)        | 0.99 | 0.91 | 1.031 |

Mechanisms of DNA binding. To further unravel mechanisms of DNA binding by FadRSa, we determined the cocystal structure of the protein–DNA complex to a resolution of 3.29 Å using a duplex DNA containing the predicted FadRSa binding motif in the control region of the fadRSa gene itself, corresponding to ChIP-seq peak 2 (Table 1 and Fig. 3a, b). The asymmetric unit contained six FadRSa subunits, organized as three dimers, and two DNA duplex molecules thus representing a nonbiological assembly (Fig. 3a), although the protein–DNA molecular interactions within this structure are representative of the biologically relevant complexes (see below, DNA-binding stoichiometry of FadRSa).

In each FadRSa subunit in the cocystal structure, residues of the recognition helix α3 and the α2–α3 loop interact with the major groove of DNA with the establishment of an extensive number of contacts (Fig. 3b, Supplementary Note 1 and Supplementary Data 1). Base-specific contacts mainly consist of hydrophobic interactions between FadRSa residues Tyr47, Leu49, Tyr51, and Phe52 and methyl groups of thymines (Fig. 3c) similar as in other TetR-like regulators (Supplementary Figure 5), in addition to electrostatic interactions between Gly48 and the N7 group of guanines. The role of these residues for DNA binding was further investigated by performing site-directed alanine substitution and analyzing the mutant proteins in EMSA (Supplementary Figure 6a). FadRSaY47A, FadRSaY51A, and FadRSaY53A are all negatively affected in DNA-binding affinity and cooperativity. With FadRSaG48A, no DNA binding was observed at all demonstrating that Gly48 is a crucial residue (Supplementary Figure 6a). Besides protein–DNA contacts, a weak electrostatic protein–protein contact was also observed between Asn37 residues of FadRSa dimers bound on different sides of the DNA helix (Fig. 3d and Supplementary Note 1).

DNA-binding stoichiometry of FadRSa. To dissect the stoichiometric nature of the electrophoretically distinct FadRSa–DNA complexes, SEC was performed with the different molecular species (Fig. 4a). With a homogenous population of dimers in solution (Fig. 4a and Supplementary Figure 1) and the FadRSa–Saci_1123 complex B having an apparent molecular weight (MW) measured in SEC of 140 kDa that is only minimally exceeding that of free DNA (119 kDa), it can be concluded that FadRSa binds the Saci_1123 operator as a single dimer. The observation that the relative mobility in EMSA of FadRSa–Saci_1123 complex B is highly similar to that of the FadRSa–fadRSa complex B1 (Fig. 4b), led us to postulate that the transitional FadRSa–fadRSa complex B1 has a stoichiometry similar as for the sole FadRSa–Saci_1123 complex. In contrast, the dominantly formed complex B2 with the fadRSa operator has a larger apparent MW (179 kDa) (Fig. 4a): it can be assumed that the apparent MW attributed by the FadRSa protein itself is similar for measurements of free protein and of FadRSa–DNA complexes and that the FadRSa–fadRSa complex B2 has a stoichiometry that is twice as large as that of complex B1, thus harboring two dimers. SEC experiments with lower protein:DNA molar ratios indicate that the entire amount of FadRSa in the preparation is capable of binding DNA (Supplementary Figure 7). This excludes the possibility that a subpopulation of the protein is in a ligand-induced state lacking DNA-binding activity as suggested by the observation of acyl-CoA cocrystallizing with the protein in the apo crystal structure (Fig. 1b), assuming that acyl-CoA binding causes DNA dissociation like in bacterial FadR regulators.

Next, footprinting experiments were performed for the FadRSa–DNA complexes B1 and B2 observed in EMSAs with fadRSa and quasi-identical Saci_1106 operator probes (representing ChIP-seq peaks 1 and 2, respectively) (Fig. 4c and Supplementary Figure 8). Chemical “in gel” Cu–phenantroline footprinting demonstrated that for the fadRSa operator probe,
**Fig. 2** FadRsa interacts with four genomic loci in the Saci_1103-Saci_1126 gene cluster. 

**a** Overview of the genomic binding profile of FadRsa as monitored by ChIP-seq (IP = immunoprecipitated sample). A zoomed image of this profile is shown for the genomic region encompassing the Saci_1103-Saci_1126 gene cluster (corresponding to genomic coordinates 903,000–932,000), with indication of the four clearly visible peaks (numbered 1–4). Below the profile, a schematic representation of the genomic organization of the Saci_1103-Saci_1126 gene cluster is shown with indication of the ChIP-seq peak summit locations and of the transcription start sites as detected in the transcriptomic analysis in ref. 37. 

**b** ChIP-qPCR experiment with targeted quantification of enrichment for peaks 1 and 2 (given their close proximity to each other, these are assayed within a single fragment representing the Saci_1106/Saci_1107 intergenic region), peak 3 and peak 4. Fold enrichment is expressed relative to a genomic region within the ORF of Saci_1336 that was shown not be bound by FadRsa in the ChIP-seq profile. Error bars represent standard deviations of biological duplicates. 

**c** Sequence logo of the FadRsa binding motif representing MEME predictions of ChIP-seq enriched sequences. 

**d** Electrophoretic mobility shift assays of FadRsa binding to radiolabeled DNA probes of about 500 bp representing the ChIP-seq peaks identified in the Saci_1103-Saci_1126 gene cluster (see panel (a)). Molar protein concentrations are shown above each autoradiograph, whereas populations of free DNA (F) and FadRsa-bound DNA (B1 and B2) are indicated with an arrowhead. Apparent K_D and Hill coefficient (n) calculations are based on densitometric analysis of free DNA bands followed by binding curve analysis (Supplementary Figure 4).
both complexes are characterized by a similar protection zone roughly restricted to the predicted binding motif with a small stretch of additional protection extending upstream of the motif in complex B2 (Fig. 4c). This upstream extension was also observed in DNase I footprinting experiments for both operator probes (Fig. 4c and Supplementary Figure 8). In contrast, the protection zones observed in footprinting experiments with the *SacI*\_1123 operator probe (representing ChIP-seq peak 4) are smaller and only correspond to the binding motif (Supplementary Figure 9). These results support the notion that in the FadRS\textsubscript{a}–DNA cocrystal structure with the biologically relevant stoichiometries, we mapped the contacts on the DNA cocrystal structure, summarizing Supplementary Data 1.

