Fatty acid metabolism in *Sulfolobus acidocaldarius* and its potential as platform organism in biotechnology

Dissertation

zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

-Dr. rer. nat. –

vorgelegt von

Xiaoxiao Zhou

Arbeitskreis für Molekulare Enzymtechnologie und Biochemie

Abteilung für Umweltmikrobiologie und Biotechnologie

Fachbereich Chemie

der

Universität Duisburg-Essen

Mai 2021

Die vorliegende Arbeit wurde im Zeitraum von März 2016 bis Mai 2021 bei Prof. Dr. Bettina Siebers im Arbeitskreis für Molekulare Enzymtechnologie und Biochemie in der Abteilung für Umweltmikrobiologie und Biotechnologie an der Fakultät für Chemie der Universität Duisburg-Essen durchgeführt.

Tag der Disputation: 13.07.2021

Gutachter:	Prof. Dr. Bettina Siebers
	Prof. Dr. Peter Bayer
Vorsitzender:	Prof. Dr. Mathias Ropohl

DOI: 10.17185/duepublico/74886
URN: urn:nbn:de:hbz:465-20250108-075036-9

Content

1 Introduction

1.1 Archaea

The Archaea have initially been identified as one of three domains of life in addition to Bacteria and Eukaryotes based on 16S/18S rRNA sequences and in the three domains of life tree topology the Archaea and Eukaryotes have been regarded as sister groups [1] (Fig. 1). The Archaea share a mixture of bacterial and eukaryotic properties but possess also unique archaeal features. The prokaryotic cell organization and DNA structure (e.g. one circular chromosome, operon structures, plasmids) resemble Bacteria, whereas information processing mechanisms (e.g. replication, transcription, repair and translation) are more similar to Eukaryotes but less complex [2]. However, although Archaea are similar in metabolic complexity and genomic organization compared to Bacteria, they are characterized by the presence of unique metabolic pathways (e.g. methanogenesis) and by unusual, modified pathway versions of the “classical” routes [3]. Archaea produce distinct cell wall structures from other organisms. Unlike the bacterial cell walls composed of murein/petidoglycan, most Archaea possess surface layer proteins (S-layer) as their cell wall components, except for methanogenic species whose cell walls are made of pseudomurein employing *N*-acetyl-D-glucosamine (NAG) and *N*-acetyltalosaminuronic acid (NAT) as building blocks instead of NAG and *N*-acetyl-muramic acid (NAM) used in Bacteria [4]. Archaeal S-layers consist of mushroom-shape (glyco)proteins connected either with trans-membrane domains or lipid-modified subunits as membrane anchors. The S-layers account for up to 10% of the whole cell protein content and are considered to function as the protective coats, and play key roles in cell adhesion, surface recognition, antifouling and determination of the cell shape [5].

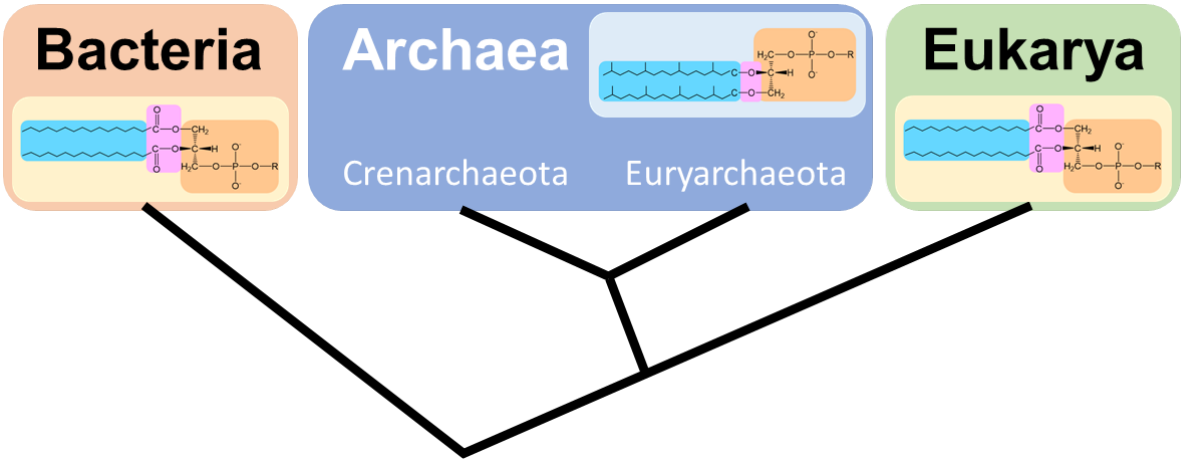


Figure 1. Phylogenetic two-domain tree of life based on 16S/18S rRNA sequences [1].

Yet a striking characteristic trait of archaea is their membrane lipid composition [8]. The membrane lipids of bacteria and eukaryotes are composed of fatty acids (FAs) ester-linked to G3P (glycerol-3-phosphate) forming membrane bilayers and are thus fundamentally different

from archaeal membrane lipids comprised of isoprenoid chains ether-linked to G1P (glycerol-1-phosphate) forming mono- or bilayer membranes (Fig. 2). Archaeal membranes exhibit high diversity in lipid compositions and contain two major groups, diether lipids (DELs) and tetraether lipids (TELs). The monopolar DELs consist of phytanyl hydrophobic chains (C20) and form bilayer structures, whereas in bipolar TELs two phytanyls fuse together to generate biphytanyl hydrophobic cores (C40) and monolayer architectures (Fig. 2C) [9]. According to distinct polar head groups, TELs are divided into glycerol dibiphytanyl glycerol tetraethers (GDGTs) and glycerol dibiphytanyl nonitol/calditol tetraethers (GDNTs/calditol-GDGTs). Both GDGTs and GDNTs connect to a phospho-myo-inositol at one end but link to a β-D-galactosyl-D-glucose or calditol D-glucose at the other end, respectively. Their hydrophobic chains usually include 0-8 cyclic pentane ring structures, sometimes even a cyclohexane ring (Fig. 2D) [9]. These bipolar TELs display high specificity and importance for Archaea and are believed to play extraordinary roles under extreme environments such as high temperatures, low pH, high salt concentrations, hydrostatic or oxidation pressure [10]. For instance, it has

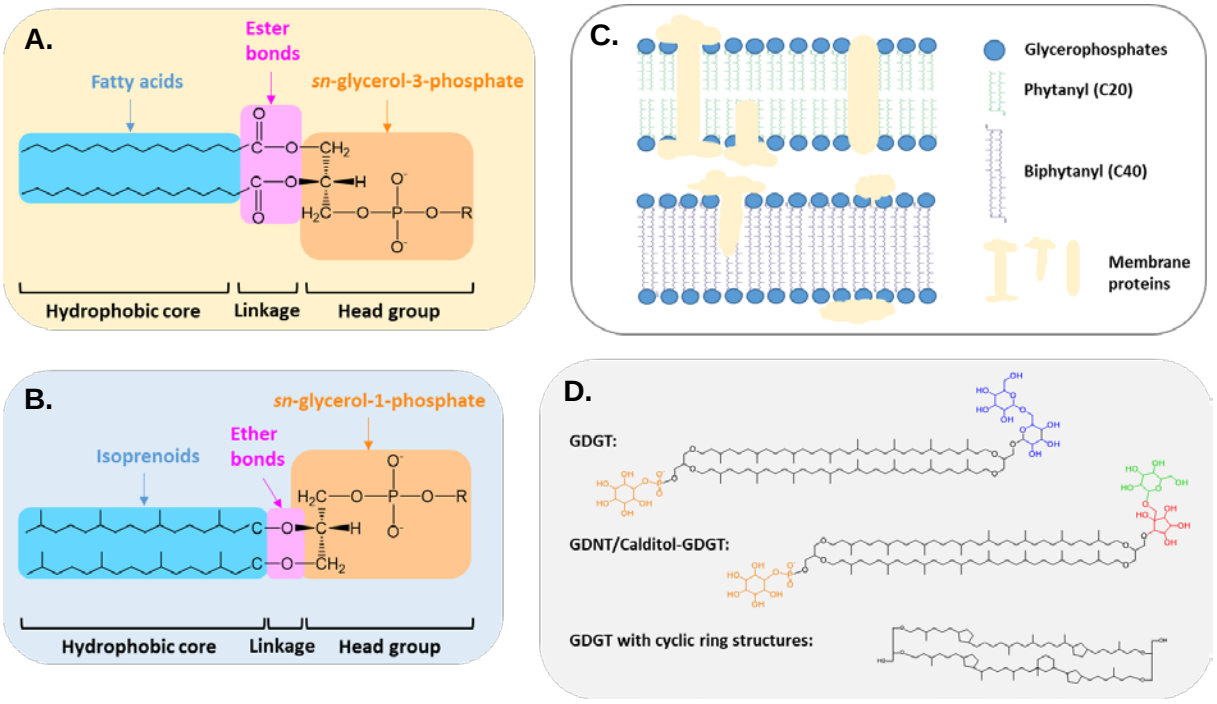


Figure 2. Lipid compositions in Bacteria, Eukaryotes (A) and Archaea (B, C & D). The cell membranes of Bacteria and Eukarya are composed of fatty acid based-lipids ester-bound to G3P (A) while Archaea produce isoprenoid-based membrane lipids ether-linked to G1P (B). The archaeal lipids include two types: the monopolar diether lipids (DELs) forming bilayer membranes and the bipolar tetraether lipids (TELs) building up monolayer structures (C). The TELs are rich in thermophilic Archaea and mainly comprise glycerol dialkyl glycerol tetraethers (GDGTs) and glycerol dialkyl nonitol tetraethers (GDNTs). They both contain 40 carbons in each poly-isoprenoid chain but differ in head groups. At one end of the hydrophobic cores, both GDGTs and GDNTs connect with a phospho-myo-inositol head group, yet at the other end, GDGTs link to β-D-galactosyl-D-glucose whereas GDNTs bind to calditol D-glucose moiety (D) [9].

been studied that in thermoacidophilic *Sulfolobus acidocaldarius* GDGTs account for 90% of its total lipid content and most of them contain four cyclic pentane rings [11, 12]. The ether linkage in archaeal lipids are chemically more stable than ester bonds while the methyl branches may help with lipid packing and compaction. The monolayer architectures as well as the cyclic rings can enhance membrane rigidity meanwhile reduce the permeability and fluidity, thereby making the membrane more stable [9, 11, 13].

In the beginning two main archaeal phyla have been identified, the Euryarchaeota and the Crenarchaeota [14] (Fig. 1), and the Archaea were regarded as mostly extremophilic or as metabolic specialists like methanogens. It has been argued that the different composition of their ether lipid membranes, which are generally regarded as more stable than ester lipids, might be due to this extremophilic life style [8]. However, in recent years metagenomics/environmental molecular biology approaches demonstrated that Archaea are ubiquitously distributed also in mesophilc habitats and that they are important players in geochemical cycles [15]. Moreover, numerous further archaeal phyla have now been deciphered which are grouped in four major clades, (i) the Euryarchaeota, (ii) the TACK (Thaum-, Aig-, Cren-, Korarchaeota) superphylum [16] (also comprising the Verstraete-, Bathy- and Geoarchaeota), (iii) the DPANN superphylum (Diapherotrites, Parv-, Aenigma-, Nano-, and Nanohaloarchaeota), and (iv) the Asgard archaea comprising the Loki-, Thor-, Odin-, and Heimdallarchaeota [17] (Fig. 3). Most of these exciting, newly identified organisms are not yet culturable. Phylogenomic analyses including these novel taxa have profoundly changed the tree of life. The data suggest that Bacteria and Archaea represent the only primary evolutionary lineages and that the Eukaryotes originated later on from within the Archaea namely from the Asgard Archaea [17] (Fig. 3). This means that during evolution of Eukaryotes from within the Archaea the membrane constitution must have fundamentally changed from archaeal to bacterial-type phospholipids. However, since one of the main and perhaps most important function of FAs as key constituent of cell membrane phospholipids and thus of cell structure in Bacteria and Eukaryotes is substituted by isoprenoids in Archaea the presence and the function of FAs in archaeal representatives is obscure.

Previous findings indicated that at least some archaea might be able to utilize FAs as carbon and energy source (e.g. haloarchaea and *Archaeoglobus fulgidus*) and homologues of all enzymes from the bacterial β oxidation have been identified in several archaeal species although detailed analyses have not been reported [18]. Also, the ability of several archaea to degrade FA based lipids by means of esterases and lipases *in vivo* was recently analysed in detail using an ABPP approach established for *Sulfolobus acidocaldarius*, *Saccharolobus solfataricus*, and also for *Haloferax volcanii* [19]. In addition, *Saccharolobus solfataricus* P1 was shown to degrade complex FA based lipids [20]. The results showed that these organisms cleave exogenous FA based triglycerides into FAs and

glycerol. The growth and the degradation pathway of glycerol has previously been reported for *H. volcanii* [21] and was analysed in some detail in *S. acidocaldarius* in course of the present thesis (data not shown).

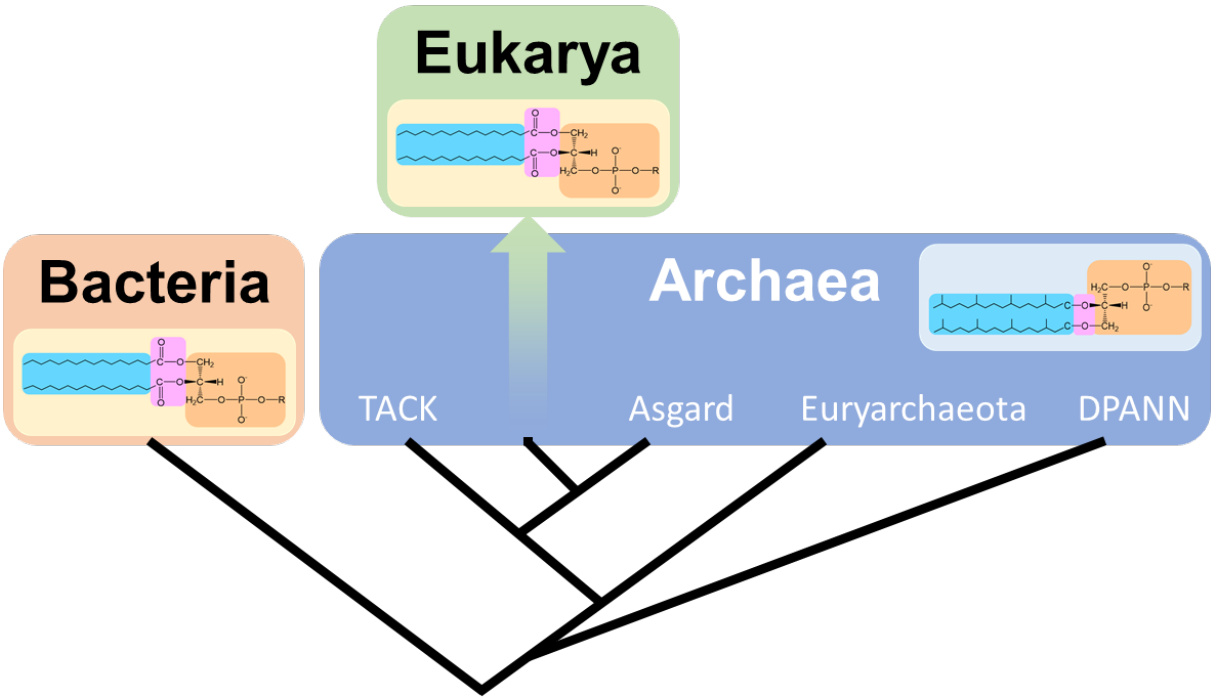


Figure 3. Phylogenetic three-domain tree of life based on 16S/18S rRNA sequences ([6, 7]).

1.2 *Sulfolobus acidocaldarius*

Sulfolobus acidocaldarius belongs to the crenarchaeal phylum [14] and has been isolated from acid hot springs in Yellowstone National Park. *S. acidocaldarius* possesses an obligate aerobic, thermoacidophilic life style with optimal growth around 80°C and pH 2-5 [22, 23]. The central carbon and energy metabolism of *S. acidocaldarius* is well understood [3, 23]. Furthermore, the genome sequences and a versatile genetic tool box for *S. acidocaldarius* is available [23]. The genetic system enables the construction of in frame markerless deletion mutants, ectopic integration of foreign DNA and provides a homologous expression system [24]. Also, systems biology approaches, genome scale metabolic models, kinetic models, have been developed and a wealth of polyomics and physiological information is available [25]. Together with the ease of cultivation under aerobic/microaerophilic conditions on many different substrates, this qualifies *S. acidocaldarius* as a model organism to study (cren)archaeal biology.

1.3 A transcriptional regulator acting on a cluster of FAs related genes in *S. acidocaldarius*

Interestingly, the genes encoding the characterized esterases (Saci_1105, Saci_1116) involved in triacylglycerol degradation as shown by the above mentioned ABPP approach, are

organized in a gene cluster (*saci_1103-saci_1126*) (Fig. 4C) together with several β oxidation homologues in *S. acidocaldarius* which thus may be responsible for the lipid degradation in this organism.

One gene within this gene cluster, i.e. *saci_1107*, in *S. acidocaldarius* encodes a transcriptional regulator characterized in detail in the study presented in chapter 3.1 of this thesis. It was shown that the regulator from the TetR family has its specific binding sites exclusively in this cluster and regulates its own expression as well as that of many of the β oxidation homologues. This was also confirmed by transcriptomic studies comparing the regulator deletion mutant with the isogenic wild type (WT, MW001 [24]) showing that only genes from this gene cluster

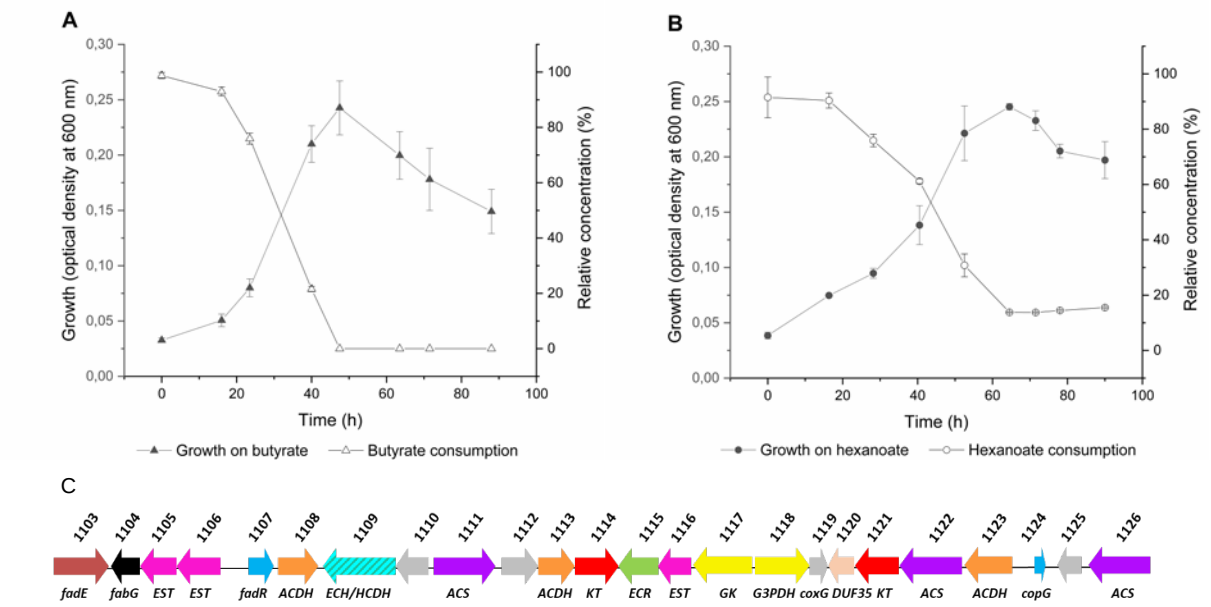


Figure 4: Growth and residual substrate concentration of *S. acidocaldarius* MW001 on short or medium chain fatty acids (A, B) as well as the genomic organization of the *saci_1103-saci_1126* gene cluster related to lipid and fatty acid metabolism (C). A. Growth curve (solid triangle) and butyrate consumption (hollow triangle) of the parental strain MW001 at pH 3, 76 °C in the minimal Brock medium supplemented with 2 mM butyrate. B. Growth curve (dot) and hexanoate consumption (circle) of the parental strain MW001 at pH 3, 76 °C in the minimal Brock medium supplemented with 2 mM hexanoate. The experiment was carried out in triplicate or quadruplicate and the error bars indicated the standard deviation. Within the fatty acid operon (C), several copies of lipases/esterases (*saci_1105, saci_1106* and *saci_1116*, shown in pink) are present, two of them have been characterized. Two genes are predicted to be involved in glycerol metabolism (*saci_1117* and *saci_1118*, displayed in yellow). The genes encoding AMP-forming ACSs, which can synthesize acyl-CoAs as precursors for β oxidation, were shown in purple (*saci_1111, saci_1122* and *saci_1126*). One or more paralogs for each steps of the bacterial-type β oxidation could be found as well, for instance, three acyl-CoA dehydrogenases (*saci_1108, saci_1113* and *saci_1123*, colored in orange), a bifunctional enoyl-CoA hydratase/3-hydroxyacyl-CoA dehydrogenase (*saci_1109*, in blue-green upward diagonal) and two β-ketothiolases/acety-CoA C-acetyltransferases (*saci_1114* and *saci_1121*, in red). Moreover, a fabG homolog (*saci_1104*, in black) and an enoyl-CoA reductase (*saci_1115*, in green) are thought to be responsible for elongation of fatty acyl chains. The genes with blue color encode transcription regulators (*saci_1107* and *saci_1124*).

were differentially regulated in both strains. This analysis also showed that the regulator is a repressor and that derepression occurs when FA-CoA esters bind to the protein. The crystal structure of the regulator was determined including the binding to the DNA as well as the binding mode of the acyl-CoA and it turned out to be different from that seen in other regulators of the TetR family.

To further support the anticipated function of the regulator and the gene cluster in Lipid/FA degradation, the contribution of this work to the publication in chapter 3.1 was to analyse the growth of *S. acidocaldarius* on FAs (Fig. 4A, B). It was clearly shown that *S. acidocaldarius* can use butyrate and hexanoate as sole carbon and energy source for growth (Fig. 4A, B) and likely utilize the β oxidation enzymes (Fig. 4C) for the degradation of FAs. In the exponential growth phase on hexanoate, the regulator disruption mutant displayed a shorter doubling time (t_d) of 20.5 h than that of WT (t_d is 26.3 h) supporting the idea of the regulator being a repressor of FA catabolism.

1.4 FA degradation via a largely bacterial like β oxidation in *S. acidocaldarius*

In the work presented in chapter 3.2 the FA degradation (Fig. 5A) in *S. acidocaldarius* was further analysed in the first part. As already mentioned above the gene cluster (Fig. 4C) identified in chapter 3.1 comprises several genes for FA degradation via β oxidation (Fig. 5A). To channel FAs into β oxidation they first need to be activated to the corresponding CoA esters usually carried out by AMP-forming acyl-CoA synthetases (Saci_1111, Saci_1122 and Saci_1126). Also, three homologues of acyl-CoA dehydrogenases have been identified (Saci_1108, Saci_1113 and Saci_1123). As a first step in the β oxidation, these enzymes introduce a double bond at the β position of the straight chain acyl-CoAs oxidizing them to the respective enoyl-CoA ester. During this process electrons will be transported to quinone and further to molecular oxygen in the respiratory chain via an electron transferring flavoprotein (ETF). Luckily, a candidate of this intermediate electron acceptor encoded by *saci_0315* was found from the *S. acidocaldarius* genome and characterized in chapter 3.2 of this work. The next step in β oxidation is the hydration of the enoyl-CoA to the corresponding 3(S)-hydroxyacyl-CoA catalysed by crotonase superfamily hydratases followed by the 3(S)-hydroxyacyl-CoA oxidation to 3-ketoacyl-CoA carried out by hydroxyacyl-CoA dehydrogenase superfamily enzymes, respectively. For both, only one gene (*saci_1109*) exists in the cluster and both are fused to a single gene encoding a fusion enzyme with an N-terminal dehydrogenase and a C-terminal crotonase domain. Finally, the 3-ketoacyl-CoA is thiolytically cleaved by ketothiolases (Saci_1114, Saci_1121) to yield acetyl-CoA and a saturated acyl-CoA ester shortened by two C atoms. However, the coding function of these genes have not been confirmed so far. The acyl-CoA synthetase (Saci_1122, ACS) as well as the β oxidation enzyme homologues encoded by *saci_1123, saci_1109* and *saci_1114* were recombinantly

produced in *E. coli*, purified and characterized in detail with special respect to kinetic constants and substrate/chain length specificity. Saci_1122 was characterized to be ATP and CoA-dependent and utilize a broad range of saturated, straight chain FAs with the length ranging from C3 to C10 but favour medium-chain substrates (C5-C8). The Saci_1123 protein displayed acyl-CoA dehydrogenase (ACAD) activity depending on FAD as cofactor in presence of the electron transferring flavoprotein (ETF) Saci_0315 or an artificial electron acceptor ferrocenium ($FcPF_6$)/2,6-dichlorophenolindophenol (DCPIP), and preferred saturated, straight medium-chain fatty acyl-CoAs (C4-C8) as substrates. Both enoyl-CoA hydratase and 3-hydroxyacyl-CoA dehydrogenase activities were observed with the bi-functional enzyme Saci_1109 (HCDH/ECH) *in vitro* using crotonoyl-CoA and 3-hydroxybutyryl-CoA as substrate, respectively, and NAD^+ as cofactor. Besides, enoyl derivatives other than crotonoyl-CoA, e.g. decenoyl-CoA (C10:1) or hexadecenoyl-CoA (C16:1) were tested and the enzyme showed no activity with C16:1 but high activity with C4:1 and C10:1 indicating a preference to medium-chain substrates. This enzyme is specific to 3(S)-hydroxybutyryl-CoA showing a high specific acitivity of 48 U mg^{-1} and no activity with the (R)-stereoisomer. Additionally, Saci_1109 was able to convert acetoacetyl-CoA into 3-hydroxybutyryl-CoA in presence of NAD(P)H as well meaning that the enzyme functions reversibly. This enzyme exhibited similar bi-functional activities as the bacterial FadB homologues, e.g. from *E. coli*, but the domain organisations are different. FadB harbors an N-terminal crotonase domain and a C-terminal dehydrogenase whereas Saci_1109 possesses an opposite architecture as mentioned above. The β-ketothiolase (KT) activity of the recombinant Saci_1114 was spectrophotometrically confirmed with the substrate acetoacetyl-CoA. In addition, activity was also obtained in the opposite direction with acetyl-CoA meaning that the enzyme in principle operates reversibly. It was reported in Bacteria, e.g. *E. coli*, that FadB and FadA, which are the respective homologues of Saci_1109 and Saci_1114 proteins, occur as a heterotetrameric complex containing two copies of each protein [26]. Therefore, we speculated that the ketothiolase in *S. acidocaldarius* might also form a complex with the Saci_1109 HCDH/ECH. Therefore, Saci_1114 was mixed equimolarly with Saci_1109 and applied to size exclusion chromatography. However, there was no complex formation observed between these two proteins under the employed *in vitro* conditions. All the four β oxidation enzymes were also analysed by HPLC. First, the complete conversion of butyryl-, hexanoyl- and octanoyl-CoA to the respective enoyl-CoA esters by ACAD was shown with ETF or with the artificial electron acceptor DCPIP or $FcPF_6$. It could also be demonstrated that bifunctional HCDH/ECH Saci_1109 converted >90% of the crotonoyl-CoA to hydroxybutyryl-CoA although the further conversion with NAD^+ to acetoacetyl-CoA could not be observed. However, when the KT was introduced nearly complete conversion of crotonoyl-CoA to acetyl-CoA could be detected without any detectable amounts of liberated acetoacetyl-CoA and with very less 3-hydroxybutyryl-CoA. In addition, the complete thiolytic cleavage of

acetoacetyl-CoA to acetyl-CoA by the KT Saci_1114 alone was obtained. Also, the reversible conversion of acetyl-CoA to crotonoyl-CoA using Saci_1114 and Saci_1109 was demonstrated although only low amounts of crotonoyl-CoA and the intermediate 3-hydroxybutyryl-CoA were observed. Furthermore, the purified enzymes were reconstituted to the complete β oxidation spiral *in vitro* and the conversion of butyryl-CoA, hexanoyl-CoA or octanoyl-CoA to acetyl-CoA was shown.

1.5 FA synthesis in *S. acidocaldarius* (?)

Furthermore, in addition to these indications that FAs serve as carbon and energy sources at least in some Archaea likely metabolized via a bacterial like β oxidation, scarce reports previously suggested that in few Archaea FAs are present including *Sulfolobus spp*.. Thus - if this is indeed the case - these archaeal organisms must have a biosynthesis machinery for FAs. However, neither a complete classical FAS II system known from bacteria nor a FAS I machinery present in animals and fungi have been identified in any archaeon. Also, the acyl carrier protein and the acyl carrier protein synthase essential for the FA biosynthesis in bacteria and eukaryotes is not present in most archaea with only very few exceptions [18]. Thus, if at least some Archaea are able to synthesize FAs then the synthesis machinery must be fundamentally different from the known systems in Bacteria and Eukaryotes. Recently, an ACP-independent FA synthesis pathway has been proposed which relies on bacterial like homologues of the FA synthesis system (FAS II) [27]. However, complete sets of such homologues could not be identified in any archaeon so far and instead, a reversed β oxidation has been proposed to be responsible for FA synthesis in Archaea [18].

However, also the reversibility of the β oxidation in general and particularly in Archaea has not been demonstrated so far. Instead, both processes known from Bacteria and Eukaryotes, the FA synthesis on the one hand and the β oxidation on the other hand which basically follow the same chemical conversions, show some remarkable differences in order to drive either of the processes in the desired direction and to separate both processes which at least in prokaryotes are localized in the cytoplasm. First of all both processes are tightly regulated on the transcriptional level but also on the protein level e.g. by feedback inhibition [26]. As mentioned above the synthesis machinery is ACP dependent which activates the FAs and transports the growing acyl chain between the involved enzymes [28]. Conversely, the β oxidation relies on CoA for FA activation. The FA synthesis is usually NADPH dependent whereas the β oxidation uses NAD^+ and FAD [29] although this cosubstrate specificity is less pronounced in Bacteria [30]. The Claisen condensation, the first reaction in the synthesis, uses malonyl-ACP as donor for chain elongation instead of acetyl-CoA, and malonyl-ACP has to be synthesized through ATP-dependent carboxylation of acetyl-CoA (and transfer to ACP). During condensation the CO_2 is liberated again. Thus, the endergonic reaction is energetically driven by ATP hydrolysis.

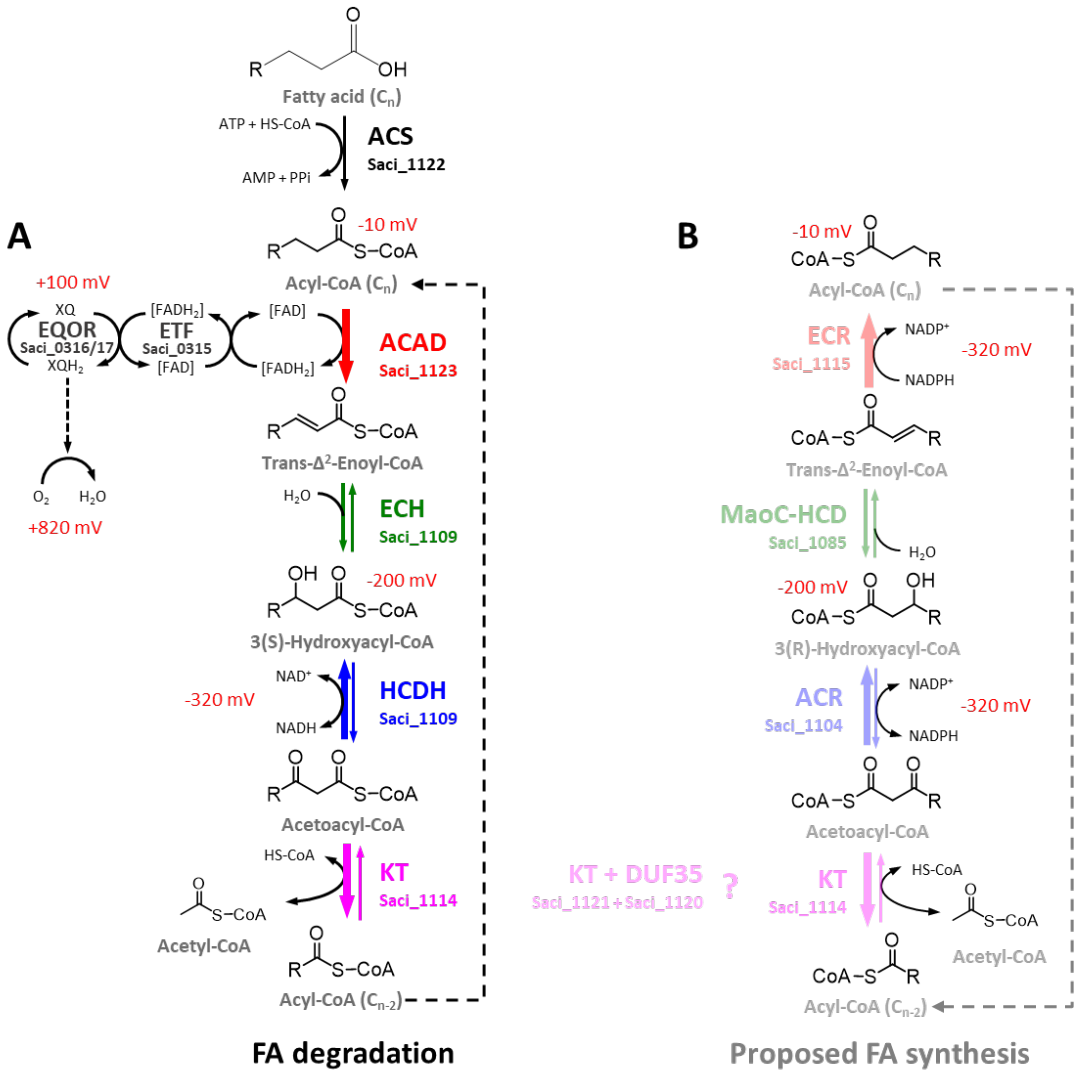


Figure 5: Reconstructed fatty acid degradation (A) and proposed potential fatty acid biosynthesis (B) in *S. acidocaldarius*. After channelling through the outer membrane, FAs are activated to the precursor fatty acyl-CoAs via AMP-forming acyl-CoA synthases (ACSs). Acyl-CoAs enter the first step in β oxidation spiral and are oxidized to unsaturated enoyl-CoA by acyl-CoA dehydrogenases (ACADs) simultaneously transferring the electron to oxygen of the respiratory chain through electron transferring flavoproteins (ETFs) and ETF:quinone oxidoreductases (EQORs). Then enoyl-CoAs are converted to 3(S)-hydroxyacyl-CoA and further oxidized to acetoacyl-CoA by multifunctional enzyme 3(S)-hydroxyacyl-CoA dehydrogenases/enoyl-CoA hydratases (HCDHs/ECHs). Finally, β-ketothiolases (KTs) cleave acetoacyl-CoAs into acetyl-CoA and acyl-CoAs with two carbon atoms shortened in its acyl chain. The shortened acyl-CoAs can further participate in the cycle until the acyl chains are broken down to acetyl-CoAs. In the proposed novel fatty acyl elongation pathway in *S. acidocaldarius* (B), short or medium chain acyl-CoAs (e.g. C2, C4 or C6) together with acetyl-CoA are condensed to synthesize acetoacyl-CoAs by the archaeal type KTs. The resulting intermediates are further reduced to the 3(R)-hydroxy CoA esters by SDR superfamily acetoacyl-CoA reductases (ACRs) employing NADPH as cofactor. Then MaoC-like 3(R)-hydroxyacyl-CoA dehydratases (MaoC-HCDs) remove one molecule of water from each 3(R)-hydroxy derivative to form double-bond containing enoyl-CoAs. The MDR family enoyl-CoA reductases (ECRs) catalyze the last step to reduce the metabolites to acyl-CoAs with two carbon atoms longer than before. The newly resulting acyl-CoAs will be further elongated until a desired length is reached. The subscript “n” of acyl-CoA represents the length of the carbon chain which equals 4, 6 or 8. The thickness of the arrow indicates the energetics of the respective reaction. For redox reactions the reduction potential is given in red.

Furthermore, the Claisen condensation itself catalysed by the decarboxylating ketoacyl-ACP synthase is regarded as irreversible in the synthesis direction [30, 31]. The following two reactions in the FA synthesis pathway, i.e. the reduction of oxoacyl to the hydroxyacyl intermediate and the dehydration to the enoyl moiety are in principle reversible but are specific for the R-hydroxyacyl intermediate instead of the S-stereoisomer specific β oxidation [29]. Finally, the interconversion of the saturated acyl- to the unsaturated enoyl-intermediate on the β oxidation is carried out by FAD dependent dehydrogenases channelling the electrons via the electron transferring flavoprotein (ETF) and the ETF:quinone oxidoreductase (EQOR) into the quinone pool of the respiratory chain (Fig. 5A) [32, 33]. This reaction is generally regarded as irreversible in the oxidative direction [34] and usually bypassed in the FA synthesis by NAD(P)H dependent dehydrogenases which in turn makes the reaction highly exergonic towards the reductive direction [30, 35].

In the second part of chapter 3.2 a potential FA synthesis in *S. acidocaldarius* (Fig. 5B) was analysed and firstly the reversibility of the newly characterized β oxidation was elucidated. For the acyl-CoA oxidation, the experimental results showing a complete conversion of the acyl-CoA esters to the respective enoyl-CoAs with both, ETF and ferrocenium as electron acceptors, catalyzed by ACAD strongly indicated that this reaction represents a major driving force of the β oxidation. This is in line with the reported mechanism described for the ACAD which preferentially binds the product enoyl-CoA in its reduced state and thus kinetically promotes the oxidative half-reaction, i.e. the electron transfer from the ACAD flavin to the ETF or the artificial electron acceptor. Thus, from kinetic point of view the reaction is already regarded as irreversible. This assumption is further supported by the thermodynamics of the catalyzed reaction(s). Although the redox potential of the flavins in the ACAD and the ETF might be in a similar range, the redox potential of the caldariella quinone (CQ) (+100 mV) in the respiratory chain and finally of course of oxygen (+820 mV) as terminal electron acceptor of the respiratory chain of *S. acidocaldarius*, is much more positive than that of the acyl-CoA/enoyl-CoA couple (-10 mV) (Fig. 5A) making a reversal of this reaction *in vivo* particularly unlikely.

Furthermore, although the following three reactions in the β oxidation are in principle reversible as also shown here for the *S. acidocaldarius* enzymes, the equilibrium is far on side of the degradation (>90%) which is also in accordance with the thermodynamics of the reaction(s) with a total free energy change of -11 kJ mol^{-1} for the whole crotonyl-CoA to acetyl-CoA conversion. Here, the KT reaction is the major driving force (-26.1 kJ mol^{-1}) whereas the hydroxyacyl-CoA dehydrogenase represents the bottleneck (+18 kJ mol^{-1}) in the degradative direction. From these results and considerations it becomes apparent that the β oxidation as an entire process is hardly reversible from a mechanistic and energetic point of view and that the ACAD catalyzed reaction with the CQ as “primary” electron acceptor is the major bottle neck.

Thus, to drive the whole process in the reverse namely the synthesis direction at least this reaction needs to be bypassed. In the canonical FA synthesis pathways this is usually accomplished by the use of NAD(P)H ($E^{0'}$ -320 mV) as electron donor which renders the reaction highly exergonic towards saturated acyl thioester formation (crotonyl-CoA/butyryl-CoA, $E^{0'}$ -10 mV) with a standard free energy change corresponding to -60 kJ mol^{-1}.

Strikingly, the *saci_11xx* gene cluster also contains genes which do not obviously contribute to lipid degradation or FA β oxidation, respectively. One of these genes (*saci_1115*), encoding a putative alcohol dehydrogenase from the medium chain dehydrogenase/reductase (MDR) superfamily, showed remarkable similarity to the acryloyl-CoA reductase from *Metallospaera sedula* suggesting enoyl-CoA reductase activity (Fig. 4C, Fig. 5B). Also, this gene was recombinantly expressed, the protein was purified and characterized confirming NADPH dependent enoyl-CoA reductase activity with a broad chain length specificity towards medium chain substrates C4 to C10. In accordance with the thermodynamics of the reaction, the enzyme did not operate in the oxidative direction, and thus represents a likely candidate to catalyse the irreversible enoyl-CoA reduction during FA synthesis in *S. acidocaldarius* and thus a potential driving force for this process (Fig. 5B). It is interesting to note, that in Bacteria the analogous reaction during FA synthesis is carried out by enoyl-ACP reductases (encoded by fabI, fabL, or fabV) from the short chain dehydrogenase/reductase (SDR) superfamily whereas the mitochondrial FAS II enoyl thioester reductase and the enoyl reductase components of the FAS I systems of other eukaryotes (mammalia and fungi) also belong to the MDR superfamily [36]. Correspondingly, our experimental data showed the Saci_1115 protein was also a homodimer under the native conditions.

Also – as the mitochondrial enzymes – Saci_1115 showed a clear preference for NADPH as cosubstrate. The cosubstrate preference is considered as a distinctive property between FA synthesis and degradation with NADPH preferred by the FAS systems and NAD^+ preferably used by the β oxidation. Thus, although the Saci_1109 bifunctional HCDH/ECH operates reversibly *in vitro* the NAD^+ preference might suggest a favorable function in β oxidation rather than biosynthesis. Furthermore, Saci_1109 showed a clear preference for the 3(S)-hydroxybutyryl-CoA. This stereospecificity is another distinctive feature of the β oxidation whereas the FA synthesis is specific for the 3(R) stereoisomers.

Interestingly, Saci_1104 also encoded in the *saci_11xx* gene cluster, a 3-oxoacyl-ACP reductase homologue from the SDR superfamily (and the only FAS II enzyme for which homologues are present in *S. acidocaldarius*), clearly showed hydroxybutyryl-CoA dehydrogenase activity with a clear specificity for 3(R)-hydroxybutyryl-CoA and NADPH as cosubstrate and thus appears as a suitable candidate to function as 3-oxoacyl-CoA reductase in course of FA synthesis in *S. acidocaldarius* (Fig. 4C, Fig. 5B). However, substrate specificity

of Saci_1104 towards different chain length was not yet known, which should be the focus of future work.

The presence of an NADPH dependent reductase with (R)-hydroxyacyl-CoA specificity, raised the question whether *S. acidocaldarius* also harbors (R)-specific hydroxyacyl thioester dehydratases. In Bacteria and Eukarya this reaction is carried out by dehydratases from the hotdog fold superfamily [29, 37, 38], which were not present in the *S. acidocaldarius* genome. Only distantly related MaoC dehydratases which also belong to the hotdog fold superfamily were identified (Saci_1070, Saci_1085). The recombinantly overproduced and purified Saci_1085 indeed showed ketoacyl-CoA dehydratase activity with a pronounced preference for the (R)-stereoisomer (Fig. 5B). The enzyme could be measured in both directions of the reaction and the kinetic constants were determined. The formation of crotonyl-CoA from 3(R)-hydroxybutyryl-CoA was also shown via HPLC analyses and addition of the enoyl-CoA reductase (Saci_1115) led to the formation of butyryl-CoA. Furthermore, with the reconstitution of these two enzymes with the Saci_1104 ketoacyl-CoA reductase the conversion of acetoacetyl-CoA to butyryl-CoA could be shown. However, these results together with the reversal of the ketothiolase putatively enabled by a DUF35 scaffolding protein (see chapter 3.2 for details) indicate a potential novel FA synthesis pathway in *S. acidocaldarius* comprising a bacterial like SDR superfamily R-specific fabG homolog, an MDR superfamily enoyl-CoA reductase like in Eukaryotes, both with a clear preference for NADPH as electron donor, as well as an R-specific MaoC like dehydratase.

Taken together, we herein report the first comprehensive biochemical study on the FA metabolism in Archaea, e.g. from the aerobic, thermoacidophilic crenarchaeal model organism *S. acidocaldarius*. Although the FA β oxidation at a first glance looks quite similar to the known pathway from Bacteria and mitochondria, the pathway shows some unusual features with respect to the ETF and the HCDH/ECH bifunctional enzyme together with the previously recognized “archaeal type” ketothiolases (Fig. 5A) [18]. Furthermore, our results strongly indicate that the β oxidation as entire pathway is not operating reversibly in Archaea. Instead, we propose a potential archaeal FA synthesis pathway which shows a kind of mosaic character with similarities to both bacterial (fabG, MaoC) and eukaryal (MDR enoyl thioester reductase) features mixed with unique archaeal properties (DUF35 domain/KT complexes, ACP independence) (Fig. 5B).

1.6 Response of *S. acidocaldarius* to solvent stress exemplified by butanol exposure

S. acidocaldarius is not only a suitable model organism to study archaeal biology, with its ease of cultivation, genetic tractability enabling metabolic engineering, established -omics approaches, and metabolic versatility it shows also an outstanding potential for

biotechnological applications. In consideration of global warming and the depletion of fossil resources sustainable bio-based alternatives to petroleum, coal or gas are urgently needed as source for energy and base chemicals. In this respect, waste materials from agriculture and industry like cellulose, hemicellulose, lignin, ash and lignocellulosics (straw, saw dust, wood chips etc.) become more and more interesting as cheap and readily available raw materials. Conventional manner to pre-treat these wastes is hydrolysis by high temperature and acidic processing.

Especially, the adaptation of *S. acidocaldarius* to two extremes, i.e. to high temperature and low pH, offers several advantages. (i) Contamination from the mesophilic microorganisms is greatly reduced. (ii) Solubility of the substrate is significantly improved. (iii) Cooling costs required for the mesophilic fermentations are saved. (iv) The separation of volatile compounds such as alcohols can be accomplished through high process temperature without additional steps [25].

In combination with the broad substrate specificity, missing catabolite repression, polymer degrading capabilities and metabolic versatility, *S. acidocaldarius* is a promising candidate for the production of value-added products from these waste materials including biofuels and base chemicals.

A key commodity widely used solvent or chemical feedstock is 1-butanol. So far, 1-butanol is mainly produced chemically by the Oxo process [39]. Fossil oil-derived raw materials used in the production of butanol are ethylene, propylene and triethyl aluminium or carbon monoxide and hydrogen [40]. Biobutanol represents a promising alternative as a fuel additive and biofuel for direct replacement of gasoline [41, 42]. Production of biobutanol from renewable resources is predominantly accomplished by *Clostridium* strains via acetone butanol ethanol (ABE) fermentation [39, 40, 43]. However, while ABE fermentation supplied approx. 66 % of the world butanol supply until the 1950’s, bio-based butanol production was outcompeted by petroleum-based processes afterwards [39].

A problem in the production of biobutanol is its toxicity towards microbial cells. For a vast majority of microorganisms, including *Clostridium* strains, a growth limit at 1% - 2% (v/v) of butanol in nutrient medium has been observed in liquid cultures [44-47]. There is the widely accepted notion that butanol toxicity results from its chaotropic effects on the cytoplasmic cell membrane, leading to the disruption of nutrient and ion transport and the loss of the membrane potential [48, 49]. Bacteria and eukaryotic microorganisms are able to adapt to the presence of aliphatic, toxic alcohols, including acetone, ethanol, butanol, isobutanol and propanol, with the development of an enhanced tolerance, allowing survival and growth at elevated concentrations of these compounds [45, 50, 51]. The adaptation strategies are versatile. Microorganisms can respond to alcohol exposure by changing their membrane lipid

composition to sustain membrane fluidity, called “homeoviscous adaptation” [52-54]. This process may include a shift in the ratio of unsaturated to saturated lipids, branched and unbranched lipids, changes of their isomerization and cyclization state or head-group composition [45, 48, 49, 52, 55]. In gram-negative bacteria outer membrane modifications (e.g. alterations in the lipopolysaccharide content and porin expression, interaction with divalent metal ions for stabilization and outer membrane vesicle formation) were reported [49, 52, 55]. Other cellular responses described for bacteria include the upregulation of energy-dependent efflux systems to lower the intracellular solvent concentration [51, 56, 57], and the metabolic degradation of solvents [48, 52]. Organic solvents, including butanol, were shown to enhance expression of heat shock proteins (HSPs), including molecular chaperones that assist correct protein folding and transport as well as the recycling of defective proteins [51, 58]. Cell aggregation and biofilm formation was shown to enhance tolerance to 1-butanol as observed for butanol production strains *C. acetobutylicum* and *C. beijerinckii* [59-61]. Biofilms are defined as microbial aggregates, embedded in a matrix of extracellular polymeric substances (EPS) on surfaces and other interfaces [62, 63]. The biofilm mode of life is dominant among prokaryotic microorganisms [64] and in general offers advantages for survival compared to the planktonic lifestyle, for example an enhanced tolerance against adverse environmental conditions [63] as they may be encountered in biotechnological processes due to toxic reactants or products.

Regarding organic solvent tolerance, archaeal extremophiles may offer advantages over mesophilic organisms due to their intrinsic robustness and adaptation to hostile environments [65-67]. However, the response to solvents or solvent stress has so far not been analyzed. *S. acidocaldarius* as potential thermoacidophilic platform organism for biotechnological application including bio refinery is able to form biofilms, and methods for analyzing *Sulfolobus* biofilms have recently been established. Proteins, carbohydrates and DNA have been identified as constituents of the EPS matrix of *S. acidocaldarius* [68]. The study described in chapter 3.3 investigated the natural ability of *S. acidocaldarius* to tolerate 1-butanol and its cellular response towards this industrially relevant organic solvent. This work contributes to the investigation of the planktonic growth properties of *S. acidocaldarius* exposed to certain concentrations of 1-butanol.

In response to butanol exposure, biofilm formation of *S. acidocaldarius* was induced and occurred up to 1.5% (v/v) 1-butanol, while planktonic growth was only observed up to 1% (v/v) 1-butanol. Confocal laser scanning microscopy revealed that the biofilm architecture changed with the formation of denser and higher tower-like structures and an increase in carbohydrate-containing material overlying the biofilm surface. Concomitantly, enhanced carbohydrate and protein content of the extracellular polymeric substances isolated from butanol-exposed biofilms was observed by biochemical analysis.

To study the cellular global response upon solvent exposure we performed transcriptome and proteome analysis comparing the response of planktonic and biofilm cells in absence and presence of 1-butanol. In response to 1% (v/v) 1-butanol we observed major changes in motility, cell envelope and membrane composition, cell division and/or vesicle formation, immune and defense systems, as well as metabolism and general stress response. Our findings show that the extreme life style of *S. acidocaldarius* coincided with a high tolerance to organic solvents. In addition, these studies provide first insights into biofilm formation and membrane/cell stress caused by organic solvents in *S. acidocaldarius*. This is the first detailed study on solvent stress response in a crenarchaeon highlighting the impressive robustness of the thermoacidophilic *S. acidocaldarius* towards organic solvents.

2 Scope of the thesis

In the current dissertation, the included research topics are illustrated in the following chapters and the contributions I have done for these manuscripts are clarified as below.

In chapter 3.1, the manuscript with the title "**A TetR-family transcription factor regulates fatty acid metabolism in the archaeal model organism *Sulfolobus acidocaldarius***", a TetR family transcriptional regulator from *S. acidocaldarius* was characterized. The paper included the determination of the crystal structure and DNA-binding mechanisms of this regulator. Repression of the TetR factor on the β oxidation genes as well as the derepression due to binding to fatty acyl-CoAs were demonstrated. In this research, I established and analyzed the growth of *S. acidocaldarius* strain MW001 and the TetR mutant on short/medium chain fatty acids namely butyrate or hexanoate as well as on fatty acid-based lipid tributyrin. In addition, I contributed to cell preparation and RNA extraction for qRT-PCR to study the relative gene expression.

In chapter 3.2, the manuscript with the title "**Fatty acid metabolism in *Sulfolobus acidocaldarius***", we intensively studied the fatty acid metabolism in *S. acidocaldarius* by recombinant expression, purification and characterization of relevant enzymes. A bacterial-like β oxidation as well as a novel fatty acid synthesis were reconstructed *in vitro* via HPLC and biochemical approaches for the first time in Archaea. For this manuscript, I heterologously expressed and purified the proteins encoded by *Saci_1122*, *Saci_1109*, *Saci_1114*, *Saci_1104* and *Saci_1115*, and measured their activities using different fatty acid or acyl-CoA substrates. I also made contribution to HPLC analysis of the saci_1114 activity and edition of the manuscript.

In chapter 3.3, the manuscript with the title "**Response of the thermoacidophilic Archaeon *Sulfolobus acidocaldarius* to solvent stress exemplified by 1-butanol exposure**", we unrevealed the natural ability of *S. acidocaldarius* to tolerate 1-butanol and its cellular response towards this industrially relevant organic. Employing microscopic techniques it was shown that cells formed denser biofilm structures as a response to butanol exposure (1% (v/v)). Further, CLSM as well as EPS isolation and quantification revealed increased amounts of extracellular carbohydrates and proteins at 1% (v/v) butanol. To investigate the cellular global response to solvent stress, transcriptome and initial proteome analyses were accomplished. The contribution of this thesis to the publication was to investigate the planktonic growth properties of *S. acidocaldarius* DSM639 (shaking cultures) exposed to certain concentrations of 1-butanol.

3 Manuscripts

Chapter 3.1

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

nature COMMUNICATIONS

ARTICLE

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

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

Kun Wang[1], David Sybers[2], Hassan Ramadan Maklad[2], Liesbeth Lemmens[2], Charlotte Lewyllie[2,5], Xiaoxiao Zhou[3], Frank Schult[3], Christopher Bräsen[3], Bettina Siebers[3], Karin Valegård[4], Ann-Christin Lindås[1] & Eveline Peeters[2]

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 ($FadR_{Sa}$) is involved in regulation of fatty acid metabolism in the crenarchaeon *Sulfolobus acidocaldarius*. Functional and structural analyses show that $FadR_{Sa}$ 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 $FadR_{Sa}$ binding by inducing conformational changes in the protein. Our results indicate that, despite its similarity in overall structure to bacterial TetR-family FadR regulators, $FadR_{Sa}$ displays a different acyl-CoA binding mode and a distinct regulatory mechanism.

[1] Department of Molecular Biosciences, The Wenner-Gren Institute, Stockholm University, Svante Arrhenius v. 20C, SE-10691 Stockholm, Sweden. [2] Research Group of Microbiology, Department of Bioengineering Sciences, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussels, Belgium. [3] Molekulare Enzymtechnologie und Biochemie, Biofilm Centre, ZWU, Fakultät für Chemie, Universität Duisburg-Essen, Universitätsstr. 2, 45117 Essen, Germany. [4] Molecular Biophysics, Department of Cell and Molecular Biology, Uppsala University, Uppsala, Sweden. [5] Present 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)

The phylogenetic classification of archaea as a domain of life distinct from bacteria[1] is supported by the nature of their membrane lipids having isoprenoid-based hydrocarbon chains instead of fatty acids[2]. Despite the absence in archaeal membrane lipids, small amounts of fatty acids and derivatives have been detected in archaeal cells[3–9]. The role of fatty acids for archaeal cellular physiology is not yet clear and a controversial issue of debate[10], although an involvement in the acylation or stabilization of membrane-bound energy-conversion proteins such as rhodopsin or cytochromes has been postulated[10–12]. Many archaeal genomes have extensive sets of typical bacterial-like genes encoding fatty acid synthase type II (FAS-II) complex and β-oxidation enzymes[10,13,14]. An outstanding question is whether these fatty acid metabolism genes perform anabolic or catabolic reactions, or both[9,10]. Given the absence of genes encoding acyl-carrier protein (ACP) or ACP synthase[13], it has been postulated that a β-oxidation pathway might operate in the reverse direction in conjunction with acetyl-CoA C-acetyltransferase enzymes[10]. These are abundantly encoded in archaeal genomes, sometimes in the direct neighborhood of β-oxidation genes[10,14].

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[15]. 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[16,17], whereas a TetR-family malonyl-CoA-dependent regulator FabR controls the ratio between monounsaturated and saturated fatty acids[18–20]. 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 degradation[21] and a DeoR family member FapR that regulates biosynthesis of saturated fatty acids and phospholipids[22]. 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[23–26].

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 archaeal model organism[27,28]. *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[10]. 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[29]. 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 $FadR_{Sa}$, thereby unveiling the function and mode of action of an acyl-CoA-responsive transcriptional regulator in an archaeal microorganism.

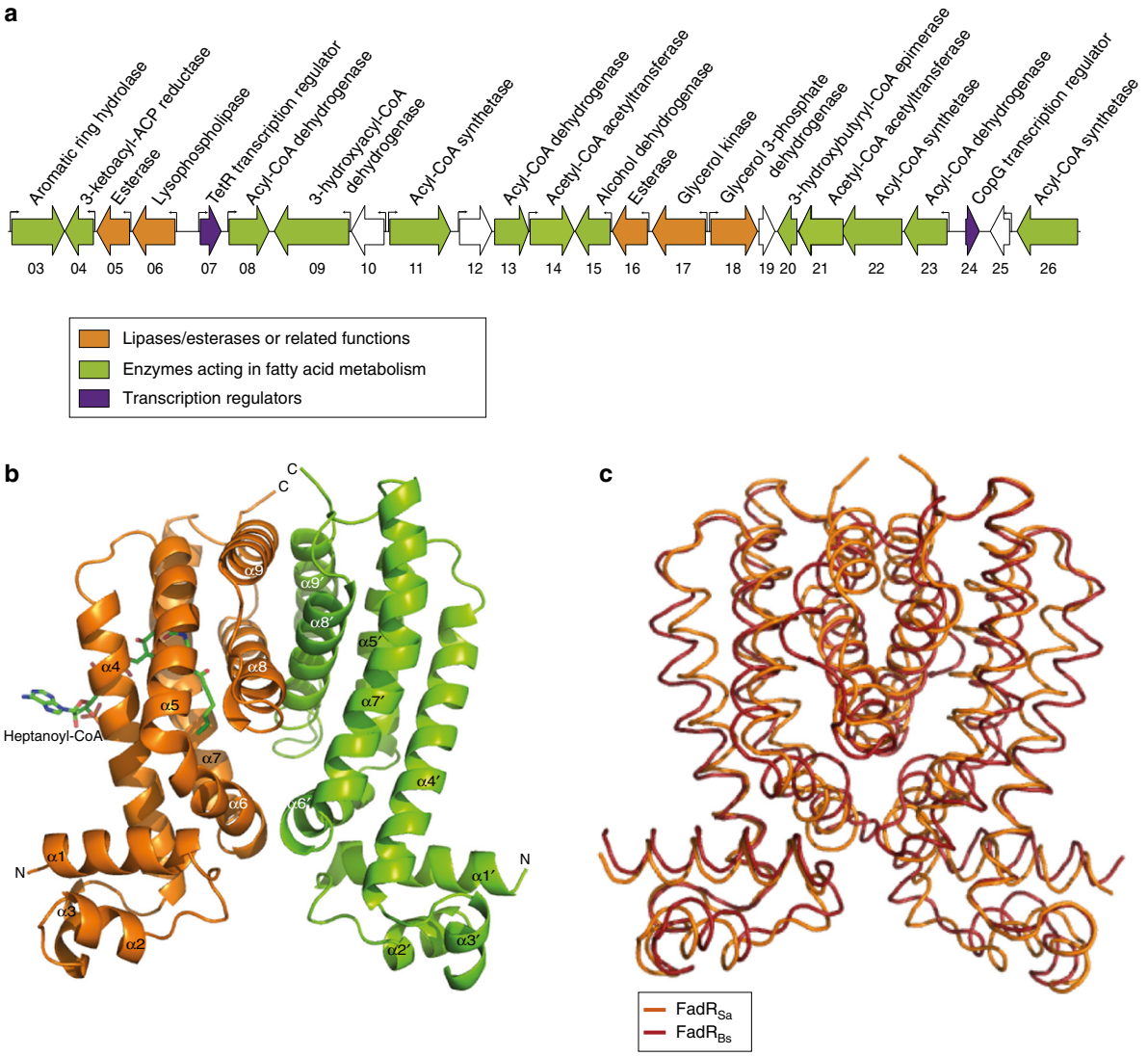
Results

$FadR_{Sa}$ 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[30].

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., $FadR_{Bs}$ (RMSD = 4.23 Å) and $FadR_{Bh}$[21,26] (RMSD = 5.88 Å), and *Thermus thermophilus*, $FadR_{Tt}$[24] (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 $FadR_{Sa}$ accordingly.

Upon solving the $FadR_{Sa}$ 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 cocrystallization of acyl-CoA with $FadR_{Sa}$ (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 $FadR_{Sa}$. As a next step toward unraveling $FadR_{Sa}$ 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 enrichments 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 $fadR_{Sa}$ 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 $fadR_{Sa}$ (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 FadR$_{Sa}$–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 FadR$_{Sa}$–DNA complexes with high affinity and positive-binding

Fig. 1 *S. acidocaldarius* harbors a TetR-family regulator structurally similar to bacterial TetR-like FadR proteins. **a** Genomic organization of the *Saci_1103-Saci_1126* gene cluster. Gene numbers are indicated by displaying the last two digits below each gene arrow. Transcriptional start sites are shown with small arrows, and are based on ref. [37]. **b** Structure of the $FadR_{Sa}$ dimer with indication of the different helices (chain A: α1′–α9′; chain B: α1–α9). The acyl-CoA molecule present in chain B is shown as a stick model. **c** Superposition of the $FadR_{Sa}$ and $FadR_{Bs}$ (PDB 3WHB) structures. **d** Structure-based sequence alignment of TetR-family FadR proteins. The alignment is based on a three-dimensional comparison between $FadR_{Sa}$ from *S. acidocaldarius* (PDB 5MWR), $FadR_{Bs}$ from *B. subtilis* (PDB 3WHB)[25], $FadR_{Bh}$ from *B. halodurans* (PDB 5GP9)[26], and FadR from *T. thermophilus* (PDB 3ANG)[24]. Regions harboring structural and sequence similarity are boxed, with identical amino-acid residues indicated as bold white letters on a red background and functionally equivalent residues indicated in red letters. Secondary structure elements and numbering for $FadR_{Sa}$ are indicated above the sequences. DNA-binding residues targeted for mutagenesis are indicated with purple triangles, ligand-binding residues with green triangles. $FadR_{Bh}$ residues important for DNA binding and ligand binding[26] are indicated below the sequences with purple and green asterisks, respectively

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

	SeMet $FadR_{Sa}$	$FadR_{Sa}$:DNA	$FadR_{Sa}$: lauroyl-CoA
Data collection			
Space group	$P2_1$	$P2_12_12_1$	$P2_1$
Cell dimensions			
a, *b*, *c* (Å)	42.0, 98.8, 56.0	54.8, 178.3, 266.7	46.0, 91.8, 53.5
α, *β*, *γ* (°)	90.0, 106.4, 90.0	90.0, 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
R_{merge}	0.046 (0.363)	0.095 (0.746)	0.065 (0.727)
$I/\sigma(I)$	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)
Redundancy	2.2	6.8	3.3
Refinement			
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			
Protein	3075	9345	4837
Ligand/ion	56	1783	122
Water	9	0	67
Wilson *B*-factors (Å^2)	53.5	97.4	34.0
R.M.S. deviation			
Bond lengths (Å)	0.009	0.008	0.002
Bond angles (°)	0.99	0.91	1.031

cooperativity. In contrast, with the low-enrichment targets only a single complex was formed with lower affinity and without apparent cooperativity (Fig. 2d and Supplementary Figure 4). These findings suggest that for different genomic targets, $FadR_{Sa}$ is capable of using distinct DNA-interaction modes differing in binding stoichiometry.

Mechanisms of DNA binding. To further unravel mechanisms of DNA binding by $FadR_{Sa}$, we determined the cocrystal structure of the protein–DNA complex to a resolution of 3.29 Å using a duplex DNA containing the predicted $FadR_{Sa}$ binding motif in the control region of the *fadR*$_{Sa}$ gene itself, corresponding to ChIP-seq peak 2 (Table 1 and Fig. 3a, b). The asymmetric unit contained six $FadR_{Sa}$ 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 $FadR_{Sa}$).

In each $FadR_{Sa}$ subunit in the cocrystal 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 $FadR_{Sa}$ residues Tyr47, Leu49, Tyr51, and Phe52 and methyl groups of thymines (Fig. 3c) similar as in other TetR-like regulators[26,30] (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). $FadR_{Sa}^{Y47A}$, $FadR_{Sa}^{Y51A}$, and $FadR_{Sa}^{Y53A}$ are all negatively affected in DNA-binding affinity and cooperativity. With $FadR_{Sa}^{G48A}$, 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 $FadR_{Sa}$ dimers bound on different sides of the DNA helix (Fig. 3d and Supplementary Note 1).

DNA-binding stoichiometry of $FadR_{Sa}$. To dissect the stoichiometric nature of the electrophoretically distinct $FadR_{Sa}$–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 $FadR_{Sa}$–*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 $FadR_{Sa}$ binds the *Saci_1123* operator as a single dimer. The observation that the relative mobility in EMSA of $FadR_{Sa}$–*Saci_1123* complex B is highly similar to that of the $FadR_{Sa}$–*fadR*$_{Sa}$ complex B1 (Fig. 4b), led us to postulate that the transitional $FadR_{Sa}$–*fadR*$_{Sa}$ complex B1 has a stoichiometry similar as for the sole $FadR_{Sa}$–*Saci_1123* complex. In contrast, the dominantly formed complex B2 with the *fadR*$_{Sa}$ operator has a larger apparent MW (179 kDa) (Fig. 4a): it can be assumed that the apparent MW attributed by the $FadR_{Sa}$ protein itself is similar for measurements of free protein and of $FadR_{Sa}$–DNA complexes and that the $FadR_{Sa}$–*fadR*$_{Sa}$ 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 $FadR_{Sa}$ 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 $FadR_{Sa}$–DNA complexes B1 and B2 observed in EMSAs with *fadR*$_{Sa}$ 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 *fadR*$_{Sa}$ operator probe,

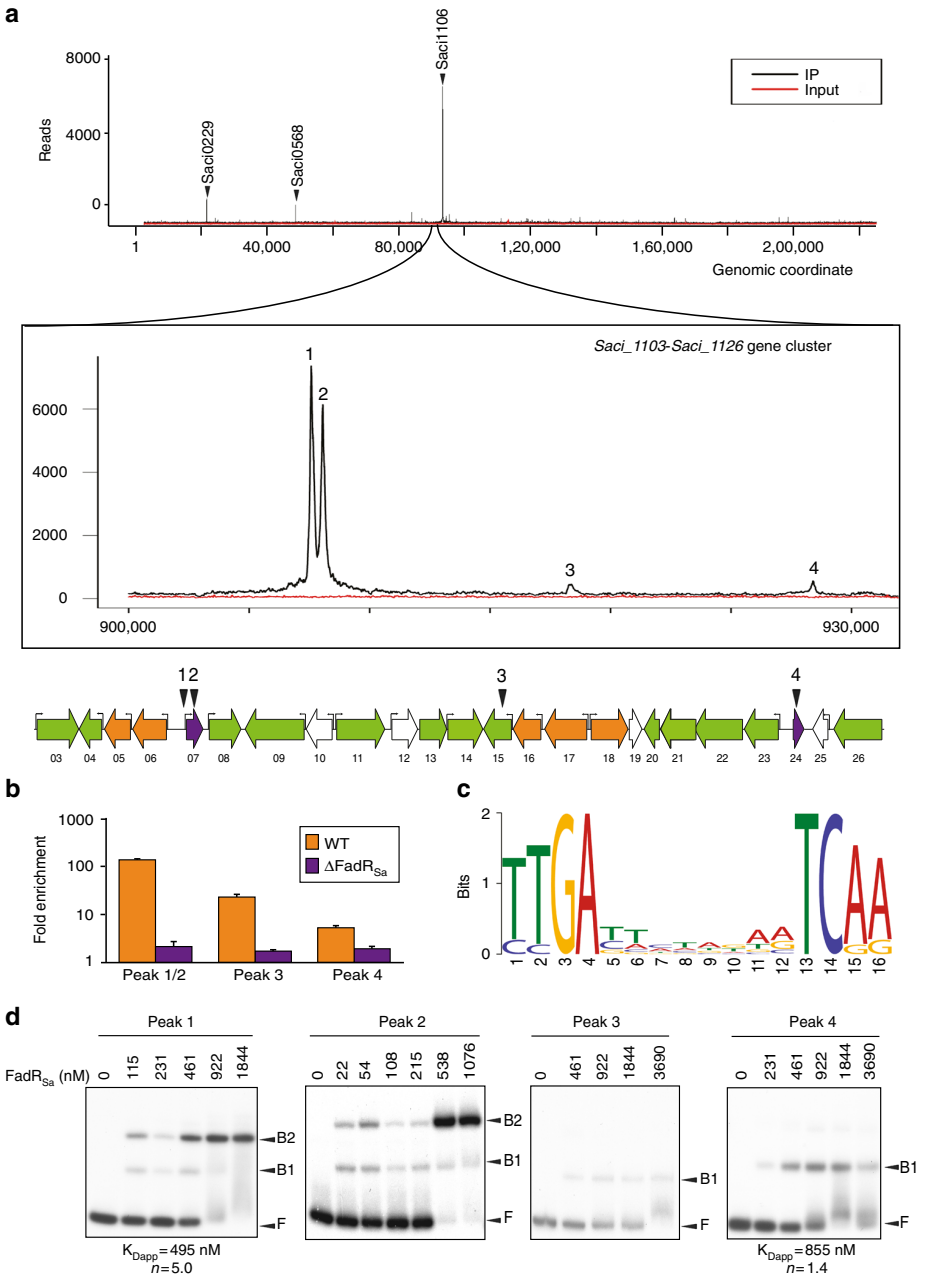


Fig. 2 $FadR_{Sa}$ interacts with four genomic loci in the *Saci_1103-Saci_1126* gene cluster. **a** Overview of the genomic binding profile of $FadR_{Sa}$ as monitored by ChIP-seq (IP = immunoprecipitated sample). A zoomed image of this profile is shown for the genomic region encompassing the *Saci_1103-Saci_1126* 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 $FadR_{Sa}$ in the ChIP-seq profile. Error bars represent standard deviations of biological duplicates. **c** Sequence logo of the $FadR_{Sa}$ binding motif representing MEME predictions of ChIP-seq enriched sequences. **d** Electrophoretic mobility shift assays of $FadR_{Sa}$ 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 $FadR_{Sa}$-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)

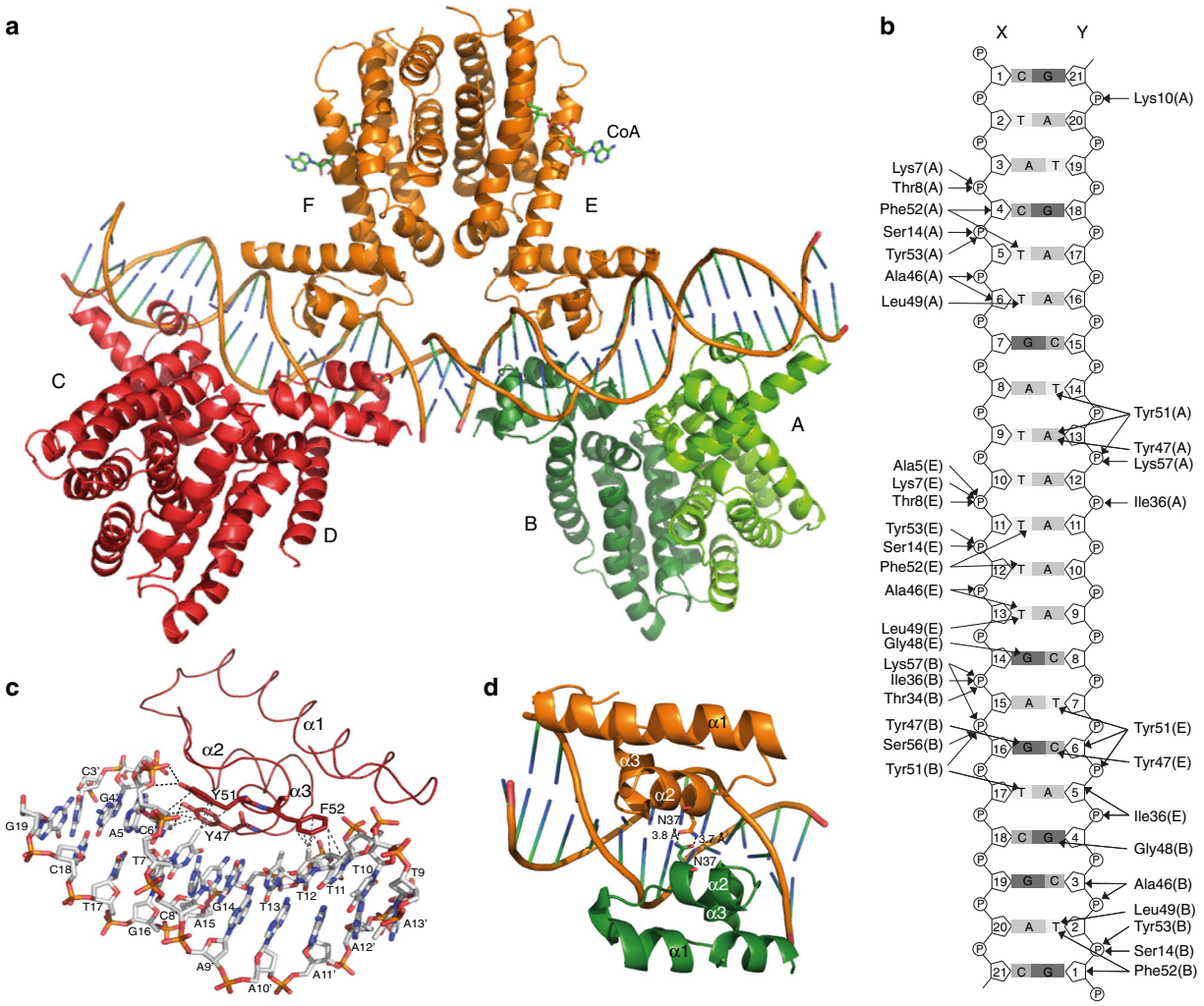


Fig. 3 A FadR$_{Sa}$–DNA cocrystal structure reveals an important role for hydrophobic interactions. **a** Cartoon representation of the cocrystal structure of FadR$_{Sa}$ in complex with the operator site in the *fadR$_{Sa}$* promoter region. Different FadR$_{Sa}$ subunits are labeled A–E, each dimer 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 FadR$_{Sa}$–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

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 FadR$_{Sa}$–*Saci_1123* complex B and the FadR$_{Sa}$–*fadR$_{Sa}$* complex B1 a single-FadR$_{Sa}$ dimer interacts with the predicted semipalindromic binding motif. A second dimer interacts with the upstream (left) side of the inverted repeat in FadR$_{Sa}$–*fadR$_{Sa}$* complex B2, suggesting a dimer-of-dimer interaction mode similar to that observed for a bacterial subclass of TetR-like regulators prototyped by QacR in *Staphylococcus aureus*, in which two overlapping inverted repeats are bound by dimers located on opposite sides of the DNA helix[31].

