

**Fatty acid and glycerol metabolism in
Sulfolobus acidocaldarius and establishing
new methods for identification of novel
biocatalysts in the white-rot fungus
*Phanerochaete chrysosporium***

Dissertation

zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften

- Dr. rer. nat. -

vorgelegt von

Christian Schmerling

Arbeitskreis für Molekulare Enzymtechnologie und Biochemie

Abteilung für Umweltmikrobiologie und Biotechnologie

Fachbereich Chemie

der

Universität Duisburg-Essen

März 2023

Die vorliegende Arbeit wurde im Zeitraum von Mai 2018 bis März 2023 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: 15.08.23

Gutachter: Prof. Dr. Bettina Siebers

Prof. Dr. Markus Kaiser

Vorsitzender: Prof. Dr. Matthias Epple

DuEPublico

Duisburg-Essen Publications online

UNIVERSITÄT
DUISBURG
ESSEN

Offen im Denken

ub | universitäts
bibliothek

Diese Dissertation wird via DuEPublico, dem Dokumenten- und Publikationsserver der Universität Duisburg-Essen, zur Verfügung gestellt und liegt auch als Print-Version vor.

DOI: 10.17185/duepublico/79090

URN: urn:nbn:de:hbz:465-20250108-092119-5

Alle Rechte vorbehalten.

Table of contents

1. Introduction	1
1.1 Preface	1
1.2 Archaea	2
1.3 FA degradation via a largely bacterial like β oxidation in <i>S. acidocaldarius</i>	6
1.4 Potential FA synthesis in <i>S. acidocaldarius</i>	10
1.5 Glycerol degradation via a largely bacterial like Glycerol kinase and an unusual Glycerol-3-phosphate dehydrogenase in <i>S. acidocaldarius</i>	15
1.6 Metabolic thermoadaptation in (hyper)thermophilic Archaea	20
1.7 Screening for novel biocatalysts secreted by the model white rot fungus <i>P. chrysosporium</i> using ABPP	22
1.8 References	27
2. Scope of the thesis	33
3. Manuscripts	35
Chapter 3.1: Fatty acid metabolism in <i>Sulfolobus acidocaldarius</i>	35
Chapter 3.2: Glycerol degradation in the thermoacidophilic crenarchaeon <i>Sulfolobus acidocaldarius</i> involves an unusual glycerol-3-phosphate dehydrogenase	109
Chapter 3.3: Enhanced underground metabolism challenges life at high temperatures – metabolic thermoadaptation in hyperthermophilic Archaea	173
Chapter 3.4: Identification of fungal lignocellulose-degrading biocatalysts secreted by <i>Phanerochate chrysosporium</i> via activity-based protein profiling	183
5. Summary	207
6. Zusammenfassung	210
Curriculum vitae	I
Acknowledgements	II

1. Introduction

1.1 Preface

In the work presented herein, two major topics were addressed and thus this thesis is separated into two main parts:

The first part (Part I) deals with the metabolism in the third domain of life, the Archaea, and is divided into three subsections: In the first subsection (Chapter 3.1), the fatty acid (FA) metabolism, which has not been studied in detail in any archaeon so far, was investigated in the thermoacidophilic crenarchaeon *Sulfolobus acidocaldarius*. It was shown that this archaeal model organism contains a fully functional largely bacterial-like β oxidation pathway for FA degradation which was biochemically characterized in detail. Furthermore, a potential novel pathway for FA synthesis acting completely independent from β oxidation was identified and characterized. In the second subsection (Chapter 3.2), the glycerol utilization in *S. acidocaldarius* was elucidated and the involved enzymes particularly including an unusual glycerol-3-phosphate dehydrogenase (G3PDH) were biochemically characterized. Finally, in the third subsection (Chapter 3.3), the current insights in the archaeal underground metabolism and metabolic thermoadaptation are discussed which enable especially (hyper)thermophilic Archaea including also *Sulfolobus* spp. to operate their unusual metabolism under conditions where many metabolic intermediates are unstable.

The second main part (Part II) (Chapter 3.4) of this thesis illustrates the high potential of the activity-based protein profiling (ABPP) approach in the discovery of novel enzymes. Using this methodology, the secretome including wood substrate-bound enzymes of the lignocellulose degrading white rot fungus *Phanerochaete chrysosporium* was analysed and identified active enzymes were characterized. This ABPP-based screening can help to identify those active enzymes in a process of interest and to find novel biocatalysts that share no sequence similarity to known counterparts.

Part I

1.2 Archaea

As mentioned in the preface this thesis takes a closer look at the metabolism of the archaeon *S. acidocaldarius*. Archaea have initially been recognized as one of the three domains of life together with Bacteria and Eukaryotes based on 16S/18S rRNA sequences and in this tree Archaea and Eukaryotes are regarded as sister groups [1]. The first two identified archaeal phyla were the Euryarchaeota and the Crenarchaeota [2] (Fig. 1A). At that time Archaea were regarded either as extremophilic or as metabolic specialists. However, in recent years different metagenomics/environmental molecular biology methods revealed that Archaea are also ubiquitously distributed in mesophilic habitats. In addition, they seem to function as important players in different environments [3]. Meanwhile, multiple additional new archaeal phyla have been found which are now grouped in four major clades, (i) the Euryarchaeota, (ii) the TACK (Thaum-, Aig-, Cren-, Korarchaeota) superphylum [4] (iii) the DPANN superphylum (Diapherotrites, Parv-, Aenigma-, Nano-, and Nanohaloarchaeota), and (iv) the Asgard archaea comprising the Loki-, Thor-, Odin-, and Heimdallarchaeota [5]. New phylogenomic analyses and the discovery of the Asgard archaea even suggest that Bacteria and Archaea represent the only two primary evolutionary lineages and that the Eukaryotes originated later on from within the Archaea [5,6] (Fig. 1B).

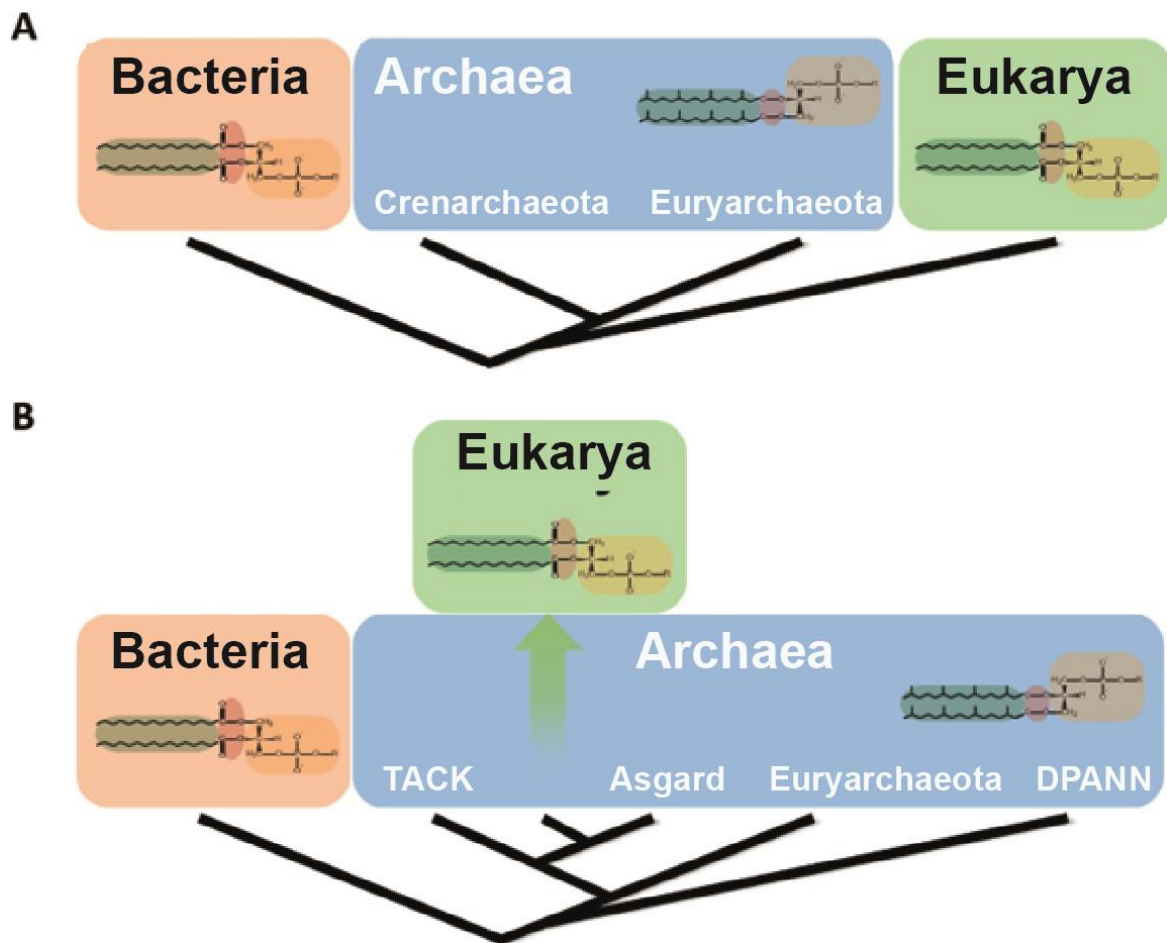


Figure 1. Evolutionary relationship of Archaea, Eukaryotes and Bacteria. (A) Phylogenetic three-domain tree of life based on 16S/18S rRNA sequences [2]. **(B)** Updated phylogenetic two-domain tree of life based on 16S/18S rRNA sequences [5,6]. The schematic lipid compositions in Bacteria, Eukaryotes and Archaea show that the cell membranes of Bacteria and Eukaryotes consists of fatty acid-based ester bound to G3P, while archaeal membrane consists of isoprenoids ether bound to G1P.

Archaea share a mixture of certain bacterial properties such as cell organization and DNA structure as well as some less complex versions of eukaryotic features such as replication, transcription, DNA-repair, and translation [7]. However, although Archaea are similar in metabolic complexity and genomic organization compared to Bacteria, they also contain specific archaeal features [8]. Regarding their lifestyle or metabolic complexity, Archaea resemble Bacteria as well as lower Eukaryotes, and chemolithoautotrophic, chemoorganoheterotrophic, and phototrophic growth was reported [9,10]. This metabolic complexity in Archaea is complemented by the absence of many classical pathways. In this regard Archaea are characterized by the presence of unique pathways, such as methanogenesis, and by some unusual, modified-pathway versions of these classical routes [8]. For example, regarding the central carbohydrate metabolism, modified variants of sugar degradation pathways such as the Embden-Meyerhof-Parnas (EMP) pathway and the Entner-Doudoroff (ED) pathway have been identified in Archaea [8]. Furthermore, the pentose phosphate pathway (PPP) is only present partly or not at all and thus pentose degradation is significantly different from what is known for bacterial model organisms [8].

In addition, unlike the bacterial cell wall that is composed of murein/petidoglycan, most Archaea possess surface layer proteins (S-layer) as their cell wall components. In some Archaea (for example, *Sulfolobus* spp.) S-layer proteins are the sole cell envelope constituent, whereas in few other Archaea the cell envelope can also contain polymers, including the polysaccharides pseudomurein and methanochondroitin (e.g., methanogenic species), while S-layers are only completely missing in some members of certain genera [11]. S-layers often consist of mushroom-shape (glyco)proteins that are membrane bound via either transmembrane domains or lipid-modified subunits. Functionally S-layers serve as protective coats, and play key roles in cell adhesion, surface recognition, antifouling and cell shape [11,12].

Another characteristic trait of Archaea is their membrane lipid composition [13]. The so called "lipid divide" describes the observation that the membrane lipid composition of Archaea is essentially different from those in Bacteria and Eukaryotes. While lipids of Bacteria and Eukaryotes are comprised of FAs ester-linked to glycerol-3-phosphate (G3P) that form membrane bilayers archaeal membrane lipids are made of isoprenoid chains ether-linked to glycerol-1-phosphate (G1P) forming either mono-or bilayer membranes (Fig. 1). This ether linkage in archaeal lipids is chemically more stable than ester bonds of bacterial/eukaryotic lipids and methyl branches of isoprenoid chains potentially support lipid packing [14-16]. Thus, since one of the main functions of FAs as key part of membrane phospholipids and thus of cell structure in Bacteria and Eukaryotes is substituted by isoprenoids in Archaea, the presence and function of FAs in Archaea remains unknown. For this reason, this works details

how FAs are degraded by the archaeal model organism *S. acidocaldarius* and explores its potential regarding FA synthesis.

S. acidocaldarius, a member of the Crenarchaeota has been originally isolated from acidic hot springs and grows heterotrophically at temperatures around 80°C and a pH of 3.0 [2]. *S. acidocaldarius* is known for its ease of cultivation, genetic tractability as well as metabolic versatility allowing the organism to utilize D-glucose, D-xylose, L-arabinose, starch, dextrin, saccharose, and proteinaceous substrates like tryptone, NZ amine and single amino acids, and in addition several -omics approaches for the organism have been established [17-19]. Interestingly, it was shown that two esterases (Saci_1105, Saci_1116) are involved in triacylglyceride degradation in *S. acidocaldarius* [20] and the corresponding genes are organized in a gene cluster (*saci_1103-saci_1126*) together with several genes encoding homologues of all proteins necessary for a functional β oxidation. Furthermore, it was shown that *S. acidocaldarius* can use FAs, i.e., butyrate and hexanoate, as sole source for carbon and energy for growth [21]. One gene within this gene cluster, *saci_1107*, encodes a transcriptional regulator from the TetR-family that was shown to exclusively bind in the *saci_1103-saci_1126* gene cluster and to regulate its own expression in addition to many genes encoding for β oxidation homologues [21]. Chapter 3.1 describes the characterization of those homologous and elucidates a pathway for FA degradation in *S. acidocaldarius* and additionally indicates a novel potential pathway for FA synthesis. Furthermore, as outlined above, *S. acidocaldarius* is able to cleave triacylglycerides by means of esterases into FAs and glycerol and therefore chapter 3.2 further reexamines the utilization of glycerol by *S. acidocaldarius*, previously regarded as a non-glycerol utilizer [17] and compares this breakdown with known pathways from Bacteria and Eukaryotes. Lastly, as a thermoacidophilic archaeon *S. acidocaldarius* has evolved strategies to mitigate metabolic instabilities associated with high temperatures and chapter 3.3 discusses how the metabolism of organisms such as *S. acidocaldarius* is adapted to growth at high temperatures.

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

As described in the preface this work takes a closer look at a potential pathway for FA degradation in *S. acidocaldarius* as presented in chapter 3.1. The ability of several Archaea to degrade FA based lipids to free FAs and glycerol by means of esterases and lipases *in vivo* was recently analyzed in detail for *S. acidocaldarius*, *Saccharolobus solfataricus*, and *Haloferax volcanii* [19,20]. In addition, a previous reports suggest that some Archaea might be able to utilize FAs as carbon and energy source (e.g., haloarchaea and *Archaeoglobus fulgidus*) via homologues of a bacterial-like β oxidation pathway, and homologues of bacterial-like β oxidation genes have been identified in several Archaea, although detailed analyses have not been reported so far [22]. As mentioned, it was already shown that *S. acidocaldarius* can utilize FAs as a sole carbon and energy source and contains a gene cluster encoding for homologues of a bacterial-like β oxidation (*saci_1103-saci_1126*) that is regulated by a TetR-family transcription factor [21], however it was not shown that the enzymes in this genecluster could actually take part in FA degradation. To further study the FA metabolism this work analysed chosen 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 *Escherichia coli*, and the recombinant proteins were purified and biochemically characterized to analyse how *S. acidocaldarius* could be able to degrade FAs and the elucidated pathway is depicted in Fig. 2A.

To allow for degradation of free FAs, they first need to be activated to the corresponding acyl-CoA esters, carried out by AMP-forming acyl-CoA synthetases (ACS) using ATP which then get degraded via β oxidation, the concerted consecutive action of four different enzymes. 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 protein catalysed the HS-CoA and ATP dependent activation of FAs forming acyl-CoAs as product with the highest activity towards chain lengths between C5 and C8, while no activity was observed with FAs longer than C10 classifying the enzyme as medium chain acyl-CoA synthetase.

In the first step of the β oxidation cycle acyl-CoA dehydrogenases (ACAD) catalyse the formation of a double bound at the α - β position of an acyl-CoA oxidizing it to a respective enoyl-CoA ester [23]. Generated electrons are then transferred via the electron transfer flavoprotein (ETF) to the membrane bound electron transfer flavoprotein-ubiquinone oxidoreductase (ETF-Q) ultimately transporting the electrons to the quinone pool of the respiratory chain. Three homologues of ACADs are encoded in the described gene cluster

(Saci_1108, Saci_1113 and Saci_1123) and candidates for the ETF and the ETF-Q encoded by *saci_0315-saci_0317* as well as *saci_0290-saci_0293* were found elsewhere in the *S. acidocaldarius* genome. In detail, Saci_1123 was shown to contain ACAD activity using flavin adenine dinucleotide (FAD) as a cofactor in presence of the ETF Saci_0315 or an artificial electron acceptor e.g., ferrocenium (FcPF₆) or 2,6-dichlorophenolindophenol (DCPIP), and preferred saturated medium-chain to long chain fatty acyl-CoAs (C4-C12) as substrates with the highest preference for C8-CoA, indicating that Saci_1123 belongs to the medium chain dehydrogenase group of ACADs. In contrast sequence alignment revealed that other ACADs encoded in the analyzed genecluster, Saci_1108 and Saci_1113, seem to be distorted lacking the binding site for CoA or FAD, respectively. Biochemical characterization of Saci_0315 showed that the protein serves as an electron acceptor for the Saci_1123 catalysed acyl-CoA oxidation. Saci_0315 represents a fusion protein compromising both the α and β subunits of eukaryotic/bacterial ETFs a feature only known to exist in archaeal ETFs [24].

The next step in β oxidation is the hydration and oxidation of the enoyl-CoA to the corresponding a 3-ketoacyl-CoA via a 3(S)-hydroxyacyl-CoA intermediate catalysed by enoyl-CoA hydratases (ECH) and hydroxyacyl-CoA dehydrogenase (HCDH). For both, only one gene (*saci_1109*) exists in the cluster and both enzyme families are fused together forming an enzyme with an N-terminal HCDH domain and a C-terminal ECH domain. For Saci_1109 it was shown that the enzyme contained both enoyl-CoA hydratase and 3-hydroxyacyl-CoA dehydrogenase activity *in vitro* using crotonoyl-CoA (C4:1), decenoyl-CoA (C10:1) and 3-hydroxybutyryl-CoA as substrates, and NAD⁺ as a cofactor, while no activity was measured using hexadecenoyl-CoA (C16:1). Saci_1109 is specific for the (S)-stereoisomer catalysing the oxidation of 3(S)-hydroxybutyryl-CoA while the enzyme showed no activity with the (R)-stereoisomer. Additionally, Saci_1109 was able to convert acetoacetyl-CoA into 3-hydroxybutyryl-CoA in presence of NADH as well meaning that the enzyme functions reversibly. The observed bi-functional activities are similar to bacterial FadB homologues, e.g., from *E. coli* [25], however here the domain organization is inverted.

The final reaction in the β oxidation is catalysed by a β -ketothiolase/acetyl-CoA acetyltransferase (KT) to yield acetyl-CoA and a saturated acyl-CoA ester shortened by two C atoms. Two homologues are present in the *saci_1103-1126* gene cluster (*saci_1114* and *saci_1121*) and the coding function of purified recombinant Saci_1114 was confirmed for both directions of acetoacetyl-CoA cleavage. Canonical homologues of KT and ECH/HCDH e.g., from *Pseudomonas fragi* or *E. coli* are known to form a complex which enables substrate channelling [26-28]. However, this work indicated that in contrast to the *E. coli* complex the *S. acidocaldarius* enzymes Saci_1109 and Saci_1114 could not be reconstituted *in vitro* into a protein complex and the results might indicate that the inverted domain structure prevents or alters complex formation [29].

The function of all four enzymes was additionally confirmed *in vitro* by analysing their combined function in an enzyme cascade via HPLC and the conversion of butyryl-CoA, hexanoyl-CoA or octanoyl-CoA to acetyl-CoA was demonstrated (Fig. 2B exemplarily shows the stepwise conversion of butyryl-CoA). Firstly butyryl-, hexanoyl- and octanoyl-CoA were completely oxidized to the respective enoyl-CoA esters by Saci_1123 via either ETF or with the artificial electron acceptors DCPIP or FcPF₆. These biochemical analyses show that the first step of β oxidation, i.e., the oxidation of an acyl-CoA to enoyl-CoA by ACADs, runs to completion independent of the electron acceptor used, indicating that this reaction functions as a major driving force for the degradation for acyl-CoAs.

In general, this step has been described as irreversible for kinetic reasons: In their reduced state ACADs preferentially bind their enoyl-CoA product which kinetically promotes the enzymes oxidative half reaction i.e. the transfer of electrons to the respiratory chain [30]; additionally, the redox potential of the caldariellaquinone (+ 100 mV) as primary electron acceptor and oxygen (+ 820 mV) as the terminal electron acceptor of the respiratory chain, is significantly more positive than an acyl-CoA/enoyl-CoA couple (-10 mV) which renders the reversal of acyl-CoA oxidation by ACADs at least unlikely.

Next, Saci_1109 converted >90% of the formed crotonyl-CoA to hydroxybutyryl-CoA (in agreement with the thermodynamics of the reaction ($\Delta G^0 = -3.3 \text{ kJ mol}^{-1}$) [31]) although the further conversion with NAD⁺ to acetoacetyl-CoA could not be observed in accordance with the hydroxyacyl-CoA oxidation being the thermodynamic bottleneck (+18,1 kJ mol⁻¹) of β oxidation.

Only when Saci_1114 was introduced crotonyl-CoA was almost completely converted to acetyl-CoA which is explained by the standard free energy change of -26.1 kJ mol⁻¹ of the thiolytic cleavage rendering the crotonyl- to acetyl-CoA conversion exergonic by -11 kJ mol⁻¹ [31] and accordingly a nearly full conversion of crotonyl-CoA to acetyl-CoA by HCDH/ECH and KT was observed.

Thus, in total Saci_1123, Saci_1109 and Saci_1114 reconstituted a fully functional β oxidation spiral degrading acyl-CoAs up to chain lengths of C8. 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" KTs [22].

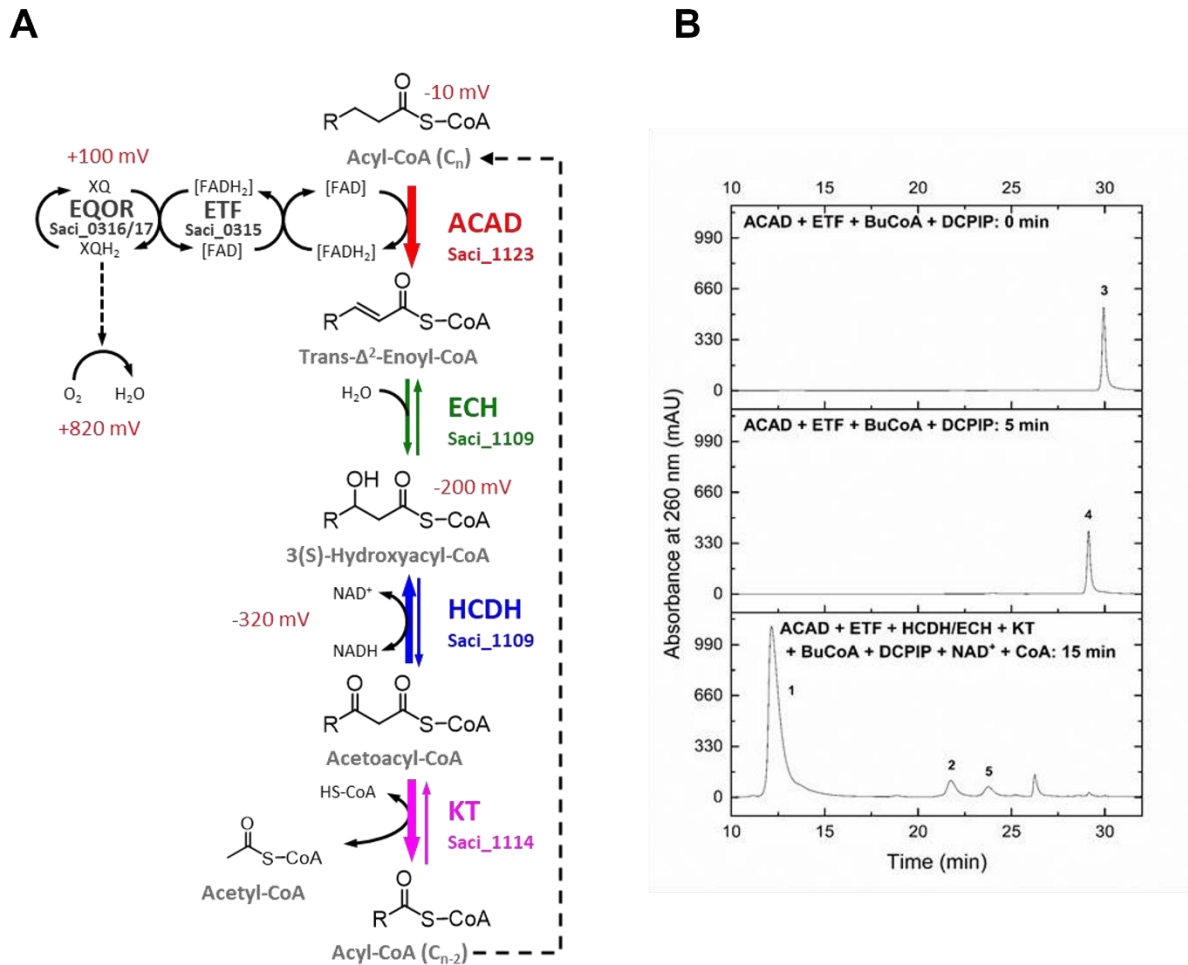


Figure 2: Reconstructed fatty acid (FA) degradation in *S. acidocaldarius*. (A) Acyl-CoAs enter the first step of a β oxidation spiral and are oxidized to unsaturated enoyl-CoA by acyl-CoA dehydrogenases (ACADs) simultaneously transferring the electron to oxygen of the respiratory chain through electron transferring flavoproteins (ETFs) and ETF:quinone oxidoreductases (ETF-Q). Then enoyl-CoAs are converted to 3(S)-hydroxyacyl-CoA and further oxidized to acetoacyl-CoA by bi-functional enzyme 3(S)-hydroxyacyl-CoA dehydrogenases/enoyl-CoA hydratases (HCDHs/ECHs). Finally, β -ketothiolases (KTs) cleave acetoacyl-CoAs into acetyl-CoA and acyl-CoAs shortened by two carbons. The shortened acyl-CoAs can further participate in the cycle until the acyl chains are broken down to acetyl-CoAs. The thickness of the arrow indicates the energetics of the respective reaction. For redox reactions the reduction potential is given in red. (B) During β oxidation enzyme cascade, butyryl-CoA (peak 3) was completely oxidized to crotonoyl-CoA (peak 4) by ACAD transferring the electron to DCPIP through ETF. 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) while the intermediate acetoacetyl-CoA was not detectable under the applied analytical conditions. 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.

1.4 Potential FA synthesis in *S. acidocaldarius*

Few reports previously suggested that in few Archaea FAs are present indicating that these archaeal organisms have a biosynthesis machinery for FAs. The presence of FAs was e.g., indicated for *S. solfataricus*, a close relative of *S. acidocaldarius*, for *Ignicoccus hospitalis* [32], for some methanogens [33], and also for some haloarchaea [34]. However, neither a complete classical bacterial FA synthesis system (FAS II) nor a FAS I machinery, that are present in animals and fungi, have been identified in any archaeon. In addition, acyl carrier proteins (ACP) and ACP-synthases essential for the FA biosynthesis in Bacteria and Eukaryotes are absent in nearly all Archaea with only very few exceptions [22]. Thus, FA containing Archaea require a synthesis machinery that is fundamentally different from known FAS I and FAS II systems in Bacteria and Eukaryotes. Recently, a phylogenetic reconstruction of archaeal genomes indicated the presence of a potentially ACP-independent FA synthesis pathway which could function using bacterial like homologues of the FAS II [35]. However, since a complete sets of FAS II homologues was not identified in any archaeon so far, instead it was proposed that a reversed β oxidation that is also ACP-independent could be used FA synthesis in Archaea [22,35]. However, the reversibility of β oxidation in general and particularly in Archaea has not been shown so far.

Even though the chemical reactions are similar, both β oxidation and FA synthesis show some remarkable differences in Bacteria and Eukaryotes, in order to drive either of the processes in the desired direction (FA degradation or synthesis, respectively) and to separate both pathways from each other, which both – at least in prokaryotes – (may) run (simultaneously) in the cytoplasm [25]. Firstly, for separation both pathways are tightly regulated on a transcriptional level and via protein level regulation e.g. by feedback inhibition [25]. In addition, the process of FA synthesis is dependent on ACP for activation and transport of growing acyl-thioester chains [36], while β oxidation exclusively relies on CoA for FA activation. Furthermore, FA synthesis is usually NADPH dependent, for both the reduction of ketoacyl-ACP and enoyl-ACP, whereas the β oxidation is dependent on NAD⁺ and FAD for oxidation [37] although this kind of cosubstrate specificity is less pronounced in Bacteria [38]. Next, FA degradation and synthesis are separated via the formation of either S- or R- hydroxyacyl-CoAs/ACPs, respectively.

As described above from a kinetic and energetic point of view, two reactions in the β oxidation are hardly reversible and need to be by-passed in the synthesis direction. Firstly, the ketothiolytic cleavage of the ketoacyl-CoAs into acetyl-CoA and acyl-CoA_(n-2) during β oxidation is thermodynamically strongly favoured by $-26.1 \text{ kJ mol}^{-1}$ [31]. This is bypassed in the FA

synthesis by using malonyl-ACP as a donor for chain elongation. Malonyl-ACP has to be synthesized through the ATP-dependent carboxylation of acetyl-CoA followed by the transfer from CoA to ACP [25,37]. During condensation with another molecule of acetyl-CoA, CO₂ is liberated again, which renders the Claisen condensation strongly exergonic, through ATP hydrolysis. Secondly, as described above first step of β oxidation, i.e., the oxidation of an acyl-CoA runs to completion and is irreversible. So, for FA synthesis this step also needs to be bypassed. In the Eukarya and Bacteria this is accomplished via enoyl-ACP reductases that use NAD(P)H (-320 mV) as an electron donor which renders the reduction of enoyl-CoAs to acyl-CoAs (-10 mV) strongly exergonic [37].

As shown the other two reactions of β oxidation are in principle reversible, however for the β oxidation in *S. acidocaldarius* the equilibrium is far on the side of degradation (>90 % conversion rate), based especially on the thiolytic cleavage by KTs (see section 1.3). Thus, even though some reactions are reversible in principle both from a mechanistic and energetic point of view, it is unlikely that the β oxidation can also function in the reverse/synthesis direction. These considerations raise the questions how FA synthesis might basically proceed in Archaea and specifically *S. acidocaldarius*, how the process is thermodynamically driven, and how FA synthesis could be separated from β oxidation, especially since ACP is absent in Archaea. The second part of chapter 3.1 describes a putative pathway in *S. acidocaldarius* as depicted in Fig. 3A that is separated from β oxidation and that fits the thermodynamic constraints to drive the conversion in the direction FA synthesis.

The first step for a functional FA synthesis pathway that does not include a 3-ketoacyl-ACP synthase is the Claisen condensation of an acetyl-CoA molecule with an ac(et)yl-CoA yielding a ketoacyl-CoA intermediate. This normally unfavourable reaction can be enabled by DUF35 scaffolding proteins (see chapter 3.1 for details) that facilitate the formation of protein complexes with enzymes that immediately convert the condensed ketoacyl-CoA. This in turn enables substrate channelling without liberating the intermediate as was recently described for a KT and 3-hydroxy-3-methylglutaryl-CoA-synthase complex [39]. Interestingly in the analysed gene cluster a homologue of DUF35 scaffolding proteins encoded by *saci_1120* was identified, located directly downstream of a gene encoding for a KT (*saci_1121*), indicating that *S. acidocaldarius* could also form KT-protein complexes to allow for this condensation reaction. The second and third step of FA synthesis require the action of a 3-oxoacyl-reductase reductase and a hydroxyacyl-CoA dehydratase. Another enzyme encoded by the *saci_1103-saci_1126* gene cluster is *Saci_1104*, a 3-oxoacyl-reductase homologue belonging to the short chain dehydrogenase (SDR) superfamily. The enzyme showed reversible 3-oxoacyl-CoA reductase activity and was specific for the 3(R)-hydroxybutyryl-CoA-

stereoisomer as well as for NADP⁺/NADPH. No activity was observed with the (S)-stereoisomer and NADH. As described above, this (R) specificity is a distinctive feature of FA synthesis known from Eukarya and Bacteria whereas β oxidation uses 3(S)-stereoisomers [37]. However, if FA synthesis in *S. acidocaldarius* indeed functions via a 3(R)-stereoisomer intermediate, this would mean that the organism also requires the presence of a 3(R)-hydroxyacyl-CoA dehydratase. In bacterial/eukaryotic FA synthesis the dehydration of an 3(R)-hydroxyacyl-ACP is carried out by enzymes belonging to the hot dog fold superfamily called FabZ and FabA [40,41], however no homologues of these enzymes can be found in the *S. acidocaldarius* genome. Only two distantly related enzymes from the hotdog fold superfamily could be found, Saci_1070 and Saci_1085, belonging to the group of MaoC dehydratases. The recombinantly produced Saci_1085 indeed showed both (R)-hydroxyacyl-CoA dehydratase and enoyl-CoA hydratase activity and the enzyme was characterized. However, Saci_1070 was found to be catalytically inactive likely due to a missing catalytically essential histidine residue [40]. Interestingly a similar observation was made in the pathogenic bacterium *Mycobacterium tuberculosis*, where a MaoC dehydratase (HadB) is part of a pathway of elongation of very long chain FAs [40,42].

The last step of this putative FA synthesis pathway involves an enoyl-CoA reductase that catalyses the formation of a saturated acyl-CoA molecule from the corresponding enoyl-CoA ester. The *saci_1103-saci_1126* gene cluster also contains a gene encoding a putative alcohol dehydrogenase, Saci_1115, that belongs to the medium chain dehydrogenase/reductase family (MDR) and shows a high similarity to the acryloyl-CoA reductase from *Metallosphaera sedula*, which indicates that Saci_1115 may possess enoyl-CoA reductase activity. Heterologous production and purification revealed Saci_1115 to be indeed able to catalyse the NADPH dependent reduction of medium chain enoyl-CoAs (C4 to C10). In accordance with the thermodynamics of the reaction (-56 kJ mol^{-1}) (see also above) the enzyme exclusively functioned in the reductive direction and could thus represent a strong driving force for FA synthesis. As already mentioned above the clear preference for NADPH is in accordance with the known cosubstrate specificity of bacterial/eukaryotic FA synthesis.

From the purified single enzymes described above, the putative FA synthesis cascade was reconstituted, and the conversion was followed stepwise via HPLC (Fig. 3B). A combination of Saci_1085 and Saci_1115 led to the formation of butyryl-CoA from 3(R)-hydroxybutyryl-CoA. Furthermore, by adding Saci_1104 to the enzyme mixture the formation of butyryl-CoA could be confirmed using acetoacetyl-CoA as a starting substrate. Together these results indicate a potential novel ACP independent FA synthesis pathway in *S. acidocaldarius*, and this work strongly indicate that the β oxidation pathway is not operating reversibly in Archaea (Fig. 2A). Instead, we propose a potential archaeal FA synthesis pathway (Fig. 3A) 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).

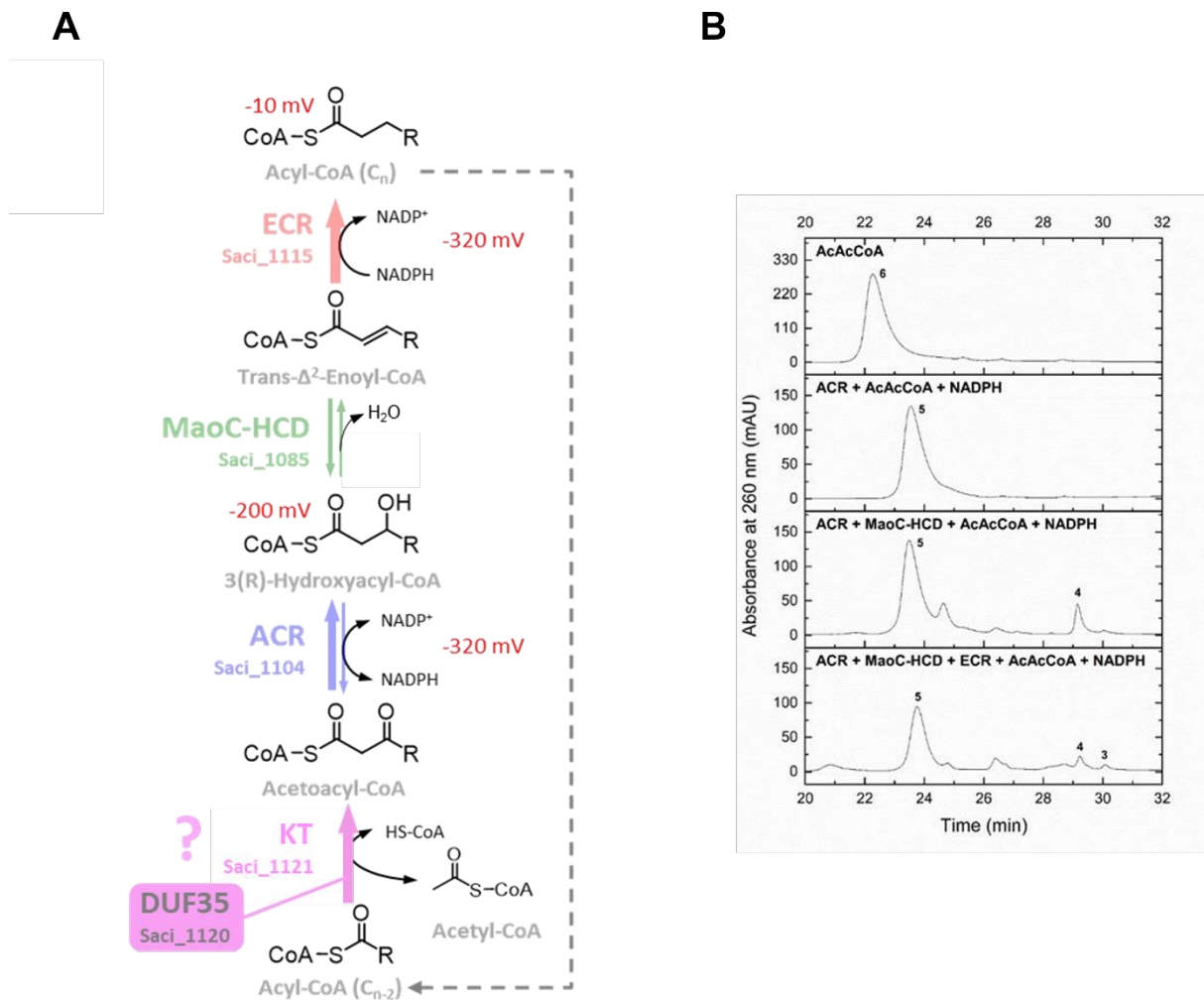


Figure 3: Reconstructed fatty acid (FA) synthesis in *S. acidocaldarius*. In the proposed novel fatty acyl elongation pathway in *S. acidocaldarius*, short or medium chain acyl-CoAs (e.g. C₂, C₄ or C₆) are condensed with acetyl-CoA to synthesize acetoacyl-CoAs by archaeal type KT. The resulting intermediates 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) catalyse the dehydration from the 3(R)-hydroxy derivative to enoyl-CoAs. The MDR family enoyl-CoA reductases (ECRs) then catalyse the reduction of the enoyl-CoAs to acyl-CoAs elongated by two carbons. The newly resulting acyl-CoAs will be further elongated until a desired length is reached. The thickness of the arrow indicates the energetics of the respective reaction. For redox reactions the reduction potential is given in red. **(B)** During FA synthesis enzyme cascade 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. 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.

1.5 Glycerol degradation via a largely bacterial like Glycerol kinase and an unusual Glycerol-3-phosphate dehydrogenase in *S. acidocaldarius*

Since it has recently been shown that *S. acidocaldarius* cleaves triacylglycerides by means of esterases [20] and to grow with the FAs as sole source of carbon and energy [21] potentially using a FA degradation pathway described above, this work reexamined the growth of *S. acidocaldarius*, on the other lipid cleavage product glycerol. Glycerol is an integral constituent as the backbone of membrane phospholipids (in all domains of life) and of storage lipids (mainly known from Bacteria and Eukaryotes) and is a highly abundant organic compound in nature. Accordingly, many Bacteria and Eukarya can utilize glycerol as carbon and energy source as depicted in Fig. 4A.

In general, for degradation glycerol needs to be transported across the cytoplasmic membrane which often involves facilitated diffusion via aqua(glycerol)porins, facilitating the rapid equilibration of glycerol concentration gradients across the membrane [43]. In Bacteria like *E. coli* these glycerol-uptake facilitator (GUF) proteins are encoded by the *glpF* gene [44]. In Bacteria, alternative glycerol transporters have (so far) only been described for *Mycoplasma mycoides* and *M. gallisepticum* in the form of ABC transporters (in addition to a GlpF homologue) [45,46], while in Eukarya, additionally glycerol/H⁺ or glycerol/Na⁺ symporters have been found (e.g. in yeast) [47,48]. Furthermore, glycerol as a small uncharged molecule is also able to enter the cell via passive diffusion [44].

Once transported, glycerol metabolism in Bacteria and Eukaryotes (e.g. *E. coli*, *P. aeruginosa*, and yeast) phosphorylation and dehydrogenation steps ultimately convert glycerol into dihydroxyacetone-phosphate (DHAP): The most widespread route as described is utilized (mainly) by respiring organisms [48-50]. Here glycerol is first converted into *sn*-glycerol-3-P (G3P) via a glycerol kinase (GK) (encoded by the *glpK* gene) using ATP as phosphoryl donor [51]. Following phosphorylation, one of two membrane bound glycerol-3-phosphate dehydrogenases (G3PDH), i.e., GlpD and GlpABC, catalyse the oxidation of G3P to dihydroxyacetone phosphate (DHAP) via the reduction of a non-covalently enzyme bound flavin adenine dinucleotide (FAD) forming FADH₂. The electrons are then further transferred from FADH₂ to the quinone pool of the respiratory chain [52-55]. The G3PDH encoded by the *glpD* gene is maximally expressed under oxic conditions (in *E. coli*) [50,52-54]. The second G3PDH (GlpABC) mainly known from Bacteria is encoded by three different genes, *glpABC*, organized as an operon [56] and the A and B subunits (encoded by *glpA* and *B*) form a soluble and active dimer [55] which is likely anchored to the membrane via the C subunit (encoded by

glpC) [56,57]. *GlpABC* expression is induced under anoxic conditions when glycerol serves as carbon and energy source (in *E. coli*) [50,56]. A third mechanism of G3P oxidation is less widespread among Bacteria and mainly restricted to aerotolerant/microaerophilic lactic acid bacteria. These organisms employ the G3P oxidase (GlpO, encoded by the *glpO* gene), which directly reduces oxygen as electron acceptor yielding H₂O₂ [58-61].

A second route of glycerol processing is less abundant and restricted to organisms growing under fermentative conditions as described for some Enterobacteriaceae including *E. coli* and few other bacterial species [62,63]. Here, instead of a quinone a metabolic intermediate is required to serve as an electron sink for glycerol oxidation. In this pathway glycerol is first oxidized to dihydroxyacetone (DHA) using NAD⁺ by a glycerol dehydrogenase encoded by the *gldA* gene followed by PEP (*dhaK*) or ATP (*glpK*, *Klebsiella pneumoniae*) dependent phosphorylation forming DHAP [50,62] (Fig. 4). The formed NADH is then reoxidized with another molecule of glycerol via the formation of 3-hydroxypropionaldehyde and 1,3-propandiol [62].

Whereas the glycerol metabolism is well understood in Bacteria and Eukarya, comparably little is known regarding glycerol degradation in Archaea. The haloarchaeon *H. volcanii* has been shown to utilize glycerol as carbon and energy source and to employ homologues of the bacterial GlpK and GlpABC proteins, although the proteins have not been studied in detail [64-67] and GK activity has been reported in crude extracts of *H. volcanii* and *Halobacterium* sp. [65,68]. However, other representatives from the archaeal domain have not been shown to grow with glycerol although GK activity has been shown in crude extracts from *Thermoplasma* strains and *Pyrococcus* strains. [68] and genes encoding putative GKs or G3PDHs were identified in various Halobacteriales [67,69] and in aerobic and anaerobic representatives of the orders Thermococcales, Thermoplasmata, Thermoproteales and Lokiarchaeota [69]. For *Thermococcus kodakarensis*, induction of a G3PDH homologue in the stationary phase of growth was reported [70]. However, although both a recombinant GK (also homologous to the bacterial/eukaryal GlpK) and G3PDH have been characterized from *T. kodakarensis* [71-74], the organism could not be grown with glycerol as carbon and energy source [73].

In this work it was demonstrated that the thermoacidophilic crenarchaeal model organism *S. acidocaldarius* grows with glycerol as sole source of carbon and energy, while previously *S. acidocaldarius* has been indicated to not grow on glycerol [17] and only GK homologues have been identified whereas G3PDH homologues appeared absent [69]. Together transcriptomic, proteomic and metabolomic analysis and crude extract enzyme measurements indicated that *S. acidocaldarius* uses a genecluster (*saci_2031-saci_2034*) for utilization of

glycerol. The gene cluster encodes a putative G3PDH (Saci_2032) with a downstream located carbon monoxide dehydrogenase subunit G (CoxG)-like protein (Saci_2031), as well as a putative GK (Saci_2033) and a glycerol uptake facilitator (GUF, Saci_2034). To confirm their activities Saci_2033 and Saci_2032 were recombinantly produced, purified, and characterized. For the homodimeric Saci_2033, GK activity was confirmed as glycerol dependent formation of ADP from ATP. Furthermore, with glyceraldehyde (GA) and dihydroxyacetone (DHA) the enzyme showed minor activities with catalytic efficiencies of <2% compared to glycerol similar to the *E. coli* GK [75] and *T. kodakarensis* GK [73]. The recombinant *S. acidocaldarius* G3PDH (Saci_2032) was shown to catalyse the oxidation of G3P transferring electrons to DCPIP or Ubiquinone Q1 as artificial electron acceptors and was found to contain one non-covalently bound FAD cofactor per monomer. Under native conditions Saci_2032 showed a homodimeric structure (α_2), similar to observations made for bacterial GlpDs e.g. from *E. coli* [54]. In addition, the *S. acidocaldarius* G3PDH was found to be highly specific for glycerol and showed no activity towards G1P, glycerol, glyceric acid, glyceraldehyde-3-phosphate, and phosphoglyceric acid.

To allow for electron transfer to the quinone of the respiratory chain, the FAD dependent G3PDHs need to interact with the membrane. The GlpA from *E. coli*, forming a catalytically active soluble dimer with its GlpB subunit, is likely anchored to the membrane via its GlpC subunit, and this membrane binding via GlpC mediates electron transfer to the quinone pool [57]. In contrast GlpD as a monotopic enzyme directly interacts with the membrane. Interestingly, upon heterologous overexpression in *E. coli*, Saci_2032 was nearly exclusively found in the cytoplasmic and not in the membrane fraction already suggesting that Saci_2032 cannot interact with the membrane alone raising the question of membrane anchoring mechanism of the G3PDH in *S. acidocaldarius*. Interestingly, *saci_2031* (as mentioned above directly encoded downstream of and cotranscribed with *saci_2032*) is annotated to encode a CoxG homologue which was reported as membrane anchor for the carbon monoxide dehydrogenase from *Oligotropha carboxidovorans*. In *S. acidocaldarius*, Saci_2031 (as a HA-tagged fusion protein) was found to a high degree in the insoluble/membrane fraction from which it could be solubilized using n-dodecyl β -D-maltoside (DDM). Furthermore, interaction of Saci_2031 and Saci_2032 was confirmed by coimmunoprecipitation using HA-tagged CoxG (Saci_2031). Mass spectrometry (MS) of interacting proteins revealed a specific interaction of Saci_2031 with Saci_2032. Thus, these results strongly indicate that the G3PDH transfers the electrons from G3P oxidation to the quinone pool of the respiratory chain (caldariellaquinone in *S. acidocaldarius*) and that the required membrane interaction is mediated by the CoxG homologue Saci_2031. However, the quinone reactivity of the Saci_2032 G3PDH dimer alone indicate that CoxG is not involved in electron transfer. Thus, the recruitment of the GlpA-like

Saci_2032 represents a novel function of CoxG homologues in Archaea and an unusual mechanism of membrane anchoring of G3PDHs in *S. acidocaldarius*.

In total, this work demonstrates that *S. acidocaldarius* utilizes glycerol as growth substrate and employs a conserved “classical” GK (homologous to GlpK) for glycerol phosphorylation. However, G3P oxidation is catalysed by a GlpA-subunit like FAD-dependent G3PDH (Saci_2032) lacking the B and C subunits of the classical, bacterial GlpABC complex. Instead, it shows an unusual type of membrane association facilitated by a small CoxG-related protein. The glycerol metabolism elucidated in *S. acidocaldarius* is illustrated in in Fig. 4 in comparison to what is known especially in Bacteria.

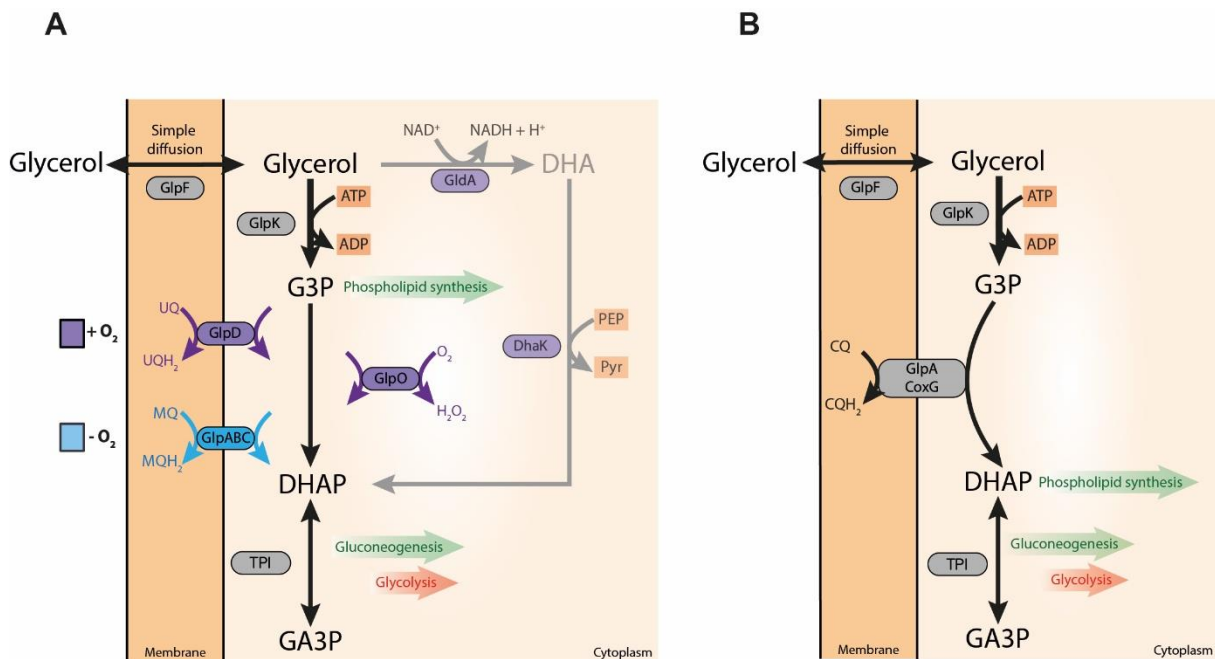


Fig. 4: Biochemical pathways involved in glycerol conversion in Bacteria (e.g., *E. coli*, *Streptococcus* or *P. putida*) and Archaea forming dihydroxyacetone phosphate (DHAP) which is channeled into central metabolism. (A) Glycerol metabolism starts with its uptake either via facilitated diffusion mediated by the glycerol uptake facilitator (GlpF) (encoded by the *glpF* gene) or via simple diffusion through the cytoplasmic membrane. Following uptake, glycerol is converted finally to DHAP via two basic pathways: In respiring organisms, glycerol is first phosphorylated by the glycerol kinase (GlpK) (encoded by the *glpK* gene) to G3P which is further oxidized by two different membrane bound FAD dependent G3P dehydrogenases (G3PDH), i.e., GlpD (encoded by the *glpD* gene) and GlpABC (encoded by the *glpA*, *B*, and *C* genes). Electrons are transferred via the G3PDH bound FAD cofactor to ubiquinone (UQ) by GlpD or to menaquinone (MQ) (GlpABC) of the respiratory chain. A third mechanism of G3P oxidation mainly known from aerotolerant lactic acid bacteria as well as *Mycoplasma* species is catalysed by a soluble, cytoplasmic FAD dependent G3P oxidase (GlpO, encoded by the *glpO* gene) which directly utilizes molecular oxygen as electron acceptor. The second basic glycerol converting pathway is known from fermentatively growing organisms. Here, glycerol is first oxidized via an NAD⁺ dependent G3PDH to dihydroxyacetone (GldA, encoded by the *gldA* gene) which is subsequently phosphorylated by dihydroxyacetone kinase (DhaK, encoded by *dhaK* gene) with phosphoenolpyruvate as phosphoryl donor. (B) Glycerol metabolism in *S. acidocaldarius* likely involves the GlpF-like glycerol uptake facilitator Saci_2034 and eventually simple diffusion through the cytoplasmic membrane. Intracellularly, glycerol is phosphorylated by GK (Saci_2033), homologous to the bacterial GlpK, forming G3P which is then oxidized by a membrane bound, unusual GlpA-like FAD dependent G3PDH (Saci_2032). Saci_2032 is anchored to the membrane by the CoxG homologue Saci_2031 forming a complex that transfers electrons to the caldariellaquinone (CQ) yielding dihydroxyacetone phosphate (DHAP).

1.6 Metabolic thermoadaptation in (hyper)thermophilic Archaea

As a (hyper)thermophilic Archaea such as *S. acidocaldarius* thrive at high temperatures above 80 °C and possess modified central metabolic pathways often involving promiscuous enzyme (see section 1.1). These organisms are particularly challenged by the so-called underground metabolism which denotes the formation of byproducts either (i) enzymatically since most enzymes are promiscuous and catalyse the conversion of alternative substrates (at low efficiency) or they catalyse alternative reactions in addition to their physiological conversions [76] or (ii) by spontaneous chemical reactions at high temperature [78-81]. Hence, the question arises how extremely thermophilic Archaea can operate their unusual metabolism at temperatures where many pathway intermediates are unstable. Chapter 3.3 discusses adaptation strategies of (hyper)thermophilic Archaea such as *S. acidocaldarius* to avoid these kinds of reactions and to dispose unstable and/or toxic products.

Many metabolites in the cell are unstable and reactive and therefore challenge cellular metabolism e.g., triose phosphates, erythrose 4-phosphate, different cofactors (e.g. NAD(P)H, thiamin diphosphate, riboflavin intermediates, tetrahydrofolates and carbamoyl phosphates), cysteine, homocysteine, glutathione, and glutamine belong to such unstable intermediates [82]. In addition, stress conditions such as heat or oxidative stress can lead to metabolite and enzyme damage and thus promote the formation of by-products [81]. These challenges posed by the underground metabolism even increase with increasing temperatures since the velocity of chemical reactions increase exponentially with temperature (according to the Arrhenius equation), and thus temperature has a dramatic effect on metabolite stability/damage [79].

However, organisms have evolved different mechanisms for mitigation including e.g. metabolite repair via specific enzymes and salvage pathways or pre-emption of damaged metabolites by degradation or export (overflow metabolism) [76,78,79,83,84]. Hence, it is assumed that many unidentified compounds in metabolomics profiles are the results of damaged metabolites and that many genes of unknown function in genomes encode proteins dedicated to metabolite damage-control systems [85]. Despite the essential importance, so far only few studies address this question. Notably, genome-scale metabolic modelling of, for example, *S. solfataricus*, a close relative of *S. acidocaldarius*, revealed that 16% of known cellular reactions had to be included without genome evidence [86] suggesting that other enzymes or pathways catalyse these reactions, or that non-catalysed reactions have replaced the enzyme's role [78]. However so far, only a few repair enzymes and salvaging pathways have been investigated in Archaea. In total it seems like (hyper)thermophilic Archaea employ

three base strategies for mitigation: (i) The concentration of unstable metabolites is reduced; (ii) different pathway topologies allow to circumvent labile intermediates, and (iii) damaged metabolites are continuously removed via novel metabolic pathways. Further findings suggest a significant influence of thermolabile intermediates as well as of promiscuous enzymes and pathways on the evolution of the metabolic network in (hyper)thermophilic Archaea.

Part II

1.7 Screening for novel biocatalysts secreted by the model white rot fungus *P. chrysosporium* using ABPP

As mentioned, in the preface, the second part of this work showcases the potential of the ABPP biocatalyst screening technology in complex and biotechnologically relevant experimental setting for studying enzyme activity under various physiological conditions, ABPP, for the discovery of new lignocellulose degrading enzymes. ABPP has emerged as a widely used chemical proteomics method for basic biology research [87-91]. During ABPP, small molecule mechanism-based inhibitors termed activity-based probes (ABPs), form a covalent, irreversible bond with a class of target proteins. This interaction then allows to label, identify, and report on active enzymes under native, physiological conditions as depicted in Fig. 5. To this end, ABPs consist of a reactive warhead, that often ensures a high enzyme class selectivity, a linker and a reporter tag such as biotin for purification and identification, fluorophores for in-gel detection or so-called 2-step reporters such as alkyne or azide moieties [92]. The warhead of ABPs is often based on established inhibitors, while the linker mostly consists of an inert carbon chain (e.g., polyethylene glycol) that assures solubility of the probe and additionally reduces the steric hindrance during labeling reactions.

One of the first ABPs was developed as a derivative of fluorophosphonate (FP) linked to a biotin moiety for affinity enrichment [93]. The Probe was termed FP-biotin and allowed for a versatile screening of serine hydrolases (SHs) under a wide range of different conditions. In the last years, intensive efforts have been undertaken, both to develop new ABPs targeting new enzyme families and to establish their use in, amongst others, drug discovery [94-99], plant biology [100] as well as microbiology [101-105]. One of the main improvements of new ABPs compared to its original probe was the introduction of alkyne or azide moieties replacing the original biotin affinity reporter group. These groups allow for a two-step procedure in which enzymes are first labeled without a bulky reporter tag. After *in vivo* labeling of targeted enzymes, a reporter moiety can then be attached to the original probe via click chemistry.

New lignocellulose degrading enzymes are searched after since lignocellulosic waste material from agriculture such as wood chips or saw dust consists of polymers such as cellulose, hemicellulose and lignin that are a promising alternative to fossil fuel-based substances or classical microbial feedstocks. However conventional use of these waste products as a microbial feedstock is preceded by thermochemical treatment and these pretreatment steps are both expensive and can lead to the formation of toxic byproducts. In this regards, novel

biocatalyst for the degradation of lignocellulosic waste products, consisting of cellulose, hemicellulose and lignin, are urgently sought to allow for easier/gentler and less expensive breakdown. One advanced screening methods for the discovery of such enzymes is ABPP which allows the identification of specifically active enzymes in a sample of interest and allows for the discovery of novel biocatalysts that share no sequence similarity to known counterparts.

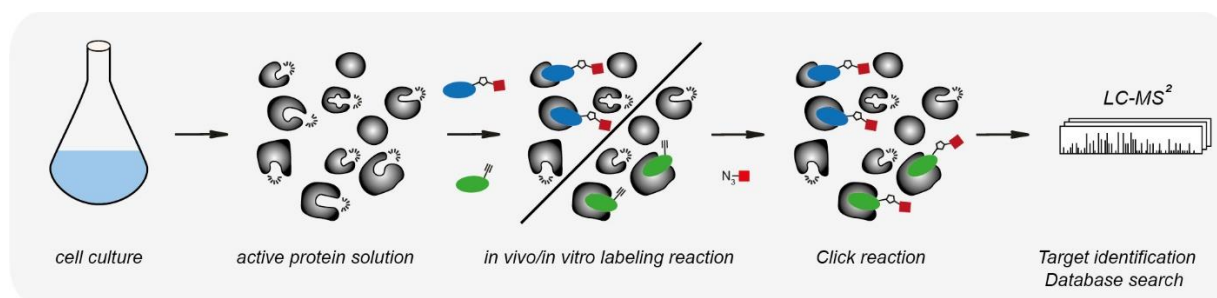


Figure 5: Example overview of an ABPP workflow for identifying target enzymes from protein samples. An ABP is added to active protein solution, followed by a standard ABPP workflow, consisting of click-attachment of a biotin-residues, affinity enrichment of labeled enzymes, trypsin digestion and subsequent MS-based protein identification. The use of enzyme-class specific ABPs therefore results in targeted identification of active biocatalysts which enables the sequence-independent identification of novel biocatalysts without similarity to known homologous.

Cellulose is a linear homopolymer composed of glucose subunits linked together by β -1,4-glycosidic bonds while hemicellulose, with its major component xylan, is a polysaccharide formed from monomeric sugars and sugars acids such as D-xylose, D-mannose, D-galactose, D-glucose, L-arabinose, 4-O-methyl-glucuronic, D-galacturonic, and D-glucuronic acids that are connected by β -1,4- and β -1,3-glycosidic bonds [106]. In contrast, lignin is a complex amorphous heteropolymer composed of phenylpropane derivatives linked to each other via C-C and C-O bonds derived from coumaryl alcohol, coniferyl alcohol and sinapyl alcohol [107] and is linked to both hemicellulose and cellulose via hydrogen bonds giving structural support, resistance against microbial degradation and confers resistance to oxidative stress. Different factors lead to the resistance of lignocellulose against chemical and enzymatic degradation such as the high crystallinity (making cellulose insoluble) and degree of polymerization of cellulose which decrease the accessible surface area, which must be modified and/or removed before degradation. Plant cell wall-degrading filamentous fungi have an important role in recycling nutrients in forest ecosystem. They are known to produce a broad variety of extracellular enzymes with diverse catalytic activities for the hydrolysis of lignocellulose-containing raw materials. To overcome the physical and chemical stability of lignocellulose,

fungi must employ large sets of enzymes, which they secrete into the environment during growth. Three groups of fungi with lignocellulose degrading abilities have been described: soft-rot, brown-rot, and white-rot fungi [106]. Of those white-rot fungi can fully decompose lignocellulose and are therefore an interesting source of lignocellulose-active enzymes to supplement commercially available mixtures of hemicellulases and cellulases [108] (Fig. 6).

This work as described in chapter 3.4 showcases an ABPP approach to identify novel enzymes in the thermotolerant white rot fungus *P. chrysosporium* a known effective degrader of dead wood and lignocellulose [109]. *P. chrysosporium* has been the most intensively studied white rot fungus and its enzymes have attracted considerable interest for application in bioprocesses such as pollutant degradation [108] and fiber bleaching [110]. Its genome encodes a large repertoire of lignocellulolytic enzymes consisting of more than 69 different carbohydrate active enzyme (CAZyme) families including a total of 166 glycoside hydrolases (GH), 14 carbohydrate esterases (CE) and 57 glycosyltransferases (GT). This enormous complexity turns this organism into a promising resource for biocatalyst discovery. However, due to its sheer size, the *P. chrysosporium* secretome cannot be explored by systematically expressing all enzymes, especially because fungal biocatalysts are highly modified upon secretion and thus are notoriously hard to heterologously produce [111,112]. Instead, a methodology to preselect on only those enzymes directly involved in lignocellulose degradation is required. Of note, expression of these biocatalysts is regulated by the presence of the lignocellulosic substrate, demanding preselection assays in the presence of insoluble wood chips and thus highly heterogenous conditions.

With the aim to rapidly identify promising lignocellulose degrading enzymes from this complex system via ABPP-based preselection, we applied FP-alkyne, a well-established serine hydrolase (SH)-targeting ABP, and the GH-targeting ABP JJB111 (*N*-alkynyl-cyclophellitol aziridine connected to biotin). This ABPP workflow was applied to all enzymes secreted by *P. chrysosporium* by labeling both soluble proteins located in the culture medium as well as DDM solubilized proteins that were bound to the solid growth substrate in the form of wood chips. For this work a fixed growth time of 5 days was chosen for all experiments since secretion of lignocellulosic enzymes is highly dependent on growth time [113-115]. The workflow allowed for identification of two CEs and twelve GHs (soluble) as well as four CEs and seven GHs (substrate bound). Of those three CEs and ten GHs were competitively inhibited by paraoxon (SH inhibitor) and KY371 (GH inhibitor), respectively. The CE1 family protein Phchr2|126075 (identified by FP-alkyne) was most enriched in the workflow and is homologous to a characterized acetyl-xylan esterase [116].