To relate the molecular interactions and binding architecture in the FadRS\textsubscript{a}–DNA cocrystal structure with the biologically relevant stoichiometries, we mapped the contacts on the DNA cocrystal structure, summarizing Supplementary Data 1.

---

**Fig. 3** A FadRS\textsubscript{a}–DNA cocrystal structure reveals an important role for hydrophobic interactions. **a** Cartoon representation of the cocrystal structure of FadRS\textsubscript{a} in complex with the operator site in the *fadR\textsubscript{a}* promoter region. Different FadRS\textsubscript{a} subunits are labeled A–E, each dimer is colored differently with subunits A and B belonging to the same dimer while being displayed in a darker and lighter green, respectively, for the sake of clarity. **b** Detailed map of interactions identified in the FadRS\textsubscript{a}–DNA cocrystal structure, summarizing Supplementary Data 1. **c** Zoom of the interface between subunit E and the X–Y DNA duplex, with emphasis on interactions established by residues Tyr47, Tyr51, and Phe52. Hydrogen bonds are indicated by black and hydrophobic interactions by gray dashed lines. Bases are labeled with those belonging to chain Y being displayed with a prime. **d** Zoom of the interaction between subunits E (in orange) and B (in green). Weak electrostatic interactions are indicated with dashed lines.
cocrystal structure (Fig. 3a). This reasoning is underscored by the observation that the introduction of a G–C and C–G bp at the indicated positions of the Saci\textsubscript{1123} operator causes the formation of two instead of one nucleoprotein complex (Fig. 5c). Given that adenines also have an N7 group, the base specificity of the Gly48–guanine interaction is possibly explained by an indirect readout of the preceding thymidine or cytosine residue in the light of YpG base pair steps being more prone to unstacking and commonly involved in sequence-specific DNA interactions in a combined direct and indirect readout\textsuperscript{32}. In addition, the mutation of Asn37 causes a diminished cooperativity in the formation of the dimer-of-dimer complex B2 with the mutated Saci\textsubscript{1123} operator (Fig. 5c), proving the involvement of this residue in a protein–protein interaction. This supports the notion that established FadRS\textsubscript{a}–DNA contacts are similar in the complex in solution as in the portion of the cocrystal structure harboring dimer AB and subunit E of dimer EF and that the relative positioning of the two dimers in the biologically relevant complex is similar to that of dimers AB and EF in the cocrystal structure.

In conclusion, while bacterial TetR proteins are subdivided in two classes depending on whether they employ a dimer or dimer-of-dimer DNA-binding mode\textsuperscript{26,30}, the archaeal FadRS\textsubscript{a} regulator is capable of using both interaction modes depending on the operator sequence. A Gly48–guanine interaction and Asn37-mediated protein–protein contacts, of which the latter have never before been observed for bacterial dimer-of-dimer binding TetR-like proteins, assist in the dimer-of-dimer interaction mode.

### Determination of the FadRS\textsubscript{a} regulon

To infer whether or not the observed genomic binding of FadRS\textsubscript{a} leads to transcriptional regulation, a comparative transcriptomic analysis was performed for the fadRS\textsubscript{a} deletion mutant versus the isogenic WT strain using an RNA-seq approach (Fig. 6a, Supplementary Note 2, 20).
Supplementary Figure 10, Supplementary Table 4 and Supplementary Data 2). The deletion of \textit{fadRSa} did not affect cell morphology and growth in a medium containing sucrose and NZamine as carbon and energy sources (Supplementary Figure 11). RNA-seq analysis revealed that thirteen genes are differentially expressed, which all belong to the \textit{Saci}_{1103}-\textit{Saci}_{1126} gene cluster. Moreover, as confirmed by quantitative reverse transcriptase PCR (qRT-PCR), all other genes of the gene cluster appear to be expressed at slightly higher levels in the mutant strain as well (Fig. 6a–b and Supplementary Data 2). We therefore conclude that Fad\textit{Rs}a is a local repressor of the entire \textit{Saci}_{1103}-\textit{Saci}_{1126} gene cluster, which is predicted to harbor genes for a complete β-oxidation pathway (Fig. 6a).

All but one of these Fad\textit{Rs}a binding sites are located at too large distances (>130 bp) from their corresponding promoters to influence transcription in a direct manner (Supplementary Figure 12), pointing to a direct repression mechanism occurring at later stages of transcription initiation than during TBP and TFB recruitment.

Besides the local regulon, several other genes, including an operon encoding a putative sulfate reduction pathway and cytochrome-encoding genes, were found to have slightly lower expression levels in the \textit{ΔfadRSa} strain pointing to an indirect activation effect (Fig. 6b, Supplementary Note 2 and Supplementary Data 2). The suggestion of this effect being indirect is corroborated by the prediction of only a limited number of putative Fad\textit{Rs}a binding sites in the genomic regions surrounding the downregulated genes, which are characterized by relative high \textit{p} values (>1.00E−05) (Supplementary Table 3) and which were not captured by ChIP-seq. Furthermore, these indirect regulatory effects hint at a reversely correlated link between the fatty acid metabolism catalyzed by the enzymes encoded in the \textit{Saci}_{1103}-\textit{Saci}_{1126} gene cluster on one hand and sulfur metabolism and cytochrome-containing membrane complexes on the other hand.