To relate the molecular interactions and binding architecture in the FadR$_{Sa}$–DNA cocrystal structure with the biologically relevant stoichiometries, we mapped the contacts on the *fadR$_{Sa}$* and *Saci_1123* operator sequences that could be hypothesized to exist based on the cocrystal structure (Fig. 5a). The presence of a G–C bp 11 positions upstream of the symmetrical C–G in the pseudopalindromic site appears to be the sole explanation of the preference of a second FadR$_{Sa}$ dimer to establish an interaction with the upstream (left) and not downstream (right) side of the pseudopalindromic site. Similarly, the absence of purine–pyrimidine and pyrimidine–purine bps on appropriate positions up- and downstream of the inverted repeat in the *Saci_1123* operator appears to explain why only a single dimer is bound. Based on the cocrystal structure, the G base of this G–C bp in the *fadR$_{Sa}$* operator can be assumed to be contacted by Gly48 that is crucial for DNA binding (Fig. 5b and Supplementary Figure 6a) with the dimer-of-dimer-complex architecture reflecting that of the architecture of the AB dimer and subunit E of the DE dimer interacting with a single-DNA duplex in the

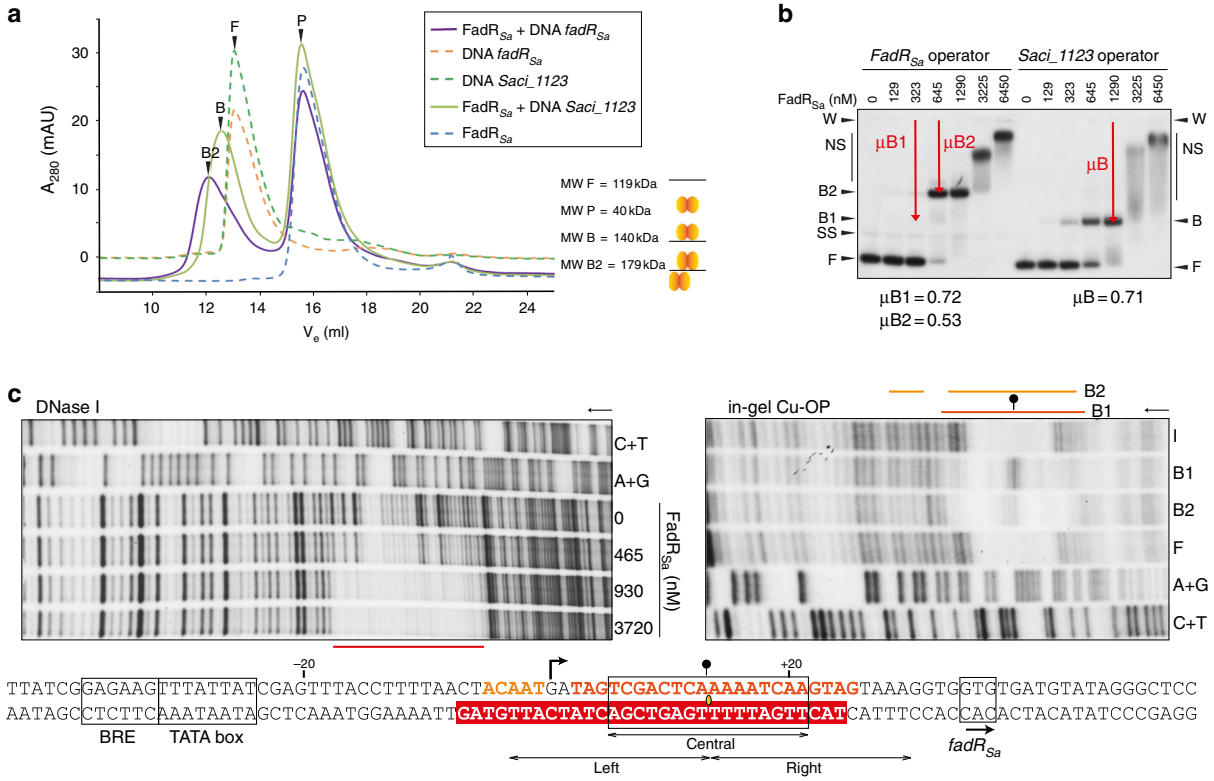


Fig. 4 $FadR_{Sa}$ interacts in different stoichiometric binding modes with different operators. **a** Size-exclusion chromatography experiment of $FadR_{Sa}$–DNA complexes with 45-bp DNA fragments. Totally, 0.2 nmol DNA and 8 nmol $FadR_{Sa}$ was used resulting in a 40:1 $FadR_{Sa}$:DNA molar ratio upon mixing. Apparent molecular weights (MWs) based on K_{av} calculations are mentioned for the peaks representing free DNA (F), $FadR_{Sa}$ protein (P) and $FadR_{Sa}$–DNA complexes B and B2. **b** Electrophoretic mobility shift assay (EMSA) monitoring interaction with probes representing the *fadR*$_{Sa}$ and *Saci_1123* operator. Molar protein concentrations are indicated and populations of single-stranded DNA (SS), free DNA (F), and $FadR_{Sa}$-bound DNA (B, B1, and B2) are indicated with an arrowhead. "W" corresponds to the well position. Relative mobilities μ are defined as the distance between the well position and complexed DNA divided by the distance between the well position and free DNA. **c** Autoradiographs of DNase I (bottom-strand labeled DNA) and "in-gel" Cu-phenantroline (Cu-OP) (top-strand labeled DNA) footprinting experiments analyzing $FadR_{Sa}$ binding to a probe representing peak 2 in the *Saci_1106*-*Saci_1107* intergenic region (*fadR*$_{Sa}$ operator). A + G and C + T denote purine- and pyrimidine-specific Maxam–Gilbert sequencing ladders, respectively. I denotes input DNA (taken from a protein-free lane in the EMSA), F denotes the population of free DNA, while B1 and B2 denote different $FadR_{Sa}$:DNA complex populations in accordance with the notation in the corresponding EMSA (second autoradiograph in Fig. 2d). The I and F samples generated a sequence-dependent cleavage profile. Protected zones are indicated with a horizontal line while a hyperreactivity site is pointed out with a ball-and-stick symbol. Below the autoradiographs, the nucleotide sequence is shown with a summary of the observed protection zones and hyperreactivity effect, the latter indicating $FadR_{Sa}$-induced DNA bending and being more pronounced for B1 than for B2. White letters in a red background represent the protection zone observed in DNase I footprinting, whereas orange letters represent the protection zone observed in Cu-OP footprinting. The predicted pseudopalindromic $FadR_{Sa}$ recognition site is boxed with indication of the center of dyad symmetry. The transcriptional start site, indicated with an arrow, is based on observations in ref. [37]

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_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 protein–DNA interactions in a combined direct and indirect readout[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_1123* operator (Fig. 5c), proving the involvement of this residue in a protein–protein interaction. This supports the notion that established $FadR_{Sa}$–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[26,30], the archaeal $FadR_{Sa}$ 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 $FadR_{Sa}$ regulon. To infer whether or not the observed genomic binding of $FadR_{Sa}$ leads to transcriptional regulation, a comparative transcriptomic analysis was performed for the *fadR*$_{Sa}$ deletion mutant *versus* the isogenic WT strain using an RNA-seq approach (Fig. 6a, Supplementary Note 2,

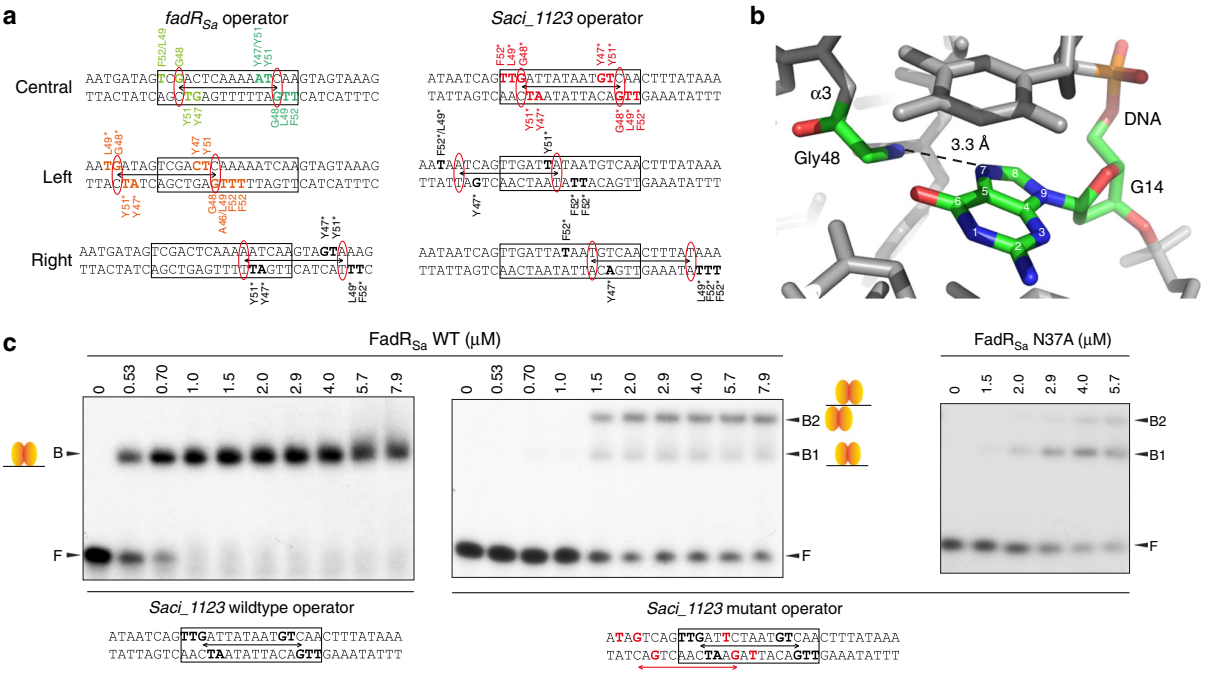


Fig. 5 Specific protein–DNA contacts determine stoichiometry of the interaction. **a** Map of observed and hypothesized interactions for the *fadR*$_{Sa}$ and *Saci_1123* operators. Colored base and amino acid residues without an asterisk represent contacts identified in the FadR$_{Sa}$–DNA cocrystal structure (using the same color code for the subunits as in panel (**a**)). Colored residues with an asterisk are hypothesized to be established in the natural operators. Black residues in bold are those that theoretically could support binding in case a GC (AT) or CG (TA) would have been present on the crucial Gly48-interacting positions, which are indicated with red ovals. **b** Zoom of the Gly48–G14 interaction in the interface between subunit E and the X–Y DNA duplex in the FadR$_{Sa}$–DNA cocrystal structure. **c** EMSAs with 45-bp DNA fragments representing a wildtype or mutated variant of the *Saci_1123* operator region, performed with wildtype or N37A mutant FadR$_{Sa}$. Molar protein concentrations are shown above the autoradiograph, whereas populations of free DNA (F) and FadR$_{Sa}$-bound DNA (B, B1, and B2) are indicated with an arrowhead, as well as with schematical representations of the binding stoichiometry

Supplementary Figure 10, Supplementary Table 4 and Supplementary Data 2). The deletion of *fadR*$_{Sa}$ 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 *Saci_1103-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 FadR$_{Sa}$ is a local repressor of the entire *Saci_1103-Saci_1126* gene cluster, which is predicted to harbor genes for a complete β-oxidation pathway (Fig. 6a).

All but one of these FadR$_{Sa}$ binding sites are located at too large distances (>130 bp) from their corresponding promoters to hypothesize a classical repression mechanism that involves direct interaction with the components of the basal transcription machinery. As an exception, the *fadR*$_{Sa}$ control region harbors a binding site just downstream of the transcription start site. For this target, it is shown that FadR$_{Sa}$ binding stimulates the interaction with basal transcription factors TATA binding protein (TBP) and transcription factor B (TFB) (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 *ΔfadR*$_{Sa}$ 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 FadR$_{Sa}$ binding sites in the genomic regions surrounding the downregulated genes, which are characterized by relative high *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 *Saci_1103-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 *Saci_1103-Saci_1126* gene cluster strongly suggests that FadR$_{Sa}$ 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 (*Saci_1105* and *Saci_1116*) led to a phenotype lacking the ability to perform tributyrin hydrolysis[29], we performed a similar phenotypic assay with the *fadR*$_{Sa}$ deletion mutant (Supplementary Figure 13a). Despite the higher expression levels of both esterase genes in the *fadR*$_{Sa}$ 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 *S. acidocaldarius* in a liquid medium containing hexanoate as a sole carbon and energy source, the *fadR*$_{Sa}$ deletion mutant displayed a significantly higher growth rate in exponential growth phase with respect to the isogenic WT strain (doubling times T_ds 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 *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

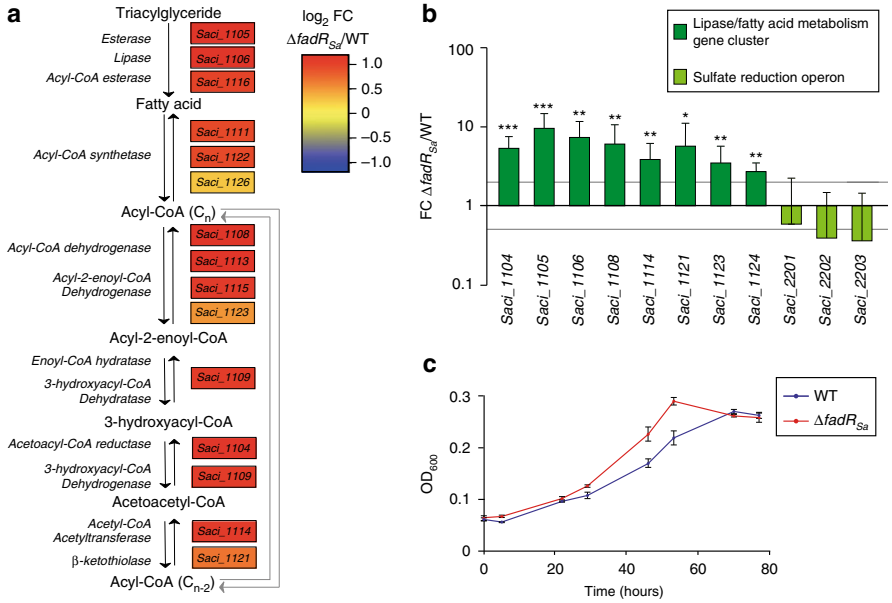


Fig. 6 $FadR_{Sa}$ exerts a repression on the transcription of the catabolic fatty acid metabolism/lipase gene cluster. **a** Reconstruction of the lipid degradation and fatty acid β-oxidation pathway with indication of the putative functions of genes in the *Saci_1103-Saci_1126* gene cluster predicted to be involved in this metabolic pathway. This is based on genomic annotations. For each of these genes, differential expression in *S. acidocaldarius* MW001 Δ$fadR_{Sa}$, as measured by RNA-seq analysis, is shown in a heat map format expressed as the log_2 of the fold change expression in the Δ$fadR_{Sa}$ versus its isogenic wildtype strain. **b** Relative gene expression as determined by qRT-PCR for a subset of the genes of the *Saci_1103-Saci_1126* gene cluster and for genes belonging to the sulfate reduction operon. Gray lines represent a fold change of 2 and 0.5, respectively. Error bars represent biological variation for triplicates (standard deviations). An asterisk indicates a *p*-value between 0.05 and 0.01, a double asterisk between 0.01 and 0.001, and a triple asterisk smaller than 0.001, as determined in a statistical *t* test. **c** Growth curves of MW001 and MW001 Δ$fadR_{Sa}$ strains in Brock medium containing 2 mM hexanoate as sole carbon source. Values are averages of four biological replicates with error bars representing standard deviations. Doubling times (T_ds) were calculated by modeling the exponential section of the growth curves. Representative curves are shown for multiple independently performed experiments

Saci_1103-Saci_1126 gene cluster. Furthermore, the $FadR_{Sa}$ 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 $FadR_{Sa}$-mediated derepression given the relative short-chain length of these acyl-CoA molecules (see below, "$FadR_{Sa}$–ligand interactions").

$FadR_{Sa}$–ligand interactions. Next, we prepared $FadR_{Sa}$ crystals in the presence of lauroyl–CoA (C12:0-CoA) and solved the $FadR_{Sa}$–lauroyl–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 $FadR_{Sa}$ as compared to the characterized bacterial FadR proteins[24–26] (Fig. 7a). In contrast to the ligand entering the pocket from within the dimer interface as in bacterial FadR regulators, in the $FadR_{Sa}$ 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 $FadR_{Sa}$ 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 $FadR_{Sa}$ while it is bent in $FadR_{Bs}$ (Supplementary Figure 14b).

Upon zooming into the ligand-binding pocket, a large number of specific lauroyl–CoA–$FadR_{Sa}$ 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, $FadR_{Sa}$ does not contain a hydrophilic patch in the acyl-binding pocket similarly to the *Bacillus* counterparts in which it affects ligand-binding specificity[25,26].

EMSAs demonstrated that acyl-CoA molecules, but not acetyl-CoA, CoA and free fatty acids, strongly abrogate $FadR_{Sa}$–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

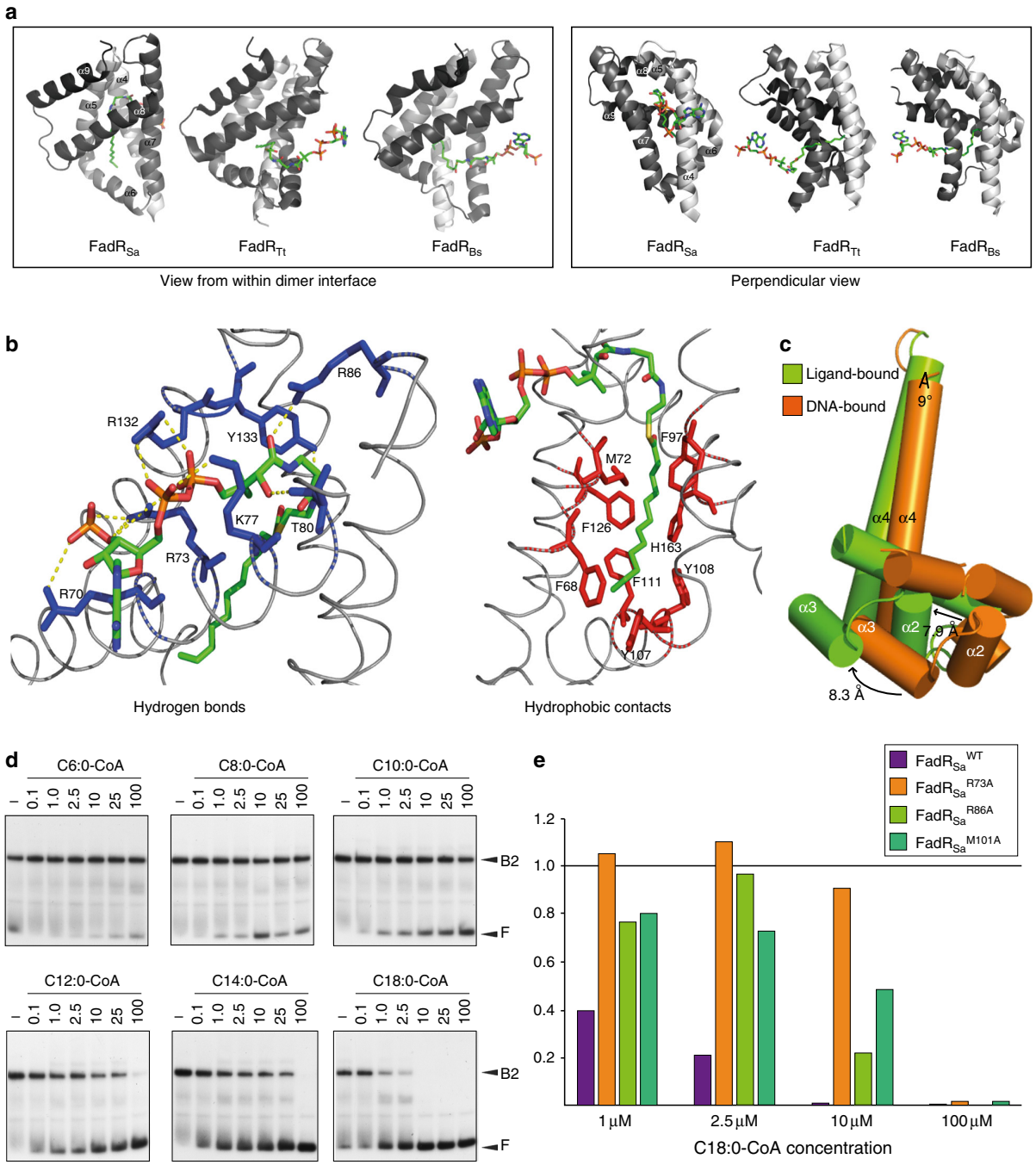


Fig. 7 Acyl-CoA disrupts $FadR_{Sa}$–DNA complexes proportional to acyl chain length. **a** Structural comparison of ligand-binding modes of lauroyl-CoA-bound $FadR_{Sa}$ (PDB: 6EL2), lauroyl-CoA-bound $FadR_{Tt}$ (PDB: 3ANG)[24], and lauroyl-CoA-bound $FadR_{Bs}$ (PDB: 3WHB)[25]. Only the C-terminal domains are shown in cartoon representation, with α-helices α4–α9 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 $FadR_{Sa}$ 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 α1–α4 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 $FadR_{Sa}$–DNA interaction using a 154-bp probe representing the *Saci_1106* binding site. Acyl-CoA concentrations are shown above each autoradiograph in µM. Populations of free DNA and complexed DNA are indicated with F and B2, respectively. The intermediate complex B1 is hardly formed due to the cooperativity of the interaction. **e** Graphical representation of ligand response measured in EMSAs performed with ligand-binding mutants $FadR_{Sa}^{R73A}$, $FadR_{Sa}^{R86A}$, and $FadR_{Sa}^{M101A}$ (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 $FadR_{Sa}$–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 $FadR_{Sa}$, 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 $FadR_{Sa}$ 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 $FadR_{Sa}$–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 $FadR_{Bh}$[26].

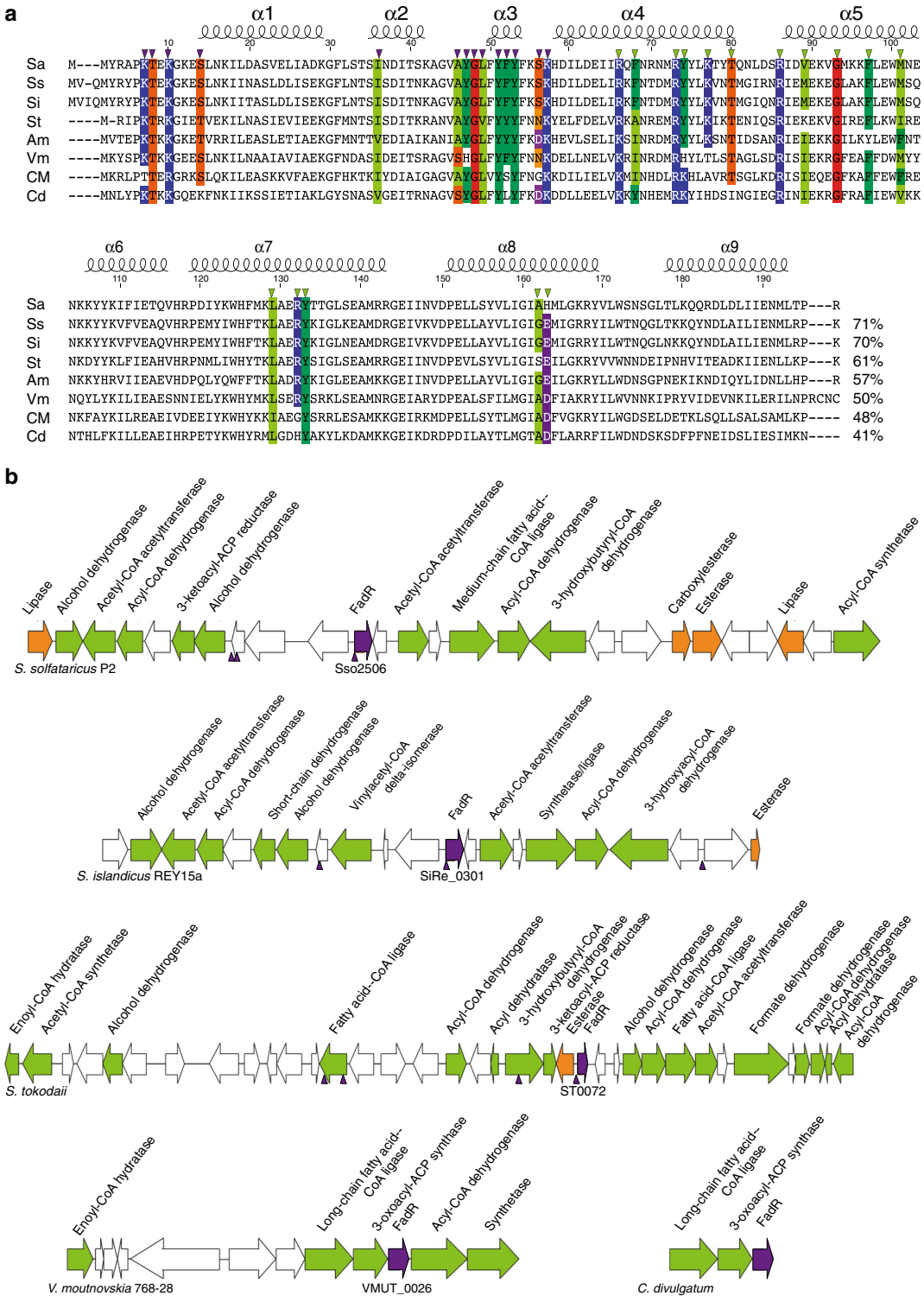
Occurence of FadR in archaea. $FadR_{Sa}$ is not restricted to *S. acidocaldarius* but is also represented in all other *Sulfolobus* species, in other Crenarchaeota belonging to *Sulfolobales* (*Acidianus* spp.) and *Thermoproteales* (*Cuniculiplasma divulgatum*), and also in species belonging to Euryarchaeota (*Thermoplasmatales*) and in the recently discovered Marsarchaeota[33] (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 these proteins are $FadR_{Sa}$ orthologs with similar functional characteristics (Fig. 8a). Moreover, as is the case for $FadR_{Sa}$, several of these homologs are encoded in genomic environments abundant in genes coding for enzymes involved in fatty acid metabolism or for enzymes with lipase functions (Fig. 8b). Although gene synteny is not strictly conserved, the extent of some of these gene clusters, especially in other *Sulfolobus* species, suggests the potential existence of similar FadR-mediated acyl-CoA responsive repression. This hypothesis is supported by the prediction of putative FadR binding sites in the neighborhood of *fadR* promoters and at distant locations, either in intergenic regions or in ORFs, for the gene clusters in other *Sulfolobus* species (Fig. 8b).

Discussion

It is well-established that archaea harbor typical bacterial-like transcription regulators[34,35], which are proposed to result from shared ancestry as well as from extensive horizontal gene transfers, especially from bacteria to archaea[36]. $FadR_{Sa}$, of which we show that it displays structural similarities with bacterial TetR-like FadR regulators, is an archaeal member of the widespread prokaryotic TetR family. Despite these similarities, there are pronounced differences between the archaeal FadR regulator and the bacterial counterparts which point to a complete absence of shared ancestry. For example, the acyl-CoA binding function of $FadR_{Sa}$ appears to have arisen through convergent evolution with respect to the bacterial regulators. In contrast to $FadR_{Bs}$ and $FadR_{Tt}$, in which the crystal structure revealed medium-chain lauroyl–CoA in one of the two dimeric subunits[24,25], we observed the presence of a short-chain acyl-CoA in the native $FadR_{Sa}$ structure. This difference likely reflects different acyl-CoA-binding specificities and might be explained by the absence of a hydrophilic patch in the part of the ligand-binding pocket that surrounds the first 10–12 carbon atoms of the acyl chain, as observed for the *Bacillus* counterparts[25,26]. As a consequence, $FadR_{Sa}$ has a different ligand specificity, which is expected to have consequences for the biological function and to reflect different biological roles of fatty acids for cellular physiology in bacteria 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 archaea[10].

$FadR_{Sa}$ uses two distinct DNA-binding modes operator-dependently with a Gly–guanine interaction being the crucial determinant and that dimer-of-dimer complex formation with the quasi-identical high-affinity $fadR_{Sa}$ and *Saci_1106* operators occurs in a cooperative manner with the transitional formation of a dimer-bound complex in which $FadR_{Sa}$-mediated DNA bending is more pronounced than in the dimer-of-dimer complex. By binding to a total of only four binding sites in the *Saci_1103-Saci_1126* gene cluster, $FadR_{Sa}$ is capable of repressing transcription of the entire 23-gene cluster containing at least 17 transcription units[37]. 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 repressors[38]. $FadR_{Sa}$ thus employs a nonparadigmatic repression mechanism that could be hypothesized to be dependent on long-range interactions. The observation of cocrystallization of $FadR_{Sa}$ with two individual DNA duplexes (Fig. 3a), as was also observed for the bacterial TetR member CgmR in *Corynebacterium glutamicum*[39], further supports the possibility that the regulator is capable of co-associating with different DNA segments.

The finding that $FadR_{Sa}$ 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 lipids[29]. Fatty acid oxidation adds to the chemoorganotrophic capabilities of *Sulfolobus* spp. that appear more important than the originally described chemolithotrophic sulfur-oxidizing metabolism[40]. 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



concentrations[9,10]. Possibly, acyl-CoA-responsive $FadR_{Sa}$-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 $FadR_{Sa}$ orthologs in other thermophilic archaea, such as *Thermoplasmatales* and Marsarchaeota, it

Fig. 8 $FadR_{Sa}$ homologs are found in other archaeal genomes. **a** Multiple sequence alignment, prepared with t-coffee[68], of different archaeal $FadR_{Sa}$ homologs identified through BLAST analysis. Sa = $FadR_{Sa}$ *Sulfolobus acidocaldarius*, Ss = AAK42639.1 *Sulfolobus solfataricus* P2, Si = PVU77113.1 *Sulfolobus islandicus*, St = WP_010978011.1 *Sulfolobus tokodaii*, Am = ARM75525.1 *Acidianus manzaensis*, Vm = WP_013603407.1 *Vulcanisaeta moutnovskia*, CM = PSN82693.1 Candidatus Marsarchaeota G1 archaeon OSP_D, Cd = WP_077076373.1 *Cuniculiplasma divulgatum*. Indication of secondary structure elements and of position numbers is with respect to the $FadR_{Sa}$ sequence. Colors indicate conservation of amino acid residues that are involved in DNA (purple triangles) or ligand (green triangles) interactions, based on $FadR_{Sa}$ cocrystal structures. Sequence identities with $FadR_{Sa}$ are mentioned behind the sequence alignment. **b** Schematic depiction of gene organization in the genomic neighborhood of selected $FadR_{Sa}$ homologs. Color coding is as follows: green = genes encoding enzymes involved in fatty acid metabolism, orange = genes encoding lipases/esterases, purple = genes encoding transcriptional regulators. Gene numbers are mentioned for the $FadR_{Sa}$ homologs. The *Sulfolobus spp*. gene cluster sequences were subjected to a binding motif prediction using RSAT[69] with a position weight matrix based on $FadR_{Sa}$ binding site sequences. Predicted binding sites are indicated with a blue triangle

can be assumed that this type of transcriptional regulation is not restricted to *Sulfolobus* spp.

Methods

Microbial strains and growth conditions. *S. acidocaldarius* strains MW001 and its derivatives were cultivated at 75 °C in Brock basal salts medium[40] supplemented with 0.2 (w/v)% sucrose, 0.2 (w/v)% NZamine and 10 μg ml^{-1} uracil. For ChIP experiments, a 200 ml *S. acidocaldarius* DSM639 culture was grown at 80 °C in Brock basal salts medium supplemented with 0.1% 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, precultures of *S. acidocaldarius* MW001 and its derivatives were first grown in Brock medium supplemented with 0.2 (w/v)% NZamine followed by multiple transfers to Brock medium supplemented with 2 mM butyrate or hexanoate. Cultivations of the third and fourth generation were considered to lack residual NZamine and to be representative for growth on fatty acids as sole carbon source. For growth on plates, Brock medium was solidified by the addition of 0.6 (w/v)% gelrite, 10 mM $MgCl_2$ and 3 mM $CaCl_2$. To detect the hydrolysis of tributyrin, cells were inoculated on tributyrin-containing plates composed of 1.4% gelrite dissolved in Brock basal salt medium supplemented with 20 mM $MgCl_2 \times 6H_2O$, 6 mM $CaCl_2$, 0.2% (w/v) NZamine, 0.4% (w/v) dextrine and 1% (v/v) tributyrin[29]. Cavities were made into the plates using a 5 ml pipette tip followed by addition of 20 μl of cell culture grown up to an optical density measured at 600 nm (OD_{600}) of 0.7. The plates were incubated up to 7 days.

E. coli strains DH5α and Rosetta DE3 were used for the propagation of plasmid DNA constructs and heterologous protein overexpression, respectively, and were grown at 37 °C in lysogeny broth (LB) medium supplemented with 50 μg ml^{-1} ampicillin (DH5α) or with 50 μg ml^{-1} ampicillin and 34 μg ml^{-1} chloramphenicol (Rosetta (DE3)). An overview of all strains used in this work is given in Supplementary Table 5.

A markerless in-frame $fadR_{Sa}$ deletion strain (*S. acidocaldarius* Δ$fadR_{Sa}$) was generated of the uracil auxotrophic *S. acidocaldarius* strain (MW001) using the classical "pop-in pop-out" method with a suicide disruption vector containing up- and downstream flanking regions and the gene-of-interest region besides the *pyrEF* selection marker genes (pSVA406xΔ*fadR*)[28]. An overview of all plasmid vectors and oligonucleotide sequences is given in Supplementary Table 6 and Supplementary Data 4, respectively. Successful deletion of the gene was confirmed by PCR analysis and sequencing of genomic DNA (gDNA) (Supplementary Figure 2).

Protein expression and purification. The $fadR_{Sa}$ coding region was amplified by PCR from *S. acidocaldarius* gDNA and cloned into a pET45b expression vector using BamHI and HindIII restriction sites (pET45b*x*$fadR_{Sa}$) resulting in an N-terminally His-tagged construct. Site-directed mutagenesis was performed with the overlap PCR mutagenesis approach[41] using pET45b*x*$fadR_{Sa}$ as a template and complementary mutagenic primers to remove an NdeI restriction site in the $fadR_{Sa}$ ORF with a silent mutation. This enabled cloning in the NdeI/XhoI sites of pET24a yielding a C-terminally His-tagged construct (pET24a*x*$fadR_{Sa}$*Ndenull*). For the construction of N37A, Y47A, G48A, Y51A, Y53A, R73A, R86A, and M101A variants of $FadR_{Sa}$, site-directed mutagenesis was performed in an identical approach using pET24a*x*$fadR_{Sa}$*Ndenull* as a template.

Heterologous overexpression of the recombinant $FadR_{Sa}$ proteins was accomplished in *E. coli* Rosetta (DE3) by growing a culture until reaching an OD_{600} between 0.6 and 0.7, incubating the cells on ice during 30 min and inducing gene expression by the addition of 0.4 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Subsequently, the culture was further incubated at 37 °C during 20 h, pelleted by centrifugation, resuspended in binding buffer (20 mM sodium phosphate pH 7.4, 0.5 M NaCl, 20 mM imidazole) and lysed by sonication. Finally, the cell extract was subjected to heat treatment (80 °C during 30 min) followed by centrifugation at 23,000×*g* during 45 min. Purification of His-tagged FadRSa was performed with immobilized metal ion affinity chromatography by applying the supernatant to a His GraviTrap system (GE Healthcare) equilibrated with binding buffer (20 mM sodium phosphate (pH 7.4), 0.5 M NaCl and 20 mM imidazole). The column was washed with binding buffer where after the protein was eluted with 3 × 1 ml of 20 mM sodium phosphate (pH 7.4), 0.5 M NaCl and 500 mM imidazole. N- and C-terminally His-tagged proteins behave similar in DNA-binding assays.

For the preparation of selenomethionine (SeMet)-substituted $FadR_{Sa}$ protein, cells were allowed to grow overnight at 37 °C in 2 ml 2 × YT medium containing 50 μg ml^{-1} ampicillin. The overnight culture was used to inoculate 50 ml M9 minimal medium (50 mM Na_2HPO_4, 3 g l^{-1} KH_2PO_4, 0.5 g l^{-1} NaCl, 1 g l^{-1} NH_4Cl) and growth was continued at 37 °C until an OD_{600} of 0.5 was reached. Thereafter, 50 ml of the culture was added to 700 ml M9 medium supplemented with 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 0.4% glucose, and 50 μg ml^{-1} ampicillin, and continue growth at 37 °C until an OD_{600} of 0.4–0.6. Following l-amino acids were then added: Lys, Phe, and Tyr (100 mg l^{-1}), Leu, Ile, and Thr (50 mg l^{-1}) and l-selenomethionine (Acros Organics; final concentration of 60 mg l^{-1}). At an OD_{600} of 0.6–0.8, protein expression was induced by adding 0.5 mM IPTG (isopropyl β-D-1-thiogalactopyranodside) followed by overnight incubation. SeMet-substituted protein was purified according to the same procedure as described above using a pET45b*x*$fadR_{Sa}$ harboring strain. All $FadR_{Sa}$ protein preparations were essentially pure, as judged by SDS-polyacrylamide gel electrophoresis (PAGE) and by SEC (Supplementary Figure 1).

The ORFs of TBP (*Saci_1336*) and TFB1 (*Saci_0866*) were PCR-amplified, digested and ligated into pET30a (Novagen) using the restriction exonucleases NdeI and XhoI and transformed into *E. coli* Rosetta 2 (DE3). Cells harboring pET30a*xtbp* were grown until reaching an OD_{600} of approximately 0.6, followed by an induction with 1 mM IPTG and further growth during 4 h at 37 °C. Cells 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* during 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-HCl pH 7.5, 300 mM NaCl. Similarly, *E. coli* Rosetta 2 (DE3) cells expressing C-terminally hexahistidine-tagged TFB1 were grown in medium supplemented with 10 mM $MgCl_2$ and 100 μM $ZnSO_4$, induced at an OD_{600} of approximately 0.6 by adding 0.2 mM IPTG and followed by further growth at 23 °C overnight. Cells were harvested by centrifugation and subsequently resuspended in modified N-buffer (25 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 100 μM $ZnSO_4$, 1 mM tris(2-carboxyethyl)phosphine (TCEP)) supplemented with 1 M NaCl and protease inhibitor. After cell disruption via French pressure cell (three passages at 12,000 psi), ultracentrifugation (30,000×*g* during 45 min) and heat treatment at 75 °C for 15 min the cleared lysate was applied to a Ni-TED column (Macherey and Nagel) for affinity purification. Modified N-buffer containing 300 mM NaCl was used for equilibration and washing, the bound target protein was eluted with the same buffer containing 250 mM imidazole. Pure TFB1 protein was finally obtained by performing SEC with a HiLoad superdex 26/60 200 prep grade gel filtration column (GE Healthcare). TBP and TFB1 were used in EMSAs (Supplementary Figure 12).

Crystallization and data collection. Crystallization of SeMet-derivated $FadR_{Sa}$ was performed at 20 °C using the hanging-drop vapor diffusion method by mixing equal volumes of protein solution (3 mg ml^{-1}) and reservoir solution consisting of 20% (w/v) PEG3350, 0.2 M sodium nitrate and 0.1 M Bis–Tris propane, pH 8.5. Appropriately sized crystals were obtained after 6–8 weeks. The crystals belong to space group $P2_1$, with unit-cell parameters $a = 41.9$, $b = 98.7$, $c = 55.9$ Å, and $\beta = 106.4°$, and two molecules per asymmetric unit, giving a Matthews coefficient of 2.51 $Å^3$ Da^{-1} and 51% solvent content. The complexes of $FadR_{Sa}$:DNA and $FadR_{Sa}$:lauroyl–CoA were obtained by cocrystallization of $FadR_{Sa}$ with a 21-bp duplex DNA and 1 mM lauroyl–CoA (lithium salt), respectively. Prior to data collection, crystals were soaked in a cryo-solution containing 20% glycerol, 10% (w/v) PEG3350, 0.1 M sodium nitrate and 0.05 M Bis–Tris propane, pH 8.5 followed by immediate flash-cooling in liquid nitrogen.

High-resolution X-ray data (Table 1) were collected at 100 K at European Synchroton Radiation Facility (ESRF) beamlines ID23-1 (SeMet-substituted $FadR_{Sa}$) and ID29 ($FadR_{Sa}$:DNA and $FadR_{Sa}$:lauroyl–CoA).

Structure determination and refinement. The structure of $FadR_{Sa}$ was determined using the SAD method with selenomethionine-substituted protein. Diffraction data were processed and scaled using the *XDS* program package[42]. A set of

5% of the reflections was set aside and used to calculate the quality factor R_{free}[43]. The structure was solved using AutoSol in PHENIX[44]. Refinement was performed with PHENIX altered with manual rebuilding in O[45]. The structure was refined to $R_{fac} = 20.3\%$ and $R_{free} = 25.3\%$, respectively (Table 1). The structure of the dsDNA and lauroyl–CoA complexes were solved by molecular replacement with Phaser[46], using the SeMet-subsituted $FadR_{Sa}$ structure as a model. All structures were evaluated using wwPDB Validation Server[47]. Refinement statistics are presented in Table 1. The coordinates and structure factors have been deposited in the PDB database with accession codes 5MWR, 6EN8, and 6EL2. PDBsum[48] was used to identify protein–DNA and protein–ligand interactions, supplemented with a manual inspection employing PyMOL[49]. All figures displaying protein structures were prepared with PyMOL[49].

Chromatin immunoprecipitation. ChIP was performed by growing *S. acidocaldarius* DSM639 to early exponential growth phase (an OD_{600} between 0.2 and 0.3) and adding formaldehyde to the culture to a final concentration of 1%[50,51]. After a 5-min incubation, glycine was added to a final concentration of 125 mM. Fixed cells were harvested by centrifugation at 8000×*g* during 10 min and the pellet was resuspended in 3 ml IP buffer (50 mM Hepes-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. One aliquot was incubated with anti-$FadR_{Sa}$ antibody coated M-280 Sheep Anti-Rabbit Dynabeads (Invitrogen) and the other was incubated with pre-immune serum coated Dynabeads, which served as a non-specific binding control (mock control). The bead-antibody complexes were prepared by mixing 80 µl Dynabeads with either $FadR_{Sa}$-specific antibodies (produced by immunizing a rabbit with purified recombinant $FadR_{Sa}$ (Innovagen)) or rabbit pre-immune serum. Precipitation reactions were performed according to manufacturer's instruction. After overnight incubation at 4 °C, the Dynabeads were collected and the captured DNA was eluted and purified by using the iPURE DNA extraction kit (Diagenode) according to the manufacturer's instruction.

DNA purified from ChIP, input and mock samples were sequenced (1 × 51 bp) by a Miseq sequencer (Illumina) at ScilifeLab, Stockholm, Sweden. Sequence reads were mapped to the *S. acidocaldarius* DSM639 genome (NC_007181.1) with Burrows–Wheeler Aligner (BWA 0.7.10)[52] using default settings and MACS2 (2.1.0)[53] was employed for peak calling. To generate sufficient sequencing reads, reads of two mock samples were combined before performing the analysis. The ChIP-seq experiment was done in biological duplicate and only peaks called in both experiments were retained; this was followed by a manual curation. Finally, ChIP-seq results were visualized by IGV version 2.3.59[54]. DNA sequences of enriched regions were extracted by BEDTools' getfasta function[55] and subjected to a binding motif search with MEMEsuite (4.10.0)[56]. The FIMO tool of MEMEsuite was used for the prediction of binding motifs in other genomic regions.

For ChIP-qPCR, 20-ml cultures of *S. acidocaldarius* MW001 and *S. acidocaldarius* MW001Δ$fadR_{Sa}$ were crosslinked and harvested in mid exponential growth phase (OD_{600} of about 0.4) and ChIP was performed as described above. Primers were designed with Primer3 Plus software[57] (Supplementary Data 5) and were chosen to amplify fragments around the ChIP-seq peak summit regions and with a length between 150 and 200 bp. qPCR was performed with a My-iQ™ Single color Real-time PCR system (Bio-Rad) and GoTaq qPCR Master Mix (Promega) was done with thermal cycling conditions: 10 min at 94 °C and 40 cycles of 30 s at 94 °C and 30 s at 60 °C. Fold-enrichment calculations were performed with the $2^{-\Delta\Delta Ct}$ method[58] using an irrelevant genomic region (the ORF of *Saci_1336*) as a nonbinding reference. Cultures were assayed in biological duplicate and the *S. acidocaldarius* Δ$fadR_{Sa}$ strain was used as mock experiment.

Electrophoretic mobility shift and footprinting assays. ^{32}P-labeled DNA was prepared by 5′-end-labeling of oligonucleotides using [γ-^{32}P]-ATP (Perkin Elmer) and T_4 polynucleotide kinase (Thermo Scientific). Each of these labeled oligonucleotides were then used together with a non-labeled oligonucleotide (Supplementary Data 4) in a PCR reaction with *S. acidocaldarius* gDNA as a template or, in case of 45-bp probes, in a hybridization reaction with the nonlabeled reverse complementary oligonucleotide. Labeled DNA fragments were subsequently purified by polyacrylamide gel electrophoresis. EMSAs were performed[59] with approximately 0.1 nM ^{32}P-labeled DNA probe and an excess of nonspecific competitor DNA. Dimethylsulfoxide was used as a solvent to dissolve acyl-CoA, but did not affect the protein–DNA interaction (Supplementary Figure 16). Binding reactions were prepared in binding buffer (20 mM Tris-HCl (pH 8.0), 1 mM $MgCl_2$, 0.1 mM dithiothreitol (DTT), 12.5% glycerol, 50 mM NaCl, 0.4 mM EDTA) and allowed to equilibrate at 37 °C prior to electrophoresis on 6% acrylamide gels in TEB buffer (89 mM Tris, 2.5 mM EDTA, and 89 mM boric acid).

DNase I footprinting was performed by the method of Galas and Schmitz[60] in the same binding buffer as used for EMSA and with Maxam–Gilbert treated samples as sequencing ladders[61]. "In gel" Cu–phenantroline (OP) chemical footprinting experiments, enabling to analyze distinct populations of nucleoprotein complexes separately, were performed by performing EMSA as described above and immersing an entire EMSA acrylamide gel in 200 ml of 10 mM Tris (pH 8.0)[59]. To initiate chemical cleavage reactions, 20 ml of solution A (2 mM OP, 0.45 mM $CuSO_4$) was added followed by 20 ml of solution B (58 mM mercaptopropionic acid). After 15 min, reactions were quenched by the addition of 20 ml 30 mM neocuprine hydrate. This mixture was allowed to equilibrate during 5 min, after which the gel was thoroughly rinsed with distilled H_2O and exposed to an autoradiograph film. Different DNA populations were excised from the gel, recovered by precipitation and analyzed by denaturing acrylamide gel electrophoresis with Maxam–Gilbert treated samples as sequencing ladders[61].

Size-exclusion chromatography. For stoichiometry experiments of $FadR_{Sa}$–DNA complexes, SEC was performed on a Superdex 200 Increase10/30 GL column with an ÄKTA FPLC system (GE Healthcare Life Sciences). A total of 4–40 nM $FadR_{Sa}$ protein was mixed with 0.2–1 nM 45-bp DNA fragments encompassing the *fadR*$_{Sa}$ operator or *Saci_1123* operator, respectively, which were prepared before by hybridization. After an incubation of the reaction mixtures in 20 mM Na_2HPO_4 (pH 7.4), 150 mM NaCl during 25 min at 37 °C, they 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), con-albumin (75 kDa), and ferritin (440 kDa), all from a Gel Filtration Calibration Kit (GE Healthcare Life Sciences).

RNA-sequencing analysis. Total RNA was prepared from duplicate MW001 and MW001Δ*fadR*$_{Sa}$ cultures in early exponential growth phase at an OD_{600} between 0.2 and 0.3 using a miRNeasy Mini Kit (Qiagen). Libraries were prepared with a TruSeq Stranded mRNA Library Prep Kit (Illumina). Paired-end (2 × 125 bp) RNA sequencing was performed using a Hiseq2500 system (Illumina) at SciLifeLab, Stockholm, Sweden. Sequence reads were first trimmed to remove sequencing adapters by cutadapt 1.9.1[62] and reads shorter than 20 nt were discarded. Processed reads were then mapped to the *S. acidocaldarius* DSM639 genome (NC_007181.1)[63] with Tophat 2.0.12[64]. For each gene, the FPKM was calculated with Cufflinks 2.2.1[65]. Finally, read counts were obtained by the featureCounts function in the Subread package 1.5.0[66] and only genes having at least one count in all samples were used for differential gene expression analysis with DESeq2[67].

Quantitative RT-PCR. RNA was extracted at an OD_{600} between 0.2 and 0.3 by stabilization with RNA Protect Reagens (Qiagen) and by using an SV Total RNA Isolation System (Promega). Residual gDNA 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[57]. qRT-PCR analysis was performed in a Bio-Rad iCycler using SYBR Green Master Mix (Bio-Rad) for amplification and detection. Each reaction mixture contained approximately 10 ng of template and 200 nM of each primer in a total volume of 25 µl. The temperature program was as follow: 10 min at 94 °C and 40 cycles of 30 s at 94 °C, 30 s at 60 °C[50]. Relative expression ratios were calculated using the delta–delta C_t method[58] for biological triplicates and by normalizing with respect to the *tbp* reference gene. Data were statistically analyzed by performing a *t* test with the software package Prism 6.0 (GraphPad).

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 5MWR, [has been superseded with 6EL2] (native $FadR_{Sa}$ structure), 6EL2 (lauroyl–CoA-bound $FadR_{Sa}$ structure) and 6EN8 (DNA-bound $FadR_{Sa}$ structure). All raw data for the ChIP-seq and RNA-seq studies presented in this work have been deposited in the Gene Expression Omnibus (GEO) databank with accession codes GSE108039 and GSE108018, respectively.

Received: 8 January 2018 Accepted: 12 March 2019
Published online: 04 April 2019

References

1. Woese, C. R. & Fox, G. E. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc. Natl Acad. Sci. USA* **74**, 5088–5090 (1977).
2. De Rosa, M., Gambacorta, A. & Gliozzi, A. Structure, biosynthesis, and physicochemical properties of archaebacterial lipids. *Microbiol. Rev.* **50**, 70–80 (1986).
3. Kates, M., Wassef, M. K. & Kushner, D. J. Radioisotopic studies on the biosynthesis of the glyceryl diether lipids of *Halobacterium cutirubrum*. *Can. J. Biochem.* **46**, 971–977 (1968).

4. Langworthy, T. A., Mayberry, W. R. & Smith, P. F. Long-chain glycerol diether and polyol dialkyl glycerol triether lipids of *Sulfolobus acidocaldarius*. *J. Bacteriol.* **119**, 106–116 (1974).
5. Tornabene, T. G. et al. Phytanyl-glycerol ethers and squalenes in the archaebacterium *Methanobacterium thermoautotrophicum*. *J. Mol. Evol.* **11**, 259–266 (1978).
6. Carballeira, N. M. et al. Unusual fatty acid compositions of the hyperthermophilic archaeon *Pyrococcus furiosus* and the bacterium *Thermotoga maritima*. *J. Bacteriol.* **179**, 2766–2768 (1997).
7. Nishihara, M., Nagahama, S., Ohga, M. & Koga, Y. Straight-chain fatty alcohols in the hyperthermophilic archaeon *Pyrococcus furiosus*. *Extremophiles* **4**, 275–277 (2000).
8. Gattinger, A., Schloter, M. & Munch, J. C. Phospholipid etherlipid and phospholipid fatty acid fingerprints in selected euryarchaeotal monocultures for taxonomic profiling. *FEMS Microbiol. Lett.* **213**, 133–139 (2002).
9. Hamerly, T. et al. Characterization of fatty acids in Crenarchaeota by GC-MS and NMR. *Archaea* **2015**, 472726 (2015).
10. Dibrova, D. V., Galperin, M. Y. & Mulkidjanian, A. Y. Phylogenomic reconstruction of archaeal fatty acid metabolism. *Environ. Microbiol.* **16**, 907–918 (2014).
11. Corcelli, A. et al. Role of palmitic acid on the isolation and properties of halorhodopsin. *Biochim. Biophys. Acta* **1281**, 173–181 (1996).
12. Colella, M., Lobasso, S., Babudri, F. & Corcelli, A. Palmitic acid is associated with halorhodopsin as a free fatty acid. Radiolabeling of halorhodopsin with ^{3}H-palmitic acid and chemical analysis of the reaction products of purified halorhodopsin with thiols and $NaBH_4$. *Biochim. Biophys. Acta* **1370**, 273–279 (1998).
13. Lombard, J., López-García, P. & Moreira, D. An ACP-independent fatty acid synthesis pathway in archaea: implications for the origin of phospholipids. *Mol. Biol. Evol.* **29**, 3261–3265 (2012).
14. Villanueva, L., Schouten, S. & Damsté, J. S. S. Phylogenomic analysis of lipid biosynthetic genes of Archaea shed light on the 'lipid divide'. *Environ. Microbiol.* **19**, 54–69 (2017).
15. Fujita, Y., Matsuoka, H. & Hirooka, K. Regulation of fatty acid metabolism in bacteria. *Mol. Microbiol.* **66**, 829–839 (2007).
16. 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).
17. My, L., Ghandour Achkar, N., Viala, J. P. & Bouveret, E. Reassessment of the genetic regulation of fatty acid synthesis in *Escherichia coli*: global positive control by the dual functional regulator FadR. *J. Bacteriol.* **197**, 1862–1872 (2015).
18. Zhang, Y.-M., Marrakchi, H. & Rock, C. O. The FabR (YijC) 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. Schujman, G. E., Paoletti, L., Grossman, A. D. & de Mendoza, D. FapR, 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* FadR controls fatty acid degradation. *Microbiology* **157**, 1589–1601 (2011).
25. Fujihashi, M. et al. Structural characterization of a ligand-bound form of *Bacillus subtilis* FadR 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 FadR. *Nucleic Acids Res.* **45**, 4244–4254 (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. Zweerink, 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 QacR. *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. Kyrpides, 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. Itou, H., Watanabe, N., Yao, 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., Belly, R. T. & Weiss, R. L. *Sulfolobus*: a new genus of sulfur-oxidizing bacteria living at low pH and high temperature. *Arch. Mikrobiol.* **84**, 54–68 (1972).
41. Higuchi, R., Krummel, B. & Saiki, R. K. A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. *Nucleic Acids Res.* **16**, 7351–7367 (1988).
42. Kabsch, W. XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
43. Brünger, A. T. Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472–475 (1992).
44. Terwilliger, 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., Jabłońska, J., Pravda, L., Vařeková, R. S. & Thornton, J. M. PDBsum: Structural summaries of PDB entries. *Protein Sci.* **112**, 535 (2017).
49. DeLano, W. L. PyMOL. (2002).
50. Nguyen-Duc, T., Peeters, E., Muyldermans, S., Charlier, D. & Hassanzadeh-Ghassabeh, G. Nanobody(R)-based chromatin immunoprecipitation/micro-array analysis for genome-wide identification of transcription factor DNA binding sites. *Nucleic Acids Res.* **41**, e59–e59 (2013).
51. Liu, H., Wang, K., Lindås, 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 genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).
55. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities 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. & Skaletsky, 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 $2-\Delta\Delta CT$ method. *Methods* **25**, 402–408 (2001).
59. Peeters, E., Thia-Toong, T.-L., Gigot, 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).

62. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* **17**, 10–12 (2011).
63. Chen, L. et al. The genome of *Sulfolobus acidocaldarius*, a model organism of the Crenarchaeota. *J. Bacteriol.* **187**, 4992–4999 (2005).
64. Kim, D. et al. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36 (2013).
65. Trapnell, C. et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511–515 (2010).
66. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
67. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
68. Notredame, C., Higgins, D. G. & Heringa, J. T-Coffee: a novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* **302**, 205–217 (2000).
69. Nguyen, N. T. T. et al. RSAT 2018: regulatory sequence analysis tools 20th anniversary. *Nucl. Acids Res.* **46**, W209–W214.

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 purification, 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.

SUPPLEMENTARY INFORMATION

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

Wang et al

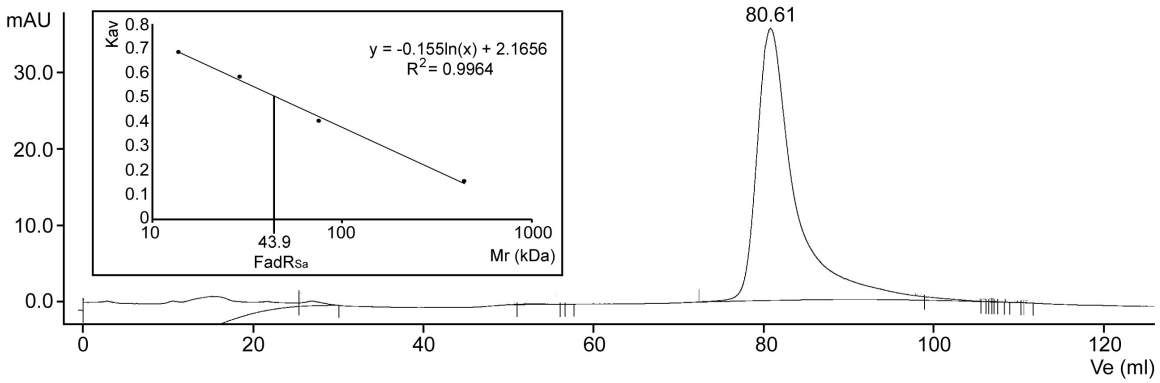
Supplementary Note 1. Detailed description of $FadR_{Sa}$-DNA contacts in the cocrystal structure.

A map of the observed protein-DNA interactions illustrates the extensive number of contacts that are established between each of the two DNA duplexes and the three interacting $FadR_{Sa}$ monomers (Figure 4a; Supplementary Data 1). These contacts take place in the major groove of the DNA and involve mainly $FadR_{Sa}$ residues from the α3 recognition helix (Tyr51, Phe52 and Tyr53) and from the loop between α2 and α3 (Ala46, Tyr47, Gly48 and Leu49). They consist of backbone contacts with the sugar or phosphate moiety in addition to base-specific contacts with 13 of the 21 bps of each DNA duplex. Tyr53 and Lys57, highly conserved in FadR-like or TetR-like proteins, respectively (Supplementary Figure 5a), form hydrogen bonds with the phosphate backbone thereby anchoring $FadR_{Sa}$ to the DNA so that the recognition helices are appropriately oriented for interactions. The base-specific contacts made by each of the $FadR_{Sa}$ monomers are essentially established by identical amino acid residues although the local nucleic acid environment encountered by each of the monomers is diverse (Figure 4b; Supplementary Data 1). The majority of the base-specific contacts are hydrophobic with involvement of Tyr47, Leu49, Tyr51 and Phe52: they form a hydrophobic plane docking into the major groove (Figure 4c). It can be postulated that the contacts established between Leu49, Tyr51 and Phe52 on one hand and methyl groups of thymines on the other hand (T5, T6, T11, T12, T13 and T17 of DNA chain X; T2, T7 and T14 of DNA chain Y; Figure 4b) endow a certain degree of sequence specificity to the interaction. Besides the hydrophobically interacting residues, Gly48 of subunit B forms a hydrogen bond between its amide nitrogen and the N7 of a guanine (Figure 4b; Supplementary Data 1). Although not categorized as hydrogen bonds, weak electrostatic interactions also exist between the N7 atom of G7 of the X chain and Gly48 of subunit A thereby categorizing it as a sequence-specific contact for all subunits. Besides protein-DNA contacts, a protein-protein interaction was also observed in the cocrystal structure between $FadR_{Sa}$ dimers bound on different sides of the DNA helix involving Asn37. The orientation of this residue in opposing monomers (e.g. subunits E and B) enables the formation of weak electrostatic interactions that could further stabilize the protein-DNA complex. The establishment of this protein-protein interaction might be possible because of the relative orientation of the two DNA-bound dimers to each other. This relative orientation is characterized by an angle of 122°, which is smaller than other TetR proteins that bind as a pair of dimers (Supplementary Figure 5b).

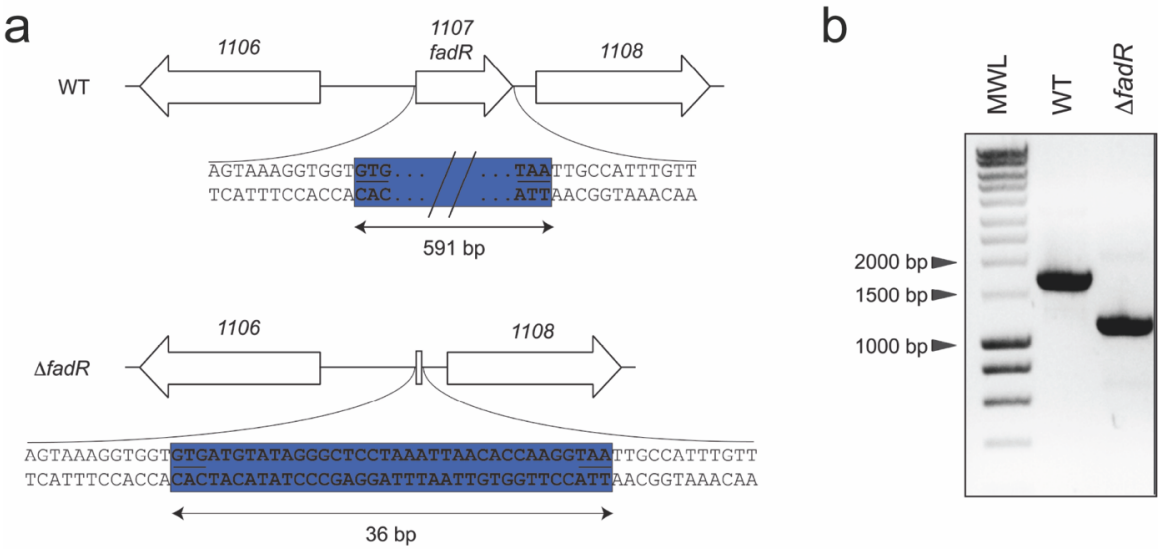
Supplementary Note 2. Detailed description of the interpretation of RNA-seq and qRT-PCR analyses.
Quantitative reverse transcriptase PCR (qRT-PCR) confirmed higher expression levels in the *ΔfadR*$_{Sa}$ strain for a subset of the *Saci_1103-Saci_1126* genes, including the *Saci_1105*, *Saci_1123* and *Saci_1124* genes, that were not reaching the significance threshold in the RNA-seq analysis (Figure 6b). The lack of detecting regulation of the CopG regulator-encoding *Saci_1124* with RNA-seq is explained by the gene being expressed at very low, almost undetectable levels. Nevertheless, qRT-PCR analysis demonstrated that the expression of this gene is 2.74 fold higher in the *fadR*$_{Sa}$ mutant strain (with a *P*-value of 0.0013) (Figure 6b). We can therefore conclude that FadR$_{Sa}$ is a local regulator of the entire *Saci_1103-Saci_1126* gene cluster (Figure 6a). RNA-seq analysis did not reveal differential expression for any of the genes adjacent to the other ChIP-seq identified genomic binding regions. However, several other genes were found to have a slightly lower expression in the *ΔfadR*$_{Sa}$ strain, pointing to an indirect regulatory effect (Supplementary Data 2). These genes include an operon encoding a putative sulfate reduction pathway (*Saci_2198-Saci_2203*) and cytochromes (*Saci_1858*, *Saci_1859* and *Saci_1861*). The observation that several of these genes are transcribed in operons strengthens the assumption that the observed small expression changes are relevant despite the determined fold-changes being unreliable (as confirmed by qRT-PCR (Figure 6b)).

Supplementary Note 3. Description of conformational differences in $FadR_{Sa}$ dimers in the $FadR_{Sa}$-DNA cocrystal structure.

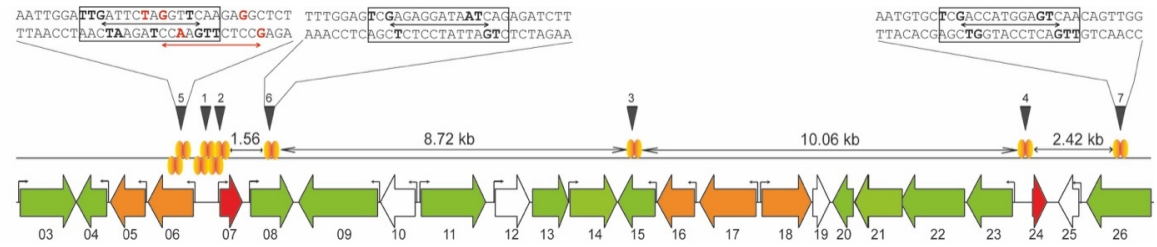
Intriguingly, the central (subunits E and F) but not the two flanking dimers (subunits A, B, C and D) in the asymmetric unit of the $FadR_{Sa}$-DNA structure harbored additional electron density in the ligand-binding pockets (Figure 3a). The exact nature of these ligands is unknown, although they are modeled to contain a CoA moiety. This can be explained by conformational differences in the central *versus* flanking dimers, which is also reflected in the distance between the two α3 recognition helices, which is 36.8 Å and 37.1 Å for the flanking dimers AB and CD, respectively, and much larger, 45.3 Å, for the central dimer (EF). The latter resembles the conformation of lauroyl-CoA-bound $FadR_{Sa}$, in which recognition helices are separated by 43.2 Å. This indicates that the ligand-bound E and F subunits of the central dimer are not well-positioned for simultaneous interaction with major groove segments of a single DNA molecule but that DNA binding can be accommodated for the two α3 helices by two distinct DNA duplexes. This creates an artificial situation with $FadR_{Sa}$ in a single complex both having ligand molecules bound and being bound to DNA at the same time.



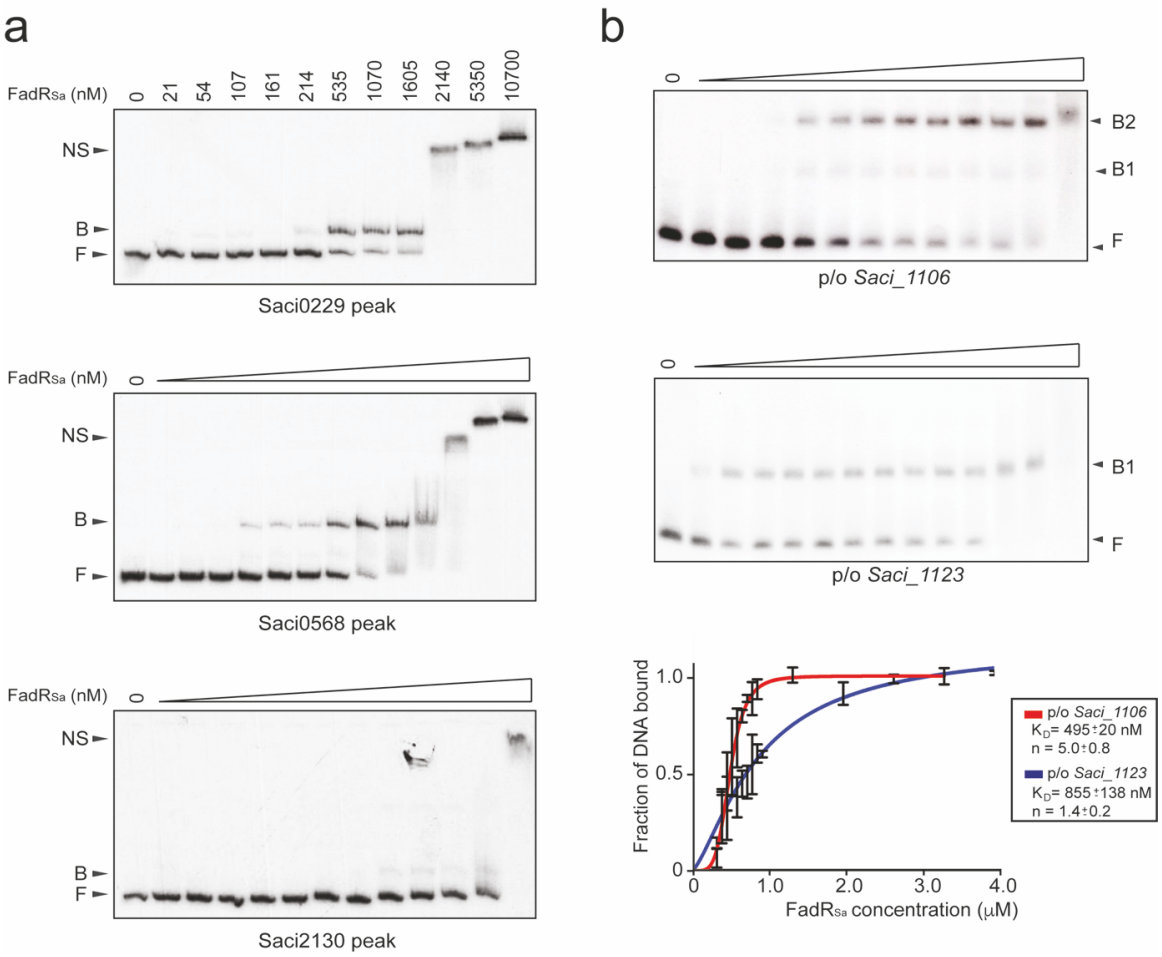
Supplementary Figure 1. Oligomeric state of $FadR_{Sa}$ in solution. The oligomeric state of native $FadR_{Sa}$ was analyzed by size exclusion chromatography using a Hiload 16/600 Superdex 200 pg column (GE Healthcare) equilibrated with 0.05 M sodium phosphate buffer (pH 7.4) containing 0.15 M NaCl. The calibration curve shown in the inset was prepared with ribonuclease A (13.7 kDa), carbonic anhydrase (29 kDa), conalbumin (75 kDa) and ferritin (440 kDa). After injecting 0.5 ml of purified recombinant $FadR_{Sa}$ at a concentration of 2.5 $\mu g\ ml^{-1}$, a single peak was observed with an elution volume (Ve) of 80.61 ml, corresponding to a molecular weight of 43.9 kDa as indicated on the calibration curve. Considering a molecular weight of 24 kDa for a His-tagged $FadR_{Sa}$ subunit, this observation supports a homogenous population of $FadR_{Sa}$ homodimers.



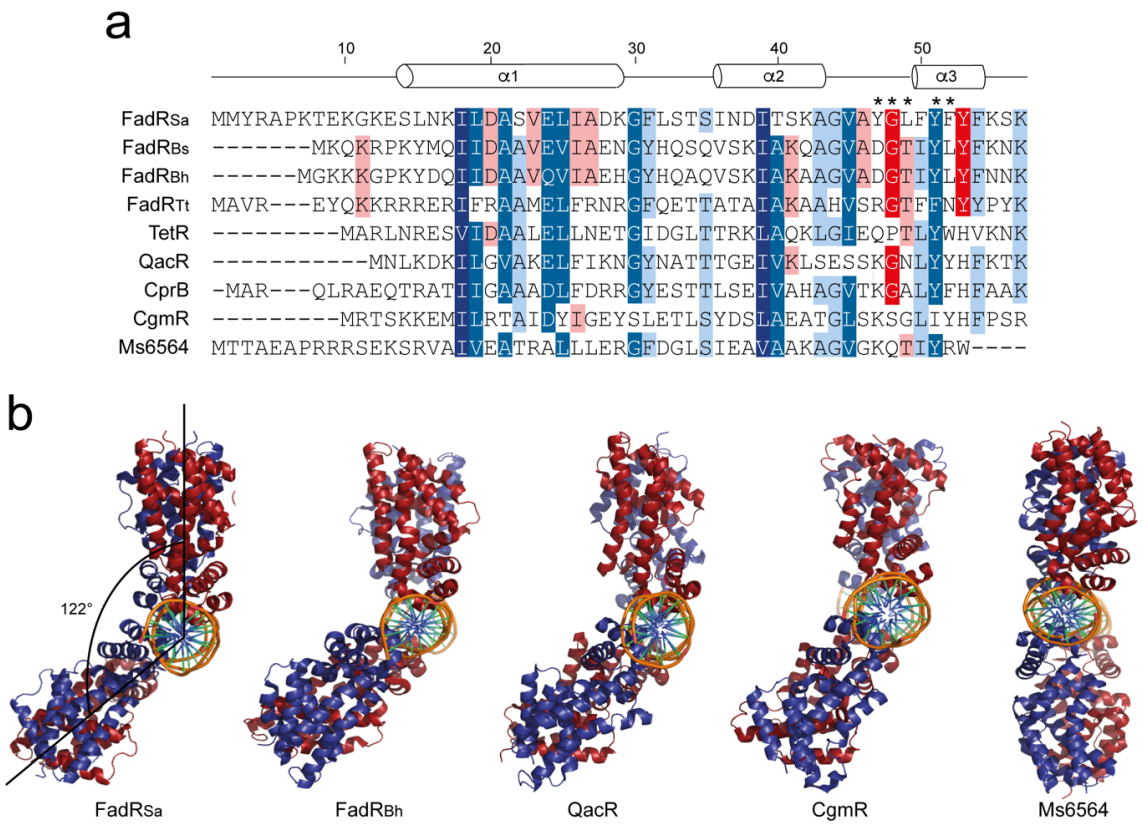
Supplementary Figure 2. Construction of an in-frame markerless $FadR_{Sa}$ deletion mutant. (**a**) Schematic representation of the genomic environment of the *$fadR_{Sa}$* gene in a *S. acidocaldarius* wild-type (MW001) and *S. acidocaldarius* MW001 Δ*$fadR_{Sa}$* strain, respectively. The length of the intact or remainder of the disrupted gene is mentioned, the latter resulting in a nonfunctional protein product of 12 amino acids. (**b**) PCR analysis of genomic DNA of the wild-type and Δ*$fadR_{Sa}$* strain with primers ep397 and ep398 (Supplementary Data 4) demonstrating a successful deletion. Product sizes are 1716 bp (WT) and 1161 bp (Δ*$fadR_{Sa}$*). MWL = molecular weight ladder.



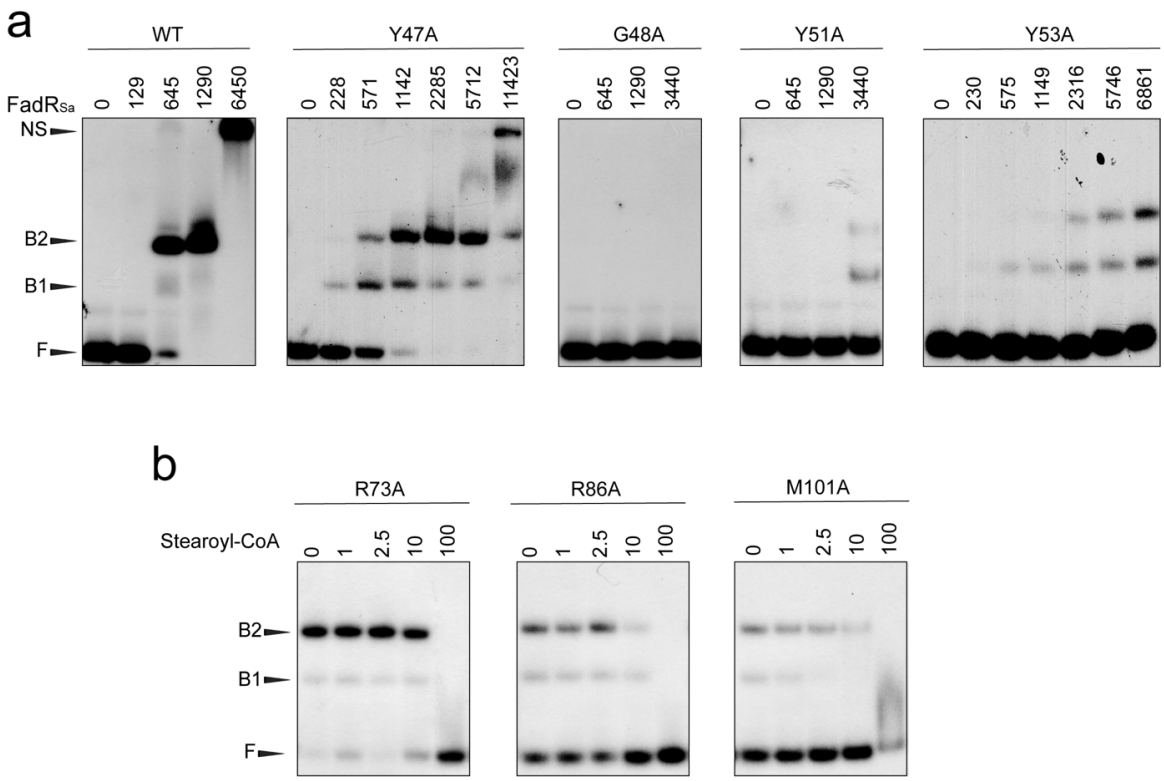
Supplementary Figure 3. Schematic overview of all (putative) $FadR_{Sa}$ binding sites in the *Saci_1103-Saci_1126* gene cluster. This scheme is depicted with indication of distances between them (in kb) and of whether the site is bound in a dimer or dimer-of-dimer mode (based on observations described in the section "DNA-binding stoichiometry of $FadR_{Sa}$"). Sequences of the newly predicted binding sites 5, 6 and 7 are shown, with indication of nucleotides that are assumed to be important for the interaction in bold (black = central operator; red = left or right operator for binding of second dimer).