Expression, purification, and characterization showed Phchr2|126075 to be an acetyl-xylan esterase that is able to cleave pNP-Acetate. Acetyl-xylan esterases play an important role in the hydrolysis of xylan, as the acetyl side-groups can interfere with the approach of enzymes that break the backbone by steric hindrance, and their elimination thus facilitates the action of endo-1,4- β -xylanases [117]. Few acetyl-xylan esterases have been purified and characterized until now and not much is known about their physicochemical properties.

Furthermore, the GH5 Phchr2|2915237 (identified by JJB111) was expressed and characterized as a promiscuous thermostable (up to 70°C) polysaccharide-cleaving β -glucanase that shows activity against both xylan and lichenan as well as against the artificial nitrophenol conjugates *p*-nitrophenyl-glucopyranosid and *p*-nitrophenyl-xylopyranosid. Analysis of degradation products by thin layer chromatography further revealed that Phchr2|2915237 functions as an β -endoglucanase and can therefore contribute to the breakdown of lignocellulose in *P. chrysosporium*. In agreement with our biochemical results GH5 family members are mostly a class of endo-cleaving glycoside hydrolases that are further divided into a total 56 subfamilies with characterized members involved in polysaccharide degradation of e.g., cellulose or xylan.

Finally, ABPP was used to assign functions to proteins of unknown functions. As such during ABPP a four DUF-domain (DUF4964, DUF5127, DUF4965 and DUF1793) containing protein Phchr2|3002168 was identified using a GH specific probe, for which no other characterized homologue has ever been described although previous analysis suggested that fungal DUF5127 proteins could function as glycoside hydrolases [118]. However, Phchr2|3002168 was found to be insoluble during heterologous production in *E. coli*, so instead a homologue protein was identified via Blast (WP_074995790 from *Streptomyces misionensis*, a known degrader of cellulose [119]), and the enzyme was recombinantly produced and characterized. For this protein β -galactosidase activity was confirmed even though the total specific activity was low. Based on the labeling approach, sequence analysis and enzyme assays, this work proposes that these four DUF domain containing proteins such as Phchr2|3002168 or WP_074995790 function as GHs, although their physiological function remains unknown. In conclusion, this ABPP approach helps to overcome a persisting challenge in biocatalyst discovery: the difficulty to link data from functional screenings directly to sequence information. The presented ABPP approach shortcuts this and allows to narrow the analysis to active enzymes targeted by the enzyme selective ABP probe.

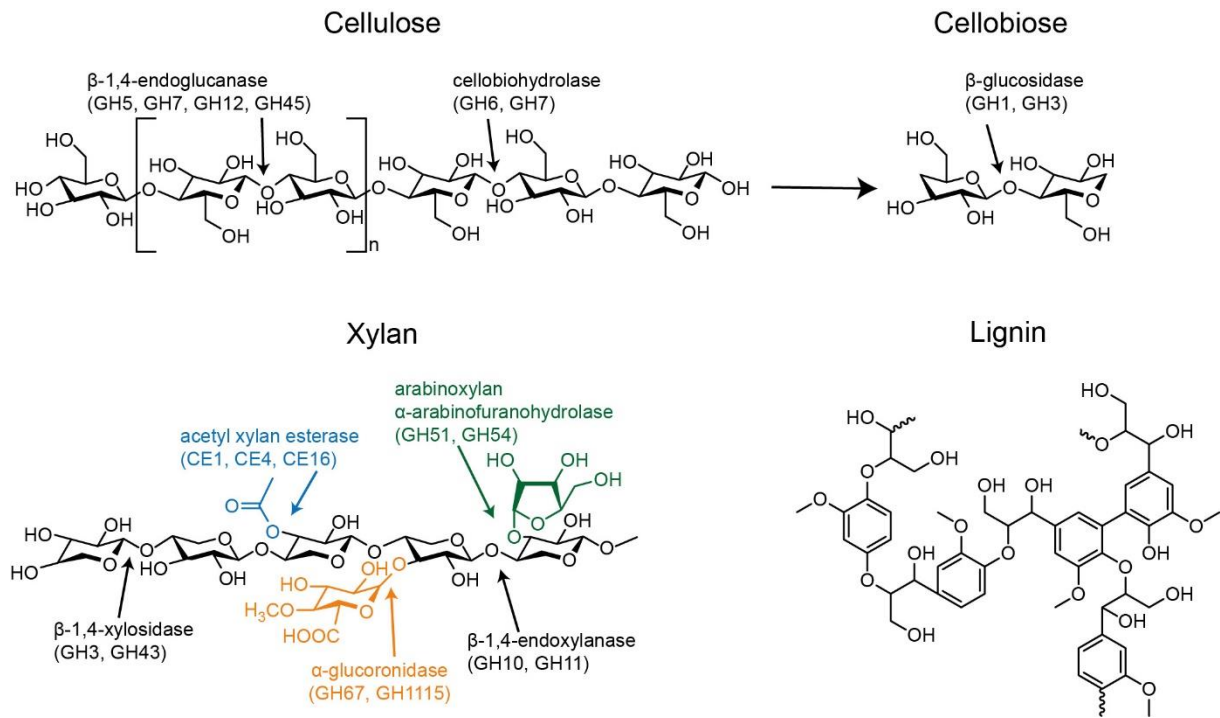


Figure. 6: Components of lignocellulose and the common enzymes involved in their hydrolysis.

Lignocellulose is a complex and recalcitrant polymer built up from cellulose, xylan (hemicellulose) and lignin and its degradation requires the synergistic action of various enzymes as depicted. GH: glycoside hydrolases; CE: carbohydrate esterases.

1.8 References

1. Woese CR, Fox GE: **Phylogenetic structure of the prokaryotic domain: the primary kingdoms.** *Proceedings of the National Academy of Sciences* 1977, **74**:5088-5090.
2. Woese CR, Kandler O, Wheelis ML: **Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya.** *Proceedings of the National Academy of Sciences* 1990, **87**:4576-4579.
3. DeLong EF, Pace NR: **Environmental diversity of bacteria and archaea.** *Systematic biology* 2001, **50**:470-478.
4. Guy L, Ettema TJ: **The archaeal 'TACK'superphylum and the origin of eukaryotes.** *Trends in Microbiology* 2011, **19**:580-587.
5. Eme L, Spang A, Lombard J, Stairs CW, Ettema TJ: **Archaea and the origin of eukaryotes.** *Nature Reviews Microbiology* 2017, **15**:711-723.
6. Spang A, Saw JH, Jørgensen SL, Zaremba-Niedzwiedzka K, Martijn J, Lind AE, Van Eijk R, Schleper C, Guy L, Ettema TJ: **Complex archaea that bridge the gap between prokaryotes and eukaryotes.** *Nature* 2015, **521**:173-179.
7. Grohmann D, Werner F: **Recent advances in the understanding of archaeal transcription.** *Current Opinion in Microbiology* 2011, **14**:328-334.
8. Bräsen C, Esser D, Rauch B, Siebers B: **Carbohydrate metabolism in archaea: Current insights into unusual enzymes and pathways and their regulation.** *Microbiology and Molecular Biology Reviews* 2014, **78**:89-175.
9. Bryant DA, Frigaard N-U: **Prokaryotic photosynthesis and phototrophy illuminated.** *Trends in Microbiology* 2006, **14**:488-496.
10. Schönheit P, Schäfer T: **Metabolism of hyperthermophiles.** *World Journal of Microbiology and Biotechnology* 1995, **11**:26-57.
11. Albers S-V, Meyer BH: **The archaeal cell envelope.** *Nature Reviews Microbiology* 2011, **9**:414-426.
12. Sleytr UB, Schuster B, Egelseer E-M, Pum D: **S-layers: principles and applications.** *FEMS Microbiology Reviews* 2014, **38**:823-864.
13. Caforio A, Driessen AJ: **Archaeal phospholipids: Structural properties and biosynthesis.** *Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids* 2017, **1862**:1325-1339.
14. Chong PL-G: **Archaeobacterial bipolar tetraether lipids: Physico-chemical and membrane properties.** *Chemistry and Physics of Lipids* 2010, **163**:253-265.
15. Tourte M, Schaeffer P, Grossi V, Oger PM: **Functionalized membrane domains: an ancestral feature of Archaea?** *Frontiers in Microbiology* 2020, **11**:526.
16. Siliakus MF, van der Oost J, Kengen SW: **Adaptations of archaeal and bacterial membranes to variations in temperature, pH and pressure.** *Extremophiles* 2017, **21**:651-670.
17. Grogan DW: **Phenotypic characterization of the archaeobacterial genus *Sulfolobus*: Comparison of five wild-type strains.** *Journal of Bacteriology* 1989, **171**:6710-6719
18. Schocke L, Bräsen C, Siebers B: **Thermoacidophilic *Sulfolobus* species as source for extremozymes and as novel archaeal platform organisms.** *Current Opinion in Biotechnology* 2019, **59**:71-77.
19. Lewis AM, Recalde A, Bräsen C, Counts JA, Nussbaum P, Bost J, Schocke L, Shen L, Willard DJ, Quax TEF, et al.: **The biology of thermoacidophilic archaea from the order Sulfolobales.** *FEMS Microbiology Reviews* 2021.
20. Zweerink S, Kallnik V, Ninck S, Nickel S, Verheyen J, Blum M, Wagner A, Feldmann I, Sickmann A, Albers SV, et al.: **Activity-based protein profiling as a robust method for enzyme identification and screening in extremophilic Archaea.** *Nature Communications* 2017, **8**:15352.
21. Wang K, Sybers D, Maklad HR, Lemmens L, Lewyllie C, Zhou X, Schult F, Brasen C, Siebers B, Valegard K, et al.: **A TetR-family transcription factor regulates fatty acid metabolism in the archaeal model organism *Sulfolobus acidocaldarius*.** *Nature Communications* 2019, **10**:1542.
22. Dibrova DV, Galperin MY, Mulikidjanian AY: **Phylogenomic reconstruction of archaeal fatty acid metabolism.** *Environmental Microbiology* 2014, **16**:907-918.

23. Kunau W-H, Dommes V, Schulz H: **β -Oxidation of fatty acids in mitochondria, peroxisomes, and bacteria: a century of continued progress.** *Progress in Lipid Research* 1995, **34**:267-342.
24. Garcia Costas AM, Poudel S, Miller AF, Schut GJ, Ledbetter RN, Fixen KR, Seefeldt LC, Adams MWW, Harwood CS, Boyd ES, Peters JW: **Defining electron bifurcation in the electron-transferring flavoprotein family.** *Journal of Bacteriology* 2017, **199**:00440-00417.
25. Fujita Y, Matsuoka H, Hirooka K: **Regulation of fatty acid metabolism in bacteria.** *Molecular Microbiology* 2007, **66**:829-839.
26. Ishikawa M, Tsuchiya D, Oyama T, Tsunaka Y, Morikawa K: **Structural basis for channelling mechanism of a fatty acid β -oxidation multienzyme complex.** *The EMBO Journal* 2004, **23**:2745-2754.
27. Sah-Teli SK, Hynönen MJ, Sulu R, Dalwani S, Schmitz W, Wierenga RK, Venkatesan R: **Insights into the stability and substrate specificity of the *E. coli* aerobic β -oxidation trifunctional enzyme complex.** *Journal of Structural Biology* 2020, **210**:107494.
28. Venkatesan R, Wierenga RK: **Structure of mycobacterial β -oxidation trifunctional enzyme reveals its altered assembly and putative substrate channeling pathway.** *ACS Chemical Biology* 2013, **8**:1063-1073.
29. Liu L, Huber H, Berg IA: **Enzymes catalyzing crotonyl-CoA conversion to acetoacetyl-CoA during the autotrophic CO₂ fixation in *Metallosphaera sedula*.** *Frontiers in Microbiology* 2020, **11**:354.
30. Ghisla S, Thorpe C: **Acyl-CoA dehydrogenases: a mechanistic overview.** *European Journal of Biochemistry* 2004, **271**:494-508.
31. Flamholz A, Noor E, Bar-Even A, Milo R: **eQuilibrator—the biochemical thermodynamics calculator.** *Nucleic acids research* 2012, **40**:D770-D775.
32. Hamerly T, Tripet B, Wurch L, Hettich RL, Podar M, Bothner B, Copié V: **Characterization of fatty acids in crenarchaeota by GC-MS and NMR.** *Archaea* 2015, **2015**.
33. Gattinger A, Schloter M, Munch JC: **Phospholipid etherlipid and phospholipid fatty acid fingerprints in selected euryarchaeotal monocultures for taxonomic profiling.** *FEMS Microbiology Letters* 2002, **213**:133-139.
34. Kolbe M, Besir H, Essen L-O, Oesterhelt D: **Structure of the light-driven chloride pump halorhodopsin at 1.8 Å resolution.** *Science* 2000, **288**:1390-1396.
35. Lombard J, López-García P, Moreira D: **An ACP-independent fatty acid synthesis pathway in archaea: implications for the origin of phospholipids.** *Molecular Biology and Evolution* 2012, **29**:3261-3265.
36. Byers DM, Gong H: **Acyl carrier protein: structure–function relationships in a conserved multifunctional protein family.** *Biochemistry and Cell Biology* 2007, **85**:649-662.
37. Bhaumik P, Koski MK, Glumoff T, Hiltunen JK, Wierenga RK: **Structural biology of the thioester-dependent degradation and synthesis of fatty acids.** *Current Opinion in Structural Biology* 2005, **15**:621-628.
38. White SW, Zheng J, Yong-Mei Z, Rock CO: **The structural biology of type II fatty acid biosynthesis.** *Annual Review of Biochemistry* 2005, **74**:791.
39. Vögeli B, Engilberge S, Girard E, Riobé F, Maury O, Erb TJ, Shima S, Wagner T: **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**:3380-3385.
40. Sacco E, Covarrubias AS, O'Hare HM, Carroll P, Eynard N, Jones TA, Parish T, Daffé M, Bäckbro K, Quémard A: **The missing piece of the type II fatty acid synthase system from *Mycobacterium tuberculosis*.** *Proceedings of the National Academy of Sciences* 2007, **104**:14628-14633.
41. Pidugu LS, Maity K, Ramaswamy K, Suroliya N, Suguna K: **Analysis of proteins with the 'hot dog' fold: Prediction of function and identification of catalytic residues of hypothetical proteins.** *BMC Structural Biology* 2009, **9**:1-16.
42. Lefebvre C, Frigui W, Slama N, Lauzeral-Vizcaino F, Constant P, Lemassu A, Parish T, Eynard N, Daffé M, Brosch R: **Discovery of a novel dehydratase of the fatty acid synthase type II critical for**

- ketomycolic acid biosynthesis and virulence of *Mycobacterium tuberculosis*.** *Scientific Reports* 2020, **10**:2112.
43. Stroud RM, Miercke LJ, O'Connell J, Khademi S, Lee JK, Remis J, Harries W, Robles Y, Akhavan D: **Glycerol facilitator GlpF and the associated aquaporin family of channels.** *Current Opinion in Structural Biology* 2003, **13**:424-431.
 44. Richey DP, Lin EC: **Importance of facilitated diffusion for effective utilization of glycerol by *Escherichia coli*.** *Journal of Bacteriology* 1972, **112**:784-790.
 45. Blötz C, Stülke J: **Glycerol metabolism and its implication in virulence in *Mycoplasma*.** *FEMS Microbiology Reviews* 2017, **41**:640-652.
 46. Mahdizadeh S, Masukagami Y, Tseng C-W, Markham PF, Souza DPD, Nijagal B, Tull D, Tatarczuch L, Browning GF, Sansom FM, et al.: **A *Mycoplasma gallisepticum* glycerol ABC transporter involved in pathogenicity.** *Applied and Environmental Microbiology* 2021, **87**:03112-03120.
 47. Wille U, Schade B, Duszenko M: **Characterization of glycerol uptake in bloodstream and procyclic forms of *Trypanosoma brucei*.** *European Journal of Biochemistry* 1998, **256**:245-250.
 48. Klein M, Swinnen S, Thevelein JM, Nevoigt E: **Glycerol metabolism and transport in yeast and fungi: established knowledge and ambiguities.** *Environmental Microbiology* 2017, **19**:878-893.
 49. Lin ECC: **Glycerol dissimilation and its regulation in bacteria.** *Annual Review of Microbiology* 1976, **30**:535-578.
 50. Poblete-Castro I, Wittmann C, Nikel PI: **Biochemistry, genetics and biotechnology of glycerol utilization in *Pseudomonas* species.** *Microbial Biotechnology* 2020, **13**:32-53.
 51. Yeh JI, Charrier V, Paulo J, Hou L, Darbon E, Claiborne A, Hol WG, Deutscher J: **Structures of enterococcal glycerol kinase in the absence and presence of glycerol: correlation of conformation to substrate binding and a mechanism of activation by phosphorylation.** *Biochemistry* 2004, **43**:362-373.
 52. Schryvers A, Lohmeier E, Weiner JH: **Chemical and functional properties of the native and reconstituted forms of the membrane-bound, aerobic glycerol-3-phosphate dehydrogenase of *Escherichia coli*.** *Journal of Biological Chemistry* 1978, **253**:783-788.
 53. Yeh JI, Chinte U, Du S: **Structure of glycerol-3-phosphate dehydrogenase, an essential monotopic membrane enzyme involved in respiration and metabolism.** *Proceedings of the National Academy of Sciences of the United States of America* 2008, **105**:3280-3285.
 54. Austin D, Larson TJ: **Nucleotide sequence of the glpD gene encoding aerobic sn-glycerol 3-phosphate dehydrogenase of *Escherichia coli* K-12.** *Journal of Bacteriology* 1991, **173**:101-107.
 55. Schryvers A, Weiner JH: **The anaerobic sn-glycerol-3-phosphate dehydrogenase of *Escherichia coli*. Purification and characterization.** *Journal of Biological Chemistry* 1981, **256**:9959-9965.
 56. Cole ST, Eiglmeier K, Ahmed S, Honore N, Elmes L, Anderson WF, Weiner JH: **Nucleotide sequence and gene-polypeptide relationships of the glpABC operon encoding the anaerobic sn-glycerol-3-phosphate dehydrogenase of *Escherichia coli* K-12.** *Journal of Bacteriology* 1988, **170**:2448-2456.
 57. Varga MER, Weiner JH: **Physiological role of GlpB of anaerobic glycerol-3-phosphate dehydrogenase of *Escherichia coli*.** *Biochemistry and Cell Biology* 1995, **73**:147-153.
 58. Colussi T, Parsonage D, Boles W, Matsuoka T, Mallett TC, Karplus PA, Claiborne A: **Structure of α -glycerophosphate oxidase from *Streptococcus* sp.: A template for the mitochondrial α -glycerophosphate dehydrogenase.** *Biochemistry* 2008, **47**:965-977.
 59. Finnerty CM, Charrier V, Claiborne A, Karplus PA: **Crystallization and preliminary crystallographic analysis of the soluble α -glycerophosphate oxidase from *Streptococcus* sp.** *Acta Crystallographica* 2002, **58**:165-166.
 60. Zotta T, Parente E, Ricciardi A: **Aerobic metabolism in the genus *Lactobacillus*: impact on stress response and potential applications in the food industry.** *Journal of Applied Microbiology* 2017, **122**:857-869.

61. Parsonage D, Luba J, Mallett TC, Claiborne A: **The Soluble α -Glycerophosphate oxidase from *Enterococcus casseliflavus*: sequence homology with the membrane-associated dehydrogenase and kinetic analysis of the recombinant enzyme.** *Journal of Biological Chemistry* 1998, **273**:23812-23822.
62. Clomburg JM, Gonzalez R: **Anaerobic fermentation of glycerol: a platform for renewable fuels and chemicals.** *Trends in Biotechnology* 2013, **31**:20-28.
63. Clomburg JM, Cintolesi A, Gonzalez R: ***In silico* and *in vivo* analyses reveal key metabolic pathways enabling the fermentative utilization of glycerol in Escherichia coli.** *Microbial Biotechnology* 2022, **15**:289-304.
64. Rawls KS, Yacovone SK, Maupin-Furlow JA: **GlpR represses fructose and glucose metabolic enzymes at the level of transcription in the haloarchaeon *Haloferax volcanii*.** *Journal of Bacteriology* 2010, **192**:6251-6260.
65. Sherwood KE, Cano DJ, Maupin-Furlow JA: **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**:4307-4315.
66. Rawls KS, Martin JH, Maupin-Furlow JA: **Activity and transcriptional regulation of bacterial protein-like glycerol-3-phosphate dehydrogenase of the Haloarchaea in *Haloferax volcanii*.** *Journal of Bacteriology* 2011, **193**:4469-4476.
67. Williams TJ, Allen M, Tschitschko B, Cavicchioli R: **Glycerol metabolism of haloarchaea.** *Environmental Microbiology* 2017, **19**:864-877.
68. Nishihara M, Yamazaki T, Oshima T, Koga Y: **sn-glycerol-1-phosphate-forming activities in Archaea: separation of archaeal phospholipid biosynthesis and glycerol catabolism by glycerophosphate enantiomers.** *Journal of Bacteriology* 1999, **181**:1330-1333.
69. Villanueva L, Schouten S, Damsté JSS: **Phylogenomic analysis of lipid biosynthetic genes of Archaea shed light on the 'lipid divide'.** *Environmental Microbiology* 2017, **19**:54-69.
70. Gagen EJ, Yoshinaga MY, Garcia Prado F, Hinrichs K-U, Thomm M: **The proteome and lipidome of *Thermococcus kodakarensis* across the stationary phase.** *Archaea* 2016, **2016**:5938289.
71. Hokao R, Matsumura H, Katsumi R, Angkawidjaja C, Takano K, Kanaya S, Koga Y: **Affinity shift of ATP upon glycerol binding to a glycerol kinase from the hyperthermophilic archaeon *Thermococcus kodakarensis* KOD1.** *Journal of Bioscience and Bioengineering* 2020, **129**:657-663.
72. Koga Y, Katsumi R, You D-J, Matsumura H, Takano K, Kanaya S: **Crystal structure of highly thermostable glycerol kinase from a hyperthermophilic archaeon in a dimeric form.** *The FEBS Journal* 2008, **275**:2632-2643.
73. Koga Y, Konishi K, Kobayashi A, Kanaya S, Takano K: **Anaerobic glycerol-3-phosphate dehydrogenase complex from hyperthermophilic archaeon *Thermococcus kodakarensis* KOD1.** *Journal of Bioscience and Bioengineering* 2019, **127**:679-685.
74. Koga Y, Morikawa M, Haruki M, Nakamura H, Imanaka T, Kanaya S: **Thermostable glycerol kinase from a hyperthermophilic archaeon: gene cloning and characterization of the recombinant enzyme.** *Protein Engineering, Design and Selection* 1998, **11**:1219-1227.
75. Hayashi S-I, Lin ECC: **Purification and Properties of Glycerol Kinase from *Escherichia coli*.** *Journal of Biological Chemistry* 1967, **242**:1030-1035.
76. Bommer GT, Van Schaftingen E, Veiga-da-Cunha M: **Metabolite repair enzymes control metabolic damage in glycolysis.** *Trends in Biochemical Sciences* 2020, **45**:228-243.
77. D'Ari R, Casadesús J: **Underground metabolism.** *Bioessays* 1998, **20**:181-186.
78. Keller MA, Piedrafita G, Ralser M: **The widespread role of non-enzymatic reactions in cellular metabolism.** *Current Opinion in Biotechnology* 2015, **34**:153-161.
79. Hanson AD, Henry CS, Fiehn O, de Crécy-Lagard V: **Metabolite damage and metabolite damage control in plants.** *Annual Review of Plant Biology* 2016, **67**:131-152.
80. Rosenberg J, Commichau FM: **Harnessing underground metabolism for pathway development.** *Trends in Biotechnology* 2019, **37**:29-37.

81. Piedrafita G, Keller MA, Ralser M: **The impact of non-enzymatic reactions and enzyme promiscuity on cellular metabolism during (oxidative) stress conditions.** *Biomolecules* 2015, **5**:2101-2122.
82. Lerma-Ortiz C, Jeffryes JG, Cooper AJ, Niehaus TD, Thamm AM, Frelin O, Aunins T, Fiehn O, de Crécy-Lagard V, Henry CS, et al.: **'Nothing of chemistry disappears in biology': the Top 30 damage-prone endogenous metabolites.** *Biochemical Society Transactions* 2016, **44**:961-971.
83. Linster CL, Van Schaftingen E, Hanson AD: **Metabolite damage and its repair or pre-emption.** *Nature Chemical Biology* 2013, **9**:72-80.
84. de Crécy-Lagard V, Haas D, Hanson AD: **Newly-discovered enzymes that function in metabolite damage-control.** *Current Opinion in Chemical Biology* 2018, **47**:101-108.
85. Ellens KW, Christian N, Singh C, Satagopam VP, May P, Linster CL: **Confronting the catalytic dark matter encoded by sequenced genomes.** *Nucleic Acids Research* 2017, **45**:11495-11514.
86. Ulas T, Riemer SA, Zaparty M, Siebers B, Schomburg D: **Genome-scale reconstruction and analysis of the metabolic network in the hyperthermophilic archaeon *Sulfolobus solfataricus*.** *PLOS ONE* 2012, **7**:e43401.
87. Sanman LE, Bogoy M: **Activity-based profiling of proteases.** *Annual Review of Biochemistry* 2014, **83**:249-273.
88. Bennis HJ, Wincott CJ, Tate EW, Child MA: **Activity- and reactivity-based proteomics: Recent technological advances and applications in drug discovery.** *Current Opinion in Chemical Biology* 2021, **60**:20-29.
89. Fang H, Peng B, Ong SY, Wu Q, Li L, Yao SQ: **Recent advances in activity-based probes (ABPs) and affinity-based probes (AfBPs) for profiling of enzymes.** *Chemical Science* 2021, **12**:8288-8310.
90. Li N, Overkleeft HS, Florea BI: **Activity-based protein profiling: an enabling technology in chemical biology research.** *Current Opinion in Chemical Biology* 2012, **16**:227-233.
91. Cravatt BF, Wright AT, Kozarich JW: **Activity-based protein profiling: from enzyme chemistry to proteomic chemistry.** *Annual Review of Biochemistry* 2008, **77**:383-414.
92. Verhelst SH, Bongers KM, Willems LI: **Bioorthogonal reactions in activity-based protein profiling.** *Molecules* 2020, **25**:5994.
93. Liu YS, Patricelli MP, Cravatt BF: **Activity-based protein profiling: The serine hydrolases.** *Proceedings of the National Academy of Sciences of the United States of America* 1999, **96**:14694-14699.
94. Niphakis MJ, Cravatt BF: **Enzyme inhibitor discovery by activity-based protein profiling.** *Annual Review of Biochemistry* 2014, **83**:341-377.
95. Deng H, Lei Q, Wu YP, He Y, Li WM: **Activity-based protein profiling: Recent advances in medicinal chemistry.** *European Journal of Medicinal Chemistry* 2020, **191**.
96. Nomura DK, Dix MM, Cravatt BF: **Activity-based protein profiling for biochemical pathway discovery in cancer.** *Nature Reviews Cancer* 2010, **10**:630-638.
97. Kok BP, Ghimire S, Kim W, Chatterjee S, Johns T, Kitamura S, Eberhardt J, Ogasawara D, Xu J, Sukiasyan A, et al.: **Discovery of small-molecule enzyme activators by activity-based protein profiling.** *Nature Chemical Biology* 2020, **16**:997-1005.
98. Parker CG, Galmozzi A, Wang YJ, Correia BE, Sasaki K, Joslyn CM, Kim AS, Cavallaro CL, Lawrence RM, Johnson SR, et al.: **Ligand and target discovery by fragment-based screening in human cells.** *Cell* 2017, **168**:527.
99. Wright MH, Sieber SA: **Chemical proteomics approaches for identifying the cellular targets of natural products.** *Natural Product Reports* 2016, **33**:731-733.
100. Morimoto K, van der Hoorn RA: **The increasing impact of activity-based protein profiling in plant science.** *Plant and Cell Physiology* 2016, **57**:446-461.
101. Keller LJ, Babin BM, Lakemeyer M, Bogoy M: **Activity-based protein profiling in bacteria: Applications for identification of therapeutic targets and characterization of microbial communities.** *Current Opinion in Chemical Biology* 2020, **54**:45-53.
102. Lentz CS: **What you see is what you get: activity-based probes in single-cell analysis of enzymatic activities.** *Biological Chemistry* 2020, **401**:233-248.

103. Krysiak J, Sieber SA: **Activity-based protein profiling in bacteria.** *Methods in Molecular Biology* 2017, **1491**:57-74.
104. Heal WP, Tate EW: **Application of activity-based protein profiling to the study of microbial pathogenesis.** *Activity-Based Protein Profiling* 2012, **324**:115-135.
105. Steuten K, Kim H, Widen JC, Babin BM, Onguka O, Lovell S, Bolgi O, Cerikan B, Neufeldt CJ, Cortese M, et al.: **Challenges for targeting SARS-CoV-2 proteases as a therapeutic strategy for COVID-19.** *Acs Infectious Diseases* 2021, **7**:1457-1468.
106. Sanchez C: **Lignocellulosic residues: biodegradation and bioconversion by fungi.** *Biotechnology Advances* 2009, **27**:185-194.
107. Chen H: **Biotechnology of lignocellulose.** *Theory and Practice. China: Chemical Industry Press and Springer* 2014.
108. Bumpus JA, Aust SD: **Biodegradation of environmental pollutants by the white rot fungus *Phanerochaete chrysosporium*: involvement of the lignin degrading system.** *BioEssays* 1987, **6**:166-170.
109. da Silva RR, Pedezi R, Souto TB: **Exploring the bioprospecting and biotechnological potential of white-rot and anaerobic *Neocallimastigomycota* fungi: peptidases, esterases, and lignocellulolytic enzymes.** *Applied Microbiology and Biotechnology* 2017, **101**:3089-3101.
110. Michel Jr F, Dass SB, Grulke E, Reddy CA: **Role of manganese peroxidases and lignin peroxidases of *Phanerochaete chrysosporium* in the decolorization of kraft bleach plant effluent.** *Applied and Environmental Microbiology* 1991, **57**:2368-2375.
111. Lambertz C, Garvey M, Klinger J, Heesel D, Klose H, Fischer R, Commandeur U: **Challenges and advances in the heterologous expression of cellulolytic enzymes: a review.** *Biotechnology for Biofuels* 2014, **7**:1-15.
112. Zoglowek M, Lübeck PS, Ahring BK, Lübeck M: **Heterologous expression of cellobiohydrolases in filamentous fungi—an update on the current challenges, achievements and perspectives.** *Process Biochemistry* 2015, **50**:211-220.
113. MacDonald J, Suzuki H, Master ER: **Expression and regulation of genes encoding lignocellulose-degrading activity in the genus *Phanerochaete*.** *Applied Microbiology and Biotechnology* 2012, **94**:339-351.
114. Macdonald J, Master ER: **Time-dependent profiles of transcripts encoding lignocellulose-modifying enzymes of the white rot fungus *Phanerochaete carnos* grown on multiple wood substrates.** *Applied and Environmental Microbiology* 2012, **78**:1596-1600.
115. Malgas S, Thoresen M, van Dyk JS, Pletschke BI: **Time dependence of enzyme synergism during the degradation of model and natural lignocellulosic substrates.** *Enzyme and Microbial Technology* 2017, **103**:1-11.
116. Komiya D, Hori A, Ishida T, Igarashi K, Samejima M, Koseki T, Fushinobu S: **Crystal structure and substrate specificity modification of acetyl xylan esterase from *Aspergillus luchuensis*.** *Applied and Environmental Microbiology* **83**:e01251-01217.
117. Biely P, Westereng B, Puchart V, De Maayer P, Cowan DA: **Recent progress in understanding the mode of action of acetylxylan esterases.** *Journal of Applied Glycoscience* 2014, **61**:35-44.
118. Arntzen MO, Bengtsson O, Varnai A, Delogu F, Mathiesen G, Eijsink VGH: **Quantitative comparison of the biomass-degrading enzyme repertoires of five filamentous fungi.** *Scientific Reports* 2020, **10**:20267.
119. Franco-Cirigliano MN, Rezende Rde C, Gravina-Oliveira MP, Pereira PH, do Nascimento RP, Bon EP, Macrae A, Coelho RR: ***Streptomyces misionensis* PESB-25 produces a thermoacidophilic endoglucanase using sugarcane bagasse and corn steep liquor as the sole organic substrates.** *BioMed Research International* 2013, **2013**:584207.

2. Scope of the thesis

In the current dissertation my research topics are described in the following chapters and the contributions I have made for these manuscripts are outlined below.

In chapter 3.1 of the manuscript with the title **“Fatty acid metabolism in *Sulfolobus acidocaldarius*”**, the FA metabolism of *S. acidocaldarius* was extensively studied by the recombinant production, purification, and characterization of relevant enzymes. Here a β oxidation with some unusual features and a novel FA synthesis pathway were reconstructed via HPLC and biochemical approaches in *S. acidocaldarius*. For this manuscript I heterologously produced and purified the proteins Saci_1123, Saci_1085, and Saci_0315 and confirmed their activity using different acyl-CoA substrates and contributed to the characterization of Saci_1114 and Saci_1115. In addition, I performed all HPLC measurements of single enzymes as well as the reconstruction of enzyme cascades *in vitro* and did the bioinformatic analysis of proteins encoded by *saci_1123*, *saci_1113*, *saci_1108* and *saci_0315*.

In chapter 3.2 of the manuscript with the title **“Glycerol degradation in the thermoacidophilic crenarchaeon *Sulfolobus acidocaldarius* involves an unusual glycerol-3-phosphate dehydrogenase”** it was elucidated how glycerol is degraded in *S. acidocaldarius* using multiple omics approaches as well as the recombinant production and characterization of relevant enzymes. Here it was shown that *S. acidocaldarius* employs a pathway for glycerol dissimilation using a GK and unusual G3PDH. For this work I interpreted and analyzed all omics data, analyzed relevant gene clusters and abundant proteins. In addition, I purified and extensively characterized Saci_2032 and established the interaction and membrane localization of Saci_2031 and Saci_2032 and performed part of the structural analysis of different G3PDHs.

Chapter 3.3 with the title **“Enhanced underground metabolism challenges life at high temperatures – metabolic thermoadaptation in hyperthermophilic Archaea”** discusses how Archaea circumvent metabolic problems associated with growth at high temperatures such as formation and accumulation of toxic by/side products. For this work I participated in writing and performed literature search as well as image generation.

In chapter 3.4 of the manuscript with the title: **“Identification of fungal lignocellulose-degrading biocatalysts secreted by *Phanerochaete chrysosporium* via activity-based protein profiling”** we established an ABPP approach in the white-rot fungus *P. chrysosporium* and identified a set of GHs and CEs involved in lignocellulose degradation. For this soluble and substrate bound proteins secreted by *P. chrysosporium* were incubated with different

ABPs, followed by affinity purification and target identification. Purification and characterization of two target proteins confirmed their expected activity and the analysis and expression of a third protein indicated the presence of GH activity in DUF5127 domain proteins. For this work I performed all cultivation experiments and the isolation of active protein fractions. Furthermore, I analyzed and annotated gathered proteome data, to confirm the function of target proteins. For enzymatic measurements I produced and purified Phchr2|126075 and WP_074995790 and performed the characterizations of Phchr2|126075, Phchr2|2915237 and WP_074995790.

3. Manuscripts

Chapter 3.1: Fatty acid metabolism in *Sulfolobus acidocaldarius*

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, #

^aMolecular 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

^bDepartment of Molecular Biosciences, The Wenner-Gren Institute, Stockholm University, SE-106 91, Stockholm, Sweden

^cBio-engineering Sciences, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussels, Belgium

^dChemical 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/enoylCoA hydratase – (KT) β -ketothiolase or acetyl-CoA C acetyltransferase – (ACR) acetoacyl-CoA reductase – (MaoC-HCD) MaoC like 3(R)-hydroxyacyl-CoA dehydratase – (ECR) enoylCoA 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_cPF_6) ferrocenium hexafluorophosphate – (DTT) dithiothreitol – (ACN) acetonitrile

Abstract

Fatty acids are main building blocks for bacterial and eukaryotic membrane lipids and together with other functions e.g., in energy supply and signaling 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)archaeal 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 thiolitically 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 be 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 CO₂ 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 channeling 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 DH5 α for cloning and Rosetta (DE3) for expression were cultivated independently in Luria Bertani (LB) medium containing appropriate antibiotics (100 μ g/ml ampicillin for strains containing plasmid pET15b or pET45b, 50 μ g/ml kanamycin for strains including plasmid pET28b and 50 μ g/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 NaH₂PO₄ 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 flash 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 addition 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 SDSpolyacrylamide 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 $\text{mM}^{-1} \text{cm}^{-1}$) 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 $\text{mM}^{-1} \text{cm}^{-1}$) [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 (acetylCoA, 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 $\text{mM}^{-1} \text{cm}^{-1}$ [35]) while the acylCoA compounds were thus oxidized. The reaction mixture (500 μl) contained 1 mM FcPF_6 , 0.13 $\mu\text{g}/\mu\text{l}$ 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 $\mu\text{g}/\mu\text{l}$ ACAD and 0.01 $\mu\text{g}/\mu\text{l}$ ETF with 0.4

mM of acyl-CoAs (butyryl-CoA, hexanoyl-CoA or octanoyl-CoA) were incubated for 5 min. Moreover, 0.8 mM FcPF₆ was also used as an electron acceptor 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 iodinitrotetrazolium 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 crotonyl-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 3HBCoA 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 crotonyl-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 Mg²⁺ (extinction coefficient of Mg²⁺- 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 MgCl₂, 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.

Acetoacetyl-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) containing 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 µl).

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 FcPF₆, 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 µg 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 EppendorfTM 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 ScientificTM DionexTM 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. 1C). 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 ($\text{FA} + \text{ATP} + \text{HS-CoA} \rightarrow \text{FA-CoA} + \text{AMP} + \text{PP}_i$) [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 PP_i 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 PP_i-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 isoalloxazine-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 homotetrameric 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 Nterminal 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 μ M (crotonyl-CoA) and 36 μ M (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)-hydroxybutyrylCoA to acetoacetyl-CoA with a V_{\max} of 48 U mg⁻¹ and was specific for the (S)-enantiomer (K_m 43 μ M) 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 μ M) over NADPH (K_m 94 μ M) (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 μ M), CoA (V_{\max} 2.5 U mg⁻¹, K_m 3 μ M) 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 thiolistical 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*, *Pseudomonas 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 octanoylCoA to the respective enoyl-CoA esters by ACAD was shown with FcPF₆ 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-hydroxybutyrylCoA) (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 μ M, K_m (NADPH) 7 μ M) (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. acidocaldarius*. 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 homododecameric 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

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 ($C \geq 12$) [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 regarded 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 dinucleotide (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. acidocaldarius* 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 Cterminal, respectively) of known ETFs [54] (Fig. S5) and the monomeric structure is thus in line with the reported heterodimeric ($\alpha\beta$) 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 harbour 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 iodinitrotetrazolium (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 2,3decenoyl-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 ($8\text{-}35 \text{ U mg}^{-1}$ for *M. sedula* and 6 U mg^{-1} for *F. placidus*) [5961] 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 $-11.9 \text{ kJ mol}^{-1}$ (with the thiolase as major driving force of $-26.1 \text{ kJ mol}^{-1}$), the NAD^+ dependent hydroxyacyl-CoA oxidation is endergonic by $+18 \text{ kJ mol}^{-1}$ 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-3methylglutaryl-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 ($\Delta G^0 = -3.3 \text{ kJ mol}^{-1}$). The addition of NAD^+ did not result in any detectably acetoacetyl-CoA formation which is also in accordance with the ΔG^0 of $+18 \text{ kJ mol}^{-1}$ which makes the two-step crotonoyl- to acetoacetyl-CoA conversion endergonic by nearly $+15 \text{ kJ mol}^{-1}$. 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 $-26.1 \text{ kJ mol}^{-1}$ 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 $+100 \text{ mV}$ [52], that of crotonoyl-CoA/butyryl-CoA -10 mV [75]) as electron acceptor (and finally oxygen ($E^0 +820 \text{ mV}$)), 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 butyrylCoA 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.

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° -320 mV) as electron donor which renders the reaction highly exergonic towards saturated acyl thioester formation (crotonoyl-CoA/butyrylCoA E° -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 harbours 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 perfectly 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 nondecarboxylative 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 CO₂ 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 enoylCoA 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 distributa*), 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-hydroxybutyrylCoA, 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.

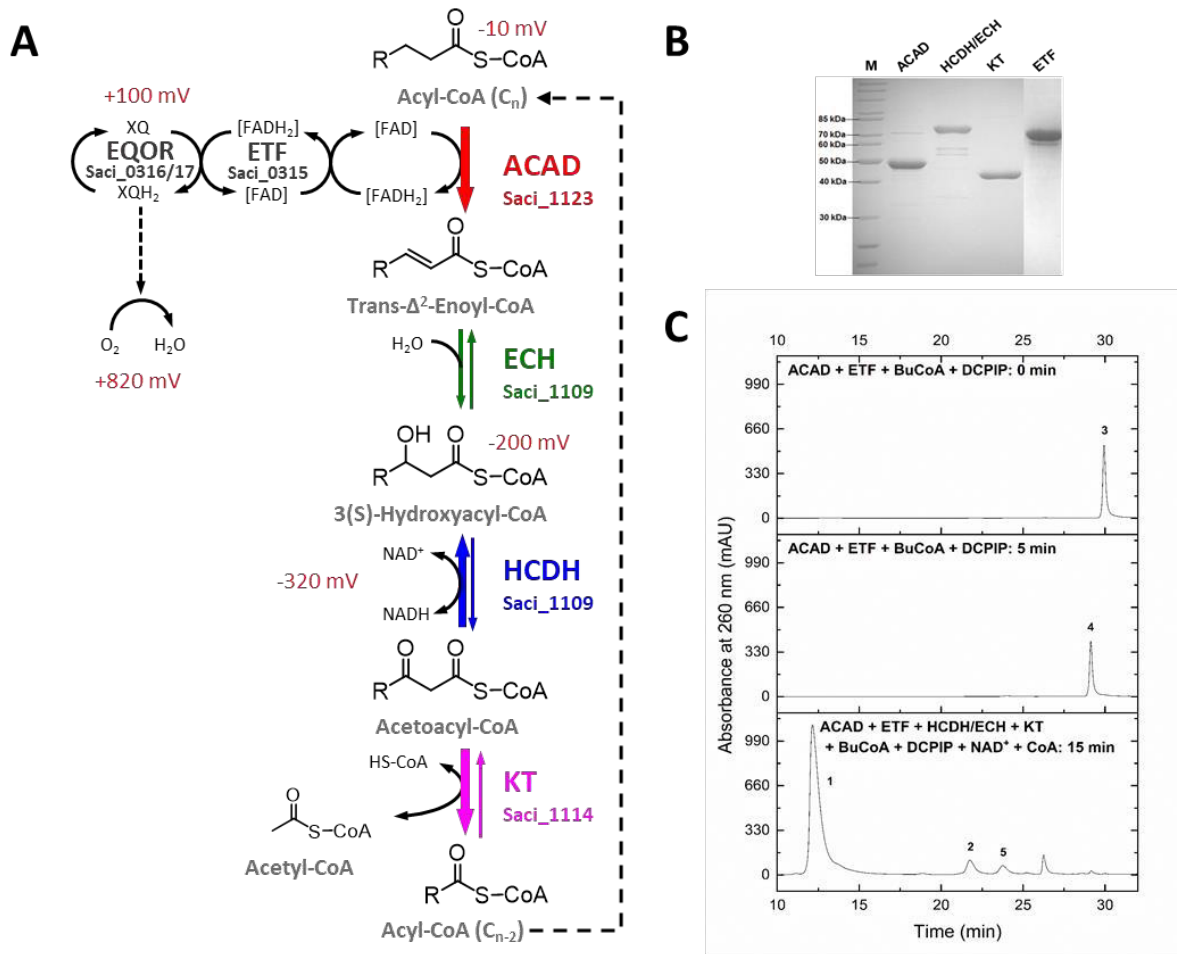


Figure 1. **Reconstructed β oxidation pathway for FA degradation in *S. acidocaldarius*.** The β oxidation reactions (A), the purified recombinant proteins that catalyze these reactions (SDS-PAGE and Coomassie Blue staining) (B) as well as the HPLC chromatogram of the β oxidation enzyme cascade for butyryl-CoA 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.

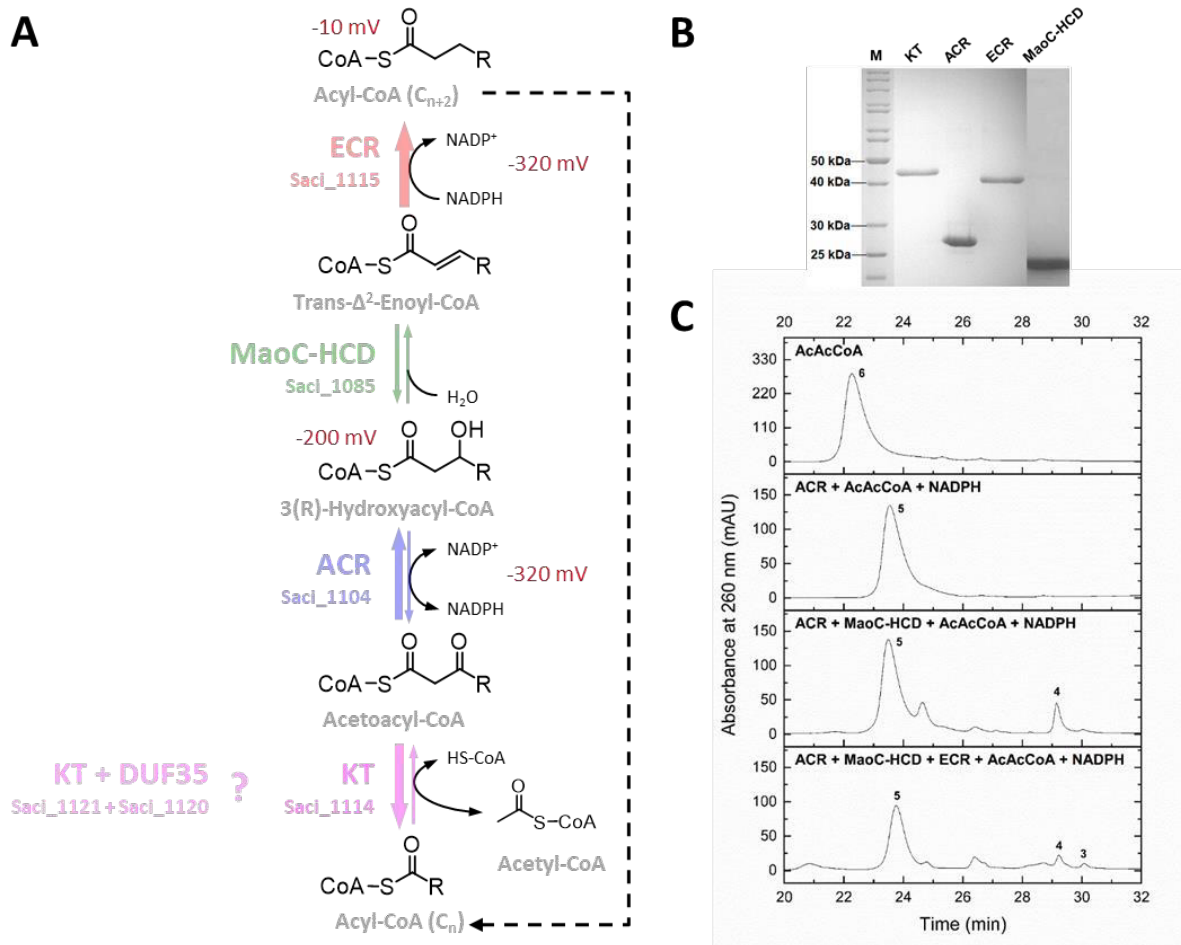


Figure 2. Proposed FA biosynthetic pathway in *S. acidocaldarius*. The FA synthetic reactions (A), the purified recombinant proteins that catalyze these conversions (SDS-PAGE and Coomassie Blue staining) (B) as well as the enzyme cascade for butyryl-CoA synthesis via HPLC (C) are shown. The abbreviations 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-ketoacylCoA 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)	Temp	Substrate/Cofactor	Km (mM)	Vmax (U mg ⁻¹)	kcat (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	ND	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 Enoyl-CoA hydratase/3-hydroxyacylCoA 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
NADH	0.028 ± 0.007					
NADPH	*0.094 ± 0.22					
Saci_1114 β-Ketothiolase (KT)	88 (homodimer)	23°C	Acetoacetyl-CoA (C4)	0.033 ± 0.01	1.676 ± 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)	Temp	Substrate/Cofactor	Km (mM)	Vmax (U mg ⁻¹)	kcat (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 (homododecamer)	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 octenoyl-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 archaeobacterium 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 fatty 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**: 815820
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 Methanotheroxobacterium*. Journal of bacteriology, 1989. **171**(10): 5430-5435.
47. Bräsen, C. and Schönheit, 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. 35565.
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 Schäfer, G., *Cytochrome aa³ 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. 38943899.
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 beta-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 beta-oxidation trifunctional enzyme complex*. Journal of Structural Biology, 2020. **210**(3): p. 107494.
68. Venkatesan, R. and Wierenga, R.K., *Structure of Mycobacterial beta-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 CO₂ Fixation in Metallosphaera sedula*. Frontiers in Microbiology, 2020. **11**(354).
70. Bartlett, K. and Eaton, S., *Mitochondrial beta-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 beta-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/4hydroxybutyrate CO₂ fixation cycle in extremely thermoacidophilic archaea*. Metabolic Engineering, 2016. **38**: p. 446-463.
77. Teufel, R., et al., *3-Hydroxypropionyl-Coenzyme A Dehydratase and AcryloylCoenzyme A Reductase, Enzymes of the Autotrophic 3-Hydroxypropionate/4Hydroxybutyrate 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 beta-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 shortchain 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. 1306413073.
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. 46376.
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. 30311.
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. 36963710.

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 S1: 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)
<i>Saci_1122</i>	<i>Saci_1122_fd_NdeI</i>	GCGTGA CATATG AGCAATGAGTATTACG
	<i>Saci_1122_rv_BamHI</i>	TTCACT GGATCC TCACTTGTTCTTCT
<i>Saci_1123</i>	<i>Saci_1123_fd_HindIII</i>	ATA AAGCTT ATGGTTTTGCCTTTTAAAAC
	<i>Saci_1123_rv_HindIII</i>	ATA AAGCTT TTACATTTTTATGCCAAATAA
<i>Saci_0315</i>	<i>Saci_0315_fd</i>	TAGCAGCCGGATCCTCGAGCAGCCCCTTCTTTTAAATTAAC
	<i>Saci_0315_rv</i>	CCTGGTGCCGCGCGGCAGCCATATGGCAGAGCTTAAAATTGTCG
<i>Saci_1109</i>	<i>Saci_1109_fd_BamHI</i>	TAT GGATCC CATGAAAGTAGAAGATATTAAGAAA
	<i>Saci_1109_rv_BamHI</i>	GAT GGATCC TTATTCTCCTTTGAACTGTG
<i>Saci_1114</i>	M13_fw	TGTA AACGACGGCCAGT
	M13_rv	CAGGAAACAGCTATGACC
<i>Saci_1104</i>	<i>Saci_1104_fd_NdeI</i>	GCTCGC CATATG TAGTCTCTTAAAGAC
	<i>Saci_1104_rv_BamHI</i>	CTAGCT GGATCC TTAAGCAATTCCT
<i>Saci_1115</i>	<i>Saci_1115_fw_NdeI</i>	CCTACG CATATG ATGAAAGCTGTAATTCTTC
	<i>Saci_1115_rv_BamHI</i>	CGAGCT GGATCC TTATGGCTTATAAGAATTTTAC
<i>Saci_1085</i>	<i>Saci_1085_fw_NdeI</i>	GCCCG CATATG TCAGAGCAGGGTCC
	<i>Saci_1085_rv_BamHI</i>	CGGGC GGATCC TCATTGTGTTTGTGAGTAC
Sequenced genes	Oligonucleotides for sequencing	Sequences (5'→3')
All above except for <i>saci_1085</i>	T7-promoter	TAATACGACTCACTATAGGG
	T7-terminator	GCTAGTTATTGCTCAGCGG
<i>Saci_1085</i>	FXara-fd	CAGCGTTTATAACGTTTAAACATG
	FXsulf-rv	CCATTTAATAGTTTGTATGGTCTACCC

Table S2. Plasmids and strains used for cloning and expression of the FA metabolic enzymes from *S. acidocaldarius*.

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

Table S3: Expression conditions used for the selected genes from *S. acidocaldarius*.

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

Table S4: 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 S5: 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 \pm represent standard error (SE); ND: not detected).

ORF/Enzyme	Native mass (kDa) (Oligomer)	Temp	Substrate/Cofactor	Km (mM)	Vmax (U mg ⁻¹)	kcat (S ⁻¹)
Saci_1122 Acyl-CoA synthase (ACS)	117 (homodimer)	55°C	Butyric acid (C4)	ND	1.34	ND
			Valeric acid (C5)	*0.046 \pm 0.009		
			CoA	0.575 \pm 0.134	*2.106 \pm 0.056	*2.35
			ATP	*0.965 \pm 0.151		
			Octanoic acid (C8)	*0.01 \pm 0.002	*3.282 \pm 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

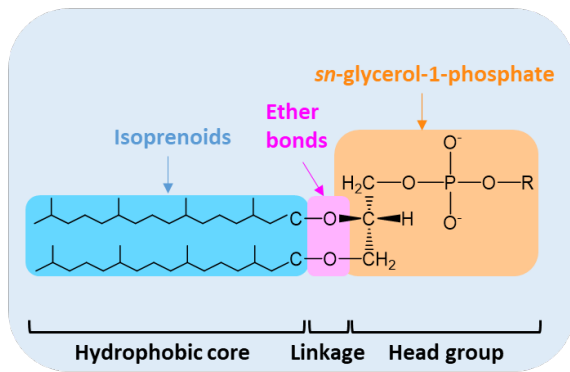
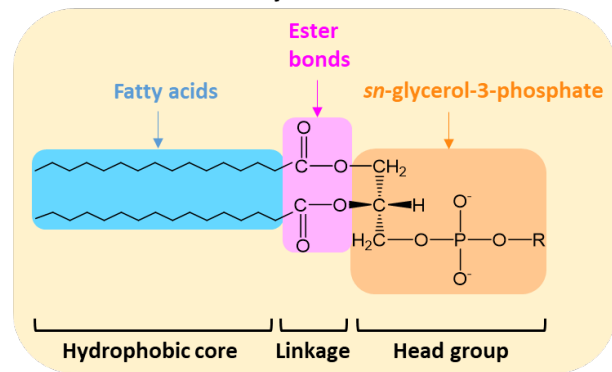
A. Archaea:**B. Bacteria & Eukarya:**

Figure S 1. **Lipid compositions in Archaea (A), Bacteria and Eukaryotes (B).** Archaea produce isoprenoidbased 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]

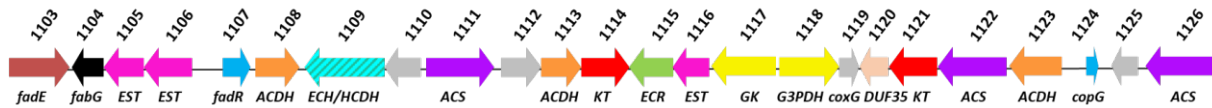


Figure S 2. **Genomic organization of the *saci_1103-saci_1126* gene cluster related to lipid and fatty acid metabolism in *S. acidocaldarius*.** Within this fatty acid operon, several copies of lipases/esterases (*saci_1105*, *saci_1106* and *saci_1116*, shown in pink) are present, two of them have 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/acetyl-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]

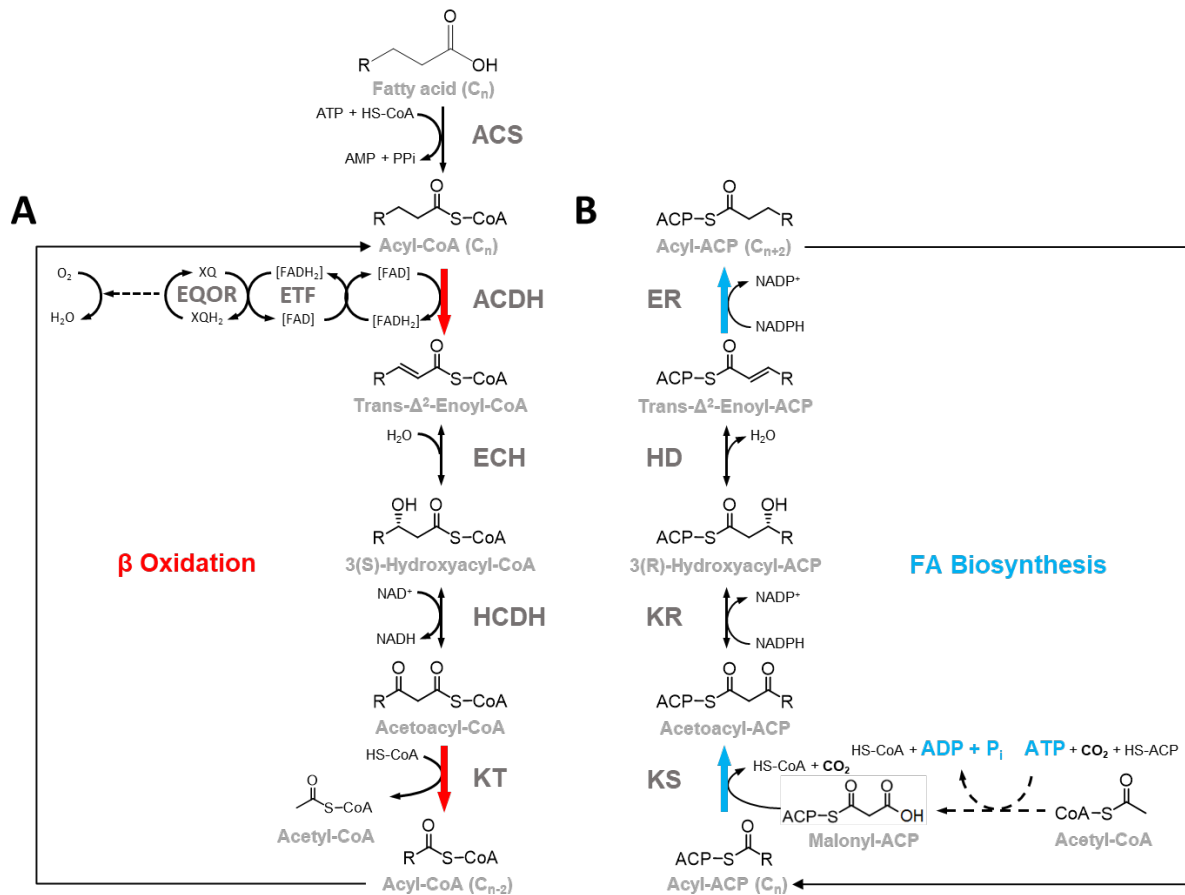


Figure S 3. **The classical fatty acid degradation and biosynthesis pathway.** Fatty acids are initially activated into acyl-CoAs by an AMP-forming acyl-CoA synthetase (ACS) and then degraded via β oxidation (A) as the following steps. Fatty acyl CoAs are oxidized to enoyl-CoAs by acyl-CoA dehydrogenase (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 acylCoA 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 acetylCoA 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 enoyl-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.

#

FadE27	112	VSAVSGDRILTVALDGEMGEGPVQAAG	138	150	TQVGYGPV	157
FadE29	112	PAILAGEAHFAIGYT-EPEAG-TDLAS	136	155	VFTTGAHD	162
saci_1108	112	RRILTAEDIWCQGFSEPHAG-SDLAS	136	155	IWSSYAHL	162
saci_1113	105	DKLFSGEVKIAVSDS-NYVPG-ADQAD	129	131	ILIDNT--	136
FadE26	116	PRIAAGDLHFSIGYS-EPGAG-TDLAN	140	159	MWTSLIQY	166
FadE28	94	AGVAKGGVLTALN-EPGAALPDRPA	120	131	VGVGYAEQ	138
MCAD	147	GRMTEEPLMCAYCVT-EPGAG-SDVAG	171	190	MWITNGGK	197
SCAD	141	TPFTSGDKIGCFALS-EPGNG-SDAGA	165	184	AWITNAWE	191
LCAD	159	PQMTAGKCIGAIAMT-EPGAG-SDLQG	183	202	VFISNGSL	209
IBD	147	PPLCTMEKFASYCLT-EPGSG-SDAAS	171	190	AFISGAGE	197
IVD	151	PPLCTMEKFASYCLT-EPGSG-SDAAS	175	194	FWITNGPD	201
saci_2217	121	TPVARGDKVAAFANT-EPQAG-SDVAG	145	164	IFITNGGI	171
saci_1123	121	TPVAKGDKVAAFANT-EPQAG-SDVAG	145	164	IFITNGGI	171
FadE27	224	LSTLSRT-----AFQ	233	247	YARTREQF-DRPIGSFOAVG	265
FadE29	235	TTQLNNEKRVMLG--P	247	261	WASVPGGN-GVTPIDHDDVK	279
saci_1108	235	MSTLNYESLNIGTIL	259	261	-----TGYKGESLI-----	270
saci_1113	175	-----SILFASQM	182	197	YSKERIAF-GKPIGSYQAIK	215
FadE26	241	TNQLNHEKVALVSPA	255	267	AQNTKDAGGTRLIDSEWVQ	286
FadE28	199	QLLAL----AVMGAYA	209	233	YVANRKQF-GKPLSTFQTV	251
MCAD	274	MGAFDKTPVVAAGA	288	302	YALERKTF-GKLLVEHQAI	320
SCAD	265	MQTLDMGRIGIASQA	279	293	YAENRMAF-GAPLTKLQVI	311
LCAD	285	MKELPQERLLIADVA	299	313	YVKQRKAF-GKTVAHLQTV	331
IBD	270	VRGLNGGRINIASC	284	298	HLNVRKQF-GEPLASNOYL	316
IVD	277	MSGLDLERLVLAGGP	291	305	YLHVREAF-GQKIGHFQLM	323
saci_2217	248	MSGLDLERLVLAGGP	262	276	YSVQSAF-GSPLLGFQMV	294
saci_1123	248	MSTFDASRVGVAGQA	262	276	YSVQSAF-GSPLLGFQMV	294
FadE27	323	HTIVHVGQVGV	334	349	QTEFALGGATGQLRR	363
FadE29	327	RLAEEIVGKYGN	338	360	VITF-GGGVNEVMRE	373
saci_1108	316	ESAFNTMGPEAL	327	348	SITI-AGGTSEILRN	361
saci_1113	258	LSGIQVHGGIGF	269	285	LSKI-YNGKVDISEF	298
FadE26	339	RLLMEVLGTAAT	350	373	ILTF-GGGTNEVQRD	386
FadE28	291	QICHHLHGGMGM	302	318	LTRL-LGGPSHRLEL	331
MCAD	370	TDAVQILGGNGF	381	396	IYQI-YEGTSQIQRL	409
SCAD	361	HQAIQILGGMGY	372	387	ITEI-YEGTSEIQRL	400
LCAD	381	YDCVQLHGGWGY	392	407	VQPI-YGGTNEIMKE	420
IBD	367	NQALQMHHGGYGY	378	393	VHQI-LEGSNEVMRI	406
IVD	373	LDGIQCFGGNGY	384	399	LYEI-GAGTSEVRR	412
saci_2217	346	IRAITVHGGYGV	357	372	VMKI-YEGANDIQKL	385
saci_1123	346	IRAITVHGGYGV	357	372	IMKI-YEGANDIQKL	385






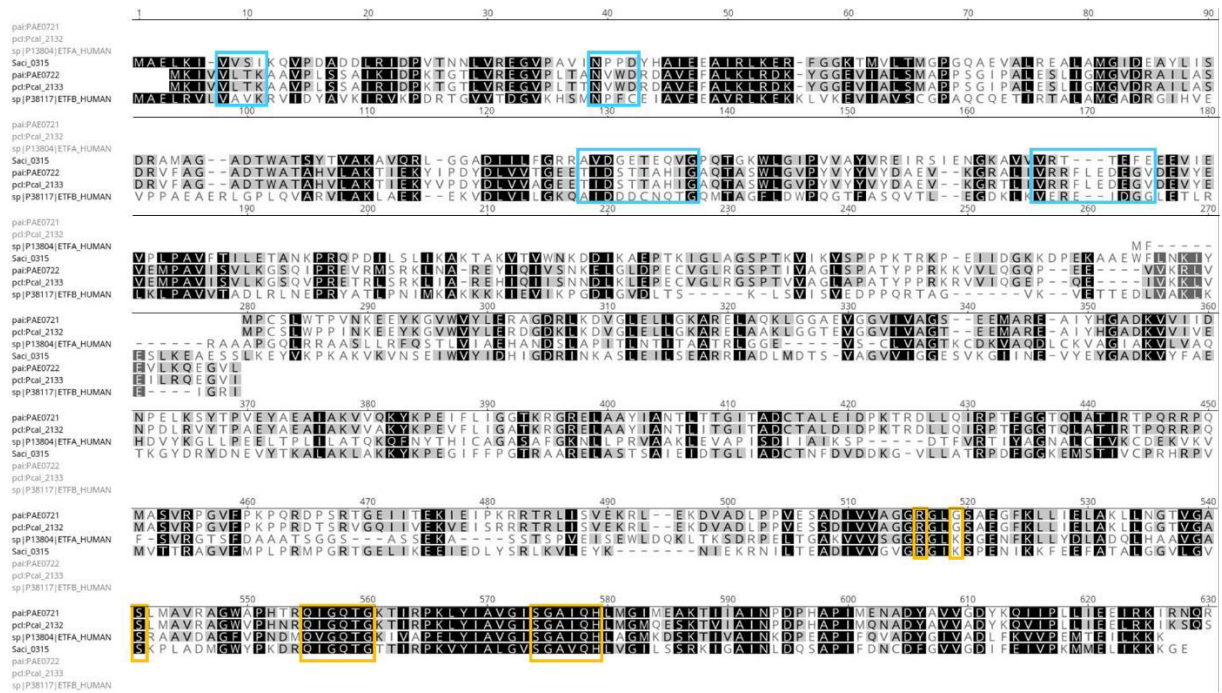
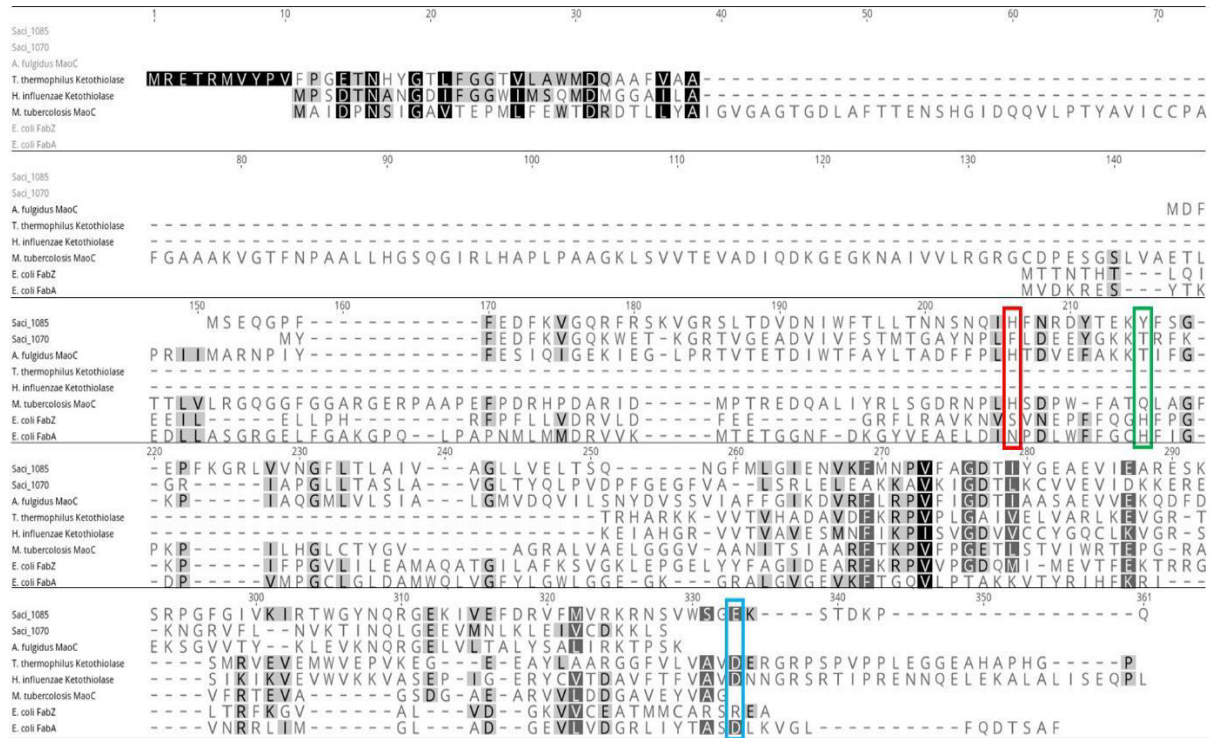
	Isoalloxazine (FAD) binding site		Adenosine (FAD) binding site
	CoA binding site		Catalytic base
	Catalytic base + Tyrosine residue (specific for SCAD and MCAD)		

Figure S 4. **Sequence alignment of various acyl-CoA dehydrogenases.** FadE26, 27, 28 and 29 represent the *Mycobacterium* 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 longchain 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 isoalloxazine of FAD) and dark blue (arginines, glutamines and glycines for adenosine of FAD), respectively.