The observed transcriptional regulation of the \textit{Saci}_{1103}-\textit{Saci}_{1126} gene cluster strongly suggests that Fad\textit{Rs}a is implicated in the regulation of fatty acid and lipid metabolism. Since it was previously observed that simultaneous deletion of both esterase-encoding genes in the gene cluster (\textit{Saci}_{1105} and \textit{Saci}_{1116}) led to a phenotype lacking the ability to perform tributyrin hydrolysis, we performed a similar phenotypic assay with the \textit{fadRSa} deletion mutant (Supplementary Figure 13a). Despite the higher expression levels of both esterase genes in the \textit{fadRSa} deletion mutant (Fig. 6a–b), we did not observe a difference in time-dependent halo formation upon growth on tributyrin (Supplementary Figure 13a). In contrast, upon growing \textit{S. acidocaldarius} in a liquid medium containing hexanoate as a sole carbon and energy source, the \textit{fadRSa} deletion mutant displayed a significantly higher growth rate in exponential growth phase with respect to the isogenic WT strain (doubling times \textit{T}_{\text{doub}} of 20.5 and 26.3 h, respectively; Fig. 6c). As this effect was not observed during growth on the shorter-chain butyrate (Supplementary Figure 13b), it correlates to fatty acid chain length. These experiments support that \textit{S. acidocaldarius} is capable of degrading fatty acids to sustain growth and that this catabolic metabolism may be at least partly catalyzed by enzymes encoded in the...
Saci_1109-Saci_1126 gene cluster. Furthermore, the FadRsa regulator represses this catabolic fatty acid metabolism as its deletion, thereby causing a derepression of the gene cluster, results in a faster growth rate (Fig. 6c). The observation of this difference can be explained by hexanoate only causing a partial FadRsa-mediated derepression given the short chain length of these acyl-CoA molecules (see below, “FadRsa-ligand interactions”).

**FadRsa-ligand interactions.** Next, we prepared FadRsa crystals in the presence of lauroyl-CoA (C12:0-CoA) and solved the FadRsa-laurol-CoA cocrystal structure to a resolution of 1.90 Å (Table 1). In contrast to the initial structure showing an acyl-CoA derivative bound to only one of the subunits, in this structure, both subunits of the dimer harbor a lauroyl-CoA molecule (Supplementary Figure 14a). The orientation of the ligand-binding pockets within the protein and the binding mode of the ligand is completely different in FadRsa as compared to the characterized bacterial FadR proteins24–26 (Fig. 7a). In contrast to the ligand entering the pocket from within the dimer interface as in bacterial FadR regulators, in the FadRsa structure the ligand enters the protein from the outer surface of the protein completely opposite to the dimer interface. Further, for each ligand-binding pocket only a single FadRsa subunit is involved in ligand interaction in contrast to two subunits in the bacterial FadR regulators. Consequently, ligand conformation is different and the acyl chain has a rather straight conformation in FadRsa while it is bent in FadR (Supplementary Figure 14b).

Upon zooming into the ligand-binding pocket, a large number of specific lauroyl-CoA–FadRsa interactions are observed (Fig. 7b and Supplementary Data 3). Whereas the adenine moiety is located on the outside of α4 and appears not to be contacted by the protein, the remainder of the lauroyl-CoA molecule enters the protein in between α4 and α7 with the establishment of electrostatic interactions between polar residues (a.o. Arg73, Lys77, Arg86, and Arg132) and the CoA moiety, especially with the CoA phosphate groups (Fig. 7b). Upon comparison of residue conformations in the DNA-bound and lauroyl-CoA structures, it became apparent that the orientation of the α5 residue Met101 is significantly altered (Supplementary Figure 15a). The lauroyl chain is deeply buried into a tunnel-like binding pocket formed in between helices α4–α7 and is entirely surrounded by hydrophobic residues such as Phe68, Phe97, Phe111, and Phe126. Although the nature of these ligand-interaction residues (polar residues for CoA-interactions, hydrophobic residues for side chain interactions) is similar as in bacterial TetR-like FadR regulators, they are not homologous as shown on a structure-based sequence alignment (Fig. 1d). Furthermore, FadRsa does not contain a hydrophilic patch in the acyl-binding pocket similarly to the Bacillus counterparts in which it affects ligand-binding specificity25,26.

EMSA demonstrated that acyl-CoA molecules, but not acetyl-CoA, CoA and free fatty acids, strongly abrogate FadRsa-DNA complex formation with the extent of the abrogation effect correlating with the length of the acyl chain (Fig. 7d and Supplementary Figure 16a). Competition assays confirmed that the addition of acetyl-CoA or hexanoyl-CoA does not affect sensitivity of the protein to oleoyl-CoA and thus that the inhibition effect reflects binding specificity (Supplementary Figure 16b). Alanine substitution of the CoA-interacting residues...
Acyl-CoA disrupts FadR<sub>Sa</sub>-DNA complexes proportional to acyl chain length. a Structural comparison of ligand-binding modes of lauroyl-CoA-bound FadR<sub>Sa</sub> (PDB: 6EL2), lauroyl-CoA-bound FadR<sub>Tt</sub> (PDB: 3ANG)<sup>24</sup>, and lauroyl-CoA-bound FadR<sub>Bs</sub> (PDB: 3WHB)<sup>25</sup>. Only the C-terminal domains are shown in cartoon representation, with α-helices α<sub>4</sub>-α<sub>9</sub> colored from light to dark gray; the lauroyl-CoA molecule is shown as a stick model. Two different viewpoints are shown: a view from within the monomer-monomer interface in the dimer and a view perpendicular to that. b Close-up view of the ligand-binding pocket with indication of FadR<sub>Sa</sub> residues involved in establishing polar contacts with lauroyl-CoA (in blue) and of residues involved in hydrophobic interactions (in red). c Schematic representation of a close-up view of helices α<sub>1</sub>-α<sub>4</sub> with indication of relative conformational differences. A detailed structural superimposition is shown in Supplementary Figure 15b. d Electrophoretic mobility shift assays (EMSAs) demonstrating the effect of acyl-CoA on the FadR<sub>Sa</sub>-DNA interaction using a 154-bp probe representing the SacI<sub>T106</sub> binding site. Acyl-CoA concentrations are shown above each autoradiograph in μM. Populations of free DNA and complexed DNA are indicated with F and B<sub>2</sub>, respectively. The intermediate complex B<sub>1</sub> is hardly formed due to the cooperativity of the interaction. e Graphical representation of ligand response measured in EMSAs performed with ligand-binding mutants FadR<sub>Sa</sub>R<sup>73A</sup>, FadR<sub>Sa</sub>R<sup>86A</sup>, and FadR<sub>Sa</sub>M<sup>101A</sup> (Supplementary Figure 6b). The Y-axis represents the fraction of bound DNA with respect to a protein-free control lane.
Arg73 and Arg86 confirmed their importance for the ligand response (Fig. 7e and Supplementary Figure 6b). Likewise, alanine mutation of the allosterically altered Met101 residue desensitizes the protein to oleoyl-CoA. In conclusion, acyl-CoA binds the regulator thereby causing dissociation of FadRsa–DNA complexes with the affinity and extent of the effect correlating to the acyl chain length.