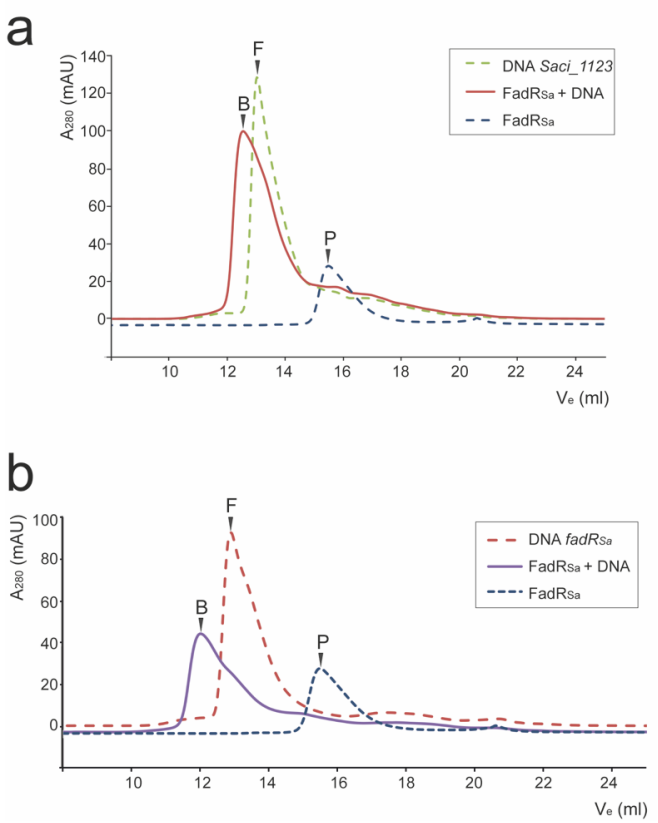
Supplementary Figure 4. *In vitro* DNA-binding analysis of $FadR_{Sa}$ to fragments representing ChIP-seq peaks. (**a**) Electrophoretic mobility shift assays (EMSAs) were performed with radiolabeled DNA probes representing high-enrichment binding regions identified in the ChIP-seq experiment located outside the lipid/fatty acid metabolism gene cluster (Supplementary Table 2). Experimental procedures are described in the Methods section. The same protein concentrations were used in the three experiments. Nucleic acid populations are indicated as follows: F = free DNA, B = specifically bound DNA, NS = non-specifically bound DNA. Note that for the Saci2130 peak fragment only a minor fraction of specific $FadR_{Sa}$-DNA complex was detected, which is hardly visible. (**b**) Binding parameter analysis of the interaction between $FadR_{Sa}$ and the promoter/operator (p/o) regions of *Saci_1106* and *Saci_1123*. Representative concentration-gradient EMSAs are shown that were used for the densitometric analysis and construction of binding curves. Below autoradiographs, binding curves are displayed that were fitted with a Hill equation on densitometric data with indication of the equilibrium dissociation constant K_D and Hill coefficient n as a measure for cooperativity. Averages are made for technical triplicates; error bars represent standard deviations.



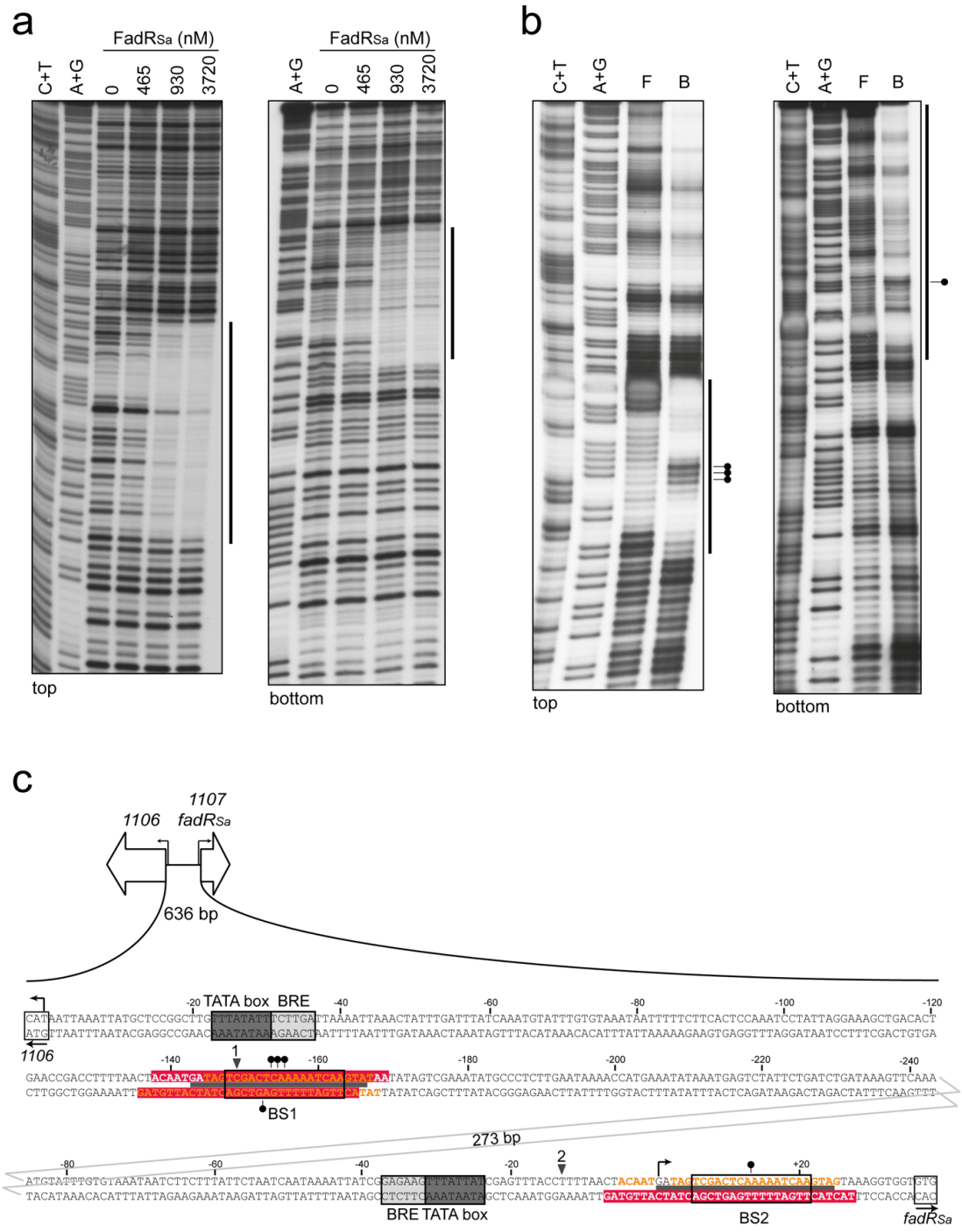
Supplementary Figure 5. Comparison of the DNA-binding mode of TetR family proteins. (**a**) Sequence alignment of the DNA-binding domains of the following TetR-like proteins: $FadR_{Sa}$ of *Sulfolobus acidocaldarius*, $FadR_{Bh}$ of *Bacillus halodurans*, $FadR_{Tt}$ of *Thermus thermophilus*, TetR of *Escherichia coli*, QacR of *Staphylococcus aureus*, CprB of *Streptomyces coelicolor*, CmgR of *Corynebacterium glutamicum* and Ms6564 of *Mycobacterium smegmatis*. Residues are indicated in blue when they are conserved in all (darkest blue), all but one (dark blue) or all but two (light blue) proteins (identical or similar residues). Residues are indicated in dark and light red when they are conserved in all or all but one FadR proteins, respectively (identical residues). Residue position numbers and secondary structure elements are indicated for $FadR_{Sa}$. Asterisk symbols denote $FadR_{Sa}$ residues in the DNA-binding domain that establish base-specific interactions. (**b**) Structural comparison of DNA-bound complexes of TetR-like proteins that bind in a dimer-of-dimer interaction mode (PDB codes: 6EN8 ($FadR_{Sa}$); 5GPC ($FadR_{Bh}$)[1]; 1JT0 (QacR)[2]; 2YVH (CgmR)[3]; 4JL3 (Ms6564)[4]). Each monomeric subunit of a dimer is colored differently (red or blue). In the case of $FadR_{Sa}$, only a single DNA molecule (XY) is shown with its two interacting dimers AB and EF.



Supplementary Figure 6. ***In vitro*** **DNA-binding and ligand-response analysis of $FadR_{Sa}$ mutants.** (**a**) Electrophoretic mobility shift assays (EMSAs) with DNA-binding mutants of $FadR_{Sa}$. Mutant proteins were prepared by site-directed mutagenesis using an overlap PCR approach followed by heterologous expression and purification with the same procedure as used for the WT protein. EMSAs were performed with radiolabeled DNA probes representing the high-affinity binding region in the promoter region of *Saci_1106*. Experimental procedures are described in the Methods section. Protein concentrations are expressed in nM units. Nucleic acid populations are indicated as follows: F = free DNA, B1 and B2 = specifically bound DNA, NS = non-specifically bound DNA. (**b**) EMSAs with ligand-binding mutants. In each of the lanes, an identical protein concentration was used. Stearoyl-CoA concentrations are mentioned in μM.

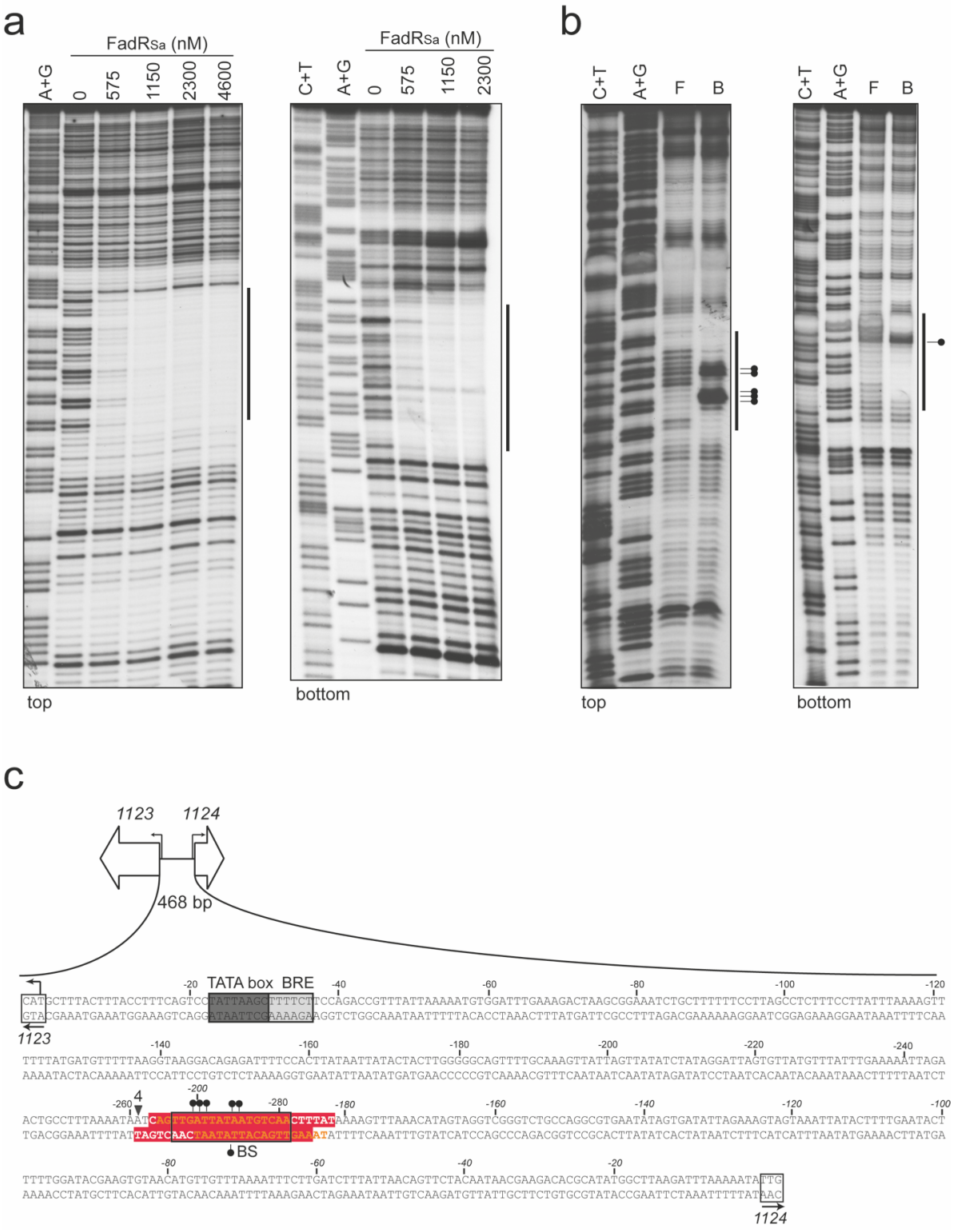


Supplementary Figure 7. Additional SEC analyses of $FadR_{Sa}$-DNA complexes. Size exclusion chromatography (SEC) was performed with 45-bp duplex oligonucleotides harboring the *Saci_1123* (**a**) or *$fadR_{Sa}$* (**b**) operator sequence. Following molar amounts were used: 1 nmol DNA, 8 nmol $FadR_{Sa}$ and 4 nmol $FadR_{Sa}$ mixed with 1 nmol DNA in the case of the *Saci_1123* operator (4:1 molar ratio) or 4 nmol $FadR_{Sa}$ mixed with 0.5 nmol DNA in the case of the *$fadR_{Sa}$* operator (8:1 molar ratio). Peaks are indicated as follows: P = unbound protein, F = unbound DNA and B = protein-DNA complex. Determination of molecular weights of these molecular species is presented in Figure 4a.



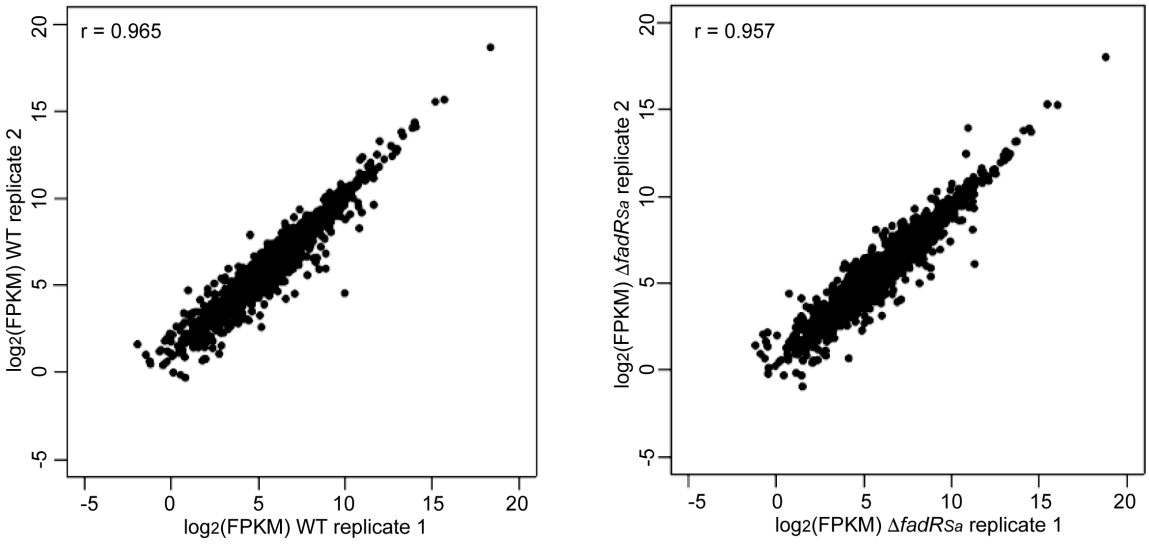
Supplementary Figure 8. *In vitro* contact probing of the interaction between $FadR_{Sa}$ and binding sites in the *Saci_1106-Saci_1107* intergenic region. (**a**) Autoradiographs of DNase I footprinting experiments analyzing $FadR_{Sa}$ binding to a probe representing peak 1 in the *Saci_1106-Saci1107* intergenic region. This was performed for a top-strand and bottom-strand labeled fragment (defined with respect to the orientation of *$fadR_{Sa}$* transcription), as indicated. A+G and C+T denote purine- and pyrimidine-specific Maxam-Gilbert sequencing ladders, respectively. Protected zones are indicated with a vertical line next to the autoradiographs. (**b**) Autoradiographs of chemical 'in-gel' Cu-phenantroline (Cu-OP) footprinting experiments analyzing $FadR_{Sa}$ binding to a probe representing peak 1. F and B indicate free and bound DNA populations, with the bound population corresponding to the slowest migrating complex in EMSA (B2) (Figure 4b and Supplementary Figure 4b). We do not consider the upper part of the footprint autoradiograph of the experiment with the top strand

labeled as a region protected by protein for two reasons: i) this protection is not observed in the corresponding region when performing the experiment with the bottom strand labelled (see right-hand panel of the figure), while all other footprinting protection zones are confirmed by protection observed for both DNA strands; ii) a nonspecific decrease in band intensity for the larger fragments can be explained by the chemical footprinting reaction conditions leading to excessive DNA cleavage in this specific experiment. Additionally, for the bottom-strand labeled fragment delineation of the protection zone is only accurate for the 5' end. (**c**) Nucleotide sequence of the *Saci_1106-Saci_1107* intergenic region with indication of the protection zones and hyperreactivity sites identified in the footprinting experiments presented in panels (**a**) and (**b**) (peak 1), and in Figure 3e (peak 2). Transcription start sites are indicated with an arrow and based on the observations in [5]. Translational start codons, putative TATA box and factor B recognition element (BRE) promoter elements are boxed, as well as the predicted pseudopalindromic binding sites (BSs). Residues corresponding to the ChIP-seq peak summits are indicated with triangle symbols. White letters in a red background represent protection zones observed in DNase I footprinting experiments, whereas orange letters represent protection zones observed in Cu-OP footprinting experiments. Ball-and-stick symbols represent hyperreactivity effects. A 24-bp repeat that harbors each of the binding sites is indicated with a grey line in between the top and bottom strand.

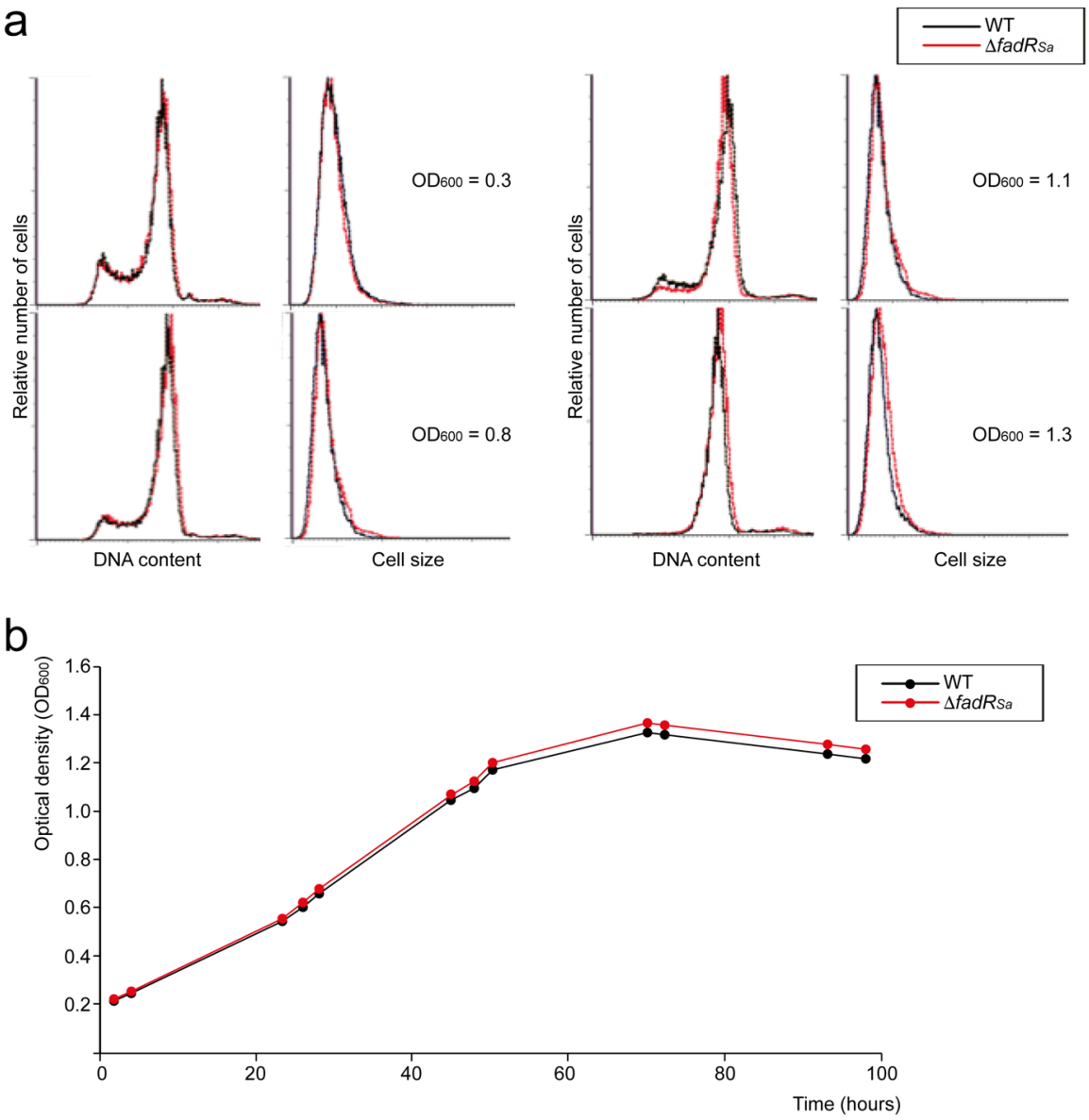


Supplementary Figure 9. *In vitro* contact probing of the interaction between $FadR_{Sa}$ and its binding site in the *Saci_1123-Saci_1124* intergenic region. (**a**) Autoradiographs of DNase I footprinting experiments analyzing $FadR_{Sa}$ binding to a probe representing peak 4 in the *Saci_1123-Saci1124* intergenic region. This was performed for a top-strand and bottom-strand labeled fragment (defined with respect to the orientation of *Saci_1124* transcription), as indicated. A+G and C+T denote purine- and pyrimidine-specific Maxam-Gilbert sequencing ladders, respectively. Protected zones are indicated with a vertical line next to the autoradiographs. (**b**)

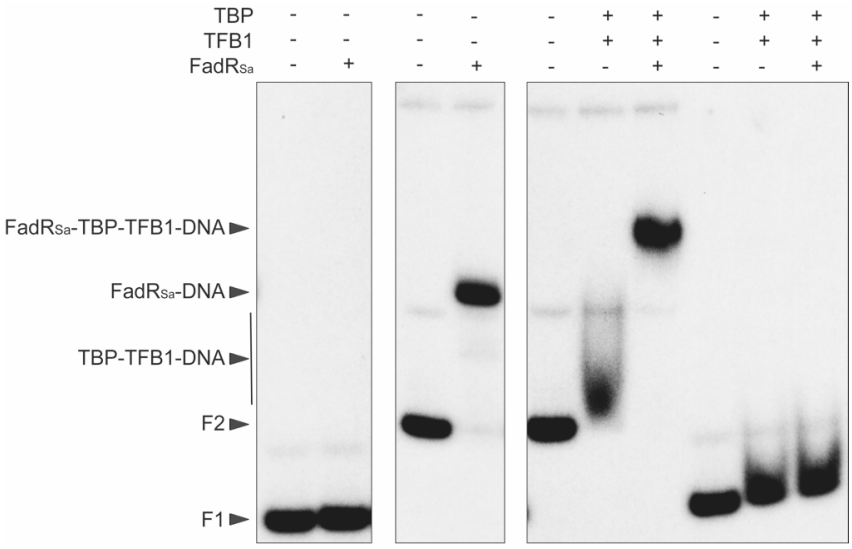
Autoradiographs of chemical 'in-gel' Cu-phenantroline (Cu-OP) footprinting experiments analyzing $FadR_{Sa}$ binding to a probe representing peak 1. F and B indicate free and bound DNA populations. (**c**) Nucleotide sequence of the *Saci_1123-Saci_1124* intergenic region with indication of the protection zones and hyperreactivity sites as identified in the footprinting experiments presented in panels (**a**) and (**b**). The *Saci_1123* transcription start site is indicated with an arrow and based on the observations in[5]. Translational start codons, putative TATA box and factor B recognition element (BRE) promoter elements are boxed, as well as the predicted pseudopalindromic binding site (BS). Numbering with respect to the *Saci_1123* transcription start site is in regular font while that with respect to the *Saci_1124* start codon is in italic font. The residue corresponding to the ChIP-seq peak summit is indicated with a triangle symbol. White letters in a red background represent protection zones observed in DNase I footprinting experiments, whereas orange letters represent protection zones observed in Cu-OP footprinting experiments. Ball-and-stick symbols represent hyperreactivity effects.



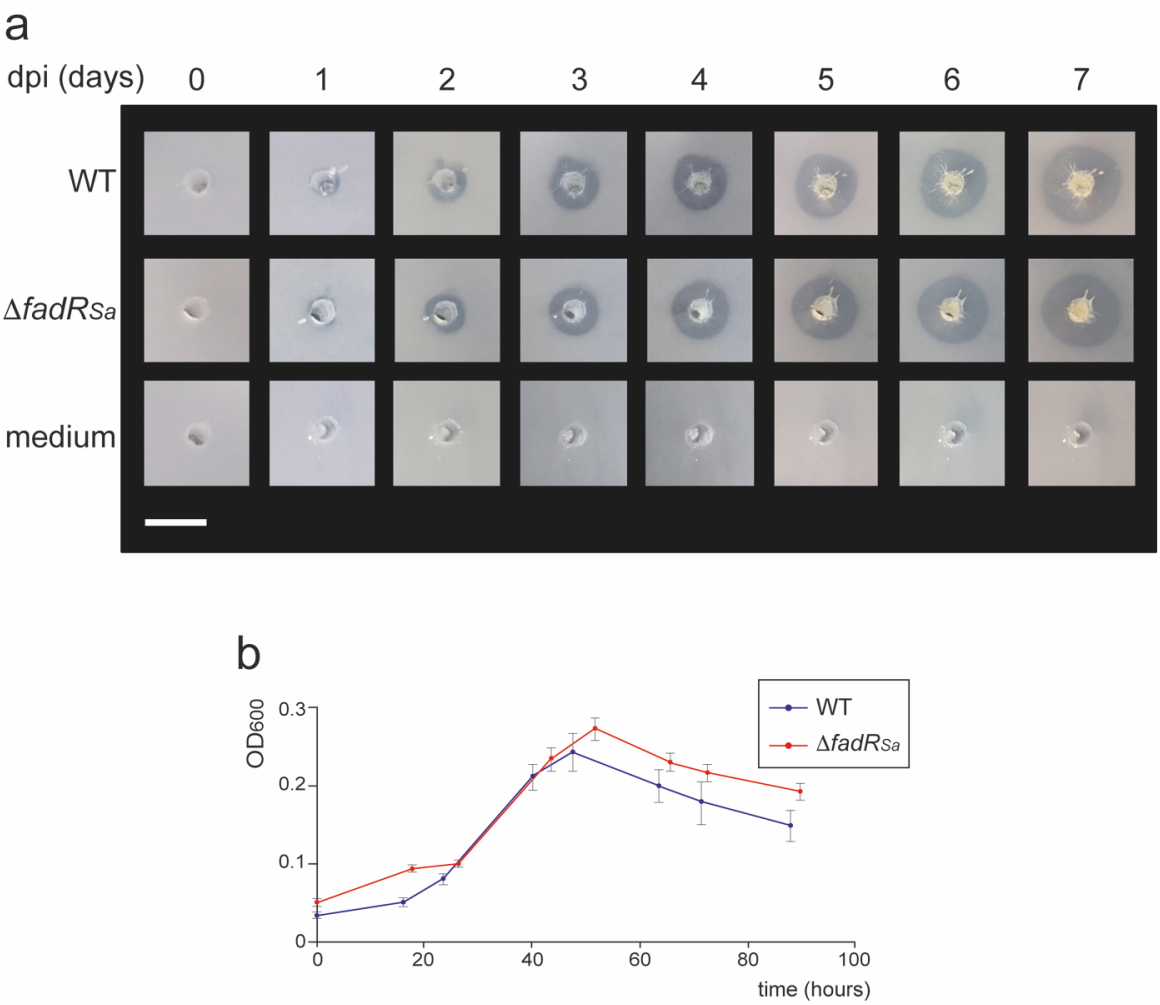
Supplementary Figure 10. Reproducibility of the RNA-seq data. Comparison of the $\log_2$(FPKM) values of the data points between each of the replicates, both for the wild-type samples as for the *ΔfadR*$_{Sa}$ samples.



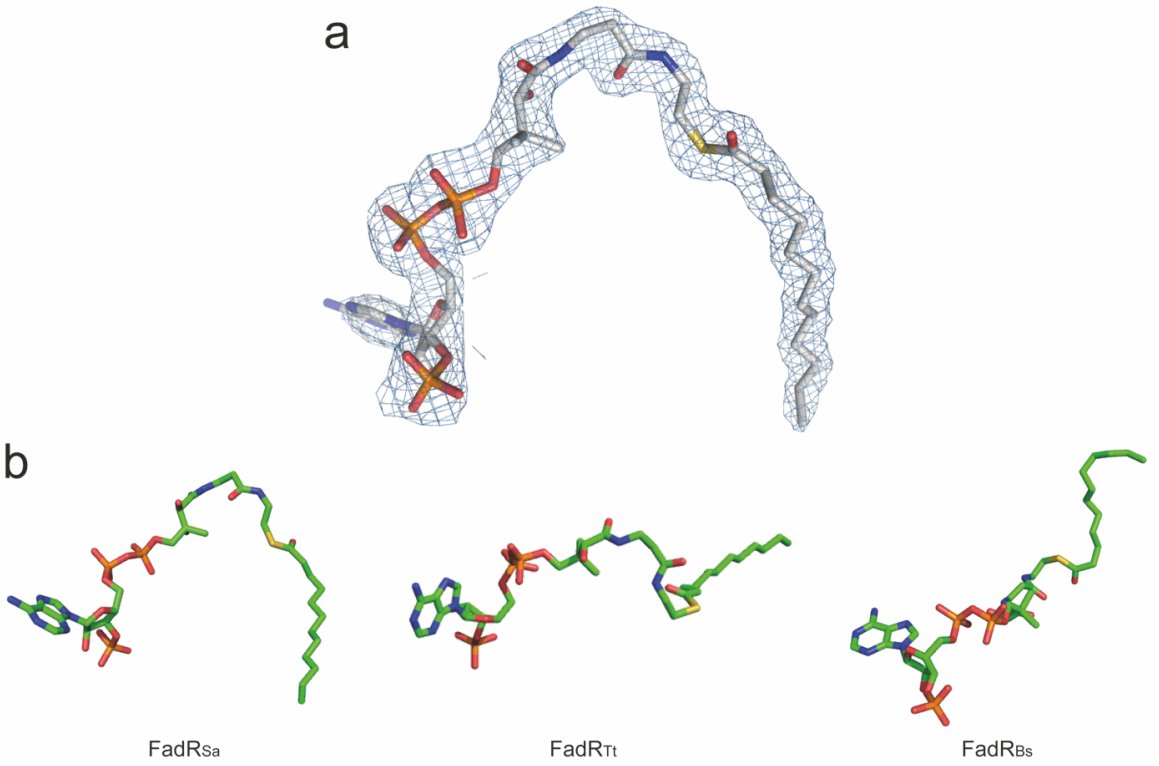
Supplementary Figure 11. Cell morphology and growth behaviour of the FadR$_{Sa}$ deletion mutant. (**a**) Flow cytometry histograms of DNA content and cell size for MW001 and MW001 Δ*fadR*$_{Sa}$ strains in different growth phases. Cell samples were collected at four time points during cell culturing representing early (OD_{600} = 0.3), middle (OD_{600} = 0.8) and late (OD_{600} = 1.1) exponential phase and stationary phase (OD_{600} = 1.3). (**b**) Growth curves of MW001 and MW001 Δ*fadR*$_{Sa}$ strains in Brock medium containing NZamine and sucrose. Cultures were grown in 50-ml volumes. This experiment was replicated multiple times; a representative curve is shown.



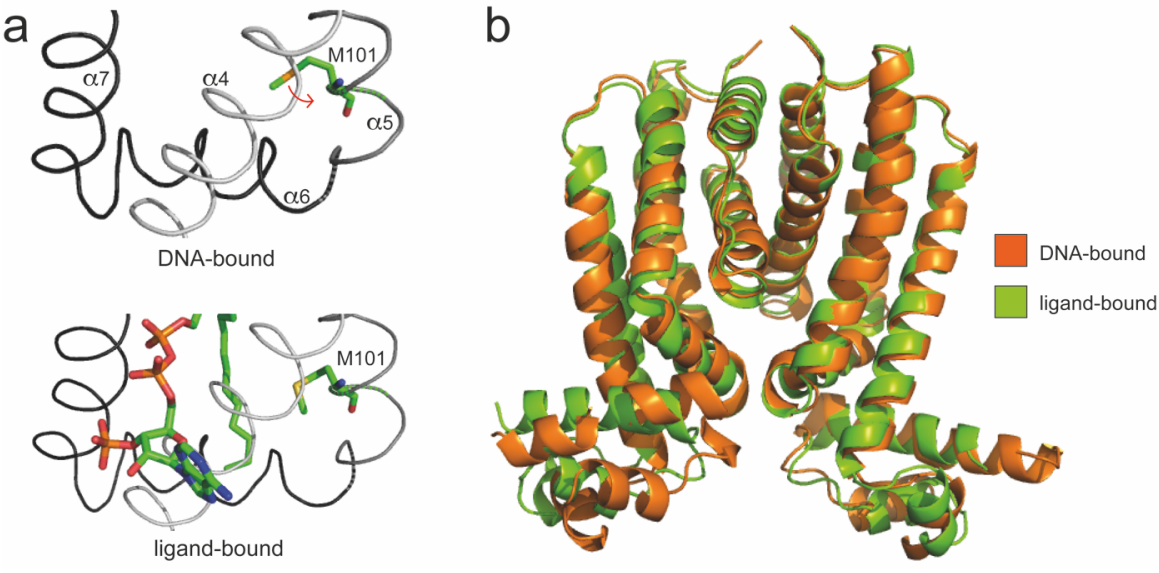
Supplementary Figure 12. *In vitro* DNA-binding analysis of TBP, TFB1 and $FadR_{Sa}$ to the *$fadR_{Sa}$* operator. Electrophoretic mobility shift assay (EMSA) analysis was performed for a radiolabelled negative control probe (F1) encompassing part of an ORF and for a radiolabelled probe containing the *$fadR_{Sa}$* control region (F2). The experimental procedure is described in the Methods section, with supplementation of a 5-fold higher concentration of non-specific competitor DNA (125 ng/μl) to minimize nonspecific interactions between TBP and TFB1 proteins on one hand and DNA on the other hand. Following protein concentrations were used: 974 nM $FadR_{Sa}$, 1.79 μM TBP and 864 nM TFB1. The supplementation of TBP and TFB1 causes the formation of an unstable nucleoprotein complex, while simultaneous addition of TBP, TFB1 and $FadR_{Sa}$ enables the formation of a stable $FadR_{Sa}$-TBP-TFB1-DNA complex with lower migration velocity that that of a $FadR_{Sa}$-DNA complex. This observation points to $FadR_{Sa}$ stimulating TBP and TFB1 binding. As such, the regulator does not employ a repression mechanism in which TBP and TFB1 binding is sterically inhibited, but transcriptional repression can be assumed to occur in later stages of transcription initiation (RNA polymerase binding or open complex formation).



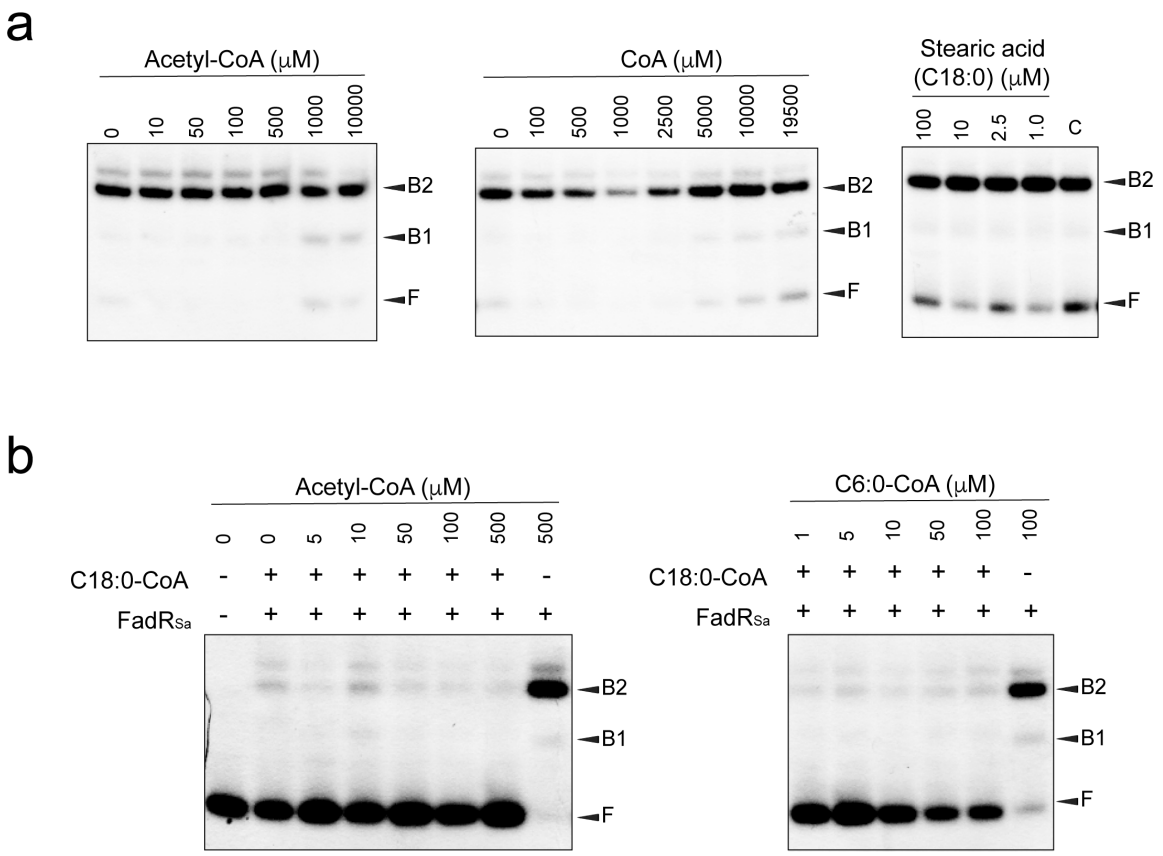
Supplementary Figure 13. Growth behavior of the $FadR_{Sa}$ deletion mutant in the presence of fatty acids and lipids. (**a**) *In vivo* esterase activity assays in which 20 µl of cultures with an OD_{600} around 0.7 were spotted on plates containing 1% (v/v) tributyrin, which were incubated at 78 °C for 7 days. After incubation esterase activity is visible by the appearance of a halo surrounding the original inoculation site indicating the hydrolysis of tributyrin. The esterase plate assay with the *S. acidocaldarius* MW001 (parent strain) and single knockout mutant Δ$fadR_{Sa}$ strain both show esterase activity. dpi: days post inoculation. The white bar corresponds to 1 cm. (**b**) Growth curves of MW001 and MW001 Δ$fadR_{Sa}$ strains in Brock medium containing 2 mM butyrate as sole carbon source. Values are averages of four biological replicates with error bars representing standard deviations. Representative curves are shown for multiple independently performed experiments.



Supplementary Figure 14. Comparison of the ligand conformation in FadR proteins. (**a**) Electron density map of the bound lauroyl-CoA in the $FadR_{Sa}$-lauroyl-CoA cocrystal structure. (**b**) Conformations of lauroyl-CoA in different cocrystal structures of $FadR_{Sa}$ (PDB: 6EL2), $FadR_{Tt}$ (PDB: 3ANG)[6] and $FadR_{Bs}$ (PDB: 3WHB)[7].



Supplementary Figure 15. Conformational differences between DNA-bound and ligand-bound $FadR_{Sa}$. (**a**) Close-up view of the ligand-binding pocket in DNA-bound and lauroyl-CoA bound $FadR_{Sa}$, with indication of the conformationally altered Met101 residue. (**b**) Superposition of the lauroyl-CoA-bound $FadR_{Sa}$ dimer (green) and subunits A and B of DNA-bound $FadR_{Sa}$ (orange), yielding an RMSD of 1.01 Å considering both subunits in each dimer.



Supplementary Figure 16. Effect of putative ligands on the $FadR_{Sa}$-DNA interaction. Electrophoretic mobility shift assays demonstrating the effect of putative ligands and combinations thereof on the interaction between $FadR_{Sa}$ and a 154-bp radiolabeled DNA probe representing the *Saci_1106* control region. Experimental procedures are described in the Methods section. (**a**) Analysis of the effect of acetyl-CoA, CoA and free fatty acids on the $FadR_{Sa}$-DNA interaction. In all binding reactions, a $FadR_{Sa}$ concentration of 645 nM was applied. Populations of free DNA (F) and of $FadR_{Sa}$-DNA complexes (B1 and B2) are indicated with arrowheads. In some assays, a minor population was observed of an additional nucleoprotein complex migrating somewhat higher than B2. C indicates a control reaction, in which the effect of 2% dimethylsulfoxide (used to dissolve stearic acid) in the binding buffer was tested. Acetyl-CoA and CoA molecules only displayed small ligand-induced inhibition effects at physiologically irrelevant high concentrations of 1 and 5 mM, respectively, while free fatty acids did not affect the $FadR_{Sa}$-DNA interaction at all. (**b**) Competition assay between long- and short-chain acyl-CoA ligands, which were added simultaneously to the binding reactions. In all binding reactions, + indicates the addition of 645 nM $FadR_{Sa}$ or 10 μM stearoyl-CoA, respectively.

Supplementary Table 1. Pairwise amino acid sequence identities and similarities between $FadR_{Sa}$ and bacterial acyl-CoA-dependent TetR members. Following bacterial proteins are considered: $FadR_{Bs}$ of *Bacillus subtilis*, $FadR_{Bh}$ of *Bacillus halodurans*, $FadR_{Tt}$ of *Thermus thermophilus* and $FabR_{Ec}$ of *Escherichia coli*. Alignments were performed with EMBOSS Needle. Values are expressed as %. N-terminal domains are defined as follows: residues 1-55 ($FadR_{Sa}$), 1-48 ($FadR_{Bs}$), 1-49 ($FadR_{Bh}$), 1-52 ($FadR_{Tt}$), 1-72 ($FabR_{Ec}$). C-terminal domains are defined as follows: residues 56-196 ($FadR_{Sa}$), 49-194 ($FadR_{Bs}$), 50-195 ($FadR_{Bh}$), 53-205 ($FadR_{Tt}$), 73-234 ($FabR_{Ec}$).

	$FadR_{Bs}$	**$FadR_{Bh}$**	**$FadR_{Tt}$**	**$FabR_{Ec}$**
Identity between FL sequences	26.0	18.8	18.8	17.0
Similarity between FL sequences	41.0	33.2	34.9	33.6
Identity between N-terminal domains	38.2	31.0	27.3	25.7
Similarity between N-terminal domains	58.2	50.0	49.1	43.2
Identity between C-terminal domains	9.0	15.1	11.0	6.6
Similarity between C-terminal domains	16.1	28.1	17.8	15.5

Supplementary Table 2. Summary of ChIP-seq data.

Peak summit coordinate	Fold enrichment	Nearest open reading frame (ORF)	Annotation	Peak summit location[a]	*In vitro* binding[b]	Predicted binding motif	Motif P-value	Distance between motif and summit
191020	2.51	*Saci_0229*	Hypothetical protein	G	++	CTGACTGTAGAATCAA	$6.46E^{-07}$	-1
224943	1.74	*Saci_0266*	Hypothetical protein	G	N.D.	TTGACAAGCTAATCAA	$8.14E^{-06}$	-13
460981	2.21	*Saci_0568*	Hypothetical protein	G	++	TTGAGTCAGTAATCAG	$5.33E^{-06}$	-5
813158	1.74	*Saci_1014*	Aminotransferase	G	N.D.	TTGATACCTGAGTCAA	$6.46E^{-07}$	21
907585	4.24	*Saci_1106*	Acyl-CoA esterase	I	+++	TCGACTCAAAAATCAA	$8.14E^{-06}$	-7
908025	4.24	*Saci_1107*	TetR family transcriptional regulator	I	+++	TCGACTCAAAAATCAA	$8.14E^{-06}$	-26
910364	1.63	*Saci_1109*	Enoyl-CoA hydratase	G	N.D.	N.A.	N.A.	N.A.
918302	2.03	*Saci_1115*	Acyl-2-enoyl-CoA reductase	G	++	TTGACAGAGGGATCAA	$1.19E^{-05}$	-48
928397	2.17	*Saci_1124*	CopG family transcription regulator	I	++	TTGACATTATAATCAA	$2.52E^{-06}$	-14
1161910	1.47	*Saci_1359*	3-hydroxy-3-methylglutaryl coenzyme A reductase	G	N.D.	CTGATAGTATAGTCAA	$7.35E^{-06}$	39
1166706	1.76	*Saci_1364*	Hypothetical protein	I	N.D.	TTGACCCTTTAATCAA	$1.07E^{-06}$	9
1611635	1.83	*Saci_1843*	Cyclase	I	N.D.	TTGATGATATAATCAA	$1.44E^{-06}$	-20
1929449	1.85	*Saci_2107*	Hypothetical protein	G	N.D.	CTGAGTCGGATATCAA	$2.79E^{-05}$	-16
1956504	1.83	*Saci_2130*	Oxidoreductase	G	+	TTGATGCAATGGTCAA	$6.57E^{-06}$	1

[a] G = genic (inside ORF); I = intergenic; [b] *in vitro* binding results.

+ = specific but low-affinity binding, ++ = intermediate-affinity binding, +++ = high-affinity binding.

N.D. = not determined. N.A. = not applicable.

Supplementary Table 3. Binding motif predictions in genomic regions not captured by ChIP-seq.
FIMO predictions of putative $FadR_{Sa}$ binding sites in genomic regions encompassing genes downregulated in the RNA-seq analysis presented in Supplementary Data 2.

Predicted motif	Motif *P*-value	Genomic coordinate	Location with respect to ORFs
CTGACCCTTTGGTCAA	$1.71E^{-05}$	2037883	Inside *Saci_2200*
TTGAGAGTGTCCTCAA	$3.19E^{-05}$	1634798	Inside *Saci_1859*
CTGATAGTCAAATCAG	$4.69E^{-05}$	2093127	Inside *Saci_2250*
TCGACAGGAACGTCAA	$8.62E^{-05}$	380550	Inside *Saci_0447*
CTGAAGACATAATCAG	$9.20E^{-05}$	2094465	Inside *Saci_2252*
CTGACTATAGAGTCTA	$9.30E^{-05}$	2039429	Inside *Saci_2201*

FIMO predictions of putative $FadR_{Sa}$ binding sites in the *Saci_1103-Saci_1126* genomic region (genomic coordinates: 902793-932298) not retrieved in previous characterizations and/or predictions presented in Supplementary Table 2.

Predicted motif	Motif *P*-value	Genomic coordinate	Location with respect to ORFs
TTGATTCTAGGTTCAA	$3.55E^{-06}$	906727	Inside *Saci_1106*
TCGACCATGGAGTCAA	$8.89E^{-06}$	930823	Inside *Saci_1126*
TCGAGAGGATAATCAG	$6.93E^{-05}$	909604	Inside *Saci_1108*

Supplementary Table 4. Summary statistics and alignment information of the paired-end RNA-seq experiment. WT = *S. acidocaldarius* MW001; KO = *S. acidocaldarius* MW001 $\Delta fadR_{Sa}$.

Sample	Raw read length (bp)	Read pairs	Mapped read pairs	Mapping rate (%)
WT1	124	8639675	8313946	96.23
WT2	124	10597695	10196655	96.20
KO1	124	11779481	11330497	96.20
KO2	124	7618684	7343047	96.40

Supplementary Table 5. Overview of strains used in this work.

Name	Description/purpose	Reference or source
Escherichia coli DH5α	Plasmid propagation strain	Gibco
E. coli Rosetta (DE3)	Protein overexpression strain	Novagen
E. coli Rosetta 2 (DE3)	Protein overexpression strain	Novagen
E. coli ER1821	Strain used for plasmid methylation	New England Biolabs
Sulfolobus acidocaldarius DSM639	Wild-type strain	DSMZ
S. acidocaldarius MW001	Uracil auxotrophic strain for genetic experiments	[8]
S. acidocaldarius MW001Δ*fadR*	Markerless *fadR* gene deletion mutant	This work

Supplementary Table 6. Overview of plasmids used in this work.

Name of plasmid	Description/purpose	Reference
pET45b	Protein overexpression vector	Novagen
pET24a	Protein overexpression vector	Novagen
pET30a	Protein overexpression vector	Novagen
pET45b*xfadR$_{Sa}$*	pET45b containing the *fadR$_{Sa}$* open reading frame	This work
pET24a*xfadR$_{Sa}$Ndenull*	pET24a containing *fadR$_{Sa}$Ndenull*	This work
pET24a*xfadR$_{Sa}^{Y47A}$Ndenull*	pET24a containing *fadR$_{Sa}$Ndenull* with Y47A mutation	This work
pET24a*xfadR$_{Sa}^{G48A}$Ndenull*	pET24a containing *fadR$_{Sa}$Ndenull* with G48A mutation	This work
pET24a*xfadR$_{Sa}^{Y51A}$Ndenull*	pET24a containing *fadR$_{Sa}$Ndenull* with Y51A mutation	This work
pET24a*xfadR$_{Sa}^{Y53A}$Ndenull*	pET24a containing *fadR$_{Sa}$Ndenull* with Y53A mutation	This work
pET24a*xfadR$_{Sa}^{R73A}$Ndenull*	pET24a containing *fadR$_{Sa}$Ndenull* with R73A mutation	This work
pET24a*xfadR$_{Sa}^{R86A}$Ndenull*	pET24a containing *fadR$_{Sa}$Ndenull* with R86A mutation	This work
pET24a*xfadR$_{Sa}^{M101A}$Ndenull*	pET24a containing *fadR$_{Sa}$Ndenull* with M101A mutation	This work
pET30a*xtbp*	pET30a containing the *tbp* open reading frame	This work
pET30a*xtfb1*	pET30a containing the *tfb1* open reading frame	This work
pSVA431	Backbone plasmid for gene disruption construct, harbouring a *pyrEF* selection marker	[8]
pSVA431x*ΔfadR$_{Sa}$*	Suicide *fadR$_{Sa}$* gene disruption construct	This work

SUPPLEMENTARY REFERENCES

1. Yeo, H. K., Park, Y. W. & Lee, J. Y. Structural basis of operator sites recognition and effector binding in the TetR family transcription regulator FadR. *Nucleic Acids Res.* **45**, 4244-4254 (2017).
2. Schumacher M. A. *et al.* Structural basis for cooperative DNA binding by two dimers of the multidrug-binding protein QacR. *EMBO J.* **21**, 1210-1218 (2002).
3. Itou H., Watanabe, N., Yao, M., Sirakihara, Y. & Tanaka, I. Crystal structures of the multidrug binding repressor *Corynebacterium glutamicum* CgmR in complex with inducers and with an operator. *Journal of Molecular Biology* **403**, 174-184 (2010).
4. Yang S. *et al.* Structural basis for interaction between *Mycobacterium smegmatis* Ms6564, a TetR family master regulator, and its target DNA. *J. Biol. Chem.* **288**, 23687-23695 (2013).
5. Cohen, O. *et al.* Comparative transcriptomics across the prokaryotic tree of life. *Nucleic Acids Res.* **44,** W46–53 (2016).
6. Agari, Y., Agari, K., Sakamoto, K., Kuramitsu, S. & Shinkai, A. TetR-family transcriptional repressor *Thermus thermophilus* FadR controls fatty acid degradation. *Microbiology* **157,** 1589–1601 (2011).
7. Fujihashi, M. *et al.* Structural characterization of a ligand-bound form of *Bacillus subtilis* FadR involved in the regulation of fatty acid degradation. *Proteins* **82,** 1301–1310 (2014).
8. Wagner, M. *et al.* Versatile genetic tool box for the crenarchaeote *Sulfolobus acidocaldarius*. *Front Microbiol* **3,** 214 (2012).

Chapter 3.2

Fatty acid metabolism in *Sulfolobus acidocaldarius* – a potential archaeal pathway for fatty acid synthesis

Fatty acid metabolism in *Sulfolobus acidocaldarius* – a potential archaeal pathway for fatty acid synthesis

Authors

Xiaoxiao Zhou[a], Christian Schmerling[a], Kun Wang[b], David Sybers[c], Till Kessenbrock[d], Ann-Christin Lindås[b], Eveline Peeters[c], Markus Kaiser[d], Christopher Bräsen[a], Bettina Siebers[a, #]

[a]Molecular Enzyme Technology and Biochemistry (MEB), Environmental Microbiology and Biotechnology (EMB), Centre for Water and Environmental Research (CWE), Department of Chemistry, University of Duisburg-Essen, Universitätsstr. 5, 45141 Essen, Germany

[b]Department of Molecular Biosciences, The Wenner-Gren Institute, Stockholm University, SE-106 91, Stockholm, Sweden

[c]Bio-engineering Sciences, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussels, Belgium

[d]Chemical Biology, Centre of Medical Biotechnology (ZMB), Department of Biology, University of Duisburg-Essen, Universitätsstr. 2, 45117 Essen, Germany

[#]Address correspondence to: Bettina Siebers, bettina.siebers@uni-due.de

Keywords

Archaea – *Sulfolobus acidocaldarius* – Fatty acids – β oxidation – Fatty acid biosynthesis

Abbreviations

(FAs) fatty acids – (ACS) acyl-CoA synthetase – (ACAD) acyl-CoA dehydrogenase – (ETF) electron transfer flavoprotein – (HCDH/ECH) 3(S)-hydroxyacyl-CoA dehydrogenase/enoyl-CoA hydratase – (KT) β-ketothiolase or acetyl-CoA C acetyltransferase – (ACR) acetoacyl-CoA reductase – (MaoC-HCD) MaoC like 3(R)-hydroxyacyl-CoA dehydratase – (ECR) enoyl-CoA reductase – (3-HBCoA) 3-hydroxybutyryl-CoA – (AcAcCoA) acetoacetyl-CoA – (SDR) short-chain dehydrogenases/reductases superfamily – (MDR) medium-chain dehydrogenases/reductases superfamily – (FAD) flavin adenine dinucleotide – (DCPIP) 2,6-dichlorophenolindophenol – (INT) iodonitrotetrazolium – (DTNB) 5,5'-dithiobis-(2-nitrobenzoic acid) or Ellman's reagent – (F_cPF6) ferrocenium hexafluorophosphate – (DTT) dithiothreitol – (ACN) acetonitrile

Word count for the abstract: 296

Word count for the text: 15565

Abstract

Fatty acids are main building blocks for bacterial and eukaryotic membrane lipids and together with other functions e.g. in energy supply and signalling central to cellular structure. However, in Archaea this central function as membrane component is substituted by isoprenoid groups which are ether-connected to the glycerol 1-phosphate backbone and therefore, the presence and metabolism of FAs in Archaea remained elusive. Scarce reports indicated the presence and hence the ability at least of some Archaea to synthesize FAs in an acyl carrier protein independent manner probably via a reversed β oxidation. Also, the possibility of some Archaea to utilize FAs as carbon and energy source was demonstrated. However, detailed biochemical studies have not been performed. We herein studied the FA metabolism in *S. acidocaldarius* by characterisation of the β oxidation homologues encoded in one gene cluster (*saci_1103-1126*). From the characterized single enzymes a fully functional β oxidation spiral was reconstituted *in vitro* and the complete oxidation of FAs to acetyl-CoA up to a chain length of C8 was demonstrated. The data further indicated that the β oxidation is not fully reversible. Instead, the potential of *S. acidocaldarius* to synthesize FAs via a novel CoA dependent pathway acting independently from β oxidation enzymes is shown. The pathway comprises a bacterial like SDR superfamily R-specific fabG homolog, an MDR superfamily enoyl-CoA reductase like in some Eukaryotes, both with a clear preference for NADPH as electron donor. Furthermore, an R-specific MaoC like dehydratase was identified. The enzymes were biochemically characterized and catalysed the synthesis of medium chain FA-CoA esters up to chain lengths of C8. These results provide a basic understanding of the FA metabolism in Archaea and thus pave the way for the further understanding of the presence and significance of FA in Archaea and its evolutionary implications.

Introduction

The “lipid divide” denotes that the membrane lipid composition of Archaea is fundamentally different from those in Bacteria and Eukaryotes. The membrane phospholipids of Bacteria and Eukaryotes are composed of fatty acids ester-linked to glycerol 3-phosphate (G3P) forming membrane bilayers whereas archaeal membrane lipids are comprised of isoprenoid chains ether-linked to glycerol 1-phosphate (G1P) (Fig. S1) forming mono- or bilayer membranes. Thus, the membrane lipids are one of the most striking characteristic traits of Archaea. Therefore, since one of the main and perhaps most important function of fatty acids (FAs) as key constituent of cell membrane phospholipids and thus of cell structure in Bacteria and Eukaryotes is substituted by isoprenoids in Archaea, the presence and the function of FAs in archaeal representatives remained obscure.

Previous findings indicated the ability of several Archaea to degrade FA based lipids into FAs and glycerol by means of esterases and lipases recently analysed in detail using an ABPP approach newly established for *Sulfolobus acidocaldarius*, *Saccharolobus solfataricus*, and also for *Haloferax volcanii* [1]. In addition, *Saccharolobus solfataricus P1* was shown to degrade complex FA based lipids [2]. Also, at least in some Archaea FAs might serve as carbon and energy source (e.g. haloarchaea and *Archaeoglobus fulgidus*) [3, 4] and homologues of all enzymes from the bacterial β oxidation have been identified in several archaeal species although detailed analyses have not been reported [3, 5]. Interestingly, the genes encoding the characterized esterases in *S. acidocaldarius* (*saci_1105, saci_1116*) involved in triacylglycerol degradation are organized in a gene cluster (*saci_1103-saci_1126*) (Fig. S2) together with several β oxidation homologues which thus may be responsible for the lipid degradation in this organism [1, 6-8]. *S. acidocaldarius* is an obligately aerobic, thermoacidophilic (optimal growth around 80°C and pH around 3) (cren)archael model organism whose central carbohydrate and energy metabolism is well understood [9-11]. The genome sequence as well as genetic tools are available [10, 12]. One gene within this *S. acidocaldarius saci_1103-1126* gene cluster, i.e. *saci_1107*, encodes a transcriptional regulator from the TetR family which binds specifically and exclusively to DNA recognition sequences within this cluster and regulates its own expression as well as that of many of the β oxidation homologues [8]. This analysis also showed that the regulator is a repressor and that derepression occurs when FA-CoA esters bind to the protein. The crystal structure of the regulator was determined including the binding to the DNA as well as the binding mode of the acyl-CoA. Furthermore, the growth of *S. acidocaldarius* on the FAs butyrate and hexanoate as sole carbon and energy source was clearly shown [8]. In addition to the esterases and the TetR regulator, the cluster comprises genes putatively encoding enzymes for FA degradation [8]. To channel FAs into β oxidation they first need to be activated to the corresponding CoA esters usually carried out by AMP-forming acyl-CoA synthetases (ACSs) [13-15] (*saci_1111,*

saci_1122, saci_1126). Also, three homologues of acyl-CoA dehydrogenases (ACADs) have been identified (*saci_1108, saci_1113, saci_1123*). As a first step in the β oxidation, these enzymes introduce a double bond in β position of the straight chain acyl-CoAs oxidizing them to the respective enoyl-CoA esters. The next step in β oxidation is the hydration of the enoyl-CoA to the corresponding 3(S)-hydroxyacyl-CoA followed by the 3(S)-hydroxyacyl-CoA oxidation to 3-ketoacyl-CoA carried out by crotonase superfamily hydratases (ECHs) and hydroxyacyl-CoA dehydrogenase (HCDHs) superfamily enzymes, respectively. For both, only one gene (*saci_1109*) exists in the cluster and both are fused to encode a single polypeptide chain. Finally, the 3-ketoacyl-CoA is thiolytically cleaved by ketothiolases (KTs) to yield acetyl-CoA and a saturated acyl-CoA ester shortened by two C atoms (*saci_1114, saci_1121*). However, the coding function of these genes have not been confirmed so far.

Furthermore, scarce reports previously suggested that in few Archaea FAs are present including *Sulfolobus* spp. although the organisms were grown and adapted to FA free medium [16-19]. And thus, if this is indeed the case, these archaeal organisms must have a biosynthesis machinery for FAs. However, neither a complete classical FAS II system known from bacteria nor a FAS I machinery present in animals and fungi have been identified in any archaeon [5]. Also, the acyl carrier protein and the acyl carrier protein synthase essential for the FA biosynthesis in bacteria and eukaryotes is not present in most Archaea [20]. Thus, if at least some archaea are able to synthesize FAs then the synthesis machinery must be fundamentally different from the known systems in Bacteria and Eukaryotes. Recently, an ACP-independent FA synthesis pathway has been proposed which relies on bacterial like homologues of the FA synthesis system (FAS II) [20]. However, complete sets of such homologues could not been identified in any archaeon so far and instead, a reversed β oxidation has been proposed to be responsible for FA synthesis in Archaea [5]. However, also the reversibility of the β oxidation in general and particularly in Archaea has not been demonstrated so far. Instead, both processes known from Bacteria and Eukaryotes, the FA synthesis on the one hand and the β oxidation on the other hand which basically follow the same chemical conversions, show some remarkable differences in order to drive either of the processes in the desired direction and to separate both processes which at least in prokaryotes are localized in the cytoplasm (Fig. S3). First of all both processes are tightly regulated on the transcriptional level but also on the protein level e.g. by feedback inhibition [21]. As mentioned above the synthesis machinery (Fig. S3B) is ACP dependent which activate the FAs and transports the growing acyl chain between the involved enzymes [22]. Conversely, the β oxidation (Fig. S3A) relies on CoA for FA activation. The FA synthesis is usually NADPH dependent whereas the β oxidation uses NAD^+ and FAD [23] although this cosubstrate specificity is less pronounced in Bacteria [24]. The Claisen condensation, the first reaction in the synthesis, uses malonyl-ACP as donor for chain elongation instead of acetyl-CoA, and

malonyl-ACP has to be synthesized through ATP-dependent carboxylation of acetyl-CoA (and transfer to ACP). During condensation the CO_2 is liberated again. Thus, the endergonic reaction is energetically driven by ATP hydrolysis. Furthermore, the Claisen condensation itself catalysed by the decarboxylating ketoacyl-ACP synthase is regarded as irreversible in the synthesis direction [24, 25]. The following two reactions in the FA synthesis pathway, i.e. the reduction of oxoacyl to the hydroxyacyl intermediate and the dehydration to the enoyl moiety are in principle reversible but are specific for the R-hydroxyacyl intermediate instead of the S-stereoisomer specific β oxidation [23]. Finally, the interconversion of the saturated acyl- to the unsaturated enoyl-intermediate on the β oxidation is carried out by FAD dependent dehydrogenases channelling the electrons via the electron transferring flavoprotein (ETF) and the ETF: quinone oxidoreductase (EQOR) into the quinone pool of the respiratory chain [26, 27]. This reaction is generally regarded as irreversible in the oxidative direction [28] and usually bypassed in the FA synthesis by NAD(P)H dependent dehydrogenases which in turn makes the reaction highly exergonic towards the reductive direction [24, 29].

We herein studied the FA metabolism in *S. acidocaldarius* by characterisation of the β oxidation homologues encoded in the *saci_1103-1126* gene cluster. From the characterized single enzymes a fully functional β oxidation spiral was reconstituted *in vitro* and the complete oxidation of FAs to acetyl-CoA up to a chain length of C8 was demonstrated (Fig. 1, S24). The data further indicated that the β oxidation is not fully reversible. Instead, the potential of *S. acidocaldarius* to synthesize FAs via a novel CoA dependent pathway acting independently from β oxidation enzymes is shown (Fig. 2A). The pathway comprises a bacterial like SDR superfamily R-specific fabG homolog (ACR, *saci_1104*), an MDR superfamily enoyl-CoA reductase (ECR, *saci_1115*) like in Eukaryotes, both with a clear preference for NADPH as electron donor, as well as an R-specific MaoC like dehydratase (MaoC-HCD, *saci_1085*). The enzymes were biochemically characterized and catalysed the synthesis of medium chain FA-CoA esters up to chain lengths of C8.

Material and Methods

The CoA esters including HS-CoA, acetyl-CoA and butyryl-CoA employed in this study were purchased from Millipore Sigma® Merck KGaA in Darmstadt, Germany. The commercial hexanoyl-CoA, octanoyl-CoA, crotonoyl-CoA and mixed 3(S/R)-hydroxybutyryl-CoA were produced by Santa Cruz Biotechnology (Dallas, USA). The decenoyl-CoA, hexadecenoyl-CoA as well as the individual 3(S)-hydroxybutyryl-CoA and 3(R)-hydroxybutyryl-CoA were chemically synthesized as described before [30]. The compound acetoacetyl-CoA was achieved either from Millipore Sigma® Merck KGaA or via chemical synthesis. The substrates hexenoyl-CoA and octenoyl-CoA were enzymatically synthesized as described in the assay for enoyl-CoA reductase (Saci_1115).

Strains and growth conditions

The *E.coli* strains DH5α for cloning and Rosetta (DE3) for expression were cultivated independently in Luria Bertani (LB) medium containing appropriate antibiotics (100 µg/ml ampicillin for strains containing plasmid pET15b or pET45b, 50 µg/ml kanamycin for strains including plasmid pET28b and 50 µg/ml chloramphenicol for Rosetta (DE3), respectively) at 37°C (or under distinct conditions for expression of different proteins) (Table S3).

The *S. acidocaldarius* strain MW001 (uracil auxotrophic) were cultured aerobically in Brock´s basal medium at 75°C, pH 3, 180 rpm [31] supplemented with 0.1% (w/v) N-Z-amine and 0.3% (w/v) D-xylose. The Gelrite-Brock solid medium was prepared as mentioned before [32].

Cloning, expression and purification of the recombinant proteins

Open reading frames (ORFs) *saci_1122, saci_1123*, *saci_0315, saci_1109*, *saci_1104*, *saci_1115* and *saci_1085* were PCR amplified using genomic DNA of *S. acidocaldarius* as template (the employed primer pairs were listed in Table S1) and cloned into distinct plasmid vectors, which were purchased by Novagen, USA (the employed primer pairs were listed in Table S2). The recombinant plasmids were individually transformed into *E. coli* strain Rosetta (DE3) for heterologous overproduction induced by isopropyl-β-d-thiogalactopyranoside (IPTG) or into *S. acidocaldarius* MW001 for homologous expression induced by D-xylose. The detailed information for cloning and expression was described in the supplementary Table S1-S3.

The cell pellets were harvested by centrifugation at 4°C, 7000 rpm for 15 min. Except for the cells maintaining ACAD (*saci_1123*) or ETF (*saci_0315*) protein, the other cell pellets were resuspended in 50 mM NaH_2PO_4 plus 300 mM NaCl at pH 8 in a ratio of 1 g wet cells per 3 ml buffer. Cell suspensions were lysed either by sonication (0.5 cycle, 55 amplitude for 15 min) or via French press (3 times, 1200 psi). After centrifugation (4°C, 15000 rpm, 45 min) the supernatant was collected and protein samples except Saci_1085 were incubated at 70°C in

a water bath for 20 min to perform heat precipitation. The thermostable proteins were then separately from protein precipitations via centrifugation (4°C, 15000 rpm for 30 min). Afterwards, the his-tagged proteins were purified from the supernatant via Protino® Ni-TED (tris-carboxymethyl ethylene diamine) affinity chromatography (Machery & Nagel, Düren, Germany) whereas the strep-tagged Saci_1085 was purified by Strep-Tactin®XT system (IBA Solutions for Life Science, Göttingen Germany) according to the provided instructions. An extra wash step was performed for Saci_1085 using 500 mM NaCl and 0.1% SDS to remove the impure proteins at 70 kDa or 100 kDa and thus obtain pure Saci_1085. The protein samples were then applied to size exclusion chromatography (Superdex 200 prep grad, HiLoad 26/60 or 16/60, GE Healthcare Life Sciences, Freiburg, Germany). A buffer containing 50 mM HEPES/NaOH plus 300 mM NaCl pH 7.2 was applied for equilibrium and elution of the proteins Saci_1122, Saci_1109, Saci_1104 and Saci_1115, separately. For Saci_1114, 50 mM HEPES/KOH pH 7.2 plus 300 mM KCl and 3 mM DTT was used as elution buffer. Finally, all the purified proteins were stored in 25% glycerol at -70°C for further use. The pure Saci_1114 was flush frozen with liquid nitrogen in prior to long-term storage.

The ACAD (Saci_1123) or ETF (Saci_0315) containing cells were resuspended in 50 mM Tris plus 150 mM KCl, 10 mM imidazole and 0.1% Triton X-100 pH 7.5 and were disrupted by sonication. The cell debris was subsequently removed through centrifugation (4°C, 15000 rpm, 40 min) and the supernatant was added with 1 mM FAD to improve the cofactor binding. Purification of ACAD and ETF proceeded *via* metal-ion affinity chromatography, which was carried out with the Äkta PrimePlus (GE Healthcare) system. The crude extracts were filtered (0.45 µM polyvinylidene fluoride membrane, Carl Roth, Karlsruhe, Germany) and applied to 1 or 5 ml Nickel-IDA-Sepharose column according to the user´s manual. The wash buffer contained 50 mM Tris, 150 mM KCl and 10 mM imidazole pH 7.5. The protein samples were eluted out with the same buffer in additional to 400 mM imidazole. Afterwards, imidazole was removed using a 10 or 30 kDa cut-off centrifugal concentrator (Sartorius, Goettingen, Germany). The concentrated pure proteins were stored in 25 mM PIPES, 20 mM KCl, 10% glycerol pH 6.5 followed by flash freezing in liquid nitrogen and stored at -70°C.

Protein concentrations were determined with TECAN reader at 450 nm and 595 nm according to the standard Bradford assay [33]. Purification and molecular weight were visualized by SDS-polyacrylamide gel electrophoresis and the Coomassie Brilliant Blue staining. The purified proteins were employed for further enzyme characterization.

Reconstitution of protein complex between HCDH/ECH Saci_1109 and KT Saci_1114

To study the complex formation between the recombinant HCDH/ECH and KT proteins, 0.015 µmol of each pure protein were mixed together and incubated on ice for 4 hours. Afterwards, the protein mixture was applied to the Superdex 200 prep grad HiLoad 16/60 gel

filtration column (mentioned above). A buffer containing 50 mM HEPES/NaOH plus 300 mM NaCl pH 7.2 was adopted to elute the protein samples.

Enzyme assays

The activities of all the enzymes were determined using two distinct assay systems. The continuous assays were performed via a spectrophotometer while the discontinuous assays were analysed via HPLC. All the reaction mixtures were pre-incubated under the assay conditions for 2 min, afterward reactions were started by addition of the measured enzymes.

Acyl-CoA synthetase (ACS) – ACS activity of Saci_1122 was determined at 55°C in 100 mM HEPES/NaOH pH 7 containing 20 mM $MgCl_2$, 2 mM CoA, 6 mM PEP, 5 mM ATP, 0.2 mM NADH and 6.7 µg ACS enzyme in couple with a series of auxiliary enzymes including 11.4 U myokinase (MK), 4.6 U pyruvate kinase (PK) and 4.2 U lactate dehydrogenase (LDG). All the auxiliary enzymes were extracted from rabbit muscle and purchased from Merck, Darmstadt, Germany. Fatty acids with chain length C2-C12 were tested in a total volume of 500 µl and the assay was monitored via the decrease of absorbance due to NADH (extinction coefficient: 6.22 mM^{-1} cm^{-1}) consumption at 340 nm. The reverse activity of ACS was measured at 70°C, 412 nm by CoA-dependent increase in absorbance due to its colour reaction with DTNB to release yellow 2-nitro-5-thiobenzoate (TNB^{2-}, extinction coefficient: 14.15 mM^{-1} cm^{-1}) [34]. The assay mixture (0.5 ml) contained 100 mM HEPES/NaOH pH 7.5 with 10 µg ACS protein, 1 mM $MgCl_2$, 20 mM PPi, 0.5 mM AMP, 0.1 mM DTNB and 0.5 mM of different acyl-CoAs (acetyl-CoA, butyryl-CoA, octanoyl-CoA and palmitoyl-CoA).

Acyl-CoA dehydrogenase (ACAD) – ACAD activity of Saci_1123 was measured at 65°C, pH 6.5 in 50 mM HEPES/KOH buffer containing 20 mM KCl. Continuous spectroscopic assays were based on the reduction of the artificial electron acceptor ferrocenium ($FcPF_6$) to its reduced species ferrocene at 300 nm (extinction coefficient: 4.3 mM^{-1} cm^{-1} [35]) while the acyl-CoA compounds were thus oxidized. The reaction mixture (500 µl) contained 1 mM $FcPF_6$, 0.13 µg/µl Saci_1123 and 0.4 mM acyl-CoAs (butyryl-CoA, hexanoyl-CoA, octanoyl-CoA and palmitoyl-CoA). Variable concentrations of octanoyl-CoA (0-0.15 mM) were used for determination of V_{max} and K_m. For oxidation of palmitoyl-CoA 2.5% DMSO and 0.1% Triton were added to increase solubility of the substrate. The discontinuous assay (50 µl) including 50 mM MES/KOH, 20 mM KCl, 0.4 mM DCPIP, 0.02 µg/µl ACAD and 0.01 µg/µl ETF with 0.4 mM of acyl-CoAs (butyryl-CoA, hexanoyl-CoA or octanoyl-CoA) were incubated for 5 min. Moreover, 0.8 mM $FcPF_6$ was also used as an electron accpetor instead of ETF and DCPIP.

Electron transfer flavoprotein (ETF) – Determination of ETF Saci_0315 activity was performed by the artificial electron acceptor DCPIP, which exhibits the highest absorbance at 600 nm (extinction coefficient: 21 mM^{-1} cm^{-1} [36]). Discontinuous assays were based on that ACADs

are not able to directly transfer electrons to DCPIP but instead need an intermediate electron acceptor in the form of ETF. The assay mixture (0.5 ml) containing 50 mM MES/KOH, 20 mM KCl, 0.2 mM DCPIP, 0.2 mM of different acyl-CoAs (butyryl-CoA, hexanoyl-CoA or octanoyl-CoA), 1.7 µg ACAD and 3.8 µg ETF was incubated at 65°C, pH 6.5 for 5 min. Afterwards, reduction of DCPIP by ETF was monitored by detecting the absorption spectrum at the wavelength varying from 400 to 800 nm. Moreover, to detect the NADH-linked EtfAB activity of Saci_0315, the assay was performed in 50 mM HEPES/NaOH (pH 7.5) containing 100 mM NaCl, 0.2 mM iodonitrotetrazolium chloride (INT) and 0.015 µg/µl ETF protein with 0-1 mM NADH for K_m measurement. Then the activity was determined by monitoring the release of the red formazan at 500 nm (extinction coefficient 19.3 mM^{-1} cm^{-1}).

3-Hydroxyacyl-CoA dehydrogenase/enoyl-CoA hydratase (HCDH/ECH) – The coupling activity of the bifunctional protein HCDH/ECH Saci_1109 was determined at 75°C in 100 mM Tris/HCl pH 7 with 0.4 mM crotonyl-CoA, 0.2 mM NAD^+ or $NADP^+$ and 0.69 µg protein (total volume 0.5 ml). K_m values for crotonoyl-CoA and NAD^+ were determined by varying the concentrations from 0-1.6 mM and 0-0.3 mM, respectively. The single oxidation step of 3-HBCoA was measured using the commercial mixed 3(S/R)-HBCoA or the synthesized single 3(S)- or 3(R)-HBCoA instead of crotonyl-CoA in the assay. K_m values for mixed 3(S/R)-HBCoA (0-0.7 mM), NAD^+ (0-1.5 mM) or single 3(S)-HBCoA (0-1.25 mM) were measured. The reversed reduction of AcAcCoA was measured at 35°C due to instability of the substrate. The assay solution (0.5 ml) contained 100 mM Tris/HCl pH 7, 0.6 mM AcAcCoA, 0.2 mM NADH or NADPH and 4.05 µg HCDH/ECH enzyme. K_m values for AcAcCoA (0-0.7 mM), NADH (0-1.5 mM) or NADPH (0-1.25 mM) were determined. Discontinuous assay (50 µl) for the combined activity of HCDH/ECH was carried out in 50 mM MES/KOH (pH 6.5 at 65°C), 20 mM KCl, 0.0144 µg/µl protein and 0.4 mM crotonoyl-CoA with or without 2 mM NAD^+ and was incubated for 15 min. Moreover, to measure the combined last three steps of β oxidation, 1.6 mM CoA and 0.054 µg/µl KT were included into the assay.