A



B

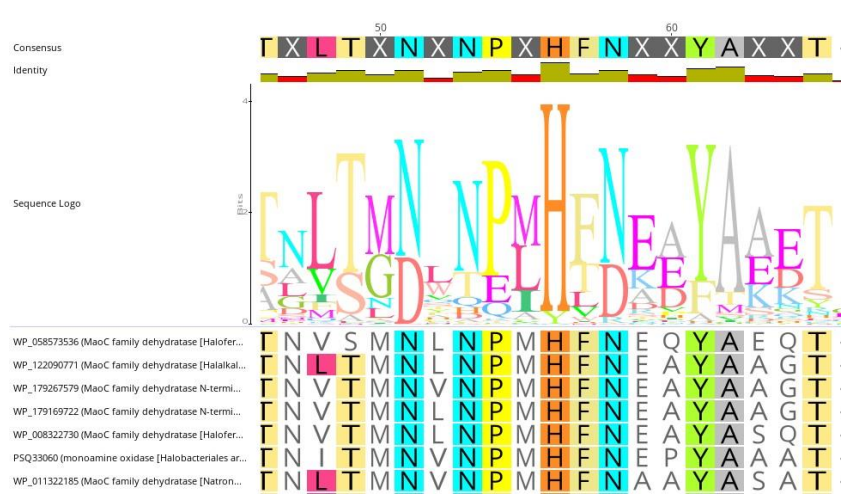


Figure S 6. Catalytic sites for different hot-dog fold enzymes (A) and conserved motifs for MaoC like dehydratase homologues (B). Different sequences were aligned and hot-dog fold enzymes of *S. acidocaldarius* were screened for conserved catalytic motifs. The catalytic histidine residue of FabA and FabZ is absent in all other hot-dog fold sequences from Bacteria like *Thermus thermophilus*, *Haemophilus influenzae* 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).

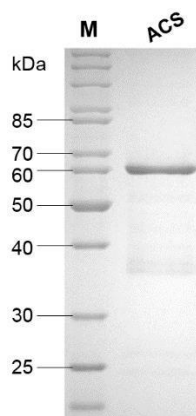


Figure S 7. **The recombinant acyl-CoA synthetase from *S. acidocaldarius* after purification.** ACS: SDS-PAGE (12.5%, Coomassie Blue staining) of 3 μ g of the purified AMP-forming acyl-CoA synthetase Saci_1122; M: 5 μ l PageRuler™ Unstained Protein Ladder (Fermentas).

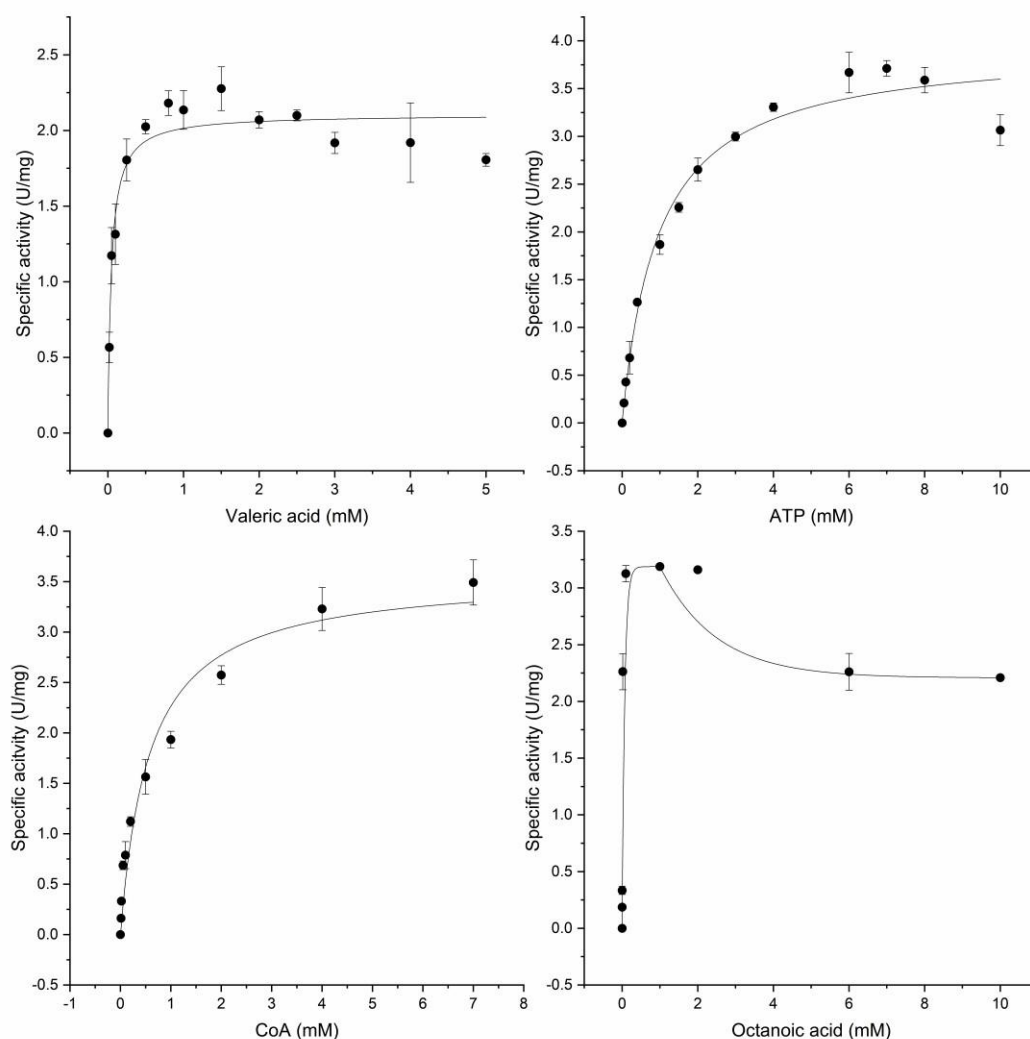


Figure S 8. **Determination of kinetic parameters of the recombinant ACS Saci_1122 from *S. acidocaldarius*.** The ACS activity was determined in couple with the auxiliary enzymes myokinase (MK), pyruvate kinase (PK) and lactate dehydrogenase (LDH) from rabbit muscle by monitoring NADH oxidation at 340 nm. The assay was performed at 55°C, pH 7 in 100 mM HEPES/NaOH buffer plus 20 mM MgCl₂ 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).

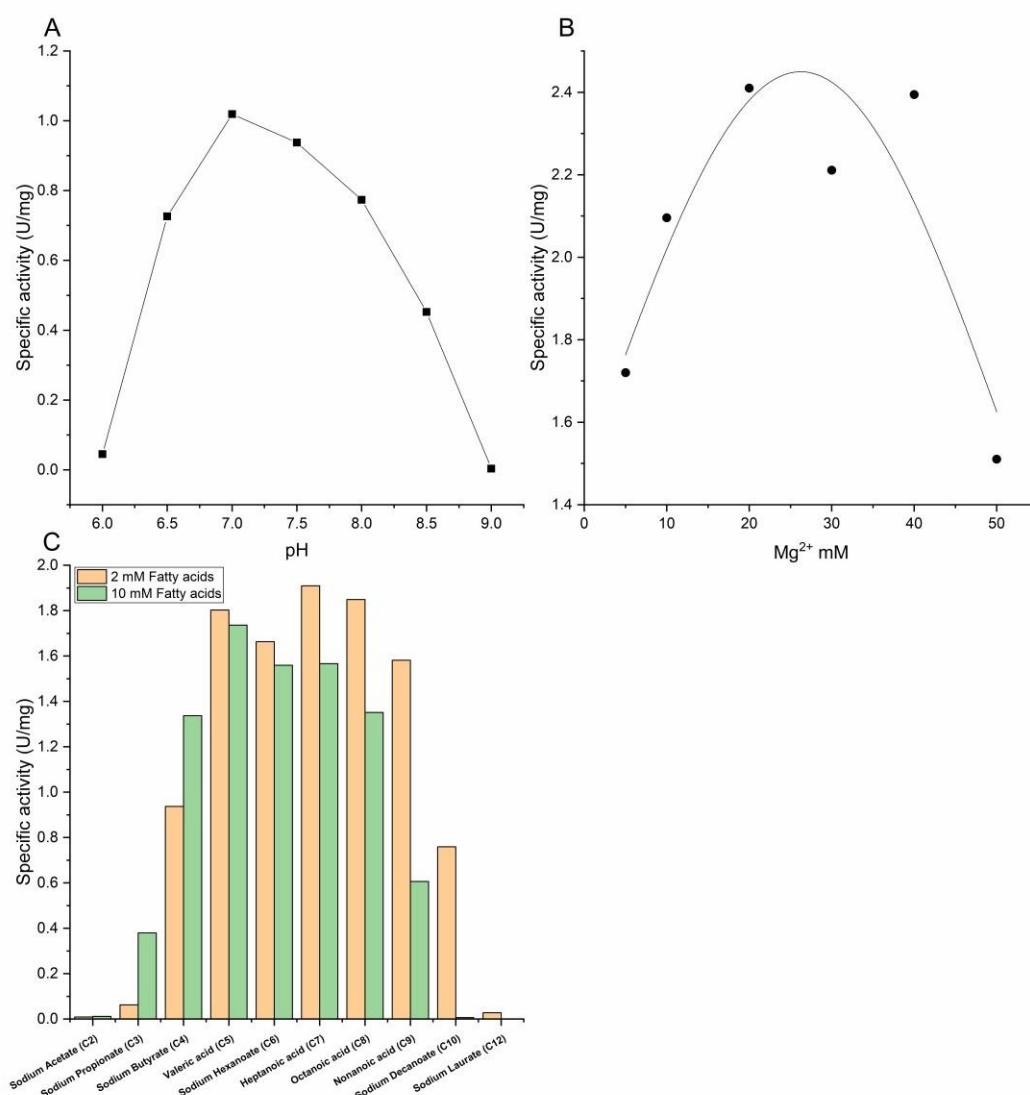


Figure S 9. **Determination of the optimal pH (A), Mg²⁺ dependency (B) and substrate spectrum of the recombinant ACS Saci_1122 from *S. acidocaldarius*.** The optimal pH was identified as 7 at 55°C with a mixed buffer of 50 mM MES, 50 mM HEPES and 50 mM Tris setting the pH values from 6 to 9. The mixture contained 5 mM MgCl₂, 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. Mg²⁺ dependency was studied in 100 mM HEPES/NaOH buffer (pH 7) with the same assay but varying the Mg²⁺ concentration from 0 to 50 mM. In the assay for analysis of the substrate specificity 10 mM MgCl₂, 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.

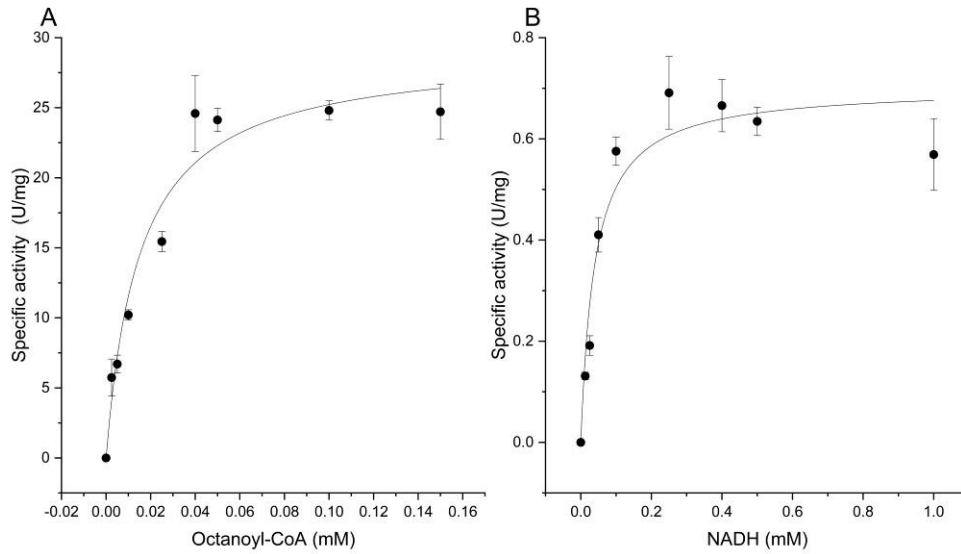


Figure S 10. Investigation of kinetic properties of the recombinant ACAD Saci_1123 (A) and ETF Saci_0315 (B) from *S. acidocaldarius*. The specific activity and K_m of Saci_1123 was tested at 65°C, pH 6.5 in 50 mM HEPES/KOH buffer plus 20 mM KCl, 0.13 $\mu\text{g}/\mu\text{l}$ ACAD with 1 mM FcPF₆ and 0-0.15 mM octanoyl-CoA. Reduction of FcPF₆ was monitored at 300 nm (extinction coefficient 4.3 $\text{mM}^{-1} \text{cm}^{-1}$). 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 idonitrotetrazolium chloride (INT) and 0.015 $\mu\text{g}/\mu\text{l}$ 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 $\text{mM}^{-1} \text{cm}^{-1}$). The independent measurements were performed in triplicate and error bars indicate the standard error of the mean (SEM).

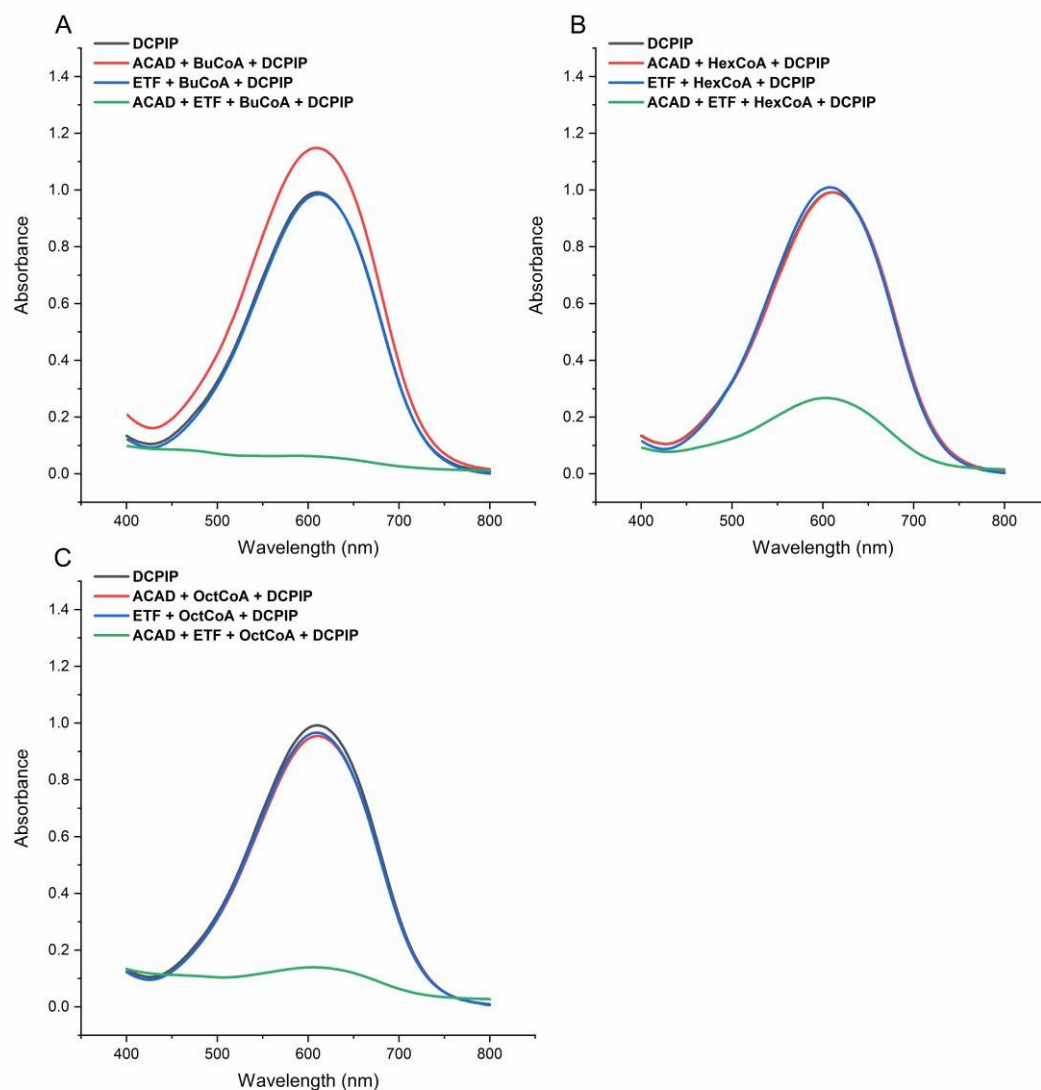


Figure S 11. **Investigation of electron transfer by ACAD and ETF.** The assay mixture (0.5 ml) contained 50 mM MES/KOH, 20 mM KCl, 0.2 mM DCPIP, 0.2 mM of different acyl-CoA (butyryl-CoA (A), hexanoyl-CoA (B) or octanoyl-CoA(C)), 1.7 μ g ACAD as well as 3.8 μ g ETF and was incubated at 65°C, pH 6.5 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.

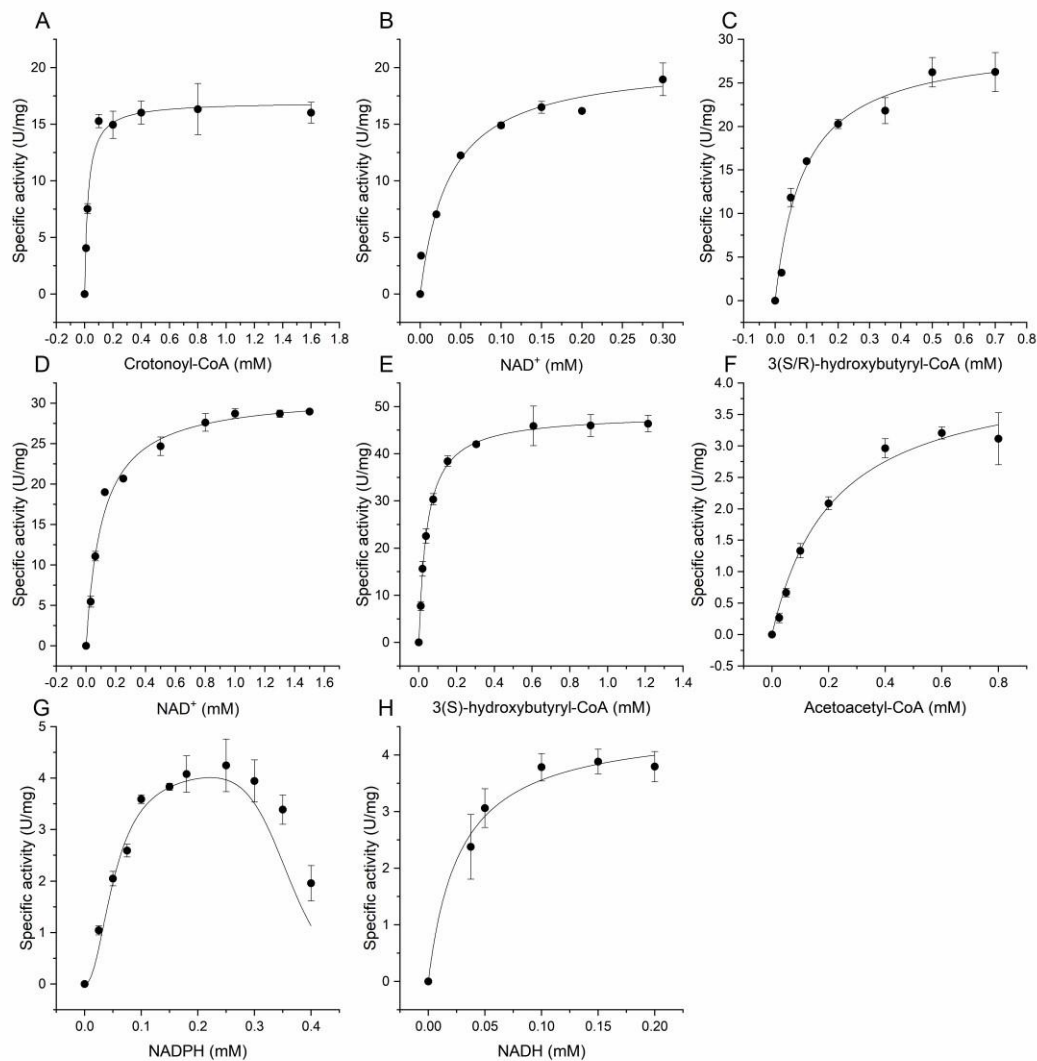


Figure S 12. **Investigation of kinetic properties of the recombinant HCDH/ECH Saci_1109 from *S. acidocaldarius*.** The enzymatic activity of HCDH/ECH was determined at pH 7 in 100 mM Tris/HCl (500 μ l). The formation or consumption of NAD(P)H was monitored at 340 nm. The assay for oxidative reactions was performed at 70°C with 0.69 μ g 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 crotonoyl-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 μ g 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).

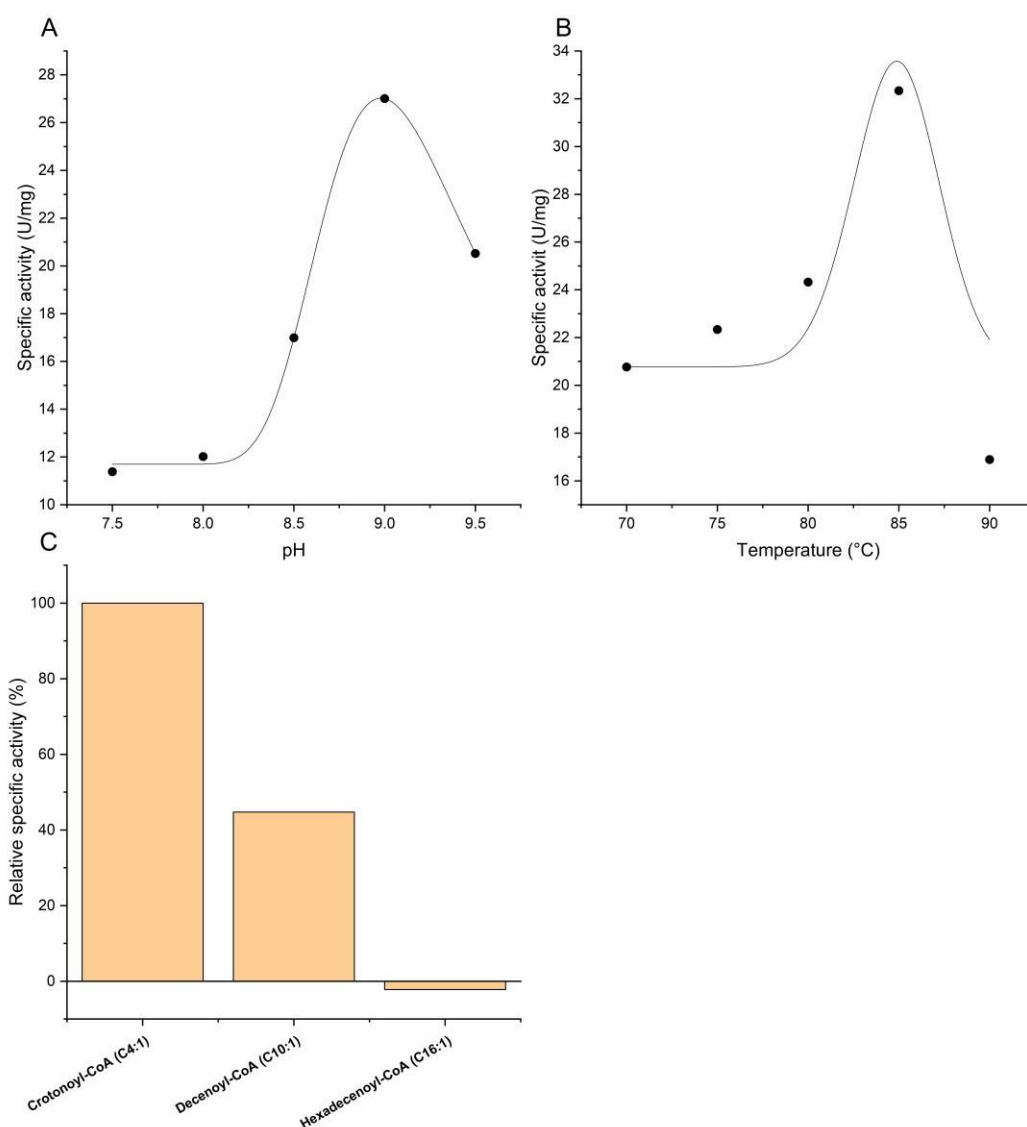


Figure S 13. **Determination of the optimal pH (A) and temperature (B) as well as the substrate spectrum (C) of the recombinant HCDH/ECH Saci₁₁₀₉ from *S. acidocaldarius*.** The optimal pH was determined in a mixed buffer of 50 mM MES, 50 mM HEPES and 50 mM Tris at 70°C. The assay contained 0.4 mM crotonoyl-CoA, 0.2 mM NAD⁺ and 0.69 µg Saci₁₁₀₉. 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.

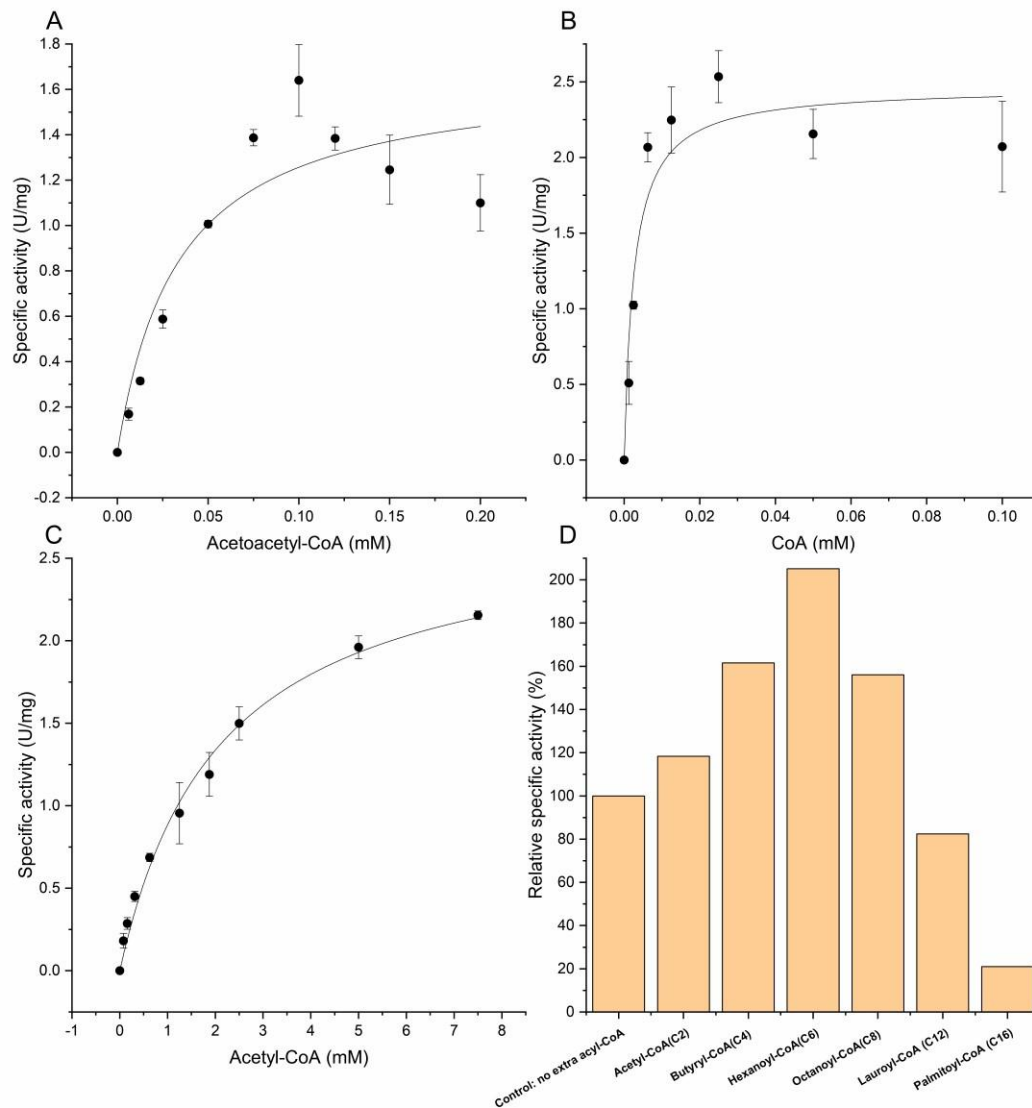


Figure S 14. **Determination of kinetic parameters and the substrate spectrum of the recombinant KT from *S. acidocaldarius*.** The specific activity of KT was photometrically tested at room temperature (23°C), pH 8 by monitoring the decrease of Mg^{2+} -AcAcCoA chelates (extinction coefficient of $21.4 \text{ mM}^{-1} \text{ cm}^{-1}$) 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 $2.7 \mu\text{g}$ 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, $17.1 \mu\text{g}$ Saci_1109, $10.8 \mu\text{g}$ 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).

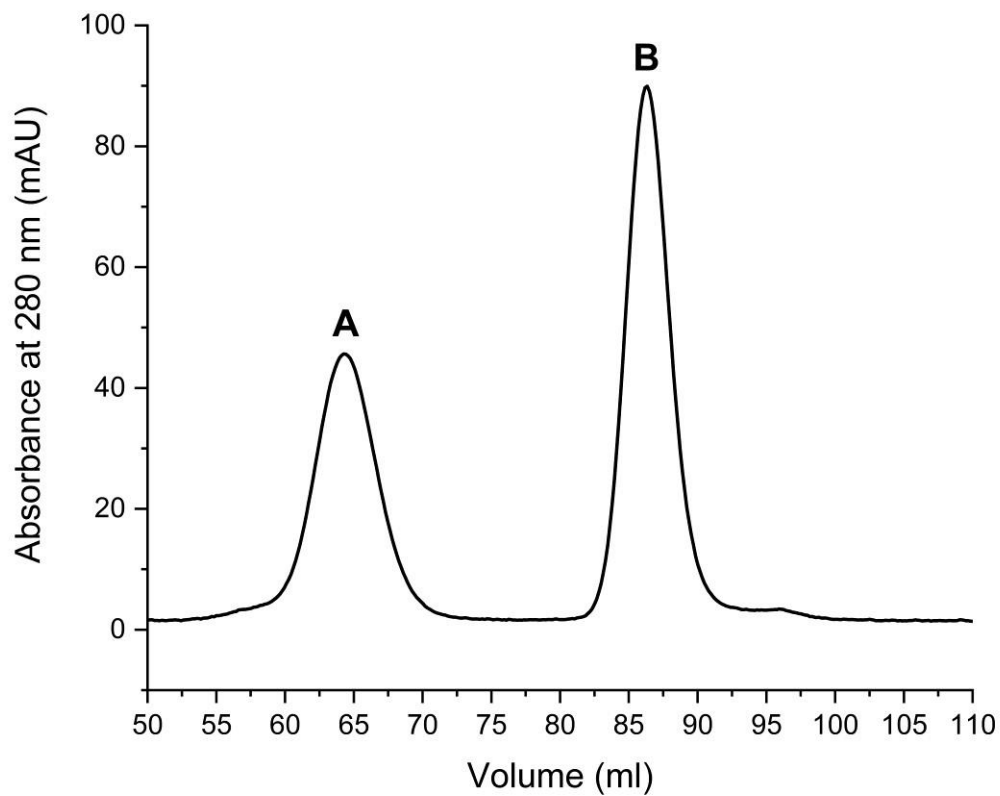


Figure S 15. **Gel filtration of the recombinant ECH/HCDH Saci_1109 and KT Saci_1114 from *S. acidocaldarius*.** Around 0.015 μmol of each of the purified recombinant ECH/HCDH and KT proteins were mixed and incubated on ice for 4 hours. Then, the protein mixture was applied to a Superdex 200 prep grad 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.

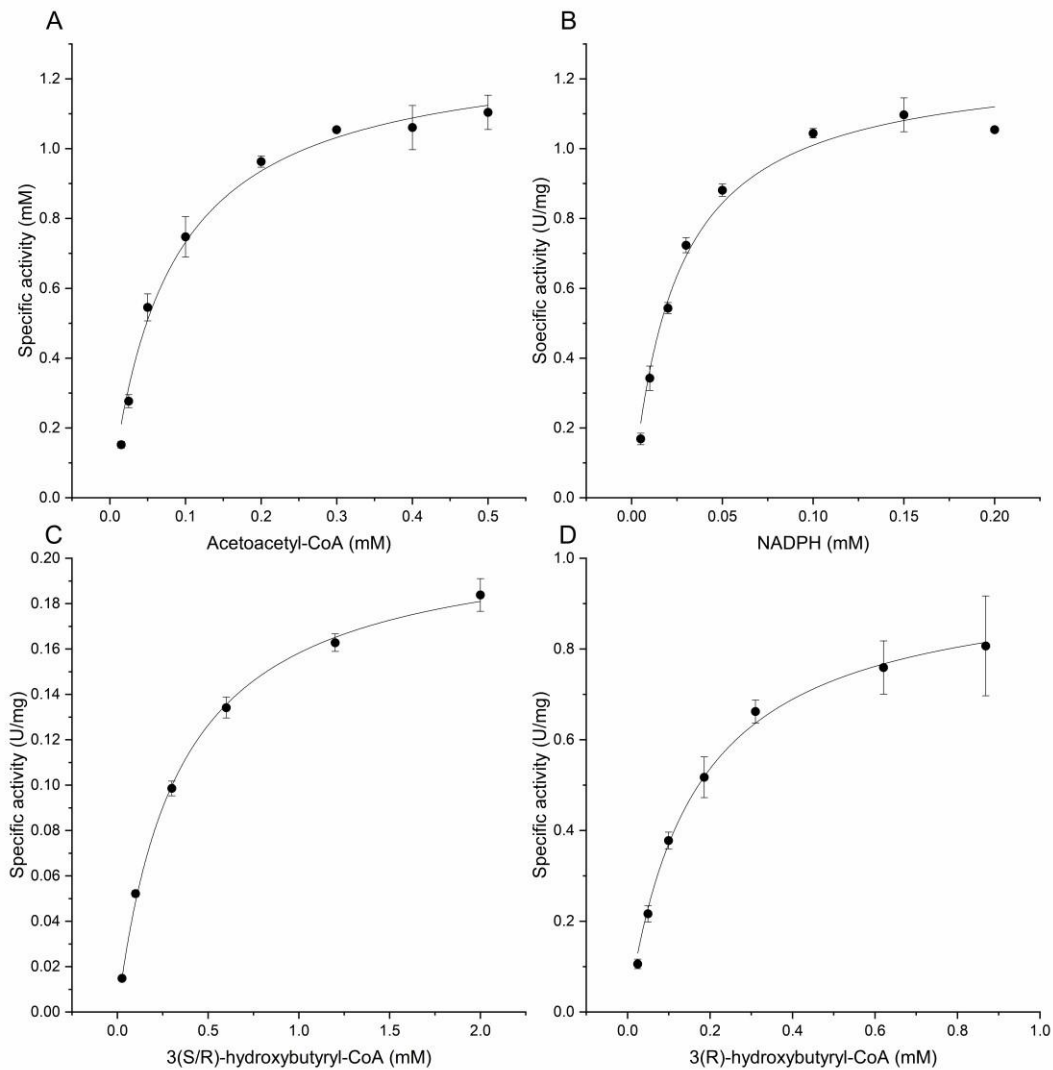


Figure S 16. **Investigation of the kinetic parameters of the recombinant ACR Saci_1104 from *S. acidocaldarius*.** The ACR activity was determined in 100 mM Tris/HCl with 0.3 mM AcAcCoA, 0.2 mM NADH/NADPH and 4.03 μ g pure protein at 35°C, pH 7 (340 nm). The K_m values were determined for AcAcCoA (A) 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 μ g 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).

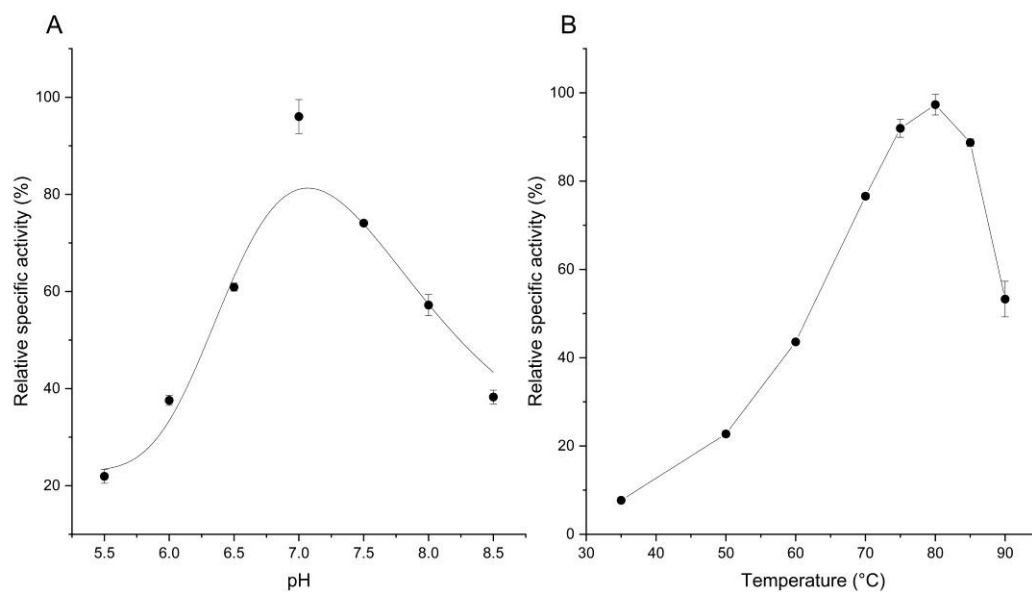


Figure S 17. **Identification of the optimal pH (A) and temperature (B) of the recombinant ACR Saci_1104 from *S. acidocaldarius*.** The optimal pH was determined in the direction of 3-hydroxybutyryl-CoA formation at 35°C using 100 mM MES/NaOH (pH 5.5-6.5) and 100 mM Tris /HCl (pH 7.0-8.5) as buffers. 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

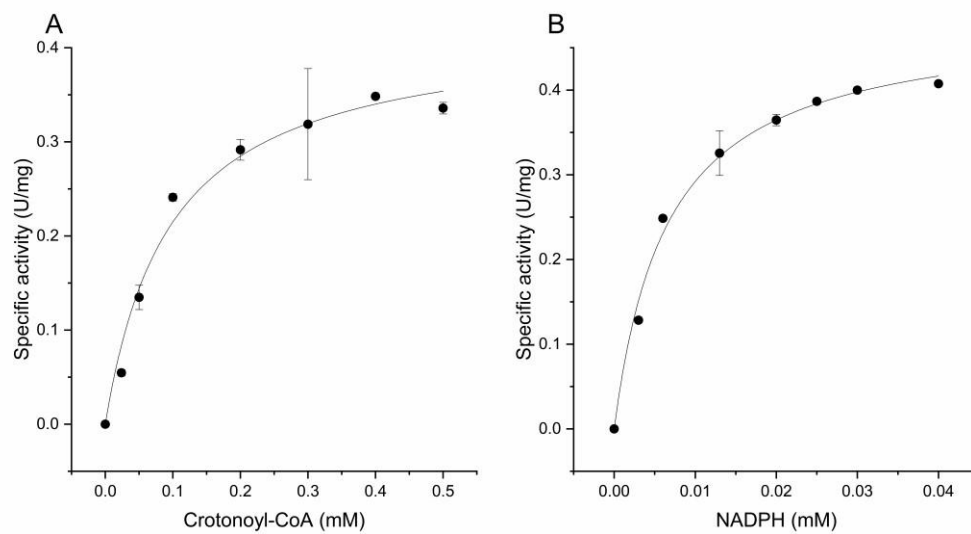


Figure S 18. **Determination of kinetic parameters of the recombinant ECR Saci_1115 from *S. acidocaldarius*.** The ACR activity was determined at 70°C, pH 7.5 (340 nm) with 12 µg pure protein in 100 mM HEPES/NaOH including 10 mM KCl, 0.3 mM NADPH and 0.4 mM crotonoyl-CoA in a total volume of 0.5 ml. 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).

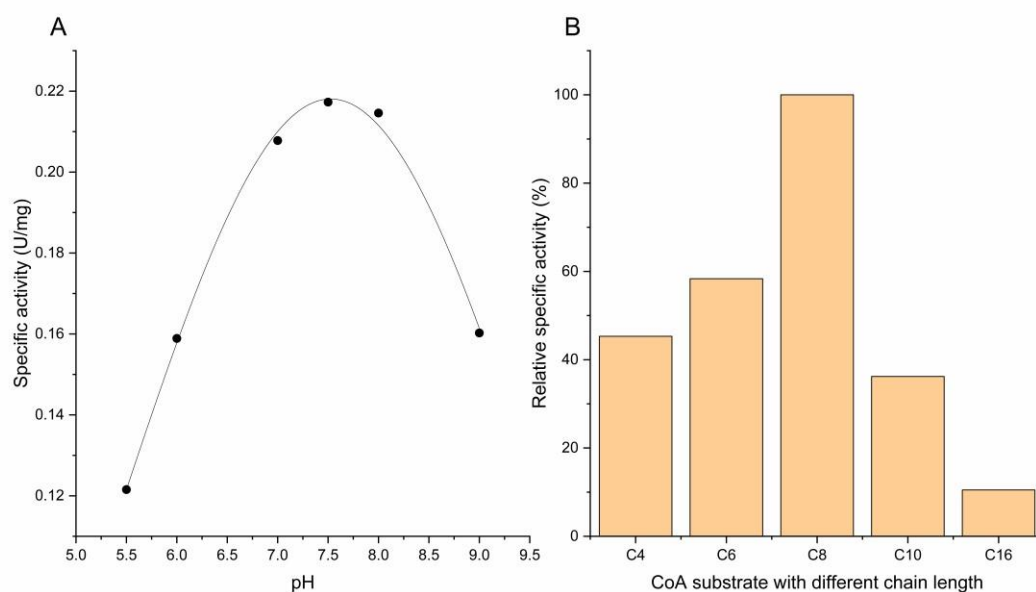


Figure S 19. **Determination of the pH optimum (A) and the substrate spectrum (B) of the recombinant ECR Saci_1115 from *S. acidocaldarius*.** The optimal pH for ECR was determined as 7.5 at 70°C using mixed buffer of 0.5 M HEPES, 0.5 M Tris and 0.5 M MES in presence of 10 mM KCl, 0.4 mM crotonoyl-CoA, 0.3 mM NADPH and 8 µg enzyme. The substrate specificity of ECR towards different 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 µg protein.

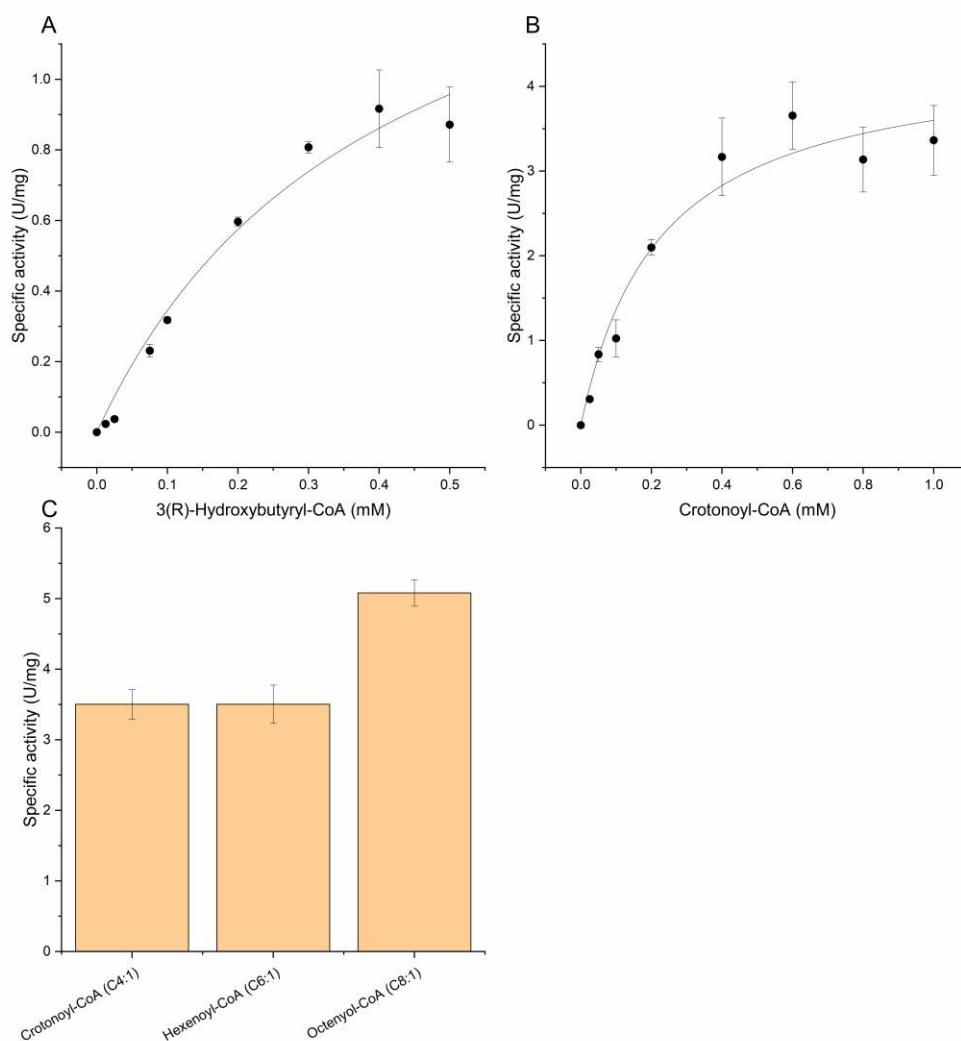


Figure S 20. **Investigation of kinetic properties (A, B) and substrate spectrum (C) of the recombinant MaoCHCD Saci_1085 from *S. acidocaldarius*.** The activity of MaoC-HCD Saci_1085 was tested at 65°C, pH 6.5 via a discontinuous assay (50 μ l) containing 50 mM MES, 20 mM KCl and 0.0675 μ g/ μ l protein. 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 μ g/ μ l 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).

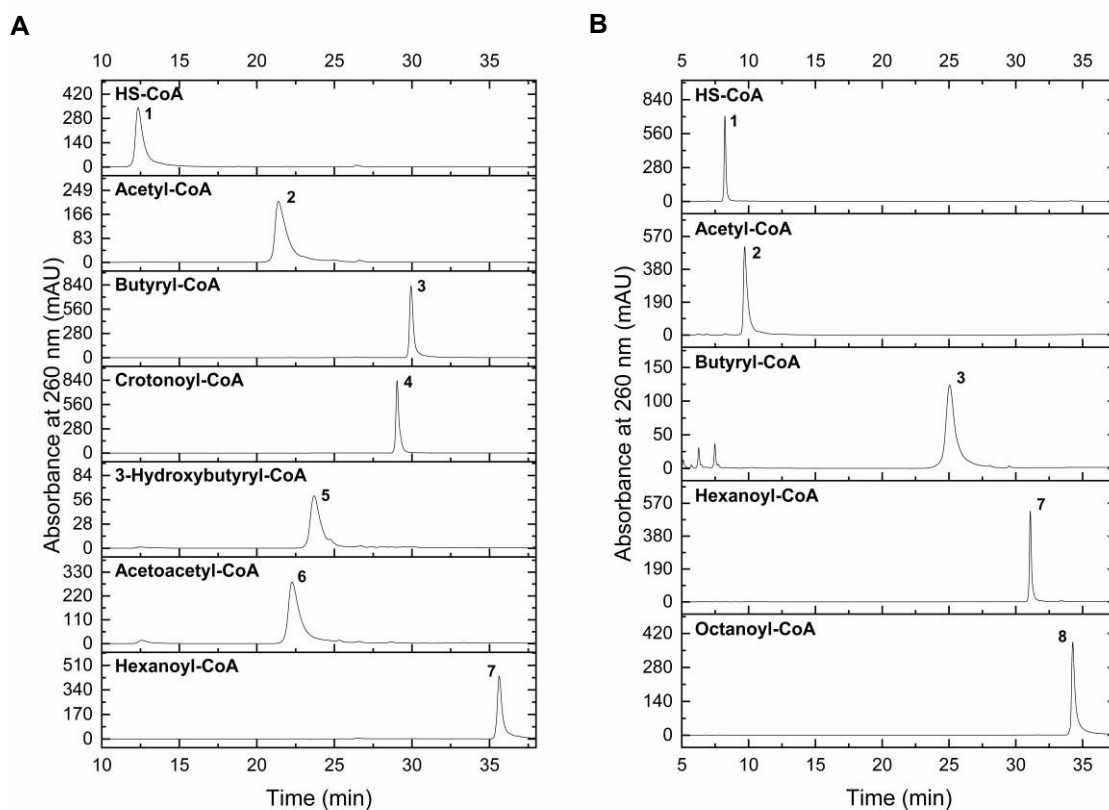


Figure S 21. **HPLC chromatogram of CoA ester standards involved in FA metabolism.** Two distinct programs with different acetonitrile (ACN) concentration gradients were applied for analyzing different chain lengths of CoA esters. The program “4-30% ACN program” (A) was used for shorter chain acyl-CoAs 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. AcetylCoA; 3. Butyryl-CoA; 4. Crotonoyl-CoA; 5. 3-Hydroxybutyryl-CoA; 6. Acetoacetyl-CoA; 7. Hexanoyl-CoA; 8. Octanoyl-CoA.

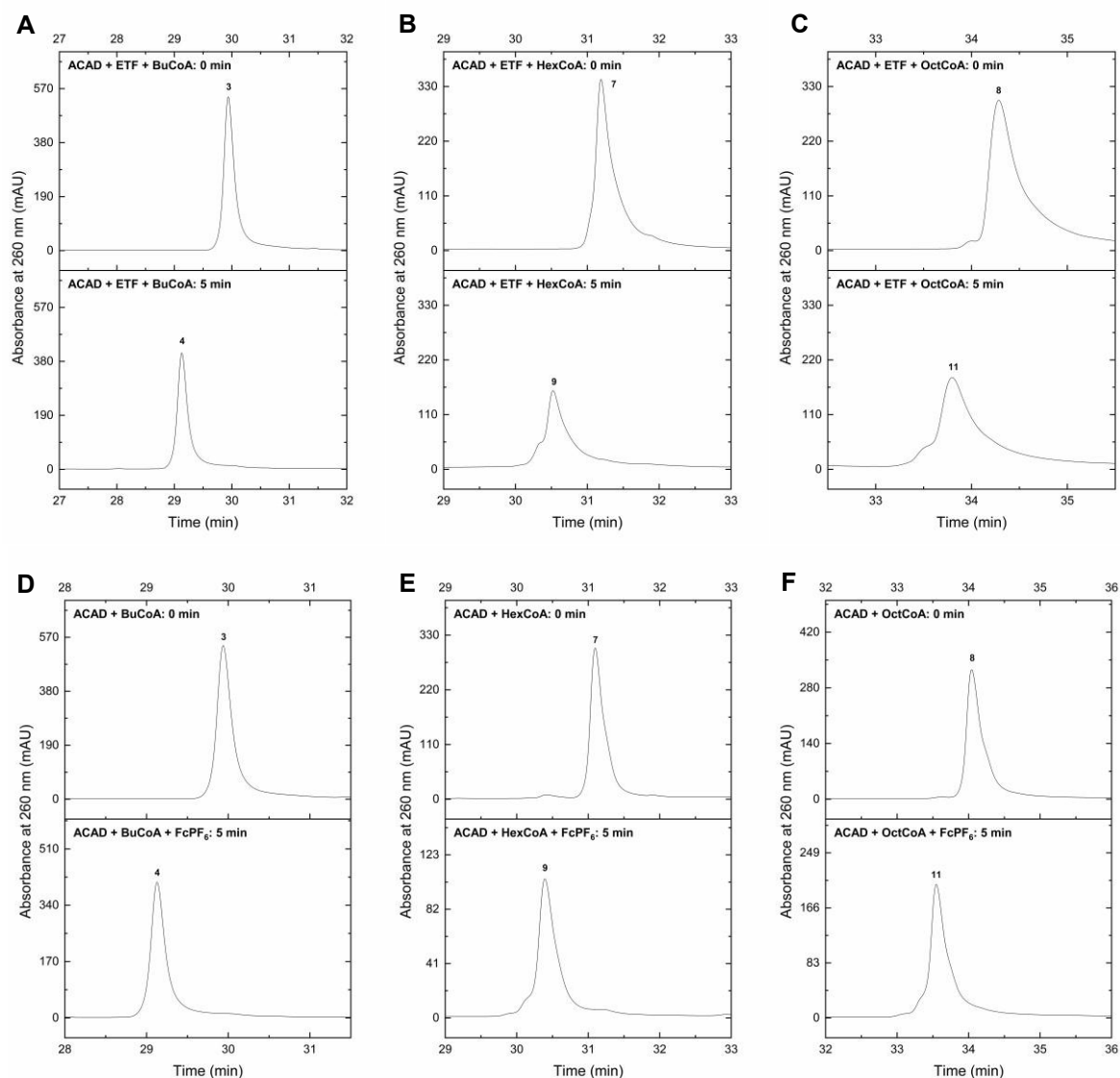


Figure S 22. **HPLC chromatograph of acyl-CoA oxidation to enoyl-CoA.** The oxidation of saturated acyl-CoAs to enoyl-CoA derivatives was analyzed by discontinuous assays at 65°C, pH 6.5. The 0.02 µg/µl ACAD and 0.01 µg/µl ETF were incubated in 50 mM MES/KOH (50 µl) with 20 mM KCl, 0.4 mM DCPIP and 0.4 mM of acyl-CoAs (butyryl-CoA, hexanoyl-CoA or octanoyl-CoA) for 5 min (A, B & C). Moreover, 0.8 mM FcPF₆ 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).

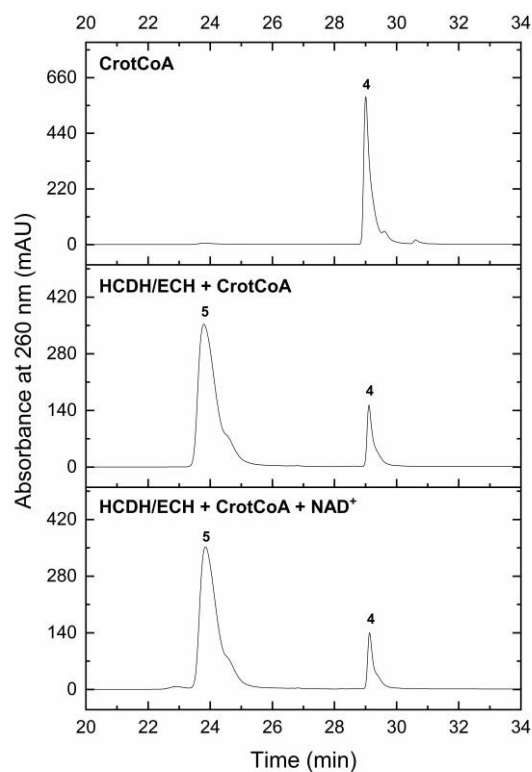


Figure S 23. **HPLC analysis of crotonoyl-CoA conversion into acetoacetyl-CoA by HCDH/ECH from *S. acidocaldarius*.** The discontinuous assay was carried out in 50 mM HEPES/NaOH with 20 mM KCl, 0.4 mM crotonoyl-CoA, 0.0144 $\mu\text{g}/\mu\text{l}$ Saci_1109 in absence or presence of 2 mM NAD^+ . The 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 3hydroxybutyryl-CoA (peak 4) from crotonoyl-CoA (peak 5) could be shown. However, further production of acetoacetyl-CoA was not observed in presence of NAD^+ .

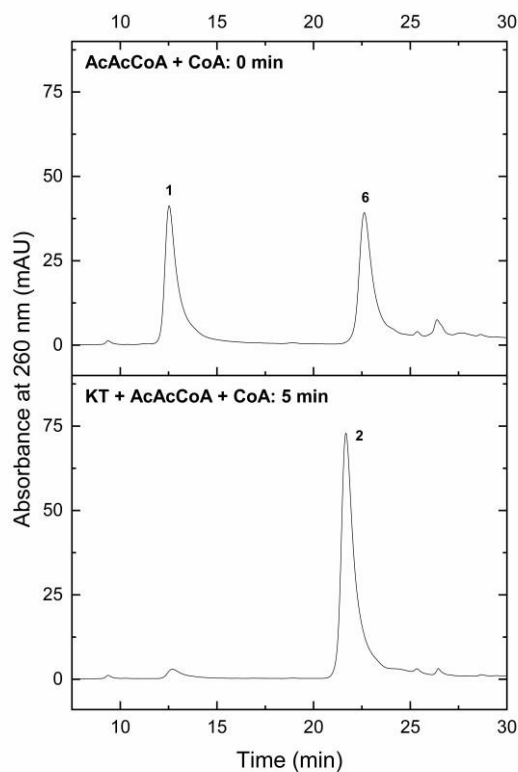


Figure S 24. **HPLC analysis of cleavage of acetoacetyl-CoA by KT.** To analyze the last thiolytic step, reaction components (400 μ l) including 0.2 mM CoA and 0.1 mM AcAcCoA were incubated in 50 mM MES/KOH (pH 6.5) at 23°C for 2 min. The reaction was then initiated by addition of 2.7 μ g of KT followed 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).

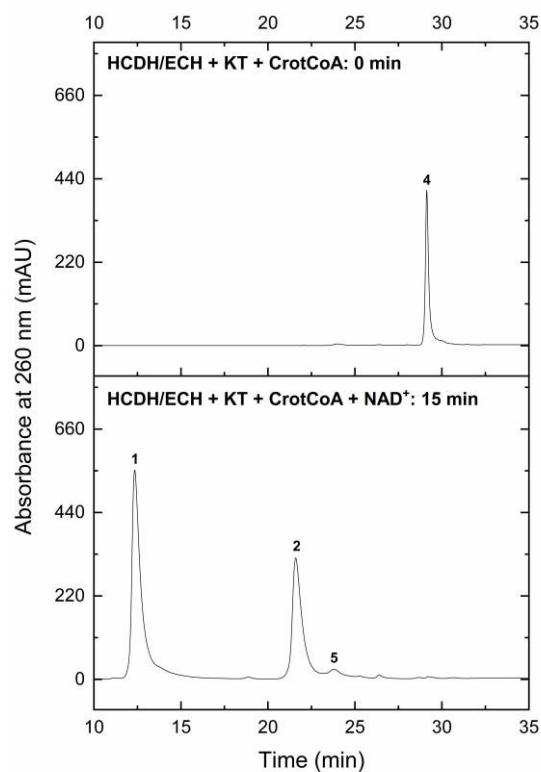


Figure S 25. **HPLC investigation of the last three steps by the β oxidation enzymes HCDH/ECH and KT.** The assay was performed in 50 mM MES/NaOH at 65°C, pH 6.5 by adding 20 mM KCl, 0.4 mM crotonoyl-CoA, 2 mM NAD⁺, 0.0144 $\mu\text{g}/\mu\text{l}$ HCDH/ECH, 1.6 mM CoA and 0.054 $\mu\text{g}/\mu\text{l}$ 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.