Molecular mechanism of ligand response. To learn more about the allosteric regulatory mechanism employed by FadRsa, we compared the ligand-bound and DNA-bound structures (Fig. 7c and Supplementary Figure 15b). Both structures were superimposed with an RMSD of 1.01 Å (Supplementary Figure 15b). Subtle differences were noted in the relative orientation of the HTH motifs within a dimer with the binding of lauroyl–CoA causing the distance between the two α3 helices to be enlarged from an average 37.0–43.2 Å (measured as the Cα–Cα distance of Tyr51 residues located in α3). As a consequence, the increased distance between the α3 helices makes the dimeric FadRsa conformation suboptimal for interaction in consecutive major groove segments. Besides the α3 helix, the α2 helix is shifted by an average distance of 7.9 Å and the α4 helix is displaced with an angle of 9° (Fig. 7c). Intriguingly, one of the three dimers in the FadRsa–DNA cocrystal structure appeared to have a ligand-bound conformation distinct from the other two dimers; this is reflected by the distance between the two α3 helices being 45.3 Å for this central dimer EF versus an average of 37.0 Å for the flanking dimers (Fig. 3a; Supplementary Note 3). In conclusion, ligand binding allosterically opens up the dimer thereby causing it to dissociate from the DNA, similarly as the mechanism observed for FadRth26.

Occurrence of FadR in archaea. FadRsa is not restricted to S. acidocaldarius but is also represented in all other Sulfolobus species, in other Crenarchaeota belonging to Sulfolobales (Acidianus spp.) and Thermoproteales (Cuniculiplasma divulgatum), and also in species belonging to Euryarchaeota (Thermoplasmales) and in the recently discovered Marsarchaeota33 (Fig. 8). These organisms have in common that they are all thermophiles, some with a (facultative) aerobic metabolism, others with an anaerobic metabolism. Conservation of residues that are involved in DNA or ligand binding indicate that in this family and archaea. Interestingly, the observation of an inverse correlation between the expression of the Saci_1103-Saci_1126 gene cluster and that of cytochrome-encoding genes further supports the suggested function of fatty acids stabilizing membrane complexes in archaea39. FadRsa uses two distinct DNA-binding modes operator-dependently with a Gly–guanine interaction being the crucial determinant and that dimer–dimer complex formation with the quasi-identical high-affinity fadRsa and Saci_1106 operators occurs in a cooperative manner with the transitional formation of a dimer-bound complex in which FadRsa–mediated DNA bending is more pronounced than in the dimer–dimer complex. By binding to a total of only four binding sites in the Saci_1103-Saci_1126 gene cluster, FadRsa is capable of repressing transcription of the entire 23-gene cluster containing at least 17 transcription units37. Furthermore, with the exception of the autoregulatory binding site that is located just downstream of the TTS and that is expected to result in repression through direct interaction with the basal transcription initiation machinery, the other binding sites are located at least 130 bp upstream of the corresponding promoters, which is a noncanonical position as compared to most previously characterized archaeal repressors38.

FadRsa thus employs a nonparadigmatic repression mechanism that could be hypothesized to be dependent on long-range interactions. The observation of cocrystallization of FadRsa with two individual DNA duplexes (Fig. 3a), as was also observed for the bacterial TetR member Cgmr in Corynebacterium glutamicum39, further supports the possibility that the regulator is capable of co-associating with different DNA segments.

The finding that FadRsa represses the Saci_1103-Saci_1126 gene cluster and that it is responsive to acyl-CoA molecules acting as inducers in vitro strongly suggests that intracellularly present acyl-CoA molecules cause a derepression and thus higher transcriptional expression of the gene cluster in vivo. The observation that deletion of the regulator causes cells to display a faster growth on hexanoate as sole energy and carbon source supports that the β-oxidation enzymes encoded in this gene cluster minimally have a degradation function; this is in line with the logic behind the regulatory strategy. A catabolic function of the β-oxidation enzymes is also in agreement with the function of the co-regulated esterase enzymes encoded by Saci_1105 and Saci_1116, which enable cells to grow on lipids39. Fatty acid oxidation adds to the chemoorganotrophic capabilities of Sulfolobus spp. that appear more important than the originally described chemolithotrophic sulfur-oxidizing metabolism40. A full picture of the functioning of fatty acid metabolism in Sulfolobus, and whether the enzymes encoded in the Saci_1103–1126 function only the catabolic or also anabolic direction, awaits the biochemical and genetic characterization of the enzymes. An intriguing hypothesis has been put forward stating that fatty acid metabolism enzymes in archaea do not have a catalytic bias but are instead regulated by the relative substrate and product
possibly, acyl-CoA-responsive FadRα-mediated induction of the expression of the Saci_1103-Saci_1126 gene cluster is only a single element in a more complex regulatory system in which other regulatory mechanisms are also in place to enable the fine-tuned expression and activity of promiscuous enzymes in response to relative concentrations of a variety of fatty acid-related metabolic signals. This postulated regulatory complexity might reflect the employment of the same enzymes for catabolic and anabolic reactions instead of the use of distinct dedicated pathways as in bacteria and eukaryotes. Furthermore, based on the occurrence of FadRα orthologs in other thermoophilic archaea, such as Thermoplasmatales and Marsarchaeota, it
can be assumed that this type of transcriptional regulation is not restricted to Sulfolobus spp.