β-Ketothiolase (also known as acetyl-CoA C-acetyltransferase) (KT) – Cleavage of acetoacetyl-CoA into acetyl-CoA by KT was investigated by monitoring the chelation of AcAcCoA with Mg^{2+} (extinction coefficient of Mg^{2+}- AcAcCoA complex: 21.4 mM^{-1} cm^{-1} [37]) at 23°C, 303 nm under the UV light. A decrease in absorbance should be observed due to the consumption of AcAcCoA. The reaction mixture (0.5 ml) contained 100 mM Tris/HCl pH 8, 20 mM $MgCl_2$, 0.2 mM CoA, 0.1 mM AcAcCoA and 2.7 µg enzyme. Variable concentrations of AcAcCoA (0-0.2 mM) and CoA (0-0.01 mM) were used for identifying K_m values, respectively. Discontinuous assay (400 µl) including 50 mM MES/KOH (pH 6.5), 0.2 mM CoA and 0.1 mM AcAcCoA was pre-incubated at 23°C for 2 min. The reaction was initiated by adding 2.7 µg KT and incubated for further 5 min.

Detection of the reversed activity was coupled with the HCDH/ECH at 75°C. The activity was tested by detecting the NADH oxidation at 340 nm. The reaction components (0.5 ml) contained 100 mM MOPS/NaOH pH 6.5, 0.3 mM NADH, 17.1 µg ECH/HCDH, 10.8 µg KT and 0.075-7.5 mM acetyl-CoA was included for determining V_{max} and K_m. Furthermore, 0.5 mM of distinct acyl-CoAs (acetyl-CoA, butyryl-CoA, hexanoyl-CoA, octanoyl-CoA, lauroyl-CoA or palmitoyl-CoA) was added to the assay in addition to 2.5 mM acetyl-CoA to investigate the substrate specificity. Reaction mixture for discontinuous assay (400 µl) contained 100 mM MOPS/NaOH buffer pH 6.5 with 1 mM acetyl-CoA, 0.3 mM NADH, 17.1 µg HCDH/ECH and 1 mM of distinct acyl-CoA (acetyl-CoA, butyryl-CoA, hexanoyl-CoA or octanoyl-CoA) and was incubated at 65°C for 2 min. Afterwards 16.2 µg KT was added to initiate reaction followed by further incubation for 5 min.

β Oxidation enzyme cascades for HPLC analysis – The complete enzyme cascades were performed in two steps. The first oxidation step by ACAD and ETF was done as above described. In the second step, 2 mM NAD^+, 0.0144 µg/µl HCDH/ECH, 1.6 mM CoA and 0.054 µg/µl KT were successively added followed by incubation for 15 min.

Acetoacyl-CoA reductase (ACR) – ACR activity was determined according to the AcAcCoA dependent oxidation of NADH/NADPH at 340 nm, 35°C. The assay mixture included 100 mM Tris/HCl pH 7, 0.3 mM AcAcCoA, 0.2 mM NADH/NADPH and 4.03 µg protein. The concentration of AcAcCoA or NADPH was varied from 0-0.5 mM or 0-0.2 mM, respectively, for K_m determination while concentrations of the other components were kept constant. Reversed activity of ACR was tested at 70°C using 3-HBCoA as substrate and 2 mM NAD^+/$NADP^+$ as cofactor in 100 mM Tris/HCl (pH 7) with 20.16 µg protein. Concentrations of 0-2 mM were used for K_m measurement for the mixed 3(S/R)-HBCoA while 0-1 mM for single 3(R)-HBCoA in presence of 2 mM $NADP^+$.

MaoC like 3-hydroxyacyl-CoA dehydratase (MaoC-HCD) – The activity of MaoC-HCD Saci_1085 was tested at 65°C, pH 6.5 via a discontinuous assay (50 µl) containinig 50 mM MES, 20 mM KCl and 0.0675 µg/µl protein. K_m for 3(R)-hydroxybutyryl-CoA was determined by varying its concentration from 0-0.5 mM while a variable concentration of 0-1 mM was used for measuring K_m for crotonoyl-CoA. Activities toward substrates with different chain lengths (C4, C6 or C8) (C) were determined by incubating 0.09 µg/µl protein with 0.4 mM enoyl-CoA namely crotonoyl-CoA, hexenoyl-CoA or octenoyl-CoA. Here, the substrate hexenoyl-CoA or octenoyl-CoA was enzymatically produced by ACAD, ETF and DCPIP as above mentioned. To study the stereochemical specificity towards the 3-hydroxyacyl-CoA intermediates, 0.4 mM of individual 3(S)- or 3(R)-hydroxybutyryl-CoA was used and the reaction was run for 30 min. Afterwards, the reaction was stopped by mixing the sample with acetonitrile in a ratio of 1:3

(v/v) at different time points and then freezing the mixture. The formation of the relevant product was analysed via HPLC and thus the specific activities were calculated.

Enoyl-CoA reductase (ECR) – ECR activity was checked by monitoring NADH/NADPH consumption at 340 nm. The enzymatic assay (500 µl) was done in 100 mM HEPES/NaOH pH 7.5 at 70°C and included 10 mM KCl, 0.3 mM NADH/NADPH, 0.4 mM crotonyl-CoA and 12 µg ECR. Moreover, 0-5 mM crotonoyl-CoA or 0-0.04 mM NADPH was applied to determined K_m values, respectively. The activity for oxidation of butyryl-CoA was checked with 0.2 mM butyryl-CoA as substrate and 1 mM NAD^+/$NADP^+$ as cofactor under the same assay conditions (500 ul).

Substrate spectrum of ECR was also checked with different CoA esters. The unsaturated hexenoyl-CoA or octenoyl-CoA was enzymatically synthesized by ACAD (Saci_1123) from the saturated derivatives. The reaction mixture contained 100 mM HEPES/NaOH (7.5, 70°C), 0.3 mM of hexanoyl-CoA or octanoyl-CoA, 0.6 mM of $FcPF_6$, then 7.54 µg ACAD (Saci_1123) was introduced to produce the corresponding enoyl derivatives. The continuous assay for reduction of the enoyl-CoAs with the carbon chain length of C4, C6, C8, C10 or C16 was performed in 100 mM HEPES/NaOH (7.5, 70°C) in presence of 0.2 mM NADPH, 10 ug protein and 0.3 mM enoyl-CoA. K_m for crotonoyl-CoA or NADPH was determined by using a variable concentration of 0-0.5 mM or 0-0.04 mM, respectively.

FA synthesis enzyme cascades for HPLC analysis – To investigate the conversion of 3(R)-hydroxyacyl-CoA into acyl-CoA, 0.4 mM 3(R)-hydroxybutyryl-CoA (3(R)-HBCoA) was incubated with 0.84 µg MaoC-HCD Saci_1085, 1.23 µg ECR Saci_1115 and 2 mM NADPH in 50 µl of 50 mM MES/KOH plus 20 mM KCl at 65°C, pH 6.5 for 30 min. Furthermore, time dependent formation of crotonoyl-CoA or butyryl-CoA was monitored by analysing samples at different time points (0, 30, 60, 120, 180 and 240 min). For reconstruction of a potential FA synthetic pathway for producing butyryl-CoA from acetoacetyl-CoA (AcAcCoA), 0.4 mM AcAcCoA, instead of 3(R)-HBCoA was introduced to the same assay as the initial substrate and additional 1.04 µg of ACR Saci_1104 was included. Then the samples were analysed by HPLC.

***In vitro* analysis of the FA metabolic reactions via HPLC**

Acyl-CoA esters were extracted from β oxidation reaction mixtures and meanwhile proteins were removed through cold acetone precipitation. The reaction mixtures were mixed with acetone (1:3 v/v), and then incubated at -80°C for 20 min or at -20°C overnight. Afterwards, the samples were centrifuged at 4°C for 30 min to precipitate proteins. The supernatants were transferred in clean tubes and the liquid was evaporated completely at 30°C using an Eppendorf™ Concentrator Plus with Membrane Vacuum Pump. The precipitation was

resuspended in ultrapure water and applied to the Thermo Scientific UltiMate 3000 HPLC system (Thermo Fisher Scientific US). Detection and separation of CoA esters were accomplished via a reversed phase NUCLEODUR C18 Pyramid analytical column (MACHEREY-NAGEL GmbH & Co. KG, Germany). Acyl-CoAs with different chain lengths were analyzed with different concentration gradients of acetonitrile (ACN). On one hand, for shorter chain CoA esters ($C \leq 4$), the analytical column was pre-equilibrated with 96% of Buffer A (0.2 M ammonium acetate, pH 5) and 4% of Buffer B (ACN), CoA and CoA esters were eluted employing three linear gradients of ACN: 0-20 min, 4-7% ACN; 20-35 min, 7-30% ACN; 35-35.5 min, 30-4% followed by an isocratic flow with 4% Buffer B for 9.5 min. In the current study, this program for separating shorter chain CoA compounds was regarded as the “4-30% ACN” program. On the other hand, to detect longer chain CoA esters ($C \geq 4$), the column was pre-equilibrated with 99% of Buffer A and 1% of Buffer B, then acyl-CoA compounds were eluted firstly with a linear gradient of 1 to 10% ACN for 5 min followed by an isocratic flow with 10% ACN for 16 min, then two linear gradients of 10 to 60% for 20 min and 60-1% for 0.5 min were used. Afterwards, the system was run with 1% ACN for further 19.5 min. Here, the program for detecting longer chain CoA esters was regarded as “1-60% ACN” program. All chromatography was carried out at 8°C and the flow rate was 1 ml/min. Data collection and processing was done by Thermo Scientific™ Dionex™ Chromeleon 7 Chromatography Data System (CDS) Software (Thermo Fisher Scientific US).

Results

It has recently been demonstrated, that FAs serve as sole carbon and energy source for growth of *S. acidocaldarius* on minimal media [8]. To further study the FA metabolism we first analysed the enzymes encoded in the *saci_1103-1126* gene cluster and their functional role in β oxidation. The respective genes were cloned and heterologously expressed in *E. coli*, and the recombinant proteins were purified (Fig. 1B) and biochemically characterized (Tab. 1), including the HPLC based determination of the respective CoA-ester intermediates as well as the essential cofactor NAD^+/NADH via coupled enzymatic assays (Fig. 1 C). The first step in FA degradation (after likely passively entering the cell) is their activation to the corresponding CoA esters usually catalysed by AMP-forming acyl-CoA synthetases (FA + ATP + HS-CoA → FA-CoA + AMP + PP_i) [15, 38]. Three homologues (*saci_1111, saci_1122*, and *saci_1126*) are encoded within the gene cluster and the coding function of *saci_1122* was confirmed. The recombinant homodimeric protein (61 kDa denaturing, 117 kDa native conditions) catalyzes the HS-CoA and FA dependent conversion of ATP yielding AMP and PPi as product with the highest activity towards chain lengths between C5 and C8 (Fig. S9) and the kinetic constant were determined with valeric acid or octanoic acid as substrate (Tab. S5, Fig. S8). The AMP- and PPi-dependence of the enzyme was confirmed in both directions of the reaction (Tab. S5). No activity was observed with FAs longer than C10 identifying the enzyme as medium chain acyl-CoA synthetase (MACS) (Fig. S9).

Following the FA activation to the corresponding CoA ester, the acyl-CoA then enters the β oxidation and in the first step the saturated acyl-CoA esters are oxidized to yield the corresponding enoyl-CoA esters. The sequences of the three acyl-CoA dehydrogenase homologues catalyzing this reaction (*saci_1108, saci_1113, saci_1123*) were thoroughly analysed revealing that Saci_1108 (missing adenosine binding site) and Saci_1113 (no CoA- and isoalloxaxine-binding sites, no catalytic base) lack catalytically essential features (Fig. S4). Accordingly, both recombinant enzymes were inactive. However, Saci_1123 harbours all relevant sequence features (Fig. S4) and catalysed the acyl-CoA oxidation with ferrocenium as artificial electron acceptor with highest activity towards octanoyl-CoA (V_{max} 29 U mg^{-1}, K_m 0.0151 mM), and to a lesser extend also with hexanoyl-CoA and butyryl-CoA. Palmitoyl-CoA (C16) was not converted classifying the enzyme as medium chain specific, fitting well to the results of the sequence analyses (Tab. 1) (Fig. S4). The yellow colour of the enzyme preparation already indicated that the homotretrameric 170 kDa protein (subunits 43.9 kDa) (Tab. 1) contains FAD as a cofactor and the content and function for catalysis was shown spectroscopically indicating that 62% of the purified protein contained an FAD cofactor and from that one FAD cofactor per monomer was concluded. However, the native electron acceptor for the reduced FAD cofactor is usually the electron transferring flavoprotein (ETF) which transfers the electrons to the membrane bound ETF dehydrogenase finally reducing

quinones of the respiratory chain. In *S. acidocaldarius* two homologues for the ETF (*saci_0315*, and *saci_0290/0291*) were identified and both form operon like structures with genes encoding ETF dehydrogenases (*saci_0316/0317*, *saci_0292/0293*). The *saci_0290/0291* encodes an ETF comprised of two different subunits whereas *saci_0315* represents a fusion of both genes encoding one single fusion protein. Saci_0315 was recombinantly produced in *E. coli*. The purified 68 kDa monomeric protein (calculated 66.9 kDa) could be shown to take over the electrons from and thus re-oxidize the FAD cofactor of the ACAD and further transport the electrons to DCPIP (Fig. S11). A direct reduction of a quinone derivative (Q10) by ETF could not be observed. This shows that the ETF acts as native electron acceptor for the ACAD and that the ETF dehydrogenase is a further essential part of the electron transfer from the saturated acyl-CoA finally to the quinones of the respiratory chain. However, an electron transfer from NAD(P)H to ETF or further to the ACAD was not observed under aerobic conditions suggesting that ACAD/ETF only work in the oxidative direction.

For the next two steps in the β oxidation, e.g. the hydration of the enoyl-CoA to the 3(S)-hydroxyacyl-CoA and its oxidation to 3-oxoacyl-CoA, one fused gene homologue *saci_1109* is present in the *saci_1103-1126* gene cluster encoding a single polypeptide chain with an N-terminal hydroxyacyl-CoA dehydrogenase and a C-terminal enoyl-CoA hydratase domain (HCDH/ECH). The molecular mass of the purified recombinant protein composed of 73 kDa subunits could not unequivocally be determined due to the formation of higher oligomerization states with molecular masses above 500 kDa as judged by size exclusion chromatography and MS (Tab. 1). However, according to its fused structure the protein showed the combined activity, i.e. the crotonyl-CoA dependent NAD^+ reduction to NADH, with a V_{max} of 17 U mg^{-1} and K_m values of 24 µM (crotonyl-CoA) and 36 µM (NAD^+) (Tab. 1). The enzyme also catalysed the conversion of decenoyl-CoA (C10:1) (V_{max} around 45% of that with crotonyl-CoA) but was not active with hexadecenoyl-CoA (C16:1) indicating a medium chain length specificity (Tab. 1, Fig. S13C). Moreover, the enzyme catalysed the reversible oxidation of 3(S)-hydroxybutyryl-CoA to acetoacetyl-CoA with a V_{max} of 48 U mg^{-1} and was specific for the (S)-enantiomer (K_m 43 µM) while the (R)-enantiomer was not converted. Additionally, no activity could be observed using the co-substrate $NADP^+$ demonstrating the NAD^+–dependent property of the recombinant protein. Furthermore, also in the reductive direction Saci_1109 showed a clear preference for NADH (K_m 28 µM) over NADPH (K_m 94 µM) (Tab. 1). The V_{max} (7 U mg^{-1}) was similar for both cosubstrates (in the reductive direction, the reaction had to be measured at 35°C due to the pronounced instability of acetoacetyl-CoA), although at higher NADPH concentrations a pronounced inhibition was observed which did not occur with NADH (Fig. S12G, H).

The final reaction in the β oxidation is catalysed by the β-ketothiolase/acetyl-CoA C-acetyltransferase (KT) and two homologues are present in the *saci_1103-1126* gene cluster (*saci_1114* and *saci_1121*). The coding function was confirmed for the purified recombinant Saci_1114 in both direction of acetoacetyl-CoA cleavage and the kinetic parameters were determined for acetoacetyl-CoA (V_{max} 1.7 U mg^{-1}, K_m 30 µM), CoA (V_{max} 2.5 U mg^{-1}, K_m 3 µM) and acetyl-CoA (V_{max} 2.7 U mg^{-1}, K_m 2.1 mM), respectively (Tab. 1). The low K_m values for acetoacetyl-CoA and CoA as well as the relatively high K_m for acetyl-CoA already suggest that the thiolytical cleavage of the 3-ketoacyl-CoA is the physiologically favorable direction. The native molecular mass of 88 kDa and a subunit size of 43 kDa of Saci_1114 suggested a homodimeric structure. It has been reported that in Bacteria (i.e. *Escherichia coli*, *Psuedomonas fragi* or *Mycobacterium tuberculosis*) the homologues of this β-ketothiolase often form complex with the bifunctional protein involved in β oxidation [13, 39, 40]. In order to reconstitute the trifunctional protein complex of *S. acidocaldarius* β oxidation in vitro, equal molar amount of HCDH/ECH and KT was mixed and analysed via gel filtration, as a result there was no complex formation between these two proteins indicating a distinct mechanism of the *S. acidocaldarius* β oxidation enzymes (Fig. S15). To elucidate the chain length specificity of Saci_1114 ketothiolase it was monitored if the addition of substrates with different chain length to an assay only containing 2.5 mM of acetyl-CoA could enhance the activity. An activity increase was observed with acetyl-, butyryl-, hexanoyl-, and octanoyl-CoA again indicating a short to medium chain length specificity (Fig S14D).

All four enzyme activities were also confirmed by HPLC following the consumption and formation of the respective CoA ester substrates and intermediates, respectively. First, the complete conversion of butyryl-CoA to crotonoyl-CoA as well as of hexanoyl- and octanoyl-CoA to the respective enoyl-CoA esters by ACAD was shown with $FcPF_6$ or the ETF as electron acceptors which then transfers the electrons further to DCPIP (Fig. S22). For the bifunctional HCDH/ECH Saci_1109 it could be demonstrated that crotonoyl-CoA is converted by >90% to hydroxybutyryl-CoA although the further conversion with NAD^+ to acetoacetyl-CoA could not be observed (Fig. S23). However, when the KT was added to the Saci_1109 reaction the nearly complete conversion of crotonoyl-CoA to acetyl-CoA could be detected without any detectable amounts of liberated acetoacetyl-CoA (and only low amounts of 3-hydroxybutyryl-CoA) (Fig. S25). In addition, the complete thiolysis of acetoacetyl-CoA into acetyl-CoA by the KT Saci_1114 alone was observed (Fig. S24). Also, the reversible conversion of acetyl-CoA to crotonoyl-CoA using Saci_1114 and Saci_1109 was demonstrated although only low amounts of crotonoyl-CoA were observed (and also 3-hydroxybutyryl-CoA as intermediate) (Fig. S27A). Finally, the complete β oxidation was reconstituted by first incubating the respective saturated acyl-CoA esters (C4-C8) with the ACAD and ETF/DCPIP for 5 min leading to complete conversion to the respective enoyl esters. Upon addition of HCDH/ECH and the

KT as well as the required coenzymes NAD^+ and HS-CoA further formation of acetyl-CoA (Fig. 1C), acetyl- and butyryl-CoA (Fig. S26A), as well as acetyl- and hexanoyl-CoA (Fig. S26B) was observed, respectively. The additional peaks visible in the chromatograms could not undoubtedly be identified but likely correspond to longer chain hydroxyacyl or ketoacyl intermediates, respectively. Together, the results show that the saturated acyl-CoAs can completely be converted to acetyl-CoA by the reconstituted β oxidation enzyme cascade and that the whole pathway conversion is sufficiently exergonic to run the acyl-CoA conversion to acetyl-CoA to completion with the ACAD as the major driving force. The other three reactions were shown to operate reversibly in vitro also as a three-activity-cascade although with an equilibrium far on side of the acetyl-CoA formation, i.e. the oxidative/cleavage direction.

Interestingly, in the *saci_1103-1126* gene cluster a gene (*saci_1115*) was identified putatively encoding a dehydrogenase of the medium chain dehydrogenase/reductase (MDR) superfamily which showed remote sequence similarity (59.46%) to the acryloyl-CoA reductase from *Metallosphaera sedula*. We heterologously overproduced and purified the 69 kDa homodimeric enzyme (subunit size 36 kDa) and its characterization revealed crotonoyl-CoA reductase activity with NADPH as electron donor (V_{max} 0.422 U mg^{-1}, K_m (crotonoyl-CoA) 96 µM, K_m (NADPH) 7 µM) (Tab. 2). Also, NADH could serve to reduce crotonoyl-CoA although with 4.5 fold reduced activity compared to NADPH. The enzyme also showed activity towards hexenoyl-CoA, octenoyl-CoA and decenoyl-CoA (C6, C8 and C10 2-enoyl-CoAs) with the highest specific activity for octenoyl-CoA (0.93 U mg^{-1}) (Tab. 2, Fig. S19). This indicates that Saci_1115 is a medium chain specific enoyl-CoA reductase with a clear preference for NADPH.

The NADPH specificity is regarded as distinctive property of the FA synthesis compared to the NAD-specific β oxidation (see above) [23, 24]. Since the HCDH/ECH (Saci_1109) showed a clear preference for NAD^+, we sought for the presence of NADPH dependent ketoacyl thioester reductases which might be operative in FA synthesis in *S. acidocaldairus*. These ketoacyl thioester reductases of the canonicle FA synthesis pathways usually belong to the SDR superfamily [23]. The *S. acidocaldarius* genome harbors in total 11 SDR homologues which are annotated as fabG. One of these SDR homologues, e.g. Saci_1104, is also encoded in the *saci_1103-1126* gene cluster and was therefore characterized after heterologous overexpression in *E. coli* followed by purification. The recombinant Saci_1104 catalyzed the reversible and strictly NADPH dependent interconversion of acetoacetyl-CoA to 3(R)-hydroxybutyryl-CoA. No activity was obtained neither with NADH nor with the 3(S)-hydroxybutyryl-CoA stereoisomer. The kinetic constants were determined in both directions of the reaction. In the direction of 3(R)-hydroxybutyryl-CoA formation the V_{max} was 1.3 U mg^{-1} and the K_m values were 0.077 mM and 0.024 mM for AcAcCoA and NADPH respectively (measured at 35°C due to the instability of AcAcCoA). For 3(R)-hydroxybutyryl-CoA the K_m

was 0.16 mM and the V_{max} 0.97 U mg^{-1} (measured with 2 mM $NADP^+$ at 70°C) (Tab. 2). Thus, the catalytic efficiency in the reductive direction was already much higher although measured at much lower temperature compared to the oxidative direction. This strict stereospecificity for the 3(R)-hydroxyacyl intermediates is also a known characteristic trait of FAS systems compared to the (S) specificity of the β oxidation [23]. And the presence of an NADPH dependent reductases with (R)-hydroxyacyl-CoA specificity, raised the questions whether *S. acidocaldarius* also harbor (R)-specific hydroxyacyl thioester dehydratases. In Bacteria and Eukarya this reaction is carried out by dehydratases from the hotdog fold superfamily [11], which were not present in the *S. acidocaldarius* genome. Only distantly related MaoC dehydratases which also belong to the hot dog fold superfamily were identified (*saci_1070, saci_1085*). The recombinantly overproduced and purified Saci_1085 formed homo-dodecameric structures (native mass 244 kDa, calculated subunit size 21 kDa) and indeed showed 3-hydroxyacyl-CoA dehydratase activity with a pronounced preference for the (R)-stereoisomer (Tab. 2, Fig. S29A). The enzyme could be measured in both directions of the reaction and the kinetic constants were determined for conversion of 3(R)-hydroxybutyryl-CoA (V_{max} 1.72 U mg^{-1}, K_m 0.4 mM) or crotonoyl-CoA (V_{max} 4.4 U mg^{-1}, K_m 0.22 mM). Interestingly, the enzyme preferred C8 enoyl-CoA over C4 or C6 derivatives, which is similar to Saci_1115 (Fig. S20C). The formation of crotonoyl-CoA from 3(R)-hydroxybutyryl-CoA was also shown via HPLC analyses and addition of the enoyl-CoA reductase (Saci_1115) led to the formation of butyryl-CoA (Fig. S29). Furthermore, with the reconstitution of these two enzymes with the Saci_1104 ketoacyl-CoA reductase the conversion of acetoacetyl-CoA to butyryl-CoA could be shown (Fig. 2C). The conversion of acetoacetyl-CoA was complete in the reported time frame whereas 3(R)-hydroxybutyryl-CoA was only partially converted to crotonoyl-CoA and butyryl-CoA, respectively, suggesting that the final enoyl-CoA reduction do not have the expected “pulling” character for the whole enzyme cascade. However, these results together with the reversal of the ketothiolase putatively enabled by the DUF35 scaffolding protein (see below in the discussions section) indicate a potential novel FA synthesis pathway in *S. acidocaldarius*.

Discussion

β oxidation

It has recently been shown that *S. acidocaldarius* grows on FAs as sole carbon and energy source [41] and this organism belongs to those archaeal species harboring full sets of β oxidation homologues (with several paralogous copies for each of the four steps). Also, for e.g. halophilic Archaea and for *Archaeoglobus* species the presence of a complete β oxidation pathway has been reported on sequence level and also the growth on FAs has been indicated [3-5]. This suggests that indeed the presence of these genes correlates with the ability to utilize FAs as carbon and energy source. We here confirmed the coding function of a FA activation protein and several of these β oxidation related genes in *S. acidocaldarius* organized in a recently identified gene cluster, i.e. *saci_1103-1126* [41], and characterized them as ACS (Saci_1122), ACAD (Saci_1123), bifunctional HCDH/ECH (Saci_1109) and KT (Saci_1114). This gene cluster including the characterized encoded enzymes was shown to be repressed by a TetR like regulator which is released from its DNA target sequences upon binding of medium- to long chain acyl-CoA esters (C6-C18) [41]. This at least partly correlates with the substrate specificity of the characterized proteins which all show a preference for medium chain fatty acid/acyl intermediates. This also applies for the two esterase enzymes encoded in the same cluster which were characterized previously (*saci_1105, saci_1116*) [42, 43] and cover a medium chain acyl length spectrum with the highest activity. However, the presence of further genes encoding additional paralogous copies of the respective enzymes might suggest that the substrate spectrum of the whole cluster is probably extended also to longer chain substrates.

At first, the *saci_1122* encoding protein has been characterized as an AMP-forming medium-chain ACS for FA activation. To our knowledge, experimental data about the medium-chain ACS (MACS) from Archaea especially Sulfolobales is limited, instead many characterized archaeal ACSs stick to short chain FA substrates like acetate and propionate [44-47]. In contrast the *E. coli* fadD homologue is responsible for activating long chain FAs (C ≥12) [38]. A MACS from *Methanosarcina acetivorans* (MaACS) was reported to favor branched, medium chain FA like 2-methylbutyrate [48] while the *Pyrobaculum aerophilum* and *Archaeoglobus fulgidus* enzymes were regared as unusual due to their utilization of longer chain FAs than acetate and propionate, i.e. butyrate, isobutyrate or valerate [34, 44], especially, the PaACS exhibited octameric structure, unlike the other monomeric or homodimeric ACSs [34]. Compared with these ACSs, the homodimeric Saci_1122 did not convert the branched isovaleric acid but showed highest activity toward straight chain FAs ranging from C5 to C8 (Fig. S9C). This is in line with the definition of MACS which activates FAs with C4 to C12 [49].

However, the β oxidation enzymes from Archaea have so far not been characterized in detail and surprisingly little is known about the acyl-CoA dehydrogenases. The ACADs belong to a large superfamily of flavoproteins and oxidize saturated acyl-CoAs to the corresponding unsaturated 2,3-enoyl-CoA thioesters and represent key enzymes of β oxidation and amino acid metabolism [26, 50, 51]. The primary electron acceptor is a non-covalently enzyme-bound flavin adenine dinocluotide (forming a charge transfer complex with the substrate) from which the electrons are further transferred to the electron transferring flavoprotein. This is in turn reoxidized by the membrane bound ETF: quinone oxidoreductase which reduces the quinone pool of the respiratory chain [27], in *S. acidocaldarius* the caldariellaquinone [52]. Finally, the electrons are transferred to oxygen in *S. acidocaldarius* as terminal electron acceptor. In this study the ACAD Saci_1123 from the *S. acdiocaldarius saci_1103-1126* gene cluster was characterized as medium chain ACAD which fits quite well to the sequence analyses (conserved catalytic base and tyrosine residue specific for short and medium chain ACAD, Fig. S4). Also, the highest activity towards octanoyl-CoA, the FAD content per monomer of likely one, as well as the homotetrameric structure are well in line with known medium chain ACADs (MCADs) [26]. The only archaeal ACAD reported so far is the isovaleryl-CoA dehydrogenase from *Halobacterium salinarum* which was however not biochemically characterized in detail [53]. Also, from Bacteria, only little biochemical information is available. MCADs are best characterized from mammalian like e.g. *Homo sapiens* or *Rattus norvegicus* and the kinetic constants with K_m values for the octanoyl-CoA in the micromolar range (2-8 µM) and V_{max} of 10-25 U mg^{-1} (Brenda database, 2021) are at least in a similar range as observed for the *S. acidocaldarius* enzyme. However, the physiological electron acceptor for most of the ACADs is the ETF and two paralogues have been identified in the *S. acidocaldarius* genome (see above) and we could experimentally show that the Saci_0315 is reduced by the MCAD Saci_1123. Saci_0315 represents a fusion protein of the β and α subunit (N- and C-terminal, respectively) of known ETFs [54] (Fig. S5) and the monomeric structure is thus in line with the reported heterodimeric (αβ) structures of the latter. BLAST searches revealed that these fused ETFs are restricted to Archaea and occur mainly in the Sulfolobaceae [54] and some Thermoplasmatales. They mostly cooccur in gene clusters with the genes encoding the ETFCX (annotated as fixCX) ETF: quinone oxidoreductase [54] which thus likely funnels electrons from the ETF to the caldariellaquinone pool of *S. acidocaldarius*. Many Sulfolobaceae species habour a second ETF encoded by separated α and β genes and it might also be involved in FA degradation. It is well known that ETFs can accept electrons from a variety of ACADs [26] and thus a function in e.g. amino acid metabolism is also likely considering that most Sulfolobaceae also grow with proteinaceous substrates as carbon and energy source [10]. However, the distinct function of both ETFs remains to be elucidated. Interestingly, as revealed by bioinformatics analyses, most of the archaeal lineages containing

complete sets of β oxidation enzymes also contain such ETF encoding genes whereas those not containing ETFs do also not contain a complete FA degradation machinery. This indicates and may thus confirm a functional interconnection of these ETFs and FA catabolism. Saci_0315 shows around 40% sequence identity to the recently characterized *Pyrobaculum aerophilum* homologue [55]. And accordingly, Saci_0315 also catalyzed the reduction of iodonitrotetrazolium (INT) with a specific activity around 0.703 U mg^{-1} which is only 1.68% of that for the *P. aerophilum* ETF (41.8 U mg^{-1}) [55]. Furthermore, first results obtained by protein denaturation and concomitant spectrophotometric determination of FAD and HPLC analyses of the AMP in the supernatant indicated that the Saci_0315 fusion ETF contains most likely two FAD and no AMP, which is a characteristic feature of bifurcating ETFs (data not shown). Also, phylogenetic analyses showed that both ETFs cluster together with other crenarchaeal ETFs which appear more closely related to bifurcating ETFs than to non-bifurcating ones [54, 56]. In *P. aerophilum* the whole ETFABCX complex was shown to bifurcate electrons from NADH to ferredoxin and to menaquinone, respectively [55]. Thus, these results obtained for the Saci_0315 ETF suggest that also the complete Saci_0315-0317 ETFABCX catalyzes electron bifurcation from NADH to ferredoxin and the caldariellachinone, respectively. The physiological role and significance of this process in these aerobic to microaerophilic organisms appears not fully established. For *P. aerophilum* it has been discussed that the complexation of NAD^+ to the ETF might prevent electron transfer to oxygen forming harmful superoxide [55]. However, we herein clearly showed that the ETF functions as electron acceptor for the ACAD and most likely further transfers these electrons via the ETFCX oxidoreductase to the caldariellaquinone.

As already mentioned above, there is some overlap of the β oxidation to other metabolic pathways like the 3HP/4HB and the dicarboxylate/4HB cycles for CO_2 fixation found in Cren- and Thaumarchaeota, i.e. the conversion of crotonoyl-CoA to two acetyl-CoA in the 4HB part of both pathways [57, 58]. Also, degradation pathways of aromatic compounds involve the crotonoyl-CoA to acetyl-CoA conversion as well as other analogous reactions as described in Archaea for the Euryarchaeon *Ferroglobus placidus* [59]. Herein, for Saci_1109, the conversion of crotonoyl-CoA to acetoacetyl-CoA via 3(S)-hydroxybutyryl-CoA was demonstrated. Also, in *Metallosphaera sedula* (CO_2 fixation) and *F. placidus* this conversion has been reported to be catalyzed by a similar bifunctional enzyme like Saci_1109 comprised of an N-terminal HCDH and a C-terminal enoyl-CoA hydratase (crotonase superfamily) domain [60, 61]. However, the activity with other substrates than crotonoyl-CoA was not analyzed for the latter proteins as shown for the Saci_1109 which was active also with straight chain 2,3-decenoyl-CoA indicating a broader function in the β oxidation of various FAs of at least medium chain length. The combined specific activity, i.e. the conversion of crotonoyl-CoA to acetoacetyl-CoA with NAD^+, of 17 U mg^{-1} for Saci_1109 was in a similar range as reported for

the other two archaeal enzymes (8-35 U mg^{-1} for *M. sedula* and 6 U mg^{-1} for *F. placidus*) [59-61] although the K_m values were somewhat lower for the *S. acidocaldarius* enzyme. Only recently, the *M. sedula* enzyme was reevaluated, also showing activity towards the medium chain C8 enoyl-CoA substrate and a function in β oxidation rather than in CO_2 fixation was discussed. Interestingly, sequence comparison revealed that in archaea nearly exclusively the N-terminal HCDH/C-terminal ECH domain organization occurs if such fusion proteins are present. In contrast, in the canonical β oxidation known so far from bacteria and mitochondria the crotonoyl- to acetoacetyl-CoA conversion is also carried out by a fusion protein which however show an inverted domain organization [62]. As mentioned above, the dehydrogenase activity of *S. acidocaldarius* (30 U mg^{-1}) is around double of that for the hydratase (17 U mg^{-1}) while for the *M. sedula* enzyme, both activities were shown almost to be equal (16 U mg^{-1} for the dehydrogenase and 20 U mg^{-1} for the hydratase). However, the hydratase activity of the *E. coli* β oxidation complex was reported to be 5-10 fold higher than the dehydrogenase activity [63, 64]. This differences between the hydratase and dehydrogenase activities of the fusion protein occurred in Archaea and Bacteria might also be caused by the distinct domain structures. The Saci_1109-like domain organization is also quite widespread in bacteria but only one such homologue has been characterized so far from *Cupriavidus necator* [65]. However, the ratio behind this domain rearrangement has not been analyzed. The canonical enzymes e.g. from *Pseudomonas fragi* or *E. coli* are known to form a complex with the ketothiolase catalyzing the next step in the β oxidation spiral enabling substrate channeling [66-68]. First results obtained from gel filtration analyses of Saci_1109 with the Saci_1114 ketothiolase similarly carried out as for the *E. coli* complex [67], indicated that in contrast to the *E. coli* complex the *S. acidocaldarius* enzymes could not be reconstituted *in vitro* into a protein complex. Although it cannot be ruled out that Saci_1109 forms complexes with other ketothiolase paralogues than Saci_1114, the results might indicate that the altered domain structure prevent or alter complex formation. [69]. The substrate channeling enabled by complex formation appear reasonable in the last three reactions of the β oxidation because, although the overall thermodynamics of the three reactions is exergonic by -11.9 kJ mol^{-1} (with the thiolase as major driving force of -26.1 kJ mol^{-1}), the NAD^+ dependent hydroxyacyl-CoA oxidation is endergonic by +18 kJ mol^{-1} and represents the energetic bottle neck of the β oxidation spiral [29]. However, in the mitochondrial pathway except for the membrane bound long chain specific complex the reaction sequence is carried out by single enzymes which do not form protein complexes [70, 71].

The ketothiolase Saci_1114 was active as a stand-alone enzyme whereas the *E. coli* enzyme is only active as part of the complex [67]. The specific activity of Saci_1114 with 1.7 U mg^{-1} was much lower than previously reported for the *M. sedula* and *Pyrobaculum neutrophilum* enzymes (Msed_0656, 141 U mg^{-1}; Tneu_0249, 55 U mg^{-1}) [60, 61]. However, due to the

pronounced instability of acetoacetyl-CoA observed herein the assay temperature was lowered to 23°C and the activity at the physiological temperature of 75°C can be roughly estimated to 55 U mg^{-1} using the van't Hoff rule. With the K_m value of the Saci_1114 ketothiolase of 0.033 mM for acetoacetyl-CoA it indicates a much higher catalytic efficiency than the *M. sedula* or *P. neutrophilum* (0.15-0.18 mM) enzymes. Thiolases are subdivided into two main groups the degradative and biosynthetic thiolases [72]. The degradative thiolases are characterized by a broader substrate spectrum also converting longer acyl chain substrates whereas biosynthetic thiolases are specific for short chain acyl-CoAs ≤C4. Although recently the structural basis for these different substrate spectra has been indicated [72], it appears difficult to deduce the biosynthetic or degradative function of a thiolases simply by sequence. However, biosynthetic archaeal ketothiolases seem to cooccur downstream with a DUF35 domain encoding gene (see below) and the DUF35 domains were shown to act as scaffold protein in complex formation [73]. This has been shown for the ketothiolase and the 3-hydroxy-3-methylglutaryl-CoA synthase from *Methanothermococcus thermolithotrophicus* and has also been indicated for haloarchaea [73]. However, the DUF35 domain protein encoding gene is not present downstream of *saci_1114* which might therefore support the degradative function of Saci_1114. The archaeal ketothiolases have been indicated as phylogenetically ancestral [74]. Furthermore, both degradative and biosynthetic archaeal ketothiolases appear more related to each other than to their bacterial (and perhaps also eukaryotic) counterparts [5, 20], which might further explain, together with the inverted domain structure of the HCDH/ECH fusion proteins, the differences in complex formation.

From the single enzymes the β oxidation cascade was reconstituted (Fig. 1) in a stepwise manner and the substrate and product formation was followed via HPLC analyses. The complete conversion of butyryl-CoA to crotonyl-CoA observed with both, ferrocenium and ETF as artificial and natural electron acceptor, respectively, is in line with the mechanisms described for the ACAD which preferentially binds the product enoyl-CoA in its reduced state and thus kinetically promotes the oxidative half-reaction, i.e. the electron transfer from the ACAD flavin to the ETF or the artificial electron acceptor [26]. The further conversions observed in the β oxidation spiral coincided well with the thermodynamics [29]. The crotonoyl-CoA conversion to 3(S)-hydroxybutyryl-CoA by the bifunctional HCDH/ECH did not run to completion but showed roughly an 80% conversion in agreement with the thermodynamics of the reaction ($\Delta G^{0'}$ = -3.3 kJ mol^{-1}). The addition of NAD^+ did not result in any detectably acetoacetyl-CoA formation which is also in accordance with the $\Delta G^{0'}$ of + 18 kJ mol^{-1} which makes the two-step crotonoyl- to acetoacetyl-CoA conversion endergonic by nearly +15 kJ mol^{-1}. Only when the ketothiolase is added the crotonoyl-CoA is fully transformed to two molecules of acetyl-CoA which is again nicely explained by the standard free energy change of -26.1 kJ mol^{-1} rendering the crotonoyl- to acetyl-CoA conversion in total exergonic by -11 kJ

mol^{-1} and accordingly a nearly full conversion of crotonoyl-CoA to acetyl-CoA by HCDH/ECH and KT was observed. Together with the full conversion of butyryl- to crotonoyl-CoA with ETF, this also indicates that the fully reconstituted complete β oxidation cascade including the ACAD, HCDH/ECH and KT with the ETF and NAD^+ as electron acceptors can run to completion. We could indeed observe acetyl-CoA formation from butyryl-CoA but it was not possible to get a full conversion since the addition of ETF in sufficient amounts led to protein precipitation. On the other hand, artificial electron acceptors for the ACAD (ferrocenium) or the ETF (DCPIP) interfered with the CoA essential for the KT reaction, making the full reconstitution experimentally impossible. However, taken into account, in addition to the mechanistic of the ACAD catalyzed reaction, the thermodynamics with the caldariellquinone (CQ) (reduction potential +100 mV [52], that of crotonoyl-CoA/butyryl-CoA -10 mV [75]) as electron acceptor (and finally oxygen ($E^{0'}$ +820 mV)), which may drive the reaction further into the oxidative direction, it becomes apparent that the β oxidation as an entire process is hardly reversible from a mechanistic and energetic point of view and that the ACAD catalyzed reaction with the CQ as “primary” electron acceptor is the major bottle neck. However, with the described two step enzyme cascade with ACAD/ETF followed by the HCDH/ECH-KT enzymes the full degradation of C4, C6 and C8 saturated FA-CoA esters to acetyl-CoA as well as butyryl-CoA and hexanoyl-CoA, respectively, could be confirmed. This in vitro reconstituted β oxidation pathway opens up further experimental opportunities to study β oxidation in Archaea and in general gain a deeper understanding of this long known pathway which is however, especially in Bacteria and Archaea, not that extensively studied as one might expect.

FA synthesis in *S. acdiocaldarius*?

Previous reports suggest that at least some archaea contain FAs and are hence able to synthesize them. This was indicated for *Saccharolobus solfataricus* (formerly known as *Sulfolobus solfataricus*), a close relative of *S. acidocaldarius*, and for *Ignicoccus hospitalis* [16], for some methanogens [17], and also some haloarchaea [19]. However, some of the characteristic components of the known FA synthesis machineries especially the acyl-carrier protein and the acyl-carrier protein synthase are nearly completely absent in Archaea [20] and also no archaeal organism has been shown so far to contain a full set of FA synthesis enzymes known from Bacteria and Eukaryotes [5]. Thus, if Archaea really synthesize FAs then the synthesis pathway must be fundamentally different and particularly ACP independent [20], and the β oxidation has been proposed to be used in the reverse direction [5]. However, the studies discussed above on the FA degradation in *S. acidocaldarius* strongly suggested that the FA β oxidation is not just a reversible process which can simple be used in one or the other direction. Instead, the β oxidation employs the acyl-CoA oxidation to the enoyl-CoA as one major driving force. Thus, to drive the whole process in the reverse i.e. the synthesis direction at least this reaction needs to be bypassed. In the canonical FA synthesis pathways this is usually

accomplished by the use of NAD(P)H ($E^{0'}$ -320 mV) as electron donor which renders the reaction highly exergonic towards saturated acyl thioester formation (crotonoyl-CoA/butyryl-CoA $E^{0'}$ -10 mV) with a standard free energy change corresponding to -60 kJ mol^{-1} [29]. Combined with the lower part of the β oxidation this would turn such a pathway exergonic by -49 kJ mol^{-1}. Saci_1115 encoded in the *saci_1103-1126* gene cluster clearly showed such a NADPH dependent enoyl-CoA reductase activity with a specificity for medium-chain substrates between C4 and C10. The closest characterized homologue to this protein is the acryloyl-CoA reductase from *M. sedula*, also catalyzing a similar double bond reduction on the C3 derivative acryloyl-CoA in course of autotrophic growth utilizing the 3HP/4HB cycle for CO_2 fixation. Although the *M. sedula* enzyme was shown not to utilize crotonoyl-CoA, the catalytic efficiency with acryloyl-CoA was reported to be more than 100 fold higher than that of the Saci_1115 with crotonoyl-CoA [76, 77]. One might argue that such high activities are required to enable the high growth rates during autotrophic growth, whereas the low activities observed for Saci_1115 could already be sufficient to ensure satisfactory synthesis of FAs in *S. acidocaldarius* which are likely present only in minor quantities in archaeal organisms. Saci_1115 (and also the *M. sedula* enzyme) is a member of the medium chain dehydrogenase/reductase (MDR) superfamily whereas the bacterial enoyl-ACP reductases (encoded by fabI, fabL, or fabV) belong to the short chain dehydrogenase/reductase (SDR) superfamily [24, 78, 79]. Interestingly, the mitochondrial FAS II enoyl thioester reductase and also the enoyl reductase components of the FAS I systems of other eukaryotes (mammalia and fungi) also belong to the MDR superfamily [80]. As the *S. acidocaldarius* enzyme, the mitochondrial enzymes also show a clear preference for NADPH as cosubstrate. In contrast, cosubstrate specificity in Bacteria for the enoyl-ACP reductases appear less strict and e.g. the *E. coli* enzyme accept both substrates with similar efficiencies [81, 82].

However, the cosubstrate preference is considered as a distinctive property between FA synthesis and degradation with NADPH preferred by the FAS systems and NAD^+ preferably used by the β oxidation. Thus, although the Saci_1109 bifunctional HCDH/ECH operates reversibly *in vitro* the NAD^+ preference might suggest a favorable function in β oxidation rather than biosynthesis. Furthermore, Saci_1109 showed a clear preference for the 3(S)-hydroxybutyryl-CoA. This stereospecificity is another distinctive feature of the β oxidation whereas the FA synthesis is specific for the 3(R) stereoisomers [23]. In this context Saci_1104 with its specificity for 3(R)-hydroxybutyryl-CoA and NADPH as cosubstrate appears as a suitable candidate to function as 3-oxoacyl-CoA reductase in course of FA synthesis in *S. acidocaldarius*. As the classical bacterial and eukaryal 3-oxoacyl-ACP reductases the Saci_1104 belongs to the short chain dehydrogenase/reductase (SDR) superfamily and is accordingly also annotated as *fabG* in the genome [23]. However, most of the SDR members belong to only one big subfamily within the superfamily including the fabGs and hence the

annotation might in these cases not be that meaningful [83, 84]. The SDR superfamily is very large and although the diversity in Archaea is much more limited compared to Eukaryotes and especially Bacteria, SDR members are found in most archaeal genomes often on several paralogous copies. The *S. acidocaldarius* genome habours in total 11 SDR members of which only Saci_1232 was characterized and the crystal structure has been solved [85]. The enzyme catalyzed the stereospecific reduction of benzil to (R)-benzoin and also utilized other cyclic and aromatic substrates. From extremely halophilic *Haloarcula hispanica* it was shown that one of in total six SDRs functions in (R)-hydroxybutyryl-CoA formation from acetoacetyl-CoA in course of polyhydroxyalkanoate production [86]. The kinetic constants determined for Saci_1104 (k_{cat} 0.6 s^{-1} at 35°C corresponding to ~10 s^{-1} at physiological temperature of 75°C, K_m 0.08 mM) are in a similar range as reported e.g. for *Mycobacterium tuberculosis* (7 s^{-1}, 0.165 mM) [87], although the range in the kinetic constants appear rather broad with reported K_m values for acetoacetyl-CoA between 0.006 mM and 2.2 mM (Brenda database). Moreover, a native molecular mass of 84 kDa (subunit size 27 kDa) was obtained and would perfect correspond to a homotrimeric structure. However, a similar result was also observed for the *Pseudomonas aeruginosa* homologue showing a molecular weight between dimer and tetramer [89]. The explanation for it was that the ion strength of the buffer affected the oligomeric state of the protein and led to appearance of the dimer-tetramer mixture in solution. Therefore, our results suggested that Saci_1104 might present in both homodimeric and homotetrameric structures.

Interestingly, in the genome of *Haloarcula hispanica* the fabG gene is directly adjacent to a gene annotated as maoC1. MaoCs are only distantly related to the 3(R)-hydroxyacyl-ACP dehydratases from the canonical FA synthesis pathways and both belong to the hotdog fold superfamily [23, 89]. In the PHA biosynthesis the MaoCs act as (R)-specific enoyl-CoA hydratases providing (R)-hydroxyacyl-CoAs for PHA synthesis from FAs interconnecting β oxidation and PHA biosynthesis [90, 91]. Furthermore, in *Mycobacterium* spp. the MaoC dehydratases were shown to participate in the biosynthesis of mycolic acids dehydrating 3(R)-hydroxyacyl-ACPs to the respective trans-enoyl-ACPs during FA elongation [92]. Among the seven hot dog fold proteins in *S. acidocaldarius* only Saci_1085 and Saci_1070 are similar to the MaoC dehydratases whereas the others likely represent thioesterases [93]. Accordingly, for Saci_1085 the reversible 3(R)-hydroxybutyryl-CoA dehydratase activity was confirmed and also a histidine residue catalytically important in the *Mycobacterium* enzyme is well conserved whereas the thiolase sequences show a different sequence pattern in the active site (Fig. S6). The kinetic constants for 3(R)-hydroxybutyryl-CoA were determined as V_{max} 1.72 U mg^{-1} and K_m 0.4 mM whereas the V_{max} of 4.4 U mg^{-1} toward crotonoyl-CoA was similar to that reported for the *Mycobacterium tuberculosis* enzyme (2-3 U mg^{-1}), although the K_m of 0.22 mM appeared rather high. The K_m of the *M. tuberculosis* enzyme for the C12 enoyl-CoA was as

low as 1.6 µM [94]. However, the K_m for crotonoyl-CoA was not determined and might be expected to be much higher as crotonoyl-CoA was shown to be a poorer substrate than dodecenoyl-CoA (V_{max} ~60 U mg^{-1}). The activities for the canonical enoyl thioester dehydratases with the enoyl-CoA esters like crotonoyl-CoA were reported to be much lower, e.g. for *Campylobacter jejuni* FabZ 0.18 U mg^{-1} was reported and the K_m of 0.07 mM was in a similar range as observed for the Saci_1085 herein [95]. However, the canonical enzymes convert the ACP thioesters *in vivo* and thus the catalytic efficiency might be higher with these physiological substrates [96]. For the Saci_1085 enoyl-CoA dehydratase a broader substrate spectrum also towards longer chain enoyl-CoAs such as C6 or C8 appears likely since only one additional dehydratase candidate, i.e. Saci_1070, is present in *S. acidocaldarius*, although this protein lacks the catalytically essential histidine residue. The apparent equilibrium of the reaction under the chosen conditions was roughly 2 (3(R)-hydroxybutyryl-CoA:crotonoyl-CoA) and fits thus quite well to the theoretical value of 3.7 under standard conditions [29]. To partially reconstitute the FA synthesis pathway *in vitro*, 3(R)-hydroxybutyryl-CoA was incubated with both, Saci_1085 (3(R)-hydroxybutyryl-CoA dehydratase) and Saci_1115 (enoyl-CoA reductase) and indeed butyryl-CoA formation was observed. However, the equilibrium of the dehydratase reaction was reached comparably fast whereas the further reduction of crotonoyl- to butyryl-CoA continuously proceeded, but only slowly. This slow conversion can be explained by the low catalytic efficiency of the enzyme for crotonoyl-CoA as a substrate which can be expected to be much higher with longer chain substrates based on the observed much higher activities. Also, fatty acid biosynthesis in bacteria is a tightly regulated process which also includes feedback inhibition mechanisms especially of the acetyl-CoA carboxylase, the ketoacyl thioester synthase and the enoyl-thioester reductase [21]. Although the *S. acidocaldarius* enzymes are different from the classical bacterial ones especially the enoyl thioester reductase from the MDR superfamily, such feedback mechanisms might also contribute to the low conversion rate of the enoyl-CoA reductase Saci_1115. Nevertheless, there are several further candidates of the MDR superfamily in *S. acidocaldarius* which have so far not been characterized and we can also not exclude that uncharacterized SDR homologues might also be involved in this conversion (see above). In principle, however, this experiment shows that the reductase is likely involved in the potential FA synthesis in *S. acidocaldarius*.

Also, as mentioned above, the 3-oxoacyl-CoA to 3-hydroxyacyl-CoA conversion which lies far on the side of the hydroxyacyl-CoA contributes to the overall energetics of the FA synthesis. However, in the canonical FA synthesis pathways, the initial ATP-dependent carboxylation of acetyl-CoA to malonyl-CoA and the subsequent decarboxylative synthesis of the ketoacyl thioester from malonyl and acetyl thioester bypasses the energetically unfavorable non-decarboxylative Claisen condensation of two ac(et)yl thioesters and thus represent a further

important driving force in the synthesis direction [25]. Although in some Crenarchaeota including *S. acidocaldarius* the acetyl-CoA carboxylases are present as part of the 3HP/4HB pathway for CO_2 fixation during autotrophic growth [97, 98], the decarboxylating ketothiolases (= ketoacyl thioester synthases), which are also members of the thiolase superfamily, are missing [5, 20]. Instead, another mechanism as already mentioned above to drive this reaction involving DUF35 domain mediated complex formation and substrate channeling has recently been described to occur in Archaea [73]. And also, in *S. acidocaldarius* several DUF35 domain proteins were identified. One of those is encoded in the *saci_1103-1126* gene cluster (Saci_1120) together with an upstream located second ketothiolase gene (Saci_1121) and since the biosynthetic ketothiolases were shown to mostly cooccur in this manner with the DUF35 domain genes in Archaea [73], it appears likely that this ketothiolase rather than the Saci_1114 is involved in FA synthesis in *S. acidocaldarius*. And this in turn would also mean that this mechanism might be of broader significance in Archaea. Further studies on complex formation and the role of the DUF35 domains in FA metabolism in Archaea are currently under way.

Taken together, we herein report the first comprehensive biochemical study on the FA metabolism in Archaea, i.e. from the aerobic, thermoacidophilic crenarchaeal model organism *S. acidocaldarius*. Although the FA β oxidation at a first glance looks quite similar to the known pathway from Bacteria and mitochondria, the pathway shows some unusual features with respect to the ETF and the HCDH/ECH bifunctional enzyme together with the previously recognized “archaeal type” ketothiolases [5]. Furthermore, our results strongly indicate that the β oxidation as entire pathway is not operating reversibly in Archaea (Fig. 1A). Instead, we propose a potential archaeal FA synthesis pathway (Fig. 2A) which shows a kind of mosaic character with similarities to both bacterial (fabG, MaoC) and eukaryal (MDR enoyl thioester reductase) features mixed with unique archaeal properties (DUF35 domain/KT complexes, ACP independence). In contrast to the studies on the distribution of the β oxidation homologous in Archaea including the Asgardarchaeota previously reported [5, 99], the abundance of the proposed FA synthesis pathway remains to be elucidated. This is particularly challenging since especially the SDR and MDR superfamily proteins, i.e. the ketoacyl-CoA dehydrogenases and the enoyl reductases, are difficult to identify just by sequence and there are a lot of paralogues present in many Archaea. Nevertheless, first results already suggest that the presence of MaoC homologues might correlate with the occurrence of the MDR enoyl-CoA reductase and the SDR ketoacyl-CoA reductase in Sulfolobaceae (e.g. *S. acidocaldarius, S. islandicus, Saccharolobus solfataricus, Sulfurisphaera tokodaii* and *Metallosphaera sedula*), some Thermoproteaceae (e.g. *Vulcanisaeta distibuta*), some Thaumarchaeota (Nitrososphaeria), and also some unclassified Euryarchaeota species. Further remaining open questions, the answer to which is far beyond the scope of this paper, like e.g. the physiological

significance of the newly proposed FA synthesis pathway and of FAs in Archaea in general, as well as the regulation of the FA synthesis and degradation pathways in *S. acidocaldarius* and other Archaea are currently under investigation and will finally shed further light on the still enigmatic “lipid divide”.

Author contributions

XZ and CS carried out the experiments. KW contributed to the cloning of the key enzymes Saci_1109 and Saci_1123. The contribution of TK is the chemical synthesis of the CoA containing intermediates acryloyl-CoA, crotonoyl-CoA, individual D- or L-3-hydroxybutyryl-CoA, acetoacetyl-CoA, decenoyl-CoA and hexadecenoyl-CoA. CB, XZ and CS wrote the manuscript, which was edited by CB and BS. CB and BS conceived the study. All authors approved the final manuscript.

Acknowledgements

XZ and TK received funds from the VW Stiftung in the “Life?” initiative (96725). CS is supported by the German Federal Ministry of Education and Research (BMBF). We acknowledge support by the Open Access Publication Fund of the University of Duisburg-Essen.

Figure 1

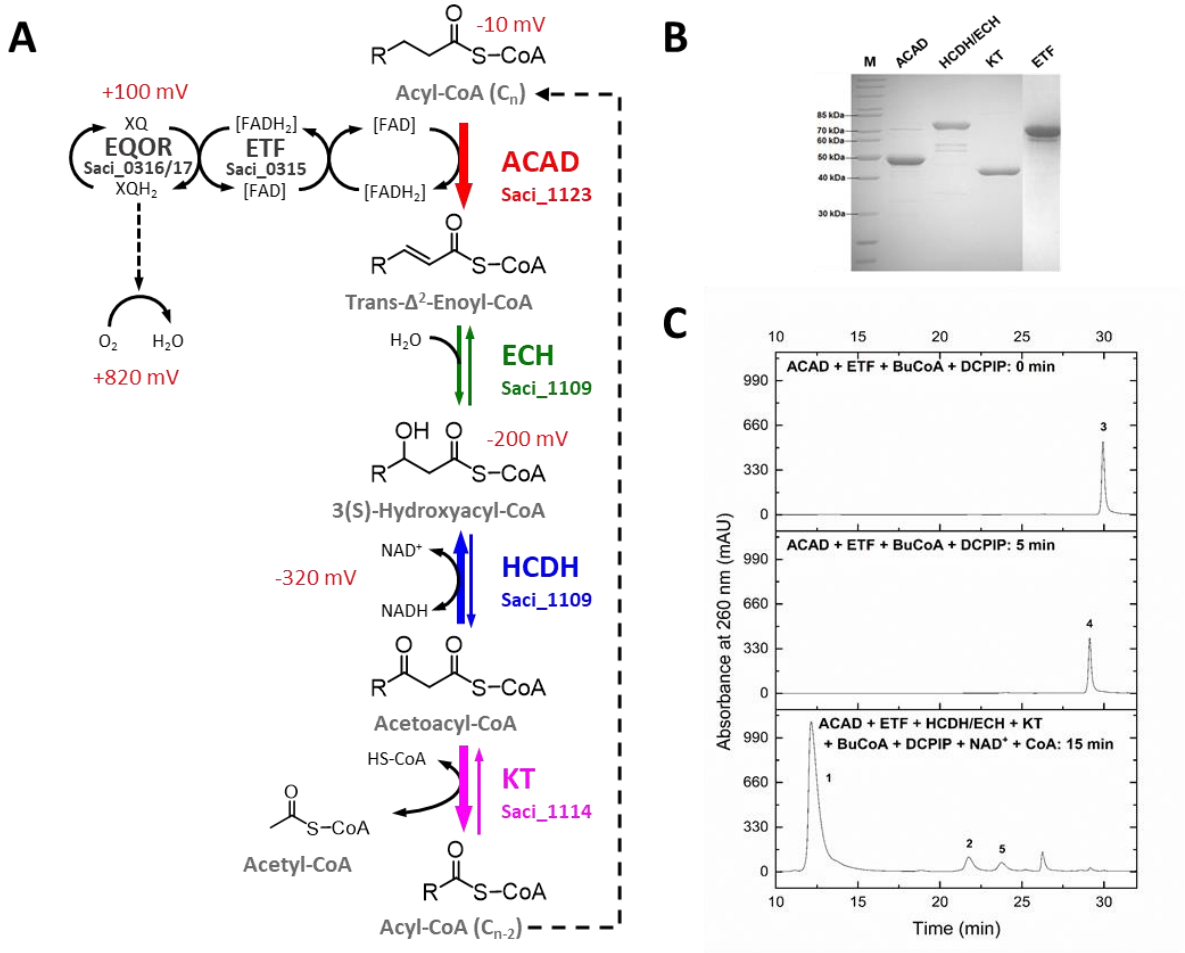


Figure 1. **Reconstructed β oxidation pathway for FA degradation in *S. acidocaldarius*.** The β oxidation reactions (A), the purified recombinant proteins that catalyze these reactions (SDS-PAGE and Coomassie Blue staining) (B) as well as the HPLC chromatogram of the β oxidation enzyme cascade for butyryl-CoA conversion (C) are shown. In Fig. A, the subscript "n" of acyl-CoA represents the length of the carbon chain which equals 4, 6 or 8. The thickness of the arrow (A) indicates the energetics of the respective reaction. For redox reactions the reduction potential is given. During β oxidation enzyme cascade (C), butyryl-CoA (peak 3) was completely oxidized to crotonoyl-CoA (peak 4) by ACAD transferring the electron to DCPIP through ETF in the first step. Then crotonoyl-CoA was further converted to 3-hydroxybutyryl-CoA (peak 5) by HCDH/ECH and finally to acetyl-CoA (peak 2) by KT in presence of free CoA (peak 1). However, the intermediate acetoacetyl-CoA released by HCDH/ECH using NAD^+ as cofactor was not detectable under the applied analytical conditions. The abbreviations: ACAD, acyl-CoA dehydrogenase; ETF, electron transferring flavoprotein; ECH, enoyl-CoA hydratase; HCDH, 3(S)-hydroxyacyl-CoA dehydrogenase; KT, β-ketothiolase or acetyl-CoA C acetyltransferase; BuCoA: butyryl-CoA; DCPIP: 2,6-dichlorophenolindophenol.

Figure 2

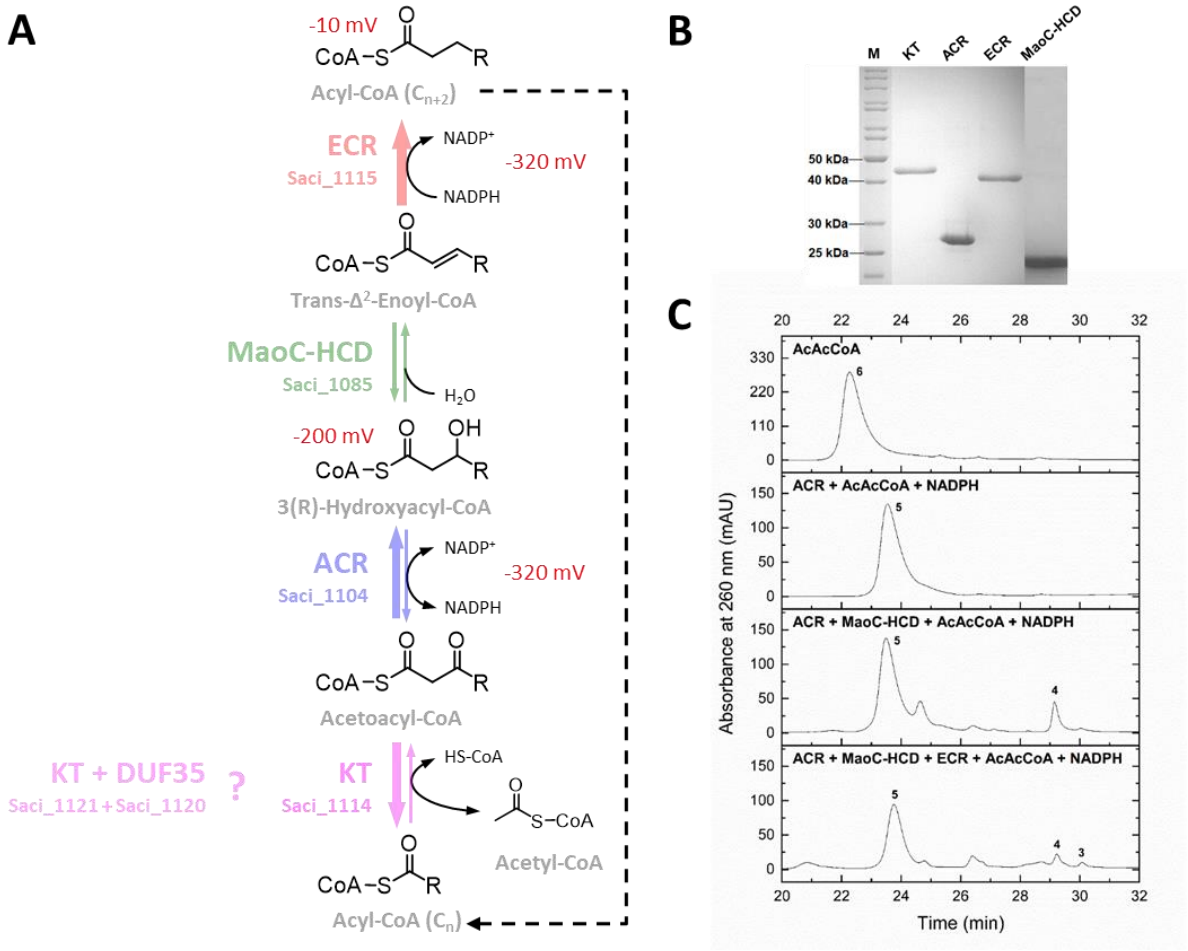


Figure 2. **Proposed FA biosynthetic pathway in *S. acidocaldarius*.** The FA synthetic reactions (A), the purified recombinant proteins that catalyze these conversions (SDS-PAGE and Coomassie Blue staining) (B) as well as the enzyme cascade for butyryl-CoA synthesis via HPLC (C) are shown. The abbreviations for enzymes: KT, β-ketothiolase or acetyl-CoA C acetyltransferase; DUF35, potential scaffold protein with uncertain function; ACR, SDR superfamily acetoacyl-CoA reductase or 3-ketoacyl-CoA reductase; HCD, MaoC like 3(R)-hydroxyacyl-CoA dehydratase; ECR, MDR family enoyl-CoA reductase. The thickness of the arrow indicates the energetics of the respective reaction. For redox reactions the reduction potential is given. In Fig. A, the subscript “n” of acyl-CoA represents the length of the carbon chain which equals 2, 4 or 6. The thickness of the arrow (A) indicates the energetics of the respective reaction. For redox reactions the reduction potential is given. During FA synthesis enzyme cascade (C), acetoacetyl-CoA (peak 6) was completely reduced to 3-hydroxybutyryl-CoA (peak 5) by ACR employing NADPH as cofactor. Then acetoacetyl-CoA was further converted to crotonoyl-CoA (peak 4) by MaoC-HCD and finally to butyryl-CoA (peak 3) was produced by the second reductase ACR in addition of NADPH coenzyme. The abbreviations: KT, β-ketothiolase or acetyl-CoA C acetyltransferase; ACR, acetoacetyl-CoA reductase; MaoC-HCD: MaoC-like 3(R)-hydroxyacyl-CoA dehydratase; ECR: enoyl-CoA reductase; AcAcCoA: acetoacetyl-CoA.

Table 1. Molecular and kinetic parameters of the β oxidation enzymes in *S. acidocaldarius.*

ORF/Enzyme	Native mass (kDa) (Oligomer)	Temperature	Substrate/Cofactor	K_m (mM)	V_{max} (U mg^{-1})	k_{cat} (S^{-1})
Saci_1123 Acyl-CoA dehydrogenase (ACAD)	170 (homotetramer)	65°C	Butyryl-CoA (C4)	ND	4.5	ND
			Hexanoyl-CoA (C6)	ND	8.24	ND
			Octanoyl-CoA (C8)	0.0151 ± 0.004	29.046 ± 2.048	21.234
			Palmitoyl-CoA (C16)	ND	0	ND
Saci_0315 Electron transfer flavoprotein (ETF)	68 (monomer)	65°C	NADH	0.039 ± 0.012	0.703 ± 0.046	0.784
			Idonitrotetrazolium chloride	ND		
Saci_1109 Enoy-CoA hydratase/3-hydroxyacyl-CoA dehydrogenase (ECH/HCDH)	512 (homoheptamer)	75°C	Aryloyl-CoA (C3:1)	ND	[a] 0.073	ND
			Crotonyl-CoA (C4:1)	0.024 ± 0.004	16.974 ± 0.479	20.56
			NAD^+	0.036 ± 0.009		
			$NADP^+$	ND	0	ND
			Decenoyl-CoA (C10:1)	ND	[a] 7.6	ND
			Hexadecenoyl-CoA (C16:1)	ND	0	ND
			3(S/R)-Hydroxybutyryl-CoA (C4)	0.092 ± 0.016	29.658 ± 1.446	35.923
			NAD^+	0.11 ± 0.012		
			$NADP^+$	ND	0	ND
			3(S)-Hydroxybutyryl-CoA (C4)	0.043 ± 0.002	48.439 ± 0.398	58.672
			3(R)-Hydroxybutyryl-CoA (C4)	ND	0	ND
		35°C	Acetoacetyl-CoA (C4)	0.076 ± 0,008	6.969 ± 0.191	8.441
			NADH	0.028 ± 0.007		
			NADPH	*0.094 ± 0.22		
Saci_1114 β-Ketothiolase (KT)	88 (homodimer)	23°C	Acetoacetyl-CoA (C4)	0.03339 ± 0.01	1.67645 ± 0.25	1.2
		75°C	CoA	0.00338	2.53375	1.8
			Acetyl-CoA (C2)	2.097 ± 0.263	2.739 ± 0.139	1.959

* The kinetic parameters were calculated in regardless of inhibition effect caused by higher concentration of the substrate.
Numbers after ± represent standard error (SE).
ND: not detected.
[a] The values of these specific activities were estimated according to the experimental measurements.

Table 2. Molecular and kinetic parameters of the enzymes involved in the new potential FA synthesis pathway in *S. acidocaldarius*.

ORF/Enzyme	Native mass (kDa) (Oligomer)	Temperature	Substrate/Cofactor	K_m (mM)	V_{max} (U mg^{-1})	k_{cat} (S^{-1})
Saci_1104 Acetoacyl-CoA reductase (ACR)	84 (homotrimer)	35°C	Acetoacetyl-CoA (C4)	0.077 ± 0.009	1.3 ± 0.042	0.586
			NADPH	0.024 ± 0.003		
		70°C	3(S/R)-Hydroxybutyryl-CoA (C4)	0.336 ± 0.02	0.211 ± 0.004	0.095
			3(S)-Hydroxybutyryl-CoA (C4)	ND	0	ND
			3(R)-Hydroxybutyryl-CoA (C4)	0.16 ± 0.012	0.965 ± 0.024	0.435
Saci_1085 MaoC like 3(R)-hydroxacyl-CoA dehydratase (MaoC-HCD)	244 (homo-dodecamer)	65°C	3(S)-Hydroxybutyryl-CoA (C4)	ND	0	ND
			3(R)-Hydroxybutyryl-CoA (C4)	0.399 ± 0.127	1.72 ± 0.301	0.549
			Crotonoyl-CoA (C4:1)	0.222 ± 0.064	4.398 ± 0.414	1.404
			Hexenoyl-CoA (C6:1)	ND	3.5	ND
			Octenoyl-CoA (C8:1)	ND	5.1	ND
Saci_1115 Enoyl-CoA reductase (ECR)	69 (homodimer)	70°C	Aryloyl-CoA (C3:1)	ND	0	ND
			Crotonoyl-CoA (C4:1)	0.096 ± 0.018	0.422 ± 0.024	0.256
			NADPH	0.007 ± 0.00076		
			NADH	ND	0.077	ND
			[b] Hexenoyl-CoA (C6:1)	ND	[a] 0.54	ND
			[b] Octenoyl-CoA (C8:1)	ND	[a] 0.93	ND
			Decenoyl-CoA (C10:1)	ND	[a] 0.337	ND
			Hexadecenoyl-CoA (C16:1)	ND	[a] 0.098	ND

The numbers after ± represent standard error (SE).
ND: not detected.
[a] The values of these specific activities were estimated according to the experimental measurements.
[b] In these enzyme assays, the saturated, straight-chain CoA-esters (hexanoyl-CoA or octanoyl-CoA) were introduced first and converted into the relevant enoyl-form (hexenoyl-CoA or octenoyl-CoA) by the acyl-CoA dehydrogenase Saci_1123, then the produced enoyl-CoAs were subsequently used by the enoyl-CoA reductase Saci_1115.

References

1. Zweerink, S., et al., *Activity-based protein profiling as a robust method for enzyme identification and screening in extremophilic Archaea.* Nature Communications, 2017. **8**: p. 15352.
2. Choi, Y.H., et al., *Purification and characterization of a new inducible thermostable extracellular lipolytic enzyme from the thermoacidophilic archaeon Sulfolobus solfataricus* P1. Journal of Molecular Catalysis B: Enzymatic, 2016. **124**: p. 11-19.
3. Falb, M., et al., *Metabolism of halophilic archaea.* Extremophiles, 2008. **12**(2): p. 177-96.
4. Khelifi, N., et al., *Anaerobic Oxidation of Fatty Acids and Alkenes by the Hyperthermophilic Sulfate-Reducing Archaeon Archaeoglobus fulgidus.* Applied and Environmental Microbiology, 2010. **76**(9): p. 3057-3060.
5. Dibrova, D.V., et al., *Phylogenomic reconstruction of archaeal fatty acid metabolism.* Environ Microbiol, 2014. **16**(4): p. 907-918.
6. Sobek, H. and Gorisch, H., *Purification and characterization of a heat-stable esterase from the thermoacidophilic archaebacterium Sulfolobus acidocaldarius.* Biochem J, 1988. **250**(2): p. 453-8.
7. Park, Y.J., et al., *A carboxylesterase from the thermoacidophilic archaeon Sulfolobus solfataricus P1; purification, characterization, and expression.* Biochimica et Biophysica Acta - General Subjects, 2006. **1760**(5): p. 820-828.
8. Wang, K., et al., *A TetR-family transcription factor regulates fatty acid metabolism in the archaeal model organism Sulfolobus acidocaldarius.* Nature Communications, 2019. **10**(1): p. 1542.
9. Brock, T.D., et al., *Sulfolobus: A new genus of sulfur-oxidizing bacteria living at low pH and high temperature.* Archiv für Mikrobiologie, 1972. **84**(1): p. 54-68.
10. Lewis, A.M., et al., *The biology of thermoacidophilic archaea from the order Sulfolobales.* FEMS Microbiology Reviews, 2021.
11. Bräsen, C., et al., *Carbohydrate Metabolism in Archaea: Current Insights into Unusual Enzymes and Pathways and Their Regulation.* Microbiology and Molecular Biology Reviews, 2014. **78**(1): p. 89-175.
12. Chen, L., et al., *The genome of Sulfolobus acidocaldarius, a model organism of the Crenarchaeota.* J Bacteriol, 2005. **187**(14): p. 4992-9.
13. Fujita, Y., et al., *Regulation of fatty acid metabolism in bacteria.* Mol Microbiol, 2007. **66**(4): p. 829-39.
14. Ford, T.J. and Way, J.C., *Enhancement of E. coli acyl-CoA synthetase FadD activity on medium chain fatty acids.* PeerJ, 2015. **3**: p. e1040.
15. Morsczeck, C., et al., *The macrophage-induced gene (mig)of Mycobacterium avium encodes a medium-chain acyl-coenzyme A synthetase.* 2001.
16. Hamerly, T., et al., *Characterization of Fatty Acids in Crenarchaeota by GC-MS and NMR.* Archaea, 2015. **2015**: p. 9.
17. Andreas, G., et al., *Phospholipid etherlipid and phospholipid fatty acid fingerprints in selected euryarchaeotal monocultures for taxonomic profiling1.* FEMS Microbiology Letters, 2002. **213**(1): p. 133-139.
18. Corcelli, A., et al., *Role of palmitic acid on the isolation and properties of halorhodopsin.* Biochimica et Biophysica Acta (BBA) - Biomembranes, 1996. **1281**(2): p. 173-181.
19. Kolbe, M., et al., *Structure of the Light-Driven Chloride Pump Halorhodopsin at 1.8 Å Resolution.* Science, 2000. **288**(5470): p. 1390-1396.
20. Lombard, J., et al., *An ACP-independent fatty acid synthesis pathway in archaea: implications for the origin of phospholipids.* Mol Biol Evol, 2012. **29**(11): p. 3261-5.
21. Fujita, Y., et al., *Regulation of fatty acid metabolism in bacteria.* Molecular Microbiology, 2007. **66**(4): p. 829-839.
22. Byers, D.M. and Gong, H., *Acyl carrier protein: structure–function relationships in a conserved multifunctional protein family.* Biochemistry and Cell Biology, 2007. **85**(6): p. 649-662.