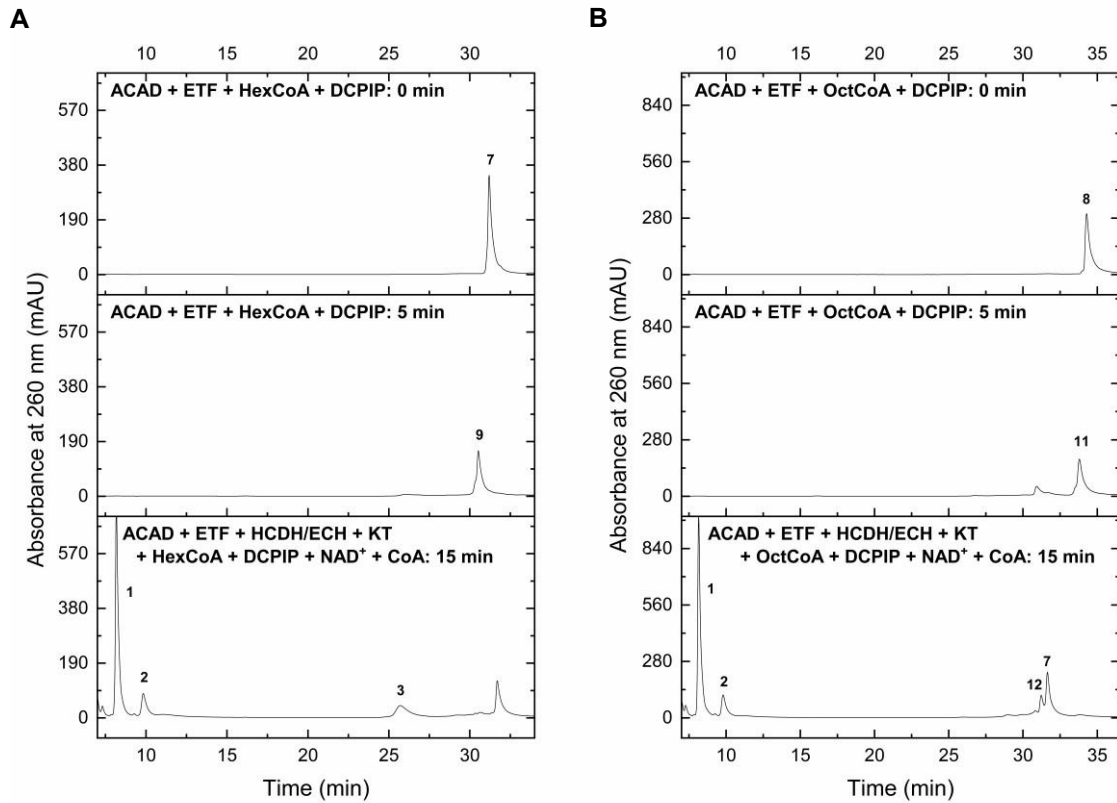


Figure S 26. **β oxidation cascades for degrading hexanoyl-CoA (A) or octanoyl-CoA (B).** The enzyme assays were carried out in two steps. The first oxidation step by 0.02 $\mu\text{g}/\mu\text{l}$ ACAD (Saci_1123) and 0.01 $\mu\text{g}/\mu\text{l}$ ETF (Saci_0315) was done in 50 mM MES/KOH (pH 6.5) at 65°C with 20 mM KCl, 0.4 mM DCPIP and 0.4 mM of acylCoA. The reaction was run for 5 min. In the second step, 2 mM NAD⁺, 0.0144 $\mu\text{g}/\mu\text{l}$ HCDH/ECH (Saci_1109), 1.6 mM CoA and 0.054 $\mu\text{g}/\mu\text{l}$ 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).

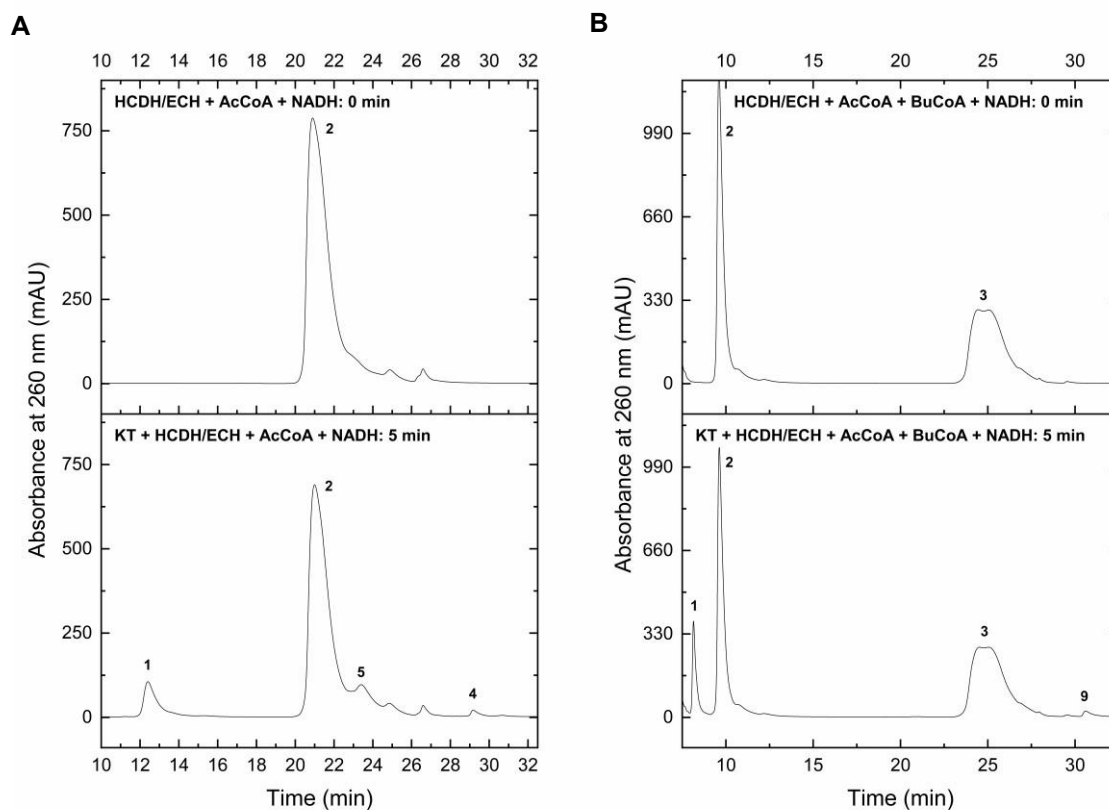


Figure S 27. **HPLC analysis of elongation of acetyl-CoA (A) or butyryl-CoA (B) by KT and HCDH/ECH.** The reaction mixture (400 μ l) included 50 mM HEPES/NaOH (65°C, pH 6.5), 2 mM acetyl-CoA, 0.3 mM NADH and 17.1 μ g HCDH/ECH (Saci_1109). After incubation for 2 min, the reaction was started by adding 16.2 μ g 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 3hydroxybutyryl-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).

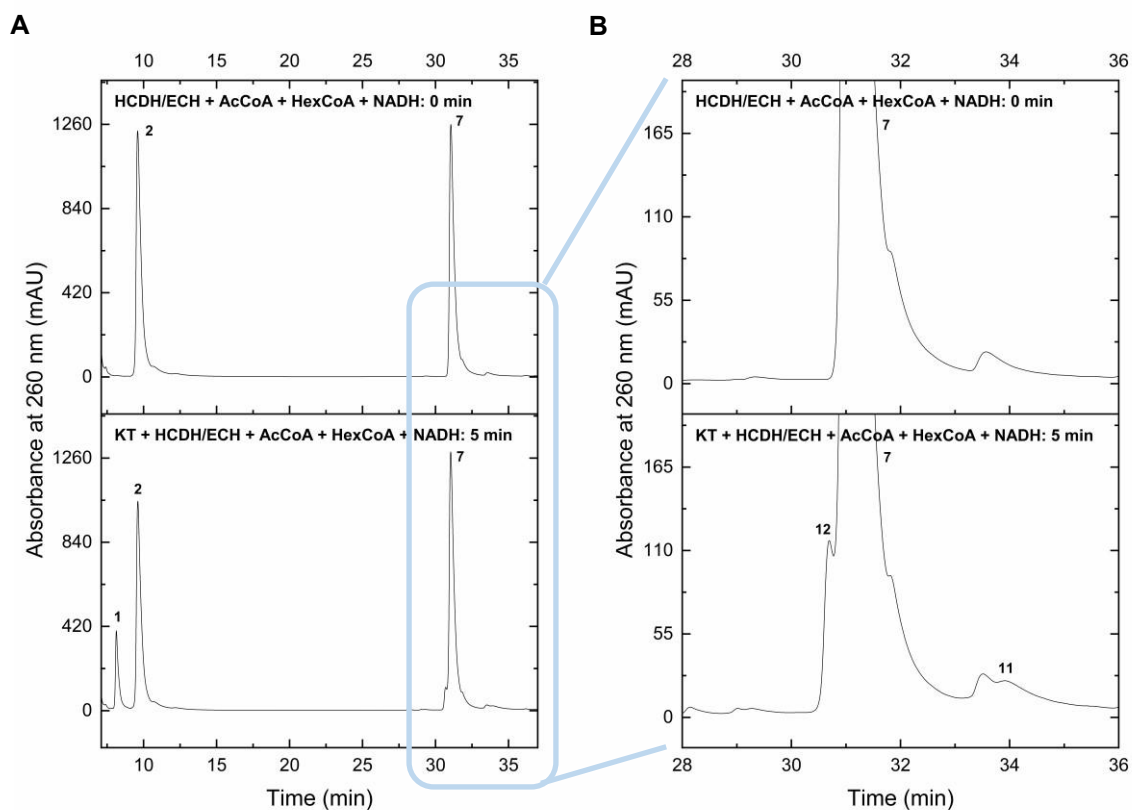


Figure S 28. **HPLC detection of condensation of hexanoyl-CoA and acetyl-CoA by KT and HCDH/ECH.** The reaction mixture (400 μ l) included 50 mM HEPES/NaOH (65°C, pH 6.5), 1 mM acetyl-CoA, 1 mM hexanoyl-CoA, 0.3 mM NADH and 17.1 μ g HCDH/ECH (Saci_1109). After incubation for 2 min, the reaction was started by adding 16.2 μ g 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.

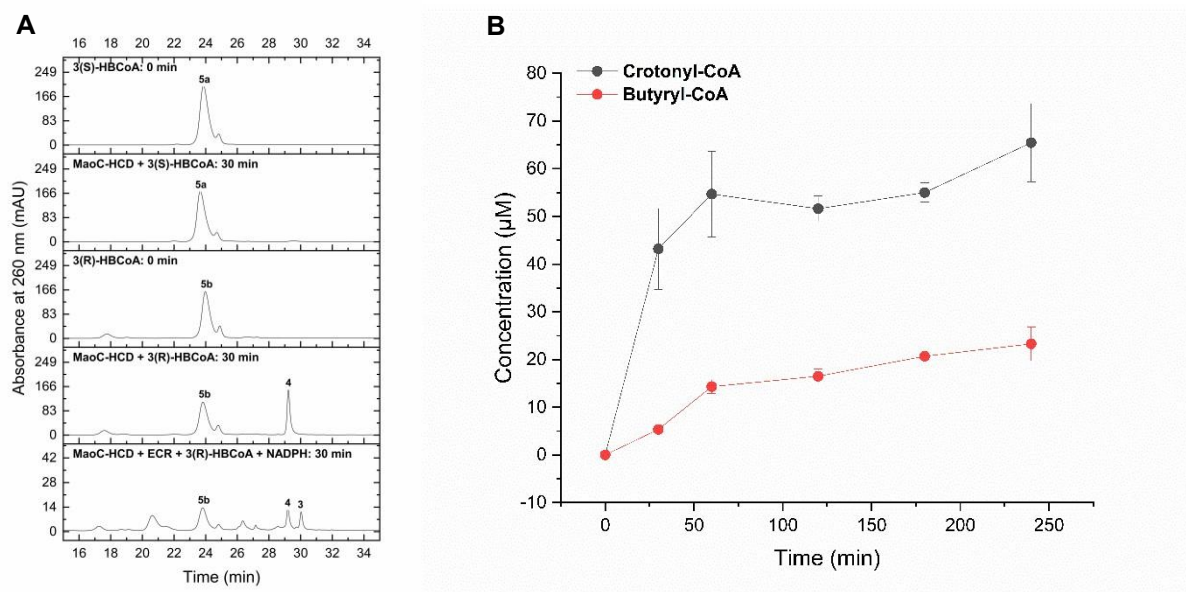


Figure S 29. **HPLC investigation of the stereospecificity of the recombinant MaoC-HCD Saci_1085 (A) and the conversion of 3(R)-hydroxybutyryl-CoA to butyryl-CoA by MaoC-HCD and ECR Saci_1115 (B).** To study the stereochemical specificity towards the 3-hydroxyacyl-CoA intermediates, 0.4 mM of individual 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 butyrylCoA (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. 48393.

**Chapter 3.2: Glycerol degradation in the
thermoacidophilic crenarchaeon
Sulfolobus acidocaldarius involves an
unusual glycerol-3-phosphate
dehydrogenase**

Glycerol degradation in the thermoacidophilic crenarchaeon *Sulfolobus acidocaldarius* involves an unusual glycerol-3-phosphate dehydrogenase

Christian Schmerling¹, Xiaoxiao Zhou¹, Carsten Schroeder¹, Tobias Busche², Jörn Kalinowski², Lidia Montero^{3,4}, Oliver Schmitz^{3,4}, Sabrina Ninck⁵, Markus Kaiser⁵, Jan Bost⁶, Alejandra Recalde⁶, Sonja V. Albers⁶, Christopher Bräsen¹, Bettina Siebers¹

¹Molecular Enzyme Technology and Biochemistry (MEB), Environmental Microbiology and Biotechnology (EMB), Centre for Water and Environmental Research (CWE), Faculty of Chemistry, University of Duisburg-Essen, Essen, Germany

²Center for Biotechnology (CeBiTec), Bielefeld University, Bielefeld, Germany

³Applied Analytical Chemistry (AAC), University of Duisburg-Essen, Essen, Germany

⁴Teaching and Research Center for Separation (TRC), University of Duisburg-Essen, Essen, Germany

⁵Chemical Biology, Center of Medical Biotechnology, Faculty of Biology, University of Duisburg-Essen, Essen, Germany

⁶Molecular Biology of Archaea, Institute of Biology II – Microbiology, University of Freiburg, Freiburg, Germany.

Abstract

Glycerol is an integral constituent of membrane phospholipids in all living organisms as well as of storage lipids and is thus a highly abundant organic compound in nature. Accordingly, the ability to utilize glycerol as a growth substrate is likewise widespread among organisms from all domains of life. In contrast to Bacteria and Eukarya, glycerol metabolism is not well understood in Archaea and despite some sequence distribution comparisons only halophiles have been analyzed in this respect to some extent. Herein, using growth experiments, polyomics approaches, and biochemical analyses it was demonstrated that the thermoacidophilic crenarchaeal model organism *Sulfolobus acidocaldarius*, previously regarded as non-glycerol utilizer, grows with glycerol as sole source of carbon and energy. After uptake likely involving facilitated diffusion, glycerol is dissimilated by ATP-dependent phosphorylation to glycerol-3-phosphate (G3P) via glycerol kinase (GK, Saci_2033) followed by oxidation mediated by a membrane associated, quinone reducing FAD-dependent G3P dehydrogenase (Saci_2032). Both enzymes were recombinantly produced, purified and characterized. Whereas the GK showed high structural and functional similarity to known enzymes from Bacteria and Eukaryotes, the G3PDH represents an unusual homologue of the bacterial GlpA subunit of the GlpABC complex with remarkable sequence differences in the C-terminus and a novel type of membrane anchoring for G3PDHs facilitated by a small CoxG-like protein (Saci_2031). Based on these findings, further sequence analyzes in archaeal organisms discovered a higher versatility of G3PDHs in Archaea with respect to interacting proteins, electron transfer, and membrane anchoring which seems to reflect tailored evolutionary solutions to meet different requirements caused by lifestyles and electron acceptors.

Introduction

Glycerol is an integral constituent of membrane phospholipids in all domains of life and also of storage lipids mainly known from bacteria and eukaryotes and is thus a highly abundant organic compound in nature. Furthermore, as a C3 molecule with a higher reduced state than sugars, it is even more attractive as energy source than sugars – at least for respiring organisms – and less challenging than e.g. C2 compounds. Accordingly, many organisms from Archaea and Bacteria to complex Eukarya are able to utilize glycerol as carbon and energy source.

Before glycerol is channeled into metabolism, it needs to be transported across the cytoplasmic membrane which often involves facilitated diffusion via aqua(glycerol)porins, transmembrane proteins catalyzing the rapid equilibration of glycerol concentration gradients across the membrane [1]. In Bacteria like *Escherichia coli* these glycerol-uptake facilitator (GUF) proteins are encoded by the *glpF* gene. However, glycerol as a small uncharged molecule can also enter the cell via passive diffusion through the cytoplasmic membrane [2]. In Bacteria, alternative glycerol transporters have so far only been described for *Mycoplasma mycoides* and *M. galliosepticum* employing ABC transporters (in addition to a GlpF homologue) with implications for transport efficiencies and pathogenicity [3, 4]. In Eukarya, also glycerol/H⁺ or glycerol/Na⁺ symport systems have been described especially in yeasts [5, 6].

Once transported, glycerol metabolism in bacteria and eukaryotes (e.g. *E. coli*, *Pseudomonas aeruginosa*, and yeast) as depicted in Fig. 1 essentially follows two possible biochemical routes in which phosphorylation and dehydrogenation steps ultimately convert glycerol into dihydroxyacetone-phosphate (DHAP): The first and most widespread route is utilized (mainly) by respiring organisms [6-8], which first convert intracellular glycerol into *sn*-glycerol-3-P (G3P) via a glycerol kinase (GK) (encoded by the *glpK* gene) using ATP as phosphoryl donor [9]. GK mediated phosphorylation represents – in addition to the concentration gradient - the driving force for the glycerol uptake since it keeps the intracellular concentration of free glycerol low and confers a negative charge to the molecule which thus cannot freely diffuse outside through the cytoplasmic membrane. Furthermore, the glycerol diffusion facilitator does not recognize G3P as a substrate, and thus the intermediate is trapped inside the cell [10-12]. GKs have been characterized from several Bacteria and Eukarya. They are dimers or tetramers and some including the *E. coli* enzyme show posttranslational regulation via FBP inhibition accompanied by dimer-tetramer transition [13].

Following phosphorylation, one of two membrane bound glycerol-3-phosphate dehydrogenases (G3PDH) hereafter designated as GlpD and GlpABC, encoded either by the *glpD* gene or by the operon comprised of the three genes *glpA*, *B*, and *C*, respectively, catalyze the oxidation of G3P to dihydroxyacetone phosphate (DHAP), both with the simultaneous reduction of a non-covalently enzyme bound flavin adenine dinucleotide (FAD) to FADH₂. The electrons are then further transferred from the FADH₂ to the quinone pool of the respiratory chain [14-17]. GlpD (also called the aerobic G3PDH) is maximally expressed under aerobic conditions for example in *E. coli*. [8, 15-17]. The monotopic, dimeric inner membrane protein transfers electrons to ubiquinone from where they are finally passed on to the terminal electron acceptors oxygen or nitrate. GlpD is also the G3P oxidizing enzyme in the mitochondria of Eukarya [16]. The second G3PDH (GlpABC) mainly known from Bacteria is induced under anaerobic conditions (the anaerobic G3PDH) when glycerol serves as carbon and energy source and accordingly menaquinone of the respiratory chain is reduced with the final acceptors nitrate or fumarate (in *E. coli*) [8, 18]. The A and B subunits (encoded by *glpA* and *B*) form a soluble and active dimer [14] which is likely anchored to the membrane via the C subunit [18, 19]. A third mechanism of G3P oxidation is less widespread among bacteria and mainly restricted to aerotolerant/microaerophilic lactic acid bacteria which are heme deficient and thus lack functional respiratory chains. These organisms employ the G3P oxidase (GlpO, encoded by the *glpO* gene), a cytosolic, soluble FAD-dependent enzyme directly reducing oxygen as electron acceptor yielding H₂O₂ which in turn is then detoxified via peroxidase or catalase [20-23] (Fig. 1). Also, the *Mycoplasma* spp. from the Mollicutes use GlpO for G3P oxidation and the resulting H₂O₂ has been discussed to contribute to pathogenicity of these organisms [24, 25]. Finally, the DHAP produced by G3P oxidation is then channeled into the central metabolism via glycolysis and gluconeogenesis, respectively. G3P can directly be used for (phosphor)lipid biosynthesis.

The second route of glycerol processing is less abundant and restricted to organisms growing under fermentative conditions as described for some Enterobacteriaceae including *E. coli* and few other bacterial species where glycerol is a challenging substrate due to its high reduction state [26, 27]. Here, an external electron acceptor is missing and instead of a quinone a metabolic intermediate needs to serve as electron sink for glycerol oxidation. In this pathway glycerol is first oxidized to dihydroxyacetone (DHA) in an NAD⁺ dependent manner (glycerol dehydrogenase encoded by the *gldA* gene) followed by PEP (*dhaK*) or ATP (*glpK*, *Klebsiella pneumoniae*) dependent phosphorylation to DHAP [8, 26] (Fig. 1). The DHAP then again enters the lower shunt of the EMP pathway finally yielding the major fermentation products of the respective organisms or is utilized for anabolism. The NADH derived from glycerol oxidation is reoxidized with another molecule of glycerol which is first dehydrated to 3-

hydroxypropionaldehyde and then reduced to 1,3-propanediol, or converted to methylglyoxal via DHAP followed by reduction to 1,2-propanediol [26].

Whereas the glycerol metabolism is quite well understood in Bacteria and Eukarya, comparably little is known regarding glycerol degradation in Archaea. The haloarchaeon *Haloferax volcanii* has been shown to utilize glycerol as carbon and energy source and to employ homologues of the bacterial GlpK and GlpABC proteins, while GlpD is absent [28-31]. Furthermore, GK activity has been reported in crude extracts of *H. volcanii* and *Halobacterium* sp. to be induced in the presence of glycerol although basal levels have also been demonstrated in its absence [29, 32]. Comparative sequence analyses revealed GK (GlpK) and G3PDH (GlpABC) homologues in various other Halobacteriales [31, 33]. However, other representatives from the archaeal domain have so far not been shown to grow with glycerol as sole source of carbon and energy although genes encoding putative GKs and G3PDHs were identified in aerobic and anaerobic representatives of the orders Thermococcales, Thermoplasmata, Thermoproteales and Lokiarchaeota [33]. In accordance with the distribution of sequence homologues, GK activities in crude extracts from *Thermoplasma* strains as well as from *Pyrococcus* strains were shown, although the cells were grown in the absence of glycerol [32]. For *Thermococcus kodakarensis*, induction of the G3PDH homologue in the stationary phase of growth was reported [34]. However, although both recombinant GK (also homologous to the bacterial/eukaryal GlpK) and G3PDH have been characterized from *T. kodakarensis* and also the GK crystal structure was solved [35-38], the organism could not be grown with glycerol as carbon and energy source [37]. Also, for the other archaeal organisms possessing GK and G3PDH homologues growth on glycerol remains to be shown.

The thermoacidophilic, aerobic crenarchaeal model organism *Sulfolobus acidocaldarius* grows heterotrophically at temperatures around 80°C and a pH of 3.0 with a variety of different organic carbon sources including mono- (e.g. pentoses and hexoses), di-, and polysaccharides, peptides and amino acids, as well as fatty acids. The central metabolic network particularly the carbohydrate metabolism is well understood which employs a modified unusual branched Entner-Doudoroff pathway (ED) and the Weimberg pathway for hexose and pentose degradation, respectively, as well as the Embden-Meyerhof-Parnas pathway (EMP) for gluconeogenesis [39, 40]. The organism's genome is fully sequenced and genetic tools have been developed (for literature see [40]). In *S. acidocaldarius* and other Sulfolobales, only GK homologues have been identified whereas G3PDH homologues appeared absent [33] and they have been indicated not to grow on glycerol [41]. Since it has recently been shown that *S. acidocaldarius* cleaves triacylglycerides by means of esterases and to grow with the fatty acids as sole source of carbon and energy, we reexamined the growth of *S. acidocaldarius* on

the other lipid cleavage product glycerol. Using growth studies, polyomics approaches, crude extract enzyme measurements, and biochemical analyses of the enzymes involved, we herein demonstrate that *S. acidocaldarius* utilizes glycerol as growth substrate and employs a conserved “classical” GK (homologous to GlpK) for glycerol phosphorylation. However, G3P oxidation is catalyzed by a truncated GlpA-subunit like G3PDH lacking the B and C subunits of the classical, bacterial GlpABC complex. Instead, it shows an unusual type of membrane association facilitated by a small CoxG-related protein.

Results

***S. acidocaldarius* glycerol catabolism involves GK and G3PDH**

To analyze growth and glycerol consumption by *S. acidocaldarius* MW001, cells were incubated in standard Brock medium containing glycerol (10 mM, 20 mM, and 40 mM) as sole carbon and energy source (Fig. 2a, b, and c, respectively). After a prolonged time of adaptation (several weeks), *S. acidocaldarius* MW001 grew exponentially with similar growth rates of 0.029386 h^{-1} (10 mM glycerol), 0.0280343 h^{-1} (20 mM glycerol), and 0.0284 h^{-1} (40 mM glycerol) up to cell densities of $\text{OD}_{600} 4$ (at 40 mM glycerol) and glycerol was completely consumed (Fig. 2a-c). The adapted strain was referred to as *S. acidocaldarius* MW00G. The final OD_{600} of 4 at 40 mM glycerol corresponds to a cell dry weight of approximately 1.33 g/l and hence a molar yield coefficient of 33.3 g CDW/mol of glycerol. Lowering the glycerol concentrations results in enhanced growth yields of 41 g CDW/mol (20 mM glycerol) and 50 g CDW/mol glycerol (10 mM glycerol) (Fig. 2b and c). For comparison, *S. acidocaldarius* was grown with 0.2 % (w/v) D-xylose as sole carbon and energy source leading to a final OD_{600} of 0.8 with slower growth rates of 0.0195 h^{-1} (Fig. 2d). This corresponded to a growth yield of 0.27 g/l CDW and thus a molar yield coefficient of 20 g CDW/mol D-xylose (Fig. 2d).

To further elucidate the glycerol catabolism in *S. acidocaldarius* MW00G both, the carbon source dependent global transcriptional and translational response to glycerol compared to D-xylose was studied using transcriptomics (RNA-Seq) and proteomics (LC-MS-MS). For a deeper analysis we decided to concentrate on genes/proteins that were differentially regulated with at least a log₂-fold change of 2 in all glycerol growth conditions both on a transcriptional and translational level. Under these conditions a total of 39 transcripts/proteins were significantly upregulated while a total of 14 transcripts/proteins were significantly downregulated (Supplemental table 3). Among them we observed a downregulation of the Weimberg pathway for pentose degradation (*saci_1938* – α -ketoglutarate semialdehyde dehydrogenase and *saci_1939* - 2-dehydro-3-deoxy-D-arabinonate dehydratase) as well as of the sugar binding subunit of the xylose/arabinose transporter *saci_2122*, while other central carbohydrate metabolic pathways (e.g. glycolytic branched Entner-Doudoroff pathway (ED), and tricarboxylic acid cycle (TCA), as well as the EMP for gluconeogenesis) were found to be unregulated. In agreement with glycerol dissimilation, a significant upregulation of the gene cluster *saci_2031* to *saci_2034* (Fig. 3a) and its encoded proteins was observed. The gene cluster encodes a putative G3PDH (*Saci_2032*) with a downstream located carbon monoxide dehydrogenase subunit G (CoxG)-like protein (*Saci_2031*), as well as a putative GK (*Saci_2033*) and a glycerol uptake facilitator (GUF, *Saci_2034*). Both gene couples *saci_2032-2031* and *saci_2033-2034* are divergently oriented as depicted in Fig. 3b. In accordance with

the observed regulation pattern, we also observed the induction of the GK and the G3PDH on activity level in crude extracts of glycerol compared to D-xylose grown cells (Fig. 3c). Whereas in D-xylose grown cells neither GK nor G3PDH activity could be determined under the applied assay conditions, a specific activity of 0.3 U mg^{-1} for GK and of 0.12 U mg^{-1} for G3PDH using 2,6-Dichlorophenolindophenol (DCPIP) as artificial electron acceptor was determined. NAD^+ dependent G3P oxidation was not detectable. Also, neither DCPIP nor NAD^+ dependent glycerol oxidizing activity could be observed, indicating that no alternative pathway for glycerol dissimilation is present. Notably, the DCPIP dependent G3PDH activity was not only located in the soluble fraction (crude extract) after cell disruption but also in the resuspended pellet/membrane fraction with 0.23 U mg^{-1} (Fig. 3c), whereas GK activity was exclusively found in the soluble fraction.

Next, we performed comparative targeted LC(HILIC)-MS/MS based metabolome analyses of glycerol and D-xylose grown cells with a special focus on the intermediates of the central metabolic pathways like modified branched ED pathway, gluconeogenesis, TCA cycle, and of course the hypothetical intermediates of the glycerol catabolism. Interestingly, free glycerol was almost absent in glycerol grown cells, while the concentrations particularly of G3P and also of DHAP (~20% compared to G3P) was significantly increased (Fig. 3d). Additionally, in agreement to our transcriptome and proteome data no significant changes in metabolites in the upper EMP pathway for gluconeogenesis (i.e. glucose-6-phosphate (G6P), glucose-1-phosphate (G1P), fructose-6-phosphate (F6), and fructose-1,6-bisphosphate (F1,6BP)) and the common lower shunt of ED and EMP (i.e. glyceraldehyde-3-phosphate (GAP) and 2-phosphoglycerate (2PG)) were observed.

Together, these results show that *S. acidocaldarius* is able to utilize glycerol as sole carbon and energy source and that the glycerol uptake facilitator (Saci_2034, GlpF-like) is likely involved in glycerol transport via facilitated diffusion. Glycerol is then subsequently phosphorylated to G3P (GK (GlpK-like), Saci_2033), and further oxidized by an at least loosely membrane associated G3PDH to yield DHAP which is then channelled either into the reversible, lower common shunt of the ED and EMP pathway for glycolysis or the upper EMP shunt for gluconeogenesis.

Purification and characterization of GK and G3PDH

To confirm the function, C-terminally Twin-Strep tagged GK (Saci_2033) was homologously produced in *S. acidocaldarius* MW001 since all expression efforts in *E. coli* were not successful. The putative GK encoded by *saci_2033* comprises 497 amino acids and has a calculated molecular mass of 55.3 kDa. This coincides well with the molecular mass of

approximately 55 kDa experimentally determined under denaturing conditions by SDS-PAGE for the recombinant Twin-Strep tagged protein (supplementary Fig. 1). The molecular weight under native conditions as determined via size exclusion chromatography was ~110 kDa, indicating a homodimeric structure (α_2) (supplementary Fig. 2). GK activity was confirmed as glycerol dependent formation of ADP from ATP in a coupled assay using PK-LDH at 50°C. Initial rate dependence on substrate concentrations followed classical Michaelis-Menten kinetics and a V_{\max} value of 85 U mg⁻¹ and K_M values of 0.024 mM (glycerol) and 0.17 mM (ATP) were determined (Fig. 4, Table 1). With glyceraldehyde (GA) and dihydroxyacetone (DHA) the enzyme showed only minor activities with catalytic efficiencies of <2% compared to glycerol (V_{\max} similar to glycerol but much higher K_M values) (supplementary Fig. 3 a and b). Other substrates like D(+)-glyceric acid, D-glucose, D-sorbitol, D-xylose, xylitol, and erythritol were not accepted as substrate. The pH optimum was at 6.5 and the temperature optimum at 75°C and thus corresponded well with the internal pH and optimal growth temperature of *S. acidocaldarius* (Table 1, supplementary figure 3c and d). Accordingly, the Saci GK was also stable against thermal inactivation with half-life times of 12 h and 2 h at 70°C and 80°C, respectively. However, no remaining activity could be observed after 1 h of incubation at 90°C (supplementary Fig. 3e). In addition to ATP also GTP and CTP could serve as phosphoryl donor, however with only 20% and 45% residual activity, respectively (supplementary Fig. 4a). Fructose-1,6-bisphosphate a known inhibitor of GK activity in e.g. *E. coli* [13] had no effect on the Saci_2033 activity (supplementary Fig. 4b).

The *saci_2032* gene is annotated to encode a FAD-dependent oxidoreductase/G3PDH consisting of 428 amino acids corresponding to a calculated molecular mass of 46.9 kDa. Heterologous expression in *E. coli* and purification of the N-terminally His-tagged enzyme yielded a soluble, yellow protein with a subunit size of 47.6 kDa (SDS-PAGE) and a native molecular mass of 94 kDa (supplementary Fig. 5a and b) determined by size exclusion chromatography thus representing a homodimer. The yellow colour already indicated the presence of a FAD cofactor and by spectrophotometric measurements (maximal absorption at 450 nm, extinction coefficient of 11.300 mM⁻¹ cm⁻¹, see materials and methods part) a FAD content of two for the homodimer, i.e., one FAD molecule per monomer, was determined (Fig. 5a). Furthermore, after reduction with G3P without the addition of an artificial electron acceptor the bound FAD cofactor remained stable in its colourless, reduced state and did thus not immediately react with oxygen which would have otherwise been indicated by the formation of its yellow oxidized form. Furthermore, the formation of hydrogen peroxide was excluded using the 2,2'-azinobis-3-ethylbenzothiazoline-6-sulfonate (ABTS) assay as described in the materials and methods section (data not shown). For G3P a K_M of 0.055 mM and a V_{\max} of 44.5 U mg⁻¹ determined with DCPIP as artificial electron acceptor (Fig. 5b). Under these

conditions, no activity could be detected with glycerol-1-phosphate (G1P), glycerol, glyceric acid, glyceraldehyde, glyceraldehyde-3-phosphate (GAP), and phosphoglyceric acid. Also, no activity could be observed with NAD⁺ or NADP⁺ as electron acceptor neither with G3P nor with glycerol, DHAP or glycerol-1-phosphate as substrate. However, ubiquinone-Q1 served as electron acceptor as followed by G3P dependent loss of absorption of the quinone at 280 nm (Fig. 5c). A K_M of 0.179 mM and V_{max} of 21.1 U mg⁻¹ for ubiquinone-Q1 (Fig. 5d) were determined. The enzyme showed a pH and temperature optimum of 6.5 and 70°C, respectively (Table 1, supplementary Fig. 6a and b). Also, the thermostability of the Saci_2032 G3PDH was analysed and no loss in activity was observed upon more than 6 h incubation at 70°C. At 80°C Saci_2032 G3PDH had a half-life time of 3 h, and no remaining activity could be observed after 1 h of incubation at 90°C (supplementary Fig. 6c).

A CoxG homologue serves as membrane anchor for *S. acidocaldarius* G3PDH

The ability of the Saci_2032 G3PDH to reduce quinones and thus likely to transfers electrons to the membrane-located respiratory chain, raised questions about the mechanism of membrane association of G3PDH in *S. acidocaldarius*. Interestingly, in addition to GUF (Saci_2034), GK (Saci_2033), and G3PDH (Saci_2032), transcriptomic and proteomic analysis of *S. acidocaldarius* also showed the upregulation of *saci_2031* annotated to encode a CoxG-homolog (Fig. 3a). As shown by the transcriptomics data, both *saci_2032* and *saci_2031* form an operon with *saci_2031* located in the same orientation one base pair downstream of and cotranscribed with *saci_2032*. However, Saci_2032 G3PDH was readily active *in vitro* in the absence of Saci_2031 indicating that Saci_2032 does not essentially rely on Saci_2031 for activity or electron transfer even to quinones (see above). Nevertheless, CoxG has been described as a catalytically non-essential subunit of the (aerobic) carbon monoxide dehydrogenase and was proposed to function as membrane anchor for the enzyme complex of *Oligotropha carboxidovorans*. [42]. A similar function was discussed for the 3-hydroxypyridine dehydrogenase of *Ensifer adhaerens* HP1 where a CoxG homolog was proposed to facilitate membrane binding [43].

To elucidate a potential role of the CoxG (Saci_2031) as membrane anchor for the G3PDH we heterologously co-overexpressed *saci_2032* and *saci_2031* in the pETDuet-1-vector in *E. coli*. As a control, *saci_2032* was expressed alone from the same vector. Afterwards membrane and soluble proteins were separated by ultracentrifugation and the localization of His tagged Saci_2032 was analyzed via western blotting and immunodetection using anti-His antibodies. Only the coexpression of *saci_2032* and *saci_2031* resulted in a localization of G3PDH at the cell membrane. In contrast, the expression of *saci_2032* in the absence of *saci_2031* did not yield any membrane bound G3PDH, instead the protein was exclusively localized in the

soluble, cytoplasmic fraction in *E. coli* (Fig. 6a). The signal detected in the cytoplasmic fraction for the G3PDH in the cells coexpressing *saci_2032* and *saci_2031* might be due to a less efficient *saci_2031* expression compared to *saci_2032*, so that not all G3PDH could be recruited to the membrane. Furthermore, upon homologous expression in *S. acidocaldarius* MW001 as HA-tagged protein, Saci_2031 was found to a high degree in the insoluble/membrane fraction from which it could be solubilized using n-dodecyl β -D-maltoside (DDM). These findings provided further evidence that the CoxG homolog is membrane associated in *S. acidocaldarius in vivo*. (Fig. 6b). To further confirm interaction between Saci_2031 and Saci_2032 a coimmunoprecipitation was performed using C-terminally HA-tagged CoxG (Saci_2031). MS analysis of interacting proteins revealed a specific interaction of Saci_2031 with Saci_2032 (Fig. 6c, supplementary Table 4). Together these results strongly indicate that the G3PDH transfers the electrons from G3P oxidation to the quinone pool of the respiratory chain (caldariellquinone in *S. acidocaldarius*) and that the required membrane interaction is mediated by the CoxG homolog Saci_2031, thus representing a novel function of CoxG homologues in Archaea and an unusual mechanism of membrane anchoring of G3PDH in *S. acidocaldarius*.

Discussion

S. acidocaldarius has previously been shown to cleave short chained triacylglycerols like tributyrin by means of esterases and was also shown to grow with short chain fatty acids like butyrate and hexanoate as sole carbon and energy sources [44, 45]. Therefore, the question arose whether this crenarchaeal model organism is also able to grow with glycerol, the other product of esterase mediated lipid breakdown. Herein, growth of *S. acidocaldarius* with glycerol as sole carbon and energy source was demonstrated. To further elucidate the glycerol degradation, transcriptomics and proteomics as well as targeted metabolomics/metabolite analyses and enzyme activity measurements in crude extracts, both in comparison to D-xylose grown cells, were performed. Furthermore, the enzymes upregulated at gene and protein level involved in glycerol degradation were recombinantly produced, purified and characterized. It was demonstrated that: (i) In both, transcriptomics and proteomics *saci_2031-2034*, encoding CoxG, G3PDH, GK, and GUF were strongly up-regulated in the presence of glycerol (compared to D-xylose); (ii) The activities of GK and G3PDH (with DCPIP as electron acceptor not with NAD) were highly induced in glycerol but not measurable on D-xylose grown cells, and also activities of alternative pathways for glycerol utilization appeared to be absent; (iii) The key intermediates of the pathway, especially G3P and also DHAP were much more abundant in glycerol grown cells; (iv) The enzymatic function of G3PDH (*Saci_2032*) and GK (*Saci_2033*), respectively, was confirmed by characterization of the recombinant proteins; (v) The G3PDH activity was at least partly associated with the membrane which is mediated by the CoxG homologue *Saci_2031*. These results demonstrate that glycerol degradation in *S. acidocaldarius* is mediated via glycerol phosphorylation catalyzed by a “classical” GK followed by FAD-dependent G3P oxidation with electron transfer to the quinone of the respiratory chain carried out by a structurally unusual G3PDH. The resulting DHAP is then channeled into central metabolism (Fig. 7).

Glycerol degradation in Archaea has so far only been analyzed in some detail in the halophile *H. volcanii* which also grows on glycerol minimal media. Similarly, to *S. acidocaldarius*, the growth of *H. volcanii* on glycerol was better than on sugars (e.g. D-xylose) in terms of growth yield per mol of substrate carbon [30, 46]. This higher growth yield and hence the higher ATP gain per mol of substrate is on the one hand due to the higher reduction state of glycerol compared to sugars like D-xylose but can likely not account alone for the observed yield differences and might hence hint at a more efficient energetic coupling during growth on glycerol. Accordingly, in the transcriptomics and proteomics data we observed changes in the respiratory chain of *S. acidocaldarius* with the SoxEFGHIM (*saci_2258-saci_2263*), one of three terminal oxidases, being strongly upregulated (supplementary Tab. 3). This terminal oxidase was shown to have a higher H⁺/e⁻ ratio [47-50] and thus likely additionally contribute

to the higher growth yield. In contrast to the results presented herein, *S. acidocaldarius* has been previously reported not to grow on glycerol [41] and in agreement G3PDH homologues were shown to be absent in Sulfolobales [33]. These findings can likely be explained by the prolonged time the organism required for adaptation to glycerol as carbon and energy source observed in this study and at sequence level by the C-terminal truncation/differences of the G3PDH of *S. acidocaldarius* and other Sulfolobales (as well as some Thermoprotei and Thermoplasmatales) (see below). Such an adaptation phase - although much less pronounced - was also reported for *Pseudomonas* spp. and was attributed to the mode of regulation with G3P as inductor [8]. Like *S. acidocaldarius*, *H. volcanii* utilizes the GK/G3PDH pathway for glycerol degradation. Although GK activity was induced in presence of glycerol, basal activity levels in absence of glycerol were detected indicating constitutive expression of the GK/G3PDH pathway in halophiles [29, 32]. In *S. acidocaldarius* the GK/G3PDH activities are only detectable when cells are grown on glycerol suggesting a much tighter regulation. This different expression pattern might reflect that glycerol in halophilic habitats is regarded as main and preferred carbon and energy source (see references in [29]) which is likely different in thermoacidophilic environments.

In accordance with previous findings [33, 51], bioinformatic/phylogenomic analyses revealed that - in addition to Halobacteriales and Sulfolobales - also other archaeal lineages harbor the GK/G3PDH pathway. However, the GK/G3PDH homologues and hence the potential ability to utilize glycerol is present only in some representatives of these lineages and appears thus only patchily distributed in archaea [33].

Glycerol uptake

For glycerol uptake, *S. acidocaldarius* harbors a GUF homologous to GlpF [1] and up-regulated in response to glycerol and thus likely involved in glycerol uptake. In Archaea, GlpF homologues seems to be restricted to few *Sulfolobus* and *Saccharolobus* spp. as well as to halophiles [30, 31] and are thus much less abundant than GlpA homologues in archaea meaning that by far not all archaea harboring the GK and G3PDH also contain a GlpF homologue. This might indicate that alternative transporters than GlpFs are utilized for glycerol uptake in other Archaea and the gene neighborhoods of the GK/G3PDH encoding genes often also comprise e.g. putative MFS (major facilitator superfamily) transporter encoding genes (e.g. in *Thermococcus* and *Pyrococcus* species, some Thermoplasmatales, *Vulcanisaeta*). As described above, alternative glycerol transporters are known in *Mycoplasma* spp. employing ABC transporters (in addition to a GlpF homologue) [3, 4] and also in yeast using H⁺ (Na⁺) symport systems. In *S. acidocaldarius*, in addition to the *glpF* gene, also a putative ABC transporter (*saci_1762-1765*) was up-regulated which could thus also be involved in glycerol

transport. In addition, glycerol uptake via simple diffusion might also take place in *S. acidocaldarius* and other Archaea. For example, in *E. coli* deletion of the *glpF* gene had no severe impact on glycerol consumption under high concentrations. Nevertheless, below 5 mM glycerol in the medium the growth rate remarkably decreased indicating that the facilitator is especially important at low substrate availability [2]. Further on, besides the concentration gradient a major driving force of the glycerol uptake was shown to be the subsequent ATP-dependent phosphorylation catalyzed by the GK [8, 10]. Accordingly, intracellular glycerol was nearly undetectable in *S. acidocaldarius*, whereas the G3P and to a lesser extend also the DHAP concentrations were significantly increased during growth on glycerol compared to D-xylose. Furthermore, the GK and G3PDH activities in crude extracts of *S. acidocaldarius* cells were clearly induced to 0.3 U mg^{-1} and 0.12 U mg^{-1} , respectively, but were not detectable in D-xylose grown cells. Similar GK activities were reported for *H. volcanii* (0.43 U mg^{-1}) [29] or for *E. coli* (0.61 U mg^{-1}) [2] and for the latter a further activation of the GK activity through interaction with the GlpF was suggested [10]. The GK and G3PDH activities agree with the much more pronounced accumulation of G3P compared to DHAP in *S. acidocaldarius* and also show that both activities are not rate limiting. The transcriptomics and proteomics analyses revealed that candidates for further DHAP conversion via glycolysis and gluconeogenesis as well as the TCA cycle are not regulated in response to glycerol which thus likely accounts for the accumulation of G3P and DHAP.

Properties of the glycerol kinase Saci_2033

The *S. acidocaldarius* GK like all other GKs characterized so far belong to the FGGY family within the sugar kinase/HSP70/actin superfamily [52] and shows a high degree of sequence conservation not only to other archaeal homologues but also to bacterial (>50% sequence identity) and even eukaryotic (~44%) enzymes and thus across the domains of life. The homodimeric structure of Saci_2033 is well in accordance with some other GKs from eukaryotes (e.g. *Trypanosoma brucei gambiense*, *Plasmodium falciparum*, *Chaetomium thermophilum* [53-55]) and bacteria (e.g. *Cellulomonas* sp. [56]). Also, homotetramers (for some Bacteria like *Thermus thermophilus*, *Elizabethkingia meningoseptica*, and also some Eukaryotes like *Saccharomyces cerevisiae* [57-59]) and even -hexamers (*T. kodakarensis* [35]) have been reported. For *E. coli* a dimer-tetramer equilibrium in solution was described [13] and also for *Haemophilus influenzae* a higher (>4mer) and a lower (<4mer) oligomerization state was reported [60]. Two different oligomeric species were clearly not observed for Saci_2033 which eluted solely as homodimer in size exclusion chromatography experiments (supplementary Fig. 2). The tetramer formation observed for the *E. coli* GK appears a

prerequisite for allosteric regulation by fructose-1,6-bisphosphate (FBP) [13] and thus appears in accordance with the missing FBP regulation observed for the homodimeric Saci_2033 enzyme. Also, for the dimeric enzymes from *P. falciparum* and *T. brucei gambiense* no allosteric control with FBP was observed [53, 61]. Similarly, the *T. kodakarensis* GK initially described as a homodimer did not respond to FBP as allosteric effector. Later on, it has been shown that in presence of glycerol the *Thermococcus* enzyme assembles into hexamers but the subunit interactions appear different than observed in the tetramer interactions in *E. coli* essential for FBP binding and seems thus also in agreement with the missing FBP regulation [35, 36]. The glycerol mediated hexamer formation of the *Thermococcus* GK increases the ATP affinity of the enzyme tenfold. Under the same glycerol concentration, however, the Saci_2033 GK did not change its oligomeric structure (supplementary Fig. 2) and the lysine 271 (K271) in *T. kodakarensis* enzyme essential for hexamer formation is not conserved in Saci_2033. Additionally, from *Enterobacteriaceae* (e.g. *E. coli*) and *Firmicutes* (e.g. *Enterococcus*, *Streptococcus*, *Bacillus*), two different GK regulation mechanisms involving components of the PEP-dependent phosphotransferase system (PTS) for sugar transport have been described (for literature see [62]). In *E. coli* the unphosphorylated Enzyme IIA^{Glc} acts as allosteric GK inhibitor and in firmicutes GK is phosphorylated by HPr~P at a conserved histidine residue (H232 in *Enterococcus casseliflavus*) located in the loop region where also FBP is bound (for literature see Yeh et al., 2009). This phosphorylation in the absence of PTS sugars stimulates the enzyme activity 10-15-fold [62]. Similar regulatory mechanisms involving components of the sugar transport PTS systems are unlikely to occur in *S. acidocaldarius* since PTS systems in this organism and in archaea in general are absent despite in some halophiles [63]. Interestingly, in *H. volcanii* one PTS component protein, a HPr homologue encoded by the ptsH2 gene (HVO_1543), is part of the chromosomal glycerol gene cluster [29, 30]. However, the phosphorylated H232 in the *E. casseliflavus* enzyme is not conserved in the *H. volcanii* GK (and also not in Saci_2033). The kinetic constants of the Saci_2033 GK with K_M values of 0.024 mM (glycerol) and 0.17 mM (ATP) and a V_{max} around 85 U mg⁻¹ (at 50°) are well in the (rather broad) range of those determined for other GKs. The V_{max} of the Saci_2033 GK at a more physiological temperature of 75°C was determined at 340 U mg⁻¹ which appears comparatively high but roughly approximates a similar range reported for the *T. kodakarensis* GK (>1000 U mg⁻¹). A relaxed substrate specificity as observed in Saci_2033 GK converting also DHA and GA although with much lower catalytic efficiencies (<2%) (supplementary Fig 3.) was also reported e.g. for the *E. coli* enzyme [64]. The temperature optimum (75°C) and the thermal stability with half-life times ($t_{1/2}$) of 12 h (70°C) and nearly 2 h at (80°) fit well with the growth conditions of *S. acidocaldarius*. Nevertheless, in accordance with the growth optimum of *T. kodakarensis* of 85°C the thermal properties of the GK from this euryarchaeon are more pronounced (T_{opt} 80°C, $t_{1/2}$ 30 min (100°C)) [38, 65]. The GK from the thermophilic

bacterium *Thermus thermophilus* on the other hand showed slightly lower T_{opt} between 50°C and 70°C and lost ~75% activity upon 30 min incubation at 70°C [57].

Properties of the G3PDH Saci_2032 and an unusual mode of membrane anchoring by CoxG Saci_2031

In contrast to the *S. acidocaldarius* GK/G3PDH pathway, the route in *H. volcanii* involves a G3PDH composed of all three GlpABC subunits homologous to the “anaerobic” route known from many especially facultatively anaerobic bacteria like *E. coli* [7, 8, 18, 19, 30]. In Bacteria and haloarchaea, the GlpC and GlpB subunits are supposed to permit membrane anchoring and electron transfer, respectively, from the catalytic subunit GlpA finally to the (mena)quinone pool in *E. coli* [18, 19]. However, except for most haloarchaea the GlpB and C subunits are absent from nearly all other archaea and also GlpD and GlpO homologues were not identified. These organisms in addition to GlpK (and frequently also GlpF) only contain GlpA homologues which were however demonstrated to be active without additional subunits herein for *S. acidocaldarius* and also previously for *T. kodakarensis* [37]. The TK1393 G3PDH catalytic subunit from *T. kodakarensis* alone showed a 7-8-fold lower V_{max} of 5-6 U mg^{-1} compared to the *S. acidocaldarius* enzyme and a 50-fold higher K_M value (2.6 mM vs. 0.055 mM). However, the G3PDH encoding genes are often downstream co-localized with different likely functionally related genes (for detailed discussion see below) which suggest – together with the absence of the GlpB and C subunits - that electron transfer in these archaeal species might be different. We herein forwarded evidence that the *S. acidocaldarius* G3PDH forms a catalytically active dimer which can directly reduce the (ubi)quinone analogue Q1, two features that are similar to characterized bacterial and eukaryotic GlpDs [16, 66, 67] which directly transfer the electrons to the quinone pool. The use of quinones (E^0 : ~+110 mV (ubiquinone) and -75mV (menaquinone)) as acceptor for the relatively high potential electrons of G3P (~-200 mV) oxidation is advantageous since $NAD(P)^+$ (-320 mV) would render this reaction endergonic [68]. Also, the FAD content of one per monomer is more similar to the GlpDs. In contrast, the catalytic heterodimer from *E. coli* (GlpAB) was described to contain only one FAD and also an FeS cluster both likely located in the GlpA subunit. How the GlpB subunit acts in the electron transport is not fully understood yet. Sequence analyses suggest that the GlpB subunit, which is homologous to several flavoproteins like the FAD subunit of the succinate dehydrogenase, might contain a further flavin cofactor involved in electron transport to the FeS cluster(s) in GlpC [18].

The *S. acidocaldarius* G3PDH showed similar specific activities with DCPIP as artificial electron acceptor as described for the GlpD from *Vibrio alginolyticus* with the artificial redox

carriers PMS-MTT or ferricyanide (see supplementary table 5). For the GlpDs from *E. coli* a roughly ten-fold lower activity with DCPIP has been reported which was even lower for the pig brain mitochondrial enzyme (0.8 U mg^{-1}). However, for the pig brain enzyme also the kinetic constants for the more native substrates, the ubiquinones-Q0 and -Q1, have been determined. For the ubiquinone-Q1 both the K_M and V_{\max} values were more than 10-fold lower ($V_{\max} 2.6 \text{ U mg}^{-1}$, $K_M 0.013 \text{ mM}$) [69] compared to the *S. acidocaldarius* enzyme ($V_{\max} 21 \text{ U mg}^{-1}$, $K_M 0.18 \text{ mM}$) resulting in a similar catalytic efficiency. The high substrate specificity observed for the *S. acidocaldarius* G3PDH not converting G1P, glycerol, glyceric acid, GAP, and phosphoglyceric acid has also been reported for two isoenzymes from *Acidiphilium* sp. [70].

However, as further detailed below the Saci_G3PDH is more similar to GlpAs. So far, the GlpAB active heterodimer from *E. coli* is the only bacterial GlpA-like enzyme which has been kinetically characterized in some detail. The V_{\max} value reported for the *E. coli* enzyme (34 U mg^{-1}) is in a similar range as the *S. acidocaldarius* G3PDH homodimer (44.5 U mg^{-1}) although the K_M for G3P is 6-fold lower for the latter (0.055 mM vs. 0.34 mM for the *E. coli* GlpAB heterodimer) [14]. Conversely, the direct reduction of quinone analogues was not described for the *E. coli* enzyme likely because it relies on the GlpC subunit for both, membrane anchoring and quinone reduction (electron transfer) (also underscored by the presence of a FeS cluster(s) in the C subunit) [19]. In *S. acidocaldarius* the genes encoding G3PDH (*saci_2032*) and the CoxG homologue (*saci_2031*) were shown to form an operon (up-regulated in response to glycerol) and the proteins were shown to interact with each other. The CoxG homologue recruits the G3PDH active dimer to the membrane, but as demonstrated by the quinone reactivity of the Saci_2032 G3PDH dimer alone the CoxG (subunit) is not involved in electron transfer and thus only anchors the protein to the membrane. The only functionally analyzed CoxG homologue so far is the protein from *O. carboxydovorans* which recruits the carbon monoxide dehydrogenase (from the aerobic xanthin oxidase type) to the membrane but does not seem to have any enzymatic implications [42]. In contrast, the homodimeric “aerobic” G3PDH (GlpD), e.g. from *E. coli*, does not require any membrane anchoring and/or additional electron transferring subunits but as a monotopic membrane protein directly interacts with the membrane and transfers electrons to the ubiquinone pool. Furthermore, these GlpDs were also shown to require membrane interaction, or phospholipids or nondenaturing detergents for conformational integrity and thus full activity ([16] and the literature cited therein). This was not the case for the Saci_G3PDH and might further reflect the different mode of membrane association (supplementary Fig. 7).

Sequence and structural comparisons of G3PDHs

As indicated by sequence comparisons (Fig. 8), conserved domain analyses, and structural comparison including modeling using the “ColabFold: AlphaFold2 using MMseqs2” online tool [71] (supplementary Fig. 8), the Saci_G3PDH and other archaeal homologues mentioned above are homologous in their first N-terminal ~360 amino acids with GlpDs, GlpAs and also with GlpOs. This part of the proteins constitutes the FAD-dependent oxidoreductase domain (pfam01266) which belong to the D-amino acid oxidase (DAAO) superfamily (including e.g. also the glycine oxidase from *Bacillus subtilis* (1ng3) [72]). These proteins are composed of a ‘glutathione-reductase-2’ type FAD-binding domain and an antiparallel β -sheet based substrate-binding domain (for details and literature see [24]). Accordingly, the FAD binding residues and also the G3P binding site are quite conserved (Fig. 8). However, GlpD, GlpO, and GlpA proteins as well as the archaeal homologues differ in length and domain organization of their C-terminus. Glycine oxidases just constitute the DAAO fold without any C-terminal extensions. However, as revealed by the crystal structures, the C-terminus in GlpD from *E. coli* (2qcu) forms a truncated G3P oxidase domain (pfam 16901, composed of roughly 120 amino acids) shortened by an twenty amino acids comprising helix in its very C-terminal end compared to the complete G3P oxidase domain (140 aa) present in GlpO from *Streptococcus* sp. (2rgo), (supplementary Fig. 8 as well as supplementary figure 9a and b). In addition to the complete C-terminal G3P oxidase domain, GlpOs differ from GlpDs by the presence of a 50 amino acid insertion in the C-terminal part of the DAAO fold (Fig. 8). In contrast to GlpDs and GlpOs, the C-terminus of GlpAs is even larger (160 aa) containing a bfd-like domain with an FeS cluster (pfam04324) [52, 73-75] and the four highly conserved cysteine residues likely involved in FeS cluster binding are clearly visible in the C-terminus of GlpA sequences but missing in GlpDs and GlpOs (Fig. 8, supplementary Fig. 10). The presence of such an FeS cluster is in accordance with the presence of two non-heme irons per catalytically active GlpAB subunit early reported [14]. As indicated by structural predictions the C-terminal domain of GlpA resembles in its first part the C-terminal domain of GlpOs (and thus also of gGlpDs) but is even further extended by an additional α helix in its very C-terminal end compared to GlpO (supplementary Fig. 8 and 9).

Notably, all the archaeal G3PDH homologues differ from GlpD, GlpO, and GlpA enzymes in their C-terminus: The *Thermofilum pendens* related homologues nearly completely lack a C-terminal extension and the DAAO fold is only extended by a very short α helix of 5-6 amino acids (if at all) (supplementary Fig. 10a). In the *S. acidocaldarius*-like archaeal G3PDH homologues the C-terminal tail comprises roughly 50 amino acids which does not form any known or annotated domain comprising only three α helices (and shows also no similarities to GlpA, GlpO, or GlpD) (supplementary Fig. 10b). The Thermococcales G3PDHs as indicated

by model predictions seem to contain the bfd-fold (pfam04324) like in GlpA with four cysteine residues present (and align well with the GlpAs, Fig. 8), although the overall fold of the C-terminal domain appears to be different and somewhat shorter (120 aa) (supplementary Fig. 10c and d).

The structural differences in C-terminal domains of GlpDs, GlpOs and GlpAs seem to be linked to their different functions. In GlpD, the C-terminus was demonstrated to be involved in dimer formation whereas in non-membrane associated, cytoplasmic dimeric GlpOs that are reactive with oxygen (in contrast to glpDs) the C-terminus is not involved in subunit interaction [16, 20]. For GlpA - although no crystal structure is available - the presence of an FeS clusters in the C-terminus suggest that this protein moiety is involved in electron transfer to GlpB which might necessitate their physical interaction, which in turn is supported by the fact that the GlpAB forms the catalytically active unit. The ColabFold prediction of the Saci_G3PDH complex (based on the finding that Saci_2032 alone forms a soluble dimer) (see supplementary figure 11) suggests that also in the Saci_G3PDH the (dissimilar) C-terminus might be involved in dimer formation, which, however, leads to different spatial orientation of the DAAO domains to each other than observed in GlpD. This different spatial orientation of the dimer together with CoxG interaction then likely enables optimal transfer of electrons to the quinone pool in the cytoplasmic membrane in *S. acidocaldarius*.

Distribution and phylogenetic affiliation of archaeal GKs and G3PDHs

Previous studies have already addressed the question of distribution and phylogenetic affiliation of glycerol degrading genes/enzymes in Archaea [24, 33, 51]. Complete sets of glycerol degrading enzymes and thus the potential ability to utilize glycerol as sole carbon and energy source were mainly found in the euryarchaeal orders of Thermococcales, Halobacteriales, Archaeoglobales and Thermoplasmatales and in the crenarchaeal lineages of Thermoproteales, Thermofilales, Desulfurococcales, and Sulfolobales [33, 51]. However, the GK/G3PDH homologues are present only in some representatives (not in all or at least not in the majority) of these lineages and thus only patchily distributed in archaea. It was also indicated that the evolutionary history of GK and G3PDH genes in archaea is different with the GK acquired by horizontal gene transfer from bacteria whereas the G3PDH might have been already present in the universal common ancestor as well as the bacterial and archaeal common ancestors [51]. This is also reflected by the higher overall sequence identity among the archaeal GKs (40-70% identity) compared to archaeal G3PDHs (20-50% identity) as well as by the fact that the closest homologues to GK and G3PDH from *S. acidocaldarius* were identified in different archaeal lineages, i.e., in Thermococcales (61-68% identity) and

Thermoproteales (45-48% identity), respectively. The different evolutionary history of GKs and G3PDHs points towards the conserved function of GK in all the organisms whereas the G3PDH shows more functional variability with respect to interaction partners, membrane anchoring, and electron acceptors (see below).

Phylogenetic analyses of the G3PDH sequences from all domains resulted in similar tree topologies as previously reported [30, 51], however, closer inspection of structural and sequence features including gene neighborhood analysis provided deeper functional insights that have been previously missed as illustrated in Fig. 9: Firstly, the FAD dependent G3PDHs cluster according to their different C-terminal extensions: GlpAs, GlpDs, and GlpOs form distinct subgroups within one main cluster within the DAAO superfamily with GlpDs and GlpOs more closely related to each other than to GlpAs. In contrast, the mycoplasma GlpOs appear only distantly related to glpDs/As/and Os [24]. The G3PDH from Sulfolobales and Thermoplasmatales additionally including the *Caldivirga maquilingensis* and *Vulcanisaeta moutnovskia* homologues as well as those from Thermofilales form distinct subclusters separated from canonical bacterial and haloarchaeal sequences within the GlpA subgroup, justifying their annotation as GlpAs. Based on these analyses, no archaeal GlpDs could be identified. The Thermococcales sequences constitute a cluster clearly separated from *glpA/D* sequences and might thus represent a distinct family within the DAAO superfamily designated herein as GlpTk. This cluster also comprises sequences from anaerobic Crenarchaea (Desulfurococcales) as well as anaerobic Bacteria (e.g. *Thermotoga maritima* and *Clostridium perfringens*) and amitochondriate protists (*Giardia intestinalis*, *Entamoeba histolytica*, and *Spironucleus salmonicida*).

Secondly, the phylogenetic affiliation does not only reflect the structural differences of FAD-dependent G3PDHs, it is also accompanied by a conserved gene organization (Fig. 10): (i) The canonical *glpA* genes always cooccur in operons with downstream located *glpB* and *C* (Fig. 10a). (ii) In the *Thermofilum* related species three genes are located downstream of the *glpA* gene homologue encoding putative counterparts of the succinate dehydrogenase b, c and d subunits (also involved in membrane anchoring and electron transfer [76]) (Fig. 10b). (iii) All genes of the *S. acidocaldarius* related *glpA* subgroup show the *coxG* gene directly downstream of the *glpA* gene homologue (Fig. 10c). (iv) The *glpD* and *glpO* genes don't show any obvious operon like structures with any other genes which coincides with their function as homodimers irrespective of their membrane or cytoplasmic localization, respectively (Fig. 10d and e). (v) The Thermococcus related *glpTk* genes including the bacterial sequences from *T. maritima* and *C. perfringens* (see above) are all organized in a cluster with two genes downstream annotated to encode a pyridine nucleotide disulfide oxidoreductase/FAD-dependent oxidoreductase/ NADH oxidase (NOX) and a DUF1667 domain

protein/molybdopterin oxidoreductase Fe_4S_4 domain (MOX) shown to form a soluble complex upon *in vitro* reconstitution [37] (Fig. 10f). Interestingly, the GlpTk-like G3PDH enzymes from protists such as *Giardia intestinalis* [77], *Entamoeba histolytica* and *Spironucleus salmonicida* represents a fusion protein with a GlpTk, NOX, and a DUF1667 domain. Thus, the observed clustering in the phylogenetic tree is accompanied by the formation of different complexes with proteins encoded downstream of the respective G3PDH genes in operon like structures for membrane association and electron transfer to the respective acceptors. Therefore, the acquisition of glycerol utilization necessitated the concomitant evolution of tailored membrane anchoring and/or electron transfer systems to fit the requirements of the organism's lifestyle, metabolism and terminal electron acceptors.

In summary, we herein unraveled the glycerol degradation in *S. acidocaldarius* to proceed via the GK-G3PDH pathway involving a "classical" bacterial GlpK-like GK from the FGGY family. However, with respect to the G3PDH the pathway differs remarkably from those known from bacteria and haloarchaea. Although sequence comparison clearly identified the G3PDH as GlpA homologue, the enzyme shows an unusual subunit composition lacking the B and C subunits, a different domain structure in the C-terminus of the catalytic GlpA subunit, and a different mode of membrane anchoring via a CoxG homologue instead of a GlpC. Furthermore, although the physiological significance remains to be studied, sequence analyses suggest that glycerol metabolism in archaea is more versatile with respect to the G3PDHs their interaction partners, membrane association, and electron transfer mechanisms.

Table 1: Kinetic characterization of the recombinant GK and G3PDH from *S. acidocaldarius*.