**Methods**

**Microbial strains and growth conditions**. *S. acidocaldarius* strains MW01 and its derivatives were cultivated at 75 °C in Brock basal salts medium supplemented with 0.2 (w/v)% sucrose, 0.2 (w/v)% NaZnime and 10 µg ml⁻¹ uracil. For ChIP experiments, a 200 ml strain *S. acidocaldarius* DSM639 culture was grown at 80 °C in Brock basal salts medium supplemented with 1% tryptone. The pH of the medium was adjusted to 3.5 by addition of sulfuric acid. For growth experiments in the presence of fatty acids (provitamin (xenomixid) kojic acid) or ligand (green triangles) interactions, based on FadRSα cocrystal structures. Sequence identities with FadRSα are mentioned behind the expression by the addition of 0.4 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) (isopropyl)-1-thiogalactopyranoside) or with 50 µg ml⁻¹ chloramphenicol (β). Cultivations of the third and fourth generation were considered to be representative for 1 µM fatty acids as sole carbon source. For growth on plates, Brock medium was solidified by the addition of 0.6 (w/v)% gelrite. 0.2 (w/v)% NZamine and 10 µg ml⁻¹ ampicillin (DH5α). An overview of all strains used in this work is given in Supplementary Table 6 and Supplementary Table 7.

**Protein expression and purification**. The fadRα coding region was amplified by PCR from *S. acidocaldarius* gDNA and cloned into a PET45b expression vector using BamHI and HindIII restriction sites (PET45b-fadRα) resulting in an N-terminally His-tagged construct. Site-directed mutagenesis was performed with the overlap PCR mutagenesis approach using PET45b-fadRα as a template and complementary mutagenic primers to remove an Ndel restriction site in the fadRα ORF with a silent mutation. This enabled cloning into the Ndel/Xhol sites of PET24α yielding a C-terminally His-tagged construct (PET24α-fadRαNdel) (Figure 8). For the construction of N37α, Y47α, G48α, V51α, V53α, R73α, R86α, and M101α variants of fadRα, directed mutagenesis was performed in an identical approach using PET24α-fadRαNdel as a template.

Heterologous overexpression of the recombinant FadRα proteins was accomplished in *E. coli Rosetta* (DE3) by growing a culture until reaching an OD600 of 0.6-0.7 at 37 °C before induction. For further growth during 4 h at 37 °C. Cultures were harvested by centrifugation and resuspended in lysis buffer (25 mM Tris-HCl pH 7.5, 300 mM NaCl) supplemented with protease inhibitor cocktail (Roche) and disrupted via French pressure cell (Thermo Electron Corporation, USA) for three passages at 12,000 psi followed by ultracentrifugation (30,000×g for 45 min). A heat treatment (80 °C during 15 min) was again followed by ultracentrifugation and TBP protein was further purified by anion exchange chromatography using a ResourceQ column (GE Healthcare) with a salt gradient up to 1 M NaCl and SEC on a HiLoad superdex 26/60 75 prep grade column (GE Healthcare) using 25 mM Tris-HCI pH 7.5, 300 mM NaCl. Similarly, *E. coli Rosetta* (DE3) cells expressing C-terminally hexahistidine-tagged TBP were grown in medium supplemented with 10 mM MgCl₂ and 100 µM ZnSO₄ and induced at an OD₆₀₀ of approximately 0.6 by adding 0.2 mM IPTG and followed by further growth at 25 °C overnight. Cells were harvested by centrifugation and subsequently resuspended in modified N-buffer (25 mM Tris-HCl pH 7.5, 10 mM MgCl₂, 100 µM ZnSO₄, 1 mM tris(2-carboxyethyl)phosphine (TCEP) supplemented with 1 M NaCl and protease inhibitor. After cell disruption via French pressure cell (Thermo Electron Corporation, USA) for three passages at 12,000 psi followed by ultracentrifugation (30,000×g for 45 min), protein was purified using a ResourceQ column (GE Healthcare) with a salt gradient up to 1 M NaCl and SEC on a HiLoad superdex 26/60 75 prep grade column (GE Healthcare). In the purification process, FadRα protein preparations were essentially pure, as judged by SDS-polyacrylamide gel electrophoresis (PAGE) and SEC (Supplementary Figure 1).

**Cryocrytography and data collection**. The crystals of SeMet-substituted FadRα protein were considered to be representative for 1 µM fatty acids as sole carbon source. The crystals were soaked in a cryo-solution containing 20% glycerol, 10% (w/v) sucrose, 0.1 M sodium nitrate and 0.05 M Bis-Tris (pH 5.8). The crystals were soaked in a cryo-solution containing 20% glycerol, 10% (w/v) sucrose, 0.1 M sodium nitrate and 0.05 M Bis-Tris (pH 5.8) for 15 min, followed by centrifugation and resuspended in lysis buffer (25 mM Tris-HCI pH 7.5, 300 mM NaCl) supplemented with protease inhibitor cocktail (Roche) and disrupted via French pressure cell (Thermo Electron Corporation, USA) for three passages at 12,000 psi followed by ultracentrifugation (30,000×g for 45 min). A heat treatment (80 °C during 15 min) was again followed by ultracentrifugation and TBP protein was further purified by anion exchange chromatography using a ResourceQ column (GE Healthcare) with a salt gradient up to 1 M NaCl and SEC on a HiLoad superdex 26/60 75 prep grade column (GE Healthcare) using 25 mM Tris-HCI pH 7.5, 300 mM NaCl. Similarly, *E. coli Rosetta* (DE3) cells expressing C-terminally hexahistidine-tagged TBP were grown in medium supplemented with 10 mM MgCl₂ and 100 µM ZnSO₄ and induced at an OD₆₀₀ of approximately 0.6 by adding 0.2 mM IPTG and followed by further growth at 25 °C overnight. Cells were harvested by centrifugation and subsequently resuspended in modified N-buffer (25 mM Tris-HCl pH 7.5, 10 mM MgCl₂, 100 µM ZnSO₄, 1 mM tris(2-carboxyethyl)phosphine (TCEP) supplemented with 1 M NaCl and protease inhibitor. After cell disruption via French pressure cell (Thermo Electron Corporation, USA) for three passages at 12,000 psi followed by ultracentrifugation (30,000×g for 45 min), protein was purified using a ResourceQ column (GE Healthcare) with a salt gradient up to 1 M NaCl and SEC on a HiLoad superdex 26/60 75 prep grade column (GE Healthcare). In the purification process, FadRα protein preparations were essentially pure, as judged by SDS-polyacrylamide gel electrophoresis (PAGE) and SEC (Supplementary Figure 1).