23. Bhaumik, P., et al., *Structural biology of the thioester-dependent degradation and synthesis of fatty acids.* Current Opinion in Structural Biology, 2005. **15**(6): p. 621-628.
24. White, S.W., et al., The structural biology of type II fatt acid biosynthesis. Annual Review of Biochemistry, 2005. **74**(1): p. 791-831.
25. Massengo-Tiassé, R.P. and Cronan, J.E., *Diversity in Enoyl-Acyl Carrier Protein Reductases.* Cellular and molecular life sciences : CMLS, 2009. **66**(9): p. 1507.
26. Ghisla, S. and Thorpe, C., *Acyl-CoA dehydrogenases.* European Journal of Biochemistry, 2004. **271**(3): p. 494-508.
27. Watmough, N.J. and Frerman, F.E., *The electron transfer flavoprotein: Ubiquinone oxidoreductases.* Biochimica et Biophysica Acta (BBA) - Bioenergetics, 2010. **1797**(12): p. 1910-1916.
28. Dellomonaco, C., et al., *Engineered reversal of the β-oxidation cycle for the synthesis of fuels and chemicals.* Nature, 2011. **476**(7360): p. 355-359.
29. Flamholz, A., et al., *eQuilibrator--the biochemical thermodynamics calculator.* Nucleic Acids Res, 2012. **40**(Database issue): p. D770-5.
30. Peter, D.M., et al., *A Chemo-Enzymatic Road Map to the Synthesis of CoA Esters.* Molecules, 2016. **21**(4): p. 517.
31. Brock T.D., et al., *Sulfolobus: A New Genus of Sulfur-Oxidizing Bacteria Living at Low pH and High Temperature.* 1972.
32. Wagner, M., et al., *Versatile Genetic Tool Box for the Crenarchaeote Sulfolobus acidocaldarius.* Front Microbiol, 2012. **3**: p. 214.
33. Bradford, M.M., *A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding* 1976.
34. Bräsen, C., et al., *A novel octameric AMP-forming acetyl-CoA synthetase from the hyperthermophilic crenarchaeon Pyrobaculum aerophilum.* FEBS Lett, 2005. **579**(2): p. 477-82.
35. Li, F., et al., *Coupled ferredoxin and crotonyl coenzyme A (CoA) reduction with NADH catalyzed by the butyryl-CoA dehydrogenase/Etf complex from Clostridium kluyveri.* J Bacteriol, 2008. **190**(3): p. 843-50.
36. Cabello, C.M., et al., *Antimelanoma activity of the redox dye DCPIP (2,6-dichlorophenolindophenol) is antagonized by NQO1.* Biochem Pharmacol, 2009. **78**(4): p. 344-54.
37. Kiema, T.R., et al., *The crystal structure of human mitochondrial 3-ketoacyl-CoA thiolase (T1): insight into the reaction mechanism of its thiolase and thioesterase activities.* Acta Crystallogr D Biol Crystallogr, 2014. **70**(Pt 12): p. 3212-25.
38. Zhang, H., et al., *Molecular effect of FadD on the regulation and metabolism of fatty acid in Escherichia coli.* FEMS Microbiol Lett, 2006. **259**(2): p. 249-53.
39. Ishikawa, M.,et al., *Reconstitution, morphology and crystallization of a fatty acid β-oxidation multienzyme complex from Pseudomonas fragi.* Biochem. J., 1997. **328**: 815-820
40. Venkatesan, R. and Wierenga, R.K., *Structure of mycobacterial beta-oxidation trifunctional enzyme reveals its altered assembly and putative substrate channeling pathway.* ACS Chem Biol, 2013. **8**(5): p. 1063-73.
41. Wang, K., et al., *A TetR-family transcription factor regulates fatty acid metabolism in the archaeal model organism Sulfolobus acidocaldarius.* Nature Communications, 2019. **10**, 1542.
42. Sobek, H. and Gorisch, H., *Further kinetic and molecular characterization of an extremely heat-stable carboxylesterase from the thermoacidophilic archaebacterium Sulfolobus acidocaldarius.* Biochem J, 1989. **261**(3): p. 993-8.
43. Park, Y.J., et al., *A novel thermostable arylesterase from the archaeon Sulfolobus solfataricus P1: Purification, characterization, and expression.* J Bacteriol, 2008. **190**(24): p. 8086-8095.
44. Ingram-Smith C. and Smith. K.S., *AMP-forming acetyl-CoA synthetases in Archaea show unexpected diversity in substrate utilization.* Archaea, 2006. **2**: 95-107.

45. Berger, S., et al., *Acetate activation in Methanosaeta thermophila: characterization of the key enzymes pyrophosphatase and acetyl-CoA synthetase.* Archaea, 2012. **2012**: p. 315153.
46. Jetten, M.S., et al., *Isolation and Characterization of Acetyl-Coenzyme A Synthetase from Methanothrix soehngenii.* Journal of bacteriology, 1989. **171**(10): 5430-5435.
47. Bräsen, C. and Schonheit, P., *AMP-forming acetyl-CoA synthetase from the extremely halophilic archaeon Haloarcula marismortui: purification, identification and expression of the encoding gene, and phylogenetic affiliation.* Extremophiles, 2005. **9**(5): p. 355-65.
48. Meng, Y., et al., *Characterization of an archaeal medium-chain acyl coenzyme A synthetase from Methanosarcina acetivorans.* J Bacteriol, 2010. **192**(22): p. 5982-90.
49. Kuprat, T., et al., *Acetate Metabolism in Archaea: Characterization of an Acetate Transporter and of Enzymes Involved in Acetate Activation and Gluconeogenesis in Haloferax volcanii.* Front Microbiol, 2020. **11**: p. 604926.
50. Kim, J.J.P. and Miura, R., *Acyl-CoA dehydrogenases and acyl-CoA oxidases.* European Journal of Biochemistry, 2004. **271**(3): p. 483-493.
51. Shen, Y.Q., et al., *Diversity and dispersal of a ubiquitous protein family: acyl-CoA dehydrogenases.* Nucleic Acids Research, 2009. **37**(17): p. 5619-5631.
52. Anemuller, S. and Schåfer, G., *Cytochrome aa^3 from Sulfolobus acidocaldarius.* European Journal of Biochemistry, 1990. **191**(2): p. 297-305.
53. Yamauchi, N., *The Pathway of Leucine to Mevalonate in Halophilic Archaea: Efficient Incorporation of Leucine into Isoprenoidal Lipid with the Involvement of Isovaleryl-CoA Dehydrogenase in Halobacterium salinarum.* Bioscience, Biotechnology, and Biochemistry, 2010. **74**(2): p. 443-446.
54. Garcia Costas, A.M., et al., *Defining Electron Bifurcation in the Electron-Transferring Flavoprotein Family.* Journal of Bacteriology, 2017. **199**(21): p. e00440-17.
55. Schut, G.J., et al., *The catalytic mechanism of electron-bifurcating electron transfer flavoproteins (ETFs) involves an intermediary complex with NAD+.* Journal of Biological Chemistry, 2019. **294**(9): p. 3271-3283.
56. Vogt, M.S., et al., *Structural and Functional Characterization of an Electron Transfer Flavoprotein Involved in Toluene Degradation in Strictly Anaerobic Bacteria.* Journal of Bacteriology, 2019. **201**(21): p. e00326-19.
57. Berg, I.A., et al., *Autotrophic carbon fixation in archaea.* Nat Rev Microbiol, 2010. **8**(6): p. 447-60.
58. Fuchs, G., *Alternative Pathways of Carbon Dioxide Fixation: Insights into the Early Evolution of Life?* Annual Review of Microbiology, 2011. **65**(1): p. 631-658.
59. Schmid, G., et al., *Enzymes of the benzoyl-coenzyme A degradation pathway in the hyperthermophilic archaeon Ferroglobus placidus.* Environmental Microbiology, 2015. **17**(9): p. 3289-3300.
60. Ramos-Vera, W.H., et al., *Identification of Missing Genes and Enzymes for Autotrophic Carbon Fixation in Crenarchaeota.* J Bacteriol, 2011. **193**(5): p. 1201-1211.
61. Hawkins, A.B., et al., *Conversion of 4-Hydroxybutyrate to Acetyl Coenzyme A and Its Anapleurosis in the Metallosphaera sedula 3-Hydroxypropionate/4-Hydroxybutyrate Carbon Fixation Pathway.* Appl Environ Microbiol, 2014. **80**(8): p. 2536-2545.
62. Kunau, W.H., et al., *β-Oxidation of fatty acids in mitochondria, peroxisomes, and bacteria: A century of continued progress.* Progress in Lipid Research, 1995. **34**(4): p. 267-342.
63. Sah-Teli, S.K., et al., *Insights into the stability and substrate specificity of the E. coli aerobic beta-oxidation trifunctional enzyme complex.* J Struct Biol, 2020. **210**(3): p. 107494.
64. Pawar, S. and H. Schulz, *The structure of the multienzyme complex of fatty acid oxidation from Escherichia coli.* Journal of Biological Chemistry, 1981. **256**(8): p. 3894-3899.
65. Volodina, E. and Steinbüchel, A., *(S)-3-hydroxyacyl-CoA dehydrogenase/enoyl-CoA hydratase (FadB') from fatty acid degradation operon of Ralstonia eutropha H16.* AMB Express, 2014. **4**(1): p. 69.

66. Ishikawa, M., et al., *Structural basis for channelling mechanism of a fatty acid β-oxidation multienzyme complex.* The EMBO Journal, 2004. **23**(14): p. 2745-2754.
67. Sah-Teli, S.K., et al., *Insights into the stability and substrate specificity of the E. coli aerobic β-oxidation trifunctional enzyme complex.* Journal of Structural Biology, 2020. **210**(3): p. 107494.
68. Venkatesan, R. and Wierenga, R.K., *Structure of Mycobacterial β-Oxidation Trifunctional Enzyme Reveals Its Altered Assembly and Putative Substrate Channeling Pathway.* ACS Chemical Biology, 2013. **8**(5): p. 1063-1073.
69. Liu, L., et al., *Enzymes Catalyzing Crotonyl-CoA Conversion to Acetoacetyl-CoA During the Autotrophic CO_2 Fixation in Metallosphaera sedula.* Frontiers in Microbiology, 2020. **11**(354).
70. Bartlett, K. and Eaton, S., *Mitochondrial β-oxidation.* European Journal of Biochemistry, 2004. **271**(3): p. 462-469.
71. Kim, J.J.P. and Battaile, K.P., *Burning fat: the structural basis of fatty acid β-oxidation.* Current Opinion in Structural Biology, 2002. **12**(6): p. 721-728.
72. Bhaskar, S., et al., *Structural basis for differentiation between two classes of thiolase: Degradative vs biosynthetic thiolase.* Journal of Structural Biology: X, 2020. **4**: p. 100018.
73. Vögeli, B., et al., *Archaeal acetoacetyl-CoA thiolase/HMG-CoA synthase complex channels the intermediate via a fused CoA-binding site.* Proceedings of the National Academy of Sciences, 2018. **115**(13): p. 3380-3385.
74. Jiang, C., et al., *Divergent evolution of the thiolase superfamily and chalcone synthase family.* Molecular Phylogenetics and Evolution, 2008. **49**(3): p. 691-701.
75. Buckel, W. and Thauer, R.K., *Flavin-Based Electron Bifurcation, A New Mechanism of Biological Energy Coupling.* Chemical Reviews, 2018. **118**(7): p. 3862-3886.
76. Loder, A.J., et al., *Reaction kinetic analysis of the 3-hydroxypropionate/4-hydroxybutyrate CO_2 fixation cycle in extremely thermoacidophilic archaea.* Metabolic Engineering, 2016. **38**: p. 446-463.
77. Teufel, R., et al., *3-Hydroxypropionyl-Coenzyme A Dehydratase and Acryloyl-Coenzyme A Reductase, Enzymes of the Autotrophic 3-Hydroxypropionate/4-Hydroxybutyrate Cycle in the Sulfolobales.* Journal of Bacteriology, 2009. **191**(14): p. 4572-4581.
78. Kastaniotis, A.J., et al., *Mitochondrial fatty acid synthesis, fatty acids and mitochondrial physiology.* Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids, 2017. **1862**(1): p. 39-48.
79. Vick, J.E., et al., *Escherichia coli Enoyl-Acyl Carrier Protein Reductase (FabI) Supports Efficient Operation of a Functional Reversal of the β-Oxidation Cycle.* ACS Synth Biol, 2014. **1**(11): p. 541-54.
80. Hiltunen, J.K., et al., *Mitochondrial fatty acid synthesis--an adopted set of enzymes making a pathway of major importance for the cellular metabolism.* Prog Lipid Res, 2010. **49**(1): p. 27-45.
81. Massengo-Tiassé, R.P. and Cronan, J.E., *Vibrio cholerae FabV Defines a New Class of Enoyl-Acyl Carrier Protein Reductase.* Journal of Biological Chemistry, 2008. **283**(3): p. 1308-1316.
82. Bergler, H., et al., *The Enoyl-[Acyl-Carrier-Protein] Reductase (FabI) of Escherichia coli, which Catalyzes a Key Regulatory Step in Fatty Acid Biosynthesis, Accepts NADH and NADPH as Cofactors and is Inhibited by Palmitoyl-CoA.* European Journal of Biochemistry, 1996. **242**(3): p. 689-694.
83. Gräff, M., et al., *The Short-chain Dehydrogenase/Reductase Engineering Database (SDRED): A classification and analysis system for a highly diverse enzyme family.* Proteins: Structure, Function, and Bioinformatics, 2019. **87**(6): p. 443-451.
84. Persson, B. and Kallberg, Y., *Classification and nomenclature of the superfamily of short-chain dehydrogenases/reductases (SDRs).* Chemico-Biological Interactions, 2013. **202**(1): p. 111-115.
85. Pennacchio, A., et al., *Biochemical and structural characterization of recombinant short-chain NAD(H)-dependent dehydrogenase/reductase from Sulfolobus acidocaldarius*

highly enantioselective on diaryl diketone benzil. Appl Microbiol Biotechnol, 2013. **97**(9): p. 3949-64.
86. Han, J., et al., *Identification of the polyhydroxyalkanoate (PHA)-specific acetoacetyl coenzyme A reductase among multiple FabG paralogs in Haloarcula hispanica and reconstruction of the PHA biosynthetic pathway in Haloferax volcanii.* Appl Environ Microbiol, 2009. **75**(19): p. 6168-75.
87. Silva, R.G., et al., *Mycobacterium tuberculosis β-Ketoacyl-Acyl Carrier Protein (ACP) Reductase: Kinetic and Chemical Mechanisms.* Biochemistry, 2006. **45**(43): p. 13064-13073.
88. Hou, J., et al., *Dissecting the Structural Elements for the Activation of beta-Ketoacyl-(Acyl Carrier Protein) Reductase from Vibrio cholerae.* J Bacteriol, 2016. **198**(3): p. 463-76.
89. Pidugu, L.S., et al., *Analysis of proteins with the 'hot dog' fold: Prediction of function and identification of catalytic residues of hypothetical proteins.* BMC Structural Biology, 2009. **9**(1): p. 37.
90. Tsuge, T., et al., *Molecular characterization and properties of (R)-specific enoyl-CoA hydratases from Pseudomonas aeruginosa: metabolic tools for synthesis of polyhydroxyalkanoates via fatty acid beta-oxidation.* Int J Biol Macromol, 2003. **31**(4-5): p. 195-205.
91. Riedel, S., et al., *Lipid and fatty acid metabolism in Ralstonia eutropha: relevance for the biotechnological production of value-added products.* Applied Microbiology and Biotechnology, 2014. **98**(4): p. 1469-1483.
92. Sacco, E., et al., *The missing piece of the type II fatty acid synthase system from Mycobacterium tuberculosis.* Proc. Natl. Acad. Sci. USA, 2007. **104**(37): p. 14628-33.
93. Dillon, S.C. and Bateman, A., *The Hotdog fold: wrapping up a superfamily of thioesterases and dehydratases.* BMC Bioinformatics, 2004. **5**(1): p. 109.
94. Sacco, E., et al., *Rv3389C from Mycobacterium tuberculosis, a member of the (R)-specific hydratase/dehydratase family.* Biochim Biophys Acta, 2007. **1774**(2): p. 303-11.
95. Kirkpatrick, A.S., et al., *Campylobacter jejuni fatty acid synthase II: structural and functional analysis of beta-hydroxyacyl-ACP dehydratase (FabZ).* Biochemical and biophysical research communications, 2009. **380**(2): p. 407-412.
96. Liu, N., et al., *Mechanism and inhibition of the FabI enoyl-ACP reductase from Burkholderia pseudomallei.* The Journal of antimicrobial chemotherapy, 2011. **66**(3): p. 564-573.
97. Berg, I.A., et al., *A 3-Hydroxypropionate/4-Hydroxybutyrate Autotrophic Carbon Dioxide Assimilation Pathway in Archaea.* Science, 2007. **318**(5857): p. 1782-1786.
98. Zeldes, B.M., et al., *Determinants of sulphur chemolithoautotrophy in the extremely thermoacidophilic Sulfolobales.* Environmental Microbiology, 2019. **21**(10): p. 3696-3710.
99. Spang, A., et al., *Proposal of the reverse flow model for the origin of the eukaryotic cell based on comparative analyses of Asgard archaeal metabolism.* Nature Microbiology, 2019.

Supplementary Information

Table S 1. List of oligonucleotides used for cloning and sequencing of the FA related genes from *S. acidocaldarius* in this study.

Cloned genes	Oligonucleotides for cloning	Sequences (5´→3´) (restriction sites are marked in red)
Saci_1122	*Saci_1122*_fd_NdeI	GCGTGACATATGAGCAATGAGTATTACG
	*Saci_1122*_rv_BamHI	TTCACTGGATCCTCACTTGTTCTTCT
Saci_1123	*Saci_1123*_fd_HindIII	ATAAAGCTTATGGTTTTGCCTTTTAAAAC
	*Saci_1123*_rv_HindIII	ATAAAGCTTTTACATTTTTATGCCAAATAA
Saci_0315	*Saci_0315*_fd	TAGCAGCCGGATCCTCGAGCAGCCCCTTCTTTTTAATTAAC
	*Saci_0315*_rv	CCTGGTGCCGCGCGGCAGCCATATGGCAGAGCTTAAAATTGTCG
Saci_1109	*Saci_1109*_fd_ BamHI	TATGGATCCCATGAAAGTAGAAGATATTAAGAAA
	*Saci_1109*_rv_ BamHI	GATGGATCCTTATTCTCCTTTGAACTGTG
Saci_1114	M13_fw	TGTAAAACGACGGCCAGT
	M13_rv	CAGGAAACAGCTATGACC
Saci_1104	*Saci_1104*_fd_NdeI	GCTCGCCATATGTAGTCTCTTAAAGAC
	*Saci_1104*_rv_BamHI	CTAGCTGGATCCTTAAGCAATTCCT
Saci_1115	*Saci_1115*_fw_NdeI	CCTACGCATATGATGAAAGCTGTAATTCTTC
	*Saci_1115*_rv_BamHI	CGAGCTGGATCCTTATGGCTTTATAAGAATTTTAC
Saci_1085	*Saci_1085*_fw_NdeI	GCCCGCATATGTCAGAGCAGGGTCC
	*Saci_1085*_rv_BamHI	CGGGCGGATCCTCATTGTGGTTTGTCAGTAC
Sequenced genes	**Oligonucleotides for sequencing**	**Sequences (5´→3´)**
All above except for *saci_1085*	T7-promoter	TAATACGACTCACTATAGGG
	T7-terminator	GCTAGTTATTGCTCAGCGG
Saci_1085	FXara-fd	CAGCGTTTATAACGTTTAACATG
	FXsulf-rv	CCATTTAATAGTTTGTATGGTCTACCC

Table S 2**.** Plasmids and strains used for cloning and expression of the FA metabolic enzymes from *S. acidocaldarius*.

Plasmids	**Genotype or description**	**Source/Reference**
pET15b	*E.coli* expression plasmid carrying an N-terminal His Tag for cloning of *saci_1122, saci_1104, saci_1115, saci_0315* and *saci_1085*	Novagen, USA
pET28b	*E.coli* expression plasmid carrying both an N-terminal and a C-terminal His Tag for cloning of *saci_1114*	Novagen, USA
pET45b	*E.coli* expression plasmid carrying an N-terminal His Tag and a C-terminal S Tag for cloning of *saci_1123, saci_1109*	Novagen, USA
pBS-araFX-UTR-CtSS	*S. acidocaldarius* expression plasmid carrying a C-terminal Twin Strep Tag for cloning of *saci_1085*	unpublished
Strains	**Function**	**Source/Reference**
E. coli DH5 *α*	Plasmid construction	Hanahan, USA
E. coli Rosetta (DE3)	Heterologous gene expression	Stratagene, USA
S. acidocaldarius MW001	Homologous gene expression	[1]

Table S 3**.** Expression conditions used for the selected genes from *S. acidocaldarius*.

Enzymes	ORFs	Overexpression conditions
Acyl-CoA synthase (ACS)	*Saci_1122*	0.5 mM IPTG for induction; 37°C overnight
Acyl-CoA dehydrogenase (ACAD)	*Saci_1123*	0.4 mM IPTG for induction; 20°C overnight
Electron transfer flavoprotein (ETF)	*Saci_0315*	0.4 mM IPTG for induction; 20°C overnight
Enoy-CoA hydratase/3-hydroxyacyl-CoA dehydrogenase (ECH/HCDH)	*Saci_1109*	1 mM IPTG for induction; 30°C overnight
β-Ketothiolase/Acetyl-CoA acetyltransferase (KT/ACAT)	*Saci_1114*	1 mM IPTG for induction; 22°C overnight
Acetoacyl-CoA reductase (ACR)	*Saci_1104*	1 mM IPTG for induction; 37°C 4 hours
Enoyl-CoA reductase (ECR)	*Saci_1115*	1 mM IPTG for induction; 16°C overnight
MaoC 3-hydroxyacyl-CoA dehydratase (MaoC-HCD)	*Saci_1085*	0.3% D-xylose for induction; 75°C 48 hours

Table S 4**.** Retention times of the corresponding CoA ester compounds in different HPLC running programs. The relevant HPLC chromatographs were shown in Fig. S19 (ND: not detected or not detectable).

Peak no.	Compound	Retention time (min)	
		in 4-30% ACN program:	in 1-60% ACN program:
1	HS-CoA	12.5	8.2
2	Acetyl-CoA	21.5	9.7
3	Butyryl-CoA	29.9	25.5
4	Crotonoyl-CoA (C4:1)	29	ND
5	3-Hydroxybutyryl-CoA	23.5	ND
5a	3-(S)-Hydroxybutyryl-CoA	23.5	ND
5b	3-(R)-Hydroxybutyryl-CoA	23.5	ND
6	Acetoacetyl-CoA	22.5	ND
7	Hexanoyl-CoA	35.5	31
8	Octanoyl-CoA	ND	34
9	Hexenoyl-CoA (C6:1)	ND	30.3
10	3-Hydroxyhexanoyl-CoA	ND	24.5
11	Octenoyl-CoA (C8:1)	ND	33.9
12	3-Hydroxyoctanoyl-CoA	ND	30.6

Table S 5. Kinetic parameters of the AMP-forming acyl-CoA synthase from *S. acidocaldarius* (* The kinetic parameters were calculated in regardless of inhibition effect caused by higher concentration of the substrate; Numbers after ± represent standard error (SE); ND: not detected).

ORF/ Enzyme	Native mass (kDa) (Oligomer)	Temperature	Substrate/Cofactor	K_m (mM)	V_{max} (U mg^{-1})	k_{cat} (S^{-1})
Saci_1122 Acyl-CoA synthase (ACS)	117 (homodimer)	55°C	Butyric acid (C4)	ND	1.34	ND
			Valeric acid (C5)	*0.046 ± 0.009	*2.106 ± 0.056	*2.35
			CoA	0.575 ± 0.134		
			ATP	*0.965 ± 0.151		
			Octanoic acid (C8)	*0.01 ± 0.002	*3.282 ± 0.101	*3.352
		70°C	Acetyl-CoA (C2)	ND	0.007	ND
			Butyryl-CoA (C4)	ND	0.14	ND
			Octanoyl-CoA (C8)	ND	0.181	ND
			Palmitoyl-CoA (C16)	ND	0	ND

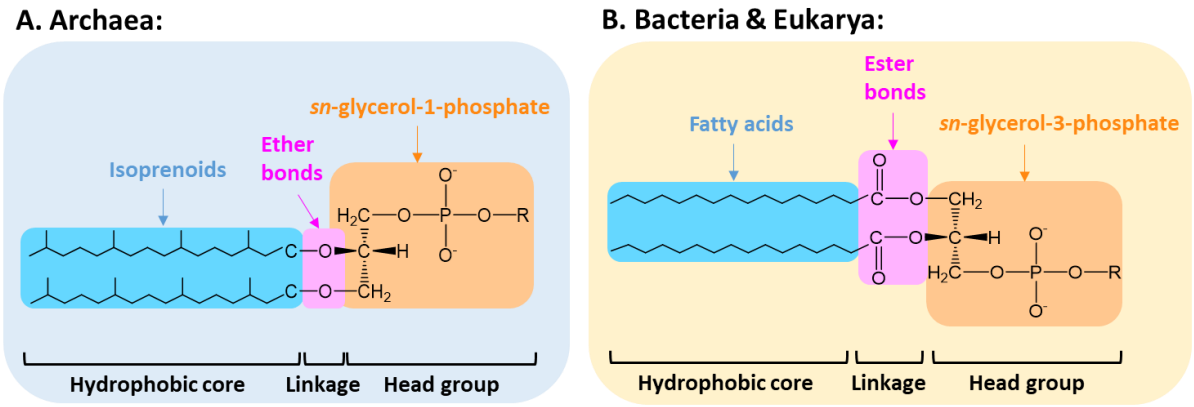


Figure S 1**. Lipid compositions in Archaea (A), Bacteria and Eukaryotes (B).** Archaea produce isoprenoid-based membrane lipids ether-linked to G1P while the cell membranes of Bacteria and Eukarya are composed of fatty acid based-lipids ester-bound to G3P. [2]

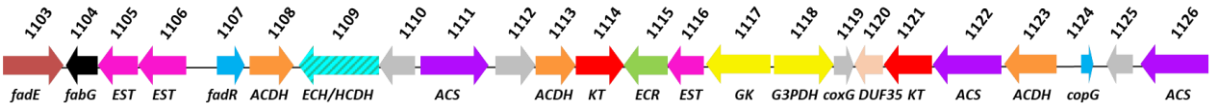


Figure S 2. **Genomic organization of the *saci_1103-saci_1126* gene cluster related to lipid and fatty acid metabolism in *S. acidocaldarius*.** Within this fatty acid operon, several copies of lipases/esterases (*saci_1105, saci_1106* and *saci_1116*, shown in pink) are present, two of them have been characterized [3]. Two genes are predicted to be involved in glycerol metabolism (*saci_1117* and *saci_1118*, displayed in yellow). The genes encode AMP-forming ACSs, which can synthesize acyl-CoAs as precursors for β oxidation, were shown in purple (*saci_1111, saci_1122* and *saci_1126*). One or more paralogs for each steps of the bacterial-type β oxidation could be found as well, for instance, three acyl-CoA dehydrogenases (*saci_1108, saci_1113* and *saci_1123*, colored in orange), a bifunctional enoyl-CoA hydratase/3-hydroxyacyl-CoA dehydrogenase (*saci_1109*, in blue-green upward diagonal) and two β-ketothiolases/acety-CoA C-acetyltransferases (*saci_1114* and *saci_1121*, in red). Moreover, a fabG homolog (*saci_1104*, in black) and an enoyl-CoA reductase (*saci_1115*, in green) are thought to be responsible for elongation of fatty acyl chains. The genes with blue color encode transcription regulators (*saci_1107* and *saci_1124*). [4]

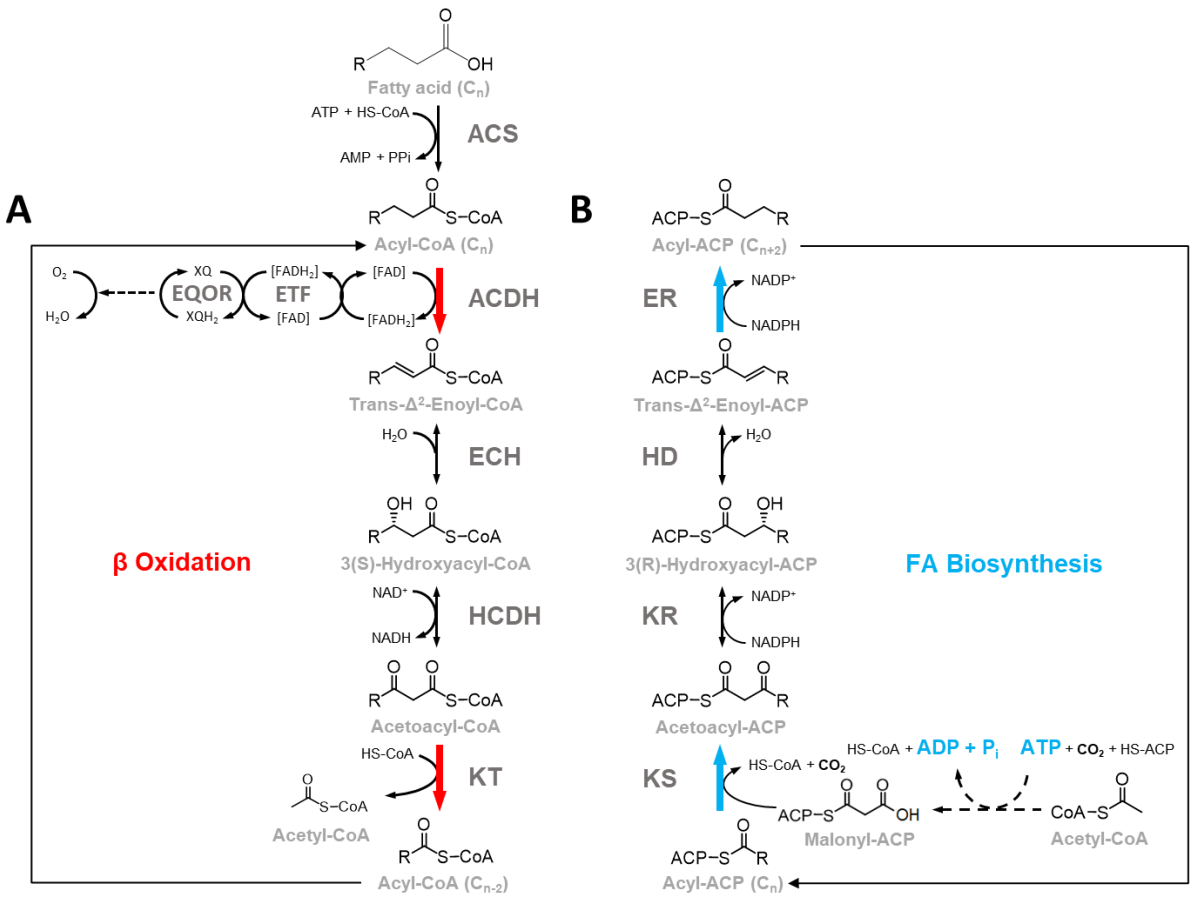


Figure S 3**. The classical fatty acid degradation and biosynthesis pathway.** Fatty acids are initially activated into acyl-CoAs by an AMP-forming acyl-CoA synthetase (ACS) and then degraded via β oxidation (A) as the following steps. Fatty acyl CoAs are oxidized to enyol-CoAs by acyl-CoA dehydrogenase (ACAD). Then, the bifunctional 3-hydroxyacyl-CoA dehydrogenase (HCDH)/enoyl-CoA hydratase (ECH) catalyzes the next hydration and the second oxidation steps and converts enoyl-CoA to acetoacyl-CoA via the 3(S)-hydroxyacyl-CoA intermediate. Afterwards, the β-ketothiolase (KT) thiolytically cleaves acetoacyl-CoA into acetyl-CoA and the acyl-CoA with two carbon atoms less than before. The resulting two-carbon shortened acyl-CoA can further enter the β oxidation cycle until all the fatty acyl chains are broken into acetyl-CoA. In the classic fatty acid biosynthesis pathway (B), malonyl-ACP is used as the elongation unit for the fatty acyl chain, so malonyl moiety is formed from acetyl-CoA via carboxylation and transferred from CoA to acyl carrier protein (ACP). In the initial reaction, acetoacyl-ACP (also named as 3-ketoacyl-ACP) is produced by the 3-ketoacyl-ACP synthase (KS) through the Claisen condensation and CO_2 is released. Next, NADPH-dependent 3-ketoacyl-ACP reductase (KR) converts acetoacyl-ACP to D-3-hydroxyacyl-ACP (also named as 3R-hydroxyacyl-ACP). In the following reaction, one molecule of water is removed from D-3-hydroxyacyl-ACP to form enoyl-ACP by 3-hydroxyacyl-ACP dehydratase (HD). Finally, enoyl-ACP is reduced to acyl-ACP by enyol-ACP reductase (ER). The forming acyl-ACP can be further elongated by the synthetic cycle until the desired chain length is reached. [5] Abbreviation: ETF, electron transfer flavoprotein; EQOR: ETF: quinone oxidoreductase.

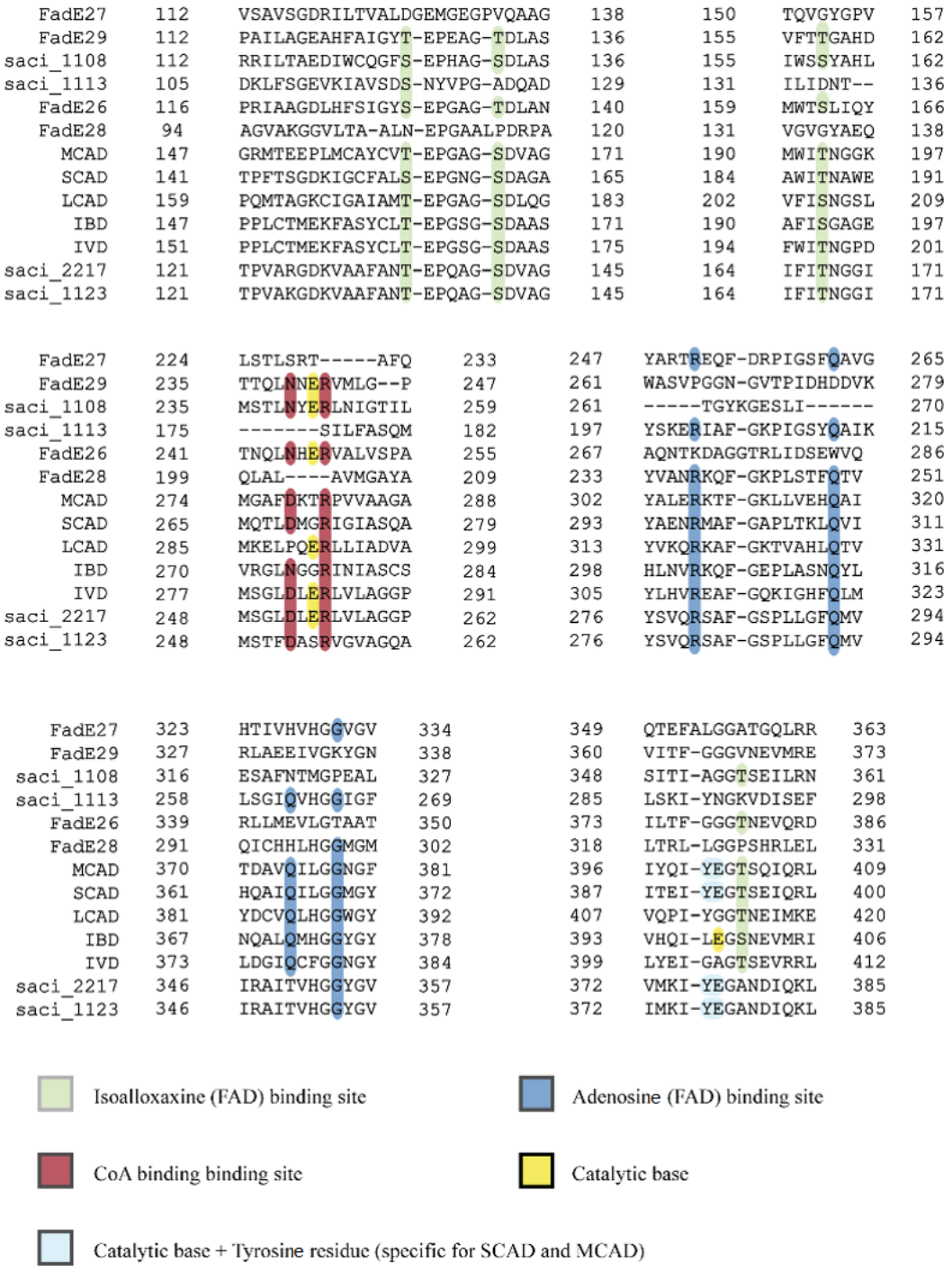


Figure S 4. **Sequence alignment of various acyl-CoA dehydrogenases.** FadE26, 27, 28 and 29 represent the *Mycobabterium* acyl-CoA dehydrogenases [6, 7] whereas Saci_1108, Saci_1113, Saci_1123 and Saci_2217 are the candidates from *S. acidocaldarius*. SCAD, MCAD or LCAD is short for the short-chain, medium-chain or long-chain acyl-CoA dehydrogenase from mammals, respectively [6, 8]. IBD (isobutyryl-CoA dehydrogenase) and IVD (isovaleryl-CoA dehydrogenase) denote the homologues from human [6]. The general catalytic base for acyl-CoA dehydrogenases is glutamic acid (in yellow) while the ones especially for the short- and medium-chain homologues are tyrosine and glutamic acid (in light blue). Amino acid residues shown in red (N/DXXR) are coenzyme A binding sites. The binding sites for cofactor FAD are shown in light green (residues threonines and serines for isoalloxaxine of FAD) and dark blue (aginines, glutamines and glycines for adenosine of FAD), respectively.

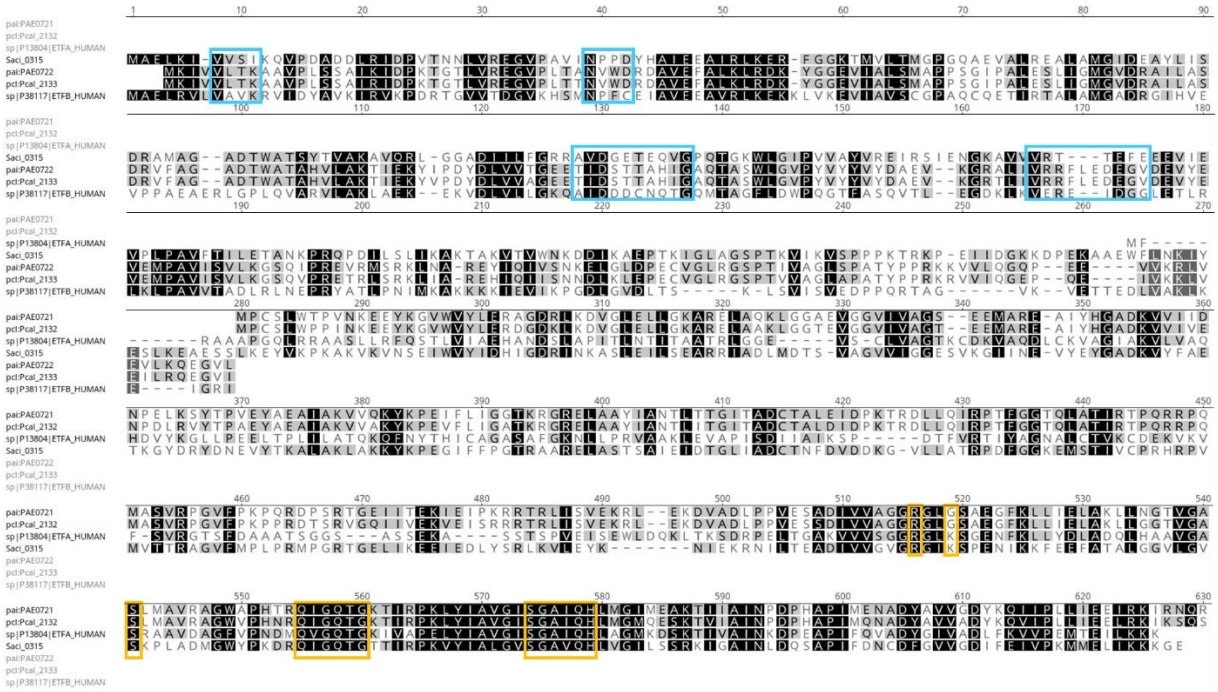


Figure S 5. **Sequence alignment of various electron transfer flavoproteins (ETFs).** The amino acid sequences pai:PAE0721, pcl:Pcal_2132 and sp[P13804]ETFA_HUMAN represent the α subunit of the ETF proteins from *Pyrobaculum aerophilum, Pyrobaculum calidifontis and human, respectively whereas* pai:PAE0722, pcl:Pcal_2133 and sp[P38117]ETFB_HUMAN indicate the corresponding β subunits. The *S. acidocaldarius* fusion ETF containing both subunits was regarded as Saci_0315. Blue boxes displayed the binding sites for AMP or one FAD while the highly conserved motifs for the other FAD was indicated in yellow boxes. Thus, the fusion ETF protein Saci_0315 includes both active sites for the cofactors.

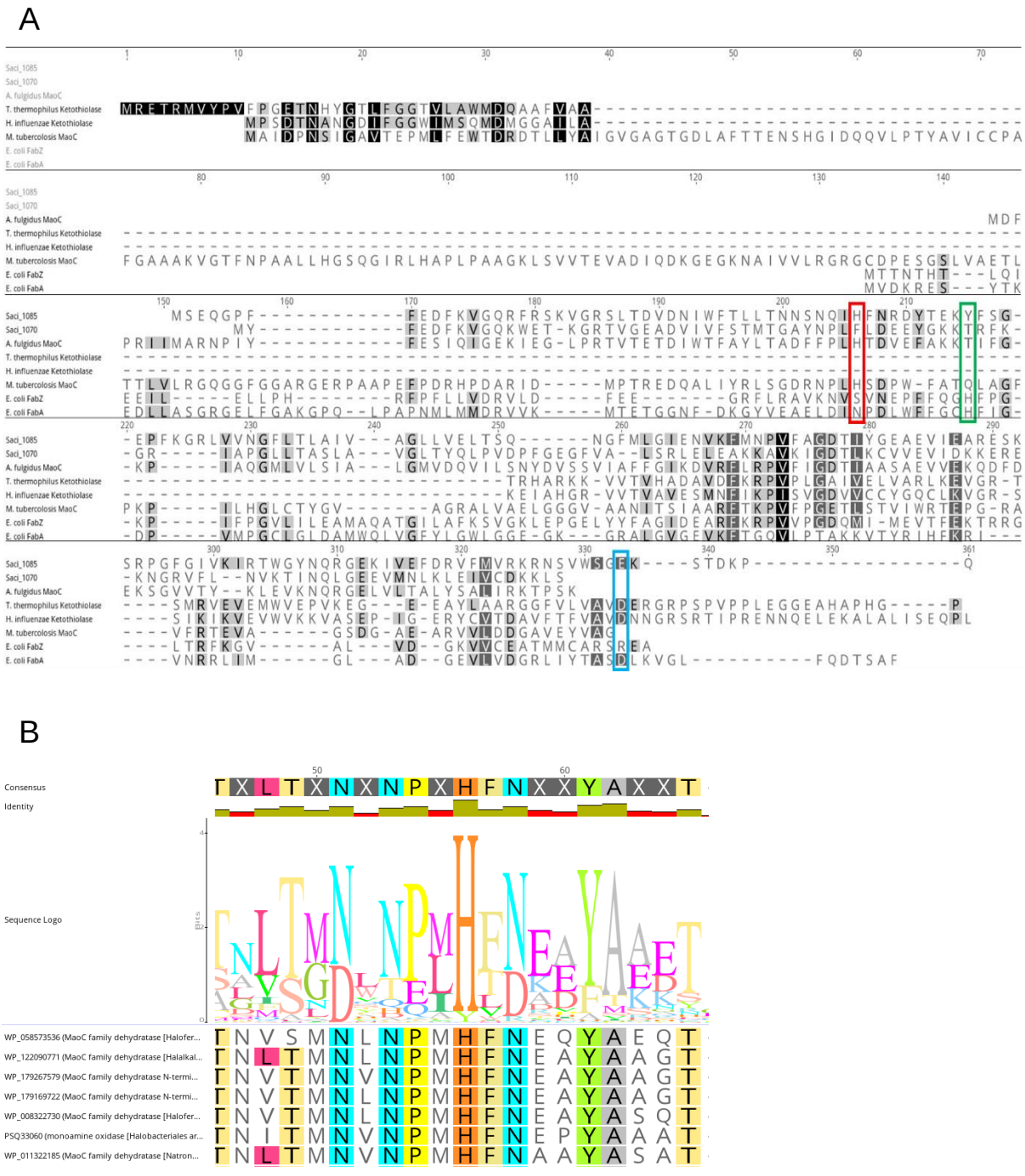


Figure S 6. **Catalytic sites for different hot-dog fold enzymes (A) and conserved motifs for MaoC like dehydratase homologues (B).** Different sequences were aligned and hot-dog fold enzymes of *S. acidocaldarius* were screened for conserved catalytic motifs. The catalytic histidine residue of FabA and FabZ is absent in all other hot-dog fold sequences from Bacteria like *Thermus thermophiles, Haemophilus influenza or Mycobacterium tuberculosis* and Archaea such as *Archaeoglobus fulgidus* or *S. acidocaldarius* (A, green box). The hot-dog fold thiolases contain an aspartic acid residue as the catalytic motif (A, blue box) while Saci_1085 contains a conserved histidine residue characteristic for the hot-dog fold enoyl-CoA hydratases (A, red box), together with either an aspartate or asparagine residue, which is in line with almost all the MaoC dehydratases (B). However, a second *S. acidocaldarius* MaoC dehydratase candidate Saci_1070 lacks this key histidine motif (A, red box).

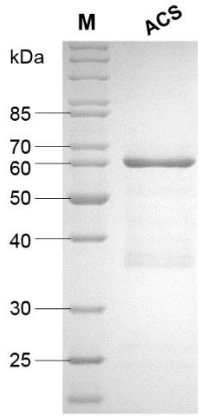


Figure S 7**. The recombinant acyl-CoA synthetase from *S. acidocaldarius* after purification.** ACS: SDS-PAGE (12.5%, Coomassie Blue staining) of 3 µg of the purified AMP-forming acyl-CoA synthetase Saci_1122; M: 5 µl PageRuler™ Unstained Protein Ladder (Fermentas).

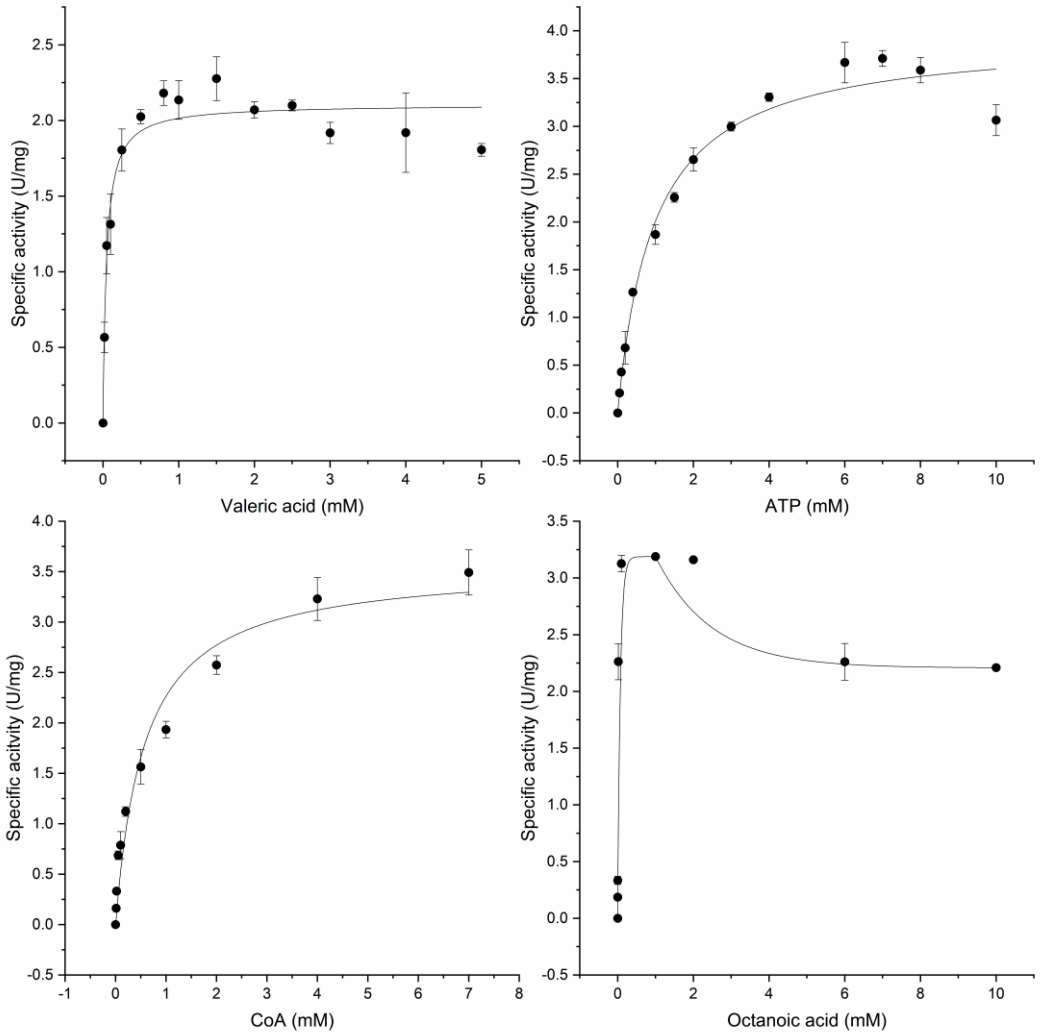


Figure S 8**. Determination of kinetic parameters of the recombinant ACS Saci_1122 from *S. acidocaldarius*.** The ACS activity was determined in couple with the auxiliary enzymes myokinase (MK), pyruvate kinase (PK) and lactate dehydrogenase (LDH) from rabbit muscle by monitoring NADH oxidation at 340 nm. The assay was performed at 55°C, pH 7 in 100 mM HEPES/NaOH buffer plus 20 mM $MgCl_2$ with 2 mM CoA, 5 mM ATP, 6 mM PEP, 0.2 mM NADH, 11.4U MK, 4.6 U PK, 4.2 U LDH and 6.7 µg pure Saci_1122 in addition to 2 mM FA. The K_m values for CoA (C), ATP (B), valeric acid (A) or octanoic acid (D) were separately determined. The independent measurements were performed in triplicate and error bars indicate the standard error of the mean (SEM).

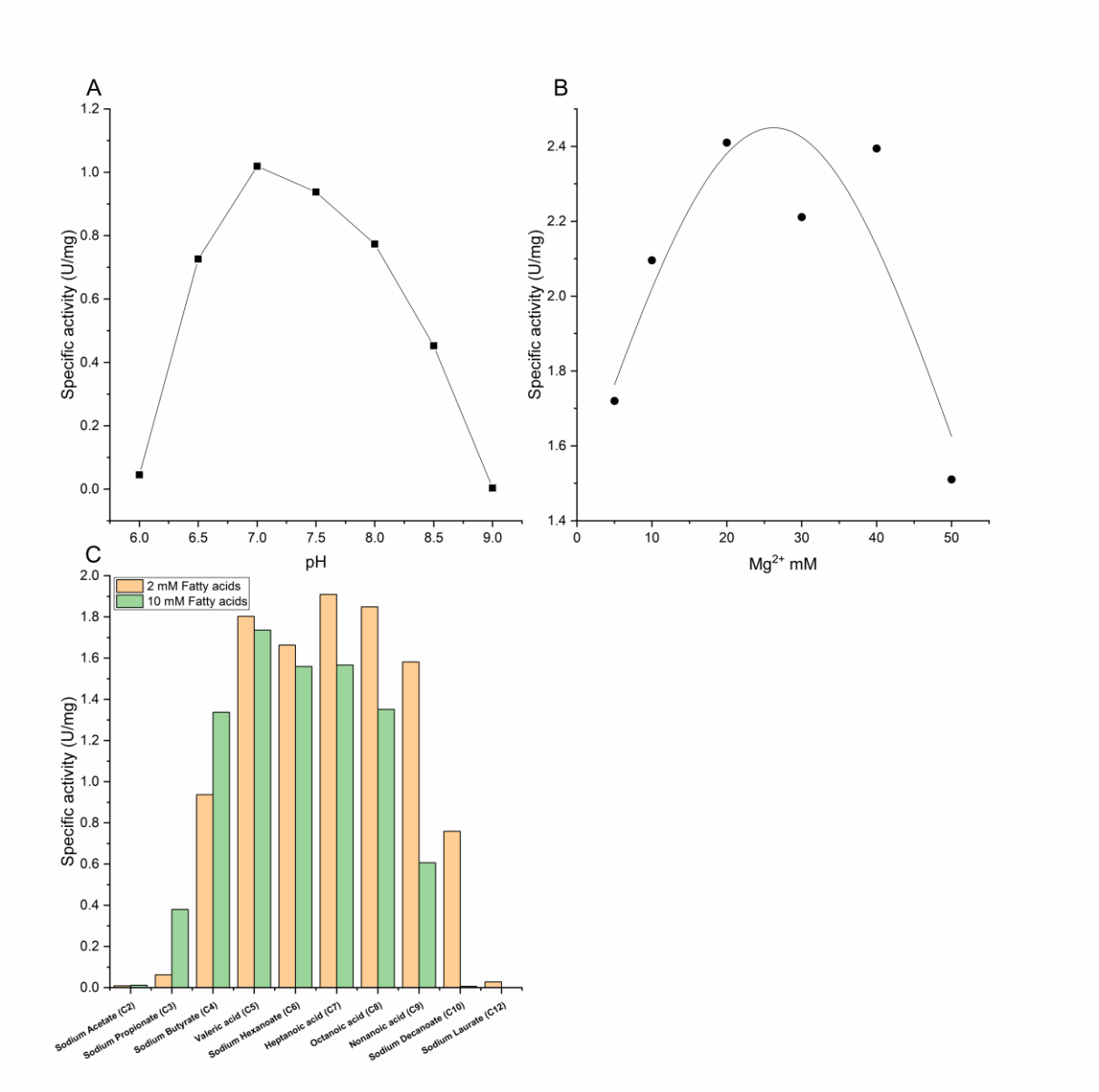


Figure S 9**. Determination of the optimal pH (A), Mg^{2+} dependency (B) and substrate spectrum of the recombinant ACS Saci_1122 from *S. acidocaldarius*.** The optimal pH was identified as 7 at 55°C with a mixed buffer of 50 mM MES, 50 mM HEPES and 50 mM Tris setting the pH values from 6 to 9. The mixture contained 5 mM $MgCl_2$, 10 mM valeric acid, 2 mM CoA, 5 mM ATP, 6 mM PEP, 0.4 mM NADH with 5.7 U MK, 2.3 U PK, 2.1 U LDH and 10.18 µg Saci_1122. Mg^{2+} dependency was studied in 100 mM HEPES/NaOH buffer (pH 7) with the same assay but varying the Mg^{2+} concentration from 0 to 50 mM. In the assay for analysis of the substrate specificity 10 mM $MgCl_2$, 0.2 mM NADH, 11.4 U MK, 4.6 U PK, 4.2 U LDH and 20.35 µg Saci_1122 as well as 2 or 10 mM of different FAs were included.

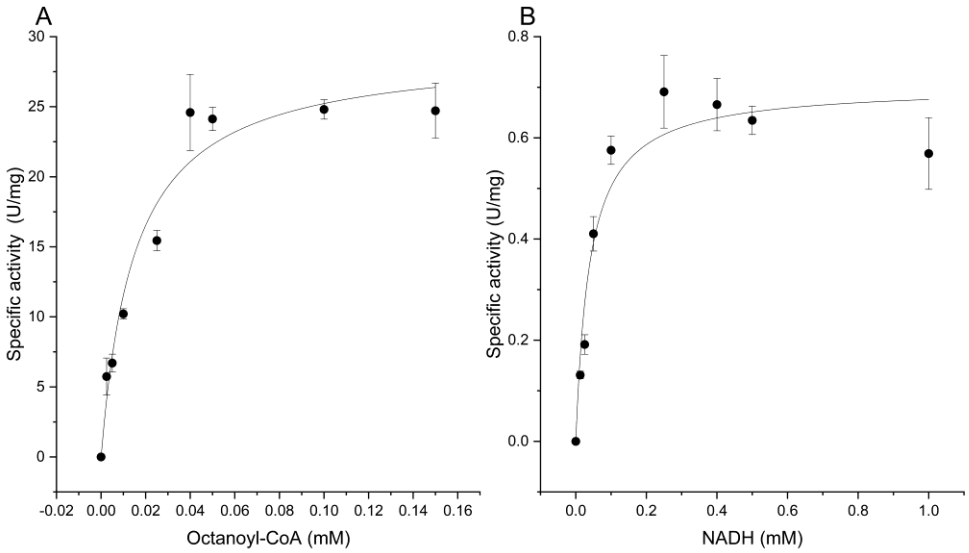


Figure S 10. **Investigation of kinetic properties of the recombinant ACAD Saci_1123 (A) and ETF Saci_0315 (B) from *S. acidocaldarius*.** The specific activity and K_m of Saci_1123 was tested at 65°C, pH 6.5 in 50 mM HEPES/KOH buffer plus 20 mM KCl, 0.13 µg/µl ACAD with 1 mM $FcPF_6$ and 0-0.15 mM octanoyl-CoA. Reduction of $FcPF_6$ was monitored at 300 nm (extinction coefficient 4.3 mM^{-1} cm^{-1}). For detecting the NADH-linked EtfAB activity of Saci_0315, the assay was carried out in 50 mM HEPES/NaOH (pH 7.5) containing 100 mM NaCl, 0.2 mM iodonitrotetrazolium chloride (INT) and 0.015 µg/µl ETF protein with 0-1 mM NADH for K_m measurement. The activity was determined by monitoring the release of the red formazan at 500 nm (extinction coefficient 19.3 mM^{-1} cm^{-1}). The independent measurements were performed in triplicate and error bars indicate the standard error of the mean (SEM).

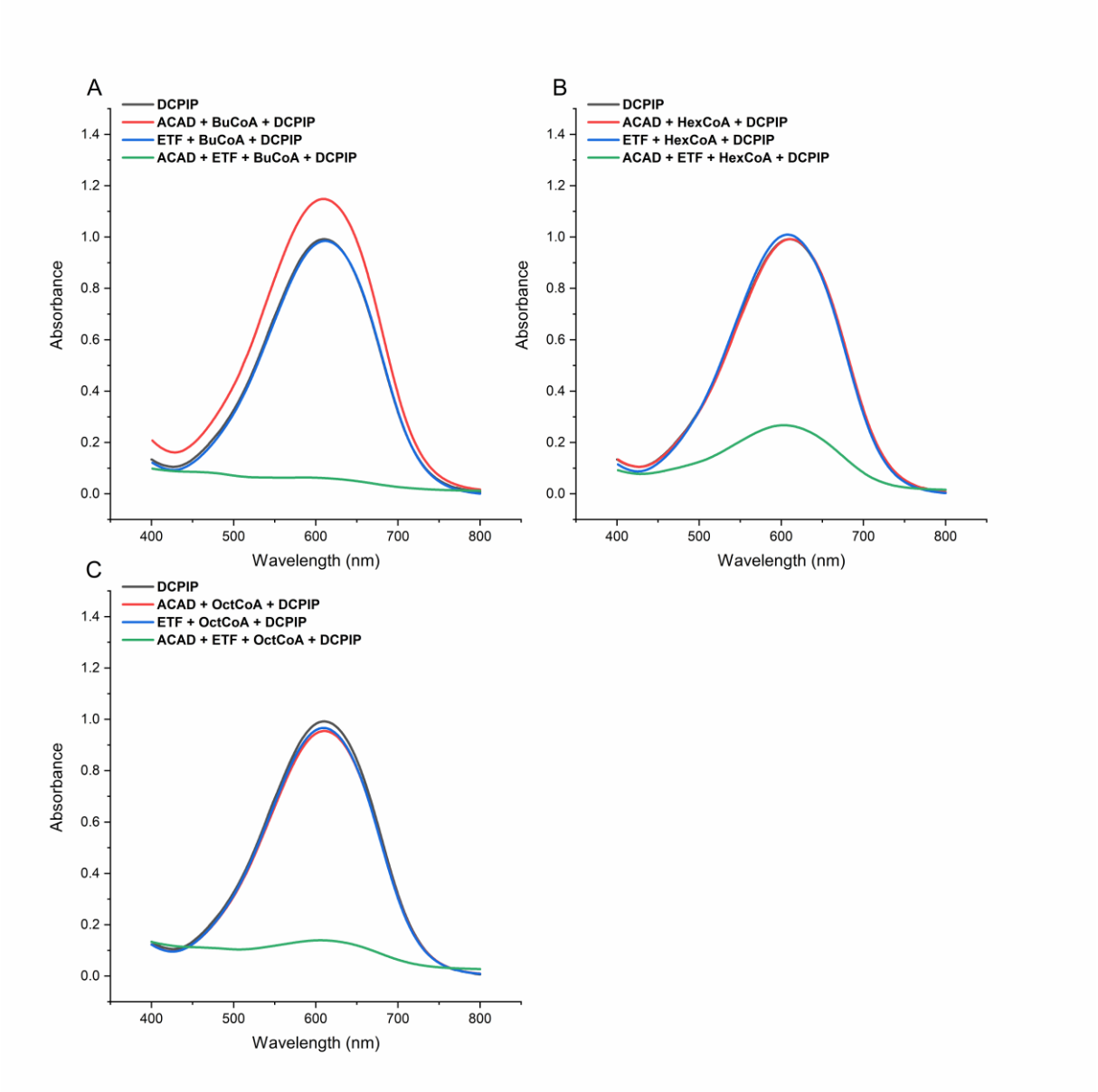


Figure S 11. **Investigation of electron transfer by ACAD and ETF.** The assay mixture (0.5 ml) contained 50 mM MES/KOH, 20 mM KCl, 0.2 mM DCPIP, 0.2 mM of different acyl-CoA (butyryl-CoA (A), hexanoyl-CoA (B) or octanoyl-CoA(C)), 1.7 µg ACAD as well as 3.8 µg ETF and was incubated at 65°C, pH 6.5 for 5 min. Afterwards, the DCPIP spectrum in each sample was determined under the wavelength from 400 to 800 nm. DCPIP exhibits the highest absorbance at 600 nm. Therefore, loss of absorbance at 600 nm indicated depletion of DCPIP due to the electron transfer from ACAD to DCPIP through ETF. The results suggested that all the employed acyl-CoAs require presence of both ACAD and ETF for transporting electrons.

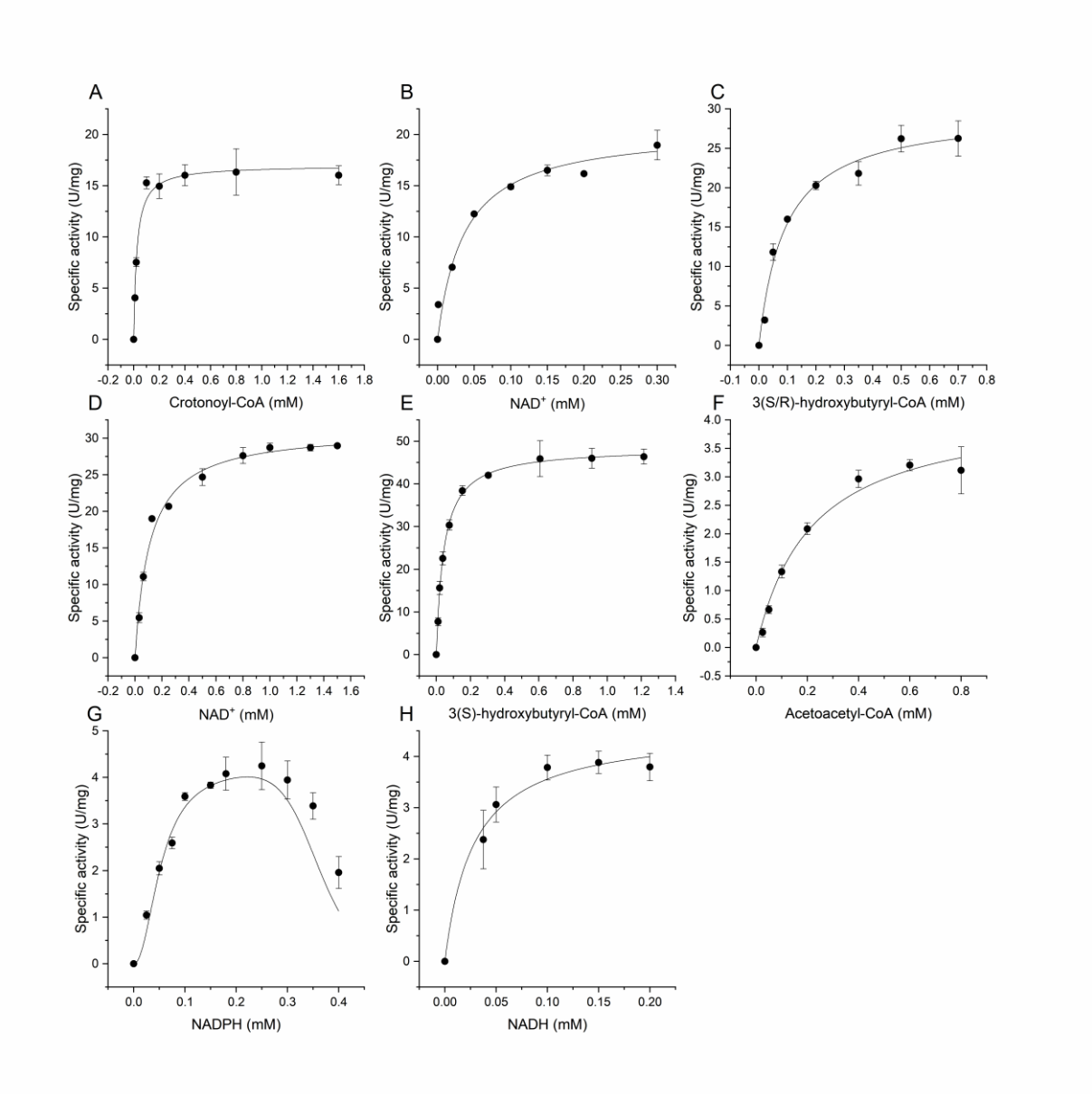


Figure S 12. **Investigation of kinetic properties of the recombinant HCDH/ECH Saci_1109 from *S. acidocaldarius*.** The enzymatic activity of HCDH/ECH was determined at pH 7 in 100 mM Tris/HCl (500 µl). The formation or consumption of NAD(P)H was monitored at 340 nm. The assay for oxidative reactions was performed at 70°C with 0.69 µg protein, 0.2 mM NAD^+ and 0.4 mM of crotonoyl-CoA. For single oxidation of 3-HBCoA, mixed 3(S/R)-HBCoA or the single 3(S)-HBCoA was applied instead of crotonyl-CoA. The K_m values for crotonoyl-CoA (A), 3(S/R)-HBCoA (C), 3(S)-HBCoA (E) and NAD^+ (B with crotonoyl-CoA as substrate; D with substrate 3(S/R)-HBCoA) were individually determined. For detection of the reverse activity at 35°C, 0.6 mM acetoacetyl-CoA was employed as substrate in addition of 0.2 mM NADPH and 4.05 µg protein. The K_m values for AcAcCoA (F), NADPH (G) and NADH (H) were measured, respectively. The independent measurements were performed in triplicate and error bars indicate the standard error of the mean (SEM).

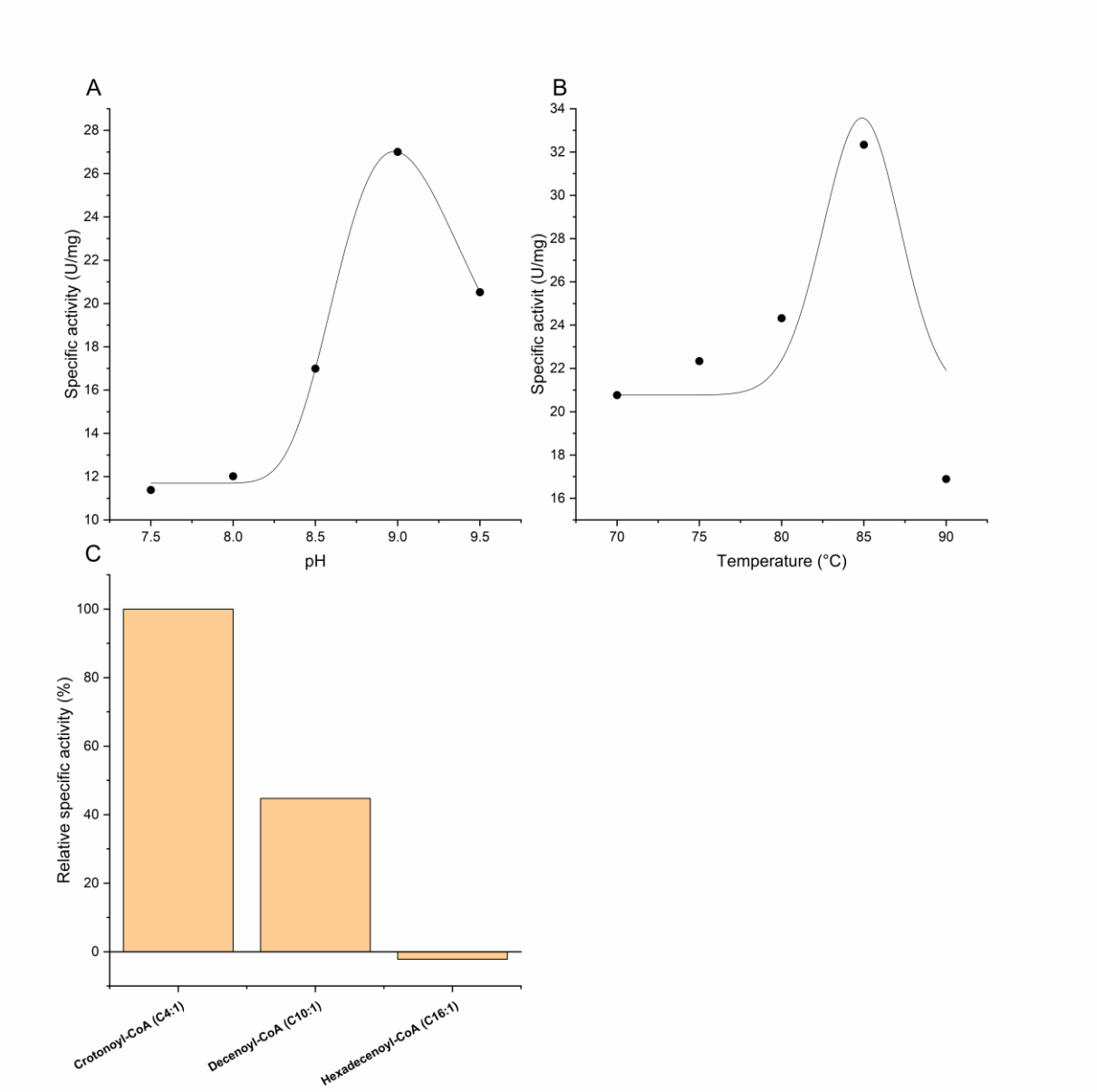


Figure S 13**. Determination of the optimal pH (A) and temperature (B) as well as the substrate spectrum (C) of the recombinant HCDH/ECH Saci_1109 from *S. acidocaldarius*.** The optimal pH was determined in a mixed buffer of 50 mM MES, 50 mM HEPES and 50 mM Tris at 70°C. The assay contained 0.4 mM crotonoyl-CoA, 0.2 mM NAD^+ and 0.69 µg Saci_1109. The catalytic temperature optimum was detected in 100 mM HEPES/NaOH (pH 8) utilizing the same assay. The substrate spectrum was performed against 0.3 mM of different enoyl-CoAs (crotonoyl-CoA, decenoyl-CoA or hexadecenoyl-CoA). The assay (400 µl) was done at 75°C, pH 7 with 0.2 mM NAD^+ and 0.36 µg protein.

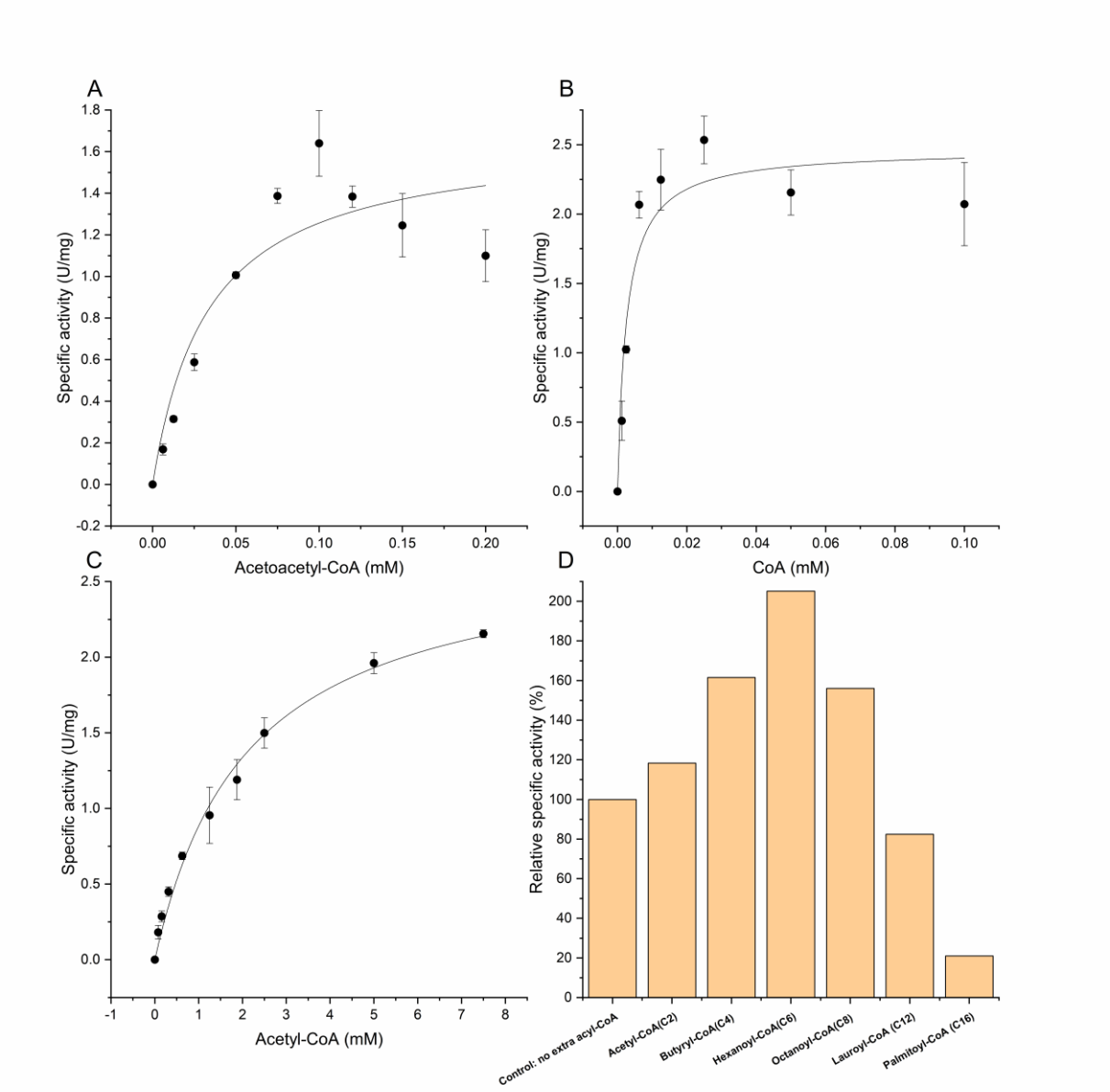


Figure S 14**. Determination of kinetic parameters and the substrate spectrum of the recombinant KT from *S. acidocaldarius*.** The specific activity of KT was photometrically tested at room temperature (23°C), pH 8 by monitoring the decrease of Mg^{2+}-AcAcCoA chelates (extinction coefficient of 21.4 mM^{-1} cm^{-1}) at 303 nm under the UV light. The reaction mixture contained 100 mM Tris/HCl, 20 mM $MgCl_2$, 0.2 mM CoA, 0.1 mM AcAcCoA and 2.7 µg Saci_1114. To calculate K_m values, variable concentrations of AcAcCoA (0-0.2 mM) (A) or CoA (0-0.1 mM) (B) were employed. The reversed activity was determined in couple with HCDH/ECH Saci_1109 by detecting the NADH oxidation at 340 nm. The enzyme assay included 100 mM MOPS/NaOH (pH 6.5 at 75°C), 0.3 mM NADH, 17.1 µg Saci_1109, 10.8 µg Saci_1114 and 0-7.5 mM acetyl-CoA for K_m measurement (C). The independent measurements were performed in triplicate and error bars indicate the standard error of the mean (SEM). The substrate preference (D) was determined by including 2.5 mM acetyl-CoA and 0.5 mM of extra acyl-CoA with different chain lengths (acetyl-CoA (C2), butyryl-CoA (C4), hexanoyl-CoA (C6), octanoyl-CoA (C8), lauroyl-CoA (C12) and palmitoyl-CoA (C16), separately).