Enzyme	Substrate	K_M [mM]	V_{max} [U mg ⁻¹]	k_{cat}/K_M [mM ⁻¹ s ⁻¹]	Calculate d mass [kDa]	Native mass [kDa]	Oligomeric state	Temperatur e optimum	pH optimum
Sacj_2033 ^a (GK)	Glycerol	0.024	339.7	13045					
	ATP	0.17	97,3	527	55.3	110	Homodimer	75°C ^a	6.5
Sacj_2032 ^b (G3PDH)	G3P	0.055	44.5	642					
	Ubiquinone -Q1	0.179	21.1	197	47.6	94	Homodimer	70°C	6.5

^a, kinetic constant for the Sacj_2033 GK were determined at 50°C.

^b, kinetic constants for the Sacj_2032 G3PDH were determined at 70°C.

n.d.: not detectable

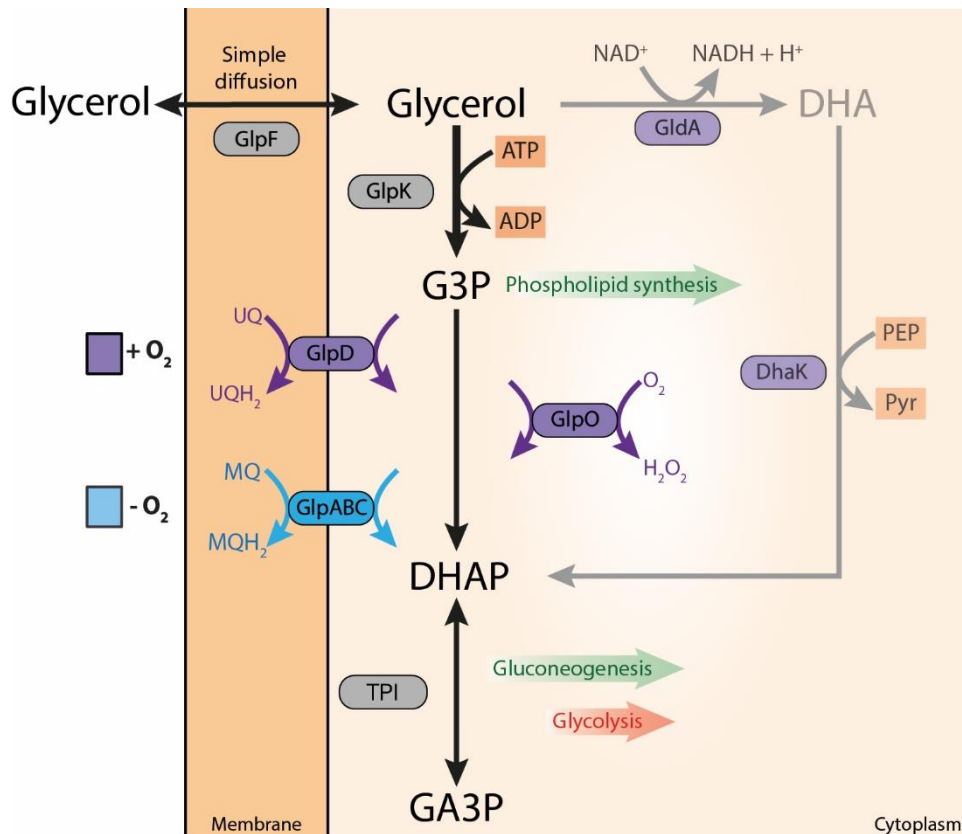


Fig. 1: Biochemical pathways involved in glycerol conversion in Bacteria (e.g., *E. coli*, *Streptococcus* or *Pseudomonas putida*) and Eukarya forming dihydroxyacetone phosphate (DHAP) which is channeled into central metabolism. Glycerol metabolism starts with its uptake either via facilitated diffusion mediated by the glycerol uptake facilitator (GlpF) (encoded by the *glpF* gene) or via (protein independent) simple diffusion through the cytoplasmic membrane. Following uptake, glycerol is converted finally to DHAP via two basic pathways: In respiring organisms, glycerol is first phosphorylated by the glycerol kinase (GlpK) (encoded by the *glpK* gene) to G3P which is further oxidized by two different membrane bound FAD dependent G3P dehydrogenases (G3PDH), i.e. GlpD (encoded by the *glpD* gene) and GlpABC (encoded by the *glpA*, *B*, and *C* genes). Electrons are transferred via the G3PDH bound FAD cofactor to ubiquinone (UQ) by GlpD or to menaquinone (MQ) (GlpABC) of the respiratory chain. A third mechanism of G3P oxidation mainly known from aerotolerant lactic acid bacteria as well as *Mycoplasma* species is catalyzed by a soluble, cytoplasmic FAD dependent G3P oxidase (GlpO, encoded by the *glpO* gene) which directly utilizes molecular oxygen as electron acceptor. The second basic glycerol converting pathway is known from fermentatively growing organisms. Here, glycerol is first oxidized via an NAD⁺ dependent G3PDH to dihydroxyacetone (GldA, encoded by the *gldA* gene) which is subsequently phosphorylated by dihydroxyacetone kinase (DhaK, encoded by *dhaK* gene) with phosphoenolpyruvate (less frequently ATP) as phosphoryl donor. In all cases, DHAP is further degraded via the lower common shunt of the ED and EMP pathway or utilized for gluconeogenesis. G3P serves as building block for phospholipid/membrane synthesis.

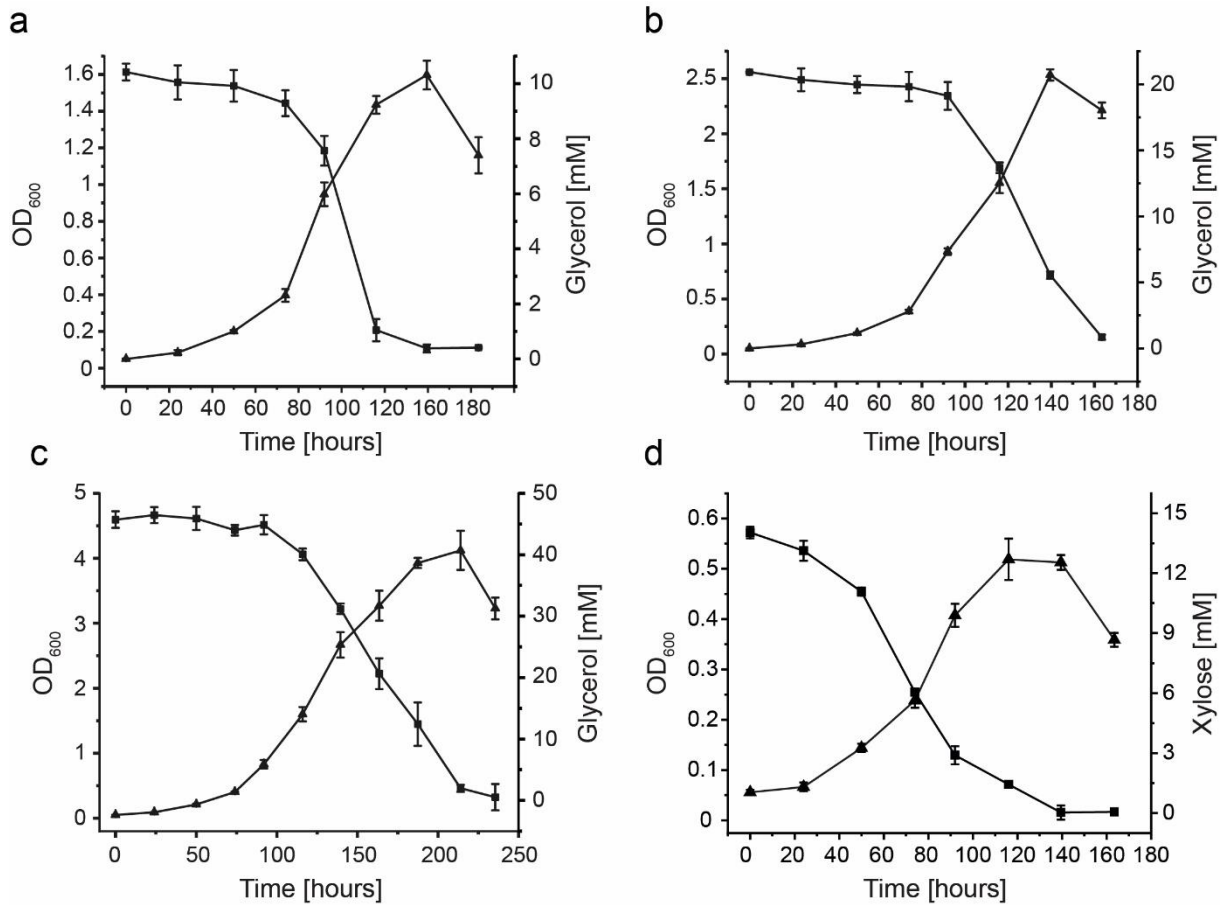


Figure 2: Growth and substrate consumption of *S. acidocaldarius* MW00G on Brock's basal medium containing 10 mM (a), 20 mM (b), and 40 mM (c) glycerol as sole carbon and energy source under aerobic conditions. For comparison the growth with 14 mM D-xylose (d) was studied. Growth was monitored as increase in OD₆₀₀. Substrate concentrations in the cell free supernatant were determined enzymatically as described in the materials and methods part. Experiments were performed in triplicate and error bars indicate the SD of the mean.

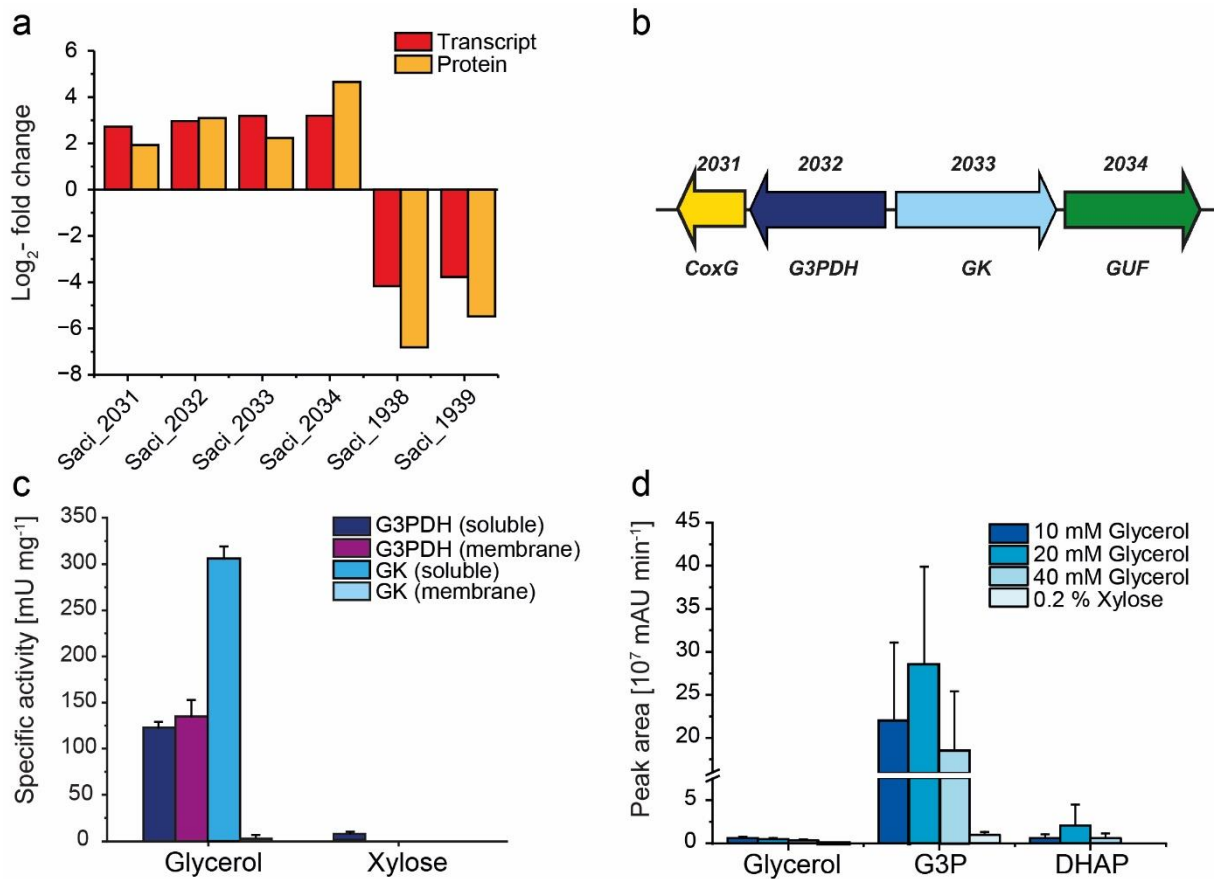


Figure 3: Poly-omics and activity-based analyses of glycerol degradation in *S. acidocaldarius*.

(a) Log₂ fold changes of transcript and protein levels of *saci_2031-saci_2034* and *saci_1938-saci-1939* of *S. acidocaldarius* MW00G grown on 40 mM of glycerol in comparison to 0.2 % (w/v) of D-xylose. (b) Divergent orientation of the potential glycerol degradation gene cluster *saci_2031-saci_2034* consisting of a putative glycerol uptake facilitator (GUF), glycerol kinase (GK, GlpK), glycerol-3-phosphate dehydrogenase (G3PDH, GlpA) and CoxG homolog. (c) GK and G3PDH activity of *S. acidocaldarius* MW00G grown on 40 mM glycerol in comparison to 0.2 % (w/v) D-xylose in the cytoplasmic and membrane fraction after cell lysis. (d) Peak areas of glycerol, G3P and DHAP of glycerol grown *S. acidocaldarius* MW00G (10mM, 20 mM and 40 mM glycerol) compared to D-xylose grown cells determined by targeted metabolomics. All values represent the average of three (growth curve and crude extract activities) or four (poly-omics analyses) independent measurements. Error bars represent the standard deviation of the mean.

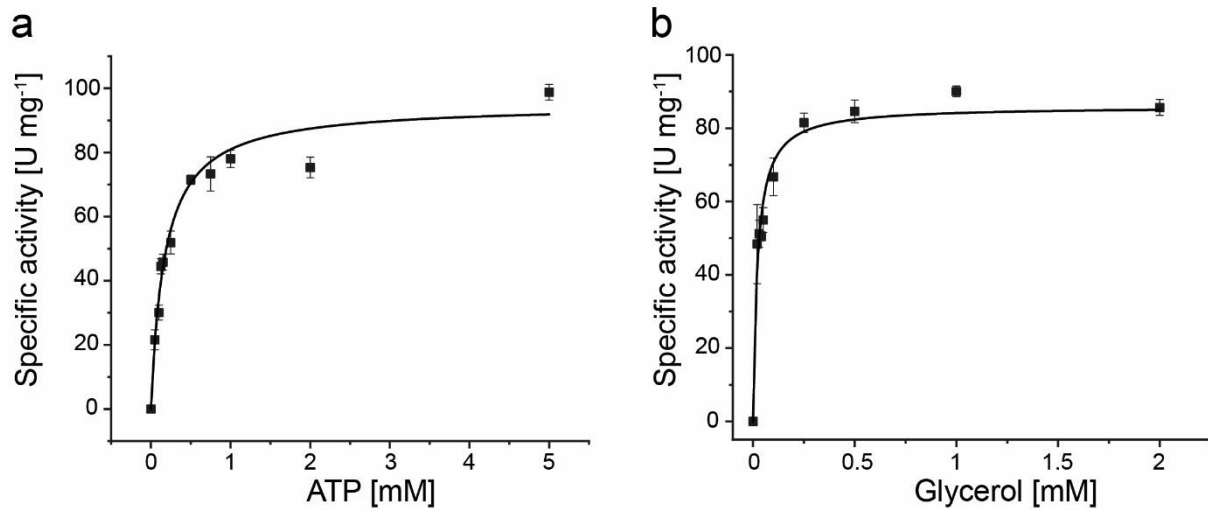


Figure 4: **Kinetic characterization of the purified recombinant GK Saci_2033.** The kinetic properties with ATP (a) and glycerol (b) were determined in a continuous assay by coupling the glycerol-dependent ADP formation from ATP to NADH oxidation via PK and LDH at 50°C as decrease in absorbance at 340 nm. Experiments were performed in triplicate and error bars indicate the SD of the mean.

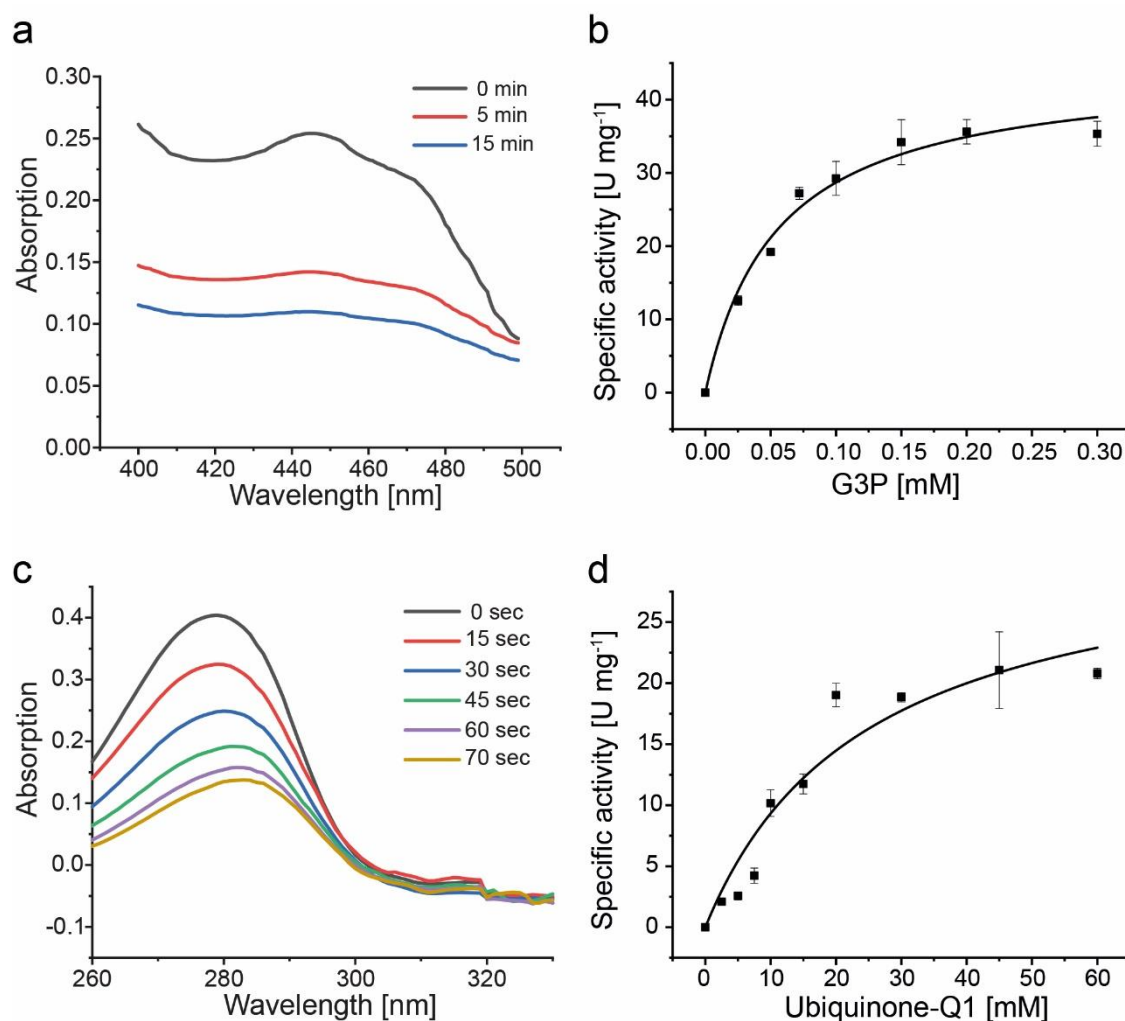


Figure 5: Biochemical and kinetic characterization of the recombinant G3PDH Saci_2032. (a) Time resolved absorption spectra of ubiquinone-Q1 in the presence of Saci_2032 and 200 μ M of G3P. Loss of absorption at 280 nm indicates reduction of ubiquinone-Q1 by Saci_2032. (b) Time resolved absorption spectrum (400-500 nm) of Saci_2032 bound FAD after addition of G3P. Loss of absorption at 450 nm indicates reduction of Saci_2032 bound FAD via G3P. Maximal absorption at 450 nm after denaturation by SDS was used to calculate the amount of bound FAD per protein. (c) The kinetic properties of recombinant Saci_2032 with G3P as substrate (0-0.3 mM) were determined in a continuous assay by coupling the oxidation of G3P to the reduction of the artificial redox active dye DCPIP as decrease of absorbance at 600 nm and 70°C. (d) Kinetic properties of Saci_2032 towards ubiquinone-Q1 were determined in a continuous assay following the G3P dependent reduction of ubiquinone-Q1 to ubiquinol-Q1 at 280 nm and 70°C. Experiments were performed in triplicate and error bars indicate the SD of the mean.

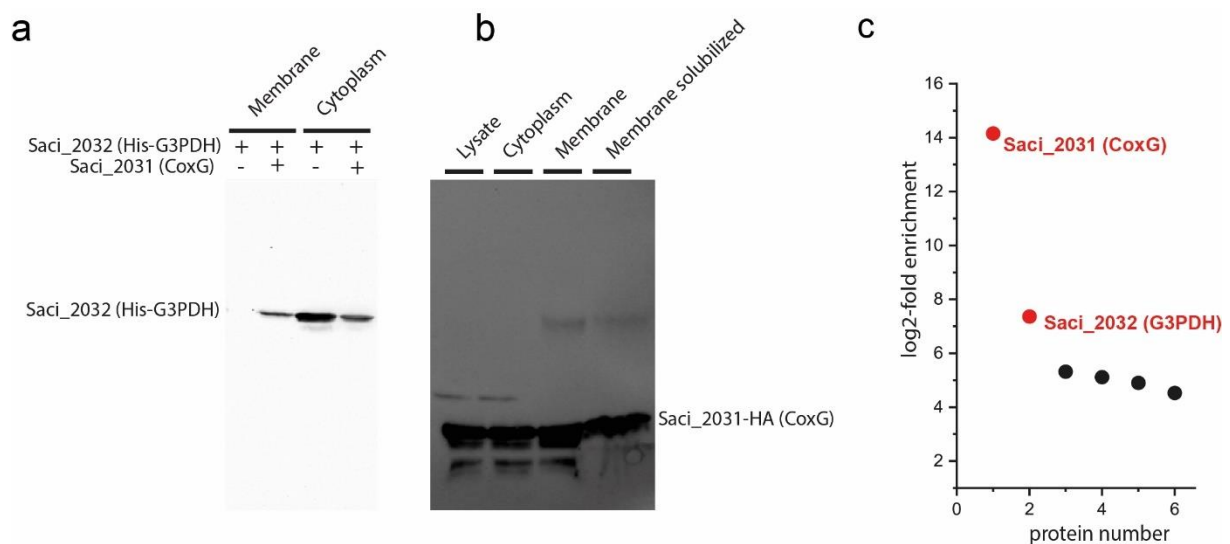


Figure 6. CoxG facilitates membrane binding Saci_2032 G3PDH *in vitro*. (a) Effect of coexpression of CoxG (Saci_2031) on the localization of His-tagged G3PDH (Saci_2032) in the cytoplasmic and membrane fraction in *E. coli* monitored via immunodetection with anti-His antibodies. Only upon coexpression with CoxG, G3PDH is found in the membrane fraction. (b) Localization of HA-tagged CoxG upon homologous overexpression using a pSVAaraFX-HA vector in *S. acidocaldarius* determined by immunodetection using anti-HA antibodies shows binding of the CoxG homolog to the cell membrane as well as localization in the cytoplasm. (c) Identification of interaction partners of HA-tagged CoxG (Saci_2031) expressed in *S. acidocaldarius* MW00G by anti-HA antibody co-immunoprecipitation and LC-MS-MS. The log₂-fold enrichment of interaction partners was determined using *S. acidocaldarius* MW00G crude extracts as a control.

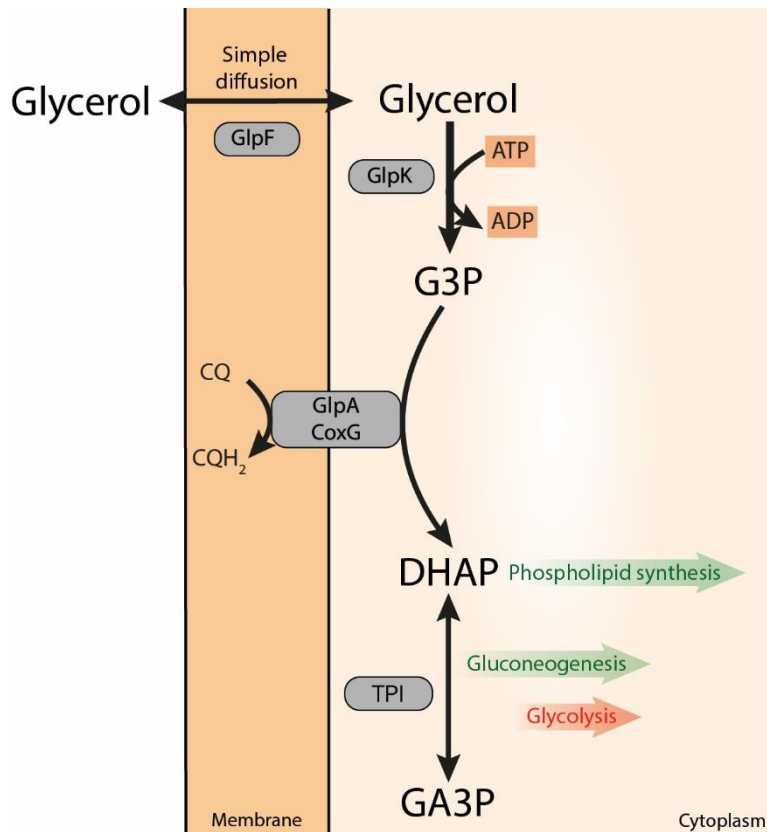


Fig. 7: Glycerol metabolism in *S. acidocaldarius*. Glycerol uptake likely involves the GlpF-like glycerol uptake facilitator Saci_2034 and eventually simple diffusion through the cytoplasmic membrane. Additional uptake systems like the ABC transporter (Saci_1762-1765) might also contribute. Intracellularly, glycerol is phosphorylated by GK (Saci_2033), homologous to the bacterial GlpK, forming G3P which is then oxidized by a membrane bound, unusual GlpA-like FAD dependent G3PDH (Saci_2032). Saci_2032 is anchored to the membrane by the CoxG homologue Saci_2031 forming a complex that transfers electrons to the caldariellaquinone (CQ) yielding dihydroxyacetone phosphate (DHAP). DHAP can be used for phospholipid synthesis (via G1P in Archaea), and - with concomitant triosephosphate isomerase (TPI, Saci_0117) mediated conversion to glyceraldehyde-3-phosphate (GA3P) - for gluconeogenesis (green arrows), or glycolysis (red arrow).

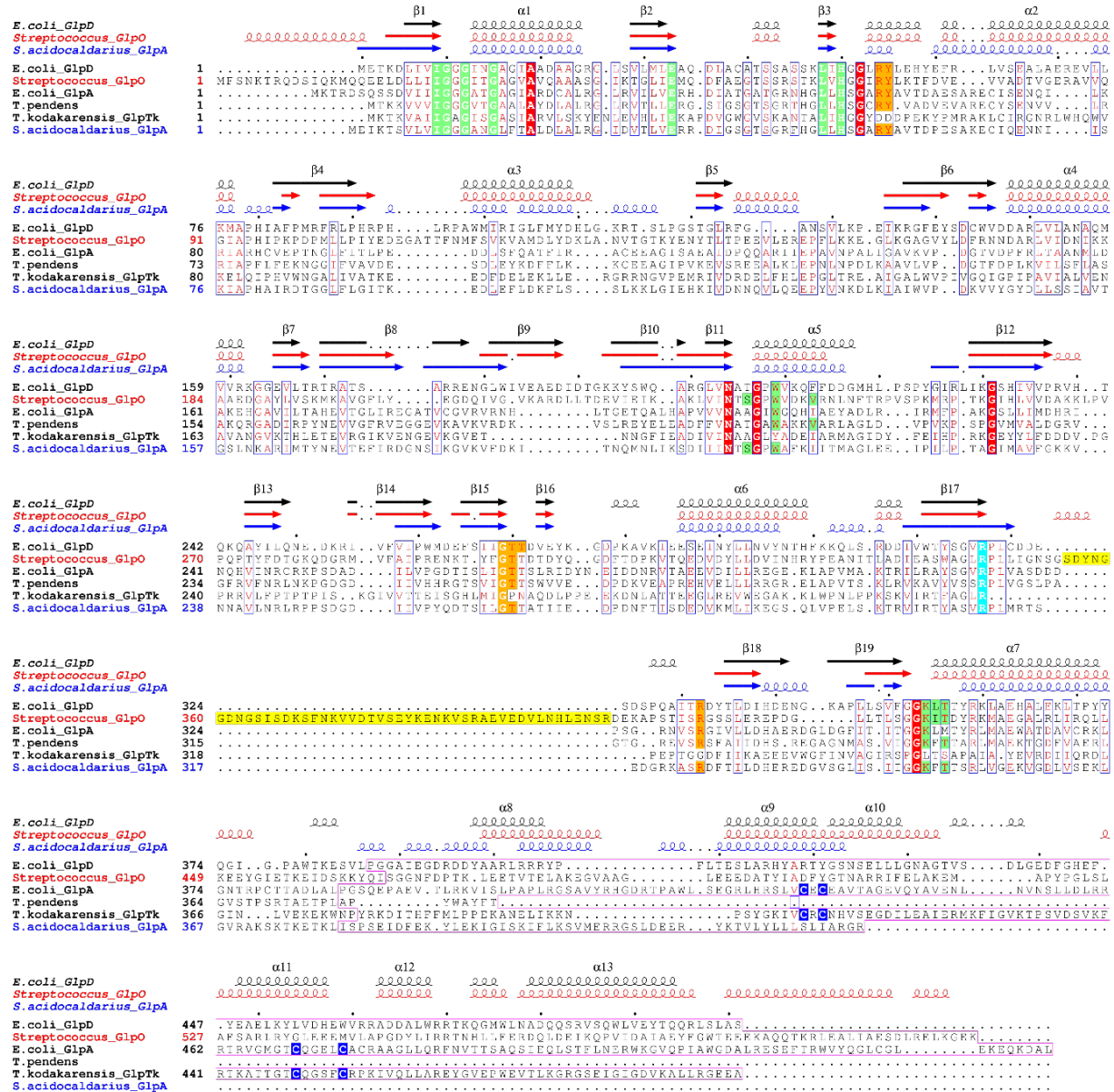


Figure 8: Sequence alignment of *S. acidocaldarius* G3PDH (Saci_2032) with *E. coli* GlpD (2qcu, P13035) and GlpA (P0A9C0), GlpO from *Streptococcus* sp. (2rgo, D0VWP7) as well as archaeal homologues from *T. pendens* (A1RZ95, Tpen_1127) and *T. kodakarensis* (Q5JGZ7, TK1393). The secondary structural elements from the available crystal structures of GlpD and GlpO as well as the predicted structure for *S. acidocaldarius* G3PDH by alphaFold based modelling are depicted above the sequences. The alignment illustrates the sequence and structural conservation in the N-terminal 370-390 aa residues. The residues shown to be involved in FAD cofactor binding and substrate (G3P/DHAP) binding in the GlpD and GlpO structures as well as their conservation in the other sequences are shaded green and orange, respectively. The arginine residue supposed to act as general base in catalysis is highlighted in cyan. The sequence insertion distinctive for GlpOs is shaded yellow. The amino acid residues comprising the C-terminal domains are boxed in pink with the cysteine residues forming the FeS clusters in GlpA and the *Thermococcus* G3PDH marked by blue shading. The figure was generated using ESPrnt 3.0 [78].

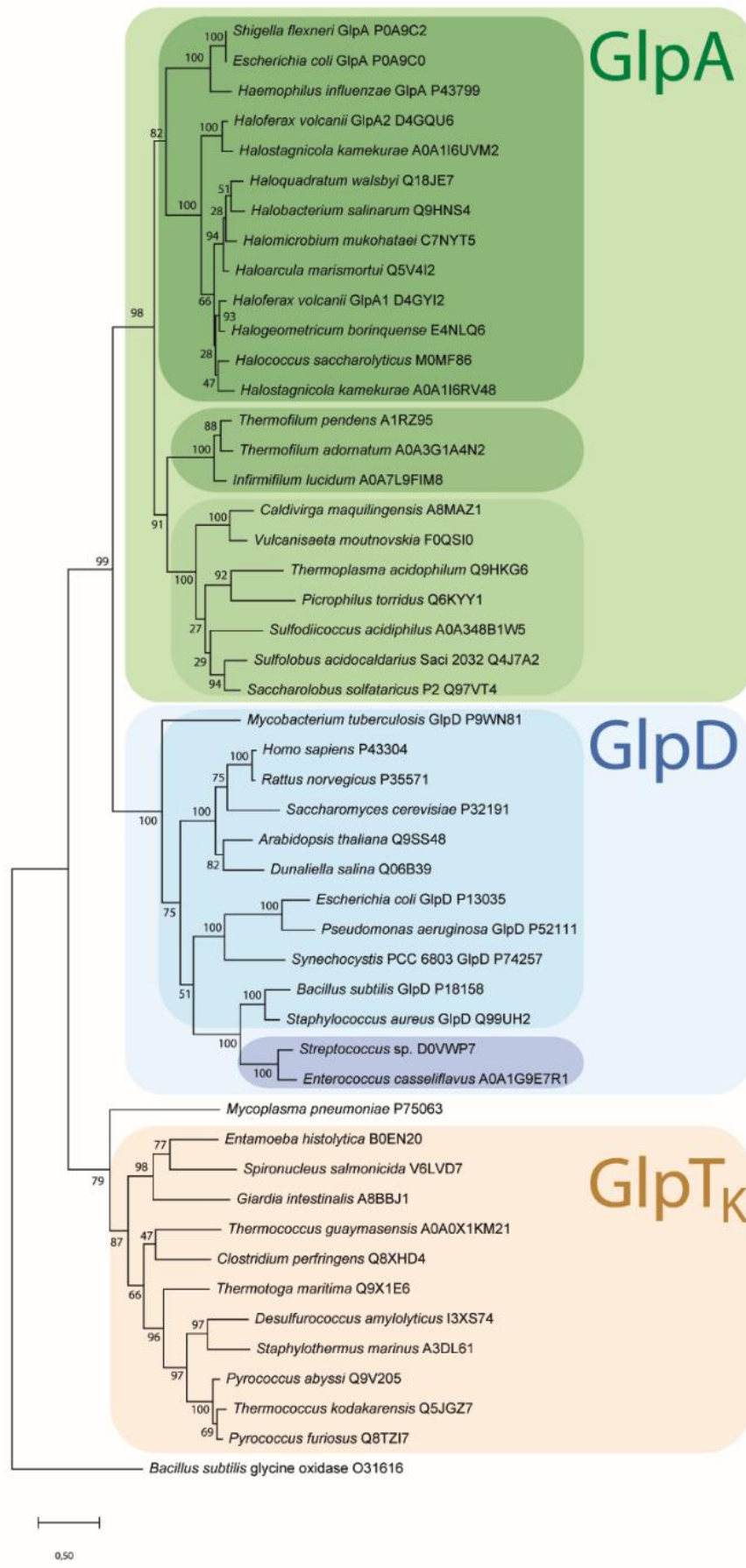


Fig. 9: Phylogenetic affiliation of the *S. acidocaldarius* G3PDH, selected GlpDs, GlpOs, and GlpAs from other Archaea, Bacteria and Eukarya (organism names and uniprot accession numbers are given). GlpA homologous are shaded green, with the canonical bacterial and haloarchaeal GlpAs shown in dark green, the *Thermophilum* homologues in green and the *S. acidocaldarius* homologues in light green. The GlpD cluster shaded blue comprises the GlpDs and GlpOs highlighted in light and dark blue, respectively. The Thermococcus-like G3PDHs, designated as GlpTk are shown in orange. The evolutionary history was inferred by using the Maximum Likelihood method and JTT matrix-based model [79]. The tree with the highest log likelihood (-20222.59) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. This analysis involved 49 amino acid sequences. All positions containing gaps and missing data were eliminated (complete deletion option) resulting in a total of 283 positions in the final dataset. Evolutionary analyses were conducted in MEGA X [80].

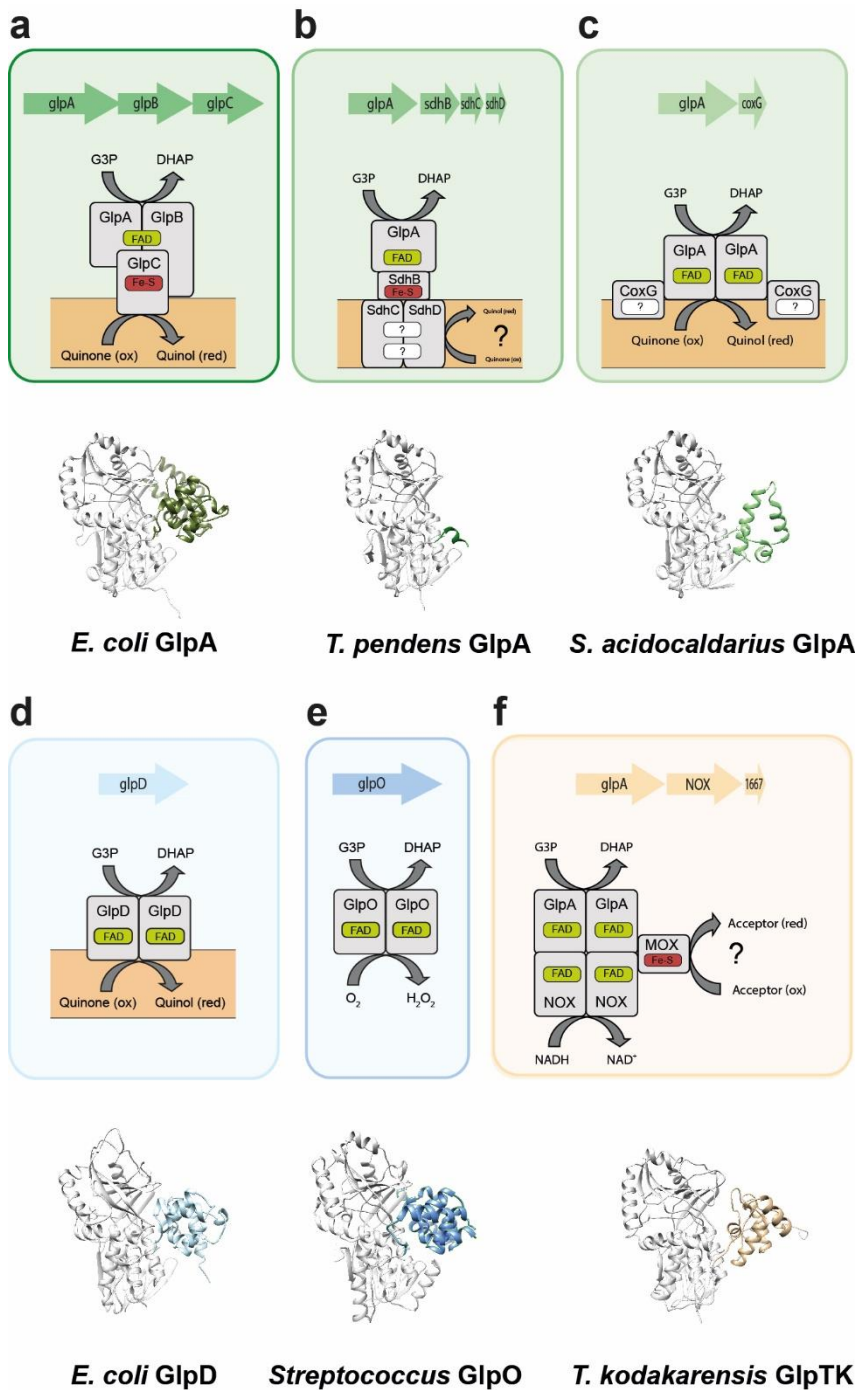


Fig. 10: Comparison of different G3PDHs from Bacteria and Archaea. The respective genome environment, the derived reconstructed protein complexes with potential interaction partners, membrane or cytoplasmic localization and cofactor content, as well as the catalytic oxidoreductase-monomer 3D-structure of the G3PDHs GlpABC (*E. coli* (a), GlpA-SdhBCD (*T. pendens*) (b), GlpA-CoxG (*S. acidocaldarius*) (c), GlpD (*E. coli*) (d) GlpO (*Streptococcus*) (e) and GlpTk-NOX-MOX (*T. kodakarensis*) (f) are shown. For colour code see Fig. 9. For all 3D structures the principal FAD-binding G3P-oxidoreductase domain is coloured in grey while the highly divergent C-terminus is coloured according to Fig. 9.

Material and Methods

Cultivation of *S. acidocaldarius* strains

Standard growth of *S. acidocaldarius* MW001 (uracil auxotroph mutant) [81] was performed at 75°C and pH 3 in Brock's basal salt medium, [82] supplemented with 0.1% (w/v) N-Z-Amine, 0.2% (w/v) dextrin, and 10 µg/ml of uracil. For growth on glycerol, *S. acidocaldarius* MW001 were adapted for several weeks in Brock's basal salt medium with 10 µg/ml uracil in the presence of 40 mM glycerol as sole carbon and energy source until growth was observed. For strain purification, cells were streaked on Brock's basal salt medium plates containing 0.1% (w/v) N-Z-Amine, 0.2% (w/v) dextrin, 10 µg/ml of uracil, and 0.7 % (w/v) Gelrite (for solidification, Carl Roth, Germany) and a single colony was regrown in Brock's basal salt medium containing 10 mM glycerol. This adapted strain denoted as *S. acidocaldarius* MW00G was stored as glycerol stock culture. To this end, *S. acidocaldarius* MW00G was grown on 10 mM glycerol to an OD₆₀₀ of 0.6, cells were pelleted at 4,300xg and 21°C for 10 min, and resuspended in standard Brock medium containing 50% (v/v) glycerol to a calculated OD₆₀₀ of 10 and stored at -80°C until used.

Growth experiments with the glycerol-adapted *S. acidocaldarius* MW00G were performed in triplicates at 76°C and 120 rpm under constant shaking using Innova®44 (New Brunswick, Germany) in the presence 10 mM glycerol, 20 mM glycerol or 40 mM glycerol as sole carbon and energy source. For comparison the same strain was cultivated with 0.2 % (w/v) D-xylose. Cells were inoculated into Brock's basal salt medium containing the respective carbon source to an OD₆₀₀ of 0.05 and growth was monitored as increase in OD₆₀₀ over time. The glycerol concentration in the medium was determined enzymatically using the recombinantly produced Saci_2033 glycerol kinase (see below) and ADP formation was quantitatively assayed spectrophotometrically as oxidation of NADH via pyruvate kinase (PK) (Merck, Germany) and lactate dehydrogenase (LDH) (Merck, Germany) at 340 nm (42°C) in a TECAN infinite M200 microplate reader (Tecan Trading AG, Switzerland). Thus, per mole of glycerol one mole of NADH was reduced. The reaction mix contained 0.1 M MOPS-KOH buffer pH 6.5, 0.6 mM NADH, 1 mM PEP, 5 mM ATP, 5 mM MgCl₂, 25.5 U LDH, 2.8 U PK, 0.4 U Saci_2033, and 20 µL glycerol containing sample corresponding to a maximal glycerol concentration of 0.5 mM in a total volume of 0.2 mL. For polyomic analysis, *S. acidocaldarius* MW00G was pre-grown in Brock's basal salt medium containing the respective glycerol concentration for two passages in 50 ml and 200 ml, respectively. With these precultures, 400 mL main cultures (four replicates) for polyomic analysis with the respective carbon sources were inoculated to a starting OD₆₀₀ of 0.05 and grown to exponential phase (OD₆₀₀ of 0.8). Cells were cooled down

with liquid nitrogen and harvested at 9,000xg, 15 min, 4°C. Cell pellets were stored at -70°C and used for RNA-seq, LC-MS-MS and LC(HILIC)-MS/MS for determination of metabolites.

Transcriptome

Transcriptomic analysis was performed similarly to what was described before [83].

Proteome

Proteomic analysis was performed similarly to what was described before [84].

Metabolome analyses

Metabolite extraction - The protocol for the metabolite extraction was based on the method described by *Sellick et al.* with slight modifications [85]. It consisted of a first disruption cell wall by resuspending the cell pellet in 500 µL of prechilled (-80°C) methanol and the addition of 20 µL of internal standards (fructose 6-phosphate, arginine 13C6, and succinic acid d4). The mixture was vortexed during 2 min and sonicated again for 2 min. After that, the sample was freeze at -80°C for 5 min and the vortex and sonication steps were repeated. The methanol was evaporated by vacuum concentration (Concentrator plus, Eppendorf, Hamburg, Germany). When the extract was completely dry, 250 µL of water were added and the process was repeated. During the whole process the cells were kept on ice. Then, after the drying the water, the extract was resuspended in 100 µL ACN/water (85:15, v/v), sonicated 2 min, and vortexed 2 min. The mixture was centrifugated at 12000 rpm during 2 min (MiniSpin centrifuge, Eppendorf, Hamburg, Germany). The supernatant was transferred to a LC vial.

LC-ESI-QTOF method for the metabolomics analysis - The samples were analyzed by a 1290 Infinity II LC instrument coupled to a 6546 LC/Q-TOF high-resolution mass spectrometer, and the ionization was performed using Dual Jet Stream source in negative mode (Agilent Technologies, Waldbronn, Germany). For the LC separation, a iHILIC-(P)Classic (150 x 2.1 mm, 5 µm) (Hilicon, Umeå, Sweden). The mobile phases consisted of 5 mM NH₄AC, pH 5 (A) and ACN/5 mM ammonium acetate, 85:15 (v/v), pH5 (B). The gradient elution was established as follows: 0 min, 90% B; 22 min, 80% B, 30 min, 65% B; 35 min, 65% B. The flow rate was set at 0.2 mL/min and the column temperature was kept at 40°C.

The electrospray ion source parameters were: gas temperature, 200°C; dry gas, 10 mL/min; nebulizer, 40 psi; sheath gas temperature, 300°C; sheath gas flow, 12 L/min; fragmentor, 125 V; skimmer, 65 V; capillary voltage, 3000 V. Full scan-data dependent acquisition was used to perform the tandem MS experiments.

Crude extract enzyme measurements

For preparation of crude extracts, *S. acidocaldarius* MW00G was grown in 50 ml cultures on 40 mM glycerol and - for comparison - 0.2 % (w/v) D-xylose to an OD₆₀₀ of 0.6. Cells were collected by centrifugation (4,300xg 15 min, 21°C), and resuspended in 5 ml of 50 mM MES-KOH pH 6.5, 20 mM KCl. Cells were lysed by sonication (3 x 8 min, 60 %, 0.5 s⁻¹; on ice, UP200S, Hielscher Ultrasonics) and centrifuged again at 4,300xg, 30 min, 4°C to remove cell debris. To separate the soluble from the membrane fraction, the supernatant was then centrifuged at 150,000xg, 60 min at 4°C. After centrifugation, the supernatant represented the soluble fraction and the pelleted membrane fraction was washed by resuspension in 5 ml of 50 mM MES-KOH, 20 mM KCl, pH 6.5 and centrifuged (150,000xg, 60 min at 4°C). The pellet was resuspended in 500 µl of 50 mM MES-KOH, 20 mM KCl, pH 6.5 and sonicated with a low amplitude (20%, 0.5 s⁻¹) until membranes were fully suspended. The so prepared membrane fraction was then used for enzyme measurements. Protein concentrations in both, soluble and membrane fractions were determined by a modified Bradford assay [86] using bovine serum albumin (Carl Roth, Germany) as standard. Enzyme activities in soluble and membrane fraction were assayed spectrophotometrically using a Specord UV/visible-light (Vis) spectrometer (Analytic Jena, Germany) with protein amounts of 50 µg (crude extract) and 40 µg (membrane fraction) in a total volume of 500 µl. Assay mixtures were preincubated at the respective temperatures before starting the reaction usually with substrate (glycerol, G3P). G3PDH activity was determined as glycerol-3-phosphate (G3P) (Merck, Germany) dependent reduction of DCPIP at 600 nm (extinction coefficient 20.7 mM⁻¹ cm⁻¹) and 70°C in 50 mM MES-KOH pH 6.5, 20 mM KCl containing 50 µM DCPIP and 200 µM G3P. GK activity was assayed in 50 mM MES-KOH pH 6.5, 20 mM KCl, 1 mM MgCl₂, 2.5 mM ATP, 5 mM glycerol, 0.2 mM NADH, and 2 mM PEP by coupling the glycerol and ATP dependent formation of ADP to NADH oxidation via LDH (63.7 U) and PK (7 U) at 50°C and 340 nm (extinction coefficient 6.22 mM⁻¹ cm⁻¹). It was ensured that the auxiliary enzymes were not rate limiting. One unit (1 U) of enzyme activity is defined as 1 µmole substrate consumed or product formed per minute.

Molecular cloning

E. coli DH5α used for plasmid construction and propagation was cultivated at 37°C and 180 rpm (INFORS shaker, Switzerland) in Lysogeny broth (LB) containing appropriate antibiotics (150 µg/ml ampicillin, 50 µg/ml kanamycin or 25 µg/ml chloramphenicol). For heterologous overexpression in *E. coli* either the pET15b (*saci_2032*) or the pETDuet-1 vector (coexpression of *saci_2032* and *saci_2031*, as well as expression of *saci_2032* alone) and for homologous overexpression in *S. acidocaldarius* MW001 the pBS-Ara-albaUTR-FX-vector (*saci_2033*) (manuscript in preparation) were used. Plasmids and strains are listed in

supplementary table 1. Open reading frames (ORFs) *saci_2031*, *saci_2032* and *saci_2033* were amplified from genomic DNA of *S. acidocaldarius* DSM 639 wild type using the primers (Eurofins Genomics) listed in supplementary table 2 employing the Q5 polymerase or PhusionHigh-Fidelity polymerase (NEB, USA) following the manufacturers' instructions. After restriction digest with the respective restriction endonucleases (NEB, USA) (supplementary table 2), the purified PCR products were ligated with the empty vectors in a molar ratio of 1:4 using T4 DNA ligase (NEB, USA) at 16 °C overnight. For construction of pETDuet-1 coexpression vectors *saci_2032* and *saci_2031* were sequentially ligated into multiple cloning site 1 and 2, respectively. For homologous overexpression of *saci_2031* in course of the co-immunoprecipitation experiments (see below), the PCR-amplified gene was cloned into pSVAaraFX-HA vector [87] using the *NcoI/XhoI* restriction sites. Plasmid constructs (pET15b-*saci_2032*; pETDuet-1-*saci_2032*; pETDuet-1-*saci_2031/saci_2032*, pBS-Ara-albaUTR-FX-*saci_2033*, and pSVA13204) were propagated in *E coli* DH5 α and successful cloning was confirmed by commercially available sequencing services (Eurofins genomics).

Homologous overexpression and affinity purification of glycerol kinase

For homologous expression, the pBS-Ara-albaUTR-FX-*saci_2033* vector was transformed in *S. acidocaldarius* MW001. To prevent degradation in the recipient strain, the plasmid was methylated by transformation into *E. coli* ER1821. After reisolation the methylated plasmid was transformed into electro-competent wild type *S. acidocaldarius* MW001 cells via electroporation using a Gene Pulser Xcell (BioRad, München, Germany) with a constant time protocol with input parameters 2 kV, 25 μ F, 600 Ω in 0.5 mm cuvettes. Cells were regenerated for 45 min at 75°C in liquid Brock's basal salt medium containing 0.1% (w/v) N-Z-Amine, 0.2% (w/v) dextrin [81], and then plated on the same medium solidified by Gelrite (see above) without uracil. Successful transformation was confirmed via colony-PCR. Clones harboring the desired plasmid construct with the respective gene insert were inoculated in a 50 ml preculture of liquid Brock's basal salt medium with 0.1% (w/v) N-Z-Amine, 0.2% (w/v) dextrin. The latter was used then to inoculate a 2 L liquid Brock's basal salt medium (OD₆₀₀ 0.05) containing 0.1 % N-Z-Amine and 0.2 % (w/v) D-xylose for induction and cells were grown for two days at 76°C under constant shaking at 120 rpm. Afterwards, cells were harvested by centrifugation at 7,000 x g and 4°C for 15 minutes and stored at -80°C for further use. For protein purification cells were resuspended in 50 mM TRIS-HCl pH 8.0, containing 250 mM NaCl in a ratio of 3 ml buffer to 1 g of cells (wet weight) and disrupted by sonication (3 x 8 min, 60% amplitude, 0.5 s⁻¹; UP200S, Hielscher Ultrasonics). Cell debris was removed by centrifugation at 30,000xg and 4°C for 45 min and the resulting crude extract was applied onto a Strep-Tactin®XT 4Flow® column (IBA Lifesciences, Germany) and purified following the manufacturers' instructions.

Protein purity was analyzed by SDS-PAGE and protein concentration was determined by a modified Bradford assay [86] using bovine serum albumin (Carl Roth, Germany) as standard.

Heterologous overexpression and affinity purification of Saci_2032

For expression the pET15b-vector-*saci_2032* was transformed into *E. coli* Rosetta (DE3) [pLys] and grown in terrific broth (TB) (Carl Roth, Germany) supplemented with 150 µg/ml ampicillin and 25 µg/ml chloramphenicol. After growth at 37°C to an OD₆₀₀ of 0.8 cells were induced with 1 mM IPTG and further grown at 20°C for 17 h. Cells were harvested by centrifugation (6,000 x g, 20 min, and 4°C) and stored immediately at -80°C for further use. Chromatographic purification was carried out using an ÄktaPurifier system (Cytiva, USA). After thawing, cells were resuspended in 10 ml buffer per g cells (wet weight) containing 50 mM TRIS-HCl pH 7.8, 10 mM imidazole, 150 mM NaCl, supplemented with 1 mM FAD and disrupted by sonication (50% amplitude, 0.5 s⁻¹; UP200S, Hielscher Ultrasonics). Afterwards cell debris was removed by centrifugation (16,000xg for 45 min, 4°C), the supernatant was filtered (0.45 polyvinylidene fluoride membrane, Carl Roth, Germany) and applied onto a 1 ml Nickel-IDA column (GE Healthcare) equilibrated in 50 mM TRIS-HCl pH 7.8, 10 mM imidazole, 150 mM NaCl and proteins were eluted with a linear gradient from 10- 300 mM imidazole in the same buffer. G3PDH containing fractions as judged by SDS-PAGE and activity measurements were pooled and imidazole was removed using a centrifugal concentrator (10 kDa cutoff, Sartorius, Germany). Protein aliquots were frozen in liquid nitrogen and stored at -80°C for further use. Protein purity was analyzed by SDS-PAGE and protein concentration was determined by a modified Bradford assay [86] using bovine serum albumin (Carl Roth, Germany) as standard. To determine the FAD concentration of the purified G3PDH, the protein was denatured with 0.5 % (w/v) SDS and FAD absorbance spectra were determined from 300 nm to 600 nm. The FAD content was calculated using an extinction coefficient of 11.300 mM⁻¹ cm⁻¹ at 450 nm.

Determination of the native molecular mass of purified proteins

Size exclusion chromatography was used to determine the native molecular mass of all purified enzymes. To this end, pooled enzyme samples after affinity chromatography were concentrated using centrifugal concentrators (10 kDa cutoff) (Sartorius, Germany) and applied to size exclusion chromatography (Superdex 200 HiLoad 26/60 prep grade column or Superose 6/100 column (Cytavia, USA)) using 50 mM TRIS-HCl pH 7.5, containing 250 mM NaCl as buffer. For calibration, Gel Filtration LMW Calibration Kit and Gel Filtration HMW Calibration Kit (Cytavia, USA) were used.

Localization of G3PDH

To investigate the role of CoxG (*Saci_2031*) for membrane anchoring of the *Saci_2032* G3PDH in *S. acidocaldarius*, *saci_2032* was heterologously coexpressed with *saci_2031* from the respective pETDuet-1 expression vector construct in *E. coli* Rosetta (DE3) as described above. As a control, *saci_2032* was expressed alone from the same vector without *saci_2031*. After expression, membrane and soluble fractions were separated. To this end cells were disrupted as described above and insoluble cell debris and non-lysed cells were removed by centrifugation at 6,000xg, 4°C for 15 minutes. Afterwards, the resulting supernatant was centrifuged at 150,000xg at 4°C for 120 min, and the resulting supernatant (=soluble fraction) was separated from the pelleted cell membranes. The insoluble membrane fraction was washed twice with 50 mM TRIS-HCl, pH 7.0 and centrifuged (150,000xg, 4°C, 120 min). Finally, the membrane fraction was aliquoted, flash frozen in liquid nitrogen, and stored at -80°C for further use. For activity measurements and immunodetection, isolated membranes were thawed, mixed with 500 µl of buffer (50 mM TRIS-HCl pH 7.0) and sonicated with a low amplitude (20%, 0.5 s⁻¹) until membranes were fully resuspended. For immunodetection, 50 µg of membrane or soluble protein preparations were separated by SDS-PAGE (12.5% (w/v) polyacrylamide gels) and transferred to a PVDF membrane using the Trans-Blot Turbo (Biorad, USA) at 25 V and 1.0 A for 30 min in 25 mM TRIS, 192 mM glycine, 20% (v/v) Ethanol, 0.075% SDS. Membranes were blocked with 5 % (w/v) BSA in TBST (20 mM TRIS, 500 mM NaCl, 0.05% Tween20) and then incubated with horseradish peroxidase (HRP) conjugated anti-His antibody (1:10,000 dilution; Abcam, United Kingdom) in TBST with 5 % (w/v) BSA for 90 minutes. Afterwards membranes were washed three times with 15 ml TBST for 5 min at 21°C, and finally with 15 ml of TBS (20 mM TRIS-HCl pH 7.4, 500 mM NaCl) for 5 min. Detection of His-tagged proteins was performed with the Clarity™ Western ECL substrate (Biorad, USA) and using a Versadoc (Biorad, USA) for luminescence detection.

Co-immunoprecipitation (co-IP) Assays with anti-HA-Magnetic Beads

For each pulldown experiment, *saci_2031* was homologously expressed from the pSVA13204 vector construct transformed into *S. acidocaldarius* MW00G to yield a C-terminally HA-tagged recombinant protein. 500 ml cell cultures were grown Brock's basal salt medium with 10 mM glycerol as carbon source for 72 h to an OD₆₀₀ 0.6-0.7 and expression was induced by the addition of 0.2% (w/v) L-arabinose. After 4 h of further growth, cells were cooled down on ice and harvested via centrifugation for 20 min at 4 °C and 5,000xg at an OD₆₀₀ of 0.8-0.9. The pellet was resuspended to a calculated OD₆₀₀ of 40 in 25 mM TRIS-HCl pH 7, 150 mM KCl, 5% (v/v) glycerol, 10 mM EDTA, and 2 % (w/v) DDM and incubated for 1 h at 37 °C while

rotating. The lysed cells were then centrifuged for 30 min at 13000 rpm to remove cell debris at room temperature and the supernatant was used for Co-immunoprecipitation. Therefore, 35 μ l of Pierce™ anti-HA Magnetic Beads (Pierce HA-tag IP/co-IP kit (Pierce)) were added to 10 ml of the cell lysate and incubated for 1 h at room temperature while rotating. The beads were concentrated using a magnetic stand and washed twice with 25 mM TRIS-HCl pH 7 containing 150 mM KCl, 5% (v/v) glycerol, 10 mM EDTA and one time with LC-MS grade water (Merck, Germany). After washing, 35 μ l of LC-MS grade water was added and 10 μ l of the bead solution was taken for SDS-PAGE analysis and Western blotting including immunodetection. After centrifugation, the supernatant was removed from the remaining beads. The beads were frozen down at -20 °C and sent to mass spectrometry analysis. HA-tagged proteins were detected with anti-HA antibodies and chemifluorescence was generated using horse radish coupled secondary antibodies. The Westernblots were developed using an iBright™ FL1500 (ThermoFisher).

Characterization of purified enzymes

Enzyme assays were performed spectrophotometrically using a Specord UV/visible-light (Vis) spectrometer (Analytic Jena, Germany) in cuvettes filled with 500 μ l assay mixture (unless state otherwise). Assay mixtures were prewarmed to the respective assay temperature and the reaction was usually started by addition of substrate. Experimental data were fitted, and kinetic constants were determined using the OriginPro 2022 software package.

Glycerol kinase – GK activity was determined as described above for the crude extract measurements with following modifications: The assay was performed in 0.1 M MOPS-KOH buffer (pH 6.5 at 50°C), 5 mM MgCl₂, 5 mM ATP, 0.2 mM NADH, 2 mM PEP, 63.7 U LDH, 7 U PK and 0.1-0.5 μ g of purified Saci_2033. For K_M determination varying concentrations of glycerol (0-2 mM), DHA (0-30 mM), DL-glyceraldehyde (0-20 mM) or ATP (0-5 mM), were used at a constant concentration of 5 mM ATP or 2 mM glycerol, respectively. The pH optimum was determined using the same assay in a mixed buffer system ranging from pH 5.0 to pH 8.0 containing 50 mM MES, 50 mM HEPES and 50 mM TRIS. The temperature optimum was determined between 60°C to 80°C by coupling the GK mediated G3P formation to DCPIP reduction followed at 600 nm via purified, recombinant G3PDH (0.25 U) from *S. acidocaldarius* (Saci_2032) in an assay containing 0.1 M MOPS-KOH buffer (pH 6.5, temperature adjusted), 5 mM ATP, 5 mM MgCl₂, 0.1 mM DCPIP and 2 mM glycerol. To study the thermal stability, the Saci_2033 GK was incubated at 70°C, 80°C, and 90°C in 100 mM MOPS-KOH, pH 6.5 (temperature adjusted) at a protein concentration of 0.05 mg ml⁻¹. After regular time intervals samples were taken and the residual activity of the enzyme was determined in the coupled

assay via PK and LDH at 50°C as described above. The effect of fructose-1,6-bisphosphate (F1,6BP) on GK activity was also tested in the PK-LDH coupled assay (see above) in the presence of F1,6BP concentrations up to 1 mM as indicated. Nucleotide specificity was analyzed in the G3PDH coupled assay system using 5 mM GTP, CTP or phosphoenolpyruvate (PEP) instead of ATP at 75°C.

Glycerol-3-phosphate dehydrogenase - G3PDH activity was assayed spectrophotometrically at 70°C by following the G3P dependent reduction of either DCPIP, FAD or ubiquinone-Q1 as electron acceptors. Assays were performed in 50 mM HEPES-KOH pH 6.5, 100 mM KCl. For determination of the reductive half reaction, 70 µg of Saci_2032 protein corresponding to a final concentration of 5 µM was incubated with 50 µM of G3P. 0 min, 5 min and 15 min after addition of G3P, absorption spectra were measured between 400 and 500 nm in a 96 well plates (BRANDplates®, BRAND, Germany) in a Tecan infinite M200 plate reader (Tecan Trading AG, Switzerland). Loss of absorption at 450 nm indicates the reduction of bound FAD. To test the quinone reactivity, 0.6 µg of Saci_2032 was preincubated with 30 µM of ubiquinone-Q1. Afterwards 100 µM of G3P was added to the reaction mix and reduction of ubiquinone-Q1 was followed in 15 second intervals by recording the absorption spectra between 260 and 330 nm in the Tecan infinite M200 plate reader. Loss of absorption at 280 nm indicates the reduction of the ubiquinone-Q1. Continuous enzyme assays were performed by following the change (decrease) in absorbance upon reduction of the artificial electron acceptor DCPIP (600 nm) or of the ubiquinone-Q1 (280 nm). Kinetic parameters for G3P were determined in a continuous enzyme assay with 0.06 mM of DCPIP as electron acceptor and varying concentrations of G3P (0-0.3 mM). For determination of the kinetic constants for ubiquinone-Q1 concentrations were varied between 0 mM and 60 mM of the quinone at a fixed G3P concentration of 0.4 mM. Both, the pH and the temperature optimum were determined using 0.4 mM G3P and 0.06 mM DCPIP. For the pH optimum a mixed buffer system containing 50 mM MES, HEPES, and TRIS was used in the range of pH 5.0 to pH 8.0. The same assay system in 50 mM MES-KOH pH 6.5 adjusted at the respective temperature was applied for determination of the temperature optimum between 50°C and 80°C. To test the recombinant Saci_2032 for glycerol oxidase activity, the G3P dependent hydrogen peroxide formation was coupled to the oxidation of 2,2'-azinobis-(3-ethylbenzothiazoline-6-sulfonate) (ABTS) via horseradish peroxidase (HRP) according e.g. to [25]. The assay mixture contained 100 mM MES-KOH buffer at pH 6.5, 1 mM ABTS, 0.45 µg of protein and 0.2 U of HRP. The reaction was started by addition of 0.2 mM G3P and the hydrogen peroxide dependent ABTS oxidation was photometrically determined at 420 nm (extinction coefficient 42.3 mM⁻¹cm⁻¹). The thermostability of the Saci_2032 G3PDH was analyzed by incubating the enzyme at 70°C, 80°C, and 90°C in 50 mM MES, pH 6.5 (temperature adjusted) at a protein concentration of

0.18 mg ml⁻¹. After regular time intervals the residual activity of the enzyme was determined with G3P and DCPIP as described above. The influence of detergents and membrane lipids on Saci_2032 activity was also analysed with DCPIP at 50°C in the presence of either 0.5 % (w/v) DDM, 0.5 % (v/v) of Triton X-100, 50 µg of Phosphatidylcholine (Merck, Germany) or 50 µg of isolated *S. acidocaldarius* MW00G membrane fractions.