**Cryocrytography and data collection**. The crystals of SeMet-substituted FadRα protein were considered to be representative for 1 µM fatty acids as sole carbon source. The crystals were soaked in a cryo-solution containing 20% glycerol, 10% (w/v) sucrose, 0.1 M sodium nitrate and 0.05 M Bis-Tris (pH 5.8) for 15 min, followed by centrifugation and resuspended in lysis buffer (25 mM Tris-HCI pH 7.5, 300 mM NaCl) supplemented with protease inhibitor cocktail (Roche) and disrupted via French pressure cell (Thermo Electron Corporation, USA) for three passages at 12,000 psi followed by ultracentrifugation (30,000×g for 45 min). A heat treatment (80 °C during 15 min) was again followed by ultracentrifugation and TBP protein was further purified by anion exchange chromatography using a ResourceQ column (GE Healthcare) with a salt gradient up to 1 M NaCl and SEC on a HiLoad superdex 26/60 75 prep grade column (GE Healthcare). In the purification process, FadRα protein preparations were essentially pure, as judged by SDS-polyacrylamide gel electrophoresis (PAGE) and SEC (Supplementary Figure 1).

**Cryocrytography and data collection**. The crystals of SeMet-substituted FadRα protein were considered to be representative for 1 µM fatty acids as sole carbon source. The crystals were soaked in a cryo-solution containing 20% glycerol, 10% (w/v) sucrose, 0.1 M sodium nitrate and 0.05 M Bis-Tris (pH 5.8) for 15 min, followed by centrifugation and resuspended in lysis buffer (25 mM Tris-HCI pH 7.5, 300 mM NaCl) supplemented with protease inhibitor cocktail (Roche) and disrupted via French pressure cell (Thermo Electron Corporation, USA) for three passages at 12,000 psi followed by ultracentrifugation (30,000×g for 45 min). A heat treatment (80 °C during 15 min) was again followed by ultracentrifugation and TBP protein was further purified by anion exchange chromatography using a ResourceQ column (GE Healthcare) with a salt gradient up to 1 M NaCl and SEC on a HiLoad superdex 26/60 75 prep grade column (GE Healthcare). In the purification process, FadRα protein preparations were essentially pure, as judged by SDS-polyacrylamide gel electrophoresis (PAGE) and SEC (Supplementary Figure 1).

**Cryocrytography and data collection**. The crystals of SeMet-substituted FadRα protein were considered to be representative for 1 µM fatty acids as sole carbon source. The crystals were soaked in a cryo-solution containing 20% glycerol, 10% (w/v) sucrose, 0.1 M sodium nitrate and 0.05 M Bis-Tris (pH 5.8) for 15 min, followed by centrifugation and resuspended in lysis buffer (25 mM Tris-HCI pH 7.5, 300 mM NaCl) supplemented with protease inhibitor cocktail (Roche) and disrupted via French pressure cell (Thermo Electron Corporation, USA) for three passages at 12,000 psi followed by ultracentrifugation (30,000×g for 45 min). A heat treatment (80 °C during 15 min) was again followed by ultracentrifugation and TBP protein was further purified by anion exchange chromatography using a ResourceQ column (GE Healthcare) with a salt gradient up to 1 M NaCl and SEC on a HiLoad superdex 26/60 75 prep grade column (GE Healthcare). In the purification process, FadRα protein preparations were essentially pure, as judged by SDS-polyacrylamide gel electrophoresis (PAGE) and SEC (Supplementary Figure 1).
5% of the reflections was set aside and used to calculate the quality factor \( R_{	ext{free}} \).

The structure was solved using AutoSol in Phenix. Refinement was performed with Phenix, and manual rebuilding in Coot. The structure was refined to \( R_{	ext{work}} = 20.3\% \) and \( R_{	ext{free}} = 23.5\% \), respectively (Table 1). The structure of the dsDNA and lauryl–CoA complexes were solved by molecular replacement with Phaser,

using the SeMet-substituted FadRa structure as a model. All structures were evaluated using wwpdb Validation Server. Refinement statistics are presented in Table 1. The coordinates and structure factors have been deposited in the PDB database with accession codes 5MWV, 6EN8, and 6EL2. PSI-BOM was used to identify protein–DNA and protein–ligand interactions, supplemented with a manual inspection employing PyMOL.

All figures displaying protein structures were prepared with PyMOL.

**Chromatin immunoprecipitation.** ChiP was performed by growing S. acidocaldarius DSM39 to early exponential growth phase (an OD600 of 0.2 and 0.3) and adding formaldehyde to the culture to a final concentration of 1% [30]. After a 5-min incubation, glycine was added to a final concentration of 125 mM. Fixed cells were harvested by centrifugation at 8,000 × g for 10 min and the pellet was resuspended in 3 mL IP buffer (50 mM Heps-KOH pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS, protease inhibitor cocktail (Roche Applied Science)). Subsequently, cells were sonicated on ice until DNA fragments were obtained with an average size around 250 bp. After centrifugation during 20 min at 17,000 × g, 100 µl of the sheared DNA-containing supernatant was kept apart to use as input control and the remaining sample was divided into two aliquots. The first aliquot, used to prepare crosslinked nucleoprotein complexes separately, were performed by performing EMSA, incubated with pre-immune serum coated Dynabeads, which served as a non-binding reference. Cultures were assayed in biological duplicate.

**DNA footprinting experiments.** Enzymatic digestion for DNA footprinting was performed with ribonuclease A (13.7 kDa), carbonic anhydrase (29 kDa), conalbumin (75 kDa), and ferritin (440 kDa), all from a Gel Filtration Calibration Kit (GE Healthcare Life Sciences).

**RNA-seq.** Total RNA was prepared from culture MW001 and MW001ΔfadRa cultures in early exponential growth phase at an OD600 between 0.2 and 0.3 using a miRNeasy Mini Kit (Qiagen). Libraries were prepared with a TruSeq-stranded mRNA Library Prep Kit (Illumina). cDNA sequencing was performed on a Hiseq2500 system (Illunina) at SciLifeLab, Stockholm, Sweden. Sequence reads were first trimmed to remove adapters and then shorter than 20 nt were discarded. Processed reads were then mapped to the S. acidocaldarius DSM39 genome (NC_007181.1) with TopHat 2.0.12. For each gene, the FPKM was calculated with Cufflinks 2.2.16. Finally, reads were counted by being the featureCounts function in the Subread package 1.5.0-8 and only genes having at least one count in all samples were used for differential gene expression analysis with DESeq27.