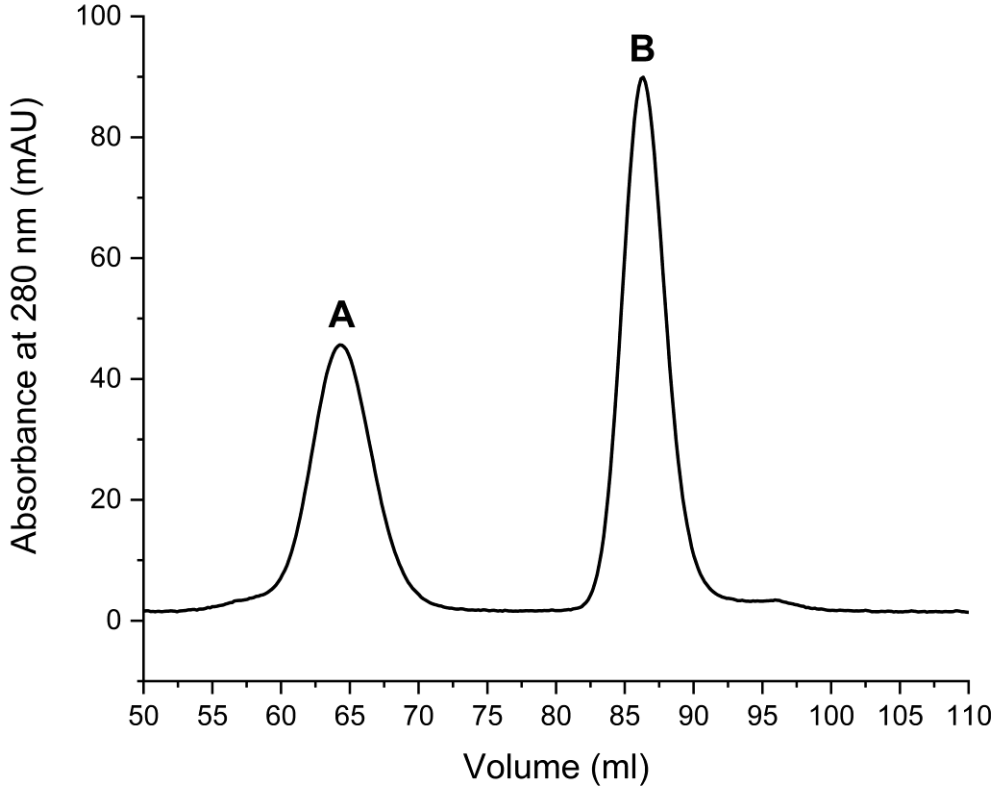


Figure S 15. **Gel filtration of the recombinant ECH/HCDH Saci_1109 and KT Saci_1114 from *S. acidocaldarius*.** Around 0.015 µmol of each of the purified recombinant ECH/HCDH and KT proteins were mixed and incubated on ice for 4 hours. Then, the protein mixture was applied to a Superdex 200 prep grad HiLoad 16/60 gel filtration column (GE Healthcare Life Sciences, Freiburg, Germany). An elution buffer containing 50 mM HEPES/NaOH and 300 mM NaCl (pH 7.2) was employed. Afterwards, two separate peaks representing the respective HCDH/ECH (A, 64.33 ml) and KT (B, 86.29 ml) were obtained indicating no complex formation between these two proteins under the experimental conditions.

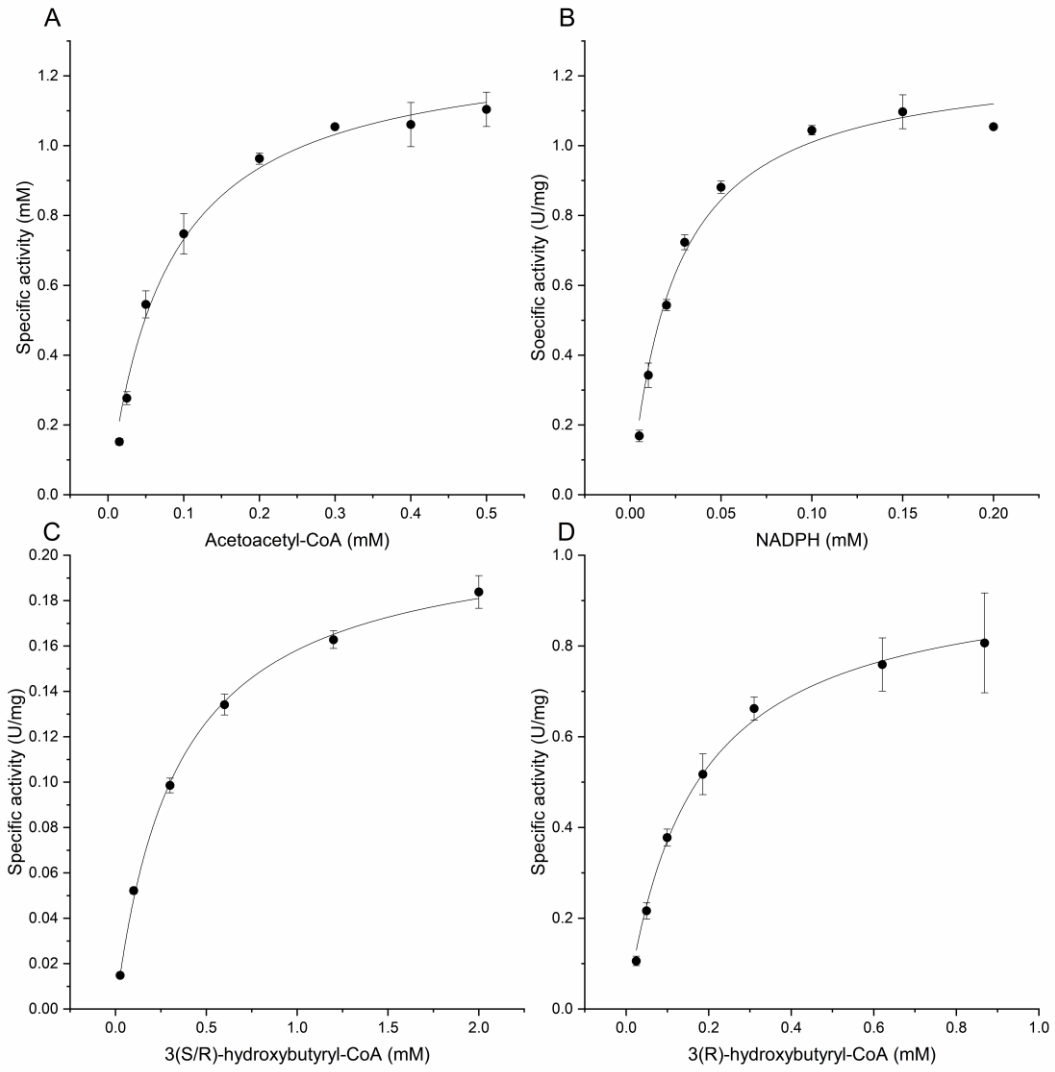


Figure S 16. **Investigation of the kinetic parameters of the recombinant ACR Saci_1104 from *S. acidocaldarius*.** The ACR activity was determined in 100 mM Tris/HCl with 0.3 mM AcAcCoA, 0.2 mM NADH/NADPH and 4.03 µg pure protein at 35°C, pH 7 (340 nm). The K_m values were determined for AcAcCoA (A) and NADPH (B), respectively. The reversed activity was determined at 70°C with the commercial, mixed 3(S/R)-hydroxybutyryl-CoA or single 3(R)-hydroxybutyryl-CoA as substrate in presence of 2 mM $NADP^+$ and 20.16 µg purified protein. The K_m values for 3-HBCoA reduction were measured with a variable concentration of 3(S/R)-HBCoA (0-2 mM) (C) or 3(R)-HBCoA (0-1 mM) (D). The independent measurements were performed in triplicate and error bars indicate the standard error of the mean (SEM).

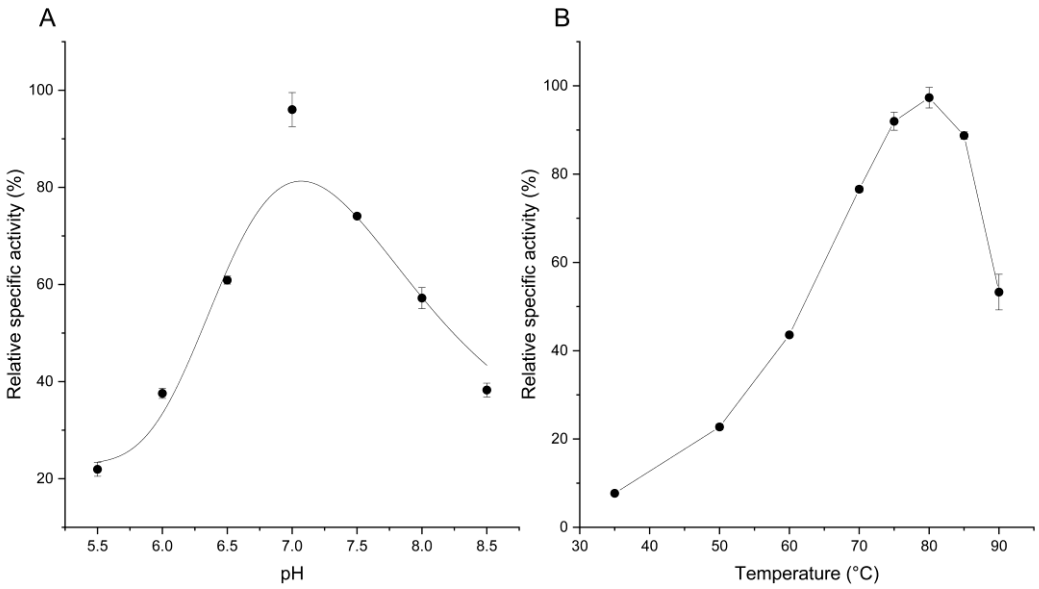


Figure S 17**. Identification of the optimal pH (A) and temperature (B) of the recombinant ACR Saci_1104 from *S. acidocaldarius*.** The optimal pH was determined in the direction of 3-hydroxybutyryl-CoA formation at 35°C using 100 mM MES/NaOH (pH 5.5-6.5) and 100 mM Tris /HCl (pH 7.0-8.5) as buffers. The assay contained 0.3 mM AcAcCoA, 0.2 mM NADPH and 4.03 µg protein. The catalytic temperature optimum was detected in the direction of acetoacetyl-CoA formation at pH 7 between 35-90°C. The assay was done in 100 mM Tris/HCl containing 2 mM $NADP^+$, 0.3 mM 3(S/R)-HBCoA and 20.16 µg protein

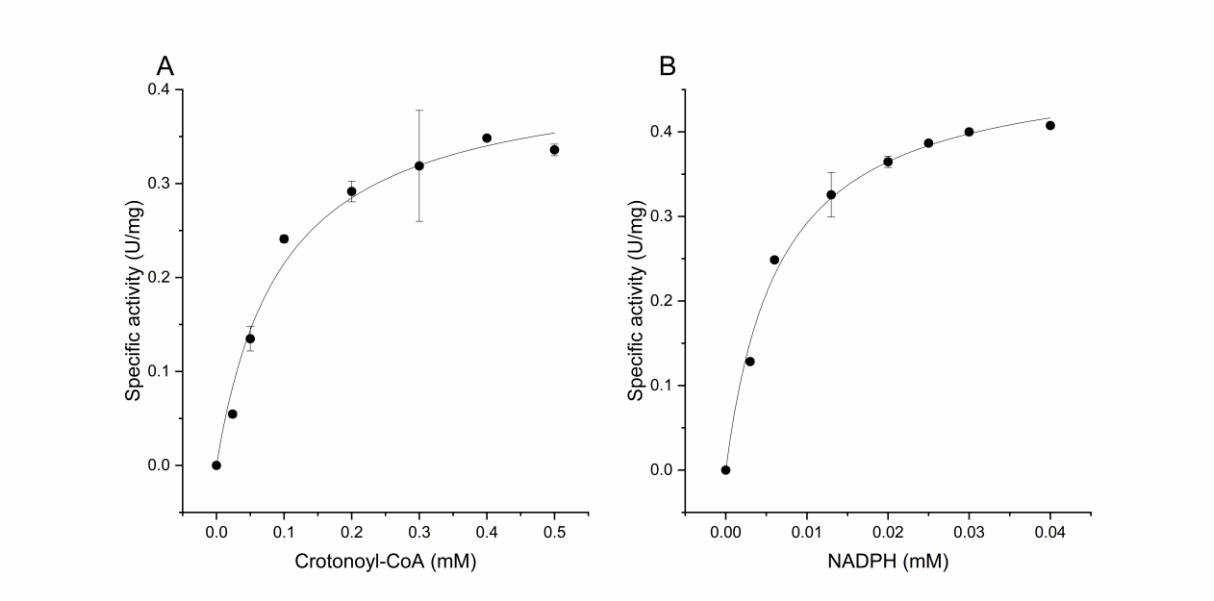


Figure S 18**. Determination of kinetic parameters of the recombinant ECR Saci_1115 from *S. acidocaldarius*.** The ACR activity was determined at 70°C, pH 7.5 (340 nm) with 12 µg pure protein in 100 mM HEPES/NaOH including 10 mM KCl, 0.3 mM NADPH and 0.4 mM crotonoyl-CoA in a total volume of 0.5 ml. The K_m values for crotonoyl-CoA and NADPH were investigated by varying the concentrations from 0-0.5 mM (A) and 0-0.04 mM (B), respectively. The independent measurements were performed in duplicate and error bars indicate the standard error of the mean (SEM).

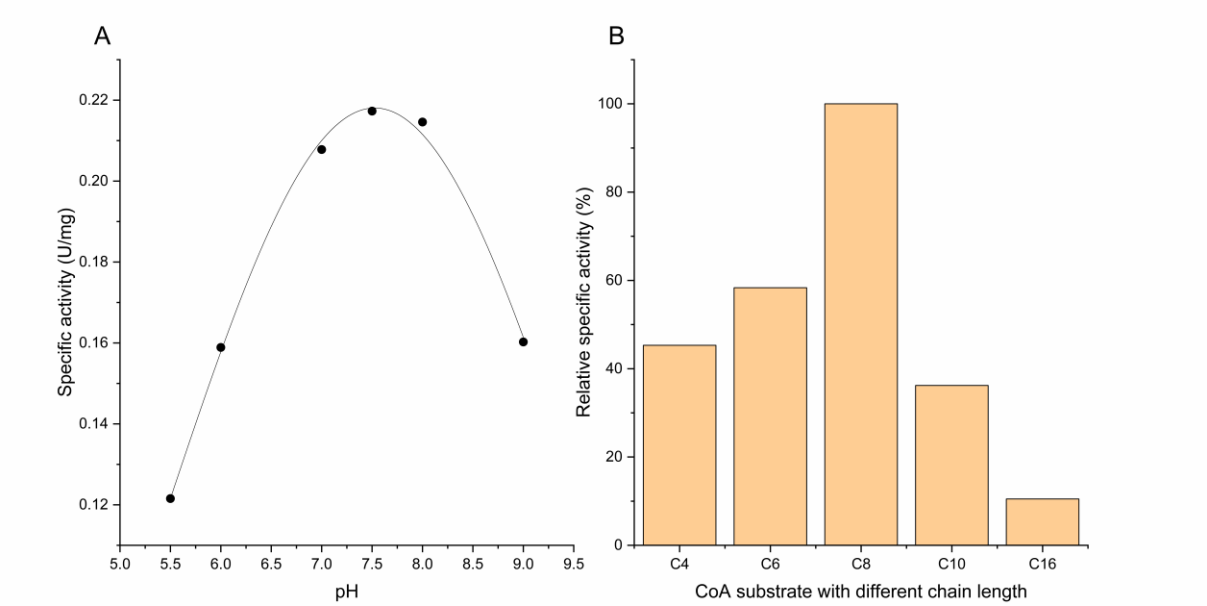


Figure S 19**. Determination of the pH optimum (A) and the substrate spectrum (B) of the recombinant ECR Saci_1115 from *S. acidocaldarius*.** The optimal pH for ECR was determined as 7.5 at 70°C using mixed buffer of 0.5 M HEPES, 0.5 M Tris and 0.5 M MES in presence of 10 mM KCl, 0.4 mM crotonoyl-CoA, 0.3 mM NADPH and 8 µg enzyme. The substrate specificity of ECR towards differenct enoyl-CoAs (crotonoyl-CoA, decenoyl-CoA or hexadecenoyl-CoA) was determined. The assay was performed in 100 mM HEPES/NaOH (pH 7.5, 70°C) with 0.3 mM of the relevant enoyl-CoA, 0.2 mM NADPH and 10 ug protein.

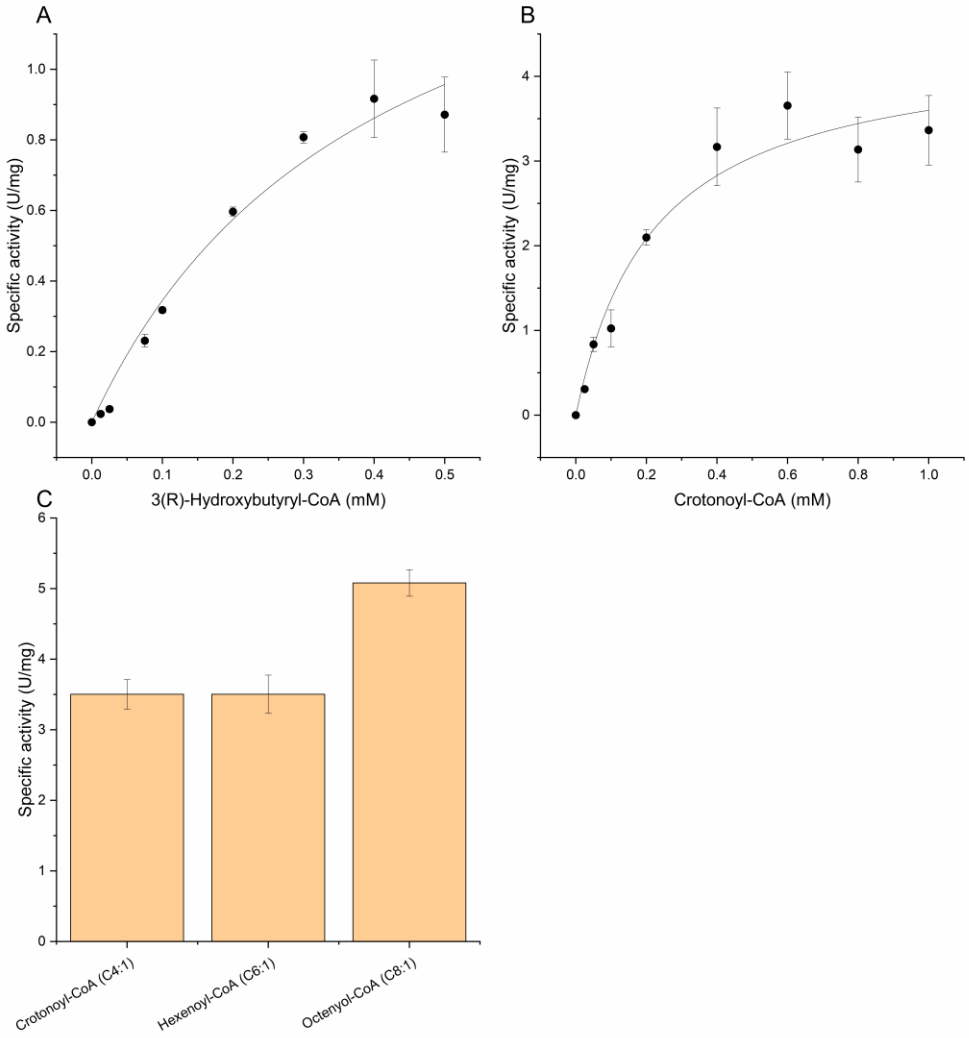


Figure S 20. **Investigation of kinetic properties (A, B) and substrate spectrum (C) of the recombinant MaoC-HCD Saci_1085 from *S. acidocaldarius*.** The activity of MaoC-HCD Saci_1085 was tested at 65°C, pH 6.5 via a discontinuous assay (50 µl) containinig 50 mM MES, 20 mM KCl and 0.0675 µg/µl protein. K_m for 3(R)-hydroxybutyryl-CoA was determined by varying its concentration from 0-0.5 mM (A) while a variable concentration of 0-1 mM was used for measuring K_m for crotonoyl-CoA (B). Activities toward substrates with different chain lengths (C4, C6 or C8) (C) were determined by incubating 0.09 µg/µl protein with 0.4 mM enoyl-CoA namely crotonoyl-CoA, hexenoyl-CoA or octenoyl-CoA. Afterwards, the reaction was stopped by mixing the sample with acetonitrile in a ratio of 1:3 (v/v) at different time points and then freezing the mixture. The formation of the relevant product was analysed via HPLC and thus the specific activities were calculated. The independent measurements were performed in triplicate and error bars indicate the standard error of the mean (SEM).

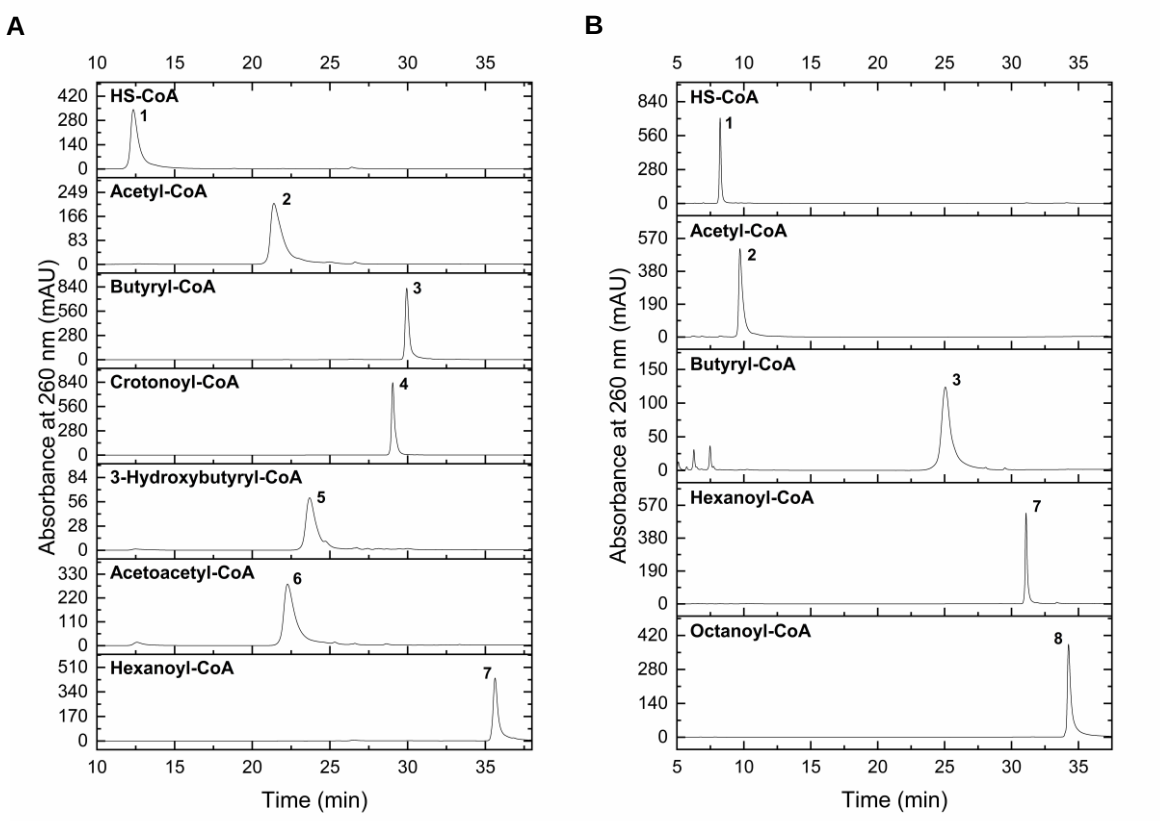


Figure S 21. **HPLC chromatogram of CoA ester standards involved in FA metabolism.** Two distinct programs with different acetonitrile (ACN) concentration gradients were applied for analyzing different chain lengths of CoA esters. The program “4-30% ACN program” (A) was used for shorter chain acyl-CoAs whereas “1-60% ACN program” (B) for longer chain CoA esters. The retention times representing the relevant CoA compounds are indicated in Table S4. The peak numbers correspond to the CoA compounds as following: 1. HS-CoA; 2. Acetyl-CoA; 3. Butyryl-CoA; 4. Crotonoyl-CoA; 5. 3-Hydroxybutyryl-CoA; 6. Acetoacetyl-CoA; 7. Hexanoyl-CoA; 8. Octanoyl-CoA.

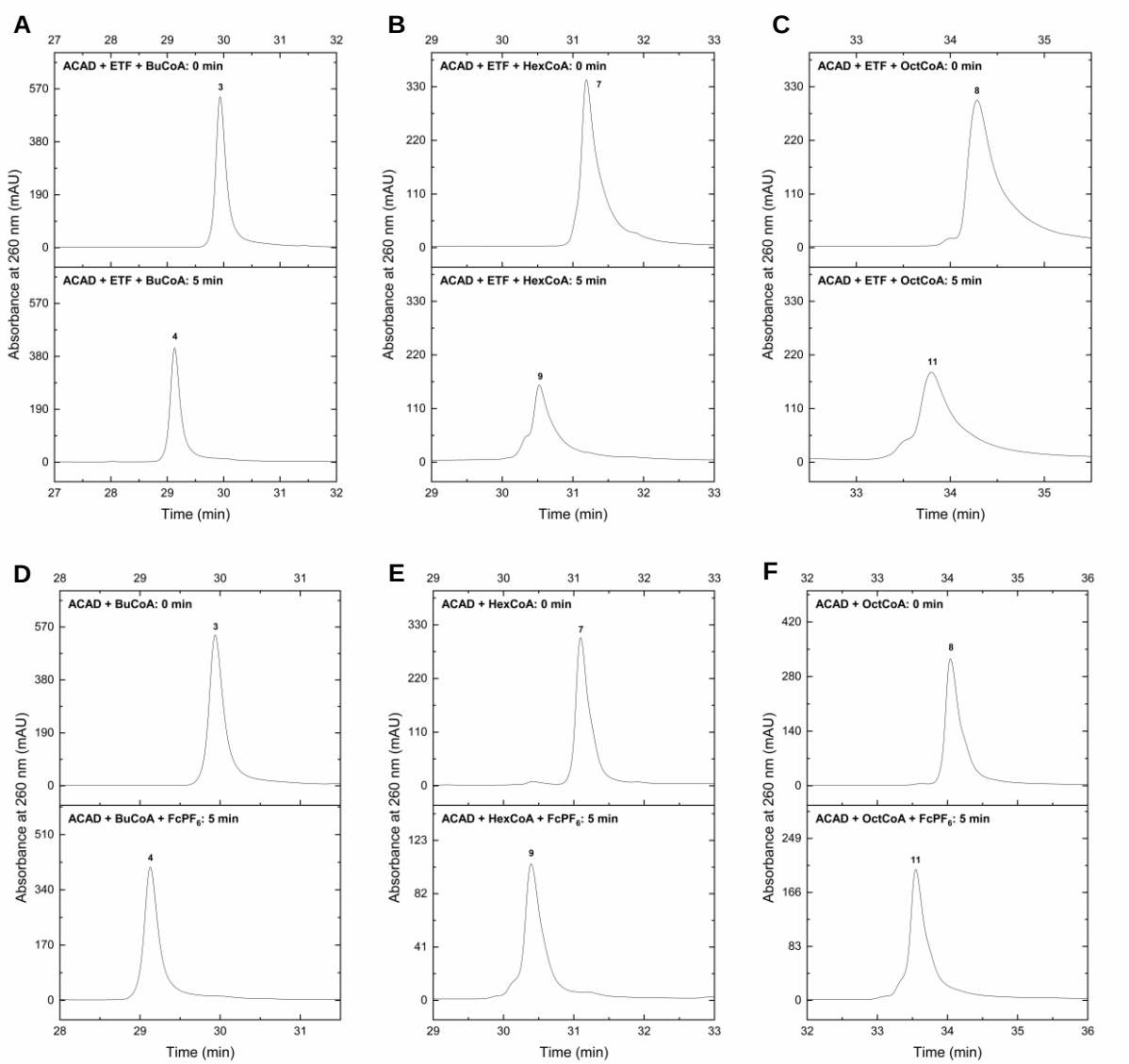


Figure S 22. **HPLC chromatograph of acyl-CoA oxidation to enoyl-CoA.** The oxidation of saturated acyl-CoAs to enoyl-CoA derivatives was analyzed by discontinuous assays at 65°C, pH 6.5. The 0.02 µg/µl ACAD and 0.01 µg/µl ETF were incubated in 50 mM MES/KOH (50 µl) with 20 mM KCl, 0.4 mM DCPIP and 0.4 mM of acyl-CoAs (butyryl-CoA, hexanoyl-CoA or octanoyl-CoA) for 5 min (A, B & C). Moreover, 0.8 mM $FcPF_6$ was used as the electron acceptor instead of ETF and DCPIP (D, E & F). Afterwards, the samples were analyzed via different HPLC programs i.e. butyryl-CoA conversion via 4-30% ACN program (A, D), oxidation of hexanoyl-CoA (B, E) or octanoyl-CoA (C,F) by 1-60% ACN program. As a result, all the tested acyl-CoAs (peak 3, 7 or 8) could be fully converted into the corresponding enoyl-CoA products (peak 4, 9 or 11).

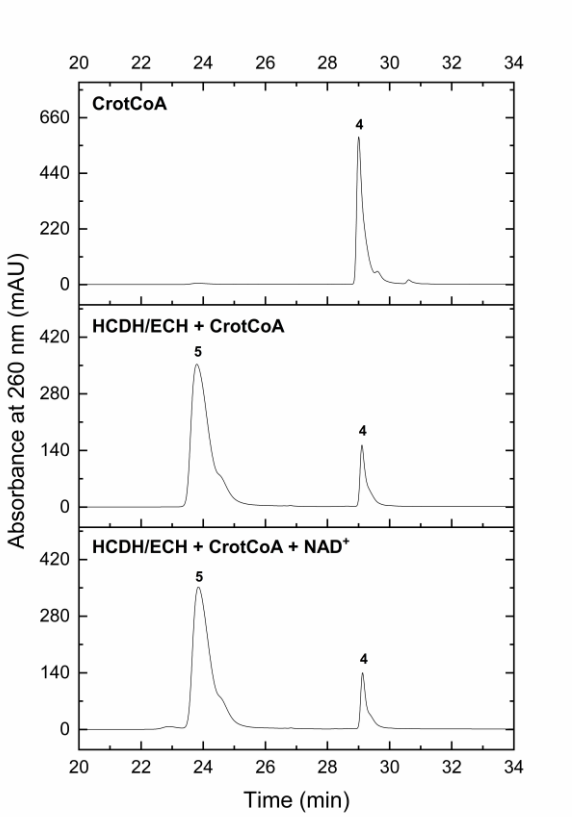


Figure S 23. **HPLC analysis of crotonoyl-CoA conversion into acetoacetyl-CoA by HCDH/ECH from *S. acidocaldarius*.** The discontinuous assay was carried out in 50 mM HEPES/NaOH with 20 mM KCl, 0.4 mM crotonoyl-CoA, 0.0144 µg/µl Saci_1109 in absence or presence of 2 mM NAD^+. The reaction mixture was incubated at 65°C, pH 6.5 for 15 min and the samples were then analyzed via the 1-60% ACN program. The formation of 3-hydroxybutyryl-CoA (peak 4) from crotonoyl-CoA (peak 5) could be shown. However, further production of acetoacetyl-CoA was not observed in presence of NAD^+.

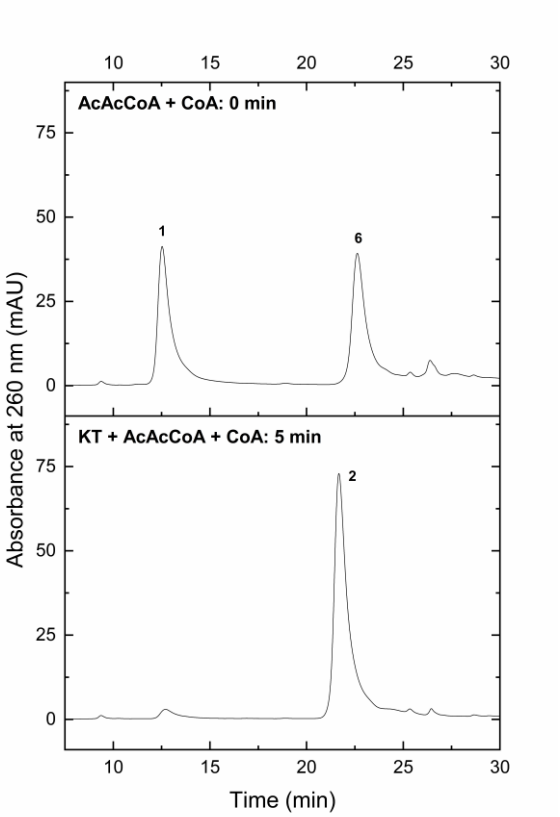


Figure S 24. **HPLC analysis of cleavage of acetoacetyl-CoA by KT.** To analyze the last thiolytic step, reaction components (400 µl) including 0.2 mM CoA and 0.1 mM AcAcCoA were incubated in 50 mM MES/KOH (pH 6.5) at 23°C for 2 min. The reaction was then initiated by addition of 2.7 µg of KT followed by incubation for 5 min. All the reaction samples were investigated adopting the 4-30% ACN HPLC program. As shown, all the employed AcAcCoA (peak 6) and CoA (peak 1) were converted to acetyl-CoA (peak 2).

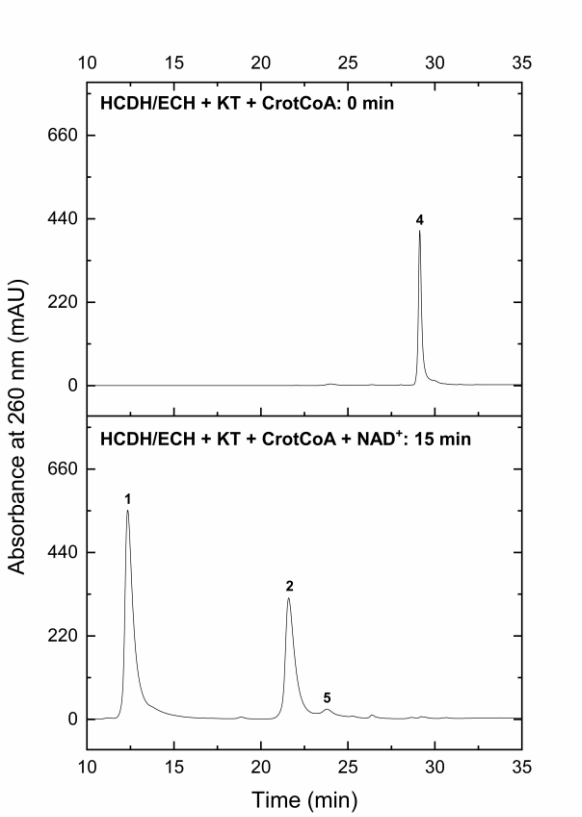


Figure S 25. **HPLC investigation of the last three steps by the β oxidation enzymes HCDH/ECH and KT.** The assay was performed in 50 mM MES/NaOH at 65°C, pH 6.5 by adding 20 mM KCl, 0.4 mM crotonoyl-CoA, 2 mM NAD^+, 0.0144 µg/µl HCDH/ECH, 1.6 mM CoA and 0.054 µg/µl KT. The assay mixture was incubated for 15 min and afterwards applied by the 4-30% ACN HPLC program. After reaction, initial substrate crotonoyl-CoA (peak 4) was fully consumed by HCDH/ECH, end product acetyl-CoA (peak 2) by KT was formed in presence of CoA (peak 1). Meantime, a limited amount of the intermediate 3-hydroxybutyryl-CoA (peak 5) was detected.

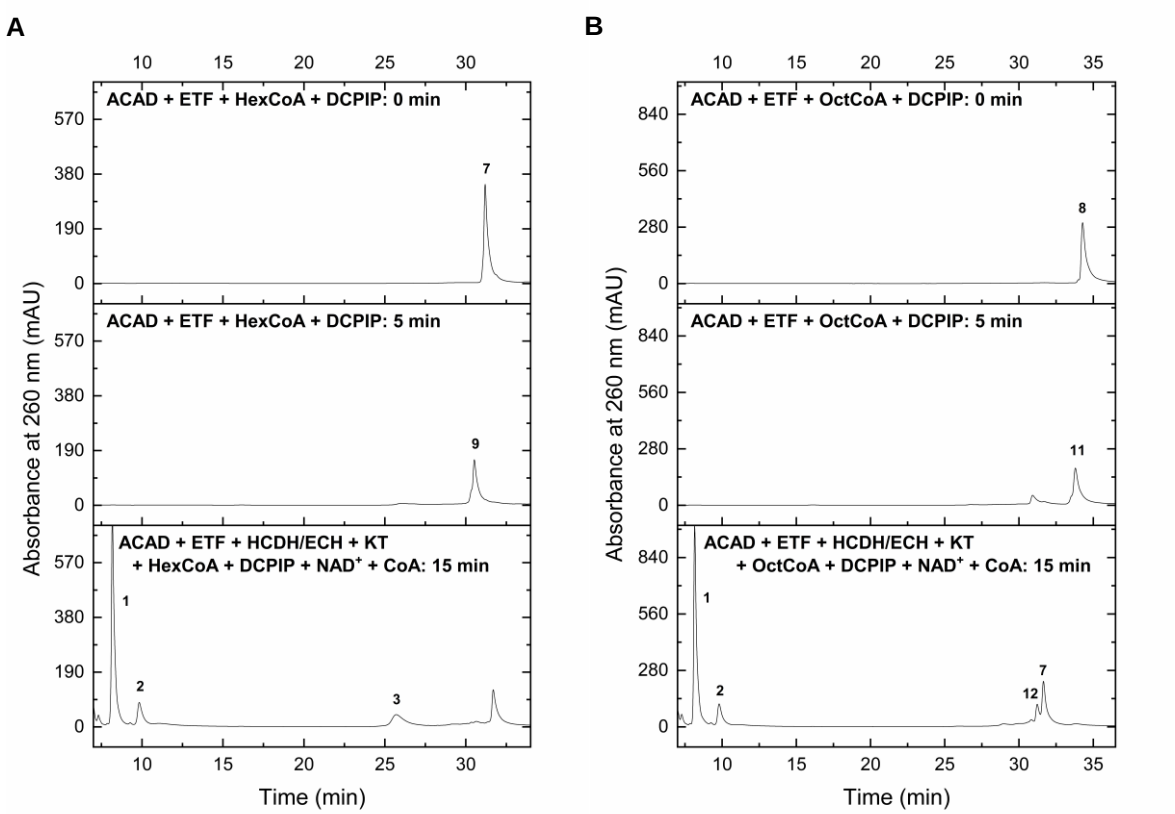


Figure S 26. **β oxidation cascades for degrading hexanoyl-CoA (A) or octanoyl-CoA (B).** The enzyme assays were carried out in two steps. The first oxidation step by 0.02 µg/µl ACAD (Saci_1123) and 0.01 µg/µl ETF (Saci_0315) was done in 50 mM MES/KOH (pH 6.5) at 65°C with 20 mM KCl, 0.4 mM DCPIP and 0.4 mM of acyl-CoA. The reaction was run for 5 min. In the second step, 2 m NAD^+, 0.0144 µg/µl HCDH/ECH (Saci_1109), 1.6 mM CoA and 0.054 µg/µl KT (Saci_1114) were successively introduced to the mixture and incubated for 15 min. At last, the β oxidation metabolites were detected by the 1-60% ACN program. After the first step, acyl-CoAs (peak 7 or 8) were completely oxidized to the relevant hexenoyl-CoA (peak 9) and octenoyl-CoA (peak 11) by ACAD and ETF in presence of an artificial electron acceptor DCPIP. Finally, hexenoyl-CoA could be fully degraded into acetyl-CoA (peak 2) and butyryl-CoA (peak 3) by HCDH/ECH and KT with presence of the essential cofactors CoA (peak 1) and NAD^+ (A) whereas the octenoyl-CoA was completely oxidized to acetyl-CoA and hexanoyl-CoA, meantime a tiny peak 12 representing the intermediate 3-hydroxyoctanoyl-CoA was found (B).

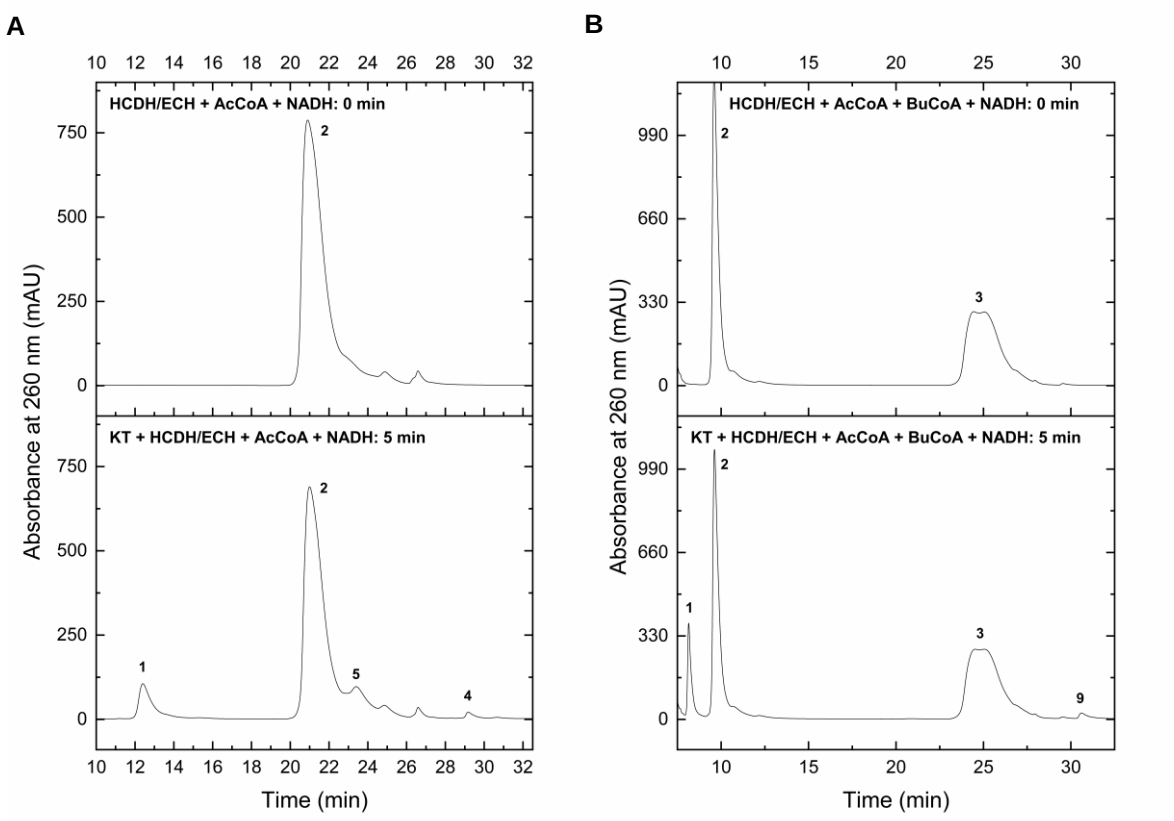


Figure S 27. **HPLC analysis of elongation of acetyl-CoA (A) or butyryl-CoA (B) by KT and HCDH/ECH.** The reaction mixture (400 µl) included 50 mM HEPES/NaOH (65°C, pH 6.5), 2 mM acetyl-CoA, 0.3 mM NADH and 17.1 µg HCDH/ECH (Saci_1109). After incubation for 2 min, the reaction was started by adding 16.2 µg KT (Saci_1114) and the mixture was incubated for another 5 min. The samples were then analyzed by distinct HPLC programs (A: 4-30% ACN; B: 1-60% ACN). Controls without KT were carried out. In both measurements, release of free CoA (peak 1) by KT was found. During acetyl-CoA elongation (A), limited amount of crotonoyl-CoA (peak 4) and 3-hydroxybutyryl-CoA (peak 5) was formed whereas less hexenoyl-CoA (peak 9) was synthesized from butyryl-CoA (peak 3) and acetyl-CoA (peak 2) by KT and HCDH/ECH (B).

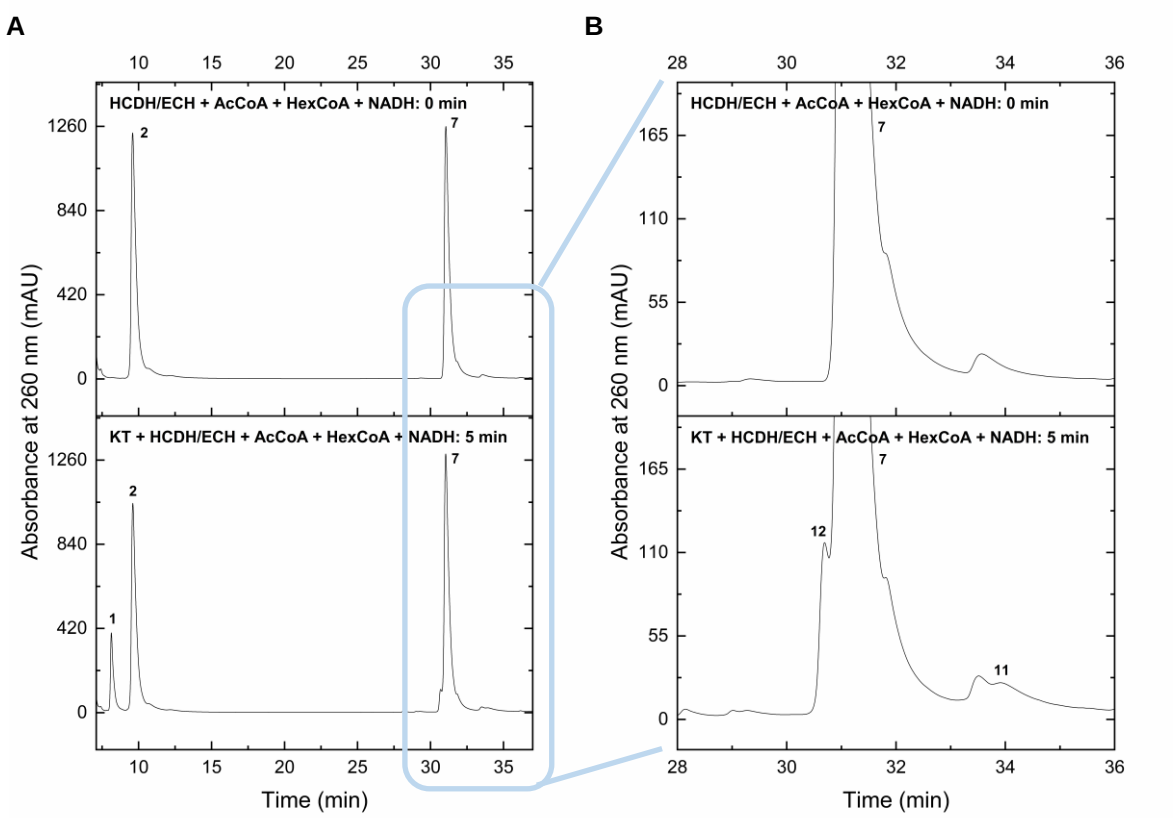


Figure S 28. **HPLC detection of condensation of hexanoyl-CoA and acetyl-CoA by KT and HCDH/ECH.** The reaction mixture (400 µl) included 50 mM HEPES/NaOH (65°C, pH 6.5), 1 mM acetyl-CoA, 1 mM hexanoyl-CoA, 0.3 mM NADH and 17.1 µg HCDH/ECH (Saci_1109). After incubation for 2 min, the reaction was started by adding 16.2 µg KT (Saci_1114). The reaction mixture was incubated for another 5 min and then analyzed by the 1-60% ACN HPLC program. Control without KT was included. The blue box of Fig. A was enlarged (B) in order to better visualize the products. Formation of CoA (peak 1) could be clearly shown. Two tiny peaks 11 and 12, which represent the respective octenoyl-CoA and 3-hydroxyoctanoyl-CoA were also observed (B) indicating the synthesis of C8-chain CoA from hexanoyl-CoA (peak 7) and acetyl-CoA (peak 2) by KT and HCDH/ECH.

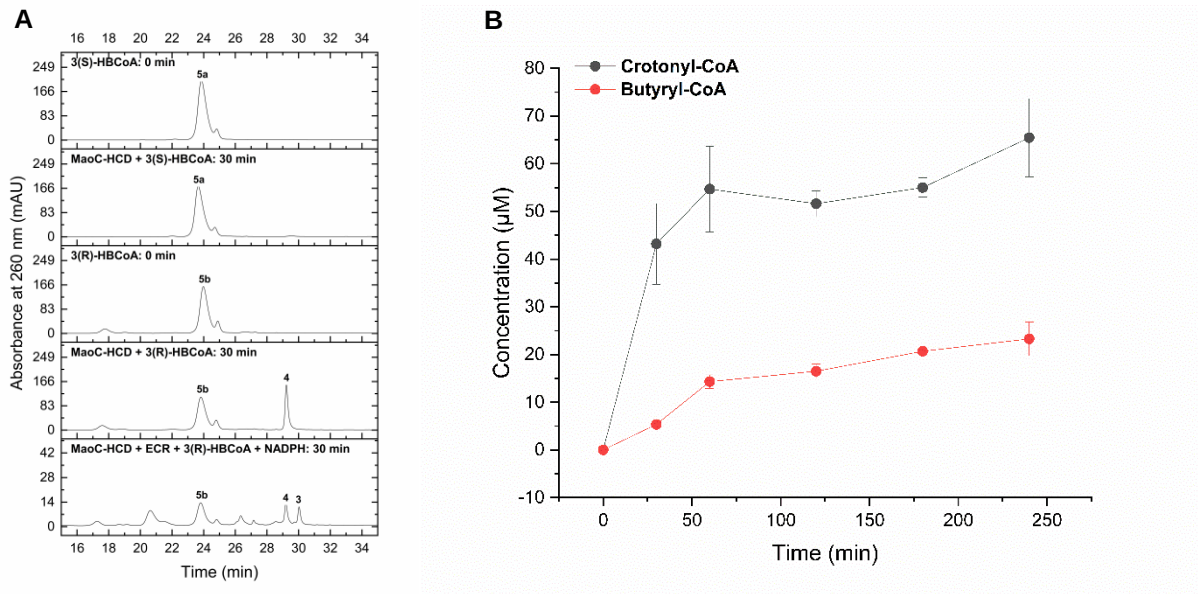


Figure S 29. **HPLC investigation of the stereospecificity of the recombinant MaoC-HCD Saci_1085 (A) and the conversion of 3(R)-hydroxybutyryl-CoA to butyryl-CoA by MaoC-HCD and ECR Saci_1115 (B).** To study the stereochemical specificity towards the 3-hydroxyacyl-CoA intermediates, 0.4 mM of individual 3(S)- or 3(R)-hydroxybutyryl-CoA was incubated with 0.0675 µg/µl Saci_1085 in 50 mM MES plus 20 mM KCl at 65°C, pH 6.5 for 30 min. The synthesis of butyryl-CoA from 3(R)-hydroxybutyryl-CoA was accomplished through two steps: dehydration of 3(R)-HBCoA via MaoC-HCD (Saci_1085) followed by reduction of crotonoyl-CoA via ECR (Saci_1115). The assay was done at 70°C, pH 6.5 in 50 mM MES, 20 mM KCl, 2 mM NADPH, 0.5 µg ECR and 0.5 µg MaoC-HCD, and was incubated for 30 min. As shown in Fig. A, 3(S)-HBCoA (peak 5a) was not converted while 3(R)-HBCoA (peak 5b) was converted to crotonoyl-CoA (peak 4) by MaoC-HCD and then further reduced to butyryl-CoA (peak 3) by ECR. In addition, a time-dependent conversion was monitored by analyzing samples at different time points: 0, 30, 60, 120, 180 and 240 min (B). Consequently, around 70 µM crotonoyl-CoA (black line) and 20 µM butyryl-CoA (red line) were produced over the time. All the assay samples were analyzed using 4-30% ACN HPLC program.

References

1. Wagner, M., et al., *Versatile Genetic Tool Box for the Crenarchaeote Sulfolobus acidocaldarius.* Front Microbiol, 2012. **3**: p. 214.
2. Lombard, J., P. Lopez-Garcia, and D. Moreira, *The early evolution of lipid membranes and the three domains of life.* Nat Rev Microbiol, 2012. **10**(7): p. 507-15.
3. Zweerink, S., et al., *Activity-based protein profiling as a robust method for enzyme identification and screening in extremophilic Archaea.* Nat Commun, 2017. **8**: p. 15352.
4. Wang, K., et al., *A TetR-family transcription factor regulates fatty acid metabolism in the archaeal model organism Sulfolobus acidocaldarius.* Nat Commun, 2019. **10**(1): p. 1542.
5. Fujita, Y., H. Matsuoka, and K. Hirooka, *Regulation of fatty acid metabolism in bacteria.* Mol Microbiol, 2007. **66**(4): p. 829-39.
6. Ruprecht, A., et al., *Characterization of novel acyl coenzyme A dehydrogenases involved in bacterial steroid degradation.* J Bacteriol, 2015. **197**(8): p. 1360-7.
7. Wipperman, M.F., et al., *Shrinking the FadE proteome of Mycobacterium tuberculosis: insights into cholesterol metabolism through identification of an alpha2beta2 heterotetrameric acyl coenzyme A dehydrogenase family.* J Bacteriol, 2013. **195**(19): p. 4331-41.
8. Kim, J.J. and R. Miura, *Acyl-CoA dehydrogenases and acyl-CoA oxidases. Structural basis for mechanistic similarities and differences.* Eur J Biochem, 2004. **271**(3): p. 483-93.

Chapter 3.3

Response of the thermoacidophilic Archaeon *Sulfolobus acidocaldarius* to solvent stress exemplified by 1-butanol exposure

AEM Accepted Manuscript Posted Online 19 March 2021
Appl. Environ. Microbiol. doi:10.1128/AEM.02988-20

Response of the thermoacidophilic Archaeon *Sulfolobus acidocaldarius* to solvent stress exemplified by 1-butanol exposure

Jens C. Benninghoff[a]*, Laura Kuschmierz[a]*, Xiaoxiao Zhou[a], Andreas Albersmeier[b], Trong Khoa Pham[c], Tobias Busche[b], Phillip C. Wright[c+], Jörn Kalinowski[b], Kira S. Makarova[d], Christopher Bräsen[a], Hans-Curt Flemming[e,f,g], Jost Wingender[e]#, Bettina Siebers[a]#

[a]Molecular Enzyme Technology and Biochemistry (MEB), Environmental Microbiology and Biotechnology (EMB), Centre for Water and Environmental Research (CWE), University of Duisburg-Essen, Universitätsstrasse 5, 45141 Essen, Germany
[b]Microbial Genomic and Biotechnology, Center for Biotechnology (CeBiTec), Bielefeld University, Universitätsstrasse 27, 33615 Bielefeld, Germany
[c]Department of Chemical and Biological Engineering, the University of Sheffield, Mappin Street, Sheffield, United Kingdom
[d]National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, Maryland 20894, USA
[e]Aquatic Microbiology, Environmental Microbiology and Biotechnology (EMB), Centre for Water and Environmental Research (CWE), University of Duisburg-Essen, Universitätsstrasse 5, 45141 Essen, Germany
[f]Singapore Center for Environmental Life Science Engineering (SCELSE), Nanyang Technological University, 60 Nanyang Drive, Singapore 637551, Singapore
[g]Water Academy, Schloss-Strasse 40, 88045 Friedrichshafen, Germany

[+]present address: Faculty of Science, Agriculture & Engineering, School of Engineering, Newcastle University, Newcastle, United Kingdom

Running Head: Response of *S. acidocaldarius* to solvent stress

*Both authors contributed equally to this work.
#Corresponding authors:
Bettina Siebers, bettina.siebers@uni-due.de
Jost Wingender, jost.wingender@uni-due.de

Keywords: Archaea, Extremophiles, *Sulfolobus acidocaldarius*, Stress Response, Organic Solvent, 1-Butanol, Biofilm, Extracellular Polymeric Substances

Abbreviations:

ABE: acetone butanol ethanol; AbfR: archaeal biofilm regulator; arCOGs: archaeal clusters of orthologous genes; BF: biofilm; Cdv: cell division protein; Cas: CRISPR-associated; CER: cation-exchange resin; CLSM: confocal laser scanning microscopy; ConA: concanavalin A; CRISPR: clustered, regularly interspaced, short, palindromic repeats; EPS: extracellular polymeric substances; ESCRT: endosomal sorting complexes required for transport; IB4: isolectin from *Griffonia simplicifolia*; iTRAQ: isobaric tags for relative and absolute quantitation; OD: optical density; PL: planktonic cells; SEM: scanning electron microscopy; TA: toxin-antitoxin; TEM: total extracellular material.

Abstract

Sulfolobus acidocaldarius is a thermoacidophilic crenarchaeon with optimal growth at 80 °C and pH 2 - 3. Due to its unique physiological properties allowing life at environmental extremes and recent availability of genetic tools, this extremophile receives increasing interest for biotechnological application. In order to elucidate the potential of tolerating process-related stress conditions, we investigated the response of *S. acidocaldarius* towards the industrially relevant organic solvent 1-butanol.

In response to butanol exposure, biofilm formation of *S. acidocaldarius* was enhanced and occurred up to 1.5% (v/v) 1-butanol, while planktonic growth was only observed up to 1% (v/v) 1-butanol. Confocal laser scanning microscopy revealed that biofilm architecture changed with the formation of denser and higher tower-like structures. Concomitantly, changes in the extracellular polymeric substances with enhanced carbohydrate and protein content were determined in 1-butanol-exposed biofilms. Using scanning electron microscopy three different cell morphotypes were observed in response to 1-butanol.

Transcriptome and proteome analyses were performed comparing the response of planktonic and biofilm cells in absence and presence of 1-butanol. In response to 1% (v/v) 1-butanol transcript levels of genes encoding motility and cell envelope structures as well as membrane proteins were reduced. Cell division and/or vesicle formation was upregulated. Furthermore, changes of immune and defence systems, as well as metabolism and general stress response were observed. Our findings show that the extreme lifestyle of *S. acidocaldarius* coincided with a high tolerance to organic solvents. This study provides first insights into biofilm formation and membrane/cell stress caused by organic solvents in *S. acidocaldarius*.

74 **Importance**

75 Archaea are unique in terms of metabolic and cellular processes as well as the
76 adaptation to extreme environments. In the past few years, the development of
77 genetic systems and biochemical, genetic and poly-omics studies have provided
78 deep insights into the physiology of some archaeal model organisms. In this study,
79 we used *S. acidocaldarius* adapted to two extremes, low pH and high temperature, to
80 study its tolerance and robustness as well as its global cellular response towards
81 organic solvents exemplified by 1-butanol. We were able to identify biofilm formation
82 as primary cellular response to 1-butanol. Furthermore, the triggered cell/membrane
83 stress led to significant changes in culture heterogeneity accompanied by changes in
84 central cellular processes such as cell division and cellular defense systems, thus
85 suggesting a global response for the protection at population level.

Introduction

Archaea are widely distributed in natural environments (1). Most cultivated Archaea are extremophiles that thrive at environmental extremes such as high temperatures, pH values, high salt concentrations, or combinations thereof (2). In particular, thermophiles and hyperthermophiles with growth optima above 60 °C and 80 °C, respectively, are of interest for biotechnological applications in high-temperature industrial processes (3, 4). They are able to produce enzymes (extremozymes/thermozymes) that are functional under extreme conditions because of enhanced enzyme rigidity and stability, and have been shown to be active in organic solvents and ionic liquids (5). In addition, Archaea possess a unique membrane lipid composition. In contrast to Bacteria and Eukarya, they use isoprenoid hydrocarbon side chains linked to *sn*-glycerol-1-phosphate via ether linkage forming monopolar diether lipids (archaeol) or membrane-spanning bipolar tetraether lipids (caldarchaeol) (6). These archaeal membranes are more stable against stressors (7).

One promising platform organism for biotechnology is the thermoacidophilic crenarchaeon *Sulfolobus acidocaldarius* (3, 4, 8, 9). *S. acidocaldarius* is an obligately aerobic organism growing optimally under two extreme conditions, including low pH values of 2.0-3.5 and high temperatures of 75 °C – 80 °C. It is genetically tractable (10), enabling metabolic engineering for potential application in industrial processes (4). *S. acidocaldarius* is able to form biofilms (11, 12), defined as microbial aggregates, embedded in a matrix of extracellular polymeric substances (EPS) on surfaces and other interfaces (13). Proteins, carbohydrates and DNA have been identified as constituents of the EPS matrix of *S. acidocaldarius* (14). The biofilm mode of life is dominant among prokaryotic microorganisms (15) and offers advantages for survival compared to the planktonic lifestyle, for example an

112 enhanced tolerance against adverse environmental conditions (13) as they may be
113 encountered in biotechnological processes due to toxic reactants or products.

114 1-butanol is a key commodity widely used as a solvent or chemical feedstock. So far,
115 1-butanol is mainly produced chemically by the Oxo process (16). Human
116 dependence on petroleum-derived fuels, the corresponding depletion of fossil
117 resources and emission of greenhouse gases, particularly CO_2, promoted the search
118 for more environmentally friendly alternatives. In this context, biobutanol represents a
119 promising alternative as a fuel additive and biofuel for direct replacement of gasoline
120 (17, 18). Production of biobutanol from renewable resources is predominantly
121 accomplished by *Clostridium* strains via acetone butanol ethanol (ABE) fermentation
122 (16). However, while ABE fermentation supplied approx. 66% of the world supply of
123 1-butanol until the 1950's, bio-based butanol production was outcompeted by
124 petroleum-based processes afterwards (16).

125 A problem in the production of biobutanol is its toxicity towards microbial cells. For a
126 vast majority of microorganisms a growth limit at 1% - 2% (v/v) of 1-butanol in
127 nutrient medium has been observed in liquid cultures (19, 20). There is the widely
128 accepted notion that 1-butanol toxicity results from its chaotropic effects on the
129 cytoplasmic cell membrane, leading to the disruption of nutrient and ion transport and
130 the loss of the membrane potential (21, 22). Bacteria and eukaryotic microorganisms
131 are able to adapt to the presence of aliphatic, toxic alcohols, including acetone,
132 ethanol, butanol, isobutanol and propanol, with the development of an enhanced
133 tolerance, allowing survival and growth at elevated concentrations of these
134 compounds (20, 23, 24). The adaptation strategies are versatile (21, 22, 25).
135 Microorganisms can respond to alcohol exposure by changing their membrane lipid
136 composition to sustain membrane fluidity, called "homeoviscous adaptation" (26).

137 This process may include a shift in the ratio of unsaturated to saturated lipids,
138 branched and unbranched lipids, a change of their isomerization and cyclization state
139 or headgroup composition (20, 22, 27). In gram-negative bacteria outer membrane
140 modifications (e.g. alterations in the lipopolysaccharide content and porin expression,
141 increased interactions with divalent metal ions for membrane stabilization and outer
142 membrane vesicle formation) were reported (27). Other cellular responses described
143 for bacteria include the upregulation of energy-dependent efflux systems to lower the
144 intracellular solvent concentration, and the metabolic degradation of solvents (28).
145 Organic solvents, including 1-butanol, were shown to enhance expression of heat
146 shock proteins, including molecular chaperones that assist correct protein folding and
147 transport as well as the recycling of defective proteins (29). Cell aggregation and
148 biofilm formation was shown to enhance tolerance to 1-butanol as observed for the
149 butanol production strain *Clostridium acetobutylicum* (30). Further, changes in the
150 composition of EPS were reported for biofilms of *C. acetobutylicum* and
151 *Pseudomonas taiwanensis* (31, 32).

152 Regarding organic solvent tolerance, archaeal extremophiles may offer advantages
153 over mesophilic organisms due to their intrinsic robustness and adaptation to hostile
154 environments (2-4, 9). This study aims to fill this gap of knowledge by investigating
155 the natural ability of the thermoacidophilic crenarchaeon *S. acidocaldarius* as an
156 archaeal model organism to tolerate 1-butanol and its cellular response towards this
157 industrially relevant organic solvent.

Results

Effect of 1-butanol on cell growth in liquid cultures

Initially, the effect of 1-butanol on *S. acidocaldarius* was investigated using liquid cultures. Growth (OD_{600nm}), D-glucose consumption and 1-butanol concentration were determined throughout the incubation period of three weeks (Fig. 1). The growth of *S. acidocaldarius* at 76 °C in the absence and presence of 0.5% (v/v) (4.05 g/L; 55 mM) 1-butanol was similar, reaching maximum OD_{600nm} values after 84 h of cultivation (Fig. 1A). On further incubation, cultures without 1-butanol showed a significant decrease of OD_{600nm} values, while OD_{600nm} values of cultures with 0.5% (v/v) 1-butanol remained unchanged for a prolonged time period of approximately 312 h (Fig. 1A). Cells exposed to 1% (v/v) (8.10 g/L; 109 mM) 1-butanol showed biphasic growth and reached the stationary phase with a considerable delay (residual growth rate of 52%, 0-72 h) compared to cells grown without or with 0.5% (v/v) 1-butanol. Concomitantly, *S. acidocaldarius* showed a substantial delay in D-glucose utilization when exposed to 1% (v/v) 1-butanol (Fig. 1B). Neither growth nor glucose degradation were observed in planktonic *S. acidocaldarius* cultures supplemented with 1.5% (v/v) (12.15 g/L; 164 mM) 1-butanol (Fig. 1A and B). The concentration of 1-butanol decreased at a similar rate in all cultures including a cell-free abiotic control with 1% (v/v) 1-butanol (Fig. 1C).

Culturability of *S. acidocaldarius* was examined using spot plates (Fig. S1). Samples of cell cultures were taken at different time points, diluted and spotted on Brock-Gelrite plates, followed by incubation of the plates at 76 °C for four days. Cells from cultures without 1-butanol as well as with 0.5 and 1% (v/v) 1-butanol were able to grow on the spot plates, while complete inhibition of culturability was observed for cells from cultures with 1.5% (v/v) 1-butanol (Fig. S1).

183 "Collars" of a slimy material developed on the inner glass surfaces of the flasks at the
184 air-liquid interface of liquid cultures exposed to 0.5 and 1% (v/v) 1-butanol within one
185 week of incubation and remained visible throughout the 3-week incubation period
186 (Fig. S2). Crystal violet staining improved the visibility of the slimy material attached
187 to the glass surface (Fig. S2C). Microscopic examination of the slime revealed high
188 numbers of cells densely packed and embedded in the slime matrix, indicating the
189 formation of biofilms upon exposure to sub-inhibitory concentrations of 1-butanol (Fig.
190 S2B). Debris of biofilms were observed in *S. acidocaldarius* cultures grown without 1-
191 butanol, while no biofilm was formed by cells exposed to 1.5% (v/v) 1-butanol (Fig.
192 S2C). Inspection of culture liquid (planktonic cells) by means of phase-contrast
193 microscopy revealed both single cells and cell aggregates in the log growth phase
194 (OD_{600nm} values between 0.5 and 1.5) independent of the absence and presence of
195 0.5 and 1% (v/v) 1-butanol (Fig. S3). At 1.5% (v/v) 1-butanol without any growth, cell
196 aggregates were not observed in the culture medium.
197 To investigate the specificity of the cell response to 1-butanol, *S. acidocaldarius* was
198 grown in the presence of other short-chain alcohols, namely ethanol, 1-propanol, and
199 isobutanol (liquid cultures). The formation of biofilms at the solid-air-liquid interphase
200 was observed for ethanol at 1-4% (v/v), 1-propanol at 0.5-2.5% (v/v) and isobutanol
201 at 0.5% and 1% (v/v) of the corresponding alcohol added to the culture media (Fig.
202 S4A). Thus, the increase of biofilm amounts at the solid-air-liquid interphase seems
203 to be a general response of *S. acidocaldarius* to short-chain alcohols.
204
205 **Effect of 1-butanol on biofilm formation and architecture**
206 Biofilm formation was quantified by a commonly used microtiter plate biofilm assay.
207 *S. acidocaldarius* was cultivated in 96-well polystyrene microtiter plates with different
208 1-butanol concentrations (0%, 0.5%, 1%, 1.5%, 2%, and 2.5% (v/v) 1-butanol) under

209 static conditions at 76 °C for four days. Growth was determined by turbidity
210 measurements (OD_{600nm}) (Fig. 2A), and biofilm biomasses were quantified based on
211 crystal violet staining of surface-attached biomass (Fig. 2B). Respiratory activity of
212 biofilms was determined using an adapted cell viability assay based on the reduction
213 of resazurin (Fig. 2C). The turbidity values (OD_{600nm}) of suspended biofilm cells
214 remained nearly constant up to 1.5% (v/v) 1-butanol (Fig. 2A). At 2% (16.2 g/L;
215 218 mM) and 2.5% (v/v) (20.25 g/L; 273 mM) 1-butanol the turbidity values of biofilm
216 cells decreased significantly. This coincided with a drop in biofilm biomass (Fig. 2B)
217 and respiratory activity of the biofilms (Fig. 2C). However, weak biofilm formation and
218 low respiratory activity were still observed, indicating the presence of metabolically
219 active cells at elevated 1-butanol concentrations of up to 2.5% (v/v).
220 Different microscopic techniques were used to visualize the distribution, architecture
221 and EPS of *S. acidocaldarius* biofilms directly on surfaces in the absence and
222 presence of 1-butanol. First, we analysed the distribution of crystal violet-stained
223 biofilm cells on glass slides in combination with light microscopy. After growth for four
224 days, a change in the pattern of cell distribution depending on the 1-butanol
225 concentration was observed (Fig. 3A). In the absence of 1-butanol the cells were
226 homogeneously distributed on the glass surface predominantly as a single layer of
227 cells, occasionally interspersed with cell aggregates (microcolonies; Fig. 3A). Under
228 the influence of 1% (v/v) 1-butanol a higher occurrence of microcolonies was
229 observed (Fig. 3A). Exposure to 1.5% (v/v) 1-butanol resulted in the formation of
230 more pronounced microcolonies with an irregular pattern of distribution (Fig. 3A).
231 Similar to the results for 1-butanol, the formation of *S. acidocaldarius* microcolonies
232 on glass substratum was also observed when the cells were exposed to certain
233 solvent-specific concentrations of ethanol, 1-propanol, and isobutanol (Fig. S4B).

234 In accordance with light microscopy, scanning electron microscopy (SEM) images
235 showed the formation of microcolonies of *S. acidocaldarius* that developed in the
236 presence of 1% and 1.5% (v/v) 1-butanol (Fig. 3B). Cell aggregates were partially
237 surrounded by extracellular material, probably comprising EPS (Fig. 3B). Thus, these
238 images confirmed that 1-butanol promoted the formation of cell aggregates with
239 concomitant production of extracellular components. The cell shape also changed
240 with increasing 1-butanol concentrations. In the absence of 1-butanol and in
241 presence of 1% (v/v) 1-butanol the cells showed two types of morphology: (i) lobe-
242 shaped with a smooth surface structure and a size of approximately 1 µm as
243 previously described by Brock et al. (1972) (8) and (ii) flat cells with a more irregular
244 structure (Fig. 3B). In the presence of 1.5% (v/v) 1-butanol, a third cell morphotype
245 was observed: the surface of *S. acidocaldarius* cells appeared more rounded with a
246 perforated surface structure (Fig. 3B).
247 Confocal laser scanning microscopy (CLSM) was used to visualize the three-
248 dimensional biofilm architecture and the occurrence of EPS. *S. acidocaldarius* was
249 incubated at 76 °C for four days in µ-dishes. Submersed biofilms were analysed by
250 SYTO 9-staining of cells and addition of fluorescently labelled lectins GS-IB4 and
251 ConA that had previously proved suitable to visualize carbohydrates as constituents
252 of *Sulfolobus* EPS (11). CLSM revealed that *S. acidocaldarius* biofilms mainly
253 consisted of tower-like structures (Fig. 4). While only a few carbohydrate-containing
254 structures were detected in *S. acidocaldarius* biofilms without 1-butanol, the presence
255 of 1% (v/v) 1-butanol led to the detection of more sugar residues bound by lectins
256 GS-IB4 and ConA (Fig. 4). Thus, the amount of carbohydrate-containing structures
257 on the surface of *S. acidocaldarius* biofilms increased in response to 1-butanol
258 exposure. Overall, the signals of ConA (Fig. 4, red) were more dominant than signals

of GS-IB4 binding (Fig. 4, blue). Some carbohydrate structures seemed to be targeted by both lectins.

Effect of 1-butanol on EPS composition

Since microscopic analyses of *S. acidocaldarius* indicated the presence and enhanced formation of extracellular material due to 1-butanol exposure, EPS isolation and quantification from biofilms grown in polystyrene Petri dishes at 76 °C for four days was performed (Fig. 5, Fig. S5). The turbidity of the isolated aqueous phase (OD_{600nm} of planktonic cells) decreased with increasing 1-butanol concentrations (Fig. S5A). The wet weight of biofilm mass harvested from the bottom of Petri dishes was the same in the absence and in the presence of 0.5 and 1% (v/v) 1-butanol (approximately 110 µg/cm^2), and the biofilm cell numbers were 6.2 to 8.3 x 10^5 cells/cm^2 (Fig. S5). Suspensions of harvested *S. acidocaldarius* biofilms were used to extract EPS with a cation-exchange resin. After the extraction procedure, the supernatant was sterile filtered to obtain cell-free total extracellular material (TEM). The final EPS fraction was obtained after removal of low-molecular-weight compounds by dialysis (cut-off of 3.5 kDa) (14, 33). The amount of carbohydrates and proteins per biofilm cell was determined in the biofilm suspensions, the TEM and the EPS fraction (Fig. 5A-C). In biofilm suspensions, the amounts of proteins as well as carbohydrates increased with an increasing concentration of 1-butanol in the growth medium (Fig. 5A). The increase in carbohydrate and protein content in response to 1% (v/v) 1-butanol was even more pronounced within the TEM and EPS fractions (Fig. 5B). The carbohydrate amount per biofilm cell increased 5-fold and the amount of proteins per cell was 19-fold higher. However, 0.5% (v/v) 1-butanol did not show a significant increase in carbohydrates and proteins compared to the control without 1-butanol. The increases in protein and carbohydrate amounts in the total

cell-free extracellular components (Fig. 5B) were due to the presence of mainly high molecular weight protein and carbohydrate compounds in the EPS (Fig. 5C). When exposed to 1% (v/v) 1-butanol the main component of the EPS were proteins (Fig. 5C). In contrast, carbohydrates were identified as the main component of biofilm cells without and with exposure to 0.5% (v/v) 1-butanol (Fig. 5C).

In summary, microscopic and biochemical analysis indicated that *S. acidocaldarius* responded to 1-butanol exposure by a significant change in EPS composition and biofilm architecture accompanied by an alteration in cell morphotypes. Overall, these results suggested that *S. acidocaldarius* might change its gene expression in response to 1-butanol concentrations above 0.5% (v/v). Therefore, transcriptome and initial proteome analyses were conducted to obtain insights into the cellular response towards 1-butanol.