References

1. Stroud, R. M., Miercke, L. J., O'Connell, J., Khademi, S., Lee, J. K., Remis, J., Harries, W., Robles, Y. & Akhavan, D. (2003) Glycerol facilitator GlpF and the associated aquaporin family of channels, *Current Opinion in Structural Biology*. **13**, 424-431.
2. Richey, D. P. & Lin, E. C. (1972) Importance of facilitated diffusion for effective utilization of glycerol by *Escherichia coli*, *Journal of Bacteriology*. **112**, 784-790.
3. Blötz, C. & Stülke, J. (2017) Glycerol metabolism and its implication in virulence in *Mycoplasma*, *FEMS Microbiology Reviews*. **41**, 640-652.
4. Mahdizadeh, S., Masukagami, Y., Tseng, C.-W., Markham, P. F., Souza, D. P. D., Nijagal, B., Tull, D., Tatarczuch, L., Browning, G. F., Sansom, F. M. & Dozois, C. M. (2021) A *Mycoplasma gallisepticum* glycerol ABC transporter involved in pathogenicity, *Applied and Environmental Microbiology*. **87**, 03112-20.
5. Wille, U., Schade, B. & Duszenko, M. (1998) Characterization of glycerol uptake in bloodstream and procyclic forms of *Trypanosoma brucei*, *European Journal of Biochemistry*. **256**, 245-250.
6. Klein, M., Swinnen, S., Thevelein, J. M. & Nevoigt, E. (2017) Glycerol metabolism and transport in yeast and fungi: established knowledge and ambiguities, *Environmental Microbiology*. **19**, 878-893.
7. Lin, E. C. C. (1976) Glycerol dissimilation and its regulation in bacteria, *Annual Review of Microbiology* **30**, 535-578.
8. Poblete-Castro, I., Wittmann, C. & Nickel, P. I. (2020) Biochemistry, genetics and biotechnology of glycerol utilization in *Pseudomonas* species, *Microbial biotechnology*. **13**, 32-53.
9. Yeh, J. I. C., Veronique; Paulo, Joao; Hou, Lihui; Darbon, E., Claiborne, A., Hol, W. G. J. & Deutscher, J. (2004) Structures of enterococcal glycerol kinase in the absence and presence of glycerol: Correlation of conformation to substrate binding and a mechanism of activation by phosphorylation, *Biochemistry*. **43**, 362-373.
10. Voegelé, R. T. S., Gaye D; Boos, Winfried; (1993) Glycerol kinase of *Escherichia coli* is activated by interaction with the glycerol facilitator, *Journal of Bacteriology* **175**, 1087-1094.
11. Lemieux, M. J., Huang, Y. & Wang, D. N. (2004) Glycerol-3-phosphate transporter of *Escherichia coli*: structure, function and regulation, *Research in Microbiology*. **155**, 623-9.
12. Weissenborn, D. L., Wittekindt, N. & Larson, T. J. (1992) Structure and regulation of the glpFK operon encoding glycerol diffusion facilitator and glycerol kinase of *Escherichia coli* K-12, *Journal of Biological Chemistry*. **267**, 6122-6131.
13. Applebee, M. K., Joyce, A. R., Conrad, T. M., Pettigrew, D. W. & Palsson, B. (2011) Functional and metabolic effects of adaptive glycerol kinase (GLPK) mutants in *Escherichia coli*, *Journal of Biological Chemistry*. **286**, 23150-23159.
14. Schryvers, A. & Weiners, J. H. (1981) The anaerobic sn-glycerol-3-phosphate dehydrogenase of *Escherichia coli*. Purification and characterization, *The Journal of Biological Chemistry*. **256**, 9959-9965.
15. Schryvers, A., Lohmeier, E. & Weiner, J. H. (1978) Chemical and functional properties of the native and reconstituted forms of the membrane-bound, aerobic glycerol-3-phosphate dehydrogenase of *Escherichia coli*, *Journal of Biological Chemistry*. **253**, 783-788.
16. Yeh, J. I., Chinte, U. & Du, S. (2008) Structure of glycerol-3-phosphate dehydrogenase, an essential monotopic membrane enzyme involved in respiration and metabolism, *Proceedings of the National Academy of Sciences of the United States of America*. **105**, 3280-3285.
17. Austin, D. & Larson, T. J. (1991) Nucleotide sequence of the *glpD* gene encoding aerobic sn-glycerol 3-phosphate dehydrogenase of *Escherichia coli* K-12, *Journal of Bacteriology*. **173**, 101-107.
18. Cole, S. T., Eiglmeier, K., Ahmed, S., Honore, N., Elmes, L., Anderson, W. F. & Weiner, J. H. (1988) Nucleotide sequence and gene-polypeptide relationships of the *glpABC* operon encoding the anaerobic sn-glycerol-3-phosphate dehydrogenase of *Escherichia coli* K-12, *Journal of Bacteriology*. **170**, 2448-2456.
19. Varga, M. E. R. & Weiner, J. H. (1995) Physiological role of GlpB of anaerobic glycerol-3-phosphate dehydrogenase of *Escherichia coli*, *Biochemistry and Cell Biology*. **73**, 147-153.

20. Colussi, T., Parsonage, D., Boles, W., Matsuoka, T., Mallett, T. C., Karplus, P. A. & Claiborne, A. (2008) Structure of α -glycerophosphate oxidase from *Streptococcus sp.*: A template for the mitochondrial α -glycerophosphate dehydrogenase, *Biochemistry*. **47**, 965-977.
21. Zotta, T., Parente, E. & Ricciardi, A. (2017) Aerobic metabolism in the genus *Lactobacillus*: impact on stress response and potential applications in the food industry, *Journal of Applied Microbiology*. **122**, 857-869.
22. Finnerty, C. M., Charrier, V., Claiborne, A. & Karplus, P. A. (2002) Crystallization and preliminary crystallographic analysis of the soluble α -glycerophosphate oxidase from *Streptococcus sp.*, *Acta Crystallographica Section D: Biological Crystallography*. **58**, 165-166.
23. Parsonage, D., Luba, J., Mallett, T. C. & Claiborne, A. (1998) The Soluble α -Glycerophosphate oxidase from *Enterococcus casseliflavus*: sequence homology with the membrane-associated dehydrogenase and kinetic analysis of the recombinant enzyme, *Journal of Biological Chemistry*. **273**, 23812-23822.
24. Elkhal, C. K., Kean, K. M., Parsonage, D., Maenpuen, S., Chaiyen, P., Claiborne, A. & Karplus, P. A. (2015) Structure and proposed mechanism of L- α -glycerophosphate oxidase from *Mycoplasma pneumoniae*, *The FEBS Journal*. **282**, 3030-3042.
25. Maenpuen, S., Watthaisong, P., Supon, P., Sucharitakul, J., Parsonage, D., Karplus, P. A., Claiborne, A. & Chaiyen, P. (2015) Kinetic mechanism of L- α -glycerophosphate oxidase from *Mycoplasma pneumoniae*, *The FEBS Journal*. **282**, 3043-3059.
26. Clomburg, J. M. & Gonzalez, R. (2013) Anaerobic fermentation of glycerol: a platform for renewable fuels and chemicals, *Trends in Biotechnology*. **31**, 20-28.
27. Clomburg, J. M., Cintolesi, A. & Gonzalez, R. (2022) *In silico* and *in vivo* analyses reveal key metabolic pathways enabling the fermentative utilization of glycerol in *Escherichia coli*, *Microbial Biotechnology*. **15**, 289-304.
28. Rawls, K. S., Yacovone, S. K. & Maupin-Furlow, J. A. (2010) GlpR represses fructose and glucose metabolic enzymes at the level of transcription in the haloarchaeon *Haloferax volcanii*, *Journal of Bacteriology*. **192**, 6251-60.
29. Sherwood, K. E., Cano, D. J. & Maupin-Furlow, J. A. (2009) Glycerol-mediated repression of glucose metabolism and glycerol kinase as the sole route of glycerol catabolism in the haloarchaeon *Haloferax volcanii*, *Journal of Bacteriology*. **191**, 4307-15.
30. Rawls, K. S., Martin, J. H. & Maupin-Furlow, J. A. (2011) Activity and transcriptional regulation of bacterial protein-like glycerol-3-phosphate dehydrogenase of the haloarchaea in *Haloferax volcanii*, *Journal of Bacteriology*. **193**, 4469-4476.
31. Williams, T. J., Allen, M., Tschitschko, B. & Cavicchioli, R. (2017) Glycerol metabolism of haloarchaea, *Environmental Microbiology*. **19**, 864-877.
32. Nishihara, M., Yamazaki, T., Oshima, T. & Koga, Y. (1999) sn-glycerol-1-phosphate-forming activities in Archaea: Separation of archaeal phospholipid biosynthesis and glycerol catabolism by glycerophosphate enantiomers, *Journal of Bacteriology*. **181**, 1330-1333.
33. Villanueva, L., Schouten, S. & Damsté, J. S. S. (2017) Phylogenomic analysis of lipid biosynthetic genes of Archaea shed light on the 'lipid divide', *Environmental Microbiology*. **19**, 54-69.
34. Gagen, E. J., Yoshinaga, M. Y., Garcia Prado, F., Hinrichs, K.-U. & Thomm, M. (2016) The proteome and lipidome of *Thermococcus kodakarensis* across the stationary phase, *Archaea*. **2016**, 5938289.
35. Hokao, R., Matsumura, H., Katsumi, R., Angkawidjaja, C., Takano, K., Kanaya, S. & Koga, Y. (2020) Affinity shift of ATP upon glycerol binding to a glycerol kinase from the hyperthermophilic archaeon *Thermococcus kodakarensis* KOD1, *Journal of Bioscience and Bioengineering*. **129**, 657-663.
36. Koga, Y., Katsumi, R., You, D.-J., Matsumura, H., Takano, K. & Kanaya, S. (2008) Crystal structure of highly thermostable glycerol kinase from a hyperthermophilic archaeon in a dimeric form, *The FEBS Journal*. **275**, 2632-2643.
37. Koga, Y., Konishi, K., Kobayashi, A., Kanaya, S. & Takano, K. (2019) Anaerobic glycerol-3-phosphate dehydrogenase complex from hyperthermophilic archaeon *Thermococcus kodakarensis* KOD1, *Journal of Bioscience and Bioengineering*. **127**, 679-685.

38. Koga, Y., Morikawa, M., Haruki, M., Nakamura, H., Imanaka, T. & Kanaya, S. (1998) Thermostable glycerol kinase from a hyperthermophilic archaeon: gene cloning and characterization of the recombinant enzyme, *Protein Engineering, Design and Selection*. **11**, 1219-1227.
39. Bräsen, C., Esser, D., Rauch, B. & Siebers, B. (2014) Carbohydrate metabolism in Archaea: Current insights into unusual enzymes and pathways and their regulation, *Microbiology and Molecular Biology Reviews*. **78**, 89-175.
40. Lewis, A. M., Recalde, A., Bräsen, C., Counts, J. A., Nussbaum, P., Bost, J., Schocke, L., Shen, L., Willard, D. J., Quax, T. E. F., Peeters, E., Siebers, B., Albers, S.-V. & Kelly, R. M. (2021) The biology of thermoacidophilic archaea from the order Sulfolobales, *FEMS Microbiology Reviews*.
41. Grogan, D. W. (1989) Phenotypic characterization of the archaeobacterial genus *Sulfolobus*: Comparison of five wild-type strains, *Journal of Bacteriology*. **171**, 6710-6719.
42. Pelzmann, A. M., Mickoleit, F. & Meyer, O. (2014) Insights into the posttranslational assembly of the Mo-, S- and Cu-containing cluster in the active site of CO dehydrogenase of *Oligotropha carboxidovorans*, *Journal of Biological Inorganic Chemistry*. **19**, 1399-1414.
43. Wang, H., Wang, X., Ren, H., Wang, X. & Lu, Z. (2020) Catabolism of 3-hydroxypyridine by *Ensifer adhaerens* HP1: a novel four-component gene encoding 3-hydroxypyridine dehydrogenase HpdA catalyzes the first step of biodegradation, *bioRxiv*, 2020-01.
44. Zweerink, S., Kallnik, V., Ninck, S., Nickel, S., Verheyen, J., Blum, M., Wagner, A., Feldmann, I., Sickmann, A., Albers, S.-V., Bräsen, C., Kaschani, F., Siebers, B. & Kaiser, M. (2017) Activity-based protein profiling as a robust method for enzyme identification and screening in extremophilic Archaea, *Nature Communications*. **8**, 15352.
45. Wang, K., Sybers, D., Maklad, H. R., Lemmens, L., Lewyllie, C., Zhou, X., Schult, F., Bräsen, C., Siebers, B., Valegård, K., Lindås, A.-C. & Peeters, E. (2019) A TetR-family transcription factor regulates fatty acid metabolism in the archaeal model organism *Sulfolobus acidocaldarius*, *Nature Communications*. **10**, 1542.
46. Johnsen, U., Dambeck, M., Zaiss, H., Fuhrer, T., Soppa, J., Sauer, U. & Schönheit, P. (2009) D-xylose degradation pathway in the halophilic archaeon *Haloferax volcanii*, *Journal of Biological Chemistry*. **284**, 27290-303.
47. Komorowski, L., Verheyen, W. & Schäfer, G. (2002) The archaeal respiratory supercomplex SoxM from *S. acidocaldarius* combines features of quinole and cytochrome c oxidases, *Biological Chemistry*. **383**, 1791-1799.
48. Gleißner, M., Kaiser, U., Antonopoulos, E. & Schäfer, G. (1997) The Archaeal SoxABCD Complex Is a Proton Pump in *Sulfolobus acidocaldarius*, *Journal of Biological Chemistry*. **272**, 8417-8426.
49. Schafer, G., Engelhard, M. & Muller, V. (1999) Bioenergetics of the Archaea, *Microbiology and Molecular Biology Reviews*. **63**, 570-620.
50. Bischof, L. F., Haurat, M. F., Hoffmann, L., Albersmeier, A., Wolf, J., Neu, A., Pham, T. K., Albaum, S. P., Jakobi, T., Schouten, S., Neumann-Schaal, M., Wright, P. C., Kalinowski, J., Siebers, B. & Albers, S.-V. (2019) Early response of *Sulfolobus acidocaldarius* to nutrient limitation, *Frontiers in Microbiology*. **9**.
51. Yokobori, S.-i., Nakajima, Y., Akanuma, S. & Yamagishi, A. (2016) Birth of archaeal cells: Molecular phylogenetic analyses of G1P dehydrogenase, G3P dehydrogenases, and glycerol kinase suggest derived features of archaeal membranes having G1P polar lipids, *Archaea*. **2016**, 1802675.
52. Lu, S., Wang, J., Chitsaz, F., Derbyshire, M. K., Geer, R. C., Gonzales, N. R., Gwadz, M., Hurwitz, D. I., Marchler, G. H., Song, J. S., Thanki, N., Yamashita, R. A., Yang, M., Zhang, D., Zheng, C., Lanczycki, C. J. & Marchler-Bauer, A. (2020) CDD/SPARCLE: the conserved domain database in 2020, *Nucleic acids research*. **48**, D265-D268.
53. Schnick, C., Polley, S. D., Fivelman, Q. L., Ranford-Cartwright, L. C., Wilkinson, S. R., Brannigan, J. A., Wilkinson, A. J. & Baker, D. A. (2009) Structure and non-essential function of glycerol kinase in *Plasmodium falciparum* blood stages, *Molecular Microbiology*. **71**, 533-545.
54. Balogun, E. O., Inaoka, D. K., Shiba, T., Kido, Y., Nara, T., Aoki, T., Honma, T., Tanaka, A., Inoue, M., Matsuoka, S., Michels, P. A., Harada, S. & Kita, K. (2013) Biochemical characterization of highly active

- Trypanosoma brucei gambiense* glycerol kinase, a promising drug target, *The Journal of Biochemistry*. **154**, 77-84.
55. Wilk, P., Kuśka, K., Wątor, E., Małeck, P. H., Woś, K., Tokarz, P., Dubin, G. & Grudnik, P. (2020) Structural characterization of glycerol kinase from the thermophilic fungus *Chaetomium thermophilum*, *International Journal of Molecular Sciences*. **21**, 9570.
56. Fukuda, Y., Abe, A., Tamura, T., Kishimoto, T., Sogabe, A., Akanuma, S., Yokobori, S.-i., Yamagishi, A., Imada, K. & Inagaki, K. (2016) Epistasis effects of multiple ancestral-consensus amino acid substitutions on the thermal stability of glycerol kinase from *Cellulomonas sp.* NT3060, *Journal of Bioscience and Bioengineering*. **121**, 497-502.
57. Huang, H.-S., Yoshida, T., Meng, Y., Kabashima, T., Ito, K., Nishiya, Y., Kawamura, Y. & Yoshimoto, T. (1997) Purification and characterization of thermostable glycerol kinase from *Thermus flavus*, *Journal of Fermentation and Bioengineering*. **83**, 328-332.
58. Thorner, J. W. & Paulus, H. (1973) 14 Glycerol and Glycerate Kinases in *The Enzymes* (Boyer, P. D., ed) pp. 487-508, Academic Press.
59. Sakasegawa, S., Yoshioka, I., Koga, S., Takahashi, M., Matsumoto, K., Misaki, H. & Ohshima, T. (1998) A novel glycerol kinase from *Flavobacterium meningosepticum*: characterization, gene cloning and primary structure, *Bioscience, Biotechnology, and Biochemistry*. **62**, 2388-2395.
60. Pawlyk, A. C. & Pettigrew, D. W. (2001) Subcloning, expression, purification, and characterization of *Haemophilus influenzae* glycerol kinase, *Protein Expression and Purification*. **22**, 52-59.
61. Balogun, E. O., Inaoka, D. K., Shiba, T., Kido, Y., Tsuge, C., Nara, T., Aoki, T., Honma, T., Tanaka, A., Inoue, M., Matsuoka, S., Michels, P. A. M., Kita, K. & Harada, S. (2014) Molecular basis for the reverse reaction of African human trypanosomes glycerol kinase, *Molecular Microbiology*. **94**, 1315-1329.
62. Yeh, J. I., Kettering, R., Saxl, R., Bourand, A., Darbon, E., Joly, N., Briozzo, P. & Deutscher, J. (2009) Structural characterizations of glycerol kinase: unraveling phosphorylation-induced long-range activation, *Biochemistry*. **48**, 346-56.
63. Pickl, A., Johnsen, U. & Schönheit, P. (2012) Fructose degradation in the haloarchaeon *Haloferax volcanii* involves a bacterial type phosphoenolpyruvate-dependent phosphotransferase system, fructose-1-phosphate kinase, and class II fructose-1,6-bisphosphate aldolase, *Journal of Bacteriology*. **194**, 3088-97.
64. Hayashi, S. & Lin, E. C. C. (1967) Purification and properties of glycerol kinase from *Escherichia coli*, *Journal of Biological Chemistry*. **242**, 1030-1035.
65. Atomi, H., Fukui, T., Kanai, T., Morikawa, M. & Imanaka, T. (2004) Description of *Thermococcus kodakaraensis sp. nov.*, a well studied hyperthermophilic archaeon previously reported as *Pyrococcus sp.* KOD1, *Archaea*. **1**, 263-7.
66. Unemoto, T., Hayashi, M. & Hayashi, M. (1981) Partial purification and properties of respiratory chain-linked L-glycerol 3-phosphate dehydrogenase from a marine bacterium, *Vibrio alginolyticus*, *The Journal of Biochemistry*. **90**, 619-628.
67. Rauchová, H., Fato, R., Drahot, Z. & Lenaz, G. (1997) Steady-state kinetics of reduction of coenzyme Q analogs by glycerol-3-phosphate dehydrogenase in brown adipose tissue mitochondria, *Archives of Biochemistry and Biophysics*. **344**, 235-241.
68. Thauer, R. K., Jungermann, K. & Decker, K. (1977) Energy conservation in chemotrophic anaerobic bacteria, *Bacteriological Reviews*. **41**, 100-180.
69. Cottingham, I. R. & Ragan, C. I. (1980) Purification and properties of L-3-glycerophosphate dehydrogenase from pig brain mitochondria, *Biochemical Journal*. **192**, 9-18.
70. Hatta, T., Inagaki, K., Sugio, T., Kishimoto, N. & Tano, T. (1989) Purification and Characterization of glycerol 3-phosphate dehydrogenase from two types of *Acidiphilium sp.*, *Agricultural and Biological Chemistry*. **53**, 651-658.
71. Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S. & Steinegger, M. (2022) ColabFold: making protein folding accessible to all, *Nature Methods*. **19**, 679-682.
72. Settembre, E. C., Dorrestein, P. C., Park, J.-H., Augustine, A. M., Begley, T. P. & Ealick, S. E. (2003) Structural and mechanistic studies on ThiO, a glycine oxidase essential for thiamin biosynthesis in *Bacillus subtilis*, *Biochemistry*. **42**, 2971-2981.

73. Wijerathne, H., Yao, H., Wang, Y., Lovell, S., Battaile, K. P. & Rivera, M. (2018) Bfd, a new class of [2Fe-2S] protein that functions in bacterial iron homeostasis, requires a structural anion binding site, *Biochemistry*. **57**, 5533-5543.
74. Yao, H., Wang, Y., Lovell, S., Kumar, R., Ruvinsky, A. M., Battaile, K. P., Vakser, I. A. & Rivera, M. (2012) The structure of the BfrB–Bfd complex reveals protein–protein interactions enabling iron release from bacterioferritin, *Journal of the American Chemical Society*. **134**, 13470-13481.
75. Marchler-Bauer, A. & Bryant, S. H. (2004) CD-Search: protein domain annotations on the fly, *Nucleic acids research*. **32**, W327-W331.
76. Unden, G., Dünwald, P. & Stewart, V. (2008) The aerobic and anaerobic respiratory chain of *Escherichia coli* and *Salmonella enterica*: Enzymes and energetics, *EcoSal Plus*. **3**.
77. Lalle, M., Camerini, S., Cecchetti, S., Finelli, R., Sferra, G., Müller, J., Ricci, G. & Pozio, E. (2015) The FAD-dependent glycerol-3-phosphate dehydrogenase of *Giardia duodenalis*: an unconventional enzyme that interacts with the g14-3-3 and it is a target of the antitumoral compound NDBHEX, *Frontiers in Microbiology*. **6**.
78. Robert, X. & Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server, *Nucleic Acids Research*. **42**, W320-W324.
79. Jones, D. T., Taylor, W. R. & Thornton, J. M. (1992) The rapid generation of mutation data matrices from protein sequences, *Bioinformatics*. **8**, 275-282.
80. Kumar, S., Stecher, G., Li, M., Nnyaz, C. & Tamura, K. (2018) MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms, *Molecular Biology and Evolution*. **35**, 1547-1549.
81. Wagner, M., van Wolferen, M., Wagner, A., Lassak, K., Meyer, B. H., Reimann, J. & Albers, S. V. (2012) Versatile genetic tool box for the crenarchaeote *Sulfolobus acidocaldarius*, *Frontiers in Microbiology*. **3**, 214.
82. Brock, T. D., Brock, K. M., Belly, R. T. & Weiss, R. L. (1972) *Sulfolobus*: a new genus of sulfur-oxidizing bacteria living at low pH and high temperature, *Archiv für Mikrobiologie*. **84**, 54-68.
83. Benninghoff, J. C., Kuschmierz, L., Zhou, X., Albersmeier, A., Pham, T. K., Busche, T., Wright, P. C., Kalinowski, J., Makarova, K. S. & Bräsen, C. (2021) Exposure to 1-butanol exemplifies the response of the thermoacidophilic Archaeon *Sulfolobus acidocaldarius* to solvent stress, *Applied and Environmental Microbiology*. **87**, e02988-20.
84. Klaus, T., Ninck, S., Albersmeier, A., Busche, T., Wibberg, D., Jiang, J., Elcheninov, A. G., Zayulina, K. S., Kaschani, F. & Bräsen, C. (2022) Activity-based protein profiling for the identification of novel carbohydrate-active enzymes involved in xylan degradation in the hyperthermophilic Euryarchaeon *Thermococcus sp.* Strain 2319x1E, *Frontiers in Microbiology*, 3868.
85. Sellick, C. A., Hansen, R., Stephens, G. M., Goodacre, R. & Dickson, A. J. (2011) Metabolite extraction from suspension-cultured mammalian cells for global metabolite profiling, *Nature Protocols*. **6**, 1241-1249.
86. Zor, T. & Selinger, Z. (1996) Linearization of the Bradford protein assay increases its sensitivity: theoretical and experimental studies, *Analytical biochemistry*. **236**, 302-308.
87. Tsai, C.-L., Tripp, P., Sivabalasarma, S., Zhang, C., Rodriguez-Franco, M., Wipfler, R. L., Chaudhury, P., Banerjee, A., Beeby, M., Whitaker, R. J., Tainer, J. A. & Albers, S.-V. (2020) The structure of the periplasmic FlaG–FlaF complex and its essential role for archaeellar swimming motility, *Nature Microbiology*. **5**, 216-225.

Supplementary information

Supplementary table 1: Strains and Plasmids used in this study

Designation	Genotype or description	Reference/Source
Strains		
MW001	<i>Sulfolobus acidocaldarius</i> strain; For growth or deletion mutant construction	[1]
DH5 α	<i>Escherichia coli</i> strain; Plasmid construction	Hanahan, USA
Rosetta (DE3)	<i>Escherichia coli</i> strain; Gene expression; pRARE (chloramphenicol resistance)	Stratagene, USA
MW00G	<i>Sulfolobus acidocaldarius</i> strain; glycerol adapted strain	This study
Plasmids		
pET15b	<i>E. coli</i> expression plasmid carrying an N-terminal His tag (ampicillin resistance)	Novagen, USA
pETDuet-1	<i>E. coli</i> coexpression plasmid carrying an N-terminal His tag (multiple cloning site 1) or S-tag (multiple cloning site 2) (ampicillin resistance)	Novagen, USA
pBS-Ara-albaUTR-FX	Expression plasmid for <i>S. acidocaldarius</i> containing an arabinose inducible promotor and N-terminal strep tag (ampicillin resistance)	unpublished
pSVAaraFX-HA	Expression plasmid for <i>S. acidocaldarius</i> containing an arabinose inducible promotor and HA-tag (ampicillin resistance)	[2]
pEt15b-saci_2032	Expression of <i>saci_2032</i> in <i>E. coli</i>	This work
pETduet1-saci_2032	Expression of <i>saci_2032</i> in <i>E. coli</i>	This work
pETduet1-saci_2032-saci_2031	Coexpression of <i>saci_2031</i> , <i>saci_2032</i> in <i>E. coli</i>	This work
pBS-Ara-albaUTR-FX-saci_2033	Expression of <i>saci_2033</i> in <i>S. acidocaldarius</i>	This work
pSVA-AraFX-HA-saci_2031	Expression of <i>saci_2031</i> in <i>S. acidocaldarius</i>	This work, [2]

Supplementary table 2: Sequences of oligonucleotide used in this study

ORFs	Name	Sequences (5'→3') (restriction sites are marked in red)
Oligonucleotides for cloning		
<i>Saci_2031</i>	<i>Fw</i> (pETDuet1)- <i>NdeI</i>	GAGGAGCATATGATGGAGATCAGCGGAGAG
	<i>Rev</i> (pETDuet1)- <i>XhoI</i>	GAGGAGCTCGAGCTAGGTAAGTCTAGAAC
	<i>Fw</i> (pSVA-AraFX-HA)- <i>NcoI</i>	CATCCATGGAGATCAGCGGAGAGTTTG
	<i>Rev</i> (pSVA-AraFX-HA)- <i>XhoI</i>	CATCTCGAGGGTAAGTCTAGAACGTATACAG
<i>Saci_2032</i>	<i>Fw</i> (pET15b)- <i>NdeI</i>	GTCCGACATATGATGGAAATAAAAACAAGCG
	<i>Rev</i> (pET15b)- <i>BamHI</i>	AATTACGGATCCCATCTCCCTCTTGCAATTAATG
	<i>Fw</i> (pETDuet1)- <i>BamHI</i>	GAGAGGGATCCATGGAAATAAAAACAAGCG
	<i>Rev</i> (pETDuet1)- <i>HindIII</i>	GAGAGAAGCTTTCATCTCCCTCTTGCAATTAATG
<i>Saci_2033</i>	<i>Fw</i> (pBS-Ara)- <i>NcoI</i>	GAAACCATGGTGGCTGAAAAATACGTGATAG
	<i>Rev</i> (pBS-Ara)- <i>XhoI</i>	GGTCTCGAGACCCCAATAGTCTTAGC

Supplementary table 3: Differentially expressed genes and proteins given as log₂-fold-changes in *S. acidocaldarius* MW00G after growth on glycerol compared to D-xylose identified via RNASeq (transcriptome) and iTRAQ (proteome) analyses.

		Transcriptome (log ₂ -fold changes)			Proteome (log ₂ -fold changes)		
		10 mM	20 mM	40 mM	10 mM	20 mM	40 mM
Glycerol concentration							
Upregulated							
Locus tag	Function						
Saci_0451	ESCRT-III	3.75	4.08	3.27	6.86	7.15	6.32
Saci_0942	CopG family transcriptional regulator	2.10	2.61	2.19	3.51	3.72	3.68
Saci_1050	ATPase, ParA family	2.96	3.00	2.85	2.61	2.56	2.71
Saci_1052	GYD domain-containing secretory protein	3.87	3.82	3.81	5.32	5.38	5.20
Saci_1054	3-(methylthio)propionyl-CoA ligase	3.39	2.44	2.20	4.63	4.52	4.68
Saci_1058	Xylulokinase	3.12	2.99	2.94	2.62	2.58	2.71
Saci_1099	Betaine-aldehyde dehydrogenase	3.00	2.32	2.17	5.40	5.58	5.50
Saci_1228	Nucleoside triphosphate hydrolase	3.37	3.10	2.56	3.84	4.00	3.39
Saci_1372	Vesicle-fusing ATPase	3.04	2.25	2.00	4.05	3.83	4.02
Saci_1416	ESCRT-III	3.91	3.99	3.25	4.42	4.45	3.85
Saci_1616	Antitoxin (DNA-binding domain)	2.08	2.54	2.76	3.29	3.17	3.20
Saci_1762	ABC-type	3.33	3.11	3.24	3.17	3.05	2.98
Saci_1763	ABC-type	3.54	2.91	3.15	3.75	3.82	3.36
Saci_1764	ABC-type	3.53	3.12	3.49	2.87	2.94	2.89
Saci_1765	ABC-type	3.39	3.17	3.38	2.89	2.98	2.89
Saci_1804	DUF2173 domain-containing protein	2.39	2.19	2.74	4.85	4.94	5.04
Saci_1810	Energy-coupling factor cobalt transporter, ATP-binding protein	2.83	2.37	2.35	3.51	3.53	3.58
Saci_1855	DUF973 domain-containing membrane protein	7.66	6.92	6.98	5.81	5.90	5.88
Saci_1856	DUF973 domain-containing membrane protein	3.74	3.88	3.90	4.13	4.15	4.07
Saci_2032	Glycerol 3-phosphate dehydrogenase	3.92	3.31	2.97	3.01	2.98	3.10
Saci_2033	Glycerol kinase	3.96	3.51	3.19	2.47	2.47	2.24
Saci_2034	Glycerol uptake facilitator	3.85	3.52	3.20	4.84	3.59	4.66
Saci_2139	CBS domain containing protein	4.62	5.54	5.81	3.51	3.38	3.05
Saci_2188	Ribonucleoside-diphosphate reductase, subunit beta	3.42	2.99	3.07	2.93	2.61	2.36
Saci_2206	Lipoate-protein ligase A	4.77	4.16	4.13	3.68	3.69	3.59
Saci_2208	3-hydroxyacyl-CoA dehydrogenase	5.30	4.72	4.56	9.34	9.38	9.45
Saci_2209	Acetyl-CoA C-acetyltransferase	4.26	3.63	3.45	2.75	2.76	2.54

Saci_2210	DUF35 OB-fold domain containing protein	5.15	4.12	4.16	6.93	6.76	7.02
Saci_2211	3-(methylthio)propionyl-CoA ligase	2.79	2.58	2.54	3.00	2.98	2.87
Saci_2212	Ribonucleoside-diphosphate reductase, subunit beta	6.32	5.23	5.38	3.91	3.95	3.37
Saci_2230	DUF35 OB-fold domain containing protein	2.94	2.42	2.68	2.80	2.37	2.74
Saci_2231	DUF35 OB-fold domain containing protein	3.29	2.63	2.93	2.95	2.85	2.96
Saci_2232	Acetyl-CoA C-acetyltransferase	3.39	2.62	2.86	2.77	2.65	2.61
Saci_2233	Acetyl-CoA C-acetyltransferase	3.24	2.81	3.07	2.80	2.73	2.71
Saci_2235	3-hydroxypropionyl-CoA synthase	5.67	5.07	5.30	8.19	8.39	7.99
Saci_2258	Membrane protein, terminal oxidase function	2.25	2.60	2.84	2.48	2.43	2.39
Saci_2259	Cytochrome B558 subunit B	2.61	2.71	3.02	4.61	4.10	4.54
Saci_2261	Rieske Fe-S protein	2.49	3.23	3.35	5.08	4.87	5.16
Saci_2263	Quinol oxidase subunit 1/3	2.97	3.05	3.13	5.85	5.46	5.73

Downregulated

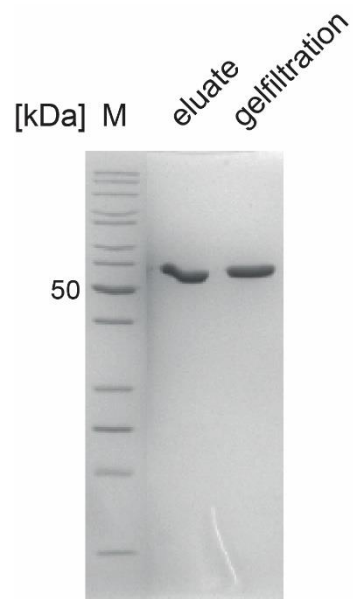
Locus tag	Function						
Saci_0137	Phosphomethylpyrimidine synthase	-4.55	-4.23	-4.57	-2.86	-2.93	-2.80
Saci_0314	Hypothetical protein	-5.43	-4.33	-4.15	-2.36	-3.06	-2.10
Saci_0526	Thiazole/oxazole-forming peptide maturase	-2.23	-2.32	-2.03	-3.01	-3.33	-3.20
Saci_0531	ABC-type Mn/Zn transport system	-3.54	-2.59	-2.54	-3.76	-3.33	-3.39
Saci_1700	2,5-dioxovalerate dehydrogenase	-6.09	-5.54	-6.06	-7.18	-7.32	-7.08
Saci_1707	ABC-type molybdate transport system	-5.53	-5.93	-5.31	-4.34	-4.79	-5.08
Saci_1830	Alpha/beta hydrolase	-4.95	-4.19	-4.07	-4.16	-4.49	-3.60
Saci_1831	Transcriptional regulator	-4.13	-3.81	-3.81	-2.34	-2.04	-2.21
Saci_1938	2,5-dioxovalerate dehydrogenase	-4.08	-4.01	-4.17	-6.83	-6.95	-6.73
Saci_1939	2-dehydro-3-deoxy-D-arabinonate dehydratase	-3.89	-3.44	-3.77	-5.70	-5.50	-5.48
Saci_2090	Rieske Fe-S protein	-3.12	-2.86	-2.62	-3.33	-3.93	-3.91
Saci_2122	Xylose/arabinose ABC-type transport system	-2.35	-2.43	-2.41	-4.07	-4.10	-4.24
Saci_2203	Sulfate adenylyltransferase	-3.61	-4.16	-4.27	-2.16	-2.49	-2.75
Saci_2204	Hypothetical protein	-3.73	-4.08	-4.25	-2.89	-3.60	-3.46

Supplementary table 4: Differentially enriched proteins via co-immunoprecipitation anti-HA antibodies coupled to magnetic beads using *S. acidocaldarius* MW00G expressing HA-*saci_2031* given as log2-fold-changes.

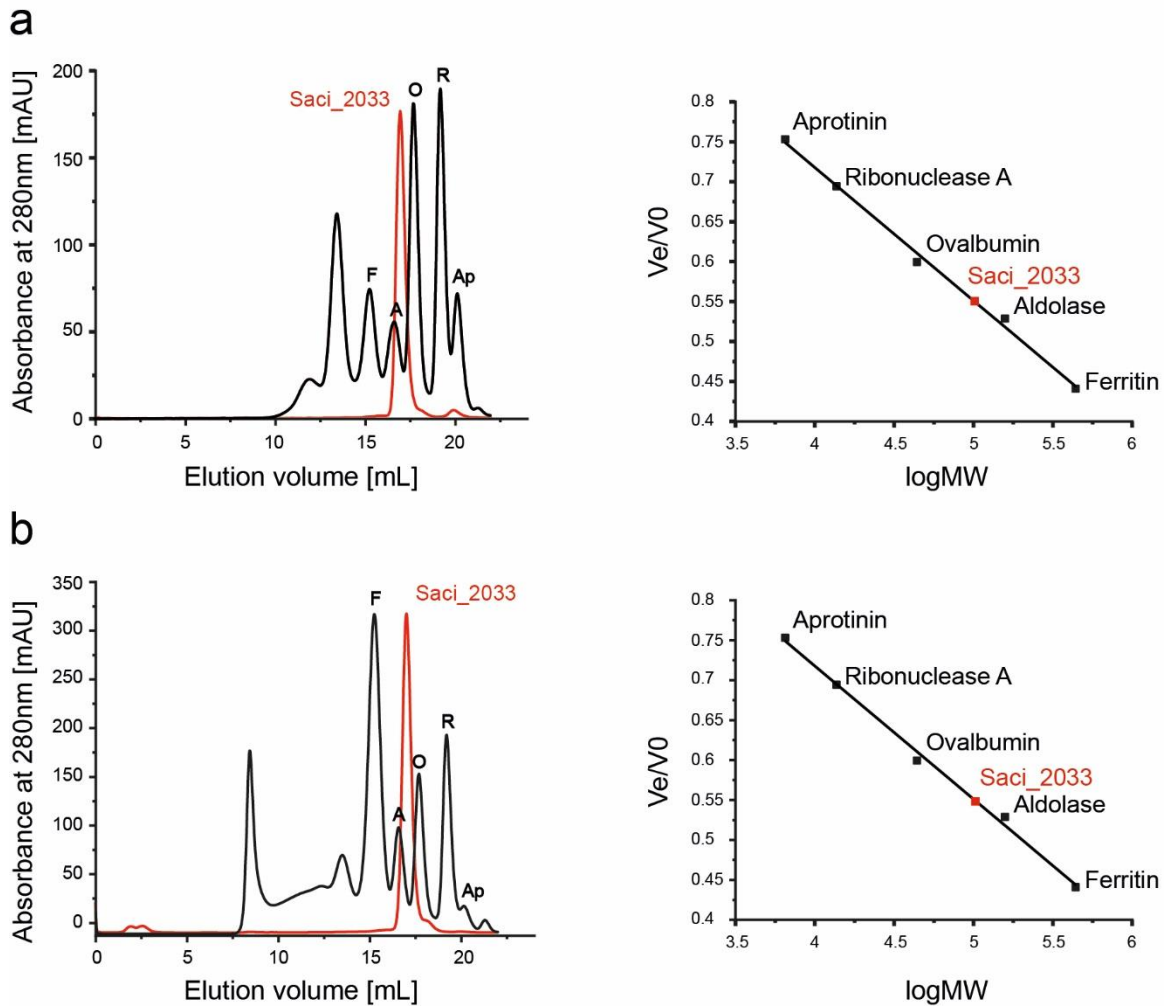
majority protein ID	gene name ID	log2-fold changes
Q4J7A4	<i>saci_2031</i>	14.15
Q4J7A2	<i>saci_2032</i>	7.36
Q4JB29	<i>rpl37ae</i>	5.31
Q4JA54	<i>EIF1A</i>	5.11
P13125	<i>saci_0362</i>	4.90
Q4JB43	<i>rpl2</i>	4.52
P28461	<i>EIF5A</i>	4.46
Q4J9B2	<i>rpl44e</i>	4.17
Q4JAS4	<i>saci_0727</i>	3.79
Q4JA43	<i>saci_0975</i>	3.68
Q4J953	<i>saci_1337</i>	3.56
Q4J974	<i>saci_1321</i>	3.42
Q4JAZ9	<i>rpl37e</i>	3.23
Q4JBQ8	<i>saci_0355</i>	3.19
Q4JB50	<i>rpl14</i>	3.06
Q4JAV2	<i>rpl24e</i>	3.03
Q4JAP0	<i>saci_0766</i>	2.57
P13123	<i>saci_0064</i>	2.42
F2Z5Y7	<i>saci_1588</i>	2.38
Q4J9Q8	<i>saci_1118</i>	2.13

Supplemental table 5: Kinetic parameters, source organism, oligomeric state, and cosubstrate spectrum of different characterized G3PDHs. DCPIP: 2,6-Dichlorophenolindophenol; FAD: flavin adenine dinucleotide; G3P: Glycerol-3-phosphate; MTT: (4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazoliumbromid; PMS: Phenazine methosulfate; UQ: Ubiquinone.

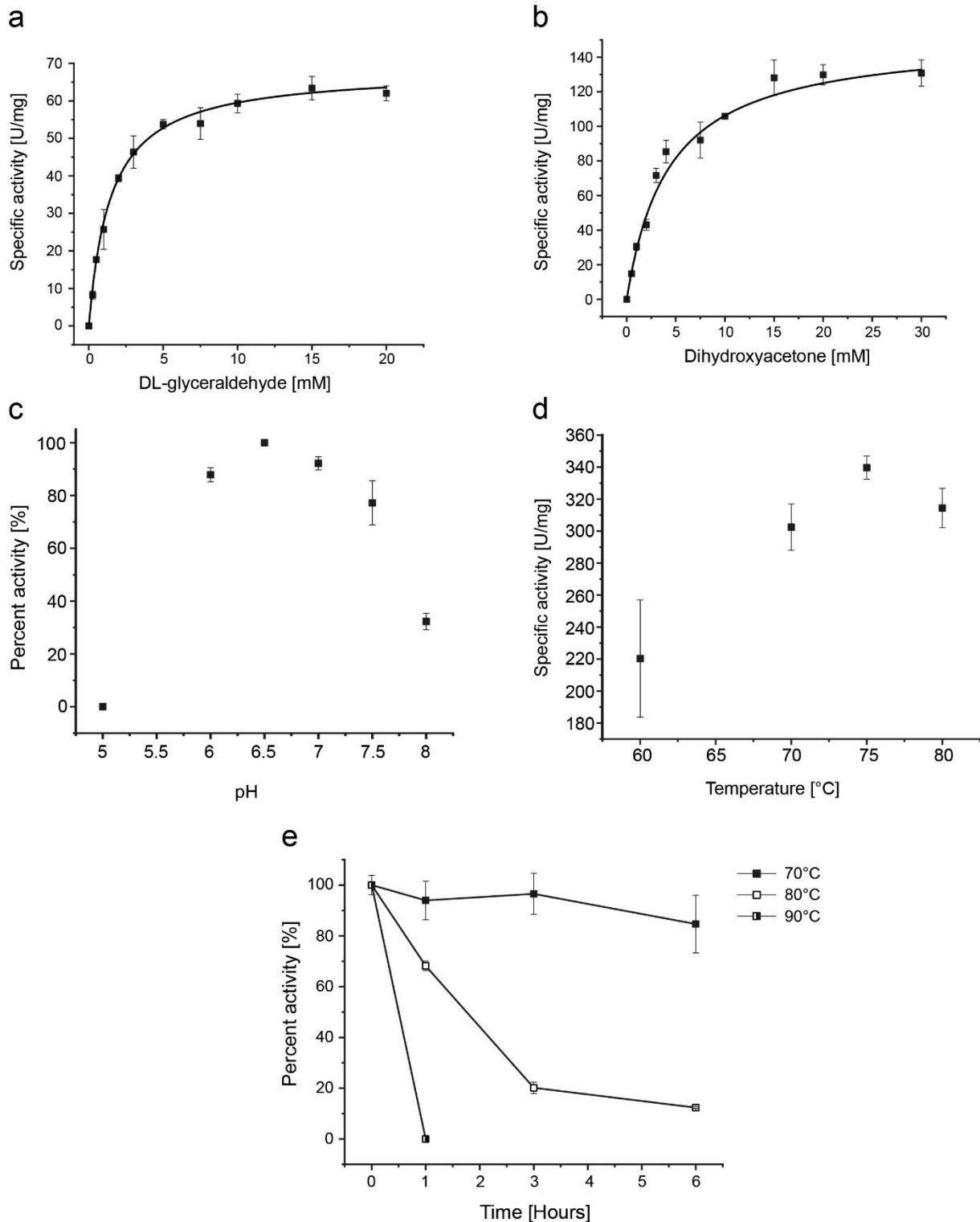
Enzyme	Organism	Cofactors per monomer	Oligomeric state	Substrate	V _{max} [U/mg]	K _M (for G3P) [mM]	K _M (electron acceptor) [mM]	Literature
Saci_2032	<i>S. acidocaldarius</i>	1 FAD	homodimer	G3P	44.5 (DCPIP)	0.055 (DCPIP)	0.179 (UQ -1)	This work
GlpAB	<i>E. coli</i>	1 FAD 2 non-heme iron	heterodimer	G3P	34.4 (PMS-MTT)	0.339 (PMS-MTT)		[3]
GlpD	<i>E. coli</i>	2 FAD	homodimer	G3P	5.8 (DCPIP)	0.8 (DCPIP)		[4-6]
GlpD	Pig brain mitochondria	FAD	homodimer	G3P	2.9 (DCPIP)	6.2 (U-Q10) 10 (DCPIP) 10 (UQ-0/1/2/6)	0.048 (UQ-0) 0.018 (UQ-1) 0.022 (DCPIP)	[7, 8]
GlpD	<i>Vibrio alginolyticus</i>	FAD	unknown	G3P	39.2 (PMS) 85.8 (ferricyanide)	4.02 (PMS) 3.7 (ferricyanide)	0.13 (PMS) 0.55 (ferricyanide)	[9]
GlpD	<i>Acidiphilium</i> sp. 63	0.58 FAD	homodimer	G3P	32 (PMS-MTT)	-	-	[10]



Supplemental Figure 1: Purification of recombinant Saci_2032 and Saci_2033. The glycerol kinase (GK) Saci_2033 was homologously produced in *S. acidocaldarius* MW001 using the pBS-Ara-albaUTR-FX and purified via Strep-tag affinity chromatography and size exclusion chromatography. Proteins (2 μ g) were separated via SDS-PAGE and stained via Coomassie Blue. Marker (M): prestained PageRuler™ (Thermo Fischer scientific, USA).

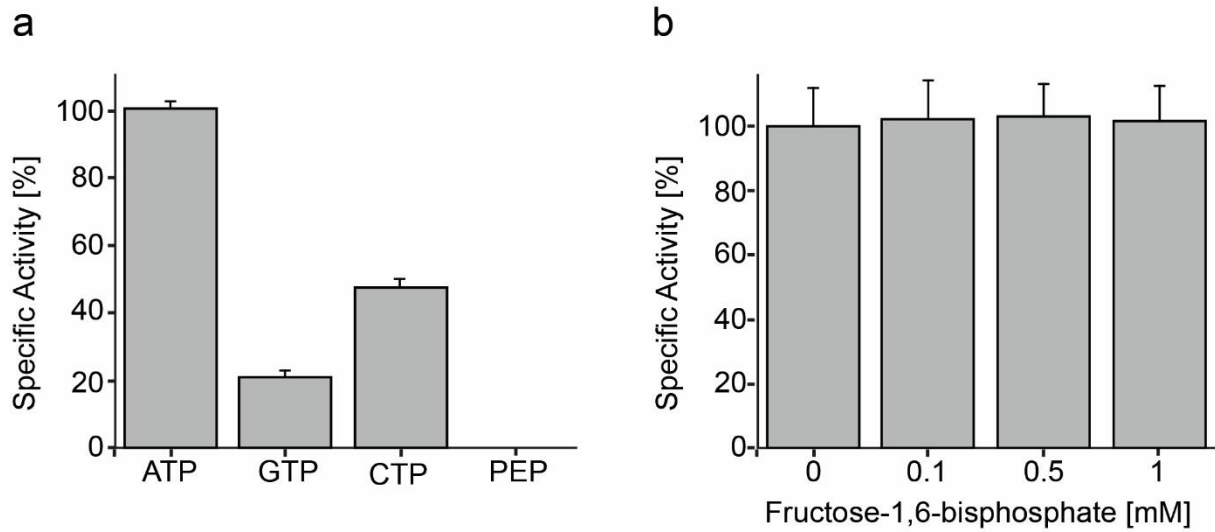


Supplementary figure 2: Size exclusion chromatography of the recombinant GK Saci_2033 from *S. acidocaldarius*. Recombinant proteins were separated on a Superose 6 10/300 column in both, the absence (**a**) and presence of 10 mM glycerol (**b**). In the elution profiles (left panels) the GK Saci_2033 from *S. acidocaldarius* is shown in red and the standard proteins for calibration in black. In the right panels the calibration curves are shown.

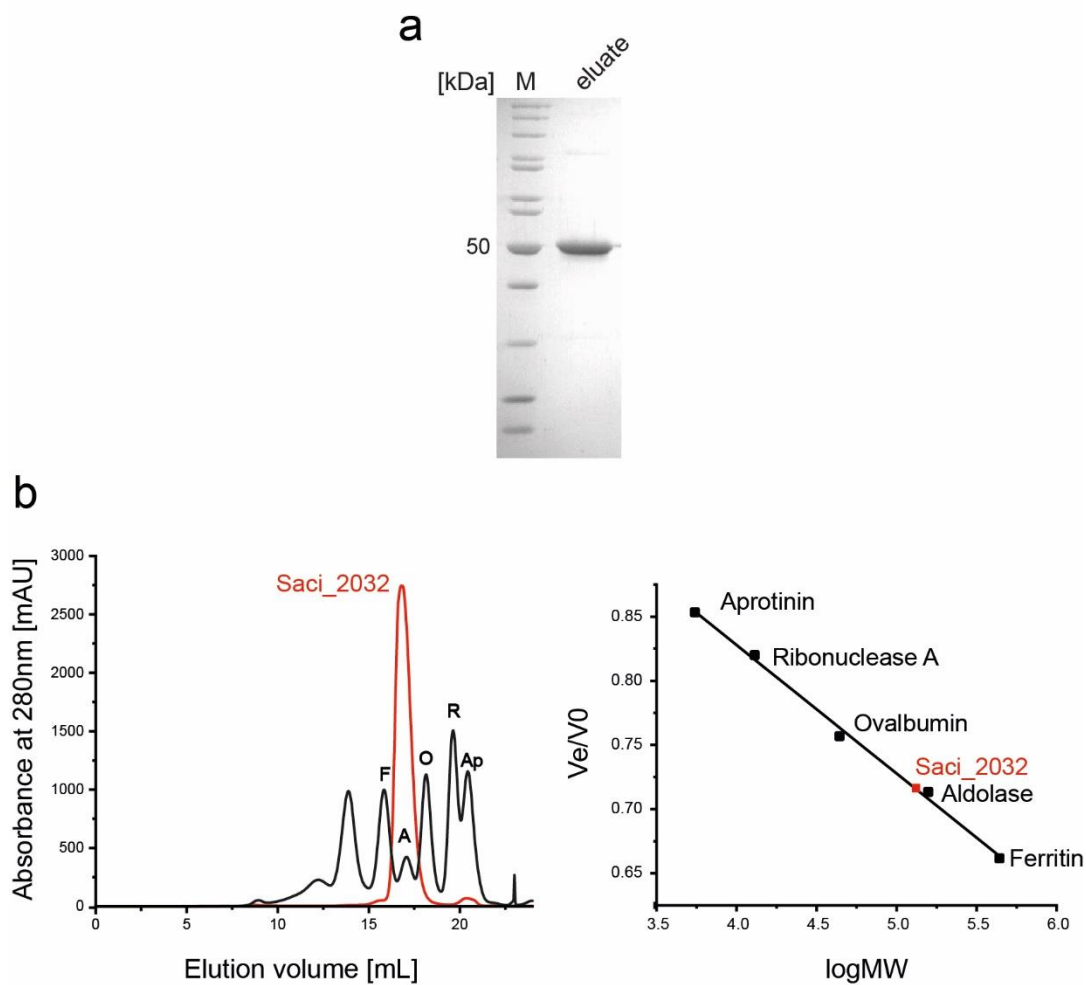


Supplementary Figure 3: Biochemical characterization of recombinant GK Saci_2033. The kinetic properties of Saci_2033 with (a) DL-glyceraldehyde and dihydroxyacetone (b) as substrate were determined in the coupled assay via PK and LDH at 50°C in 0.1 M MOPS pH 6.5. (c) The pH dependence of the Saci_2033 activity was determined in the range of pH 5.0-8.0 using a mixed buffer system (0.05 M MES, 0.05 M HEPES, 0.05 M TRIS) at 50°C. In a continuous assay coupling the formation of ADP from ATP to the oxidation of NADH via pyruvate kinase (PK) and lactate

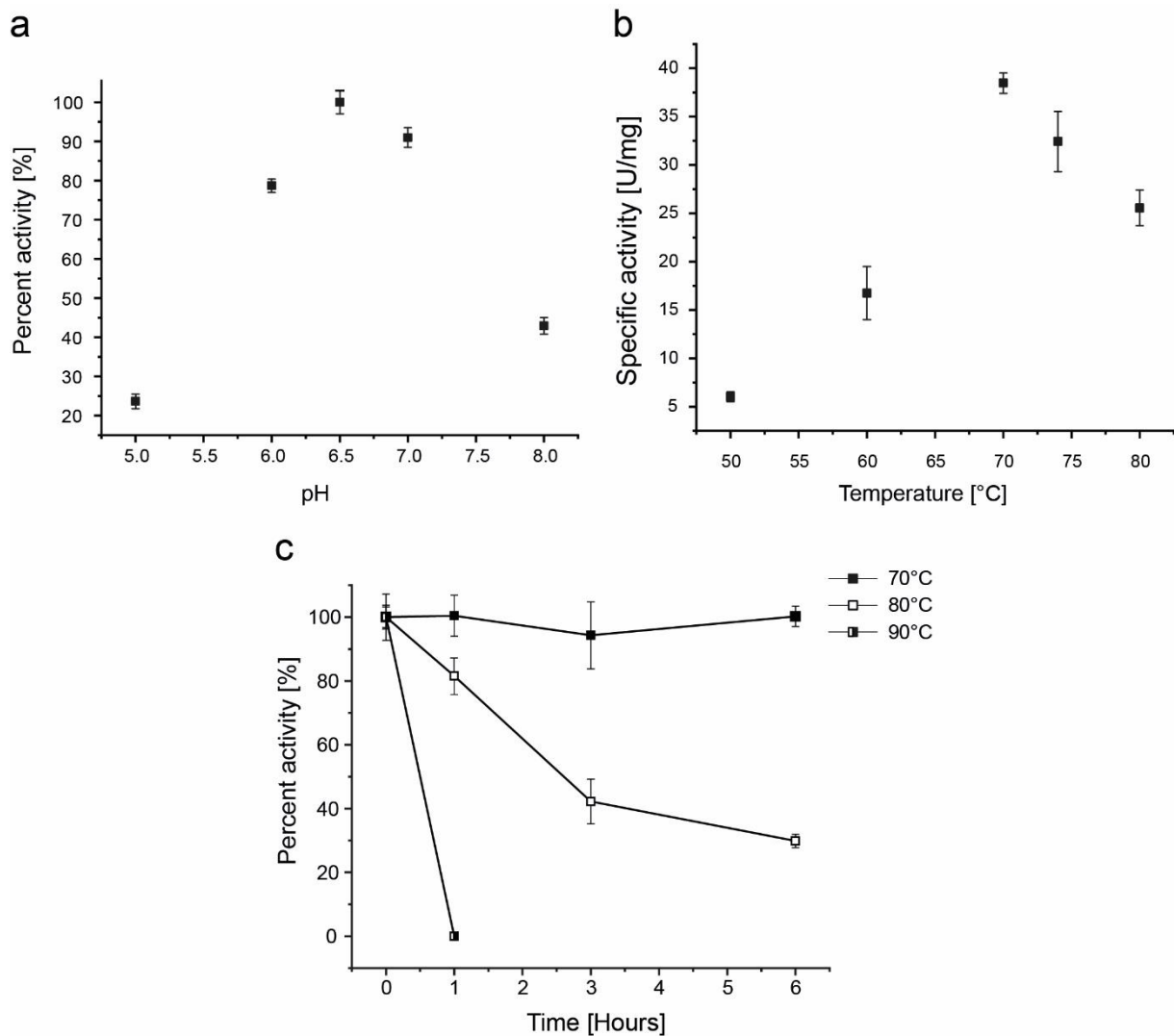
dehydrogenase (LDH) the decrease in absorbance was followed at 340 nm. **(d)** The optimum temperature of Saci_2033 was determined in a continuous assay using G3PDH (Saci_2032) and 0.1 mM DCPIP between 60°C to 80°C. **(e)** The thermal stability of the Saci_2033 GK was determined by monitoring the residual activity upon incubation at the given temperatures for the indicated time intervals. The enzyme was incubated in 100 mM MOPS-KOH, pH 6.5 (temperature adjusted) at a protein concentration of 0.05 mg ml⁻¹ and after the respective incubation period the residual activity was determined in the coupled assay via PK and LDH at 50°C as described in materials and methods. Experiments were performed in triplicate and error bars indicate the SD of the mean.



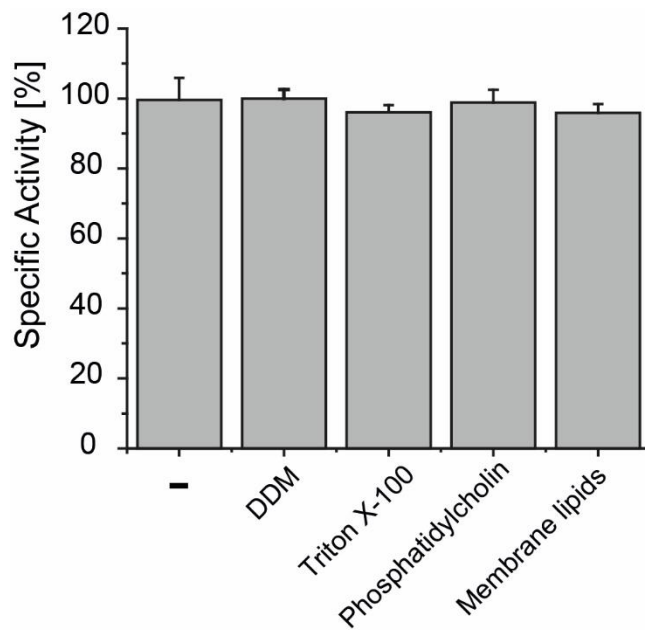
Supplementary Figure 4: Effect of fructose-1,6-bisphosphate on enzyme activity and nucleotide/phosphate donor specificity of the recombinant Saci_2033 GK. (a) The enzyme activity of Saci_2033 with 2 mM glycerol as substrate and 5 mM of different phosphate donors was determined in a continuous assay by coupling the phosphorylation of glycerol to the reduction of DCPIP via G3PDH Saci_2032 mediated oxidation of G3P to DHAP as absorbance decrease at 600 nm at 75°C. Experiments were performed in triplicate and error bars indicate the SD of the mean. (b) The enzyme activity of Saci_2033 in presence of increasing fructose 1,6-bisphosphate concentrations was determined in a continuous coupled LDH/PK assay following NADH oxidation as decrease in absorbance at 340 nm and 50°C.



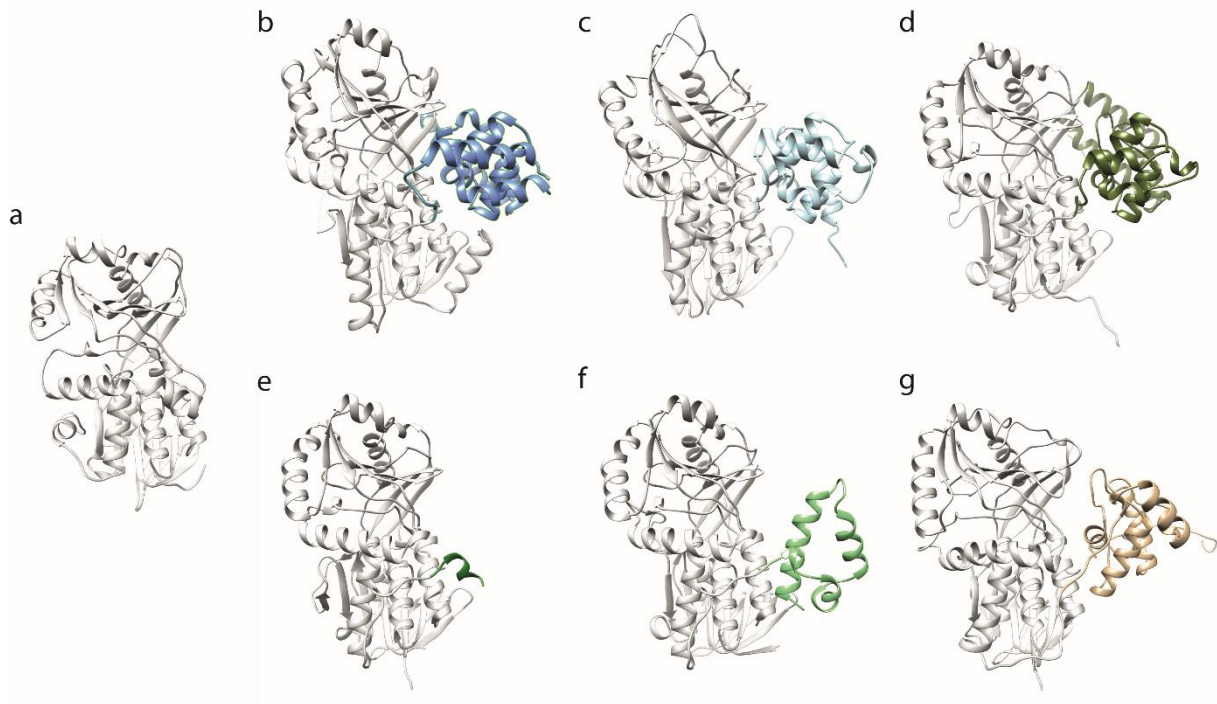
Supplementary figure 5: SDS-Page and Size exclusion chromatography of the recombinant G3PDH Saci_2032 from *S. acidocaldarius* and the corresponding column calibration. (a) The Glycerol-3-phosphate dehydrogenase (G3PDH) Saci_2032 was homologously produced in *E. coli* Rosetta using pET15b and purified via His-tag affinity chromatography. Proteins (2 μ g) were separated via SDS-PAGE and stained via Coomassie Blue. Marker (M): prestained PageRuler™ (Thermo Fischer scientific, USA). **(b)** Recombinant proteins were separated on a Superose 200 10/300 column. In the elution profiles (left panel) the G3PDH Saci_2032 from *S. acidocaldarius* is shown in red and the standard proteins for calibration in black. In the right panel the calibration curves are shown.



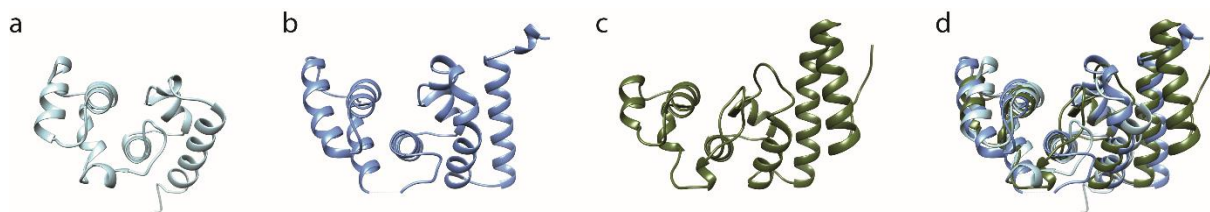
Supplementary Figure 6: Biochemical characterization of recombinant G3PDH Saci_2032. (a) The pH dependence of Saci_2032 was determined in a continuous assay as G3P dependent reduction of DCPIP in a mixed buffer system (0.05 M MES, 0.05 M HEPES, 0.05 M TRIS) at 70°C. (b) The optimum temperature of Saci_2032 was determined in the same assay system in 0.05 M MES pH 6.5 adjusted between 60°C to 80°C. (c) Heat stability of Saci_2032 was determined by monitoring the residual activity upon incubation between 70°C to 90°C for up to 6 hours. The enzyme was incubated in 50 mM MES-KOH, pH 6.5 (temperature adjusted) at a protein concentration of 0.18 mg ml⁻¹. After the respective incubation periods the residual G3PDH activity was determined with DCPIP as electron acceptor at 70°C as described above. Experiments were performed in triplicate and error bars indicate the SD of the mean.