**Quantitative RT-PCR.** RNA was extracted at an OD260 between 0.2 and 0.3 by stabilization with RNA Protect Reagens (Qiagen) and by using an SV Total RNA isolation System (Promega). Residual DNA was removed by treatment with TURBO DNase (Ambion Life Technologies). cDNA was prepared from 1 μg RNA with an iScript Select cDNA Synthesis Kit (Bio-Rad). All qRT-PCR oligonucleotides (Supplementary Data 4) were designed with Primer3 Plus software. qRT-PCR analysis was performed on a Tephys 977 (1968).

**Reporting Summary.** Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

**Data availability** All crystal structures presented in this work have been deposited in the Protein Data Bank (PDB) and are available with accession codes 5MWV, [has been superseded with 6EL2], and 6EN8. For stoichiometry experiments of FadRa DNA complexes, SEC was performed on a Superdex 200 Increase 10/30 GL column with an AKTA FPLC system (GE Healthcare Life Sciences). A total of 4–40 nM FadRa-DNA complexes were loaded with 0.2–1 nM 45-bp DNA fragments, encompassing the fadRa promoter and SacI,Δ1123 operator, respectively, which were preincubated at 25 °C for 30 min to allow hybridization and were loaded onto the column with the same buffer as mobile phase buffer. Calibration for MW calculation was performed with ribonuclease A (13.7 kDa), carbonic anhydrase (29 kDa), conalbumin (75 kDa), and ferritin (440 kDa), all from a Gel Filtration Calibration Kit (GE Healthcare Life Sciences).
10. Dibrova, D. V., Galperin, M. Y. & Mulkidjanian, A. Y. Phylogenomic
24. Agari, Y., Agari, K., Sakamoto, K., Kuramitsu, S. & Shinkai, A. TetR-family
13. Lombard, J., López-García, P. & Moreira, D. An ACP-independent fatty acid
27. Leigh, J. A., Albers, S.-V., Atomi, H. & Allers, T. Model organisms for genetics
5. Tornabene, T. G. et al. Phytanyl-glycerol ethers and squalenes in the
15. Fujita, Y., Matsuoka, H. & Hirooka, K. Regulation of fatty acid metabolism in
3. Bell, S. D. & Jackson, S. P. Mechanism and regulation of transcription in
17. Henry, M. F. & Cronan, J. E. Escherichia coli transcription factor that both activates fatty acid synthesis and represses fatty acid degradation. J. Mol. Biol. 222, 843–849 (1991).
18. Zhang, Y.-M., Marrackhi, H. & Rock, C. O. The FabR (YicJ) transcription factor regulates unsaturated fatty acid biosynthesis in Escherichia coli. J. Biol. Chem. 277, 15558–15565 (2002).
19. Zhu, K., Zhang, Y.-M. & Rock, C. O. Transcriptional regulation of membrane lipid homeostasis in Escherichia coli. J. Biol. Chem. 284, 34880–34888 (2009).
20. Feng, Y. & Cronan, J. E. Complex binding of the FabR repressor of bacterial unsaturated fatty acid biosynthesis to its cognate promoters. Mol. Microbiol. 80, 195–218 (2011).
21. Matsuoka, H., Hirooka, K. & Fujita, Y. Organization and function of the YsiA regulon of Bacillus subtilis involved in fatty acid degradation. J. Biol. Chem. 282, 5180–5194 (2007).
22. Schuiman, G. E., Paletti, L., Grossman, A. D. & De Mendoza, D. FabR, a bacterial transcription factor involved in global regulation of membrane lipid biosynthesis. Dev. Cell 4, 663–672 (2003).
23. Badger, J. et al. Structural analysis of a set of proteins resulting from a bacterial genomics project. Proteins 60, 787–796 (2005).
24. Agari, Y., Agari, K., Sakamoto, K., Kuramitsu, S. & Shinkai, A. TetR-family transcriptional repressor Thermus thermophilus FabR controls fatty acid degradation. Microbiology 157, 1589–1601 (2001).
25. Fujihashi, M. et al. Structural characterization of a ligand-bound form of Bacillus subtilis FabR involved in the regulation of fatty acid degradation. Proteins 82, 1301–1310 (2014).
26. Yeo, H. K., Park, Y. W. & Lee, J. Y. Structural basis of operator sites recognition and effector binding in the TetR family transcription regulator FabR. Nucleic Acids Res. 45, 4244–4245 (2017).
27. Leigh, J. A., Albers, S.-V., Atomi, H. & Allers, T. Model organisms for genetics in the domain Archaea: methanogens, halophiles, Thermococcales and Sulfolobales. FEMS Microbiol. Rev. 35, 577–608 (2011).
28. Wagner, M. et al. Versatile genetic tool box for the crenarchaeote Sulfolobus acidocaldarius. Front. Microbiol. 3, 214 (2012).
29. Zweerin, S. et al. Activity-based protein profiling as a robust method for enzyme identification and screening in extremophilic Archaea. Nat. Commun. 8, 15352 (2017).
30. Ramos, J. L. et al. The TetR family of transcriptional repressors. Microbiol. Mol. Biol. Rev. 69, 326–356 (2005).
31. Schumacher, M. A. et al. Structural basis for cooperative DNA binding by two dimers of the multidrug-binding protein QscR. EMBO J. 21, 1210–1218 (2002).
32. Lamoureux, J., Maynes, J. T. & Glover, J. N. M. Recognition of 5′-YpG-3′ sequences by coupled stacking/hydrogen bonding interactions with amino acid residues. J. Mol. Biol. 335, 399–408 (2004).
33. Jay, Z. J. et al. Marsarchaeota are an aerobic archaeal lineage abundant in geothermal iron oxide microbial mats. Nat. Microbiol. 3, 732–740 (2018).
34. Kyripides, N. C. & Ouzounis, C. A. Transcription in archaea. Proc. Natl Acad. Sci. USA 96, 8545–8550 (1999).
35. Bell, S. D. & Jackson, S. P. Mechanism and regulation of transcription in archaea. Curr. Opin. Microbiol. 4, 208–213 (2001).
36. Nelson-Sathi, S. et al. Origins of major archaeal clades correspond to gene acquisitions from bacteria. Nature 517, 77–80 (2015).
37. Cohen, O. et al. Comparative transcriptomics across the prokaryotic tree of life. Nucleic Acids Res. 44, W46–W53 (2016).
38. Peeters, E., Peixeiro, N. & Sezonov, G. Cis-regulatory logic in archaeal transcription. Biochem. Soc. Trans. 41, 326–331 (2013).
39. Ito, H., Watanabe, N., Tao, M., Shirakihara, Y. & Tanaka, I. Crystal structures of the multidrug binding repressor Corynebacterium glutamicum CgmR in complex with inducers and with an operator. J. Mol. Biol. 403, 174–184 (2010).
40. Brock, T. D., Brock, K. M., Kelly, R. T. & Weiss, R. L. Sulfolobus: a new genus of sulfur-oxidizing bacteria thriving at low pH and high temperature. Arch. Microbiol. 84, 54–68 (1972).
41. Higuchi, R., Krummel, B. & Saiki, R. K. A general method for in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. Nucleic Acids Res. 16, 7351–7356 (1988).
42. Kabsch, W. XDS. Acta Crystallogr. D Biol. Crystallogr. 66, 125–130 (2010).
43. Brenner, A. T. Free R values: a novel statistical quantity for assessing the accuracy of crystal structures. Nature 355, 472–475 (1992).
44. Tervilliger, T. C. et al. Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. Acta Crystallogr. D Biol. Crystallogr. 64, 61–69 (2008).
45. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A Found. Crystallogr. 47, 110–119 (1991). (Pt 2).
46. McCoy, A. J. et al. Phaser crystallographic software. J. Appl. Crystallogr. 40, 658–674 (2007).
47. Berman, H., Henrick, K. & Nakamura, H. Announcing the worldwide Protein Data Bank. Nat. Struct. Biol. 10, 980–980 (2003).
48. Laskowski, R. A., Jarlskog, J., Pravda, L., Váleková, R. S. & Thornton, J. M. PDBeSum: Structural summaries of PDB entries. Protein Sci. 11, 535 (2017).
49. Delano, W. L. PyMOL. (2002).
50. Nguyen-Duc, T., Peeters, E., Muyldermans, S., Charlier, D. & Hassanzadeh-Ghasabeh, G. Nanobody(R)-based chromatin immuno precipitation/microarray analysis for genome-wide identification of transcription factor DNA-binding sites. Nucleic Acids Res. 41, e59–e59 (2013).
51. Liu, H., Wang, K., Lindás, A.-C. & Peeters, E. The genome-scale DNA binding profile of BarR, a β-alanine responsive transcription factor in the archaeon Sulfolobus acidocaldarius. BMC Genom. 17, 569 (2016).
52. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26, 589–595 (2010).
53. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 9, R137 (2008).
54. Robinson, J. T. et al. Integrative genomic viewer. Nat. Biotechnol. 29, 24–26 (2011).
55. Quinlan, A. R. & Hall, I. M. BEDTools: a powerful tool set for comparing genomic features. Bioinformatics 26, 841–842 (2010).
56. Bailey, T. L. & Elkan, C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc. Int. Conf. Intell. Syst. Mol. Biol. 2, 28–36 (1994).
57. Rozen, S. & Skalsky, H. Primer3 on the WWW for general users and for biologist programmers. Methods Mol. Biol. 132, 365–386 (2000).
58. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2−ΔΔCT method. Methods 25, 402–408 (2001).
59. Peeters, E., This-Toong, T.-L., Cigot, D., Maes, D. & Charlier, D. Ss-LrpB, a novel Lrp-like regulator of Sulfolobus solfataricus P2, binds cooperatively to three conserved targets in its own control region. Mol. Microbiol. 54, 321–336 (2004).
60. Galas, D. J. & Schmitz, A. DNase footprinting: a simple method for the detection of protein-DNA binding specificity. Nucleic Acids Res. 5, 3157–3170 (1978).
61. Maxam, A. M. & Gilbert, W. Sequencing end-labeled DNA with base-specific chemical cleavages. Meth. Enzymol. 65, 499–560 (1980).
Acknowledgements