Genome wide transcriptional response of *S. acidocaldarius* biofilms to 1-butanol exposure

S. acidocaldarius was grown in Petri dishes (static incubation) in the absence and presence of 0.5% and 1% (v/v) 1-butanol at 76 °C for four days. Planktonic and biofilm cells were harvested and the transcriptional response towards 1-butanol was analysed. Here, regulated genes with more than four-fold changes (log2 fold change ≥ 2) are discussed in detail (Tab. 1, Tab. S1, Excel sheet RNAseq_results_20201124.xlsx). In response to lifestyle, i.e. biofilm (BF) vs. planktonic (PL) cells, only the expression of 15 and 44 genes was significantly changed in the absence (BF0/PL0) and in presence of 1% (v/v) 1-butanol (BF1/PL1), respectively, indicating that after four days of static growth both biofilm and planktonic cells are quite similar in respect to their gene expression profile (Fig. S6). In agreement with our previous experimental observations on growth of *S.*

311 *acidocaldarius* with 0.5% (v/v) 1-butanol (Fig. 1), the transcriptional response in
312 planktonic (PL05/PL0) and biofilm (BF05/BF0) cells towards 0.5% (v/v) 1-butanol was
313 quite low (8 and 16 differentially expressed genes, respectively). Major transcriptional
314 changes were observed in planktonic and biofilm cells in presence of 1% (v/v) 1-
315 butanol with 122 (PL1/PL0) and 117 (BF1/BF0) differentially expressed genes,
316 respectively (Tab. 1, Fig S6). Notably, most genes are downregulated in biofilms (74
317 genes) and planktonic cells (89 genes). Among these differentially regulated genes,
318 42 genes (16 up- and 26 downregulated genes) display a common regulation in both
319 biofilms and planktonic cells (Fig. S6).
320 For most archaeal clusters of orthologous gene (arCOG) categories no major
321 changes in gene expression were observed for biofilm and planktonic cells in
322 response to 1-butanol (1%) exposure (Fig. 6; (34)). However, for the arCOG
323 categories S (function unknown, n=437), O (uncharacterized, n=9), and N (cell
324 motility, n=18) a strong down regulation is observed in both planktonic as well as
325 biofilm cells.
326 Due to the observed changes in cell morphology via SEM (Fig. 3B), we inspected the
327 distribution of predicted transmembrane helices among these differentially regulated
328 genes in more detail (Fig. S7), revealing that 21% and 25% of the upregulated and
329 42% and 52% of the downregulated genes in biofilm and planktonic cells,
330 respectively, possess at least one transmembrane helix. For most of these predicted
331 membrane (associated) proteins the function is unknown. Also, the most highly
332 downregulated genes in biofilm upon 1-butanol exposure (log2 fold change ≥ -4) are
333 membrane proteins of unknown functions (Tab. S2).
334 Based on the great overlap of the gene regulation pattern found for planktonic and
335 biofilm cells, a detailed comparative analysis of the transcriptome data was
336 performed focussing on biofilm cells grown without and with 1% (v/v) 1-butanol. In

addition to membrane proteins genes encoding cell surface structures were strongly affected by 1-butanol exposure (Tab. S1). In line with the 1-butanol enhanced biofilm formation most genes encoding the archaellum for motility (*fla*X-*fla*J gene cluster, *saci_1172-1178*) were significantly downregulated (Fig. 7), whereas UV-induced pili (*ups* genes) for genetic DNA exchange via conjugation, and adhesive pili (*aap* genes) for cell attachment were if at all slightly affected (35). Crenarchaeota including *Sulfolobus* spp. rely on the endosomal sorting complexes required for transport (ESCRT III) machinery for cell division, vesicle formation and budding (36-39). In response to 1-butanol exposure the genes encoding one of the three CdvB paralogs (*cdvB1* gene, *saci_0451*) is significantly upregulated (38, 39).

Also for several transcriptional regulators a differential gene expression was observed, however, only for *arnR1* (PL1/PL0) a more than 2-fold downregulation was observed (Tab. S1). Saci_1171 (ArnR1) is besides ArnR (Saci_1180) one of the positive transcriptional regulators for archaellum biosynthesis (40). For the gene encoding the archaeal biofilm regulator 1 (AbfR1, Saci_0446) a slight upregulation (log2 fold change 1.71) in response to 1-butanol was observed in biofilm cells (41, 42), whereas Saci_1223 (AbfR2) was slightly downregulated (log2 fold change -1.54; (43)).

Notably, we observed a significant regulation of the CRISPR-Cas (CRISPR: clustered, regularly interspaced, short, palindromic repeats; Cas: CRISPR-associated) system (44, 45). In the presence of 1-butanol a significant downregulation of the *Sulfolobus* specific Type III system genes (*saci_1893-1899*) is observed (Tab. S1).

Like the CRISPR-Cas system also several genes of the toxin-antitoxin (TA) system (46) were downregulated in *S. acidocaldarius* cells in response to 1-butanol exposure with *saci_1056* (antitoxin, CopJ/RHH family) and *saci_1124* (CopG/RHH family DNA

binding protein, only BF1/BF0) being the most prominent ones (Tab. S1). Additionally, the operon of the HEPN-MNT system (HEPN toxin and MNT (minimal nucleotide transferase, *saci_1928*) antitoxin) showed a slight upregulation.

Concerning the central metabolism the 5-oxoprolinase, involved in the degradation of pyroglutamate, was one of the most upregulated genes in response to 1-butanol exposure (Tab. S1; (47)). In addition, the gene cluster *saci_2293* (2-keto-4-pentenoate hydratase/2-oxohepta-3-ene-1,7-dioic acid hydratase), *saci_2294* (aromatic ring hydroxylase) and *saci_2295* (catechol 2,3-dioxygenase or other lactoylglutathione lyase family enzyme) is also significantly upregulated in both lifestyles upon 1-butanol exposure. The three genes encode for proteins involved in the catechol meta-cleavage pathway for degradation of aromatic compounds/amino acids (48).

Finally, a major regulation was observed at transcript levels of the branched aerobic respiratory chain (49). Whereas the cytochrome bc1 complex (Saci_1859-1862) and one of the three terminal oxidases, the DoxBCE complex (Saci_0097-0099), is significantly downregulated in biofilm cells in response to 1-butanol, the SoxABCDL complex (Saci_2086-2089) is not regulated at all and the SoxEFGHIM complex (Saci_2258-2263) is mainly downregulated in planktonic cells upon 1-butanol exposure. Genes encoding thioredoxin (Saci_1823) and peroxiredoxin (Saci_1125) were slightly upregulated in biofilms exposed to 1-butanol (log2 fold changes of 1.4 and 1.2). In correlation, the gene *saci_1169*, encoding a thioredoxin reductase TrxB, catalysing the reduction of thioredoxin, is 6.1-fold upregulated.

In addition to the transcriptional changes, also the global proteome response of biofilm cells in absence and presence of 1% (v/v) 1-butanol was analysed. Planktonic cells could not be considered for further analysis since the cell mass retrieved after growth in presence of 1% (v/v) 1-butanol was below the limit for analysis. The

presence of 1% (v/v) 1-butanol resulted in an upregulation of 93 proteins (11 significant) and downregulation of 114 proteins (23 significant) in biofilm cells (Tab. S3). The highest upregulation in response to 1-butanol was found for the ribosomal proteins Saci_0642 (RplA/L37e) and Saci_0583 (RpsN/S14). In correspondence to the transcriptome analysis, proteins of the respiratory chain, namely two proteins of the cytochrome bc1 complex (Saci_1860 and Saci_1862) and one protein subunit 1 of the terminal oxidase complex DoxBCE (Saci_0097) were significantly downregulated (Tab. S3). In addition, the S-layer protein SlaA was found to be downregulated in both omics analyses. However, in general, no significant overlap was observed between the transcriptome and proteome data. This discrepancy was also observed upon starvation in *S. acidocaldarius* (50) and it was suggested that response to stress conditions acts on multiple layers of gene expression and regulation.

Discussion

Here, we examined the effect of 1-butanol on the thermoacidophilic archaeon *S. acidocaldarius* to study its ability to tolerate 1-butanol as well as its response towards solvent stress.

Butanol toxicity

In liquid cultures, *S. acidocaldarius* was able to grow in the presence of 1% (v/v) 1-butanol without prior adaptation, while no growth was observed with 1.5% (v/v) 1-butanol. So far 1-butanol tolerance was only investigated in a few Archaea, including mesophiles (*Natronomonas pharaonis*, *Halorubrum lacusprofundi*, *Methanosarcina acetivorans*) and hyperthermophiles (*Methanocaldococcus jannaschii* and *Aeropyrum pernix* (20)). In presence of 0.25% (v/v) 1-butanol *N. pharaonis* and *M. acetivorans* showed no growth, while *H. lacusprofundi* and *M. jannaschii* were able to grow. *A. pernix* was reported to grow in the presence of 0.5% (v/v) 1-butanol (20). The 1-butanol tolerance of *S. acidocaldarius* is in a similar range as that observed for planktonic mesophilic bacterial and yeast species commonly used as model organisms in biotechnology (1-2% (v/v), such as *Escherichia coli, Bacillus subtilis*, *C. acetobutylicum* and *Saccharomyces cerevisiae* (19, 20, 51-53)).

Thus, planktonic *S. acidocaldarius* cells possess a 2- to 4-fold higher tolerance to 1-butanol than other archaeal organisms reported and a similar tolerance as mesophilic organisms well-established in biotechnology. This is remarkable since *S. acidocaldarius* was not adapted to the solvent and challenged by both high temperature (76 °C) and low pH (pH=3.0). In the future further adaptation approaches may enhance its 1-butanol tolerance as shown for *P. putida* strains (54) and *C. acetobutylicum* (53, 55).

Biofilm formation and EPS composition

Using analytical and microscopic techniques, we demonstrated that *S. acidocaldarius* responded to 1-butanol exposure with enhanced biofilm formation. In the presence of 0.5% and 1% (v/v) 1-butanol higher amounts of adhered cells were observed on the glass surfaces of culture flasks. In contrast, only debris of biofilm was visualized for *S. acidocaldarius* cultures grown without 1-butanol. Thus, the addition of 1-butanol may promote the biofilm formation or the formation of a more robust and stable biofilm that is more resistant to the shear forces in shaking cultures.

Under static cultivation conditions in microtiter plates, biofilm cells showed enhanced 1-butanol tolerance at 1.5% (v/v) up to 2.5% (v/v). However, at concentrations in the range of 1.5% to 2.5% (v/v) 1-butanol, biomass yield and metabolic activity of *S. acidocaldarius* biofilms also decreased gradually. For *S. acidocaldarius* an enhanced biofilm formation was previously reported for other environmental stressors including non-optimal temperature (60 °C, 85 °C), increased pH values (pH 4-6) and increased pH along with higher iron concentration (pH 6.0, 0.065 g/L iron; (11)). Similarly, immobilized *C. acetobutylicum* biofilm cells showed increased butanol tolerance with growth at 1.5 % (v/v) 1-butanol and improved ABE productivity (53). While planktonic cells showed no growth at 1.5% (v/v) butanol, *C. acetobutylicum* biofilm cells showed continuous growth (53). Thus, in general, biofilm formation seems to enhance microbial tolerance to suboptimal environmental factors due to diverse protective mechanisms, commonly with a major contribution of the EPS matrix (12, 13).

In agreement with this general observation, we detected a significant increase in the amounts of EPS proteins and carbohydrates for *S. acidocaldarius* biofilms at 1% (v/v) 1-butanol. These EPS components may facilitate enhanced biofilm adhesion to the solid surface and cohesion of cells inside the biofilm. Consistently, CLSM analysis confirmed an increased concentration and changed composition of extracellular

454 polysaccharides. A drastic increase in the total amount of EPS as well as changes in
455 EPS composition with increased protein amount was also observed for *P.*
456 *taiwanensis* VLB120ΔC biofilms grown in the presence of 0.5% (v/v) 1-butanol (32).
457 In *E. coli* the upregulation of membrane modification genes involved in
458 exopolysaccharide synthesis (i.e. M-antigen, or colanic acid) was observed in stress
459 response to butanol, other industrially relevant organic solvents and organic acids
460 such as 2,4-butanediol and acetate, respectively (23). Therefore, as well established
461 in Bacteria (56), alterations of the EPS composition and their component structures
462 seems to be a typical response to environmental stresses, here exposure to 1-
463 butanol, in *S. acidocaldarius*.
464
465 Transcriptome data support the observed switch from the planktonic to the biofilm
466 mode of life. Archaella, the type IV pilus like motility structures of Archaea, are
467 involved in biofilm formation, species interactions, and adhesion (12, 35). All
468 archaellum-encoding genes and the gene encoding ArnR1, the positive regulator of
469 archaellum synthesis were downregulated (35). Protein phosphorylation has been
470 shown to play an important regulatory role in the transition from a motile (planktonic)
471 to a sessile phenotype and thus biofilm formation (40, 42). For the Hanks-type protein
472 kinase ArnC (Saci_1193) that phosphorylates the two negative regulators of motility
473 ArnA and ArnB a slight upregulation was observed (40, 57). Also, the gene encoding
474 the archaeal biofilm regulator 1 (AbfR1) was upregulated. Deletion of *abfR1* in *S.*
475 *acidocaldarius* revealed a function in repression of EPS formation and activation of
476 motility. However, in its phosphorylated form AbfR1 was shown to support biofilm
477 formation (41, 42). AbfR2, which was shown to enhance biofilm formation, was
478 slightly downregulated (41, 43). These data suggest that major players involved in
479 the complex regulatory network for motility and biofilm formation are affected by 1-

480 butanol exposure in *S. acidocaldarius*. Notably, also in Bacteria such as *E. coli,* the
481 downregulation of flagella and chemotaxis genes is reported in response to
482 industrially relevant chemicals such as organic solvents and organic acids (23).
483
484 **Effect of 1-butanol on cell morphology**
485 We analysed the cell morphology in response to 1-butanol using SEM. Beside the
486 typical lobe-shaped cells with a smooth surface structure and flat cells with an
487 irregular surface structure, we observed a third morphotype of more rounded cells
488 with a perforated surface structure upon 1.5% (v/v) 1-butanol exposure. This is an
489 interesting observation, since heterogeneity in archaeal cell communities has not
490 been well addressed so far compared to Bacteria (58). Notably, the perforated
491 morphology of *S. acidocaldarius* cells resembles SEM pictures of the *S. islandicus* S-
492 layer deletion strain (Δ*slaAB*) reported previously (59, 60). In the S-layer model SlaB
493 forms a stalk that anchors the cap (SlaA) in the cytoplasmic membrane and forms a
494 crystalline proteinaceous matrix that covers the whole cell surface. The Δ*slaA* strain
495 was shown to lack the outermost lattice layer and increased cell size as well as large
496 aggregates “bulky clumps” were observed (60). An analysis of the chromosome
497 content of single cells via flow cytometry revealed an uneven chromosome
498 distribution and an increase of chromosome numbers in Δ*slaA* cells, suggesting a cell
499 division defect in *S. islandicus*.
500 Both the removal of the S-layer as well as the addition of 1-butanol imply membrane
501 stress. This is in line with our observation that planktonic cells in shaking cultures,
502 where shear forces were applied, were more sensitive towards 1-butanol (complete
503 growth inhibition at 1.5% (v/v)) than biofilm cells grown under static cultivation
504 conditions (tolerance of up to 2.5% (v/v)). In accordance, we observed increased
505 carbohydrate and protein concentrations in the EPS in response to 1-butanol

exposure in our study. Furthermore, membrane proteins in general as well as the *slaA* and *slaB* genes were downregulated. In contrast, *cdvB1* and less pronounced other genes of the ESCRT III machinery (except *cdvA* and *cdvB3*) were upregulated suggesting an increased activity with possible function in vesicle formation, budding and/or cytokinesis (36-39, 61). For the *cdvB* paralogs *cdvB1* and *cdvB2* a function in ring formation and constriction during cell division as well as vesicle formation was shown and for *cdvB1* a more important role under stress conditions was suggested (37, 39, 61).

Archaeal ESCRT III proteins have been identified in membrane vesicles excreted by *Sulfolobus* spp. (37). Vesicle formation is also well known for bacterial biofilms (62) where they serve various functions, including the transport of toxins, plasmid DNA, small RNA, quorum signalling molecules and proteins. In addition, membrane vesicle formation was reported as multiple-stress response mechanism that enhanced cell surface hydrophobicity and biofilm formation in *P. putida* (63).

Notably also for some bacterial strains a change in cell morphology was observed (32, 51, 52). A butanol tolerant *E. coli* σ^{70} mutant showed an increased cell size and condensed cytoplasm with occasionally invaginated bodies in the presence of butanol (52). The inner membrane was still intact and not leaky upon butanol treatment and the author suggest a self-protection mechanism against damage from solvents.

Effect of 1-butanol on cell protection mechanisms

In response to 1-butanol exposure, we observed significant changes in the cells' adaptive immunity system (CRISPR) and dormancy- or cell death-inducing defence (toxin-antitoxin) system. Many of the CRISPR-Cas system related genes are downregulated in biofilms exposed to 1-butanol. The *S. solfataricus* CRISPR type III

system has been studied (45, 64) and *in vitro* studies demonstrated RNA degrading activity for Sso-IIIB (CMR), whereas Sso-IIID (CSM) cleaved both RNA and DNA (unspecific) (65). The gene encoding the Cas10 protein (Saci_1899) is significantly downregulated. Cas10 catalyzes the formation of cyclic oligoadenylates, which act as second messenger activating defence mechanisms (66). The downregulation of the cells' immune system and defence machinery (64) may increase the potential to acquire novel DNA. *S. acidocaldarius'* biofilm matrix (EPS) contains eDNA, which might enable DNA repair and horizontal gene transfer (13). However, alternative functions of CRISPR-Cas systems are well established for Bacteria in response to environmental stressors such as nutrient starvation, iron limitation and cell envelope stressors such as phage infection and high temperature (67). Furthermore, biofilm formation was shown to be regulated by the type I CRISPR-Cas system in *P. aeruginosa* (68).

In addition, significant changes in the regulation of the toxin-antitoxin (TA) system were observed (46). In *S. solfataricus* at least 26 virulence associated protein (vap) *BC* TA loci (type II) were identified. Several *vapB/C* genes are activated by heat shock and the disruption of *vapB6* (antitoxin-encoding gene) resulted in susceptibility to thermal stress (69). TA systems are ubiquitous in prokaryotes (46). In Bacteria they are supposed to provide a mechanism of cell persistence to cope with environmental stress (70). In *B. subtilis* biofilm formation, TxpA and YqcG toxins were shown to eliminate defective cells from developing biofilms upon nutrient starvation (71). Therefore, 1-butanol exposure in *S. acidocaldarius* seems to trigger changes in the adaptive immunity system (CRISPR) and dormancy- or cell death inducing defence system (TA), supporting a functional coupling of both systems in order to allow for effective protection at the population level, as proposed previously by Makarova et al. (2013) (46).

Effect of 1-butanol on cell metabolism and general stress response

Finally, major transcriptional changes were also observed in metabolism. *Sulfolobus* spp. gain energy by aerobic respiration using a branched electron transport chain. Due to their acidophilic lifestyle (pH ~3.0 outside the cell, pH ~6.5 inside the cell) one of their major challenges is the maintenance of the intracellular pH, which is directly coupled to ATP generation via the proton motive force. In response to 1-butanol exposure, we observed significant changes in the transcription of components of the respiratory chain (i.e. cytochrome bc1 complex, DoxBCE complex). Therefore, as suggested previously for *S. acidocaldarius* in response to nutrient depletion (50), the differential regulation of the components of the respiratory chain seems to allow for adaptation to different stress conditions. The genes encoding 5-Oxoprolinase, involved in the degradation of pyroglutamate, were the most upregulated genes in response to 1-butanol exposure in this study. Pyroglutamate is formed spontaneously under thermoacidophilic conditions by cyclization of glutamate. It can be used as sole carbon source by *S. acidocaldarius* with 5-oxoprolinase as key enzyme that catalyzes the ATP-dependent formation of glutamate (47). The upregulation of genes encoding proteins involved in the catechol pathway (Saci_2293, Saci_2295) indicates the degradation of aromatic amino acids (Saci_2294; (48)). Hence, the exposure to 1% (v/v) 1-butanol resulted in obvious metabolic changes in respect to aerobic respiration and degradation of pyroglutamate and aromatic amino acids.

In addition, genes encoding peroxiredoxins, thioredoxin and thioredoxin reductase were upregulated in response to 1-butanol. They serve as antioxidant proteins, eliminating reactive oxygen species (ROS) and thus protecting the cell from oxidative damage (72). In *E. coli n*-butanol stress also resulted in the perturbation of respiratory complexes and a large increase of ROS (73).

In conclusion, *S. acidocaldarius* (shaking culture) showed a high 1-butanol tolerance of up to 1% (v/v) at 76 °C and pH 3.0. In response to solvent stress (i.e. 1-butanol, ethanol, 1-propanol and isobutanol) we observed increased biofilm formation. For *S. acidocaldarius* biofilms enhanced 1-butanol tolerance and changes in the EPS composition, biofilm architecture and cell morphology with increased heterogeneity were observed. Finally, we analysed the global response to solvent exposure at gene and protein level and identified significant changes e.g. in motility, cell envelope and amount of membrane proteins, cell division and vesicle formation, immune and defence systems, as well as metabolism and general stress response that are in line with the observed phenotypic characteristics. This is the first detailed study on solvent stress response in a crenarchaeon highlighting the impressive robustness of the thermoacidophilic *S. acidocaldarius* towards organic solvents.

Materials and Methods

Strain and cultivation of liquid *S. acidocaldarius* cultures

S. acidocaldarius DSM 639 was cultivated aerobically at 76 °C in basal Brock medium, pH 3.0 (8), supplemented with 0.1% (w/v) NZ-amine (EZMix™ N-Z-Amine®, Merck, Sigma-Aldrich, Darmstadt, Germany), 0.2% (w/v) D-glucose and different concentrations of 1-butanol (0% to 2.5% (v/v); ≥99.5% p. a., Roth, Karlsruhe, Germany) or other organic solvents (ethanol, 99.9% GC, Fisher Scientific, Thermo Fisher Scientific, Waltham, Massachusetts, USA; 1-propanol, ≥99.5% GC, Merck; isobutanol, ≥99% GC, Honeywell Riedel de Haën, Fisher Scientific). Liquid cultures were incubated with agitation (180 rpm). Pre-cultures of *S. acidocaldarius* (OD_{600nm} 1.18, logarithmic-phase) were used to inoculate fresh medium without or with 1-butanol (0% to 1.5% (v/v)) to a starting OD_{600nm} of 0.05. Cell growth (OD_{600nm} values), concentrations of D-glucose and 1-butanol were determined regularly throughout three weeks of cultivation. Experiments were performed in four biological replicates.

D-Glucose and 1-butanol quantification

For 1-butanol or D-glucose quantification, 1 ml culture was removed at different time points (0 h, 48 h, 84 h, 168 h, and 336 h) and centrifuged at 16,000 x g for 10 min. Supernatants were stored at -20 °C until usage.

The D-glucose concentration was determined photometrically using glucose-6-phosphate dehydrogenase (G6PDH) from *Thermotoga maritima* following NADPH formation at 340 nm. The assay was performed in 100 mM HEPES/NaOH buffer at pH 6.5 and 70 °C with 2 mM $NADP^+$, 6 µl of G6PDH (after heterologous expression in *Escherichia coli* and purification by heat shock (20 min, 80 °C)) and diluted (up to 25-fold) culture supernatants in a total volume of 500 µl. The reaction was started by

the addition of samples. The D-glucose concentration was calculated using a standard calibration curve ranging from 0.04 to 0.5 mM D-glucose. The quantification of 1-butanol was performed enzymatically by using alcohol dehydrogenase (ADH) from *Saccharomyces cerevisiae* (Merck) following NADH formation. Briefly, the assay was conducted in sodium phosphate buffer (20 mM $Na_4P_2O_7$ x 10 H_2O, pH 8.8) with 4 mM NAD^+, 12 U ADH and diluted culture supernatant in a total volume of 500 µl at 25 °C. NADH formation was measured at 340 nm using a photometer (SPECORD 210, Analytik Jena, Jena, Germany). 1-Butanol concentrations were calculated using a standard calibration curve ranging from 0.05 to 1 mM 1-butanol.

Cell aggregation analysis

Cell aggregation was analysed as described previously (74). Briefly, 5 µl of each culture (diluted to OD_{600} = 0.2) were spotted on microscope slides coated with 1% (w/v) agar. Cell aggregation was observed using a phase-contrast microscope with 100x magnification (Leica DMLS, Leica, Wetzlar, Germany).

Spotting plates

Samples of planktonic cell cultures were collected at different growth phases. 10 µl of the culture and dilutions (10^{-1} - 10^{-6} with Brock medium, pH=5.0-5.5) were spotted on Brock-Gelrite plates containing 0.6 % gellan gum (Gelzan™, Sigma-Aldrich), 0.1% (w/v) NZ-amine and 0.2% (w/v) D-glucose. Plates were incubated at 76 °C for four days. Then, growth was inspected and documented.

Cultivation of *S. acidocaldarius* biofilms

Brock medium, supplemented with 0.1% (w/v) NZ-amine, 0.2% (w/v) D-glucose and different 1-butanol concentrations (0% to 2.5% (v/v)), was inoculated with an exponentially growing culture to an OD_{600nm} of 0.1 and was grown for four days at

76 °C in different incubation systems for biofilm formation. Growth (OD) was monitored at 600 nm.

Cultivation in 96-well microtiter plates

To evaluate the solvent tolerance of *S. acidocaldarius* biofilms towards 1-butanol, cells were grown in presence and absence of 1-butanol (see above, 150 µl total volume, 12 cavities per growth condition) in 96-well microtiter plates (Cell+, flat bottom, polystyrene, Sarstedt). Plates were sealed with a gas impermeable aluminum foil (alu-sealing tape, pierceable, Sarstedt) and cultivated inside a metal-box containing a water reservoir to reduce evaporation of medium. After incubation, OD_{600nm} was determined using a microplate reader (InfiniteM200, Tecan). Afterwards planktonic cells were transferred into a new microtiter plate and OD_{600nm} was determined again. The remaining biofilm biomass was quantified by crystal violet staining and its metabolic activity was analysed using the resazurin assay.

Cultivation on glass slides for light and scanning electron microscopy

For light and scanning electron microscopy, biofilms were grown on sterile glass coverslips (18 mm x 18 mm; Roth) placed inside the wells of a 6-well microtiter plate (6 well cell culture plate, cellstar, Greiner Bio-One International, Kremsmünster, Austria). 4.5 ml of a *S. acidocaldarius* culture were added to each well; the plate was sealed with an aluminum foil and incubated as described above. After incubation, planktonic cells were discarded, cavities were washed with 5 ml Brock medium once and biofilms were either stained with crystal violet for light microscopy or fixed for scanning electron microscopy.

Cultivation in µ-dishes for confocal laser-scanning microscopy

4 ml of a *S. acidocaldarius* culture (0% and 1% (v/v) 1-butanol), were transferred into a µ-dish (ibi-treat, 35 mm, high, ibidi, Gräfelfing, Germany), sealed with an aluminum foil and incubated as described above.

670 **Cultivation in Petri dishes for EPS extraction, transcriptome and proteome**
671 **analyses**

672 25 ml *S. acidocaldarius* culture (0%, 0.5% and 1% (v/v) 1-butanol) were transferred
673 into polystyrene Petri dishes (92 x 16 mm, Polystyrene, without cams, Sarstedt) and
674 incubated for four days at 76 °C in an air-tight chamber. For EPS extraction,
675 transcriptomics and proteomics analyses biofilms from ten Petri dishes (for each
676 growth condition) were washed with 25 ml Brock medium each, and collected from
677 the bottom of the Petri dishes by using a cell scraper.

678 **Crystal violet staining**

679 Crystal violet staining was used to visualize cells attached to abiotic surfaces after
680 cultivation (75). The planktonic fraction of *S. acidocaldarius* liquid cultures was
681 discarded and long-neck flasks were filled with 75 ml of a 0.01% (w/v) crystal violet
682 (Merck) solution and incubated for 20 min. Afterwards the flasks were washed with
683 ddH_2O three times.
684 For light microscopy biofilm on top of glass coverslips was stained with 0.01% (w/v)
685 crystal violet solution for 20 min and washed with 5 mL ddH_2O three times
686 afterwards.
687 For 96-well microtiter plates 175 µl of a 0.1% (w/v) crystal violet solution were added
688 into each cavity, incubated for 20 min and washed three times with 200 µl ddH_2O.
689 After air-drying, each well was filled with 200 µl of 95% (v/v) ethanol and the plates
690 were incubated for 30 min to release the dye from attached biofilm into solution. The
691 absorbance of crystal violet was measured at 570 nm using a microplate reader
692 (InfiniteM200, Tecan).

693 **Resazurin assay**

694 To analyze the metabolic activity of biofilm cells a modified resazurin assay was
695 established. After cultivation in microtiter plates (see above), biofilms were washed
696 with 200 µl Brock medium. Then 200 µl of 0.005% (w/v) resazurin (sodium salt,
697 Sigma-Aldrich) in Brock medium (pH=3.0), were added into the cavities. In an acidic
698 milieu (pH=3.0) the oxidoreduction dye resazurin is protonated to resorufin and
699 exhibits a red colour. Samples were incubated at 76 °C for 3 h. By *S. acidocaldarius*
700 respiratory activity resorufin is converted to the colourless product dihydroresorufin.
701 The conversion of resorufin was determined photometrically at 520 nm using a
702 microplate reader (InfiniteM200, Tecan). All experiments were performed in
703 triplicates.

704 **Light microscopy**

705 For light microscopy biofilms were grown on glass slides and stained with crystal
706 violet as described before. After staining glass slides were air-dried and used for
707 microscopy. Images were recorded using a light- and epifluorescence microscope
708 and 4x, 40x and 100x air objectives (Eclipse Ni, Nikon, Düsseldorf, Germany).

709 **Scanning electron microscopy (SEM)**

710 For SEM *S. acidocaldarius* biofilms were grown on glass slides (see above). Cells
711 were fixed by the addition of 4 ml 2% (v/v) glutardialdehyde in Brock medium and
712 incubated for 2 h at 4 °C. Then glass slides were submerged in acetone for 30 min.
713 Biofilms were dried by critical point drying and sputtered with Au/Pd (80%/20%) for
714 15 to 30 s, resulting in a metal layer of 2.5 nm to 5.5 nm in thickness. Images were
715 taken using a scanning electron microscope (QUANTA 400 FEG, FEI company,
716 Thermo Fisher Scientific).

Confocal laser scanning microscopy (CLSM)

After biofilm cultivation in µ-dishes (see above), the supernatant was removed carefully and 1 ml of Brock medium (pH=7.0) was used to wash the submersed biofilm. For staining 2 mL of a fluorescent staining solution containing 250 µM of the DNA binding dye SYTO9 (excitation: 483 nm, emission: 503 nm; Invitrogen, Thermo Fisher Scientific), 7.5 µg/mL of the fluorophore-labeled lectin concanavalin A (ConA)-Alexa-633 (excitation: 632 nm, emission: 647 nm, α-mannopyranosyl- and α-glucopyranosyl residues; Invitrogen) and 15 µg/mL of GS-IB4-Alexa-568 (isolectin from *Griffonia simplicifolia*, excitation: 578 nm, emission: 603 nm, α-D-galactosyl and N-acetyl-D-galactosamine residues; Invitrogen) in Brock medium (pH=7.0) were used. Samples were incubated for 30 min in the dark at room temperature. After staining, the supernatant was removed, the biofilm was washed twice with 1 ml Brock medium (pH=7.0) and finally 2 ml of Brock medium were added. Visualization was performed using a Zeiss LSM 510 laser scanning microscope with an 100x oil objective. Data processing was performed using the software Imaris 8.1.2 (Bitplane, Zürich, Switzerland).

Extraction of extracellular polymeric substances (EPS)

EPS extraction was performed using the cation-exchange resin (CER, Dowex® Marathon® C sodium form, Sigma-Aldrich) extraction method as described previously (14, 76). Briefly, biofilm cells were resuspended in phosphate buffer (2 mM Na_3PO_4 x 12 H_2O, 4 mM NaH_2PO_4 x 1 H_2O, 9 mM NaCl, 1 mM KCl, pH=7.0). After CER treatment of the biofilm suspension, centrifugation and sterile filtration (14) the filtrate, corresponding to total extracellular material (TEM), contained cell-free low molecular weight compounds and EPS (of high molecular weight). To obtain EPS only, the filtrate was dialyzed against deionized water by using a dialysis membrane with a molecular weight cut-off of 3.5 kDa (14).

743 **Quantification of EPS**

744 Carbohydrate and protein content was quantified in the biofilm suspension, and in the
745 TEM and EPS solutions as described before (14). Proteins were quantified using a
746 modified Lowry assay (77) with bovine serum albumin as standard (Serva
747 Electrophoresis, Heidelberg, Germany). Carbohydrates were quantified using the
748 phenol-sulfuric acid method (78) and D-glucose as standard for neutral
749 carbohydrates.

750 **Determination of total cell counts**

751 Total cell counts in biofilm suspensions were determined by 4',6-diamidino-2-
752 phenylindole (DAPI; 5 µg/mL in 0.4% formaldehyde) staining and filtration of the cells
753 on polycarbonate filters (pore size 0.2 µm, 30 mm diameter, Merck). For microscopic
754 cell counting the filters were placed on a glass slide, covered with CitiFluor AF2
755 (Citifluor, Hatfield, Pennsylvania, USA) and a coverslip. 20 grids with 20-200
756 cells/grid were counted for statistic validity.

757 **Transcriptome analysis**

758 Biofilm and planktonic cell samples (supernatant after static incubation) were
759 generated in Petri dishes (92 x 16 mm, Polystyrene, without cams, Sarstedt) as
760 described before. Cells grown under six different conditions, including Biofilm-Control
761 (BF0), Biofilm-0.5% (v/v) 1-Butanol (BF05), Biofilm-1% (v/v) 1-Butanol (BF1),
762 Planktonic-Control (PL0), Planktonic-0.5% (v/v) 1-Butanol (PL05), Planktonic-1%
763 (v/v) 1-Butanol (PL1), were separated, harvested by centrifugation (10 min, 5,000 x g,
764 4 °C) and immediately frozen at -70 °C. Isolation of cells was performed in triplicates
765 with ten technical replicas each. Samples were pooled to obtain sufficient cell mass
766 for further processing. RNA was isolated using Trizol (Thermo Fisher Scientific) as
767 described earlier (79). The obtained RNA samples were purified as described by

768 Bischof et al., 2019 (50). In accordance, sequencing libraries were prepared and
769 quantified according to Bischof et al. (2019). Sequencing was performed on a MiSeq
770 instrument (Illumina, San Diego, California, USA) using v3 chemistry with a read
771 length of 2x76 nt. Sequencing reads were mapped with Bowtie2 (80) against the
772 reference genome of *S. acidocaldarius* DSM639. Mapped reads were counted and
773 normalized as RPKM values (81) using the software ReadXplorer (82). In contrast to
774 the original value, only reads mapping to coding sequences were considered for the
775 calculation of the total number of mapped reads. For identification of differentially
776 transcribed genes, the ratios (fold change) between the RPKM values obtained in
777 different conditions for a single gene were calculated. Additionally, an A-value (signal
778 intensity value) was determined for all genes in each comparison (0.5 * log2 (RPKM
779 condition1 * RPKM condition2)). Only genes with 4-fold (log2 fold change ≥2 or ≤-2)
780 or 2-fold changes (log2 fold change ≥1 or ≤-1) (depending on the comparison) and an
781 A-value ≥ 2 were considered as differentially transcribed.

782 **Proteome analysis**

783 Biofilms were cultivated in Petri dishes (92 x 16 mm, Polystyrene, without cams,
784 Sarstedt) and harvested as described above. Frozen cells of *S. acidocaldarius*,
785 grown under two different conditions (Biofilm-Control (BF0) and Biofilm-1% (v/v) 1-
786 butanol (BF1)), were washed twice with ice-cold water. Then protein extraction was
787 performed as described by Bischof et al. (2019) (50). Briefly, cells were resuspended
788 in protein extraction buffer and lysed using ultrasonic treatment. 100 µg protein of the
789 supernatant (21,000 x g and 4 °C for 30 min) were used for an iTRAQ analysis with
790 two 8-plex-iTRAQ tags 113 and 115 labelled for BF0 and BF1, respectively. The
791 analysis was performed based on the manufacturer's instruction (SCIEX,
792 Framingham, Massachusetts, USA) as described in Bischof et al. (2019) (50). All raw
793 data files from MS analysis were submitted to MaxQuant version 1.5.3.8 for protein

identification against *S. acidocaldarius* database (consisting of 2,222 entries) downloaded in August 2015 from Uniprot (http://www.uniprot.org). All settings, data handling as well as quantitation were performed according to Bischof et al. (2019) (50).

Data availability statement

RNA-seq data have been deposited in the ArrayExpress database at EMBL-EBI (www.ebi.ac.uk/arrayexpress) under accession number E-MTAB-10093. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (83) partner repository with the dataset identifier PXD023858.

806 **Acknowledgements**

807 JCB acknowledges funding by the Mercator foundation for support with a Mercur
808 startup grant (Pr-2011-0058) and by the German Federal Ministry of Education and
809 Research (BMBF) grant SulfoSYS^{BIOTECH}, 0316188A within the e:Bio funding
810 initiative. AA was supported by grant 0316188D within SulfoSYS^{BIOTECH}. LK
811 acknowledges funding by the Deutsche Forschungsgemeinschaft (DFG, SI 642/13-
812 1). XZ acknowledges funding by the VW Stiftung in the “Life?” initiative (96725). KSM
813 is supported by the Intramural Research Program of the National Institutes of Health
814 of the USA (National Library of Medicine). TKP and PCW acknowledge the BBSRC
815 funding BB/M018172/1 and BB/M018288/1.

816

817 **Author contributions**

818 Conceptualization, CB, JW, HCF and BS; Investigation, JCB, LK, XZ, AM, TKP and
819 KSM; Writing - Original Draft, JCB and LK; Writing Review & Editing, LK, PCW, JK,
820 KSM, HCF and BS; Funding Acquisition, JW, HCF and BS; Supervision, PCW, JK,
821 CB, HCF, JW and BS.

822

823 **Conflict of Interest**

824 The authors declare no conflict of interest.

825 **References**

826 1. Adam PS, Borrel G, Brochier-Armanet C, Gribaldo S. 2017. The growing tree
827 of Archaea: new perspectives on their diversity, evolution and ecology. ISME J
828 11:2407-2425. DOI: 10.1038/ismej.2017.122
829 2. Valentine DL. 2007. Adaptations to energy stress dictate the ecology and
830 evolution of the Archaea. Nat Rev Microbiol 5:316-323. DOI:
831 10.1038/nrmicro1619
832 3. Schocke L, Bräsen C, Siebers B. 2019. Thermoacidophilic *Sulfolobus* species
833 as source for extremozymes and as novel archaeal platform organisms. Curr
834 Opin Biotechnol 59:71-77. DOI: 10.1016/j.copbio.2019.02.012
835 4. Zeldes BM, Keller MW, Loder AJ, Straub CT, Adams MW, Kelly RM. 2015.
836 Extremely thermophilic microorganisms as metabolic engineering platforms for
837 production of fuels and industrial chemicals. Front Microbiol 6:1209. DOI:
838 10.3389/fmicb.2015.01209
839 5. Cabrera MA, Blamey JM. 2018. Biotechnological applications of archaeal
840 enzymes from extreme environments. Biol Res 51:37. DOI: 10.1186/s40659-
841 018-0186-3
842 6. Caforio A, Driessen AJM. 2017. Archaeal phospholipids: Structural properties
843 and biosynthesis. Biochim Biophys Acta Mol Cell Biol Lipids 1862:1325-1339.
844 DOI: 10.1016/j.bbalip.2016.12.006
845 7. Rastädter K, Wurm DJ, Spadiut O, Quehenberger J. 2020. The cell membrane
846 of *Sulfolobus* spp.-homeoviscous adaption and biotechnological applications.
847 Int J Mol Sci 21. DOI: 10.3390/ijms21113935
848 8. Brock TDB, K. M.; Belly, R. T.; Weiss, R. L. 1972. *Sulfolobus*: A new genus of
849 sulfur-oxidizing Bacteria living at low pH and high temperature. Archiv
850 Mikrobiol 84:54-68. DOI: 10.1007/BF00408082

9. Quehenberger J, Shen L, Albers SV, Siebers B, Spadiut O. 2017. *Sulfolobus* - A potential key organism in future biotechnology. Front Microbiol 8:2474. DOI: 10.3389/fmicb.2017.02474

10. van der Kolk N, Wagner A, Wagner M, Wassmer B, Siebers B, Albers SV. 2020. Identification of XylR, the activator of arabinose/xylose inducible regulon in *Sulfolobus acidocaldarius* and its application for homologous protein expression. Front Microbiol 11:1066. DOI: 10.3389/fmicb.2020.01066

11. Koerdt A, Gödeke J, Berger J, Thormann KM, Albers SV. 2010. Crenarchaeal biofilm formation under extreme conditions. PloS ONE 5:e14104. DOI: 10.1371/journal.pone.0014104

12. van Wolferen M, Orell A, Albers SV. 2018. Archaeal biofilm formation. Nat Rev Microbiol 16:699-713. DOI: 10.1038/s41579-018-0058-4

13. Flemming HC, Wingender J. 2010. The biofilm matrix. Nat Rev Microbiol 8:623-633. DOI: 10.1038/nrmicro2415

14. Jachlewski S, Jachlewski WD, Linne U, Bräsen C, Wingender J, Siebers B. 2015. Isolation of extracellular polymeric substances from biofilms of the thermoacidophilic archaeon *Sulfolobus acidocaldarius*. Front Bioeng Biotechnol 3:123. DOI: 10.3389/fbioe.2015.00123

15. Flemming HC, Wuertz S. 2019. Bacteria and archaea on Earth and their abundance in biofilms. Nat Rev Microbiol 17:247-260. DOI: 10.1038/s41579-019-0158-9

16. Moon HG, Jang YS, Cho C, Lee J, Binkley R, Lee SY. 2016. One hundred years of clostridial butanol fermentation. FEMS Microbiol Lett 363. DOI: 10.1093/femsle/fnw001

875 17. García V, Päkkilä J, Ojamo H, Muurinen E, Keiski RL. 2011. Challenges in
876 biobutanol production: How to improve the efficiency? Renew Sust Energ Rev
877 15:964-980. DOI: 10.1016/j.rser.2010.11.008
878 18. Dürre P. 2007. Biobutanol: an attractive biofuel. Biotechnol J 2:1525-1534.
879 DOI: 10.1002/biot.200700168
880 19. Knoshaug EP, Zhang M. 2009. Butanol tolerance in a selection of
881 microorganisms. Appl Microbiol Biotechnol 153:13-20. DOI: 10.1007/s12010-
882 008-8460-4
883 20. Huffer S, Clark ME, Ning JC, Blanch HW, Clark DS. 2011. Role of alcohols in
884 growth, lipid composition, and membrane fluidity of yeasts, bacteria, and
885 archaea. Appl Environ Microbiol 77:6400-6408. DOI: 10.1128/AEM.00694-11
886 21. Weber FJ, de Bont JA. 1996. Adaptation mechanisms of microorganisms to
887 the toxic effects of organic solvents on membranes. Biochim Biophys Acta
888 1286:225-245. DOI: 10.1046/j.1462-2920.1999.00033.x
889 22. Ezeji T, Milne C, Price ND, Blaschek HP. 2010. Achievements and
890 perspectives to overcome the poor solvent resistance in acetone and butanol-
891 producing microorganisms. Appl Microbiol Biotechn 85:1697-1712. DOI:
892 10.1007/s00253-009-2390-0
893 23. Rau MH, Calero P, Lennen RM, Long KS, Nielsen AT. 2016. Genome-wide
894 *Escherichia coli* stress response and improved tolerance towards industrially
895 relevant chemicals. Microb Cell Fact 15:176. DOI: 10.1186/s12934-016-0577-
896 5
897 24. Segura A, Molina L, Fillet S, Krell T, Bernal P, Munoz-Rojas J, Ramos JL.
898 2012. Solvent tolerance in Gram-negative bacteria. Curr Opin Biotechnol
899 23:415-421. DOI: 10.1016/j.copbio.2011.11.015

25. Nicolaou SA, Gaida SM, Papoutsakis ET. 2010. A comparative view of metabolite and substrate stress and tolerance in microbial bioprocessing: From biofuels and chemicals, to biocatalysis and bioremediation. Metab Eng 12:307-331. DOI: 10.1016/j.ymben.2010.03.004

26. Sinensky M. 1974. Homeoviscous adaptation—a homeostatic process that regulates the viscosity of membrane lipids in *Escherichia coli*. Proc Nat Acad Sci USA 71:522-525. DOI: 10.1073/pnas.71.2.522

27. Eberlein C, Baumgarten T, Starke S, Heipieper HJ. 2018. Immediate response mechanisms of Gram-negative solvent-tolerant bacteria to cope with environmental stress: *cis-trans* isomerization of unsaturated fatty acids and outer membrane vesicle secretion. Appl Microbiol Biotechnol 102:2583-2593. DOI: 10.1007/s00253-018-8832-9

28. Vasylkivska M, Patakova P. 2020. Role of efflux in enhancing butanol tolerance of bacteria. J Biotechnol 320:17-27. DOI: 10.1016/j.jbiotec.2020.06.008

29. Sedlar K, Kolek J, Gruber M, Jureckova K, Branska B, Csaba G, Vasylkivska M, Zimmer R, Patakova P, Provaznik I. 2019. A transcriptional response of *Clostridium beijerinckii* NRRL B-598 to a butanol shock. Biotechnol Biofuels 12:243. DOI: 10.1186/s13068-019-1584-7

30. Zhuang W, Yang J, Wu J, Liu D, Zhou J, Chen Y, Ying H. 2016. Extracellular polymer substances and the heterogeneity of *Clostridium acetobutylicum* biofilm induced tolerance to acetic acid and butanol. RSC Adv 6:33695-33704. DOI: 10.1039/c5ra24923f

31. Liu D, Yang Z, Chen Y, Zhuang W, Niu H, Wu J, Ying H. 2018. *Clostridium acetobutylicum* grows vegetatively in a biofilm rich in heteropolysaccharides

925 and cytoplasmic proteins. Biotechnol Biofuels 11:315. DOI: 10.1186/s13068-
926 018-1316-4
927 32. Halan B, Vassilev I, Lang K, Schmid A, Buehler K. 2016. Growth of
928 *Pseudomonas taiwanensis* VLB 120Δ C biofilms in the presence of *n*-butanol.
929 J Microbial biotechnology 10:745-755. DOI: 10.1111/1751-7915.12413
930 33. Liu H, Fang HHP. 2002. Extraction of extracellular polymeric substances
931 (EPS) of sludges. J Biotechnol 95:249–256. DOI: 10.1016/S0168-
932 1656(02)00025-1
933 34. Makarova KS, Wolf YI, Koonin EV. 2015. Archaeal clusters of orthologous
934 genes (arCOGs): An update and application for analysis of shared features
935 between Thermococcales, Methanococcales, and Methanobacteriales. Life
936 5:818-840. DOI: 10.3390/life5010818
937 35. Albers SV, Jarrell KF. 2018. The archaellum: An update on the unique
938 archaeal motility structure. Trends Microbiol 26:351-362. DOI:
939 10.1016/j.tim.2018.01.004
940 36. Samson RY, Obita, T., Freund, S. M., Williams, R. L., & Bell, S. D. 2008. A
941 role for the ESCRT system in cell division in archaea. Science 322:1710-1713.
942 DOI: 10.1126/science.1165322
943 37. Ellen AF, Albers SV, Huibers W, Pitcher A, Hobel CF, Schwarz H, Folea M,
944 Schouten S, Boekema EJ, Poolman B, Driessen AJ. 2009. Proteomic analysis
945 of secreted membrane vesicles of archaeal *Sulfolobus* species reveals the
946 presence of endosome sorting complex components. Extremophiles 13:67-79.
947 DOI: 10.1007/s00792-008-0199-x
948 38. Tarrason Risa G, Hurtig F, Bray S, Hafner AE, Harker-Kirschneck L, Faull P,
949 Davis C, Papatziamou D, Mutavchiev DR, Fan C, Meneguello L, Arashiro
950 Pulschen A, Dey G, Culley S, Kilkenny M, Souza DP, Pellegrini L, de Bruin

RAM, Henriques R, Snijders AP, Saric A, Lindas AC, Robinson NP, Baum B. 2020. The proteasome controls ESCRT-III-mediated cell division in an archaeon. Science 369. DOI: 10.1126/science.aaz2532

39. Pulschen AA, Mutavchiev DR, Culley S, Sebastian KN, Roubinet J, Roubinet M, Risa GT, van Wolferen M, Roubinet C, Schmidt U, Dey G, Albers SV, Henriques R, Baum B. 2020. Live imaging of a hyperthermophilic archaeon reveals distinct roles for two ESCRT-III homologs in ensuring a robust and symmetric division. Curr Biol 30:1-8. DOI: 10.1016/j.cub.2020.05.021

40. Esser D, Hoffmann L, Pham TK, Bräsen C, Qiu W, Wright PC, Albers SV, Siebers B. 2016. Protein phosphorylation and its role in archaeal signal transduction. FEMS Microbiol Rev 40:625-647. DOI: 10.1093/femsre/fuw020

41. Orell A, Peeters E, Vassen V, Jachlewski S, Schalles S, Siebers B, Albers SV. 2013. Lrs14 transcriptional regulators influence biofilm formation and cell motility of Crenarchaea. ISME J 7:1886-1898. DOI: 10.1038/ismej.2013.68

42. Li L, Banerjee A, Bischof LF, Maklad HR, Hoffmann L, Henche AL, Veliz F, Bildl W, Schulte U, Orell A, Essen LO, Peeters E, Albers SV. 2017. Wing phosphorylation is a major functional determinant of the Lrs14-type biofilm and motility regulator AbfR1 in *Sulfolobus acidocaldarius*. Mol Microbiol 105:777-793. DOI: 10.1111/mmi.13735

43. Vogt MS, Volpel SL, Albers SV, Essen LO, Banerjee A. 2018. Crystal structure of an Lrs14-like archaeal biofilm regulator from *Sulfolobus acidocaldarius*. Acta Crystallogr D Struct Biol 74:1105-1114. DOI: 10.1107/S2059798318014146

44. Makarova KS, Wolf YI, Iranzo J, Shmakov SA, Alkhnbashi OS, Brouns SJJ, Charpentier E, Cheng D, Haft DH, Horvath P, Moineau S, Mojica FJM, Scott D, Shah SA, Siksnys V, Terns MP, Venclovas C, White MF, Yakunin AF, Yan W, Zhang F, Garrett RA, Backofen R, van der Oost J, Barrangou R, Koonin

977 EV. 2020. Evolutionary classification of CRISPR-Cas systems: a burst of class
978 2 and derived variants. Nat Rev Microbiol 18:67-83. DOI: 10.1038/s41579-
979 019-0299-x
980 45. Zhang J, White MF. 2013. Hot and crispy: CRISPR-Cas systems in the
981 hyperthermophile *Sulfolobus solfataricus*. Biochem Soc Trans 41:1422-1426.
982 DOI: 10.1042/BST20130031
983 46. Makarova KS, Wolf YI, Koonin EV. 2013. Comparative genomics of defense
984 systems in archaea and bacteria. Nucleic Acids Res 41:4360-4377. DOI:
985 10.1093/nar/gkt157
986 47. Vetter AM, Helmecke J, Schomburg D, Neumann-Schaal M. 2019. The impact
987 of pyroglutamate: *Sulfolobus acidocaldarius* has a growth advantage over
988 *Saccharolobus solfataricus* in glutamate-containing media. Archaea
989 2019:3208051. DOI: 10.1155/2019/3208051
990 48. Stark H, Wolf J, Albersmeier A, Pham TK, Hofmann JD, Siebers B, Kalinowski
991 J, Wright PC, Neumann-Schaal M, Schomburg D. 2017. Oxidative Stickland
992 reactions in an obligate aerobic organism - amino acid catabolism in the
993 Crenarchaeon *Sulfolobus solfataricus*. FEBS J 284:2078-2095. DOI:
994 10.1111/febs.14105
995 49. Auernik KS, Kelly RM. 2008. Identification of components of electron transport
996 chains in the extremely thermoacidophilic crenarchaeon *Metallosphaera*
997 *sedula* through iron and sulfur compound oxidation transcriptomes. Appl
998 Environ Microbiol 74:7723-7732. DOI: 10.1128/AEM.01545-08
999 50. Bischof LF, Haurat MF, Hoffmann L, Albersmeier A, Wolf J, Neu A, Pham TK,
1000 Albaum SP, Jakobi T, Schouten S. 2019. Early response of *Sulfolobus*
1001 *acidocaldarius* to nutrient limitation. Front Microbiol 9:3201. DOI:
1002 10.3389/fmicb.2018.03201

1003 51. Vinayavekhin N, Mahipant G, Vangnai AS, Sangvanich P. 2015. Untargeted
1004 metabolomics analysis revealed changes in the composition of glycerolipids
1005 and phospholipids in *Bacillus subtilis* under 1-butanol stress. Appl Microbiol
1006 Biotechnol 99:5971-5983. DOI: 10.1007/s00253-015-6692-0
1007 52. Si H-M, Zhang F, Wu A-N, Han R-Z, Xu G-C, Ni Y. 2016. DNA microarray of
1008 global transcription factor mutant reveals membrane-related proteins involved
1009 in *n*-butanol tolerance in *Escherichia coli*. J Biotechnol Biofuels 9:114. DOI:
1010 10.1186/s13068-016-0527-9
1011 53. Liu XB, Gu QY, Yu XB. 2013. Repetitive domestication to enhance butanol
1012 tolerance and production in *Clostridium acetobutylicum* through artificial
1013 simulation of bio-evolution. Bioresour Technol 130:638-643. DOI:
1014 10.1016/j.biortech.2012.12.121
1015 54. Rühl J, Schmid A, Blank LM. 2009. Selected *Pseudomonas putida* strains able
1016 to grow in the presence of high butanol concentrations. Appl Environ Microbiol
1017 75:4653-4656. DOI: 10.1128/AEM.00225-09
1018 55. Liu S, Qureshi N, Hughes SR. 2017. Progress and perspectives on improving
1019 butanol tolerance. World J Microbiol Biotechnol 33:51. DOI: 10.1007/s11274-
1020 017-2220-y
1021 56. Flemming HC, Wingender J, Szewzyk U, Steinberg P, Rice SA, Kjelleberg S.
1022 2016. Biofilms: an emergent form of bacterial life. Nat Rev Microbiol 14:563-
1023 575. DOI: 10.1038/nrmicro.2016.94
1024 57. Hoffmann L, Schummer A, Reimann J, Haurat MF, Wilson AJ, Beeby M,
1025 Warscheid B, Albers SV. 2017. Expanding the archaellum regulatory network–
1026 the eukaryotic protein kinases ArnC and ArnD influence motility of *Sulfolobus*
1027 *acidocaldarius*. J Microbiology 6:e00414. DOI: 10.1002/mbo3.414

1028 58. Stewart PS, Franklin MJ. 2008. Physiological heterogeneity in biofilms. Nat
1029 Rev Microbiol 6:199-210. DOI: 10.1038/nrmicro1838
1030 59. Zhang C, Phillips APR, Wipfler RL, Olsen GJ, Whitaker RJ. 2018. The
1031 essential genome of the crenarchaeal model *Sulfolobus islandicus*. Nat
1032 Commun 9:4908. DOI: 10.1038/s41467-018-07379-4
1033 60. Zhang C, Wipfler RL, Li Y, Wang Z, Hallett EN, Whitaker RJ. 2019. Cell
1034 structure changes in the hyperthermophilic Crenarchaeon *Sulfolobus*
1035 *islandicus* lacking the S-Layer. mBio 10:e01589-19. DOI: 10
1036 .1128/mBio.01589-19
1037 61. Makarova KS, Yutin N, Bell SD, Koonin EV. 2010. Evolution of diverse cell
1038 division and vesicle formation systems in Archaea. Nat Rev Microbiol 8:731-
1039 741. DOI: 10.1038/nrmicro2406
1040 62. Toyofuku M, Nomura N, Eberl L. 2019. Types and origins of bacterial
1041 membrane vesicles. Nat Rev Microbiol 17:13-24. DOI: 10.1038/s41579-018-
1042 0112-2
1043 63. Baumgarten T, Sperling S, Seifert J, von Bergen M, Steiniger F, Wick LY,
1044 Heipieper HJ. 2012. Membrane vesicle formation as a multiple-stress
1045 response mechanism enhances *Pseudomonas putida* DOT-T1E cell surface
1046 hydrophobicity and biofilm formation. Appl Environ Microbiol 78:6217-6224.
1047 DOI: 10.1128/AEM.01525-12
1048 64. Makarova KS, Wolf YI, Alkhnbashi OS, Costa F, Shah SA, Saunders SJ,
1049 Barrangou R, Brouns SJ, Charpentier E, Haft DH, Horvath P, Moineau S,
1050 Mojica FJ, Terns RM, Terns MP, White MF, Yakunin AF, Garrett RA, van der
1051 Oost J, Backofen R, Koonin EV. 2015. An updated evolutionary classification
1052 of CRISPR-Cas systems. Nat Rev Microbiol 13:722-736. DOI:
1053 10.1038/nrmicro3569

1054 65. Zhang J, Graham S, Tello A, Liu H, White MF. 2016. Multiple nucleic acid
1055 cleavage modes in divergent type III CRISPR systems. Nucleic Acids Res
1056 44:1789-1799. DOI: 10.1093/nar/gkw020
1057 66. Rouillon C, Athukoralage JS, Graham S, Gruschow S, White MF. 2018.
1058 Control of cyclic oligoadenylate synthesis in a type III CRISPR system. Elife 7.
1059 DOI: 10.7554/eLife.36734
1060 67. Ratner HK, Sampson TR, Weiss DS. 2015. I can see CRISPR now, even
1061 when phage are gone: a view on alternative CRISPR-Cas functions from the
1062 prokaryotic envelope. Curr Opin Infect Dis 28:267-274. DOI:
1063 10.1097/QCO.0000000000000154
1064 68. Cady KC, O'Toole GA. 2011. Non-identity-mediated CRISPR-bacteriophage
1065 interaction mediated via the Csy and Cas3 proteins. J Bacteriol 193:3433-
1066 3445. DOI: 10.1128/JB.01411-10
1067 69. Maezato Y, Daugherty A, Dana K, Soo E, Cooper C, Tachdjian S, Kelly RM,
1068 Blum P. 2011. VapC6, a ribonucleolytic toxin regulates thermophilicity in the
1069 crenarchaeote *Sulfolobus solfataricus*. RNA 17:1381-1392. DOI:
1070 10.1261/rna.2679911
1071 70. Buts L, Lah J, Dao-Thi MH, Wyns L, Loris R. 2005. Toxin-antitoxin modules as
1072 bacterial metabolic stress managers. Trends Biochem Sci 30:672-679. DOI:
1073 10.1016/j.tibs.2005.10.004
1074 71. Bloom-Ackermann Z, Steinberg N, Rosenberg G, Oppenheimer-Shaanan Y,
1075 Pollack D, Ely S, Storzi N, Levy A, Kolodkin-Gal I. 2016. Toxin-Antitoxin
1076 systems eliminate defective cells and preserve symmetry in *Bacillus subtilis*
1077 biofilms. Environ Microbiol 18:5032-5047. DOI: 10.1111/1462-2920.13471
1078 72. Maaty WS, Wiedenheft B, Tarlykov P, Schaff N, Heinemann J, Robison-Cox J,
1079 Valenzuela J, Dougherty A, Blum P, Lawrence CM, Douglas T, Young MJ,

1080 Bothner B. 2009. Something old, something new, something borrowed; how
1081 the thermoacidophilic archaeon *Sulfolobus solfataricus* responds to oxidative
1082 stress. PLoS One 4:e6964. DOI: 10.1371/journal.pone.0006964
1083 73. Rutherford BJ, Dahl RH, Price RE, Szmidt HL, Benke PI, Mukhopadhyay A,
1084 Keasling JD. 2010. Functional genomic study of exogenous *n*-butanol stress in
1085 *Escherichia coli*. Appl Environ Microbiol 76:1935-1945. DOI:
1086 10.1128/AEM.02323-09
1087 74. Schult F, Le TN, Albersmeier A, Rauch B, Blumenkamp P, van der Does C,
1088 Goesmann A, Kalinowski J, Albers SV, Siebers B. 2018. Effect of UV
1089 irradiation on *Sulfolobus acidocaldarius* and involvement of the general
1090 transcription factor TFB3 in the early UV response. Nucleic Acids Res
1091 46:7179-7192. DOI: 10.1093/nar/gky527
1092 75. O'Toole GA. 2011. Microtiter dish biofilm formation assay. J Vis Exp 47:e2437.
1093 DOI: 10.3791/2437
1094 76. Frølund B, Palmgren R, Keiding K, Nielsen PH. 1996. Extraction of
1095 extracellular polymers from activated sludge using a cation exchange resin.
1096 Wat Res 30:1749-1758. DOI: 10.1016/0043-1354(95)00323-1
1097 77. Peterson GL. 1977. A simplification of the protein assay method of Lowry et al.
1098 which is more generally applicable. Anal Biochem 83:346-356. DOI:
1099 10.1016/0003-2697(77)90043-4
1100 78. Dubois M, Gilles KA, Hamilton JK, Rebers PA, Smith F. 1956. Colorimetric
1101 method for determination of sugars and related substances. Anal Chem
1102 28:350-356. DOI: 10.1021/ac60111a017
1103 79. Hottes AK, Meewan M, Yang D, Arana N, Romero P, McAdams HH, Stephens
1104 C. 2004. Transcriptional profiling of *Caulobacter crescentus* during growth on

complex and minimal media. J Bacteriol 186:1448-1461. DOI: 10.1128/jb.186.5.1448-1461.2004

80. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357-359. DOI: 10.1038/nmeth.1923

81. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 5:621-628. DOI: 10.1038/nmeth.1226

82. Hilker R, Stadermann KB, Doppmeier D, Kalinowski J, Stoye J, Straube J, Winnebald J, Goesmann A. 2014. ReadXplorer-visualization and analysis of mapped sequences. Bioinformatics 30:2247-2254. DOI: 10.1093/bioinformatics/btu205

83. Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, Inuganti A, Griss J, Mayer G, Eisenacher M, Pérez E, Uszkoreit J, Pfeuffer J, Sachsenberg T, Yilmaz S, Tiwary S, Cox J, Audain E, Walzer M, Jarnuczak AF, Ternent T, Brazma A, Vizcaíno JA (2019). The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucleic Acids Res 47(D1):D442-D450 (PubMed ID: 30395289). DOI: 10.1093/nar/gky1106

1124 **Tables**

1125 **Table 1: Dataset correlation and number of differentially transcribed genes.**

1126 BF: Biofilm cells, PL: Planktonic cells, 0: Control without 1-butanol, 05: 0.5% (v/v) 1-

1127 butanol, 1: 1% (v/v) 1-butanol

Comparison		Number of regulated transcribed genes (log2 fold change ≥ 2or ≤ -2, A-value ≥ 2)			
Samples		**R^2-value**	**Down**	**Up**	**Total**
Lifestyle	BF0/PL0	0.91	13	2	15
	BF1/PL1	0.88	6	38	42
Biofilm	BF1/BF0	0.77	74	43	117
	BF05/BF0	0.91	4	12	16
	BF1/BF05	0.83	56	20	76
Planktonic	PL1/PL0	0.76	89	33	122
	PL05/PL0	0.93	3	5	8

1128 **Figure legends**

1129

1130 **Figure 1: Effect of 1-butanol on cell growth of *S. acidocaldarius* DSM 639 in**
1131 **liquid cultures.** *S. acidocaldarius* DSM 639 cells were grown in Brock medium
1132 supplemented with 0.1% (w/v) NZ-amine and 0.2% (w/v) D-glucose in the absence
1133 and presence of 1-butanol (0-1.5% (v/v), long-neck flasks, 76 °C, pH 3.0, 180 rpm).
1134 As abiotic control medium with 1% (v/v) 1-butanol was used to monitor 1-butanol loss
1135 due to evaporation. **A**. Growth of *S. acidocaldarius* determined by turbidity
1136 measurement (OD_{600nm}); **B**. D-Glucose consumption; **C**. Change of 1-butanol
1137 concentration. Experiments were carried out in four biological replicates.

1138

1139 **Figure 2: Concentration-dependent effect of 1-butanol on biofilm formation and**
1140 **cell viability of *S. acidocaldarius*.** Cells were statically incubated in 96-well
1141 microtiter plates in Brock medium containing 0.1% (w/v) NZ-amine and 0.2% (w/v) D-
1142 glucose in the presence of different 1-butanol concentrations (0-2.5% (v/v)) at 76 °C
1143 for four days. **A**. OD_{600nm} values of biofilm cells. **B**. Quantification of biofilm biomass
1144 by crystal violet staining (absorbance at 570 nm). **C**. Respiratory activity of
1145 *S. acidocaldarius* biofilm cells determined by the resazurin assay, measuring
1146 resazurin reduction by absorbance at 520 nm. Activity is expressed as decrease of
1147 absorbance over 3 h, ΔA_{520nm}. Experiments were carried out in three biological
1148 replicates.

1149

1150 **Figure 3: Effect of 1-butanol on *S. acidocaldarius* cell distribution and**
1151 **morphology analysed by light microscopy and SEM.** *S. acidocaldarius* was grown
1152 on glass surfaces for four days at 76 °C in the absence and presence of 1% and
1153 1.5% (v/v) 1-butanol by light microscopy (A) and by SEM (B). **A**. Attached cells

1154 stained with crystal violet and air-dried for subsequent analysis by light microscopy.
1155 **B**. Visualization of biofilm cell distribution and morphology by SEM.

1156

1157 **Figure 4: Effect of 1-butanol on biofilm architecture of *S. acidocaldarius***
1158 **analysed by CLSM.** Submersed biofilms were grown at 76 °C for four days under
1159 static conditions in µ-dishes (ibidi). Cells were stained with SYTO 9 (green signals),
1160 carbohydrates were visualized using the fluorescently labelled lectins GS-IB4-Alexa
1161 568 (binding to α-D-galactosyl and N-acetyl-D-galactosamine residues, blue signals)
1162 and ConA-Alexa 633 (binding to α-mannopyranosyl- and α-glucopyranosyl residues,
1163 red signals). Scale bars: 10 µm.

1164

1165 **Figure 5: Influence of 1-butanol on EPS composition.** *S. acidocaldarius* biofilms
1166 were incubated at 76 °C for four days and absolute concentration values of EPS
1167 components were normalized to the total cell counts. The amount of proteins and
1168 carbohydrates was determined in different biofilm factions (**A-C**). **A**. Total biofilms
1169 suspended in phosphate buffer containing biofilm cells and extracellular compounds.
1170 **B**. Total extracellular material (TEM): fractions after EPS extraction by the CER
1171 method and sterile filtration (CER-extracted EPS material, comprising high and low
1172 MW extracellular compounds). **C**. EPS: molecules ≥ 3.5 kDa (EPS compounds
1173 obtained after dialysis of TEM fraction using 3.5 kDa cut-off membranes).

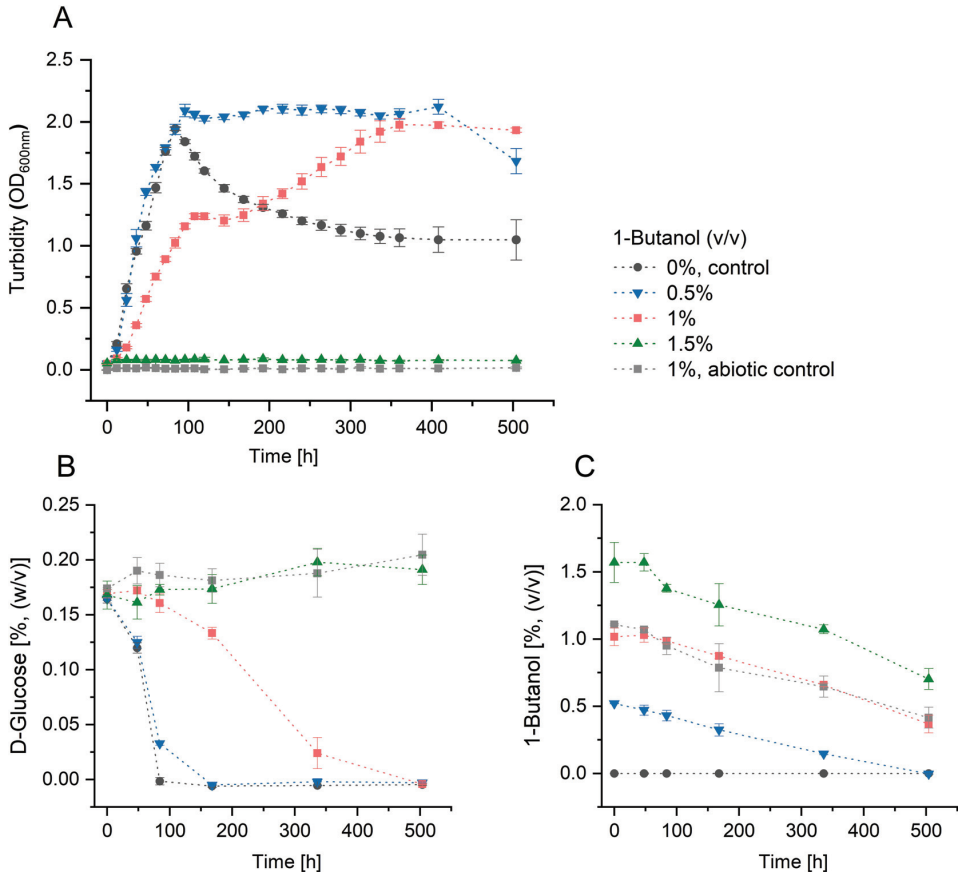
1174

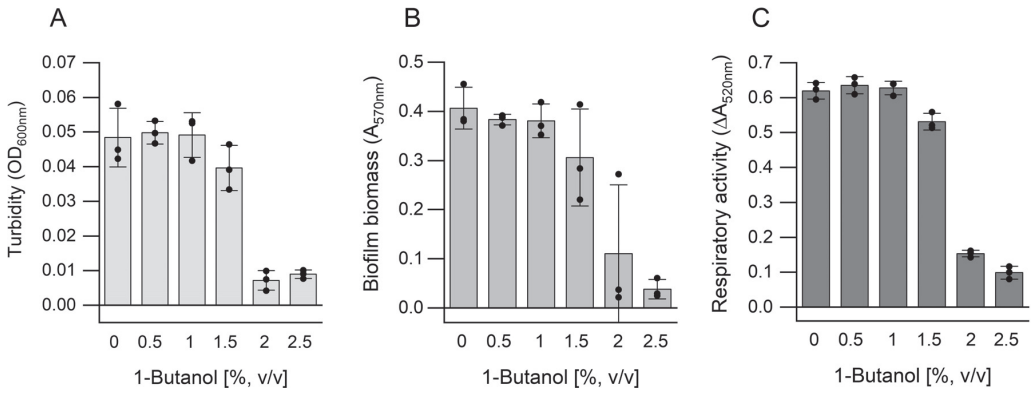
1175 **Figure 6: ArCOG classification of genes regulated in response to 1% (v/v) 1-**
1176 **butanol.** Differential regulation (log2-fold changes) of genes involved in the different
1177 arCOG categories in biofilm and planktonic cells of *S. acidocaldarius* (static
1178 cultivation in Petri dishes, 76 °C, 4 d) grown in the presence and absence of 1% (v/v)

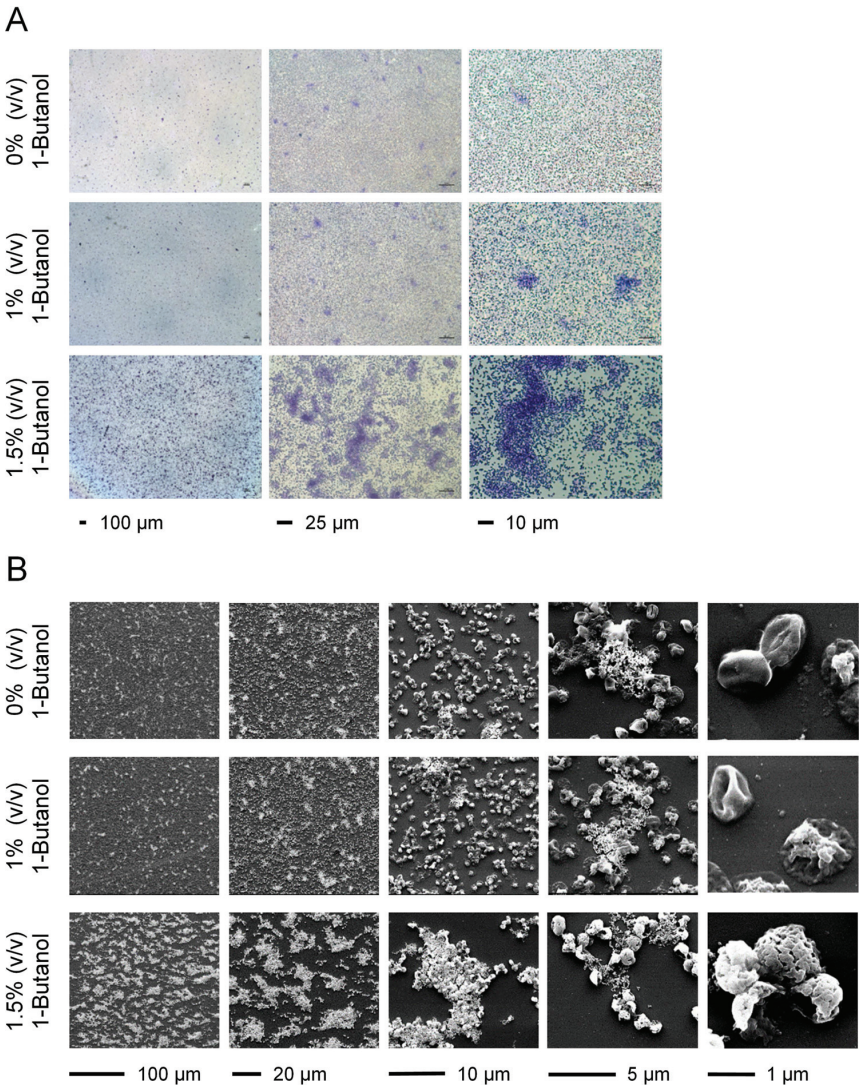
1179 1-butanol. C - Energy production and conversion; D - Cell cycle control, cell division,
1180 chromosome partitioning; E - Amino acid transport and metabolism; F - Nucleotide
1181 transport and metabolism; G - Carbohydrate transport and metabolism; H -
1182 Coenzyme transport and metabolism; I - Lipid transport and metabolism; J -
1183 Translation, ribosomal structure and biogenesis; K - Transcription; L - Replication,
1184 recombination and repair; M - Cell wall/membrane/envelope biogenesis; N - Cell
1185 motility; O - Posttranslational modification, protein turnover, chaperones; P -
1186 Inorganic ion transport and metabolism; Q - Secondary metabolites biosynthesis,
1187 transport and catabolism; R - General function prediction only; S - Function unknown;
1188 T - Signal transduction mechanisms; U - Intracellular trafficking, secretion, and
1189 vesicular transport; ; V - Defense mechanisms; X - Mobilome: prophages,
1190 transposons.

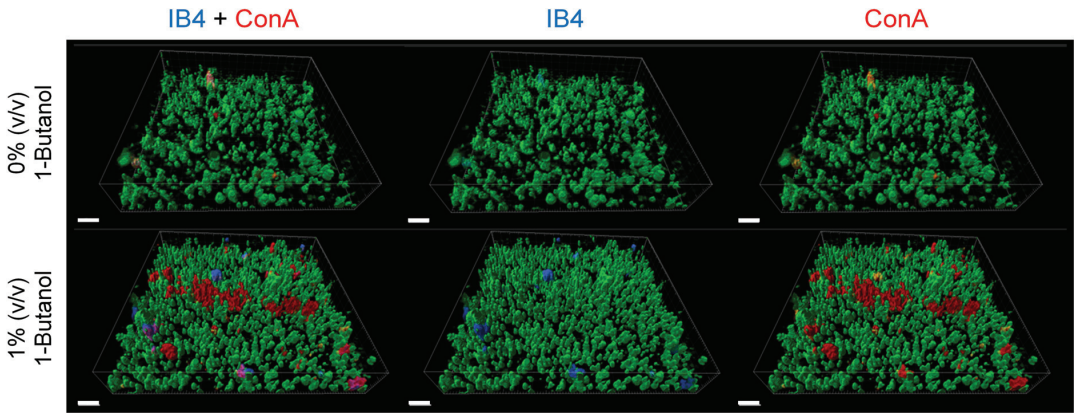
1191

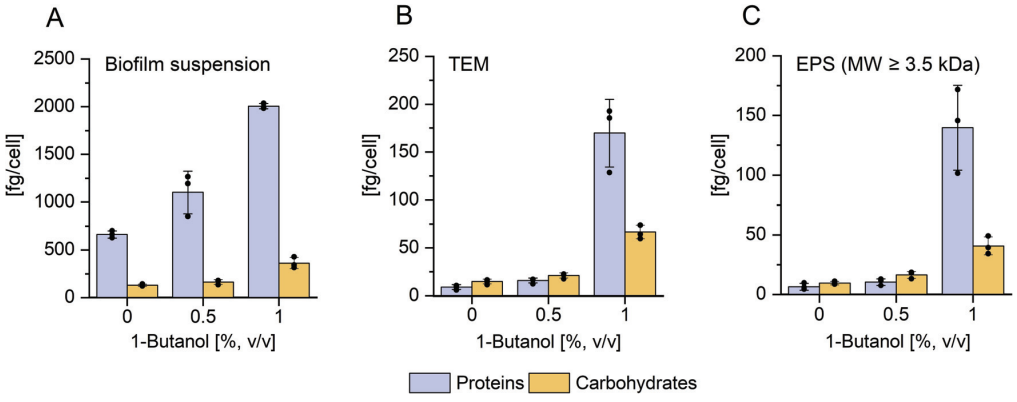
1192 **Figure 7**: **Model of 1-butanol stress response in S. acidocaldarius biofilm cells.**
1193 The increased and decreased transcription of genes in cellular structures and
1194 processes is depicted by green and red colour, respectively. Major EPS matrix
1195 components (polysaccharides, proteins and eDNA) are distributed between the cells.
1196 Genes encoding membrane proteins, the archaellum for motility, the adaptive
1197 immune system (CRISPR-Cas) the dormancy- or cell death inducing defence (toxin-
1198 antitoxin) system and components of the respiratory chain are downregulated (red).
1199 Genes encoding proteins of the ROS defence system and the ESCRT-III system for
1200 vesicle formation and/or cytokinesis are upregulated (green). Several transcriptional
1201 regulators as well as protein kinases (K) and protein phosphatases (P) for reversible
1202 protein phosphorylation were differentially expressed. For detailed discussion see
1203 text.