Supplemental Figure 7: Effect of non-ionic detergents and phospholipids on recombinant G3PDH Saci_2032. Activity of Saci_2032 in the presence of potential activators was determined in a continuous assay by coupling the oxidation of G3P to the reduction of DCPIP in a MES-KOH buffer pH 6.5 at 50°C. 0.5 % DDM, 0.5 % of Triton X-100, 50 µg of Phosphatidylcholine or 50 µg of isolated *S. acidocaldarius* membrane fractions were tested.

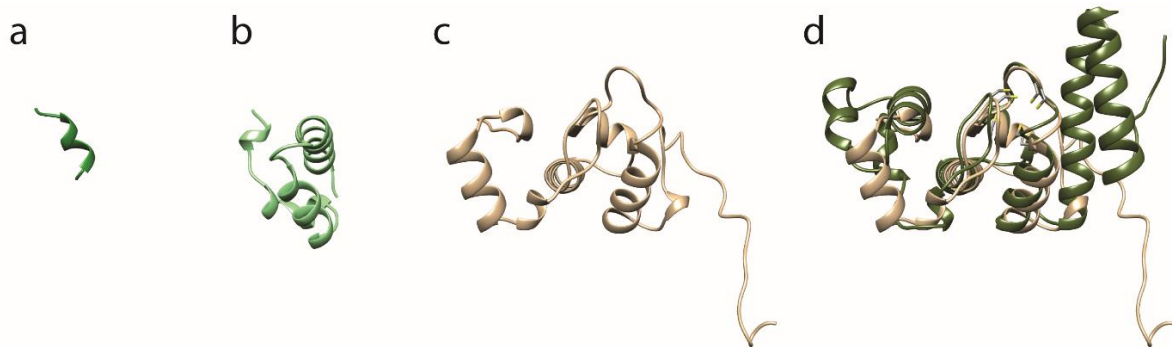


Supplementary figure 8: Comparison of the overall fold of the catalytic subunits of the different types of FAD dependent G3P oxidizing enzymes. The ribbon representations of crystal structures of (a) the *B. subtilis* glycine oxidase (1ng3) comprising the DAAO fold lacking any C-terminal extensions (shown in gray in all structures) is comparatively illustrated with those of (b) GlpO (2rgo), (c) GlpD (2qcu), as well as the structural models (generated using CoLabFold) of (d) *E. coli* GlpA, (e) the GlpA homologues of *T. pendens* and (f) *S. acidocaldarius*, and (g) the GlpTk from *T. kodakarensis* (TK1393). The C-terminal domains are highlighted using the same colour code as in Fig. 9-10 in the main text to emphasize the differences in size and fold.

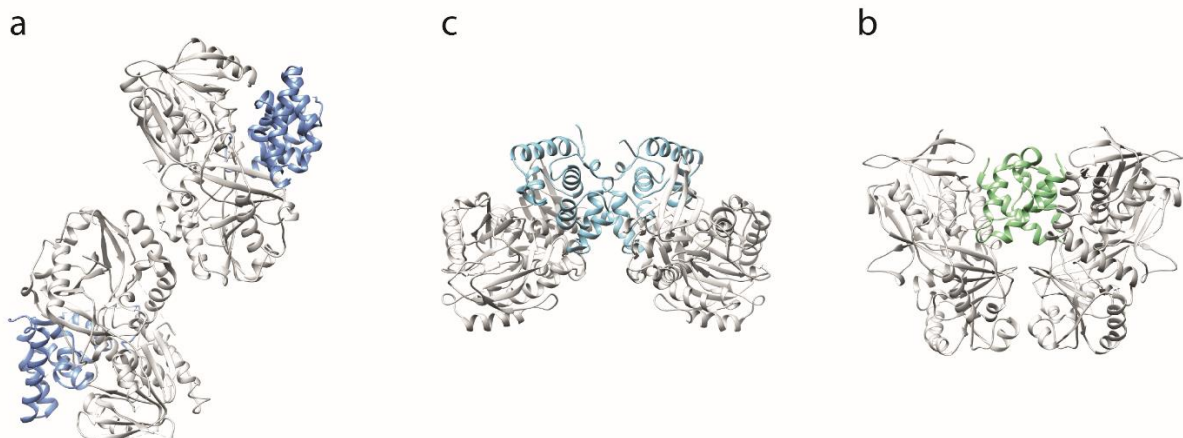


Supplementary figure 9: Comparison of C-terminal domains of GlpD (a), GlpO (b), and GlpA (c). The crystal structures of the C-termini of GlpD from *E. coli* (2qcu) (a) and GlpO from *Streptococcus* sp. (2rgo) (b) are shown, as well as the structural model of the C-terminus of GlpA from *E. coli* (c). The comparison reveals that the GlpD C-terminal fold is also present in GlpO and GlpA but extended by one and two helices, respectively. In (d) the superimposition of the three structures is shown in the

same orientation as in (a)-(c). The dashed line in (b) indicates missing structural information for M518 in the C-terminus of *Streptococcus* sp. GlpO.



Supplementary figure 10: Comparison of C-terminal extensions from GlpA-like G3PDHs from *T. pendens* (a) and *S. acidocaldarius* (Saci_2032, b), and of the GlpTk from *T. kodakarensis* (TK1393, c). The ribbon representation of the structural models of the C-terminal are shown. To illustrate the presence of the bfd-fold in GlpTk and GlpA a superimposition of both C-termini is presented (d). Both structures match in their central parts as well as in the four conserved cysteine residues shown as stick models whereas the other parts of the structure differ remarkably.



Supplementary figure 11: Comparison of the dimer formation in the crystal structures of the (a) soluble, cytoplasmic GlpO (*Streptococcus* spp., 2rgo) and (b) the monotopic, membrane bound GlpD (*E. coli*, 2qcu) with (c) the structural model of the GlpA-like Saci_2032 dimer. The three structures illustrate the different role of the C-terminus (coloured according to the preceding figures) in dimer formation as well as the resulting different spatial orientation of the DAAO-folds of the subunits to each other.

1. Wagner, M., van Wolferen, M., Wagner, A., Lassak, K., Meyer, B. H., Reimann, J. & Albers, S. V. (2012) Versatile genetic tool Box for the crenarchaeote *Sulfolobus acidocaldarius*, *Frontier in Microbiology* **3**, 214.
2. Van der Kolk, N., Wagner, A., Wagner, M., Waßmer, B., Siebers, B. & Albers, S.-V. (2020) Identification of XylR, the activator of arabinose/xylose inducible regulon in *Sulfolobus acidocaldarius* and its application for homologous protein expression, *Frontiers in Microbiology*. **11**, 1066.
3. Schryvers, A. & Weiner, J. (1981) The anaerobic sn-glycerol-3-phosphate dehydrogenase of *Escherichia coli*. Purification and characterization, *Journal of Biological Chemistry*. **256**, 9959-9965.
4. Schryvers, A., Lohmeier, E. & Weiner, J. H. (1978) Chemical and functional properties of the native and reconstituted forms of the membrane-bound, aerobic glycerol-3-phosphate dehydrogenase of *Escherichia coli*, *Journal of Biological Chemistry*. **253**, 783-788.
5. Weiner, J. H. & Heppel, L. A. (1972) Purification of the membrane-bound and pyridine nucleotide-independent L-glycerol 3-phosphate dehydrogenase from *Escherichia coli*, *Biochemical and Biophysical Research Communications*. **47**, 1360-1365.
6. Yeh, J. I., Chinte, U. & Du, S. (2008) Structure of glycerol-3-phosphate dehydrogenase, an essential monotopic membrane enzyme involved in respiration and metabolism, *Proceedings of the National Academy of Sciences*. **105**, 3280-3285.
7. Cottingham, I. R. & Ragan, C. I. (1980) Purification and properties of L-3-glycerophosphate dehydrogenase from pig brain mitochondria, *Biochemical Journal*. **192**, 9-18.
8. Dawson, A. & Thorne, C. (1969) Preparation and some properties of L-3-glycerophosphate dehydrogenase from pig brain mitochondria, *Biochemical Journal*. **111**, 27-34.
9. Unemoto, T., Hayashi, M. & Hayashi, M. (1981) Partial purification and properties of respiratory chain-linked L-glycerol 3-phosphate dehydrogenase from a marine bacterium, *Vibrio alginolyticus*, *The Journal of Biochemistry*. **90**, 619-628.
10. Hatta, T., Inagaki, K., Sugio, T., Kishimoto, N. & Tano, T. (1989) Purification and characterization of glycerol 3-phosphate dehydrogenase from two types of *Acidiphilium sp*, *Agricultural and Biological Chemistry*. **53**, 651-658.

Chapter 3.3: Enhanced underground metabolism challenges life at high temperatures – metabolic thermoadaptation in hyperthermophilic Archaea

Enhanced underground metabolism challenges life at high temperature—metabolic thermoadaptation in hyperthermophilic Archaea

Christian Schmerling¹, Theresa Kouril², Jacky Snoep^{2,3},
Christopher Bräsen¹ and Bettina Siebers¹

Abstract

The text-book picture of a perfect, well organised metabolism with highly specific enzymes, is challenged by non-enzymatic reactions and promiscuous enzymes. This, so-called ‘underground metabolism’, is a special challenge for hyperthermophilic Archaea that thrive at temperatures above 80 °C and possess modified central metabolic pathways often with promiscuous enzymes. Hence, the question arises how extremely thermophilic Archaea can operate their unusual metabolism at temperatures where many pathway intermediates are unstable? We herein discuss current insights in the underground metabolism and metabolic thermoadaptation of (hyper)thermophilic Archaea. So far, only a few repair enzymes and salvaging pathways have been investigated in Archaea. Studies of the central carbohydrate metabolism indicate that a number of different strategies have evolved: 1) reduction of the concentration of unstable metabolites, 2) different pathway topologies are used with newly induced enzymes, and 3) damaged metabolites are removed via new metabolic pathways.

Addresses

¹ 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

² Department of Biochemistry, University of Stellenbosch, Private Bag X1, Matieland 7602, South Africa

³ Department of Molecular Cell Physiology, Vrije Universiteit Amsterdam, De Boelelaan 1085, 1081 HV Amsterdam, the Netherlands

Corresponding author: Siebers, Bettina (bettina.siebers@uni-due.de)

Current Opinion in Systems Biology 2022, 30:100423

This review comes from a themed issue on **Metabolic Networks**

Edited by **Sarah-Maria Fendt** and **Markus Ralser**

For a complete overview see the [Issue](#) and the [Editorial](#)

Available online 28 April 2022

<https://doi.org/10.1016/j.coisb.2022.100423>

2452-3100/© 2022 Published by Elsevier Ltd.

Keywords

Archaea, Hyperthermophile, Underground metabolism, Metabolite instability, Metabolic thermoadaptation, Sulfolobus.

Introduction

The term underground metabolism [1] was originally applied to the *in vivo* formation of metabolic side-/by-products or damaged metabolites by side reactions of promiscuous ‘sloppy’ enzymes but is now often used in a broader sense to include spontaneous (non-enzymatically catalysed) chemical reactions [2–5]. In this short review, we will use this broader definition, which includes enzyme-catalysed and spontaneous chemical reactions. In bacterial and eukaryotic organisms, the importance of underground metabolism is becoming increasingly evident and the picture of a perfect, well-organised metabolism with precise and specific enzymes has changed. A.G. Goubelev nicely put it in a nutshell: ‘Nothing of chemistry disappears in biology’ [6]. Hence, many canonical metabolites in the cell are unstable and reactive and therefore challenge cellular metabolism. A list of ‘the 30 damage-prone endogenous metabolites’ was drawn up based on ‘lability, centrality, and prevalence’ using (bio)chemical literature, cheminformatics and genome-scale metabolic models [7]. Here, in addition to well-known metabolites such as triose phosphates, erythrose 4-phosphate and cofactors (e.g. NAD(P)H, thiamin diphosphate, riboflavin intermediates, tetrahydrofolates and carbamoyl phosphates) also previously underestimated metabolites such as cysteine, homocysteine, glutathione, and glutamine could be identified [7]. Furthermore, most enzymes are promiscuous and catalyse the conversion of alternative substrates at low efficiency (< 1%) referred to as ‘substrate promiscuity’ or they catalyse alternative ‘incorrect’ reactions in addition to their physiological conversions, called ‘catalytic promiscuity’ (e.g. hydration instead of dehydrogenation) [8]. In addition, stress conditions such as heat or oxidative stress are other causal factors that lead to metabolite and enzyme damage and thus promote the formation of by-products [5].

The damaged metabolites and side products formed either enzymatically or non-enzymatically can be quite diverse and can act as effectors of other enzymes or they might be wasteful, harmful or even toxic to the cell. Depending on the organism, such metabolites can be

part of the classical metabolic pathways ‘canonical metabolites’ or they can represent ‘non-canonical’ metabolites, that is, metabolic outsiders not connected to the respective cell metabolism. This distinction does not always seem easy, as the metabolism of the organisms is different and often not fully understood [8]. However, cells have evolved different sophisticated mitigation systems/mechanisms including metabolite repair via enzymes and salvage pathways or pre-emption of damaged metabolites, for example, by degradation or export (overflow metabolism) [2,3,8–10]. Hence, it is assumed that many unidentified compounds in metabolomics profiles can be traced back to damaged metabolites and many genes of unknown function in genomes encode proteins dedicated to metabolite damage-control systems, called the ‘catalytic dark matter’ [11]. For example, for glycolysis with 10 enzymes, so far 10 repair enzymes have been identified in humans and mice [8]. In *Escherichia coli* as a whole, over 45 enzymes are associated with metabolite damage-control/repair, and novel metabolite repair enzymes are predicted to make up a significant proportion of ‘unknowns’, which represent an estimated 15% of the genes of *E. coli* [10]. Hence, more than 260 underground reactions are suggested in *E. coli* which corresponds to 10% of metabolic reactions [12].

In the cell, damaged metabolites and side products can accumulate to substantial levels over time, and inhibit various metabolic enzymes competitively or damage proteins, lipids or nucleic acids by covalent interaction/modification. Consistently, there is increasing evidence that malfunction in metabolite repair is related to diseases such as diabetes, cancer, and aging in humans and higher vertebrates [8,13]. However, the identification of these dangerous side-products and their repair systems is not straight forward, due to their low concentration in healthy cells (with active repair systems) and problems in structure identification. Accordingly, although cofactors are unstable and typically considered to be the ‘Achilles’ heel of metabolism’ [10] longevity of major cofactors with half-lives coinciding with the doubling time of the organism has been demonstrated in *E. coli*, *Bacillus subtilis*, and *Saccharomyces cerevisiae* [14]. These findings support the presence of very effective cellular repair systems.

Underground metabolism in hyperthermophilic Archaea

Archaea were recognised as a third domain of life only in the 1970s [15]. They were initially regarded as extremophiles that inhabit hostile environments such as salt lakes, hot springs or the famous black smokers in the deep sea. However, today we know that Archaea are widely distributed in various ecological niches, and even colonise the human skin and intestines [16,17].

Archaea have a unique metabolism that is characterised by new metabolic pathways such as methanogenesis and

the reversed ribulose monophosphate (RuMP) pathway, which in many Archaea replaces the oxidative pentose phosphate pathway. Even the canonical metabolic routes like the Embden–Meyerhof–Parnas (EMP) and the Entner–Doudoroff (ED) pathway are modified in terms of their pathway topology (e.g. branched ED pathway) as well as their enzymology employing novel enzymes from different enzyme families. These modifications also cause significant changes in the regulation of the archaeal metabolic pathways (for review see [18]). Remarkably, despite these modifications most of the metabolites remain conserved in the modified archaeal metabolic pathways and correspond to those of the bacteria and eukaryotes. In addition, enzyme promiscuity often with broad substrate specificity has been reported for many archaeal enzymes [19]. However, these studies have so far not been conducted in the context of underground metabolism.

Adaptation to extremely high temperatures is particularly common in Archaea and organisms with growth up to 113 °C (*Pyrobolus fumarii*, growth optimum 106 °C) have been described [20]. Initial studies in the field focused on the stability of macromolecules such as proteins, DNA, and membranes, whereas the stability of metabolites was greatly neglected. Since the velocity of chemical reactions increase exponentially with temperature (according to the Arrhenius equation), temperature has a dramatic effect on metabolite stability/damage [3] (Table 1). This means that underground metabolism becomes increasingly important for organisms thriving under elevated temperature conditions. Thus, the fundamental question arises how (hyper)

Table 1

Half-life times of phospho-intermediates from the common lower shunt of the EMP pathway determined at different temperatures.

Temperature	Half-life (in min)				
	37 °C	65 °C	70 °C	80 °C	88 °C
GAP	n.d.	6.9	4.0/12.4 ^a	1.63	0.8
DHAP	n.d.	25.4	18.9/30.8 ^a	5.3	3.8
1,3BPG	27 ^a	n.d.	0.9 ^b	n.d.	n.d.
3PG	n.d.	n.d.	n.d.	>2 h	n.d.
2PG	n.d.	n.d.	n.d.	>2 h	n.d.
PEP	n.d.	210	89	22	11

Abbreviations of intermediates: GAP, glyceraldehyde 3-phosphate; DHAP, dihydroxyacetone phosphate; 1,3BPG, 1,3-bisphosphoglycerate; 3PG, 3-phosphoglycerate; 2PG, 2-phosphoglycerate; PEP, phosphoenolpyruvate.

The half-life time of intermediates was determined in 100 mM HEPES/KOH buffer (pH 6.5) [40] (Siebers et al. unpublished data) with few exceptions.

^a Half-life times determined in 0.1 M Tris/HCl, 20 mM MgCl₂ [37,38].

^b [3].

thermophilic Archaea deal with promiscuous enzymes and increasingly unstable metabolites in order to enable metabolism and thus life at temperatures from 60 to 116 °C. Do non-enzymatic reactions play a major role in their metabolic pathways [21] or do they possess effective repair or pre-emption strategies or other metabolic adaptations?

Cofactor salvage pathways in Archaea

Despite the essential importance, so far only few studies address this question in (hyper)thermophilic Archaea. The four-step salvage pathway for NAD⁺ [22] and the methionine salvage pathway for S-adenosyl-L-methionine (SAM) [23] have been described.

NAD⁺ is characterised by low (thermo)stability (half life time ($t_{1/2}$) of 24.2 min at 85 °C) and decomposes to ADP-ribose and nicotinamide (Nm) [22], requiring organisms to have effective salvage pathways for NAD⁺. While the four-step pathway (Preiss–Handler pathway) from Nm via nicotinic acid mononucleotide and nicotinic acid adenine dinucleotide to NAD⁺ is more common in bacteria, eukaryotes have an additional two-step pathway that converts Nm via Nm mononucleotide to NAD⁺. Studies in the hyperthermophilic Archaeon *Thermococcus kodakarensis* with enzyme characterisation, determination of intermediate thermostabilities and genetic analyses by Atomi *et al.* revealed the utilisation of the four-step pathway. Most intermediates of both salvage pathways such as ADP-ribose, Nm and nicotinic acid are very thermostable. However, since Nm mononucleotide ($t_{1/2}$ 17.1 min at 85 °C) from the two-step pathway is much more thermolabile than the corresponding nicotinic acid mononucleotide ($t_{1/2}$ 431 min at 85 °C) from the four-step pathway, they conclude that an evolution of metabolism has taken place to avoid thermolabile intermediates.

In the hyperthermophilic methanogen *Methanocaldococcus jannaschii*, the promiscuity of enzymes for the regeneration of SAM, an essential cofactor in all three domains of life, was reported by White *et al.* [23]. The enzyme characterisation of methylthioinosine phosphorylase (MTIP) and methylthioribose 1-phosphate isomerase (MTRI) are in agreement with their function in the modified anaerobic methionine salvage pathway. Both enzymes are promiscuous. Thus, MTIP was shown to have broad purine nucleoside phosphorylase activity and a possible additional function in purine salvage was proposed. MTRI was shown to have activity with 5-deoxyribose 1-phosphate (in addition to 5-methylthioribose 1-phosphate) forming 5-deoxyribulose 1-phosphate an important precursor for the synthesis of 6-deoxy-5-ketofructose 1-phosphate (DKFP) for the synthesis of aromatic amino acids.

Maillard reactions with methylglyoxal formation at high temperature: A case study

The Maillard reaction is a process occurring at moderate temperatures and is further accelerated by an increase in temperature [24]. In the Maillard cascade reducing carbohydrates such as glucose and fructose, phosphoester intermediates (e.g. glucose-6 phosphate and triose-3 phosphates) and glyoxals arising from their decomposition react spontaneously by non-enzymatic glycation with nucleophilic amines, which is particularly severe for heterotrophic organisms. In the final stage of this process a complex product mixture of Maillard products so-called advanced glycation end-products (AGE's) are formed which also include, for example, acrylamide well known from thermal food processing (60–200 °C) [25]. These AGEs are deleterious by functional impairment of proteins, nucleic acids, and lipids and are closely related with disease and aging [24,26].

The Maillard reaction has been demonstrated in hyperthermophilic Archaea by growth inhibition in presence of amino acids or peptides and monosaccharides in the culture medium for *Pyrococcus furiosus* at 100 °C [27] and for *Aeropyrum pernix* at 90 °C [28]. Further on, enzyme inactivation by carbohydrates at high temperatures (80 °C) was shown for the β -glycosidase from *P. furiosus* [29]. As in Bacteria and Eukaryotes, a DJ-1 Maillard deglycase was described in *P. furiosus* as important repair enzyme, which degrades Maillard adducts and prevents glycation damage of proteins [25]. However, these studies are controversial due to lack of *in vivo* function [8]. For other repair enzymes like NTP pyrophosphatases and lactonases (for review see [30]) so far only their phylogenetic distribution was addressed.

The best studied glycation agent in the central metabolic pathways is methylglyoxal (MG) causing so-called dicarbonyl stress. MG is formed either non-enzymatically as a spontaneous decomposition product of unstable triosephosphates, that is, dihydroxyacetone phosphate (DHAP) and glyceraldehyde 3-phosphate (GAP) [31,32], or enzymatically by side activity of the triosephosphate isomerase (TPI) and by the dedicated enzyme methylglyoxal synthase [8,33]. MG toxicity poses special challenges even under mesophilic conditions, for example, for plants, vertebrates and Bacteria, and different salvage pathways and its role as signaling molecule under stress conditions have been reported [8,33–35]. At high temperature, the half life time of triosephosphates is significantly reduced (Table 1) and MG represents a major thermal decomposition product. The metabolic challenge by metabolite instability at high temperature has been further corroborated by biochemical/enzymatic as well as in combined modelling and experimental approaches [31,36–40]. These studies revealed three major adaptation

mechanisms: (i) changes in pathway topology, (ii) the introduction of novel enzymes, and (iii) integration of damaged metabolites into novel pathways.

Novel topologies and promiscuity of pathways in (hyper)thermophilic Archaea

Investigations of the central carbohydrate metabolism in Archaea revealed significant changes in the ED pathway topology including the lower common EMP pathway [18]. For example, aerobic thermoacidophilic Archaea (exposed to oxidative stress) such as *Saccharolobus* (former *Sulfolobus*) *sofataricus* rely on a modified branched ED pathway (Figure 1) for glucose degradation. In contrast to the classical ED pathway, no phosphorylation takes place at the beginning of the pathway and 2-keto-3-deoxygluconate (KDG) is formed as the characteristic intermediate. In the semi-phosphorylative (sp) ED branch, KDG is phosphorylated by KDG kinase (KDGK) to KD(P)G, the characteristic intermediate in the classical ED pathway, and GAP and pyruvate are formed by KD(P)G aldolase. In the non-phosphorylative (np) ED branch KDG is directly cleaved by the same KD(P)G aldolase to pyruvate and glyceraldehyde, and the latter is funnelled as 2PG into the common lower EMP branch. The available KDGK mutant in *Sulfolobus sofataricus* showed that the spED pathway is not essential for growth on glucose and points to gluconeogenic function, whereas the npED branch fulfils a glycolytic function [36]. Notably, this branched pathway topology allows to circumvent the formation of thermolabile intermediates, that is, GAP, DHAP, BPG (npED) in glycolysis and of BPG (spED) in gluconeogenesis (Figure 1). In a systems biology study, a mathematical model for the branched ED was used to show that maintaining the two ED branches despite additional costs substantially contributes to pyruvate production and provides robustness to internal perturbations due to temperature changes [40]. This increased pathway flexibility/dynamics and robustness might be of special importance for organisms facing heat and oxidative stress [5].

Furthermore, it could be shown for *S. sofataricus* that the branched ED pathway is promiscuous and can be used for the degradation of D-galactose in addition to D-glucose [18]. Some of the enzymes such as glucose dehydrogenase and KD(P)G aldolase are even involved in promiscuous pentose degradation pathways, that is, the Dahms and Weimberg pathway for L-arabinose and D-xylose degradation [18,41]. These findings support the theory that enzyme promiscuity plays an important role in the evolution of metabolic pathways in organisms [12].

Novel, unusual enzymes in (hyper)thermophilic Archaea

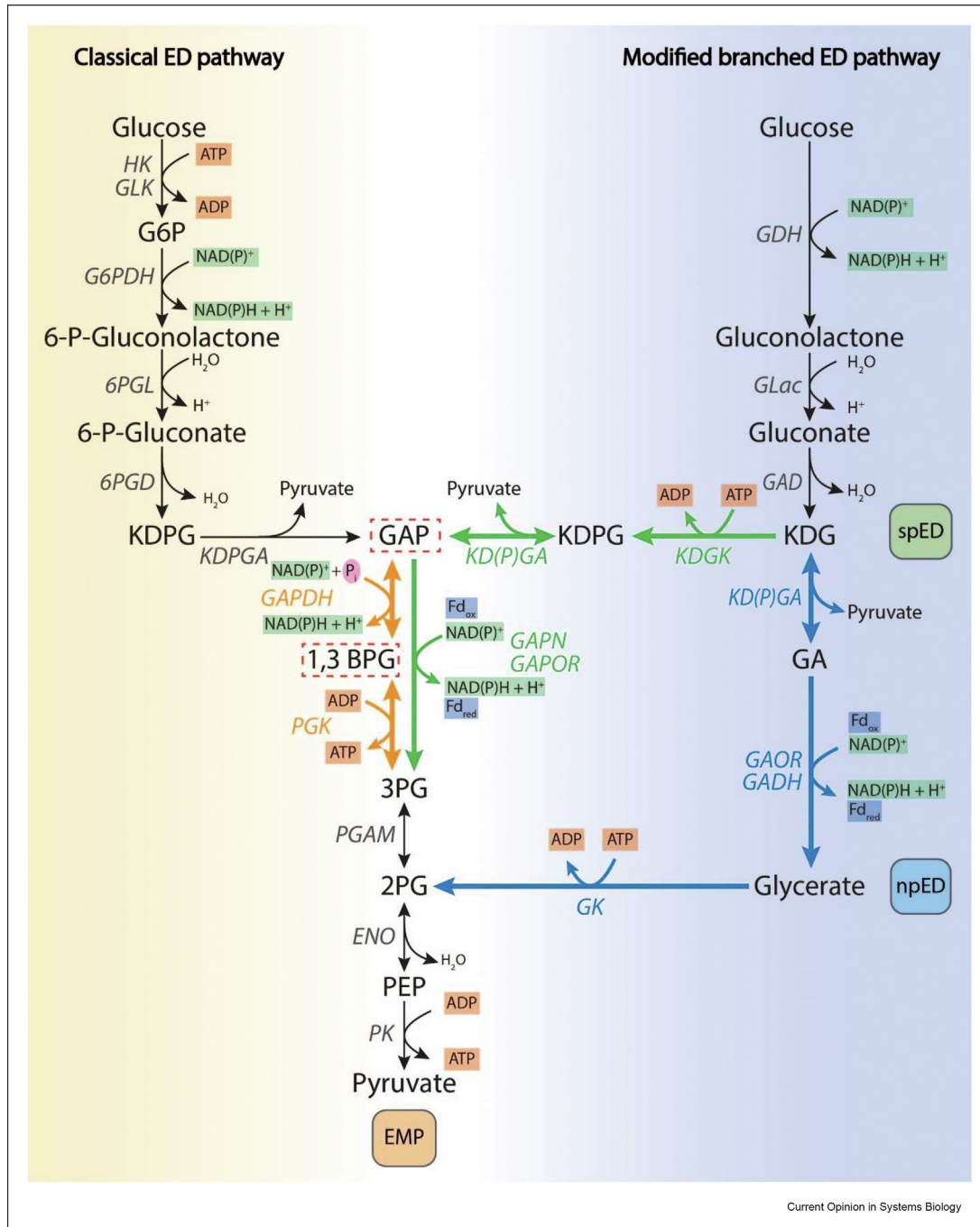
With the first discovery of the non-phosphorylating glyceraldehyde-3-phosphate (GAP) dehydrogenase (GAPN) in the hyperthermophilic Archaeon *Thermoproteus tenax* a

special role in the so-called ‘metabolic thermoadaptation’ was proposed by Reinhard Hensel [42]. GAPN like the GAP oxidoreductase (GAPOR), which is additionally found in anaerobic hyperthermophiles, catalyse the unidirectional oxidation of GAP to 3-phosphoglycerate (3PG) using NAD(P)^+ and oxidised ferredoxin (Fd_{ox}) as co-factor, respectively. Allosteric GAPN is activated by glucose 1-phosphate, an intermediate in glycogen metabolism and polymer degradation, enabling an enhanced carbon flux under glycolytic conditions. As shown for several (hyper)thermophilic Archaea (i.e. *P. furiosus*, *T. kodakarensis*, *T. tenax*, and *Sulfolobus/Saccharolobus* spp. (for review see [18]) the catabolic/glycolytic GAPN replaces the classical, reversible GAP dehydrogenase (GAPDH) and phosphoglycerate kinase (PGK) enzyme couple that only has a restricted anabolic/gluconeogenic function in these organisms. Notably, the GAP oxidation via GAPN and GAPOR is not coupled to ATP gain via substrate level phosphorylation as in the PGK reaction (Figure 2). Instead, GAPN and GAPOR circumvent the formation of the extremely thermolabile 1,3-bisphosphoglycerate (BPG), which spontaneously dephosphorylates to 3PG at high temperature. A second unusual enzyme found in (hyper)thermophilic Archaea and Bacteria is the gluconeogenic, bifunctional fructose-1,6-bisphosphate aldolase/phosphatase (FBPA/ase) [43]. The enzyme catalyses the unidirectional conversion of thermolabile GAP and DHAP to stable fructose 6-phosphate.

The effect of intermediate instability of triosephosphates (i.e. GAP, DHAP, and BPG) at high temperature was also confirmed in an *in vitro* reconstituted gluconeogenic enzyme cascade [37]. In this study a detailed mathematical model was constructed for the conversion of 3PG to F6P at 70 °C by *S. sofataricus* PGK, GAPDH, triosephosphate isomerase (TPI), FBPA/ase. The model not only included the enzyme kinetics for the reactions but also half-life-times for the intermediates. Low conversion efficiency (50%–60% of carbon in 3PG was retrieved as F6P) were observed, which could be related to the temperature-dependent degradation of GAP, DHAP, and BPG resulting in the carbon loss by formation of methylglyoxal [31]. Moreover, comparison of the yeast and *S. sofataricus* PGK and GAPDH reaction *in vitro* at 30 °C and 70 °C, respectively, using a combined modelling and experimental approach revealed futile cycling induced by the PGK leading to net ATP hydrolysis due to thermal instability of BPG (degradation to 3PG) in the thermophile, but not in yeast [38]. Addition of GAPDH at 70 °C reduced the futile cycle, by lowering the BPG concentration via rapid conversion of BPG to GAP, but in contrast to the yeast system at 30 °C only low concentration of product was formed due to the thermal instability of GAP (Figure 3).

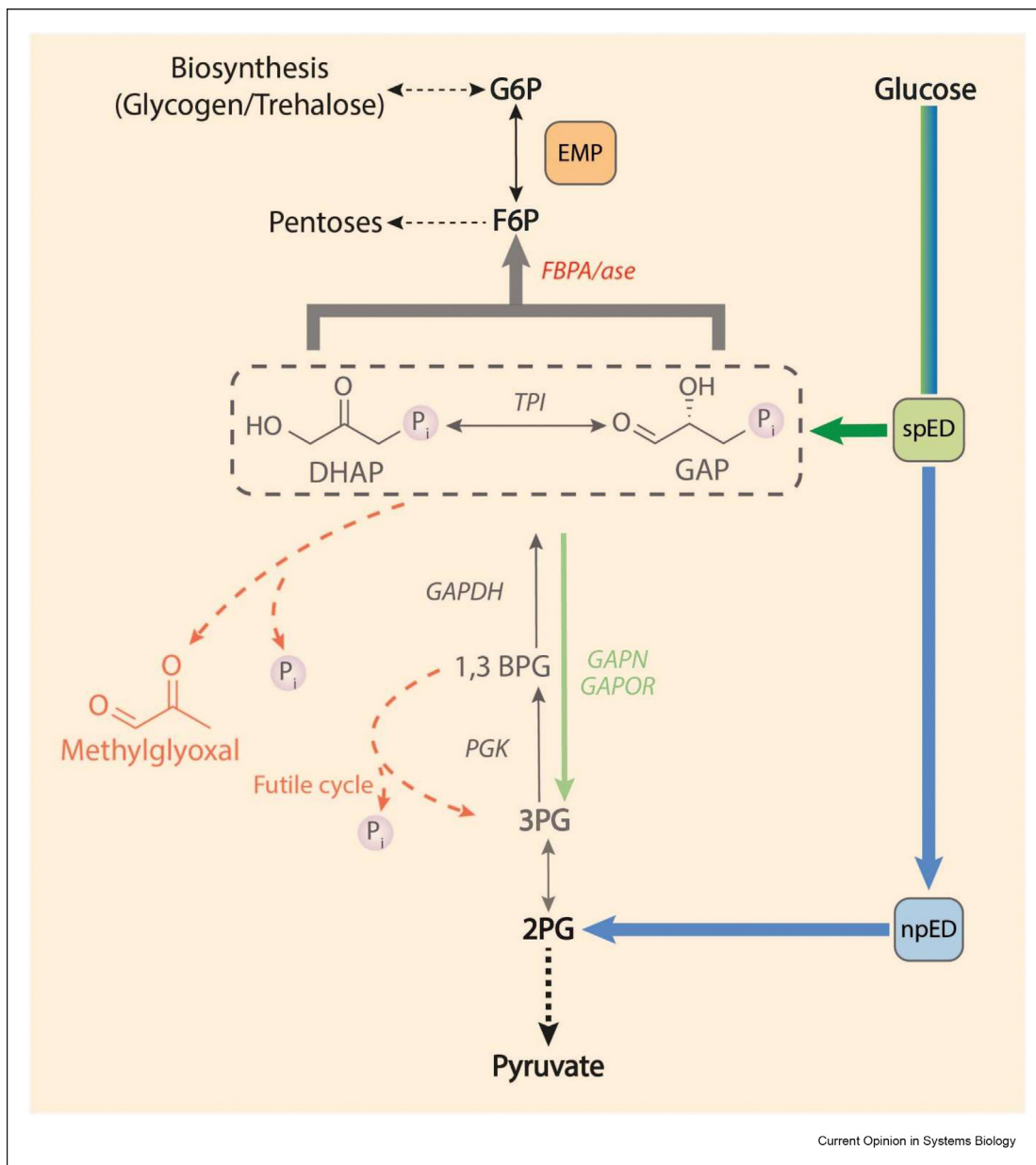
Therefore, the instability of triosephosphates causes carbon and energy loss via toxic methylglyoxal formation

Figure 1



The topology of the modified branched Entner–Doudoroff (ED) pathway allows to circumvent the formation of thermolabile intermediates. The classical bacterial ED pathway as well as the semi-phosphorylative (sp) and non-phosphorylative (npED) branch of the modified ED found in Archaea are depicted. Both pathways funnel into the lower common Embden–Meyerhof–Parnas (EMP) pathway. The thermolabile intermediates GAP and BPG are framed by red dashed lines. Enzyme abbreviations: ENO, enolase; G6PDH (ZWF), glucose-6-phosphate dehydrogenase; GAD, gluconate dehydratase; GADH, glyceraldehyde dehydrogenase; GAOR, glyceraldehyde:ferredoxin oxidoreductase; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; GAPN, non-phosphorylating GAPDH; GAPOR, GAP oxidoreductase; GDH, glucose dehydrogenase; GK, glycerate kinase; GLac, gluconolactonase; GLK, glucose kinase; HK, hexokinase; KDGK, 2-keto-3-deoxygluconate kinase; KDPGA (EDA), 2-keto-3-deoxy-6-phosphogluconate aldolase; KD(P)GA, bifunctional 2-keto-3-deoxy-(6-phospho)gluconate aldolase; PGK, phosphoglycerate kinase; PGAM, phosphoglycerate mutase; PK, pyruvate kinase; 6PGD (EDD), gluconate-6-phosphate dehydratase; 6PGL, 6-phosphoglucono-1,4-lactonase. Abbreviations of intermediates: G6P, glucose 6-phosphate; GAP, glyceraldehyde 3-phosphate; 1,3BPG, 1,3-bisphosphoglycerate; 3PG, 3-phosphoglycerate; 2PG, 2-phosphoglycerate; PEP, phosphoenolpyruvate; KDG, 2-keto-3-deoxygluconate; KDPG, 2-keto-3-deoxy-6-phosphogluconate; GA, glyceraldehyde.

Figure 2



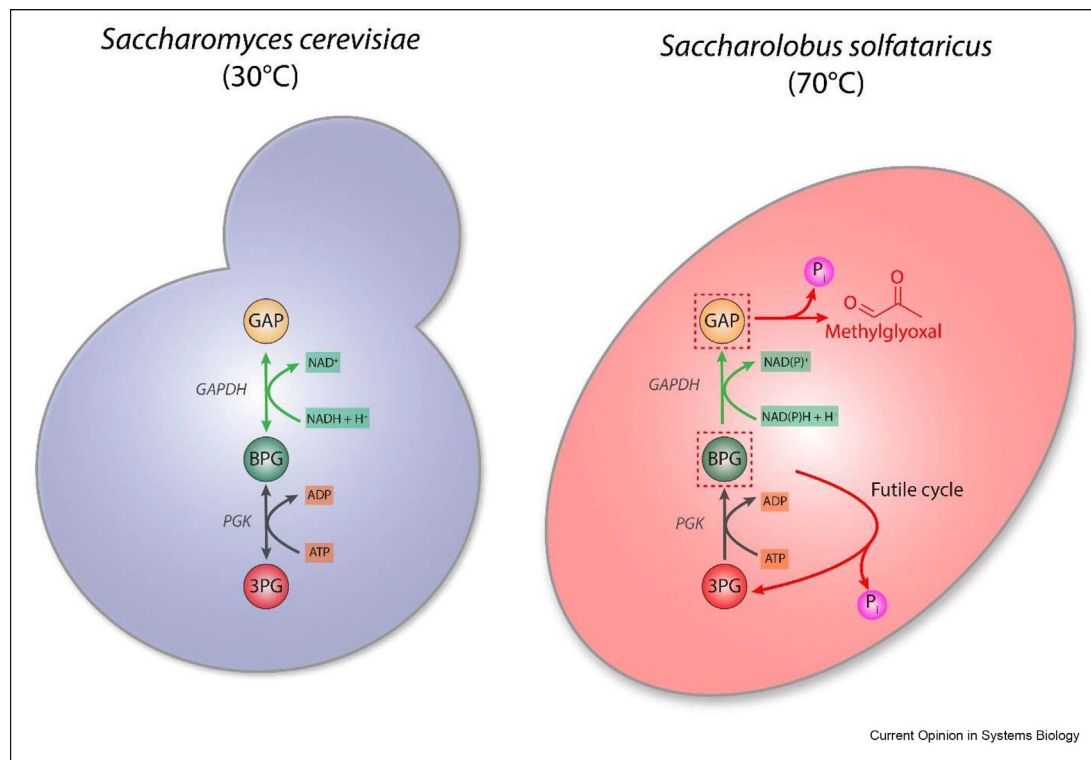
The introduction of novel enzymes into the EMP pathway of (hyper)thermophilic Archaea pre-empts thermolabile intermediate pools and supports the glycolytic and/or gluconeogenic flux. For abbreviations see figure legend 1. Additional abbreviations: FBPA/ase, bifunctional fructose-1,6-bisphosphate aldolase/phosphatase; G6P, glucose 6-phosphate; F6P, fructose 6-phosphate; DHAP, dihydroxyacetone phosphate.

(GAP, DHAP) and futile cycling (BPG). The introduction of novel enzymes, that is, glycolytic GAPN/GAPOR and gluconeogenic bifunctional FBPA/ase represents a pre-empting strategy to reduce thermolabile metabolite pools in hyperthermophilic Archaea and to support the glycolytic and gluconeogenic flux.

Integration of damaged/toxic metabolites into novel metabolic pathways

First evidence for the integration of MG into novel metabolic pathways came from studies in *M. jannaschii* with optimal growth at 85 °C [44]. The authors determined a very high concentration of MG (70 μM) in

Figure 3



Futile cycling and carbon loss challenges life at high temperature. The effect of temperature on the gluconeogenic phosphoglycerate kinase (PGK) and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) reaction was studied in a combined modelling experimental approach [38]. First the gluconeogenic PGK reaction at 30 °C using the yeast PGK was studied and due to the chemical equilibrium ($\Delta G^{\circ} 18.8 \text{ kJ mol}^{-1}$) hardly any BPG was formed. Only in presence of an effective ATP recycling system the reaction reaches an equilibrium with equal rates of 3PG and BPG formed. The *Saccharolobus (Sulfolobus) solfataricus* (Sso) PGK assayed at 70 °C showed no BPG formation even in presence of ATP recycling due to the additional constraint of thermal degradation of BPG (half life time of 56.5 s at 70 °C [38]) to 3PG leading to a futile cycling of ATP. Coupling of PGK and GAPDH activities could in both systems overcome the thermodynamic constraint and futile cycling of ATP in presence of NAD(P)⁺ recycling. However, whereas the yeast enzymes at 30 °C allowed for complete conversion of 3PG to GAP only very low GAP concentrations were observed in the Sso system at 70 °C due to the thermal instability of GAP (half life time of 12.4 min at 70 °C [37]) resulting in methylglyoxal formation and thus carbon loss.

crude extracts, which is at least 30-fold higher than in other mesophilic organisms (0.1–2 μM). In addition, they showed that MG is not just detoxified but is an essential intermediate for the biosynthesis of aromatic amino acids via 6-deoxy-5-ketofructose-1-phosphate (DKFP) and lactate for F420 biosynthesis [44–46]. The enzyme involved in methylglyoxal reduction was later on characterised [46]. Therefore, the introduction of novel pathways or pathway modifications seem to enable (hyper)thermophilic Archaea to pre-empt or reduce toxic metabolite accumulation by including them into their metabolic network.

Conclusion

(Hyper)thermophiles with a metabolic network challenged by increased metabolite instability are the perfect model organisms to study underground metabolism [10]. (Hyper)thermophilic Archaea are particularly promising here because they possess an unusual metabolic network with promiscuous enzymes and pathways.

They dominate high-temperature habitats and some representatives are established as model organisms. For several species such as *T. kodakarensis*, *P. furiosus*, *S. solfataricus*, *Sulfolobus acidocaldarius* genetic systems have been established, their central metabolism is well studied and polyomic/functional genomic data are available [18]. Notably, genome-scale metabolic modelling of, for example, *S. solfataricus* revealed that 16% of known cellular reactions had to be included without genome evidence [47] suggesting that other enzymes or pathways catalyse these reactions, or that non-catalysed reactions have replaced the enzyme's role [2]. Furthermore, the current work on central carbohydrate metabolism and salvage pathways in (hyper)thermophilic Archaea points towards a pre-emption strategy that decreases the concentration of unstable intermediates by novel pathway topologies and introduction of novel enzymes, as well as of toxic decomposition products (i.e. MG) via integration as canonical intermediates in novel biosynthesis pathways. These

findings suggest a significant influence of thermolabile intermediates as well as of promiscuous enzymes and pathways on the evolution of the metabolic network in (hyper)thermophilic Archaea.

Because of their unique metabolism (hyper)thermophilic Archaea will enable the discovery of new underground metabolites and conquer new principles in metabolic pathway design and topology as well as novel repair enzymes and salvage strategies. The gained knowledge on underground metabolism will be of particular importance for pathway development and evolution for biotechnological applications [4,12,48].

Conflict of interest statement

Nothing declared.

Acknowledgements

C.S., C.B. and B.S. acknowledge funding by the VW Stiftung in the “Life?” initiative (96725). C.S., T.K., C.B., J.S. and B.S. acknowledge funding by the Federal Ministry of Education and Research (BMBF, Germany) grants HotSysAPP (031L0078A) and HotAcidFACTORY (031B0848A). J.S. acknowledges funding from the DST/NRF (South Africa), particularly for funding the SARCHI initiative (NRF-SARCHI-82813).

References

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest

1. D’Ari R, Casadesús J: **Underground metabolism**. *Bioessays* 1998, **20**:181–186.
 2. Keller MA, Piedrafita G, Ralser M: **The widespread role of non-enzymatic reactions in cellular metabolism**. *Curr Opin Biotechnol* 2015, **34**:153–161.
 3. Hanson AD, Henry CS, Fiehn O, de Crécy-Lagard V: **Metabolite damage and metabolite damage control in plants**. *Annu Rev Plant Biol* 2016, **67**:131–152.
 4. Rosenberg J, Commichau FM: **Harnessing underground metabolism for pathway development**. *Trends Biotechnol* 2019, **37**:29–37.
- The review highlights the importance and potential of underground metabolism and especially promiscuous enzymes for novel pathway developments in metabolic engineering approaches.
5. Piedrafita G, Keller MA, Ralser M: **The impact of non-enzymatic reactions and enzyme promiscuity on cellular metabolism during (oxidative) stress conditions**. *Biomolecules* 2015, **5**: 2101–2122.
 6. Golubev A: **How could the Gompertz-Makeham law evolve**. *J Theor Biol* 2009, **258**:1–17.
 7. Lerma-Ortiz C, Jeffryes JG, Cooper AJ, Niehaus TD, Thamm AM, Frelin O, Aunins T, Fiehn O, de Crécy-Lagard V, Henry CS, et al.: **Nothing of chemistry disappears in biology’: the Top 30 damage-prone endogenous metabolites**. *Biochem Soc Trans* 2016, **44**:961–971.
 8. Bommer GT, Van Schaftingen E, Veiga-da-Cunha M: **Metabolite repair enzymes control metabolic damage in glycolysis**. *Trends Biochem Sci* 2020, **45**:228–243.

This review gives an excellent overview of the repair enzymes involved in glycolysis. The respective reactions and repair enzymes as well as the effect of deficiency on disease in humans and survival in higher vertebrates is discussed.

9. Linster CL, Van Schaftingen E, Hanson AD: **Metabolite damage and its repair or pre-emption**. *Nat Chem Biol* 2013, **9**:72–80.
 10. de Crécy-Lagard V, Haas D, Hanson AD: **Newly-discovered enzymes that function in metabolite damage-control**. *Curr Opin Chem Biol* 2018, **47**:101–108.
- The challenge by genes/enzymes of unknown function in the post-genome area and examples of newly discovered repair enzymes and damage-control systems are discussed.
11. Ellens KW, Christian N, Singh C, Satagopam VP, May P, Linster CL: **Confronting the catalytic dark matter encoded by sequenced genomes**. *Nucleic Acids Res* 2017, **45**:11495–11514.
 12. Notebaart RA, Kintsjes B, Feist AM, Papp B: **Underground metabolism: network-level perspective and biotechnological potential**. *Curr Opin Biotechnol* 2018, **49**:108–114.
- The authors describe systems biology advances to map and identify ‘underground’ activities (‘the enzyme promiscuome’) and highlight its potential to enable evolutionary adaptation to new environments as well as the design of novel (‘patchwork’) pathways in biotechnology.
13. Veiga-da-Cunha M, Van Schaftingen E, Bommer GT: **Inborn errors of metabolite repair**. *J Inherit Metab Dis* 2020, **43**: 14–24.
 14. Hartl J, Kiefer P, Meyer F, Vorholt JA: **Longevity of major co-enzymes allows minimal de novo synthesis in microorganisms**. *Nat Microbiol* 2017, **2**:17073.
 15. Woese CR, Fox GE: **Phylogenetic structure of the prokaryotic domain: the primary kingdoms**. *Proc Natl Acad Sci U S A* 1977, **74**:5088–5090.
 16. Offre P, Spang A, Schleper C: **Archaea in biogeochemical cycles**. *Annu Rev Microbiol* 2013, **67**:437–457.
 17. Borrel G, Brugère JF, Gribaldo S, Schmitz RA, Moissl-Eichinger C: **The host-associated archaeome**. *Nat Rev Microbiol* 2020, **18**:622–636.
 18. Bräsen C, Esser D, Rauch B, Siebers B: **Carbohydrate metabolism in Archaea: current insights into unusual enzymes and pathways and their regulation**. *Microbiol Mol Biol Rev* 2014, **78**:89–175.
 19. Jia B, Cheong GW, Zhang S: **Multifunctional enzymes in archaea: promiscuity and moonlight**. *Extremophiles* 2013, **17**: 193–203.
 20. Blöchl E, Rachel R, Burggraf S, Hafenbradl D, Jannasch HW, Stetter KO: **Pyrolobus fumarii, gen. and sp. nov., represents a novel group of archaea, extending the upper temperature limit for life to 113 degrees C**. *Extremophiles* 1997, **1**:14–21.
 21. Ralser M: **An appeal to magic? The discovery of a non-enzymatic metabolism and its role in the origins of life**. *Biochem J* 2018, **475**:2577–2592.
 22. Hachisuka SI, Sato T, Atomi H: **Hyperthermophilic Archaeon Thermococcus kodakarensis utilizes a four-step pathway for NAD(+) salvage through nicotinamide deamination**. *J Bacteriol* 2018, **200**.
- This is one of the few studies in hyperthermophilic Archaea where salvaging pathways are addressed. The author presents a novel four-step pathway for NAD(+) salvaging that in contrast to the common two-step pathway omits the thermolabile nicotinamide.
23. Miller DV, Rauch BJ, Harich K, Xu H, Perona JJ, White RH: **Promiscuity of methionine salvage pathway enzymes in Methanocaldococcus jannaschii**. *Microbiology (Read)* 2018, **164**:969–981.
 24. Szwergold BS: **Maillard reactions in hyperthermophilic archaea: implications for better understanding of non-enzymatic glycation in biology**. *Rejuvenation Res* 2013, **16**: 259–272.
 25. Richarme G, Marguet E, Forterre P, Ishino S, Ishino Y: **DJ-1 family Maillard deglycases prevent acrylamide formation**. *Biochem Biophys Res Commun* 2016, **478**:1111–1116.

26. Scheckhuber CQ: **Studying the mechanisms and targets of glycation and advanced glycation end-products in simple eukaryotic model systems.** *Int J Biol Macromol* 2019, **127**:85–94.
In this review the author focusses on the relevance, mode of action and molecular targets of reactive carbonyl species in model systems from fungi, plants, the nematode *Caenorhabditis elegans* and the fruit fly *Drosophila melanogaster*.
27. Driskill LE, Kusy K, Bauer MW, Kelly RM: **Relationship between glycosyl hydrolase inventory and growth physiology of the hyperthermophile *Pyrococcus furiosus* on carbohydrate-based media.** *Appl Environ Microbiol* 1999, **65**: 893–897.
28. Kim KW, Lee SB: **Inhibitory effect of Maillard reaction products on growth of the aerobic marine hyperthermophilic archaeon *Aeropyrum pernix*.** *Appl Environ Microbiol* 2003, **69**: 4325–4328.
29. Bruins ME, Van Hellemond EW, Janssen AE, Boom RM: **Maillard reactions and increased enzyme inactivation during oligosaccharide synthesis by a hyperthermophilic glycosidase.** *Biotechnol Bioeng* 2003, **81**:546–552.
30. Galperin MY, Moroz OV, Wilson KS, Murzin AG: **House cleaning, a part of good housekeeping.** *Mol Microbiol* 2006, **59**:5–19.
31. Hagedoorn P-L: **Steady-state kinetics of the tungsten containing aldehyde: ferredoxin oxidoreductases from the hyperthermophilic archaeon *Pyrococcus furiosus*.** *J Biotechnol* 2019, **306**:142–148.
32. Gauss D, Schoenenberger B, Wohlgemuth R: **Chemical and enzymatic methodologies for the synthesis of enantiomerically pure glyceraldehyde 3-phosphates.** *Carbohydr Res* 2014, **389**:18–24.
33. Booth IR, Ferguson GP, Miller S, Li C, Gunasekera B, Kinghorn S: **Bacterial production of methylglyoxal: a survival strategy or death by misadventure?** *Biochem Soc Trans* 2003, **31**: 1406–1408.
34. Booth IR: **Glycerol and methylglyoxal metabolism.** *EcoSal Plus* 2005, **1**.
35. Hoque TS, Hossain MA, Mostofa MG, Burritt DJ, Fujita M, Tran LS: **Methylglyoxal: an emerging signaling molecule in plant abiotic stress responses and tolerance.** *Front Plant Sci* 2016, **7**:1341.
36. Kouril T, Wieloch P, Reimann J, Wagner M, Zaparty M, Albers SV, Schomburg D, Ruoff P, Siebers B: **Unraveling the function of the two Entner-Doudoroff branches in the thermoacidophilic Crenarchaeon *Sulfolobus solfataricus* P2.** *FEBS J* 2013, **280**: 1126–1138.
37. Kouril T, Esser D, Kort J, Westerhoff HV, Siebers B, Snoep JL: **Intermediate instability at high temperature leads to low pathway efficiency for an in vitro reconstituted system of gluconeogenesis in *Sulfolobus solfataricus*.** *FEBS J* 2013, **280**:4666–4680.
38. Kouril T, Eicher JJ, Siebers B, Snoep JL: **Phosphoglycerate kinase acts as a futile cycle at high temperature.** *Microbiology* 2017, **163**:1604–1612.
This combined modelling-experimental approach highlights the challenge of life at high temperature resulting in carbon loss and futile cycling due to the instability of triosephosphates.
39. Zhang Y, Kouril T, Snoep JL, Siebers B, Barberis M, Westerhoff HV: **The peculiar glycolytic pathway in hyperthermophilic Archaea: understanding its whims by experimentation *in silico*.** *Int J Mol Sci* 2017, **18**.
40. Figueiredo AS, Kouril T, Esser D, Haferkamp P, Wieloch P, Schomburg D, Ruoff P, Siebers B, Schaber J: **Systems biology of the modified branched Entner-Doudoroff pathway in *Sulfolobus solfataricus*.** *PLoS One* 2017, **12**:e0180331.
41. Wagner M, Shen L, Albersmeier A, van der Kolk N, Kim S, Cha J, Bräsen C, Kalinowski J, Siebers B, Albers S-V: ***Sulfolobus acidocaldarius* transports pentoses via a carbohydrate uptake transporter 2 (CUT2)-type ABC transporter and metabolizes them through the aldolase-independent Weimberg pathway.** *Appl Environ Microbiol* 2018, **84**:e01273. 01217.
42. Brunner NA, Siebers B, Hensel R: **Role of two different glyceraldehyde-3-phosphate dehydrogenases in controlling the reversible Embden–Meyerhof–Parnas pathway in *Thermoproteus tenax*: regulation on protein and transcript level.** *Extremophiles* 2001, **5**:101–109.
43. Say RF, Fuchs G: **Fructose 1,6-bisphosphate aldolase/phosphatase may be an ancestral gluconeogenic enzyme.** *Nature* 2010, **464**:1077–1081.
44. White RH, Xu H: **Methylglyoxal is an intermediate in the biosynthesis of 6-deoxy-5-ketofructose-1-phosphate: a precursor for aromatic amino acid biosynthesis in *Methanocaldococcus jannaschii*.** *Biochemistry* 2006, **45**: 12366–12379.
45. Grochowski LL, Xu H, White RH: **Identification of lactaldehyde dehydrogenase in *Methanocaldococcus jannaschii* and its involvement in production of lactate for F420 biosynthesis.** *J Bacteriol* 2006, **188**:2836–2844.
46. Miller DV, Ruhlin M, Ray WK, Xu H, White RH: **N5,N10-methylenetetrahydromethanopterin reductase from *Methanocaldococcus jannaschii* also serves as a methylglyoxal reductase.** *FEBS Lett* 2017, **591**:2269–2278.
47. Ulas T, Riemer SA, Zaparty M, Siebers B, Schomburg D: **Genome-scale reconstruction and analysis of the metabolic network in the hyperthermophilic archaeon *Sulfolobus solfataricus*.** *PLoS One* 2012, **7**:e43401.
48. Guzmán GI, Sandberg TE, LaCroix RA, Nyerges Á, Papp H, de Raad M, King ZA, Hefner Y, Northen TR, Notebaart RA, *et al.*: **Enzyme promiscuity shapes adaptation to novel growth substrates.** *Mol Syst Biol* 2019, **15**:e8462.

**Chapter 3.4: Identification of fungal
lignocellulose-degrading biocatalysts
secreted by *Phanerochate chrysosporium*
via activity-based protein profiling**

Identification of fungal lignocellulose-degrading biocatalysts secreted by *Phanerochaete chrysosporium* via activity-based protein profiling

Christian Schmerling^{1,8}, Leonard Sewald^{2,8}, Geronimo Heilmann^{2,7,8}, Frederick Witfeld³, Dominik Begerow³, Kenneth Jensen⁴, Christopher Bräsen¹, Farnusch Kaschani^{2,5}, Herman S. Overkleef⁶, Bettina Siebers¹✉ & Markus Kaiser²✉

Activity-based protein profiling (ABPP) has emerged as a versatile biochemical method for studying enzyme activity under various physiological conditions, with applications so far mainly in biomedicine. Here, we show the potential of ABPP in the discovery of biocatalysts from the thermophilic and lignocellulose-degrading white rot fungus *Phanerochaete chrysosporium*. By employing a comparative ABPP-based functional screen, including a direct profiling of wood substrate-bound enzymes, we identify those lignocellulose-degrading carbohydrate esterase (CE1 and CE15) and glycoside hydrolase (GH3, GH5, GH16, GH17, GH18, GH25, GH30, GH74 and GH79) enzymes specifically active in presence of the substrate. As expression of fungal enzymes remains challenging, our ABPP-mediated approach represents a preselection procedure for focusing experimental efforts on the most promising biocatalysts. Furthermore, this approach may also allow the functional annotation of domains-of-unknown functions (DUFs). The ABPP-based biocatalyst screening described here may thus allow the identification of active enzymes in a process of interest and the elucidation of novel biocatalysts that share no sequence similarity to known counterparts.

¹Molecular Enzyme Technology and Biochemistry (MEB), Environmental Microbiology and Biotechnology (EMB), Centre for Water and Environmental Research (CWE), Faculty of Chemistry, University of Duisburg-Essen, Universitätsstraße 5, 45141 Essen, Germany. ²Department of Chemical Biology, ZMB, Faculty of Biology, University of Duisburg-Essen, Universitätsstraße 2, 45117 Essen, Germany. ³Evolution of Plants and Fungi, Ruhr-University Bochum, Universitätsstraße 150, 44780 Bochum, Germany. ⁴Novozymes, Biologiens vej 2, 2800 Kgs, Lyngby, Denmark. ⁵Analytics Core Facility Essen, ZMB, Faculty of Biology, University of Duisburg-Essen, Universitätsstraße 2, 45117 Essen, Germany. ⁶Department of Bio-organic Synthesis, Leiden Institute of Chemistry, Leiden University, Einsteinweg 55, 2333 CC Leiden, Netherlands. ⁷Present address: German Diabetes Center (DDZ), Leibniz Center for Diabetes Research at Heinrich Heine University Düsseldorf, Düsseldorf, Germany. ⁸These authors contributed equally: Christian Schmerling, Leonard Sewald, Geronimo Heilmann. ✉email: bettina.siebers@uni-due.de; markus.kaiser@uni-due.de

Activity-based protein profiling (ABPP) has emerged as a widely used chemical proteomics methodology for basic biology research^{1–5}. In ABPP, activity-based probes (ABPs) consisting of a reactive “warhead”, an enzyme inhibitory moiety that forms a covalent irreversible bond with their target protein(s) and often ensures a high enzyme class specificity, a linker, and a reporter tag are used to label, identify and report on active enzymes under native, physiological conditions. As reporters, biotin, fluorophores, or so-called two-step reporter tags such as alkyne or azide moieties are frequently used⁶. In the last years, intensive efforts have been undertaken, both to develop new ABPs targeting new enzyme families and to establish their use in, amongst others, drug discovery (target and lead discovery, target engagement)^{7–12}, plant biology¹³, or microbiology^{14–18}.

An evolving alternative ABPP application is its usage in biocatalyst screening and, thus, biotechnology (we here define biocatalysts as enzymes with potential industrial usage)^{19–21}. For example, ABPP can be used to uncover microbial biocatalysts able to turn over complex polymeric biomass like lignocellulose, entities much sought in the context of sustainable energy and circular bioeconomy processes^{22–24}. In contrast to sequence homology-based biocatalyst screening approaches (for instance, the analysis of genomics data²⁵), ABPP exploits the established enzyme selectivity of ABPs to identify new biocatalysts with desirable substrate preferences, in principle, without the need to assign sequence homologies (Fig. 1a)^{26,27}. In what is termed “ABP-based enrichment”, only those enzymes that are active and thus have the capability to react with an ABP are selected for ensuing identification by LC-MS/MS-based sequencing, which limits protein identification according to the target specificity of the used ABPs. Accordingly, the often-cumbersome biochemical protein expression and purification is limited to only those biocatalysts preselected by ABPP—an enormous reduction of work when compared to screening methods that rely on systematic protein expressions. This is particularly relevant in fungal biocatalyst screen campaigns, which are frequently hampered by difficult heterologous protein expression and purification, e.g., as a result of complex glycosylation patterns in the case of wood-degrading enzymes^{28,29}. Despite these intrinsic advances, ABPP-based biocatalyst discovery has been applied mostly in proof-of-concept studies using pure cultures, often of model organisms, and, more importantly, of standard culture media for their growth^{30–34}.

In the present study, we aimed to overcome this limitation and showcase the potential of the ABPP biocatalyst screening technology in a more complex and biotechnologically-relevant experimental setting. Accordingly, we applied ABPP to a suspended culture consisting of the white rot fungus *Phanerochaete chrysosporium* grown on minimal medium and solid beech wood chips as the sole carbon and energy source, similar to recent work regarding a functional ABPP approach in a set of basidiomycetes³⁵. Lignocellulose is the main component of dead wood and represents a highly recalcitrant polymeric complex built up from cellulose, xylan (hemicellulose), and lignin (Fig. 1b)³⁶. Sustainable methods for its efficient degradation are urgently sought for establishing a biotechnological conversion of non-food biomass into an industrial feedstock^{37–39}. This, however, requires a synergistic action of different biocatalysts such as glycoside hydrolases (GHs), carbohydrate esterases (CEs), polysaccharide lyases, and other enzymes belonging to the auxiliary enzyme class (auxiliary activities, AAs), such as lytic polysaccharide monoxygenases (LPMOs)^{40–42}. *P. chrysosporium* is an effective degrader of dead wood and lignocellulose in particular⁴³. Its genome harbors a large repertoire of lignocellulolytic enzymes consisting of more than 69 different carbohydrate-active enzyme (CAZyme) families, including a total

of 166 GHs, 14 CEs, and 57 glycosyltransferases (GT) (as listed in the CAZY database (www.cazy.org)^{44,45}). This enormous complexity turns this organism into a promising resource for biocatalyst discovery. However, due to its sheer size, the *P. chrysosporium* secretome cannot be explored by systematically expressing all enzymes. Instead, a methodology to preselect only those enzymes directly involved in lignocellulose degradation is required. Of note, the expression of these biocatalysts are regulated by the presence of the lignocellulosic substrate, demanding preselection assays in the presence of insoluble wood chips and thus highly heterogenous conditions⁴⁶.