We are grateful to Daniël Charlier for assistance with footprinting experiments and critical reading of the manuscript. We also would like to acknowledge support from Science for Life Laboratory, the National Genomics Infrastructure (NGI) and Uppmax in Stockholm/Uppsala, Sweden, for providing assistance in massive parallel sequencing and computational infrastructure. Research in the laboratory of E.P. was supported by the Research Council of the Vrije Universiteit Brussel (start-up funds) and by the Research Foundation Flanders (FWO-Vlaanderen) (research project G021118 and research grant 1526418N). D.S. is a recipient of an FWO-SB PhD fellowship of the Research Foundation Flanders (FWO-Vlaanderen). Research in the laboratory of A.C.L. was supported by the Swedish Research Council, Grant 621-2013-4685.

Author contributions

K.W. contributed by performing the ChIP-seq and RNA-seq experiments, protein purifications, growth experiments and the data analysis; D.S. contributed by performing the qRT-PCR, EMSA, and footprinting experiments and data analysis; H.R.M. and C.L. contributed by performing the mutant constructions, protein purifications, EMSA experiments and the data analysis; L.L. contributed by preparing the S. acidocaldarius mutant strain, performing ChIP-qPCR and the data analysis; X.Z. contributed by performing the growth experiments, RNA extractions, and data analysis; F.S. contributed by performing protein purifications; C.B. and B.S. contributed in the study design and data analysis; K.V. contributed by performing protein crystallography and data analysis; A.C. L. and E.P. contributed by conceiving and designing the study, performing data analysis, and writing the paper. All authors approved of the paper.

Additional information

Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-019-09479-1.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Journal peer review information: Nature Communications thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s) 2019