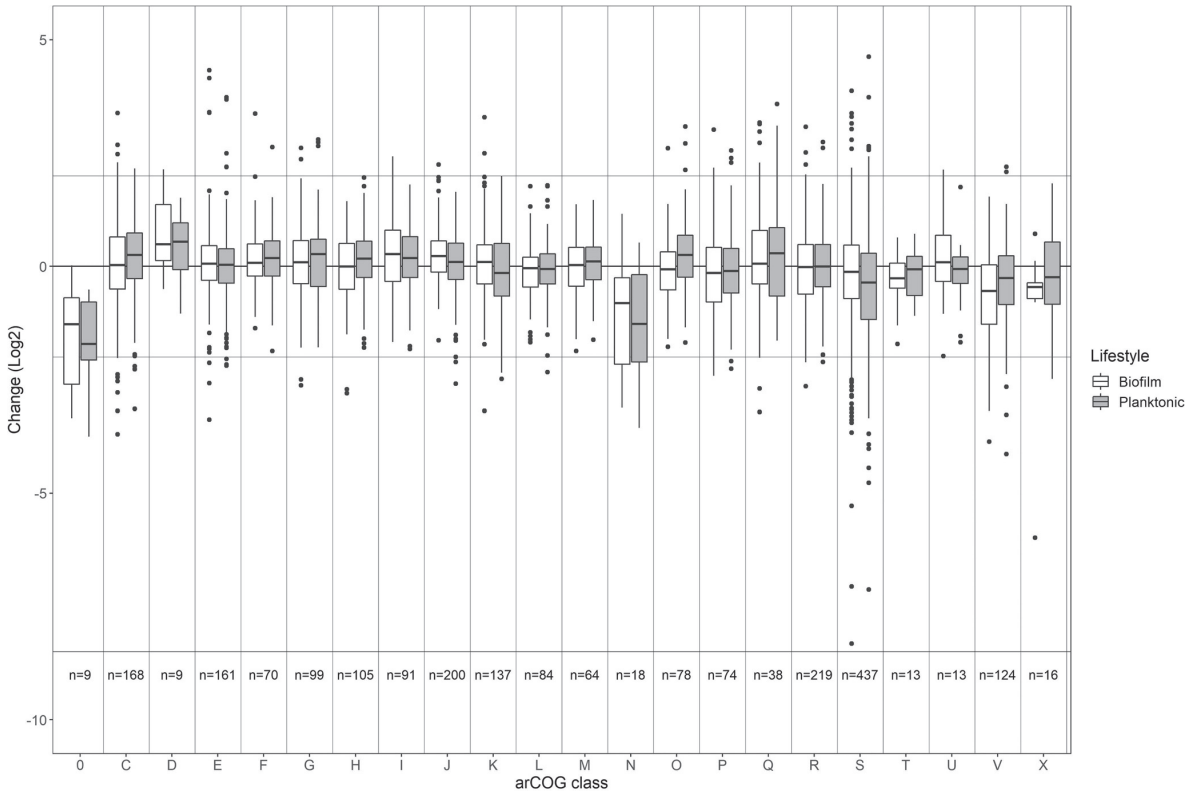


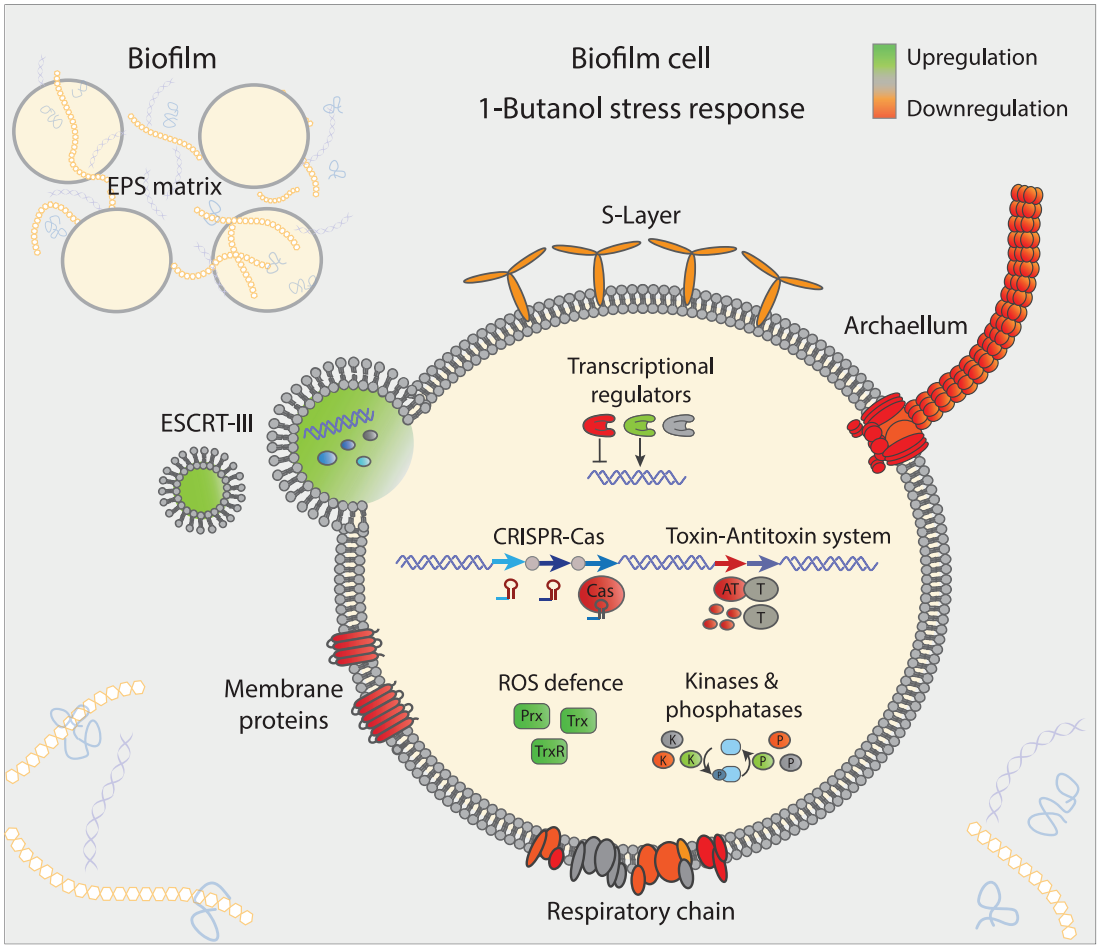












Supplementary Information

Table S1: Differentially regulated genes in *S. acidocaldarius* biofilm and planktonic cells grown statically in the presence and absence of 1% (v/v) 1-butanol (Petri dishes, 4d, 76 °C). The genes are ordered according to their involvement in cellular structures or processes. The effect of 1-butanol on gene expression in biofilm cells (BF1/BF0), planktonic cells (PL1/PL0) and the change in response to the respective lifestyle without 1-butanol exposure (BF0/PL0) is depicted. n. s.: not significantly regulated (log2 fold change < 1).

Locus	arCOG annotation	arCOG code	Biofilm (BF1/BF0) log2 fold change	Planktonic (PL1/PL0) log2 fold change	Lifestyle (BF0/PL0) log2 fold change
Cell surface structures					
S-Layer					
Saci_2354	S-layer protein SlaB	M	-1.58	n. s.	n. s.
Saci_2355	S-layer protein SlaA	M	-1.87	n. s.	n. s.
Saci_1846	Thermopsin-like protease	E	-1.90	-1.47	n. s.
Pili					
UV induced pili					
Saci_1493	Predicted component of type IV pili like system	N	n. s.	-1.36	n. s.
Saci_1494	ATPase involved in archaellum/pili biosynthesis	N	n. s.	n. s.	n. s.
Saci_1495	Pilus assembly protein TadC	N	n. s.	-1.31	n. s.
Adhesive pili					
Saci_2317	ATPase involved in archaellum/pili biosynthesis	N	n. s.	n. s.	n. s.
Saci_2318	Pilus assembly protein TadC	N	n. s.	n. s.	n. s.
Saci_2319	Pilin/Flagellin, FlaG/FlaF family	N	1.16	n. s.	-1.94
Archaellum					
Saci_1172	Archaellum assembly protein J, TadC family	N	-1.41	-2.16	n. s.
Saci_1173	ATPase involved in archaellum/pili biosynthesis	N	-2.32	-1.98	n. s.
Saci_1174	ATPase involved in biogenesis	N	-3.12	-1.91	-1.03

Locus	arCOG annotation	arCOG code	Biofilm (BF1/BF0) log2 fold change	Planktonic (PL1/PL0) log2 fold change	Lifestyle (BF0/PL0) log2 fold change
	of archaellum				
Saci_1175	Archaellum protein F, archaellin of FlaG/FlaF family	N	-1.71	-2.90	n. s.
Saci_1176	Archaellum protein G, archaellin of FlaG/FlaF family	N	-2.71	-3.42	n. s.
Saci_1177	Component of archaellum, FlaD/E family	N	-2.33	-2.80	n. s.
Saci_1178	Archaeal flagellins	N	-2.79	-3.56	n. s.
Vesicle/ESCRT system and Biofilm formation					
Vesicle/ESCRT system					
Saci_0451	Archaeal division protein CdvB1, Snf7/Vps24/ESCRT-III family	D	2.14	1.51	-1.51
Saci_1372	Cell division ATPase of the AAA+ class, ESCRT system component, CdvC	D	1.36	n. s.	n. s.
Saci_1373	Archaeal division protein CdvB, Snf7/Vps24/ESCRT-III family	D	1.26	n. s.	-1.19
Saci_1374	Archaeal division protein CdvA	D	n. s.	n. s.	n. s.
Saci_1416	Archaeal division protein CdvB2, Snf7/Vps24/ESCRT-III family	D	1.41	n. s.	-1.82
Saci_1601	Archaeal division protein CdvB3, Snf7/Vps24/ESCRT-III family	D	n. s.	n. s.	n. s.
Glycosyltransferases					
Saci_1914	Glycosyltransferase	M	-1.87	-1.21	n. s.
Saci_1915	Glycosyl transferase family 2	M	-1.15	-1.06	n. s.
Saci_1923	Glycosyltransferase	M	-1.05	n. s.	n. s.
Saci_1926	Glycosyl transferase family 2	M	n. s.	n. s.	-1.90
Regulation and stress response					
Transcriptional regulators					
Saci_0102	Transcriptional regulator,	K	-1.05	n. s.	n. s.

Locus	arCOG annotation	arCOG code	Biofilm (BF1/BF0) log2 fold change	Planktonic (PL1/PL0) log2 fold change	Lifestyle (BF0/PL0) log2 fold change
	contains HTH domain				
Saci_0446	Transcriptional regulator, contains HTH domain (AbfR1)	K	1.71	n. s.	n. s.
Saci_0665	Homolog of transcription initiation factor TFIIB, contains Zn-ribbon domain	K	n. s.	-1.41	-1.03
Saci_1171	Predicted transcriptional regulator (ArnR1)	K	-1.62	-2.01	n. s.
Saci_1209	Transcriptional regulator, contains HTH domain	K	1.17	n. s.	n. s.
Saci_1223	Transcriptional regulator, contains HTH domain (AbfR2)	K	-1.54	n. s.	n. s.
Saci_1588	DNA-binding transcriptional regulator, Lrp family	K	n. s.	1.73	n. s.
Saci_1992	CRISPR-Cas associated transcriptional regulator, contains HTH domain, lacking CARF domain	K	-3.19	-2.35	1.01
Protein phosphorylation: protein kinases and phosphatases					
Saci_0545	Protein-tyrosine phosphatase (PTP)	T	-1.30	n. s.	n. s.
Saci_0796	RIO-like serine/threonine protein kinase fused to N-terminal HTH domain (RIO2)	T	n. s.	-1.09	n. s.
Saci_1181	Membrane associated serine/threonine protein kinase (ArnS)	R	-1.19	-1.77	n. s.
Saci_1193	Membrane associated serine/threonine protein kinase (ArnC)	R	1.22	n. s.	-1.65

Locus	arCOG annotation		arCOG code	Biofilm (BF1/BF0) log2 fold change	Planktonic (PL1/PL0) log2 fold change	Lifestyle (BF0/PL0) log2 fold change
CRISPR-Cas						
Type I-D system						
Saci_1864	CRISPR-Cas system related protein, RAMP superfamily Cas6 group	cas6	V	n. s.	n. s.	n. s.
Saci_1872	CRISPR-Cas system related helicase, Cas3 (C-terminal HD nuclease domain)	cas3	V	n. s.	n. s.	n. s.
Saci_1873	CRISPR associated protein, RAMP family Cas5 group	csc1gr5	V	n. s.	n. s.	n. s.
Saci_1874	CRISPR-Cas system related protein, RAMP superfamily Cas7 group	csc2gr7	V	n. s.	n. s.	n. s.
Saci_1875	CRISPR associated protein Cas10d, large subunit of Type I-D system effector complex, contains HD family nuclease	cas10d	V	n. s.	n. s.	n. s.
Saci_1876	CRISPR-Cas associated transcriptional regulator, contains CARF and HTH domain	casR	VK	-1.23	n. s.	n. s.
Saci_1877	CRISPR-Cas system related protein, RAMP superfamily Cas6 group	cas6	V	-1.56	n. s.	n. s.
Saci_1879	CRISPR-associated protein Cas2	cas2	V	n. s.	n. s.	n. s.
Saci_1880	CRISPR-associated protein Cas4	cas4	V	n. s.	n. s.	n. s.
Saci_1881	CRISPR-associated protein Cas1	cas1	V	-1.54	n. s.	n. s.

Locus	arCOG annotation		arCOG code	Biofilm (BF1/BF0) log2 fold change	Planktonic (PL1/PL0) log2 fold change	Lifestyle (BF0/PL0) log2 fold change
***Sulfolobus* specific Type III system**						
Saci_1893	CRISPR-Cas system related protein, RAMP superfamily Cas7 group	csm3 gr7	V	n. s.	n. s.	n. s.
Saci_1896	CRISPR-Cas system related protein, RAMP superfamily Cas7 group	csm3 gr7	V	-1.14	n. s.	n. s.
Saci_1897	CRISPR associated protein, possible subunit of Type III-A effector complex	csx26	V	-2.07	n. s.	1.21
Saci_1898	CRISPR associated protein, Csm4g5-like subunit of effector complex	csm4 gr5	V	-2.12	n. s.	1.37
Saci_1899	CRISPR associated protein, Cas10-like subunit Type III-A effector complex	cas10	V	-2.60	-1.14	1.02
Adaptation/processing module						
Saci_2008	CRISPR-Cas system related protein, RAMP superfamily Cas6 group	cas6	V	-3.02	-1.44	n. s.
Saci_2010	CRISPR-associated protein Cas2	cas2	V	-1.83	n. s.	n. s.
Saci_2011	CRISPR-associated protein Cas1	cas1	V	-1.40	n. s.	n. s.
Saci_2012	CRISPR-associated protein Cas4	cas4	V	n. s.	n. s.	n. s.
Type III-D system						
Saci_2043	CRISPR-Cas system related protein, RAMP superfamily Cas7 group	csm3 gr7	V	n. s.	n. s.	n. s.
Saci_2044	CRISPR-associated	csx10	V	n. s.	n. s.	n. s.

Locus	arCOG annotation		arCOG code	Biofilm (BF1/BF0) log2 fold change	Planktonic (PL1/PL0) log2 fold change	Lifestyle (BF0/PL0) log2 fold change
	protein, RAMP family Cas5 group, signature protein for Type III-D system	gr5				
Saci_2045	CRISPR-Cas system related protein, RAMP superfamily Cas7 group	csm3 gr7	V	n. s.	n. s.	-1.09
Saci_2046	CRISPR associated protein, Cas10-like subunit Type III-A effector complex	cas10	V	-2.42	n. s.	n. s.
Saci_2048	CRISPR-Cas system related protein, RAMP superfamily Cas7 group	csm3 gr7	V	n. s.	n. s.	n. s.
Saci_2049	CRISPR-Cas system related protein, RAMP superfamily Cas7 group	csm3 gr7	V	-2.02	n. s.	n. s.
Saci_2052	CRISPR-associated protein	csm2 gr11	V	1.22	1.03	-1.63

Locus	arCOG annotation	arCOG code	Biofilm (BF1/BF0) log2 fold change	Planktonic (PL1/PL0) log2 fold change	Lifestyle (BF0/PL0) log2 fold change
Toxin-Antitoxin					
Saci_0264	Transcriptional regulator, CopG/Arc/MetJ family (DNA-binding and a metal-binding domains)	V	1.02	-1.36	n. s.
Saci_0322	CopG/RHH family DNA binding protein, antitoxin	V	1.02	n. s.	n. s.
Saci_0942	CopG/MetJ, RHH domain containing DNA-binding protein, often an antitoxin in Type II toxin-antitoxin systems	V	n. s.	-1.71	n. s.
Saci_1056	Antitoxin, CopJ/RHH family	V	-2.03	-4.14	n. s.
Saci_1124	CopG/RHH family DNA binding protein	V	-2.03	n. s.	n. s.
Saci_1812	RHH/CopG DNA binding protein	V	-1.44	-1.69	n. s.
Saci_1928	Minimal nucleotide transferase MNT, antitoxin of HEPN-MNT system	V	1.28	1.32	
Saci_1932	RHH/copG family antitoxin	V	-1.57	-1.36	n. s.
Saci_1936	RHH/CopG DNA binding protein	V	-1.53	-1.38	n. s.
Saci_1947	RHH/CopG DNA binding protein	V	n. s.	-1.11	n. s.
Saci_1952	CopG/RHH family DNA binding protein, antitoxin	V	n. s.	-1.40	n. s.
Saci_1980	RHH/CopG DNA binding protein	V	n. s.	-1.34	n. s.
Saci_2003	CopG/RHH family DNA binding protein, antitoxin	V	n. s.	n. s.	n. s.
Saci_2079	RHH/CopG DNA binding protein	V	n. s.	2.19	1.23
Metabolism					
Amino acid metabolism					
Pyroglutamate conversion					

Locus	arCOG annotation	arCOG code	Biofilm (BF1/BF0) log2 fold change	Planktonic (PL1/PL0) log2 fold change	Lifestyle (BF0/PL0) log2 fold change
Saci_2041	N-methylhydantoinase A/5-oxoprolinase, beta subunit	E	4.33	3.73	-1.28
Saci_2042	N-methylhydantoinase B/5-oxoprolinase, alpha subunit	E	4.15	3.68	-1.16
Saci_2036	N-methylhydantoinase A/5-oxoprolinase, alpha subunit	E	3.40	2.19	-2.30
Aromatic compound/amino acid conversion					
Saci_2293	2-keto-4-pentenoate hydratase/2-oxohepta-3-ene-1,7-dioic acid hydratase (catechol pathway)	Q	3.17	2.70	n. s.
Saci_2294	4-hydroxyphenylacetate 3-monooxygenase	Q	3.14	2.41	n. s.
Saci_2295	Catechol 2,3-dioxygenase or other lactoylglutathione lyase family enzyme	E	3.39	2.49	n. s.
Antioxidance defence					
Saci_1125	peroxiredoxin	O	1.20	2.13	n. s.
Saci_1169	thioredoxin reductase	O	2.61	3.08	n. s.
Saci_1823	thioredoxin	O	1.37	2.71	n. s.
Respiratory chain					
Cytochrome bc1 complex (SoxNL-CbsAB-OdsN)					
Saci_1859	Cytochrome b558/566, subunit B	C	-3.70	-3.14	1.10
Saci_1860	Rieske Fe-S protein	C	-1.66	n. s.	n. s.
Saci_1861	Cytochrome b subunit of the bc complex	C	-1.94	-1.25	n. s.
Saci_1862	Heme-degrading monooxygenase HmoA and related ABM domain proteins	H	-1.50	n. s.	n. s.
Terminal oxidase SoxABCDL complex (Saci_2086-2089): Not significantly regulated					
Terminal oxidase SoxEFGHIM complex					
Saci_2258	Predicted subunit of heme/copper-type	C	n. s.	-1.36	n. s.

Locus	arCOG annotation	arCOG code	Biofilm (BF1/BF0) log2 fold change	Planktonic (PL1/PL0) log2 fold change	Lifestyle (BF0/PL0) log2 fold change
	cytochrome/quinol oxidase				
Saci_2259	Heme/copper-type cytochrome/quinol oxidase, subunit 2	C	1.95	n. s.	-1.16
Saci_2260	Cytochrome b subunit of the bc complex	C	n. s.	-2.20	-1.22
Saci_2261	Rieske Fe-S protein	C	-1.13	-2.27	n. s.
Saci_2262	Sulfocyanin	C	n. s.	-1.98	-1.41
Saci_2263	Heme/copper-type cytochrome/quinol oxidase, subunit 1 and 3	C	n. s.	-1.53	n. s.
Terminal oxidase DoxBCE complex					
Saci_0097	Heme/copper-type cytochrome/quinol oxidase, subunit 1	C	-2.45	-1.94	n. s.
Saci_0098	Terminal oxidase, subunit doxC	C	-2.79	-1.69	n. s.
Saci_0099	Terminal oxidase, subunit doxE	C	-3.19	-1.24	n. s.

Table S2: Highly downregulated genes encoding for membrane proteins in static grown *S. acidocaldarius* biofilm cells in response to 1-butanol (1% (v/v) exposure (static cultivation in Petri dishes, 4d, 76 °C).

Locus	arCOG annotation	arCOG functional code	Regulation (BF1/BF0)		Regulation (PL1/PL0)	
			log2 fold change	A-value	log2 fold change	A-value
Saci_0301	uncharacterized membrane protein, DUF981 family	S	-8.32	14.11	-7.13	12.82
Saci_1074	uncharacterized membrane protein	S	-7.06	13.43	-3.93	12.80
Saci_1753	uncharacterized membrane protein, virus associated	X	-5.98	11.62	-2.48	11.76
Saci_0516	uncharacterized protein	S	-5.28	3.06	-2.71	3.69

Table S3: Significantly regulated proteins in static grown *S. acidocaldarius* biofilm cells in response to 1-butanol (1% (v/v) exposure (static cultivation in Petri dishes, 4d, 76 °C).

Locus	**arCOG annotation**	**arCOG functional code**	**log2 (BF1/BF0)**
Saci_0642	Ribosomal protein L37E	J	1.72
Saci_0843	Transcriptional regulator, contains N-terminal RHH domain	K	1.47
Saci_0855	Zn-ribbon protein	S	1.19
Saci_0345	Lipoate-protein ligase A	H	1.16
Saci_0107	Molybdopterin-guanine dinucleotide biosynthesis protein	H	1.14
Saci_1468	DNA-binding TFAR19-related protein, PDSD5 family	R	1.11
Saci_0583	Ribosomal protein S14	J	1.10
Saci_0356	Uncharacterized small metal-binding protein	S	1.04
Saci_1079	Threonine dehydrogenase or related Zn-	E	1.04

Locus	arCOG annotation	arCOG functional code	log2 (BF1/BF0)
	dependent dehydrogenase		
Saci_0182	Prephenate dehydrogenase	E	1.03
Saci_1261	Threonyl-tRNA synthetase	J	1.02
Saci_2322	Cobalamin biosynthesis protein CbiG	H	-1.01
Saci_1208	Predicted dithiol-disulfide isomerase involved in polyketide biosynthesis	Q	-1.05
Saci_1366	Uncharacterized protein	S	-1.09
Saci_1764	ABC-type dipeptide/oligopeptide/nickel transport system, ATPase component	E	-1.10
Saci_2119	RecB family nuclease with coiled-coil N-terminal domain	R	-1.10
Saci_1306	Uridylate kinase	F	-1.14
Saci_1308	Short-chain alcohol dehydrogenase	I	-1.15
Saci_0319	Uncharacterized protein YjgD, DUF1641 family	S	-1.15
Saci_1168	Ser-tRNA(Ala) deacylase AlaX (editing enzyme)	J	-1.18
Saci_0845	Uncharacterized protein	S	-1.20
Saci_0820	Riboflavin synthase beta-chain	H	-1.20
Saci_0177	Single-stranded DNA-specific exonuclease RecJ	L	-1.22
Saci_1862	Heme-degrading monooxygenase HmoA and related ABM domain proteins	H	-1.24
Saci_1633	Enoyl-CoA hydratase/carnithine racemase	I	-1.27
Saci_0668	Uncharacterized protein	S	-1.27
Saci_0415	Zn-dependent protease with chaperone function	O	-1.30
Saci_0097	Heme/copper-type cytochrome/quinol oxidase, subunit 1	C	-1.35
Saci_1243	Uncharacterized protein	S	-1.36
Saci_2355	S-layer protein SlaA	M	-1.39
Saci_2332	Membrane protease subunit, stomatin/prohibitin homolog	O	-1.48

Locus	arCOG annotation	arCOG functional code	log2 (BF1/BF0)
Saci_2139	CBS domain containing protein	R	-1.54
Saci_1250	Glycosyl hydrolase family 15	G	-1.74
Saci_1860	Rieske Fe-S protein	C	-1.91

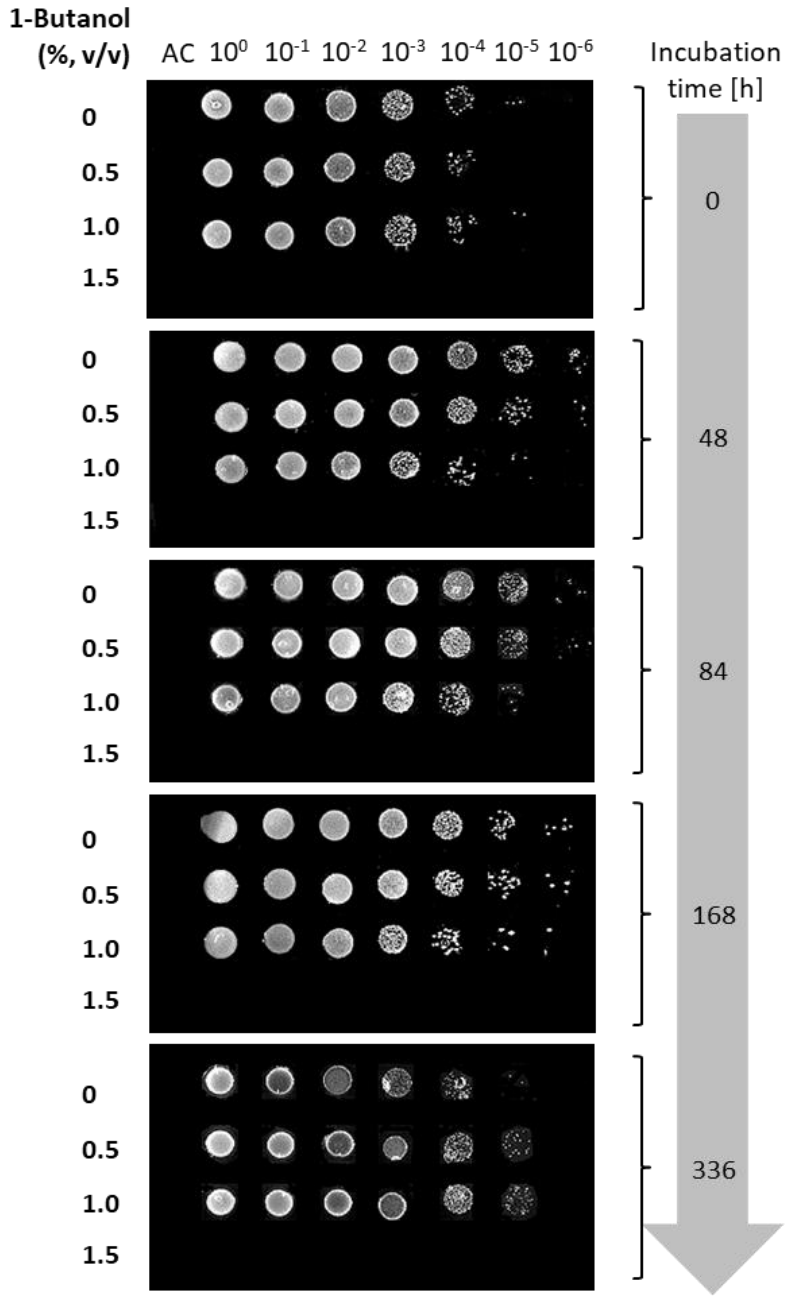


Figure S1: Culturability of *S. acidocaldarius* DSM 639 after 1-butanol exposure.

S. acidocaldarius DSM 639 liquid cultures were exposed to different concentrations of 1-butanol (0% to 1.5% (v/v)) in Brock medium supplemented with 0.1% (w/v) NZ-amine and 0.2% (w/v) D-glucose. After different cultivation times (0, 48, 84, 168 and 336 h), 10 µl of undiluted culture (10^0) or diluted culture (10^{-1}-10^{-6}) were spotted on Brock-Gelrite plates (0.1% (w/v) NZ-amine, 0.2% (w/v) D-glucose). Spot plates were incubated at 76 °C for four days. An abiotic control (AC, medium without cells) and the 10^0-10^{-6} 10-fold dilution series of *S. acidocaldarius* DSM 639 shaking cultures are shown.

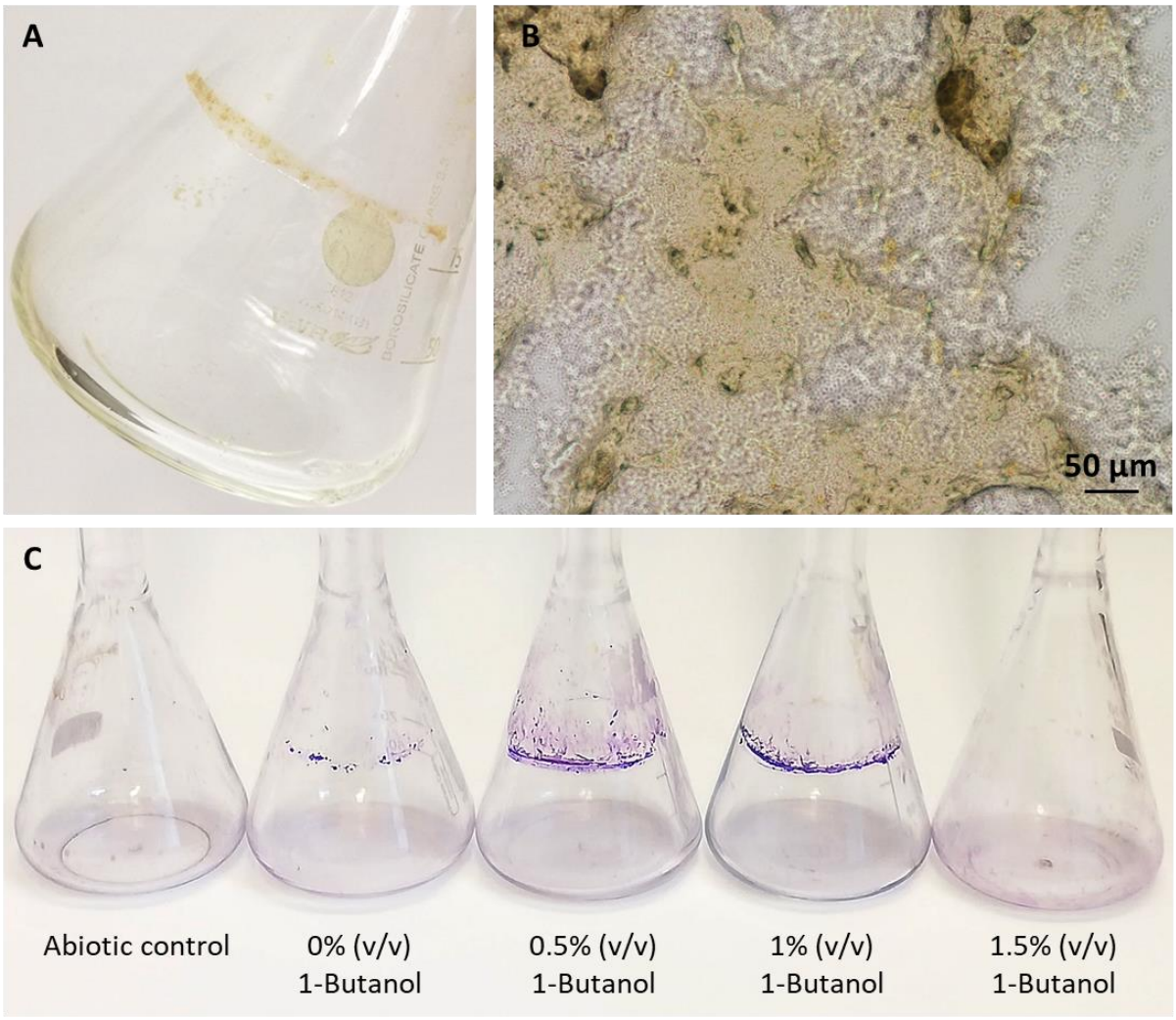


Figure S2: Adhesion of *S. acidocaldarius* DSM 639 cells after cultivation with 0.5% (v/v) and 1% (v/v) 1-butanol. Slimy material was observed at the liquid-air interfaces on the glass surface of planktonic *S. acidocaldarius* DSM 639 cultures exposed to 0.5% (v/v) and 1% (v/v) 1-butanol (Brock medium, 0.1% NZ-amine and 0.2% D-glucose). **A**. Collar of slimy material inside the Erlenmeyer flask of *S. acidocaldarius* DSM 639 culture exposed to 1% (v/v) 1-butanol. Culture fluid was discarded. The visible material was scrabbed off the glass surface using a cell scraper, applied on a cavity slide and used for light microscopy (**B**). Large aggregates of organic material surrounding *S. acidocaldarius* cells were visible. **C**. After three weeks of cultivation biofilm formation of planktonic *S. acidocaldarius* DSM 639 cultures exposed to 0% to 1.5% (v/v) 1-butanol was visualized by crystal violet staining. For biofilm visualization cultures fluid was discarded, the empty Erlenmeyer flasks were stained with 0.01% (w/v) crystal violet solution and washed with water. Experiments were carried out in three to four biological replicates.

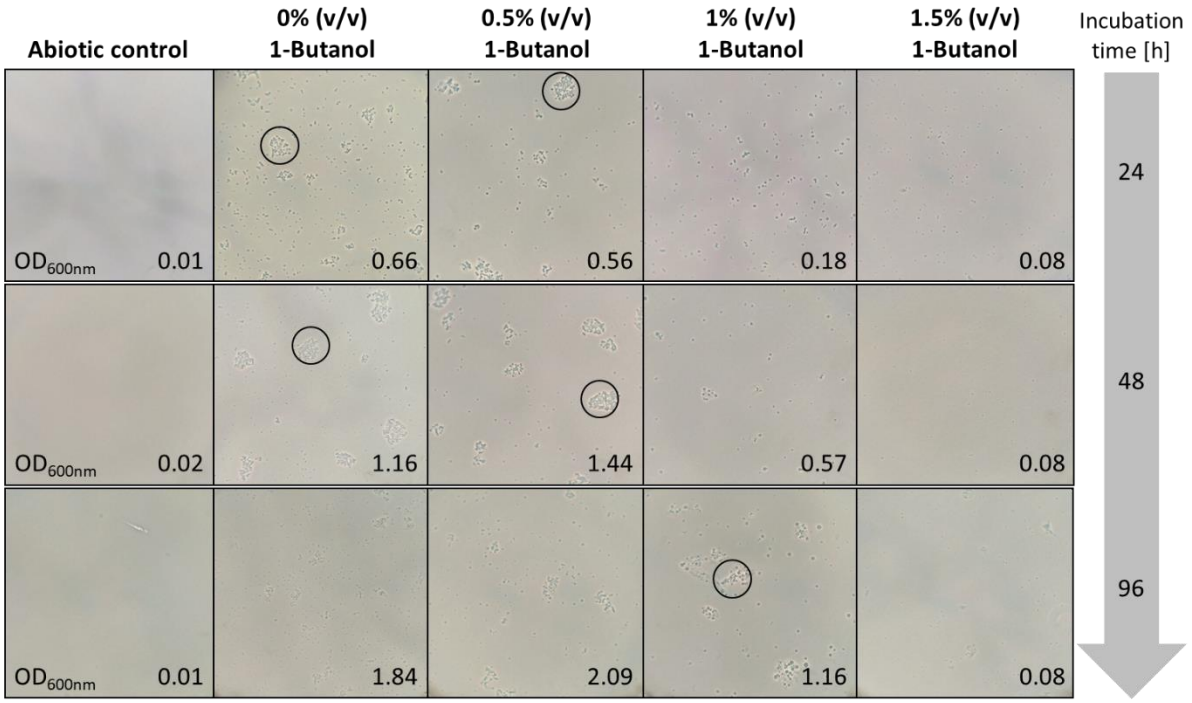


Figure S3: Cell aggregation analysis of *S. acidocaldarius* DSM 639 after 1-butanol exposure. Phase-contrast microscopy images of *S. acidocaldarius* DSM 639 shaking cultures exposed to different concentrations of 1-butanol (0% to 1.5% (v/v)) in Brock medium supplemented with 0.1% NZ-amine and 0.2% D-glucose. Circles mark examples of cell aggregates. OD_{600nm}: optical density at 600 nm

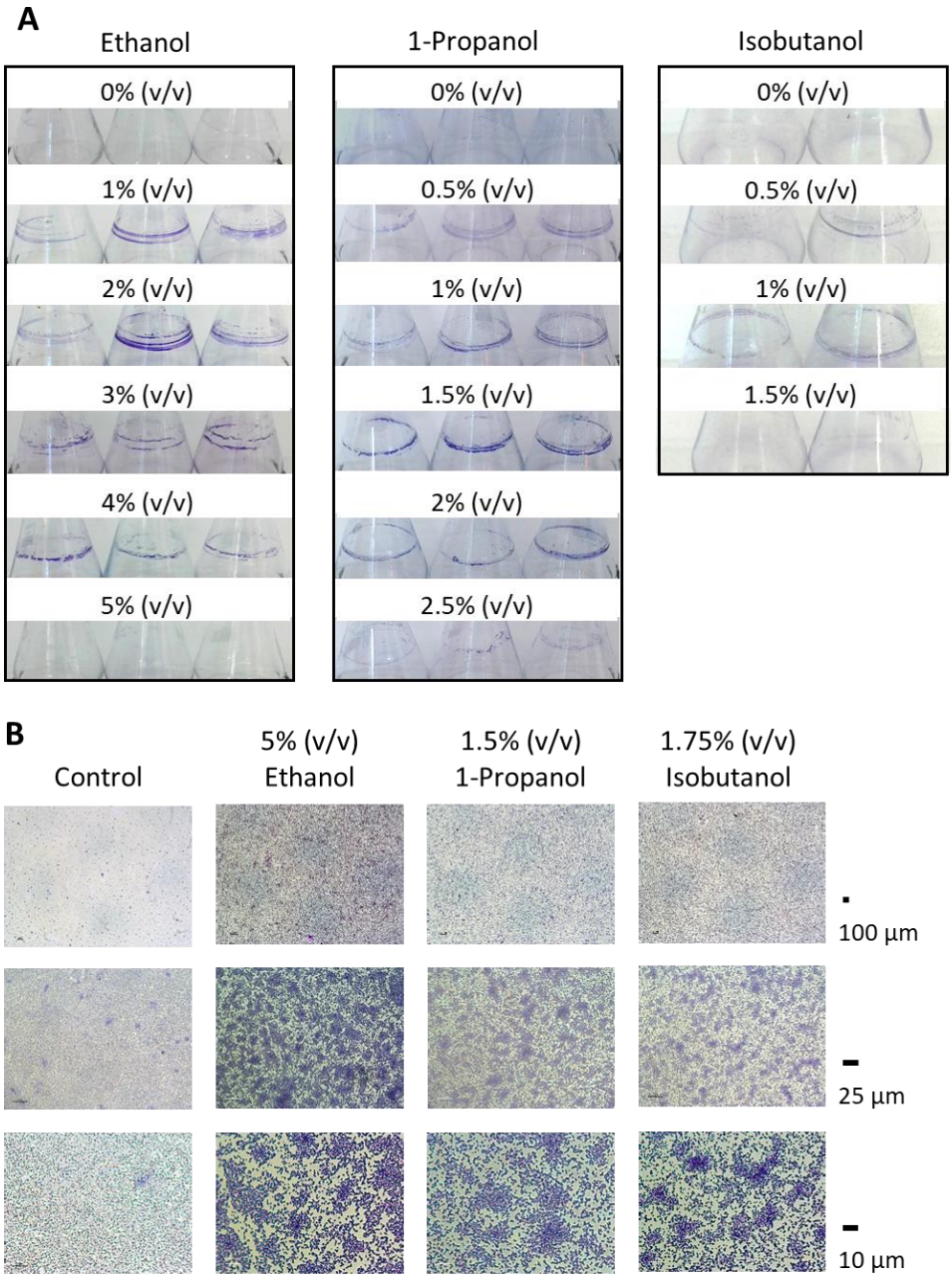


Figure S4: Effect of organic solvents on *S. acidocaldarius* cell adhesion and cell distribution. **A**. Cell adhesion of *S. acidocaldarius* DSM 639 after cultivation with different concentrations of ethanol, 1-propanol and isobutanol. Biofilms were visualized using crystal violet staining. The presence of multiple “collars” of the slimy material inside the flasks was presumably caused by medium loss due to sampling and medium evaporation, resulting in slightly decreasing culture volumes inside the flasks during the three weeks of the growth experiments. **B**. Effect of ethanol, 1-propanol and isobutanol exposure on *S. acidocaldarius* cell distribution. *S. acidocaldarius* was grown on glass surfaces for 4 d at 76 °C in presence and absence of different organic solvents. Biofilms were stained by crystal violet for subsequent analysis by light microscopy.

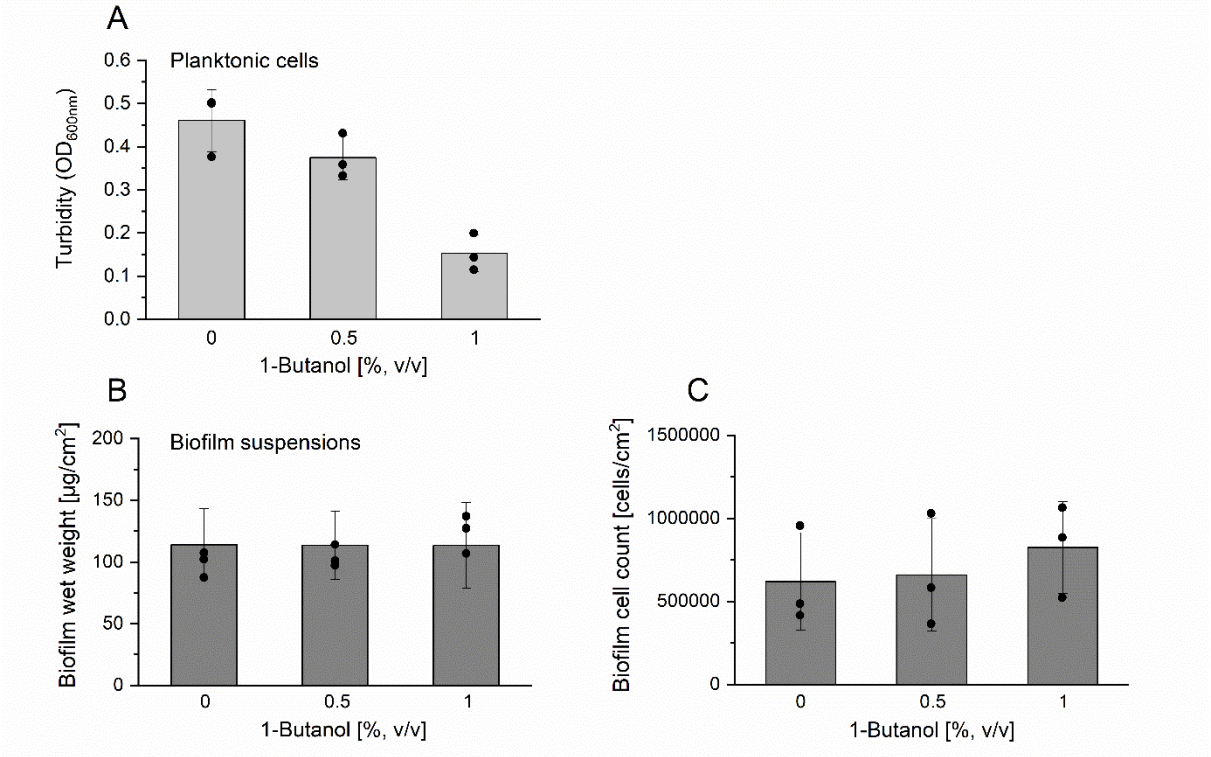


Figure S5: Influence of 1-butanol on biofilm formation. The amounts of planktonic (**A**) and biofilm cells of *S. acidocaldarius* (**B, C**) grown statically in Petri dishes were determined. Cultures were incubated at 76 °C for four days. **A**. Growth of planktonic cells was determined by turbidity measurements (OD_{600nm}; n = 3). **B**. Biofilm wet weight. The biomass was isolated from the bottom of Petri dishes for each condition and pooled biofilm samples were weighed (n = 3). **C**. Total cell counts of biofilm suspensions. Cell count was determined using the DAPI staining method (n = 3).

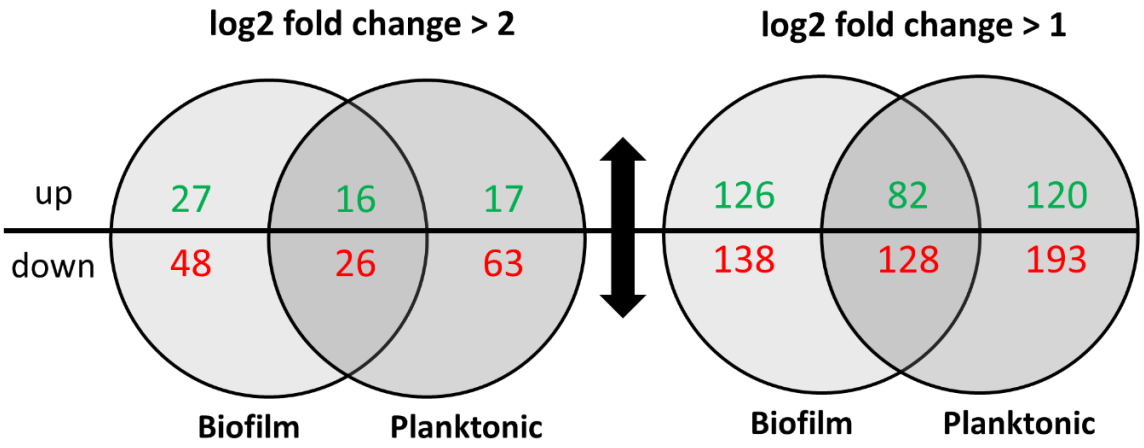


Figure S6: **Venn diagram displaying the overlap of significantly regulated genes in response to 1% (v/v) 1-butanol in biofilm and planktonic cells**. The numbers of log2 fold change > 2 and log2 fold change > 1 upregulated (green) and downregulated (red) genes are given. Intersections present the numbers of genes that are commonly regulated in both lifestyles.

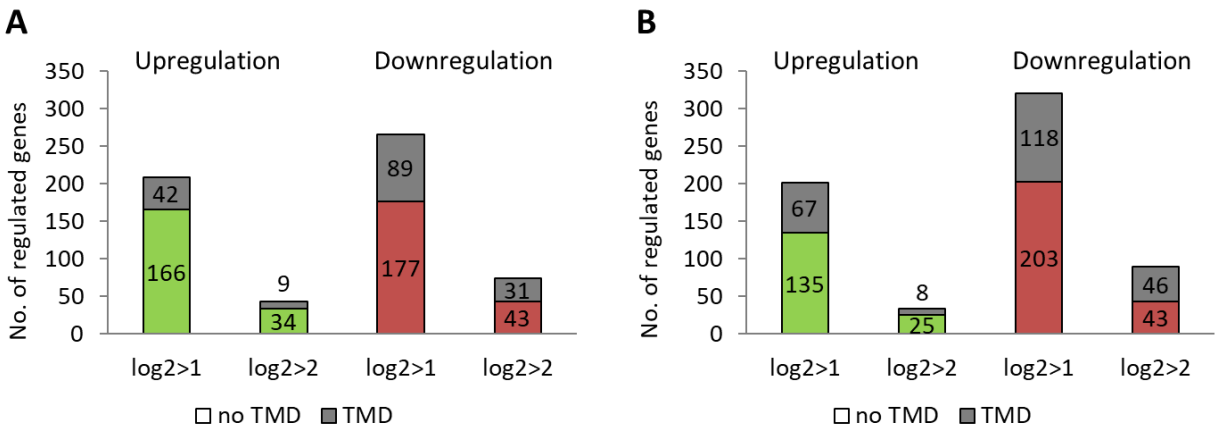


Figure S7: **Number of regulated genes encoding proteins without or with trans-membrane domains (TMD) in biofilm (A) and planktonic lifestyle (B).** The absolute numbers of >log2 fold change = 1 and >log2 fold change = 2 up- or downregulated genes are given for each lifestyle (green or red, respectively).

4 Summary

Archaea were previously classified as the third domain of life and compared to Bacteria and Eukarya, Archaea are characterized by their unique membrane lipid composition which is comprised of isoprenoid side chains ether-linked to a glycerol-1-phosphate moiety and substitutes the fatty acid based membrane structures that are ester connected to a glycerol-3-phosphate backbone in Bacteria and Eukaryotes. However, the recently proposed two domain tree of life indicated that Eukarya might have evolved from the archaeal Asgard superphylum and thus membrane lipids must have fundamentally changed from the archaeal type into the bacterial/eukaryotic one. This evolutionary differentiation of membrane structures between Archaea and Bacteria is regarded as the “lipid divide“ and raised up questions regarding the existence and the function of fatty acids in Archaea. Therefore, in course of this work the fatty acid metabolism in the crenarchaeal thermoacidophilic model organism *Sulfolobus acidocaldarius* was analysed.

In chapter 3.1, a TetR family transcription repressor (encoded by *saci_1107*) was characterized and found to regulate its own expression as well as the expression of the gene cluster *saci_1103-saci_1126* in *S. acidocaldarius.* Derepression occurred due to the binding of acyl-CoAs to the protein, upon which expression of *saci_1103-saci_1126* was increased. To further confirm the repression function in fatty acid degradation, the parental strain MW001 as well as the regulator knockout strain were shown to uptake short/medium chain fatty acids such as butyrate or hexanoate, as sole carbon and energy sources. Here, the regulator mutant displayed a shorter doubling time than MW001 when growing on hexanoate. The crystal structure of this regulator was determined and displayed a similar “omega“-shaped dimeric structure in comparison to its bacterial homologue. However, the *S. acidocaldarius* TetR possessed a different ligand binding specificity to acyl-CoAs. Furthermore, four DNA-binding sites of the regulator were found in the fatty acid related gene cluster *saci_1103-saci_1126*. Thus, the data illustrated the existence and function of a distinct archaeal type TetR transcription factor.

In chapter 3.2., the FA metabolism in *S. acidocaldarius* was studied by analysing β oxidation homologues encoded in the gene cluster (*saci_1103-1126*) identified in chapter 3.1. An AMP-forming acyl-CoA synthetase for fatty acid activation and a full set of β oxidation enzymes including a FAD dependent acyl-CoA dehydrogenase, a bifunctional 3(S)-hydroxyacyl-CoA dehydrogenase/enoyl-CoA hydratase (with an inverted domain structure/organization compared to well characterized bacterial homologues, e.g. from *E. coli*,) as well as an archaeal β-ketothiolase were characterized in detail. From the characterized single enzymes a fully functional β oxidation spiral was reconstituted *in vitro* and the complete oxidation of FAs to acetyl-CoA up to a chain length of C8 was demonstrated. The data further indicated that the β

oxidation is not fully reversible and does thus very likely not account for FA synthesis in Archaea. Instead, the potential of *S. acidocaldarius* to synthesize FAs via a novel CoA dependent pathway acting independently from β oxidation enzymes is shown. The pathway comprises a bacterial like SDR superfamily R-specific fabG homolog, an MDR superfamily enoyl-CoA reductase like in some Eukaryotes, both with a clear preference for NADPH as electron donor. Furthermore, an R-specific MaoC like dehydratase was identified. The enzymes were biochemically characterized and catalysed the synthesis of medium chain FA-CoA esters up to chain lengths of C8. These results provide a basic understanding of the FA metabolism in Archaea and thus pave the way for the further understanding of the presence and significance of FA in Archaea and its evolutionary implications.

S. acidocaldarius is not only ant model organism to study archaeal biology. With the combination of its broad substrate specificity, missing catabolite repression, polymer degrading capabilities and metabolic versatility, *S. acidocaldarius* is a promising candidate for biotechnological applications like e.g. the production of value-added products from waste materials including biofuels and base chemicals. To further elucidate this biotechnological potential, in chapter 3.3, the response of *S. acidocaldarius* to the widely used organic solvent 1-butanol was analyzed and then enhancement of biofilm formation was observed. Confocal laser scanning microscopy revealed the formation of a denser and higher biofilm with increased amounts of extracellular carbohydrates when exposed to 1% (v/v) 1-butanol. The transcriptomic and proteomic studies revealed that distinct regulations occurred in motility, cell envelope and membrane composition, cell division, vesicle formation, immune and defense systems as well as metabolism and general stress response. As a result, the extremophilic *S. acidocaldarius* displayed a high tolerance to solvent stress, exemplified by butanol exposure suggesting a great potential for industrial application.

5 Zusammenfassung

Archaeen wurden zusätzlich zu den Bakterien und Eukaryoten als dritte Domäne des Lebens identifiziert und klassifiziert. Neben anderen Unterschieden ist ein Hauptdifferenzierungsmerkmal der Archaea, dass sie im Vergleich zu Bakterien und Eukaryoten einzigartige Membranlipide besitzen. Diese sind bei Archaea aus einer Glycerin-1-phosphat-Einheit mit ether-gebundenen Isoprenoid-Seitenketten aufgebaut, wohingegen die in Bakterien und Eukaryoten vorhandenen Membranlipide aus Glycerin-3-phosphat und ester-gebundenen Fettsäuren bestehen. Neuere phylogenetische Studien haben nun zu einem revidierten Zwei-Domänen Modell des universellen phylogenetischen Stammbaumes der Organismen geführt. Danach stellen Bacteria und Archaea die beiden ursprünglichen evolutiven Hauptlinien dar, wohingegen die Eukaryoten später aus dem archaealen Asgard Superphylum hervorgegangen sind. Dieses Modell setzt jedoch einen fundamentalen Wechsel in der Membranbeschaffenheit während der Eukaryoten-Evolution vom archaealen zum bakteriellen/eukaryotischen Typus voraus. Diese evolutive Differenzierung der Membranen von Archaeen und Bakterien wurde als „Lipid Divide“ bezeichnet und wirft die Frage nach der Existenz und der Funktion von Fettsäuren in Archaeen auf. Daher wurde im Rahmen dieser Arbeit der Fettsäuremetabolismus im crenarchaealen thermoacidophilen Modellorganismus *Sulfolobus acidocaldarius* untersucht.

In Kapitel 3.1 wurde ein Transkriptionsrepressor der TetR-Familie (kodiert durch *saci_1107*) charakterisiert, der neben seiner eigenen Expression auch die Expression des Genclusters *saci_1103-saci_1126* in *S. acidocaldarius* reprimiert. Eine Aufhebung der Reprimierung und damit eine Induktion der Genexpression trat durch die Bindung von Acyl-CoAs auf. Um diese Repressor-Funktion des TetR Regulators im Fettsäureabbau weiter zu bestätigen, wurde der *S. acidocaldarius* MW001 Wildtyp-Stamm und der Regulator-Knockout-Stamm mit den Fettsäuren Butyrat und Hexanoat als einziger Kohlenstoff- und Energiequelle kultiviert. Dabei zeigte die TetR-Mutante eine kürzere Verdopplungszeit im Vergleich zum MW001 Wildtyp. Die Kristallstruktur des Regulators wurde aufgeklärt und zeigte eine "Omega"-förmige Dimerstruktur, die der Struktur des bakteriellen Homologs ähnelt. Das *S. acidocaldarius* Homolog besaß jedoch Bindungsspezifität zu verschiedenen Acyl-CoAs. Weiterhin wurden vier DNA-Bindestellen des Regulators in dem o.e. Cluster von Genen *(saci_1103-saci_1126)* identifiziert, die vermutlich Fettsäuremetabolismus-assoziierte Proteine kodieren. Insgesamt klärt diese Studie die Funktion eines bisher unbekannten besonderen archaealen TetR-Transkriptionsfaktors auf.

In Kapitel 3.2 wurde der Fettsäuremetabolismus in *S. acidocaldarius* weiter untersucht, indem die Funktion von Homologen der β-Oxidation analysiert wurde, die in dem in Kaptiel 3.1. identifizierten Gencluster kodiert sind. Die Proteine wurden rekombinant hergestellt und

gereinigt. Eine AMP-bildende Acyl-CoA-Synthetase zur Fettsäureaktivierung sowie ein vollständiger Satz der an der β-Oxidation beteiligten Enzyme zum Fettsäureabbau wurde im Detail charakterisiert, einschließlich einer FAD-abhängigen Acyl-CoA-Dehydrogenase, einer bifunktionellen 3(S)-Hydroxyacyl-CoA-Dehydrogenase/Enoyl-CoA-Hydratase mit einer im Vergleich zu gut charakterisierten bakteriellen Homologen entgegengesetzten Domänenstruktur, und einer archaealen β-Ketothiolase. Aus den einzelnen Enzymen wurde dann eine vollständige *in-vitro* Enzymkaskade zur Oxidation kurz-/mittelkettiger Acyl-CoAs einschließlich Butyryl-CoA, Hexanoyl-CoA oder Octanoyl-CoA rekonstituiert und deren Funktionalität mittels HPLC nachgewiesen. Die Ergebnisse sprechen zudem dafür, dass die β-Oxidation als Ganzes nicht reversibel arbeiten kann und deswegen sehr wahrscheinlich nicht eine eventuelle Fettsäuresynthese in Archaea katalysiert. Stattdessen indizieren im Rahmen dieser Arbeit erzielte Ergebnisse, dass *S. acidocaldarius* das Potential hat, Fettsäuren über einen neuartigen CoA-abhängigen Stoffwechselweg zu synthetisieren, der komplett unabhängig von der β-Oxidation arbeiten kann und sich wesentlich von den bakteriellen und eukaryontischen Homologen unterscheidet. Bei den beteiligten Enzymen handelt es sich zum Einen um eine Acetoacyl (ketoacyl)-CoA-Reduktase der SDR-Familie, einem Homolog des bakteriellen fabG, und zum Anderen um eine Enoyl-CoA-Reduktase der MDR-Familie, die den eukaryotischen Kandidaten ähnlicher ist. Beide Enzyme wurden als NADPH- und CoA-Ester-abhängig charakterisiert. Weiterhin wurde eine MaoC-ähnliche 3(R)-Hydroxyacyl-CoA-Dehydratase identifiziert. Mit diesen Enzymen wurde *in vitro* eine Enzymkaskade rekonstituiert, die die Synthese von gesättigten Fettsäure-CoA Estern bis zu Kettenlängen von C8 katalysieren kann. Insgesamt tragen die hier erzielten Ergebnisse zum grundsätzlichen Verständnis des Fettsäuremetabolismus in Archaea bei und werden es weiter ermöglichen, das Vorhandensein und die Funktion von Fettsäuren in Archaea und deren Bedeutung für die zelluläre Evolution zu verstehen.

S. acidocaldarius ist nicht nur ein wichtiger Modellorganismus, um die Biologie der Archaea im Allgemeinen zu verstehen. Der Organismus ist mit seiner breiten Substratspezifität, fehlender Katabolitrepression, seinen polymer-abbauenden Eigenschaften und der metabolen Vielseitigkeit zunehmend auch ein Kandidat für biotechnologische Anwendungen wie beispielsweise für die Produktion von Biotreibstoffen oder Basischemikalien. Um dieses bíotechnologische Potential von *S. acidocaldarius* weiter zu untersuchen, wurde in Kapitel 3.3 die Stressantwort von *S. acidocaldarius* auf das vielseitig und vielfach eingesetzte organische Lösemittel 1-Butanol analysiert. Konfokale Laser-Scanning-Mikroskopie offenbarte die Bildung eines dichteren und höheren Biofilms mit einer erhöhten Menge an extrazellulären Kohlenhydraten bei einer Exposition in 1% (v/v) 1-Butanol. Transkriptom- und Proteom-Studien deckten veränderte Genexpressionen zur Regulation der Motilität, Zellhülle- und Membranzusammensetzung, Zellteilung, Vesikelbildung, Immun- und Abwehrsysteme sowie

Metabolismus und allgemeiner Stressreaktionen auf. Insgesamt wies der extremophile Organismus *S. acidocaldarius* eine hohe Toleranz gegenüber Lösemittel (Butanol), womit sich ein großes Potenzial für zukünftige industrielle Anwendungen ergibt.

6 References

1. Woese, C.R. and Fox, G.E., *Phylogenetic structure of the prokaryotic domain: The primary kingdoms.* Proc. Natl. Acad. Sci. USA, 1977. **74**(11): p. 5088-5090.
2. Grohmann, D. and Werner, F., *Recent advances in the understanding of archaeal transcription.* Curr Opin Microbiol, 2011. **14**(3): p. 328-34.
3. Bräsen, C., et al., *Carbohydrate metabolism in archaea: current insights into unusual enzymes and pathways and their regulation.* Microbiology and Molecular Biology Reviews, 2014. **78**(1): p. 89-175.
4. Abolfazl, J.S. and Majid, B.S., *Archaea an all-out study.* Eurasian Journal of Biological and Chemical Sciences, 2020.
5. Sleytr, U.B., et al., *S-layers: principles and applications.* FEMS Microbiol Rev, 2014. **38**(5): p. 823-64.
6. Zaremba-Niedzwiedzka, K., et al., *Asgard archaea illuminate the origin of eukaryotic cellular complexity.* Nature, 2017. **541**(7637): p. 353-358.
7. Spang, A., et al., *Complex archaea that bridge the gap between prokaryotes and eukaryotes.* Nature, 2015. **521**(7551): p. 173-179.
8. Caforio, A. and Driessen, A.J.M., *Archaeal phospholipids: Structural properties and biosynthesis.* Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids, 2017. **1862**(11): p. 1325-1339.
9. Chong, P.L., *Archaebacterial bipolar tetraether lipids: Physico-chemical and membrane properties.* Chem Phys Lipids, 2010. **163**(3): p. 253-65.
10. Albers, S.V., et al., *Adaptations of the archaeal cell membrane to heat stress.* Front Biosci, 2000. **5**: p. D813-20.
11. Tourte, M., et al., *Functionalized Membrane Domains: An Ancestral Feature of Archaea?* Front Microbiol, 2020. **11**: p. 526.
12. Guan, Z., et al., *Gene deletions leading to a reduction in the number of cyclopentane rings in Sulfolobus acidocaldarius tetraether lipids.* FEMS Microbiol Lett, 2018. **365**(1).
13. Siliakus, M.F., et al., *Adaptations of archaeal and bacterial membranes to variations in temperature, pH and pressure.* Extremophiles, 2017. **21**(4): p. 651-670.
14. Woese, C.R., et al., *Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya.* Proceedings of the National Academy of Sciences, 1990. **87**(12): p. 4576-4579.
15. De Long, E.F. and Pace, N.R., *Environmental diversity of bacteria and archaea.* Systematic Biology, 2001. **50**(4): p. 470-478.
16. Guy, L. and Ettema, T.J.G., *The archaeal 'TACK' superphylum and the origin of eukaryotes.* Trends in Microbiology, 2011. **19**(12): p. 580-587.
17. Eme, L., et al., *Archaea and the origin of eukaryotes.* Nature Reviews Microbiology, 2017. **15**: p. 711.
18. Dibrova, D.V., et al., *Phylogenomic reconstruction of archaeal fatty acid metabolism.* Environ Microbiol, 2014. **16**(4): p. 907-18.
19. Zweerink, S., et al., *Activity-based protein profiling as a robust method for enzyme identification and screening in extremophilic Archaea.* Nat Commun, 2017. **8**: p. 15352.
20. Lewis, A.M., et al., *The biology of thermoacidophilic archaea from the order Sulfolobales.* FEMS Microbiol Rev, 2021.
21. Sherwood, K.E., et al., *Glycerol-Mediated Repression of Glucose Metabolism and Glycerol Kinase as the Sole Route of Glycerol Catabolism in the Haloarchaeon Haloferax volcanii.* Journal of Bacteriology, 2009. **191**(13): p. 4307-4315.
22. Brock, T., et al., *Sulfolobus: A new genus of sulfur-oxidizing bacteria living at low pH and high temperature.* Arch Mikrobiol, 1972. **84**(1): p. 54-68.
23. Albers, S.V. and Siebers, B., *The family sulfolobaceae*, in *The Prokaryotes: Other Major Lineages of Bacteria and The Archaea*. 2014. p. 323-346.
24. Wagner, M., et al., *Versatile Genetic Tool Box for the Crenarchaeote Sulfolobus acidocaldarius.* Front Microbiol, 2012. **3**: p. 214.
25. Quehenberger, J., et al., *Sulfolobus - A Potential Key Organism in Future Biotechnology.* Front Microbiol, 2017. **8**: p. 2474.
26. Fujita, Y., et al., *Regulation of fatty acid metabolism in bacteria.* Mol Microbiol, 2007. **66**(4): p. 829-39.
27. Lombard, J., et al., *An ACP-independent fatty acid synthesis pathway in archaea: implications for the origin of phospholipids.* Mol Biol Evol, 2012. **29**(11): p. 3261-5.

28. Byers, D.M. and Gong, H., *Acyl carrier protein: structure–function relationships in a conserved multifunctional protein family.* Biochemistry and Cell Biology, 2007. **85**(6): p. 649-662.
29. Bhaumik, P., et al., *Structural biology of the thioester-dependent degradation and synthesis of fatty acids.* Current Opinion in Structural Biology, 2005. **15**(6): p. 621-628.
30. White, S.W., et al., *The structural biology of type II fatty acid biosynthesis.* Annual Review of Biochemistry, 2005. **74**(1): p. 791-831.
31. Massengo-Tiassé, R.P. and Cronan, J.E., *Diversity in Enoyl-Acyl Carrier Protein Reductases.* Cellular and molecular life sciences : CMLS, 2009. **66**(9): p. 1507.
32. Ghisla, S. and Thorpe, C., *Acyl-CoA dehydrogenases.* European Journal of Biochemistry, 2004. **271**(3): p. 494-508.
33. Watmough, N.J. and Frerman, F.E., *The electron transfer flavoprotein: Ubiquinone oxidoreductases.* Biochimica et Biophysica Acta (BBA) - Bioenergetics, 2010. **1797**(12): p. 1910-1916.
34. Dellomonaco, C., et al., *Engineered reversal of the β-oxidation cycle for the synthesis of fuels and chemicals.* Nature, 2011. **476**(7360): p. 355-359.
35. Flamholz, A., et al., *eQuilibrator--the biochemical thermodynamics calculator.* Nucleic Acids Res, 2012. **40**(Database issue): p. D770-5.
36. Hiltunen, J.K., et al., *Mitochondrial fatty acid synthesis-an adopted set of enzymes making a pathway of major importance for the cellular metabolism.* Prog Lipid Res, 2010. **49**(1): p. 27-45.
37. Sacco, E., et al., *The missing piece of the type II fatty acid synthase system from Mycobacterium tuberculosis.* PNAS, 2007. **104**(37): p. 14628-14633.
38. Pidugu, L.S., et al., *Analysis of proteins with the 'hot dog' fold: prediction of function and identification of catalytic residues of hypothetical proteins.* BMC Struct Biol, 2009. **9**: p. 37.
39. Moon, H.G., et al., *One hundred years of clostridial butanol fermentation.* FEMS Microbiol Lett, 2016. **363**(3).
40. García, V., et al., *Challenges in biobutanol production: How to improve the efficiency?* Renewable and Sustainable Energy Reviews, 2011. **15**(2): p. 964-980.
41. Durre, P., *Biobutanol: an attractive biofuel.* Biotechnol J, 2007. **2**(12): p. 1525-34.
42. Chen, C.T. and Liao, J.C., *Frontiers in microbial 1-butanol and isobutanol production.* FEMS Microbiol Lett, 2016. **363**(5): p. fnw020.
43. Mukherjee, M., et al., *Biobutanol as a potential alternative to petroleum fuel: Sustainable bioprocess and cost analysis.* Fuel, 2020. **278**.
44. Knoshaug, E.P. and Zhang, M., *Butanol tolerance in a selection of microorganisms.* Appl Biochem Biotechnol, 2009. **153**(1-3): p. 13-20.
45. Huffer, S., et al., *Role of alcohols in growth, lipid composition, and membrane fluidity of yeasts, bacteria, and archaea.* Appl Environ Microbiol, 2011. **77**(18): p. 6400-8.
46. Fischer, C.R., et al., *Selection and optimization of microbial hosts for biofuels production.* Metab Eng, 2008. **10**(6): p. 295-304.
47. Liu, S., et al., *Progress and perspectives on improving butanol tolerance.* World J Microbiol Biotechnol, 2017. **33**(3): p. 51.
48. Ezeji, T., et al., *Achievements and perspectives to overcome the poor solvent resistance in acetone and butanol-producing microorganisms.* Appl Microbiol Biotechnol, 2010. **85**(6): p. 1697-712.
49. Weber, F.J. and de Bont, J.A.M., *Adaptation mechanisms of microorganisms to the toxic effects of organic solvents on membranes.* 1996.
50. Rau, M.H., et al., *Genome-wide Escherichia coli stress response and improved tolerance towards industrially relevant chemicals.* Microb Cell Fact, 2016. **15**(1): p. 176.
51. Segura, A., et al., *Solvent tolerance in Gram-negative bacteria.* Curr Opin Biotechnol, 2012. **23**(3): p. 415-21.
52. Isken, S. and de Bont, J.A.M., *Bacteria tolerant to organic solvents.* Extremophiles, 1998. **2**: 229-238.
53. Rastadter, K., et al., *The Cell Membrane of Sulfolobus spp.-Homeoviscous Adaption and Biotechnological Applications.* Int J Mol Sci, 2020. **21**(11).
54. Sinensky, M., *Homeoviscous Adaptation-A Homeostatic. Process that Regulates the Viscosity of Membrane Lipids in Escherichia coli.* Proc. Nat. Acad. Sci. USA, 1974. **71**(2):522-525.
55. Eberlein, C., et al., *Immediate response mechanisms of Gram-negative solvent-tolerant bacteria to cope with environmental stress: cis-trans isomerization of unsaturated fatty acids and outer membrane vesicle secretion.* Appl Microbiol Biotechnol, 2018. **102**(6): p. 2583-2593.
56. Ramos, J.L., et al., *Mechanisms of solvent tolerance in gram-negative bacteria.* Annu Rev Microbiol, 2002. **56**: p. 743-68.
57. Rutherford, S.T. and Bassler, B.L., *Bacterial quorum sensing: its role in virulence and possibilities for its control.* Cold Spring Harb Perspect Med, 2012. **2**(11).

58. Zingaro, K.A. and Papoutsakis, E.T., *Toward a semisynthetic stress response system to engineer microbial solvent tolerance.* mBio, 2012. **3**(5).
59. Ezeji, T.C., et al., *Butanol fermentation research: upstream and downstream manipulations.* Chem Rec, 2004. **4**(5): p. 305-14.
60. Zhuang, L.L., et al., *Enhanced attached growth of microalgae Scenedesmus. LX1 through ambient bacterial pre-coating of cotton fiber carriers.* Bioresour Technol, 2016. **218**: p. 643-9.
61. Dong Liu, Y.C., et al., *Biobutanol production in a Clostridium acetobutylicum biofilm reactor integrated with simultaneous product recovery by adsorption.* Biotechnology for Biofuels, 2014. **7**(5).
62. van Wolferen, M., et al., *Archaeal biofilm formation.* Nat Rev Microbiol, 2018. **16**(11): p. 699-713.
63. Flemming, H.C. and Wingender, J., *The biofilm matrix.* Nat Rev Microbiol, 2010. **8**(9): p. 623-33.
64. Flemming, H.C. and Wuertz, S., *Bacteria and archaea on Earth and their abundance in biofilms.* Nat Rev Microbiol, 2019. **17**(4): p. 247-260.
65. Schocke, L., et al., *Thermoacidophilic Sulfolobus species as source for extremozymes and as novel archaeal platform organisms.* Curr Opin Biotechnol, 2019. **59**: p. 71-77.
66. Zeldes, B.M., et al., *Extremely thermophilic microorganisms as metabolic engineering platforms for production of fuels and industrial chemicals.* Front Microbiol, 2015. **6**: p. 1209.
67. Frock, A.D. and Kelly, R.M., *Extreme Thermophiles: Moving beyond single-enzyme biocatalysis.* Curr Opin Chem Eng, 2012. **1**(4): p. 363-372.
68. Jachlewski, S., et al., *Isolation of Extracellular Polymeric Substances from Biofilms of the Thermoacidophilic Archaeon Sulfolobus acidocaldarius.* Front Bioeng Biotechnol, 2015. **3**: p. 123.

Acknowledgements

At the end of my thesis, I would like to thank all those people who made this thesis possible.

My deep gratitude goes foremost to my supervisor Prof. Dr. Bettina Siebers for offering me the opportunity to work in her group on this interesting and meaningful topic. I greatly thank for her constant support, trust and understanding throughout this study.

I would like to express my heartfelt gratitude to Dr. Christopher Bräsen, who patiently guided me and enriched my knowledge with his exceptional insights since I started this work. I deeply thank him for helping me whenever I was in need and for providing constructive ideas, experimental advice and detailed proofreading of the publications and the thesis. Without his illuminating instruction, this thesis would not have reached its present form.

I own my sincere gratitude to Prof. Dr. Peter Bayer for being my co-referee of this thesis.

I am very grateful to all my collaboration partners for their support and help, especially to Prof. Eveline Peeters, Prof. Dr. Ann-Christin Lindås, Prof. Dr. Markus Kaiser, Prof. Dr. Sonja-Verena Albers, Prof. Dr. Jörn Kalinowski, Prof. Dr. Oliver J. Schmitz, Dr. Sven Meckelmann, David Sybers, Dr. Kun Wang, Till Kessenbrock, Dr. Tobias Busche, Andreas Albersmeier and Paul E. Görs. I also thank Volkswagen for their financial support to this work.

I would like to extend thanks to Thomas Knura, Sabine Dietl and Alex Wagner for their technical assistance.

I wish to acknowledge my colleague Christian Schmerling for his support and nice teamwork during HPLC analysis, and for sharing his knowledge about bioinformatics and protein structure modeling.

Many thanks to Dr. Lu Shen, Laura Kuschmierz, Dr. Christina Stracke for their warmful encouragement and friendship during my work and my common life. I would like to thank all my dear former and current colleagues for their help and nice working environment: Svenja Höfmann, Larissa Schocke, Thomas Klaus, Agathe Materla, Dr. Benjamin Meyer, Promise Akua Dzandu, Sabine Krevet, Dr. Frank Schult, Katharina Fafenrot, Claus-Rüdiger Wallis, Dr. Jens Benninghoff and Marcel Blum.

Thanks to all my students Annika Krüger, Thorsten Meyer, Simon Kleimann, Yuteng Zhang for their hard and outstanding work.

Above all, I express my special gratitude to my family especially my parents and my husband for their unconditional support, understanding, endless love and dedication.

Erklärung der selbstständigen Verfassung der Dissertation

Hiermit erkläre ich, dass ich die vorliegende Dissertation mit dem Titel:

Fatty acid metabolism in *Sulfolobus acidocaldarius* and its potential as platform organism in biotechnology

selbstständig verfasst und keine außer den angegebenen Hilfsmitteln und Quellen benutzt habe. Alle wörtlich oder inhaltlich übernommenen Stellen habe ich als solche gekennzeichnet.

Ich versichere außerdem, dass keine vorausgegangenen Promotionsverfahren in diesem oder einem anderen Fach endgültig gescheitert sind. Ich erkläre weiterhin, dass ich diese Arbeit nur in diesem Promotionsverfahren eingereicht habe.

_____

Ort, Datum

_____

(Xiaoxiao Zhou)