With the aim to rapidly identify promising lignocellulose-degrading enzymes from this complex system via ABPP-based preselection, we applied FP-alkyne, a well-established serine hydrolase (SH)-targeting ABP^{47,48}, and the two structurally-related GH-targeting ABPs KY371 (*N*-alkynyl-cyclophellitol aziridine) and JJB111 (the biotin equivalent of KY371)^{49,50}, to *P. chrysosporium* cultures grown in minimal medium with beech wood chips as sole carbon and energy source (Fig. 1c). FP-alkyne is a reporter-tagged derivative of the well-known fluorophosphonate serine hydrolase inhibitors and thus specifically binds all SHs, i.e., serine proteases and metabolic SHs, without specificity within this enzyme class⁵¹. By contrast, JJB111 is an aziridine analog of cyclophellitol, a natural product inhibitor of GHs. JJB111 structurally mimics β -glucopyranoside moieties and, therefore, preferentially reacts with retaining β -glucosidases⁴⁹; in addition, it however also labels a variety of β -exoglycosidases^{52,53}. In hemicellulose degradation, the acetyl-xylan esterases, among them members of the SH family, cleave as one of the first steps in the overall degradation pathway acetyl groups from the carbohydrate/polysaccharide backbone⁵⁴. The GHs, in turn, are responsible for cellulose, pectin, and xylan hydrolysis. In addition to extracellular soluble enzymes found in the culture supernatant, we also applied our ABPP approach to substrate-bound enzymes isolated from wood chips. Our results thus demonstrate that ABPP allows a straightforward and technically simple targeted identification of active biocatalysts, including enzymes with previously unannotated gene sequences (composed of domains of unknown function (DUF)), directly from fungal/microbial cultures grown on complex lignocellulosic substrates.

Results

ABPP analysis of the *Phanerochaete chrysosporium* supernatant preparation. Previous studies on the identification of *P. chrysosporium* lignocellulose-degrading enzymes have been performed via ‘classical’ full proteome analyses of culture supernatants⁵⁵. To demonstrate the potential of ABPP in the rapid identification of lignocellulolytic enzymes, we grew *P. chrysosporium* cultures (DSM 1566) for 5 days at 37 °C in beech wood chips-containing minimal medium (Fig. 2a). The formation of fungal hyphae as well as macroscopic degradation of growth substrates in a submerged liquid culture confirmed efficient fungal cell growth under these conditions. The culture supernatant was filtrated, lyophilized, re-dissolved in buffer, and then labeled with 2 μ M of the two ABPs FP-alkyne or JJB111, respectively. For competitive ABPP experiments, pretreatment with either 50 μ M paraoxon in the case of FP-alkyne or 20 μ M KY358 in the case of JJB111 labeling was used. After affinity enrichment, target identification was achieved by on-bead tryptic digestion and LC-MS/MS analysis. Each identified protein was quantified using spectral intensity-based relative quantification; as a reference, DMSO- or corresponding competitor-treated samples were used. Only protein groups with a log₂ fold change (FC) of ≥ 2 were kept for further analysis in all ABPP-based labeling experiments.

Fig. 1 Activity-based protein profiling (ABPP) for identifying lignocellulose-degrading biocatalysts from *P. chrysosporium*. **a** Overview on the ABPP workflow for identifying lignocellulose-degrading enzymes from *P. chrysosporium* suspension cultures grown on minimal medium with beech wood chips. An ABP is added to either the filtrate of a *P. chrysosporium* beech wood culture (denoted as supernatant) or the dodecylmaltoside-solubilized substrate-bound fraction (denoted as SBF) after lyophilization. The pretreatment is followed by a standard ABPP workflow, consisting of click-attachment of a biotin-residue for affinity enrichment (in case of a two-step ABP), affinity enrichment of labeled enzymes, trypsin digest, and subsequent MS-based protein identification. The use of enzyme class-specific ABPs, therefore, results in targeted identification of active biocatalysts and thus functional enzyme screening and also enables the sequence-independent identification of novel biocatalysts without similarity to known homologous. The inlet image shows how *P. chrysosporium* binds to the solid wood surface during lignocellulose degradation. **b** Chemical structures of the ABPs and competitors used in this study. These are FP-alkyne (the “classical” serine hydrolase ABP) and JJB111 (GH ABP), as well as the FP competitor paraoxon and the JJB111 competitors KY371 and KY358. **c** Lignocellulose is a complex and recalcitrant polymer built up from cellulose, xylan (hemicellulose), and lignin. Its degradation requires the synergistic action of various different enzymes.

The application of JJB111 enabled the identification of twelve GHs (Table 1 and blue-labeled proteins in Fig. 2c; see Supplementary Data 2 for the complete list of identified proteins). Pretreatment with KY358 competed labeling of six of them belonging to the GH3, GH5, GH16, and GH74 families. Interestingly, in contrast to the identified GHs known to be involved in cellulose (GH3 and GH5), xylan (GH3), or xyloglucan (GH74) degradation, our analysis also revealed significant enrichment (\log_2 fold change: 5.39) of the protein Phchr2|3002168, which is annotated as a glutaminase in the JGI MycoCosm genome database⁴⁵ (Table 1 and red-labeled protein in Fig. 2c). The labeling of Phchr2|3002168 was competed by pre-incubation with KY358 and domain analysis by PFAM⁵⁷ and InterProScan⁵⁸ of its protein sequence revealed the presence of four DUF domains (DUF4964, DUF5127, DUF4965, and DUF1793) along with a secretion signal (Supplementary Fig. 1a). Structural homology analysis by HHPred⁵⁹ however predicted a β -glucosidase of the GH116 family from *Thermoanaerobacterium xylolyticum* (pdb code 5O0S⁶⁰, e-value of $7.2e^{-34}$, 14% sequence identity) and a GH52 xylosidase from *Geobacillus thermoglucosidasius* (pdb code 4C1O⁶¹, e-value of $7.1e^{-31}$, 10% sequence identity) as homologous proteins (Supplementary Fig. 1b). Additionally, analysis with InterProScan suggested the presence of a six-hairpin glycosidase domain which is characteristic for GH15, GH65, GH92 and GH116 family members. The AlphaFold⁶²-predicted 3D structure of Phchr2|3002168 showed a high secondary structure overlap to the GH52 β -xylosidase 4C1P ((α/α)₆ barrel) of 70% even despite its low sequence similarity of 12% (Supplementary Fig. 1c, d). Overall, these results indicate that Phchr2|3002168 might be a GH of a so far uncharacterized GH family.

ABPP analysis of the *P. chrysosporium* substrate-bound fraction (SBF). Our study so far demonstrates that ABPP can be used to detect active SH and GH enzymes in the fungal culture supernatant. However, lignocellulose degradation by *P. chrysosporium* involves a variety of enzymes, which partially attach to the respective carbohydrate-based substrate⁶³. Naturally, these substrate-bound enzymes are of particular interest for biotechnological applications as they are directly involved in lignocellulose degradation. During sample preparation, the lignocellulose substrate was filtered off to obtain a homogenous starting material for labeling. Indeed, such a discarding step is frequently performed in diverse functional screening approaches and a technically simple approach for targeted analysis of substrate-attached biocatalysts would be highly desirable.

To investigate whether such enzymes can be detected by a substrate-targeted ABPP approach, we again grew *P. chrysosporium* in the presence of beech wood chips. After removal of the culture supernatant and free fungal cells, the active substrate-bound enzymes were detached with 0.1% (w/v) of the MS-compatible detergent dodecyl- β -D-maltoside. The detached proteins were then lyophilized and the residue was resolved in buffer,

followed by the addition of the ABPs and the standard downstream MS sample preparation and analysis workflow (Fig. 3a).

The employment of FP-alkyne led to the identification of 53 substrate-bound proteins, of which seventeen were enriched with a \log_2 -fold change ≥ 2 . Their functional annotation revealed the presence of four CEs from the CE1 and CE15 families (Table 2 and green-labeled proteins in Fig. 3b; see Supplementary Data 3 for the complete list of identified proteins). Notably, only the labeling of the CE1 family proteins Phchr2|2983171 and Phchr2|126075 was successfully inhibited by pre-incubation with paraoxon. These proteins are of potential biotechnological interest due to their predicted potential as acetyl or feruloyl esterases. Moreover, five serine carboxypeptidases (S10) and four carboxylesterases were identified which may play a role in enzyme activation during lignocellulose degradation, for instance of cellobiose dehydrogenases⁶⁴. The labeling of three of the enriched carboxylesterases was furthermore competed by paraoxon.

The analysis of the ABPP approach with JJB111 revealed seven GHs, which were enriched with a \log_2 -FC ≥ 2 (blue-labeled proteins in Fig. 3c, see Supplementary Data 4 for the complete list of identified proteins). Of these, four were competed by pre-incubation with KY371. The proteins Phchr2|3002242, Phchr2|2945552, and Phchr2|3003144 are predicted members of the GH3 family, whereas protein Phchr2|2915237 belongs to GH5 subfamily 9; both GH families are known to catalyze cellulose or xylan degradation. By contrast, the labeling of the proteins Phchr2|3004009 (GH17), Phchr2|2895579 (GH5), and Phchr2|3038646 (GH25) was not competed by KY371. Of note, three of the identified proteins (i.e., Phchr2|291537, Phchr2|126075, and Phchr2|2912243) were also identified in the ABPP analysis of the supernatant.

Overall, these experiments demonstrate that ABPP is not only a technically simple targeted approach to identify promising lignocellulose biocatalysts but also enables to focus the analysis on relevant, active enzyme subfractions, e.g., those bound to an insoluble substrate.

Biochemical validation and characterization of selected identified enzymes. So far, our ABPP approach identified several potential lignocellulose-degrading biocatalysts. Their subsequent functional annotation was achieved by sequence homology analysis. The ABPP approach may, however, in principle, also enable the identification of enzymes of a new enzyme family, as enzyme identification is based on the ABP enzyme reactivity and not sequence homology. To demonstrate that our ABPP approach indeed resulted in the identification of lignocellulose-degrading biocatalysts, we selected three of the ABPP-identified enzymes, Phchr2|126075, Phchr2|2915237, and Phchr2|3002168, for further expression and subsequent biochemical characterization (Supplementary Fig. 2).

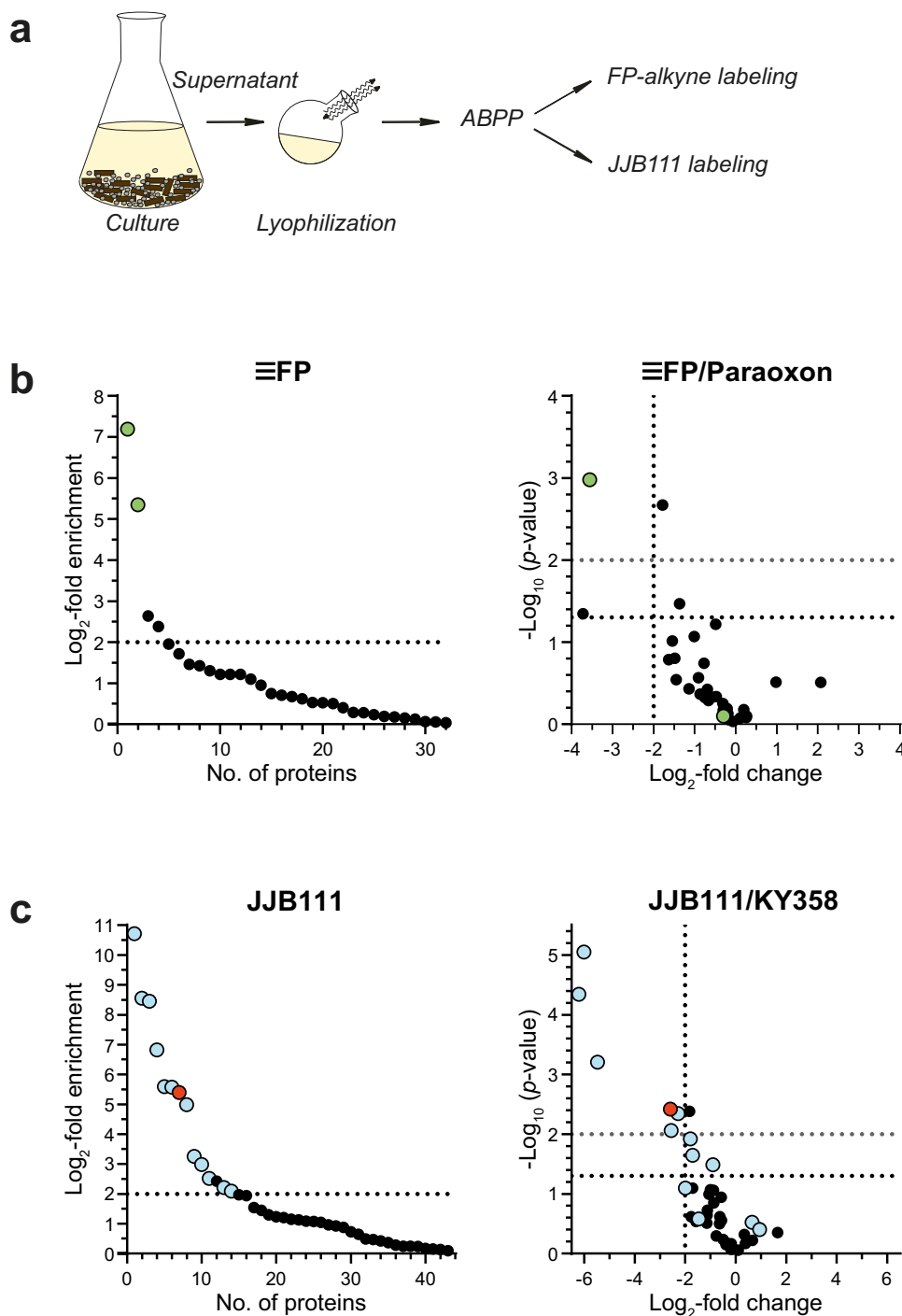


Fig. 2 ABPP of *P. chrysosporium* supernatants. **a** Workflow of the analysis. *P. chrysosporium* suspension cultures were grown for 5 days on a minimal medium supplemented with beech wood chips. The solid material was filtered off, the filtrate lyophilized, and the residue was subjected to ABPP with the corresponding probes. **b** ABPP of SHs without (left panel, \log_2 -fold change ≥ 2 compared to DMSO is indicated) or after pretreatment with 50 μM paraoxon (competition experiment, right panel) with 2 μM FP-alkyne after click chemistry ($n = 4$ biologically independent samples). Green dots indicate SHs. **c** ABPP of GHs without (left panel, \log_2 -fold change ≥ 2 compared to DMSO is indicated) or after pretreatment with 20 μM KY358 (competition experiment, right panel) with 2 μM JJB111 ($n = 4$ biologically independent samples). Blue dots indicate GHs, while the red dot represents the DUF protein with four domains of unknown function Phchr2|3002168 with potential GH activity.

Acetyl-xylan esterase activity of Phchr2|126075. The putative acetyl-xylan esterase Phchr2|126075 (338 amino acids; 35.5 kDa) was identified with FP-alkyne in the *P. chrysosporium* supernatant (\log_2 -fold enrichment of 7.18) and in the SBF (\log_2 -fold enrichment of 2.37). Phchr2|126075 belongs to the CE1 family and contains a fungal CBM1 motif for carbohydrate binding as well as

a secretion signal. The esterase domain encompasses the residues 80-288. Of note, a homologous acetyl-xylan esterase (Phchr2|129015, e-value of 0.0, 89% sequence identity) from the same species was previously characterized⁵⁶. For characterization, *phchr2|126075* was cloned into the pKLAC2 vector and over-expressed in *Kluyveromyces lactis*. As expected from the sequence

Table 1 Overview of the ABPP-identified enzymes from profiling the supernatant of *P. chrysosporium* suspension cultures grown on minimal medium supplemented with beech wood chips.

Protein ID	Annotation	Used ABP	Log ₂ fold change	MW [kDa]	Predicted function
Phchr2 126075	CE1-CBM1	FP-alkyne	7.19	35.59	acetyl-xylan esterase
Phchr2 2912243	CE15	FP-alkyne	5.35	44.29	4-O-methyl-glucuronoyl methylsterase
Phchr2 2915237	GH5 subfamily 9	JJB111	10.71	46.55	1,4-β-endoglucanase
Phchr2 3024052	GH3	JJB111	8.55	86.85	1,4-β-xylosidase
Q66NB7; Phchr2 2985730	GH5 subfamily 6	JJB111	8.45	40.41	1,4-β-glucanase
Q9URP5;O74203	CBM1-GH3	JJB111	6.82	85.56	β-glucosidase
Phchr2 2895579	GH5 subfamily 9	JJB111	5.59	73.50	1,3-β-glucosidase
Phchr2 2990154	GH30 subfamily 3	JJB111	5.58	65.13	1,6-β-endoglucanase
Phchr2 3002168	DUF	JJB111	5.39	74.43	DUF1793, DUF4964, DUF4965, DUF5127
Phchr2 3004260	GH79	JJB111	4.99	48.99	β-glucuronidase
Phchr2 2981757	GH5 subfamily 5	JJB111	3.26	42.05	1,4-β-endoglucanase
Phchr2 2927412	GH74	JJB111	2.99	79.15	exo-1,3-β-xyloglucanase/reducing end-specific cellobiohydrolase
Phchr2 2916357	GH28	JJB111	2.52	53.53	galacturonase
W5ZNX7;C6H06; Phchr2 123909	GH16	JJB111	2.21	33.57	endo-1,3-β-glucanase
Phchr2 3006243	GH18, CBM5	JJB111	2.09	54.31	endo-1,3-β-glucanase

Two SHs, 12 GHs, and one protein (DUF) with four DUFs with similarities to GHs were elucidated from FP-alkyne and JJB111 labeling, respectively.

homology, Phchr2|126075 displayed esterase activity against *p*NP-acetate and subsequent assays revealed the highest enzyme activity at a pH of 8 and 40 °C, respectively (Supplementary Fig. 3). Further kinetic characterizations then revealed a V_{\max} of 41.7 U mg⁻¹ protein and K_M of 0.67 mM for *p*NP-acetate hydrolysis (Fig. 4a).

β-glucanase activity of Phchr2|2915237. Phchr2|2915237 (422 amino acids; 46.5 kDa) was identified and enriched with JJB111 in both the soluble supernatant (log₂ fold change: 10.71) and the SBF (log₂ fold change: 4.11). An analysis with InterProScan indicates the presence of a GH5 cellulase domain encompassing residues 67-330. Phchr2|2915237 also contains an extracellular secretion signal but no transmembrane domains. A HHpred⁵⁹ analysis predicts either β-1-3 glucanases⁶⁵ (e-value: 9.7e⁻³⁵; 44% sequence identity) or β-1-4-xyloglucanases⁶⁶ (e-value: 8.8e⁻²⁴, 18% sequence identity) as the closest structural homologs. Additionally, homologs of Phchr2|2915237 are present in different wood-degrading fungal species, such as *Trametes* or *Pleurotus* as identified via BLASTP⁶⁷. The closest characterized homolog (e-value: 9e⁻¹⁰⁸, 45% sequence identity) of Phchr2|2915237 belongs to the yeast *Candida albicans* and plays a role in cell wall metabolism and restructuring⁶⁸.

We heterologously expressed Phchr2|2915237 in *Aspergillus oryzae* and studied its substrate specificity using a variety of chromogenic *p*-nitrophenol-sugar conjugates as well as different polysaccharides after purification. Phchr2|2915237 showed high GH activity if *p*NP-Glc and *p*NP-Xyl were used as substrates, while with *p*NP-Ara only residual and with *p*NP-Man, *p*NP-GlcNAc, and *p*NP-Gal no hydrolytic activity was observed (Fig. 4b). For *p*NP-Glc hydrolysis, a pH and temperature optimum of 5 and 60 °C, respectively, was elucidated (Supplementary Fig. 3). A 3, 5-dinitrosalicylic acid (DNSA) assay revealed that Phchr2|2915237 was also able to degrade both lichenan and beech wood xylan, while carboxymethyl cellulose (CMC), galactomannan, xyloglucan, and curdlan were no suitable substrates (Fig. 4b). More detailed kinetic characterizations revealed a V_{\max} of 999 U mg⁻¹ protein and a K_M of 1.82 mM as well as a V_{\max} of 612 U mg⁻¹ protein and a K_M of 6.98 mM for *p*NP-glucoopyranoside and *p*NP-xylopyranoside, respectively. For polysaccharide degradation, a V_{\max} value of 107 U mg⁻¹ protein

with a K_M value of 5.5 mg mL⁻¹ as well as a V_{\max} value of 71.6 U mg⁻¹ protein and a K_M value of 13.8 mg mL⁻¹ was determined with lichenan and beech wood xylan, respectively, showing that Phchr2|2915237 is also able to cleave natural, more complex sugar polymers (Fig. 4c). To determine if Phchr2|2915237 functions as an exo- or endo-glucanase/xylanase, we analyzed the hydrolysis products of xylan and lichenan generated by Phchr2|2915237 via thin-layer chromatography. No mono-saccharides were cleaved from the glucan chains; we instead observed the formation of polysaccharide degradation products with unknown length (Supplementary Fig. 4). In addition, a coupled assay using either a glucose or xylose dehydrogenase also showed no formation of either glucose or xylose by Phchr2|2915237 during hydrolysis of lichenan and xylan, respectively.

Activity of Phchr2|3002168 DUF family protein. Phchr2|3002168 (691 amino acids; 74.4 kDa) is an uncharacterized protein that was enriched by JJB111; its JJB111 labeling was competed by pretreatment with KY371. The four-domain protein Phchr2|3002168 is annotated as a glutaminase, but our ABPP approach, as well as our previously described additional sequence analyses suggested a possible GH activity. Moreover, a full proteome analysis of different fungal species including *P. chrysosporium* revealed that some of them secreted Phchr2|3002168-homologous proteins in the presence of lignocellulose⁶⁹.

We, therefore, tried to characterize this protein via biochemical assays. However, all our attempts to express and purify Phchr2|3002168 in *E. coli* or *A. oryzae* failed, emphasizing again the persisting difficulties of an expression- vs. an ABPP-based functional biocatalyst screening of fungal proteins. To confirm the activity of Phchr2|3002168 as a GH, we, therefore, searched for close homologs in other organisms and found WP_074995790 from *Streptomyces misionensis* (e-value: 0.0; 42% sequence identity) as a promising candidate. Please note that although WP_074995790 displays the same four domain structure as Phchr2|3002168, the InterPro protein annotation just lists the DUF5127 domain.

We, therefore, overexpressed WP_074995790 (752 amino acids; 80.5 kDa) and, although this enzyme was difficult to handle due to strong aggregation tendency in biochemical assays, we were able to perform some enzyme assays with this

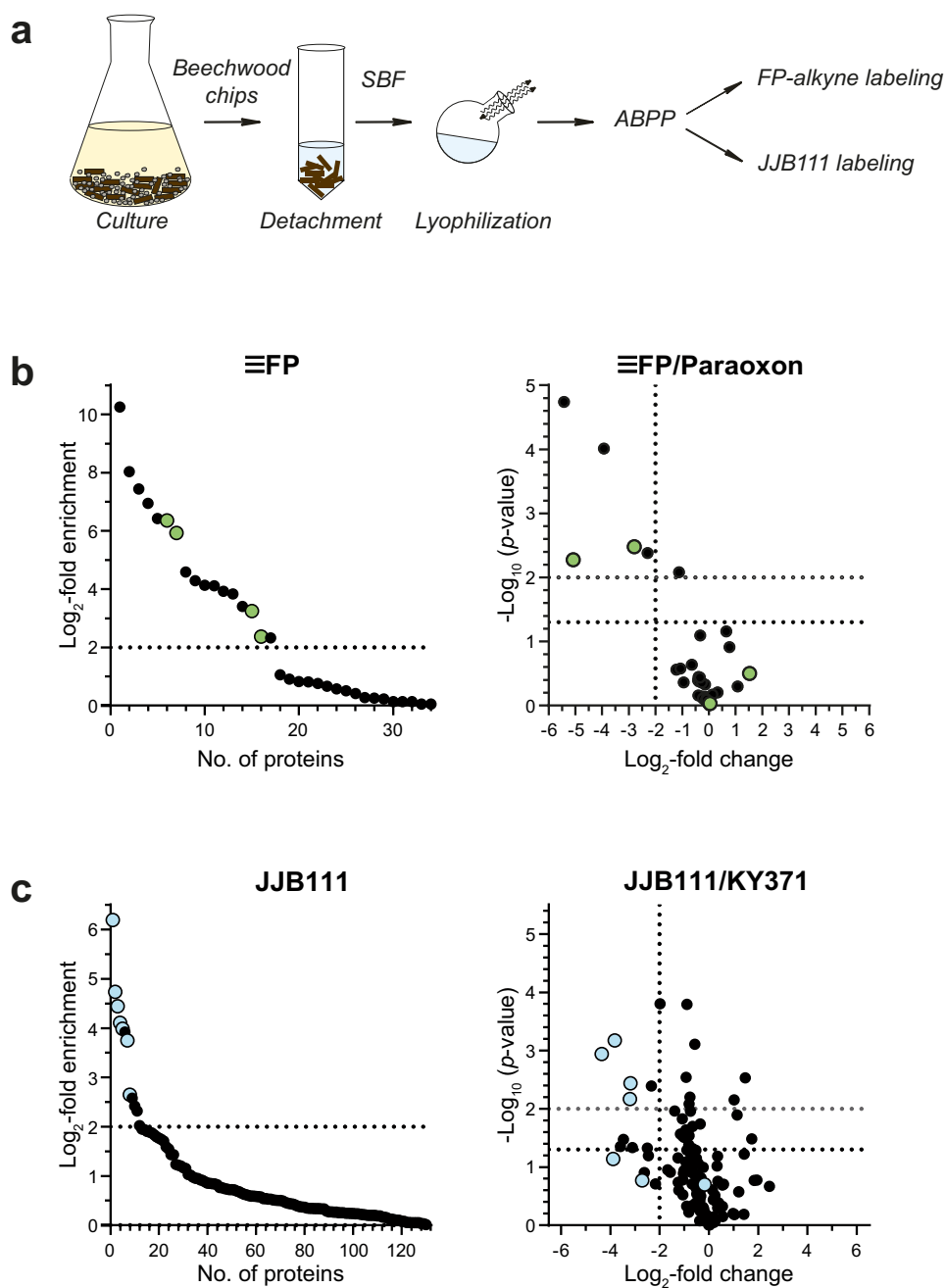


Fig. 3 ABPP of *P. chrysosporium* substrate-bound fraction (SBF). **a** Workflow of the SBF analysis. *P. chrysosporium* suspension cultures were grown for 5 days on a minimal medium supplemented with beech wood chips. The beech wood chips were isolated and substrate-bound proteins were isolated via 0.1% (w/v) dodecylmaltoside treatment. The obtained protein solution was lyophilized and the residue was subjected to ABPP with the corresponding probes. **b** ABPP of SHs without (left panel, \log_2 -fold change ≥ 2 compared to DMSO is indicated) or after pretreatment with 50 μM paraoxon (competition experiment, right panel) and 2 μM FP-alkyne after click chemistry ($n = 4$ biologically independent samples). Green dots indicate annotated CEs. **c** ABPP of GHs without (left panel, \log_2 -fold change ≥ 2 compared to DMSO is indicated) or after pretreatment with 20 μM KY371 (competition experiment, right panel) with 2 μM JJB111 ($n = 3$ biologically independent samples). Blue dots indicate annotated GHs.

preparation. Due to its original assignment as a glutaminase, we started with corresponding glutaminase enzyme assays but were unable to detect any conversion of a glutamine substrate. We next screened for GH activity by testing the same set of pNP-based sugar substrates used before with Phchr2|2915237. Satisfactorily, we were able to detect a weak β -galactosidase activity with a total specific activity of 0.85 U mg^{-1} protein. All other tested pNP-substrates were, however, not hydrolyzed (Fig. 4b). Altogether, these biochemical assays, therefore, demonstrated a GH activity of WP_074995790, although the observed overall weak β -

galactosidase activity predicts that other carbohydrate structures, e.g., more complex polymeric carbohydrates, may represent better substrates.

Discussion

White rot fungi exhibit excellent decomposition abilities and are responsible for the degradation of lignin in plant biomass; they have therefore attracted considerable interest as resources for identifying biotechnologically-relevant enzymes^{56,70}. Among them, the species *P. chrysosporium* seems to be particularly well-suited as

Table 2 Overview of the ABPP-identified enzymes from profiling the substrate-bound fraction (SBF) of *P. chrysosporium* suspension cultures grown on minimal medium supplemented with beech wood chips.

Protein ID	Annotation	Used ABP	Log ₂ fold change	MW [kDa]	Predicted function
Phchr2 3002242	GH3	JJB111	6.20	92.86	β-glucosidase
Phchr2 3004009	GH17	JJB111	4.73	33.78	1,3-β-glucanase
Phchr2 2895579	GH5	JJB111	4.44	73.50	1,3-β-glucosidase
Phchr2 2915237 ^a	GH5, subfamily 9	JJB111	4.11	46.55	1,4-β-endoglucanase
Phchr2 2945552	GH3-WSC	JJB111	3.99	98.95	β-glucosidase
Phchr2 3038646	GH25	JJB111	3.75	24.47	muramidase
Phchr2 3003144	GH3	JJB111	2.65	92.83	β-glucosidase
Phchr2 2912243 ^a	CE15	FP-alkyne	6.35	44.29	glucuronoyl methylesterase
Phchr2 2983171	CE1, CBM1	FP-alkyne	5.93	31.85	acetyl-xylan esterase
Phchr2 2918304	CE15, CBM1	FP-alkyne	3.24	47.47	glucuronoyl methylesterase
Phchr2 126075 ^a	CE1, CBM1	FP-alkyne	2.37	35.59	acetyl-xylan esterase

^aEnzymes that were also labeled in the supernatant (Table 1).

Four SHs and seven GHs were elucidated from FP-alkyne and JJB111 labeling, respectively.

a starting point for biocatalyst discovery as its enzymes are often thermostable due to its rather high growth temperature optimum of 40 °C⁷¹. Accordingly, multiple proteomic studies have investigated its secretome during growth on lignocellulose substrates with the aim of identifying lignocellulose-degrading biocatalysts, although most of the identified biocatalysts have never been validated biochemically^{46,55,69,72}.

Here we described an alternative approach for identifying biotechnologically-relevant enzymes via ABPP with enzyme class-specific ABPs that allows to focus the analysis on only active enzymes with relevance to the overall biological process. This approach can be employed to rapidly analyze extracellular soluble (supernatant) biocatalysts. More importantly, however, it can also be used, as described here for the first time, to identify substrate-bound biocatalysts, e.g., in our case, enzymes attached to solid lignocellulose in the form of beech wood chips. The ABPP approach represents a preselection step for the most promising biocatalysts. Its application to soluble enzymes or enzymes directly bound to substrates thereby allows a technically straightforward functional screening that we anticipate may find wider application in targeted biocatalyst discovery. The use of further ABPs, e.g., GH-directed probes with different α/β-specificity or different sugar selectivity⁵⁰, will thereby enable to elucidate additional enzymes for lignocellulose degradation.

To demonstrate the applicability of this approach, we biochemically validated three of the identified ABPP 'hits'. We selected one SH target of the FP-alkyne as well as one GH target of the JJB111 ABPP labeling approach. In addition, we chose to characterize a bacterial homolog of one target enzyme, Phchr2|3002168, which has not been assigned as a carbohydrate-active enzyme (CAZyme) based on sequence analyses. The FP-alkyne identified CE1 family protein Phchr2|126075 was homologous to a characterized acetyl-xylan esterase⁷³, and we were able to confirm a potent hydrolytic activity against pNP-acetate. This suggests a potential biotechnological application of this enzyme in the first steps of hemicellulose degradation, which is the cleavage of acetyl groups on lignocellulose. Phchr2|2915237 was identified as a promiscuous and highly catalytically active polysaccharide-cleaving β-endoglucanase that shows activity against both xylan and lichenan as well as against pNP-Glc and pNP-Xyl and releases neither glucose nor xylose from lichenan or xylan chains, respectively. It therefore most likely contributes to the breakdown of lignocellulose in *P. chrysosporium*. Phchr2|2915237 contains a secretion signal for extracellular transport and is predicted to belong to the GH5 subfamily 9. So far, only a few enzymes from this subfamily are known, most of which contain exo-β-1,3- or

exo-β-1,4-glucanase activity in addition to an endo-1,6-glucanase activity in some family members. A total of 17 family members have been characterized, all of them either belonging to different yeast or *Aspergillus* species, while no homolog has been described in any basidiomycete so far. Characterized homologs of Phchr2|2915237 have been found to play key roles in morphogenetic processes during development and differentiation, for example, in *Candida albicans* where exo-β-1,3-glucanases partially hydrolyze cell wall areas, enabling the insertion of new cell wall material, and can additionally also cleave 1,4 and 1,6 glycosidic bonds⁶⁸. In *S. pombe*, a GH5 subfamily 9 protein was able to hydrolyze both β-1,3 and β-1,6 glycosidic bonds^{74,75}. However, different homologs have also been shown to function as antifungal enzymes or to be involved in plant cell wall degradation^{76,77}. Interestingly, Phchr2|2915237 also shows some similarity to endo-1-4-glucanases that have been shown to catalyze the cleavage of different xylo/gluco-oligosaccharides⁶⁶. Regarding the cleavage of fungal cell walls in *P. chrysosporium*, so far, mostly GH16 and GH55 enzymes have been attributed to be involved in cell wall morphogenesis and nutrient recycling⁷⁸. Although we do not know the exact in vivo functions of Phchr2|2915237, its export seems to be induced when *P. chrysosporium* is grown on lignocellulose and its dual endo-glucanase/xylanase activity would allow the enzyme to take part in the degradation of plant cell wall material. The overall high enzymatic activity and the broad substrate specificity of Phchr2|2915237, in conjunction with its compatibility with heterologous expression in an industrial-relevant production organism such as *A. oryzae* turns this enzyme into a promising biocatalyst for efficient carbohydrate degradation. Finally, we were able to show the potential of ABPP to annotate also proteins of unknown or misassigned function. Phchr2|3002168 was identified by labeling with the GH probe JJB111. As we, however, failed to directly express this protein, we instead characterized the highly homologous protein WP_074995790 from *S. misionensis*, a bacterium also known to degrade cellulose in the form of sugarcane bagasse⁷⁹. We could confirm a β-galactosidase activity even though the total specific activity was low. This indicates that WP_074995790 might indeed possess GH activity but that we were not able to identify the native substrates for these novel enzymes so far. However, based on our labeling approach, sequence analysis, and enzyme assays, we suggest that proteins containing DUF4964, DUF5127, DUF4965, and DUF1793 domains, such as Phchr2|3002168 or WP_074995790 may function as glycoside hydrolases.

However, it should be noted that the reported ABPP approach in this study has also some limitations. The identification of

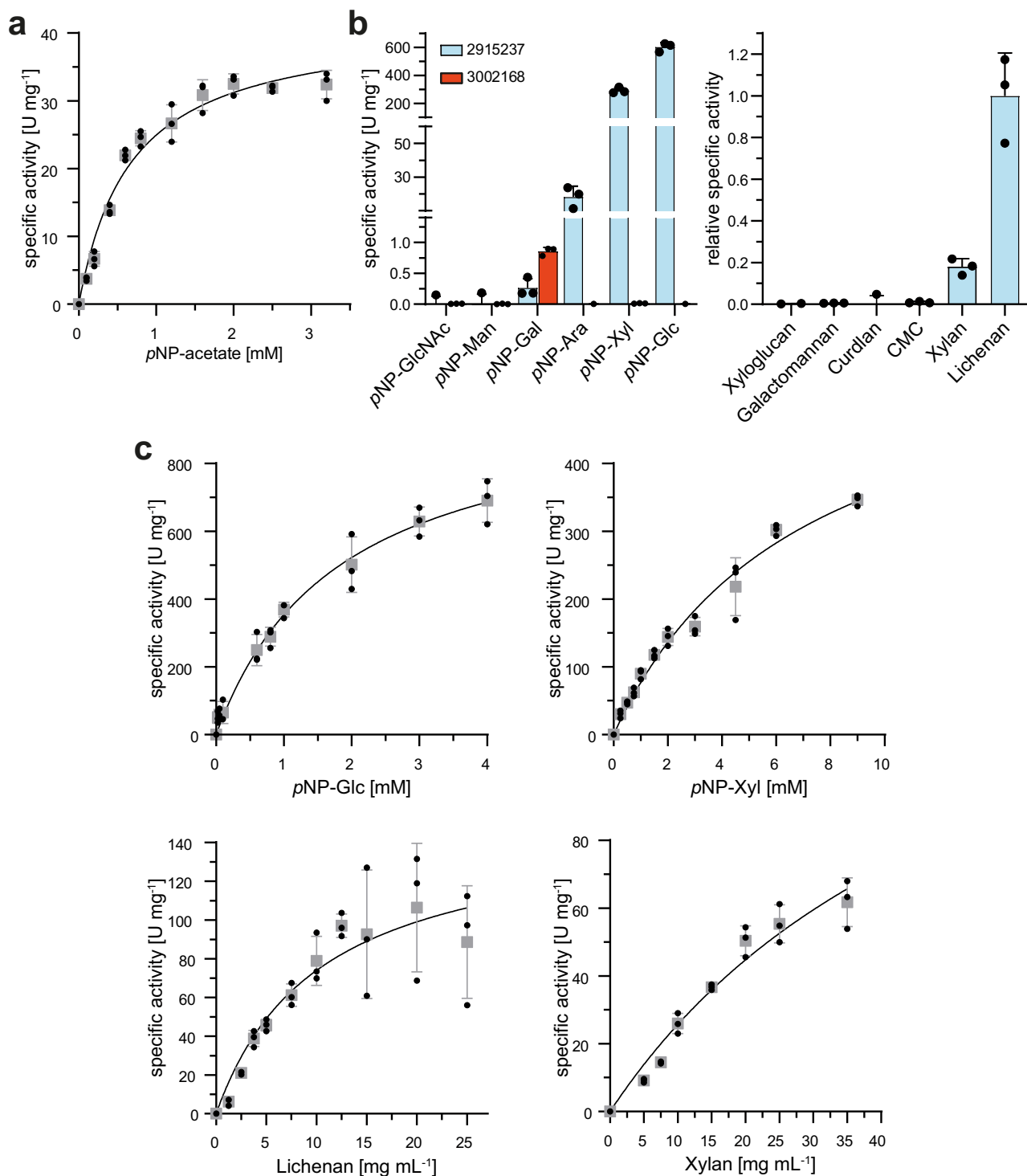


Fig. 4 Biochemical characterization of selected identified enzymes. **a** Phchr2|126075 was heterologously produced in *K. lactis* and activity was confirmed against pNP-acetate by following the release of pNP with a V_{max} of 41.7 U mg⁻¹ protein and a K_M of 0.67 mM. **b** Substrate specificity of Phchr2|2915237 and WP_074995790 from *S. missionensis*. The activity with different p-nitrophenol-based substrates was determined by measuring the release of pNP. The activity on complex polysaccharides was determined via the DNSA assay that quantifies the formation of reducing ends upon polysaccharide cleavage. Phchr2|2915237 displayed the highest activities with pNP-Glc, pNP-Xyl, lichenan, and beech wood xylan, while WP_074995790, a close homolog of Phchr2|3002168 showed activity only with pNP-Gal as a substrate with a specific activity of 0.85 U mg⁻¹ protein. **c** Kinetic characterization of Phchr2|2915237 using pNP-Glc, pNP-Xyl, lichenan, and beech wood xylan as substrates. All activity measurements were performed in triplicate ($n = 3$), mean values are shown and the error bars indicate the standard deviation (SD).

potential ABPP target proteins is influenced by the cell culturing conditions as the composition of proteins secreted by *P. chrysosporium* is dependent on culture growth time and the source and treatment of lignocellulosic biomass. Since both factors were fixed in this study, the scope of potential targets is limited to the proteins produced and secreted under these conditions. In addition, the usage of additional ABPs, preferentially with different target specificity, may allow the identification of further enzymes. ABPP labeling may also be lowered by the presence or enzyme-mediated production of competing metabolites. Finally, most ABPP methods currently require subsequent bioinformatics analysis and the verification of target hits via expression and purification due to a small proportion of unspecific labeling during ABPP profiling.

In conclusion, our ABPP approach may help to overcome a persisting challenge in biocatalyst discovery: the difficulty to link data from a functional screen to sequence information. The ABPP approach shortcuts this and allows to narrow the analysis to active enzymes targeted by the enzyme-selective ABP probe. As more and more ABPs become available, this may allow to identify biocatalyst ensembles even for the degradation of complex substrates solely by assembling together identified enzymes from different ABPP screening campaigns.

Methods

Chemicals. Chemicals for cultivation of *Escherichia coli* and *P. chrysosporium* DSM 1556, including yeast extract, malt extract, soytone, lysogeny broth, TRIS, MES, and salts for minimal media were obtained from Carl Roth (Germany). Beech wood chips for growth were obtained from J. Rettenmaier Söhne GmbH & Co. KG. (Germany). Carboxymethyl cellulose (CMC), lichenan, mannan, xyloglucan, and glucomannan were purchased from Sigma Aldrich (USA), and beech wood xylan was purchased from Carl Roth (Germany). *para*-nitrophenol (pNP), *para*-nitrophenyl- β -D-galactopyranoside (pNP-Gal), *para*-nitrophenyl-acetate (pNP-acetate), *para*-nitrophenyl- β -D-glucopyranoside (pNP-Glc), *para*-nitrophenyl- β -D-xylopyranoside (pNP-Xyl), *para*-nitrophenyl- β -D-mannose (pNP-Man), *para*-nitrophenyl- β -D-arabinofuranoside (pNP-Ara), and *para*-nitrophenyl-N-acetyl- β -D-glucosamine (pNP-GlcNAc) were purchased from Megazyme (Ireland), *n*-dodecyl β -D-maltoside (DDM) was obtained from Thermo Scientific (USA) and bovine serum albumin (BSA) from VWR Chemicals (USA). The sources of supply of more methodology-specific reagents are reported in the corresponding procedure section.

Cell growth and storage. *P. chrysosporium* DSM 1556 was obtained from the DSMZ (Germany). For long-term storage *P. chrysosporium* DSM 1556 was grown on MYP-agar plates (6 g L⁻¹ malt extract; 1 g L⁻¹ peptone from soy, 0.5 g L⁻¹ yeast extract) for 2 days at 37 °C. Afterward, cells were scraped off, resuspended in 100 μ L aliquots of sterile 50% (v/v) glycerol and stored at -80 °C as a glycerol stock. For the growth of solid cultures, 1.5% (w/v) MYP-Agar was inoculated with 20 μ L of *P. chrysosporium* DSM 1556 glycerol stock and grown for 2 days at 37 °C until the whole agar plate was covered with fungal hyphae. Afterward minimal medium containing 2.5 g L⁻¹ K₂HPO₄, 0.02 g L⁻¹ KH₂PO₄, 0.1 g L⁻¹ NaCl, 0.02 g L⁻¹ CaCl₂, 0.1 g L⁻¹ (NH₄)₂SO₄, 0.02 g L⁻¹ MgSO₄, 0.001 g L⁻¹ FeSO₄, and 40 g L⁻¹ beech wood chips at pH 5 supplemented with 100 μ g mL⁻¹ of chloramphenicol to inhibit bacterial growth were inoculated with plate grown *P. chrysosporium* DSM 1556. Suspension cultures were incubated for 5 days at 37 °C under constant shaking (180 rpm). Cell growth was tracked by following the formation of fungal hyphae as well as by the macroscopic degradation of the growth substrate. *E. coli* Rosetta DE3 was either grown in standard LB-medium (10 g L⁻¹ tryptone, 10 g L⁻¹ NaCl, 5 g L⁻¹ yeast extract) in precultures or in TB-medium (22 g L⁻¹ yeast extract, 12 g L⁻¹ tryptone, 4 mL L⁻¹ glycerol, 0.072 M K₂HPO₄, 0.017 M KH₂PO₄, pH 7.2) for heterologous overexpression of enzymes. *Kluyveromyces lactis* GG799 was grown in supplied media of the NEB *K. lactis* protein expression kit (New England Biolabs, USA) or in YPGlu media (10 g L⁻¹ yeast extract, 20 g L⁻¹ peptone, 2 % glucose, pH 7) for heterologous overexpression.

Production of supernatant and SBF. After 5 days, the supernatant was removed and sterilized by filtration through a 0.2 μ m filter (Filtropur S 0.2; Sarstedt, Germany). To obtain MS samples, 50 mL (labeling) culture supernatant was snap-frozen, lyophilized overnight, and subsequently stored at -20 °C until further analysis. For the isolation of substrate-bound proteins, *P. chrysosporium* DSM 1556 was grown on 40 g L⁻¹ beech wood chips in a minimal medium in 50 mL of submerged culture per replicate for a total of three or four biological replicates, respectively. After a growth time of 5 days, the beech wood chips were separated from the culture medium and free fungal cells by decantation and the remaining

wood chips were washed three times with 50 mL of buffer (50 mM TRIS, pH 8) by centrifugation (3000 \times g, 10 min, 4 °C) to remove all unbound proteins and cells. The pelleted wood chips were then incubated in 5 mL of 50 mM TRIS pH 8 containing 0.1 % (w/v) of the MS-compatible detergent DDM at 37 °C for 30 min under constant shaking at 180 rpm to solubilize all substrate-bound proteins. Similar to the supernatant, 5 mL of the detachment solution was snap-frozen, lyophilized overnight, and subsequently stored at -20 °C until further analysis.

Large-scale labeling for target identification. All probes and competitors were dissolved in DMSO. To identify enzyme targets of FP-alkyne and JJB111, lyophilized proteins were resuspended in 2 mL (supernatant) or 100 μ L (SBF) of either 50 mM Na₂PO₄ (pH 8.0) (FP-alkyne labeling) or 50 mM NaOAc (pH 5.0) (JJB111 labeling), respectively, and the protein concentration was determined with Roti[®]-Nanoquant (modified Bradford assay; Carl Roth, Germany). A total amount of 400 μ g (supernatant) or 100 μ g (substrate-bound fractions) protein was labeled with 2 μ M of the indicated probe (1 h, 37 °C, vigorous shaking). For the competition of labeling with indicated ABPs, either 50 μ M paraoxon (FP-alkyne labeling) or 20 μ M KY358-acyl or KY371 (JJB111 labeling) was used as indicated (30 min, 37 °C, vigorous shaking). FP-alkyne labeled proteins were subsequently subjected to a click reaction with 10 μ M TAMRA-biotin-N₃ (Jena Bioscience, Germany), 100 μ M TBTA (Sigma Aldrich, USA), 2 mM TCEP (Sigma Aldrich, USA), and 1 mM CuSO₄ (Sigma Aldrich, USA; 1 h, room temperature, in the dark).

Affinity enrichment and MS sample preparation. Prior to affinity enrichment, a modified methanol-chloroform⁸⁰ precipitation was performed to clean up proteins. Briefly, protein solutions were incubated with four equivalents of methanol (-20 °C, overnight) before one equivalent of chloroform and three equivalents of MS-grade water (VWR Chemicals, USA) were added. The precipitated proteins were washed twice with methanol, dried on air, and then dissolved in a final volume of 8 mL 0.2 % (w/v) SDS in 1 \times PBS (155 mM NaCl, 3 mM Na₂HPO₄, 1.06 mM KH₂PO₄, pH 7.4) under gentle shaking (37 °C, 30 min). For enrichment of ABP-reacted proteins, the obtained protein solution was incubated with 100 μ L of avidin bead slurry (Thermo Scientific, USA) while gently rotating (1 h, room temperature). Next, the beads were washed five times with 0.2 % (w/v) SDS (10 min, room temperature, gently rotating), followed by three washes with MS-grade H₂O (5 min, room temperature, vigorously shaking), and collected by centrifugation (400 \times g, room temperature, 5 min). The washed beads were taken up in 100 μ L of 0.8 M urea (GE Healthcare Life Sciences, USA) in 50 mM ammonium bicarbonate (ABC), the proteins were reduced with 5 mM dithiothreitol (DTT, Sigma Aldrich, USA) in 50 mM ABC (30 min, 37 °C, vigorous shaking), and subsequently alkylated by adding 10 mM iodoacetamide (IAM) in 50 mM ABC (30 min, 37 °C, in the dark). The alkylation reaction was quenched by adding DTT to a final concentration of 10 mM. For protein digestion, 1 μ g trypsin (Thermo Fisher Scientific, USA) in 50 mM acetic acid was added (37 °C, 16 h, vigorous shaking). The beads were collected by centrifugation (5 min, room temperature, 650 \times g) and the supernatant was recovered and mixed with formic acid (FA) to a final concentration of 0.5% (v/v). To wash the beads, 40 μ L 1% (v/v) FA were added (5 min, room temperature, vigorous shaking) and the supernatant was combined with the recovered digestion mix. To remove the remaining beads from the peptide solution, the mix was centrifuged (5 min, room temperature, 100 \times g) through a homemade two-disc glass microfiber membrane (GE Healthcare, USA; pore size 1.2 μ m, thickness 0.26 mm) StageTip. The cleared peptide solution was then desalted on homemade C18 StageTips as described (for the protocol used, see below).

Sample clean-up for LC-MS/MS. All peptide solutions after digestion and removal of solid matter were desalted using homemade C18 StageTips as described previously⁸¹. All centrifugation steps were performed in the range of 400–800 \times g and for 1–3 min at room temperature. Briefly, the acidified tryptic digests were passed over two-disc StageTips and the immobilized peptides were washed twice with 0.5% (v/v) FA. The peptides were eluted from the StageTips by a two-step elution with 80% (v/v) acetonitrile containing 0.5% (v/v) FA. After elution from the StageTips, samples were dried using a vacuum concentrator (Eppendorf, Germany) and the peptides were resuspended in 15 μ L 0.1% (v/v) FA. The thus prepared samples were directly used for LC-MS/MS experiments (see below for details).

LC-MS/MS. LC-MS/MS experiments were performed on an Orbitrap Fusion Lumos mass spectrometer (Thermo Fisher Scientific, USA) that was coupled to an EASY-nLC 1200 liquid chromatography (LC) system (Thermo Fisher Scientific, USA). The LC was operated in the one-column mode and the analytical column was a fused silica capillary (inner diameter 75 μ m \times 36–46 cm) with an integrated PicoFrit emitter (New Objective, USA) packed in-house with Reprosil-Pur 120 C18-AQ 1.9 μ m (Dr. Maisch, Germany). The analytical column was encased by a column oven (Sonation, Germany) and attached to a nanospray flex ion source (Thermo Fisher Scientific, USA). The column oven temperature was adjusted to 50 °C during data acquisition. The LC was equipped with two mobile phases: solvent A (0.1% (v/v) FA, in water) and solvent B (0.1 % (v/v) FA, 20% (v/v) H₂O, in acetonitrile). All solvents were of UHPLC (ultra-high-performance liquid chromatography) grade (Honeywell, Germany). Peptides were directly loaded onto

the analytical column with a maximum flow rate that would not exceed the set pressure limit of 980 bar (usually around 0.5–0.8 $\mu\text{L min}^{-1}$). Peptide solutions were subsequently separated on the analytical column using different gradients (105 min length); for details, see Supplementary File Sample_Legend_and_LC-MS_Settings, Section “LC_Settings”).

The mass spectrometer was operated using Xcalibur software v4.3.7.3.11. The mass spectrometer was set in the positive ion mode. Precursor ion scanning (MS¹) was performed in the Orbitrap analyzer (FTMS; Fourier Transform Mass Spectrometry with the internal lock mass option turned on (lock mass was 445.120025 m/z , polysiloxane))⁸². Dynamic exclusion was turned on (exclude after n times = 1; Exclusion duration (s) = 120; mass tolerance = ± 10 ppm). MS² product ion spectra were recorded only from ions with a charge bigger than +1 and in a data-dependent fashion in the ITMS (Ion Trap Mass Spectrometry). All relevant, individual MS settings (resolution, scan rate, scan range, AGC, ion acquisition time, charge states, isolation window, fragmentation type and details, cycle time, number of scans performed, and various other settings) for the individual experiments can be found in Supplementary File Sample_Legend_and_LC-MS_Settings, Section “MS_Settings”).

Protein identification using MaxQuant and Perseus. RAW spectra were submitted to an Andromeda search in MaxQuant⁸³ (version 1.6.10.43) using the default settings. Label-free quantification and match between runs was activated. MS/MS spectra data were searched against the UniProt *P. chrysosporium* (Phanerochaete_chrysosporium_Uniprot_210114.fasta; 430 entries) and Joint Genome Institute *P. chrysosporium* (Phanerochaete_chrysosporium_JGI_210114.fasta)⁴² (best-filtered model, 13602 entries) database. All searches included a contaminants database (as implemented in MaxQuant, 246 sequences). The contaminants database contains known MS contaminants and was included to estimate the level of contamination. Andromeda searches allowed the oxidation of methionine residues (16 Da), acetylation of the protein N-terminus (42 Da) as dynamic modifications, and the static modification of cysteine (57 Da, alkylation with IAM). Enzyme specificity was set to “Trypsin/P”. The instrument type in Andromeda searches was set to Orbitrap and the precursor mass tolerance was set to ± 20 ppm (first search) and ± 4.5 ppm (main search). The MS/MS match tolerance was set to ± 0.5 Da. The peptide spectrum matched FDR and the protein FDR were set to 0.01 (based on the target-decoy approach). The minimum peptide length was seven amino acids. For protein quantification, unique and razor peptides were allowed. Modified peptides with dynamic modifications were allowed for quantification. The minimum score for modified peptides was 40. Match between runs was enabled with a match time window of 0.7 min and match ion mobility window of 0.05 min⁸⁴. Further data analysis and filtering of the MaxQuant output was done in Perseus v1.6.2.3.⁸⁵ Label-free quantification (LFQ) intensities were loaded into the matrix from the proteinGroups.txt file and potential contaminants, as well as hits from the reverse database and hits only identified by peptides with a modification site, were removed. Related biological replicates were combined into categorical groups to allow comparison of different treatments or culture media. The data were transformed to the log₂-scale and only those proteins that were found in two of three or three of four replicates, respectively, were investigated separately. Prior to quantification, missing values were imputed from a normal distribution (width 0.3, downshift 1.8).

In labeling experiments, the log₂-fold enrichment of protein groups by FP-alkyne or JJB111 was calculated based on a two-sided Student’s *t*-test (permutation-based FDR: 0.05, $s = 0.1$, 250 randomizations) compared to the DMSO control. Proteins with a log₂-fold change > 2 were considered significantly enriched and all proteins with a positive fold change were plotted against their numerical order. To examine the effect of the competitor pretreatment on protein enrichment, a two-sided Student’s *t*-test (permutation-based FDR: 0.05, $s = 0.1$, 250 randomizations) was performed to calculate the difference in protein abundance between noncompetitive and pretreated probe-labeled samples and the statistical significance of the fold change. The log₂-fold change for samples preincubated with the corresponding competitors compared to noncompetition samples labeled with the respective probe was plotted against the $-\log p$ value. Proteins with a reduction of $> 75\%$ in their abundance and a p value < 0.01 were considered as primary hits, while proteins with a p value < 0.05 were reported as secondary hits. The protein ID was reported as either JGI ID or Uniprot ID.

Isolation of *P. chrysosporium* mRNA, cDNA synthesis, and cloning. For the synthesis of cDNA from *P. chrysosporium* DSM 1556, suspension cultures were grown on a minimal medium with beech wood chips for 5 days. Afterward, cells were separated from the beech wood chips by decantation and collected by centrifugation (6000 \times g, 4 °C, 30 min) and resuspended in 5 mL buffer (10 mM TRIS, pH 8). About 500 μL of cells were then transferred into a 0.1/0.5 mm bashing beads vial (Zymo Research, USA) and lysed in a bead beater (Precellys 24, VWR, USA). Afterward, cell debris was removed by centrifugation (16,000 \times g, 2 min) and RNA was isolated from 300 μL of lysed cells using the Monarch total RNA isolation kit (New England Biolabs, USA), by mixing with 300 μL of RNA lysis buffer. After RNA isolation, a cDNA library of *P. chrysosporium* DSM 1556 was synthesized using the SMARTer PCR cDNA synthesis kit (Takara Bio Europe, France). cDNA was stored at -20 °C and used for the amplification and sequencing of targeted genes. The *P. chrysosporium* gene *phchr2|126075* was amplified without introns

from cDNA using Q5[®] polymerase (New England Biolabs, USA) and the following gene-specific primers (Eurofins Genomics, Germany) 5′-ATGAGGTTGACATGTCCTCC-3′ and 5′-ACCTCCAATTCCTCGG-3′. The resulting PCR product was then used as a template for the amplification of *phchr2|126075* containing additional specific restriction sites for the pKLAC2 vector using the following primers: 5′-GAGGAGCATATGATGAGGTTGACATGTCCTCC-3′ and 5′-GAGGAGCTCGAGACCTCCAATTCCTCGG-3′ (*Nde*I and *Xho*I restriction sites underlined). Afterward, the PCR products were purified using the Wizard[®] SV Gel and PCR clean-up kit (Promega, USA). *Phchr2|126075* was cloned into the pKLAC2 vector (Novagen, USA), after restriction digest of the purified PCR products and the empty vector with the respective restriction enzymes (NEB, USA). The restricted PCR product and vector were used in a molar ratio of 1:4 for ligation using T4 DNA ligase (New England Biolabs, USA) at 16 °C overnight. *E. coli* DH5 α cells (Novagene, USA) were transformed with the obtained constructs and the presence of successfully cloned genes was confirmed by sequencing using the gene-specific primers above. pKLAC2:*phchr2|126075* was then transformed into *K. lactis* GG799 following the manufacturer’s instructions (New England Biolabs, USA). Correct integration of *phchr2|126075* was confirmed by PCR using the supplied integration primers. Transformed clones were inoculated in 2 mL of YPGlu medium to test for secretion of Phchr2|126075. Expression clones were isolated and resuspended in 250 μL of sterile 20% (v/v) glycerol and stored for further use at -80 °C. For expression of *phchr2|3002168* from *P. chrysosporium* and WP_074995790 from *Streptomyces misionensis* in *E. coli* Rosetta DE3, the coding sequences of both genes were synthesized by BioCat (Germany) without secretion signals for cloning into the pET20b-vector (with C-terminal His-tag). For heterologous overexpression, *E. coli* Rosetta DE3 was freshly transformed with the corresponding plasmid.

Heterologous overexpression of *phchr2|126075*, *phchr2|2915237*, and WP_074995790. Recombinant production of Phchr2|126075 was performed in *K. lactis* GG799 (New England Biolabs, USA). After transformation, the culture supernatant was collected by centrifugation (4000 \times g, 30 min, 4 °C) and screened for the clone with the highest activity against pNP-acetate (50 mL culture volume). To this end, 100 μL of the supernatant was incubated with 100 μL of 50 mM TRIS pH 8 and 400 μM pNP-acetate, and the release of *p*-NP was determined at 410 nm in a 96-well plate (BRANDplates[®], BRAND, Germany) using a Tecan infinite M200 plate reader (Tecan Trading AG, Switzerland). For protein expression, 50 mL of culture was inoculated with 1% (v/v) of a pre-culture and incubated for 3 days at 30 °C. Afterward, cells were centrifuged (4000 \times g, 30 min, 4 °C), and the supernatant passed through a 0.45 μm filter (Rotilabo[®] syringe filter, Carl Roth, Germany) before being used for determination of esterase activity of Phchr2|126075.

For the production of WP_074995790, a freshly inoculated *E. coli* Rosetta DE3 [pET20b::WP_074995790] culture in terrific broth (TB) medium (500 mL), supplemented with 150 $\mu\text{g mL}^{-1}$ ampicillin and 50 $\mu\text{g mL}^{-1}$ chloramphenicol, was grown to an OD₆₀₀ of 0.8 at 37 °C under constant shaking (180 rpm). Protein expression was induced by the addition of 1 mM isopropyl- β -D-thiogalactopyranoside (IPTG). The cells were then incubated at 18 °C for another 16 h, collected by centrifugation (6000 \times g, 20 min, 4 °C), and resuspended in 10 mL of buffer A (50 mM TRIS-HCl pH 7.8, 200 mM KCl, 10 mM imidazole) per gram of wet weight. Cell lysis was performed by sonication using a UP 200 S sonicator (Hielscher Ultrasonics GmbH, Germany) for 3 \times 7 min (50% amplitude, 0.5 s⁻¹) under constant cooling followed by centrifugation (14,000 \times g, 60 min, 4 °C). For further purification of WP_074995790, the cleared lysate was passed through a 0.45 μm filter (Rotilabo[®] syringe filter, Carl Roth, Germany) and applied onto a 5 mL Ni-IDA column (Cytiva, USA) equilibrated with buffer A at a flow speed of 5 mL min⁻¹. After washing with buffer A (20 column volumes), elution was performed with a linear gradient of buffer B (50 mM TRIS pH 7.8, 200 mM KCl, 400 mM imidazole). The elution buffer was exchanged for storage buffer (50 mM TRIS pH 8.0, 20 mM KCl, 10% (v/v) glycerol) using Amicon[®] centrifugal filter devices (50 kDa cutoff, Merck, Germany) by repeated concentration and dilution. For storage, proteins were flash-frozen in liquid nitrogen and stored at -80 °C.

A synthetic gene encoding Phchr2|2915237 was ordered as a gBlock from IDTdna (USA), integrated into the genome of *Aspergillus oryzae* and expressed as an extracellular enzyme as described elsewhere⁸⁶. A C-terminal his-tag (6xHis) was added to ease downstream purification. The fermentation broth was sterile filtered and 500 mM NaCl was added and adjusted to pH 7.5 by the addition of NaOH. The sample was loaded onto a Ni-Sepharose[™] 6 Fast Flow column (GE Healthcare, USA) equilibrated in 50 mM HEPES, pH 7.5 with 500 mM NaCl (buffer A). After loading, the column was washed with 10 column volumes of buffer A, and bound proteins were eluted with 500 mM imidazole in buffer A. The fractions containing the enzyme were pooled and applied to a Sephadex[™] G-25 (medium) (GE Healthcare, USA) column equilibrated and eluted in 100 mM HEPES pH 7.5. Fractions were analyzed by SDS-PAGE, and fractions containing the enzyme were combined.

Activity assay for recombinant CE1. Esterase activity was confirmed by measuring the release of *p*NP from *p*NP-acetate at 410 nm in either a discontinuous or continuous assay. To confirm the pH optimum of Phchr2|126075 1.08 μg of protein was incubated in 50 mM phosphate citrate buffer at a pH range from 5.5 to 8 with 25 mM of *p*NP-acetate in a total volume of 500 μL for 10 min at 35 °C.

Afterward, the reactions were stopped by the addition of 500 μL of 0.5 M sodium carbonate and the absorption of samples was determined at 410 nm. The production of pNP was then calculated using the established extinction coefficient of $16.1 \text{ mM}^{-1} \text{ cm}^{-1}$ for pNP⁸⁷. For determination of the temperature optimum, Phchr2|126075 was incubated in 50 mM TRIS pH 8, 20 mM KCl and release of pNP from pNP-acetate was continuously determined in a temperature range from 30 to 70 °C in a total volume of 500 μL . For the determination of kinetic constants, Phchr2|126075 was incubated with different concentrations of pNP-acetate at 40 °C in an aqueous solution containing 50 mM TRIS pH 8 and 20 mM KCl. Initial velocities of each reaction were taken and used for the calculation of activities. All enzyme assays were performed in triplicate.

Activity assay for recombinant glycoside hydrolases. The activity of Phchr2|2915237 and WP_074995790 was established against different polysaccharides and artificial pNP-conjugates. The pH and temperature optimum of Phchr2|2915237 were determined by incubating 0.5 μg of an enzyme with 200 μM of pNP-Glc or pNP-Gal for 10 min in 500 μL phosphate citrate buffer⁸⁸ at a pH range of 3–8 at 30 °C or a temperature range of 30–70 °C at pH 5, respectively. After incubation, reactions were stopped by the addition of 500 μL of 0.5 M sodium carbonate and cleavage of pNP-Glc was followed by determining the release of pNP at 410 nm as described above. For measuring the substrate specificity of Phchr2|2915237 and WP_074995790 against pNP-conjugates, either 200 μM of pNP-Ara, pNP-GlcNAc, pNP-Gal, and pNP-Man were incubated with 10 μg of Phchr2|2915237 or 11.2 μg of WP_074995790 or 200 μM of pNP-Xyl and pNP-Glc were incubated with 0.4 μg of Phchr2|2915237 or 11.2 μg of WP_074995790 as described above. Kinetic measurements against nitrophenyl substrates were performed in a discontinuous assay by incubating different concentrations of pNP-Glc and pNP-Xyl with 0.5 μg of Phchr2|2915237 while stopping the reaction after 1, 2, 5, and 10 min, respectively. Afterward, the initial velocities of the reaction were used to calculate the specific activity of Phchr2|2915237 towards nitrophenol substrates.

For measuring the substrate specificity of Phchr2|2915237 against xyloglucan, galactomannan, curdlan, CMC, xylan, and lichenan, 1 μg of enzyme was incubated with 0.5 % (w/v) of the respective polysaccharides in 500 μL of in citric acid phosphate buffer pH 5 for 1, 2, 5, and 10 min. Afterward, 500 μL of the DNSA solution (10 g L^{-1} DNSA, 2 mL L^{-1} of 0.05 g L^{-1} sodium sulfite, 200 g L^{-1} potassium sodium tartrate, and 10 g L^{-1} NaOH) was added and subsequently incubated at 100 °C for 15 min. Samples were then cooled on ice for 15 min and centrifuged (10,000 \times g, 4 °C for 30 min) before 250 μL of the supernatant was transferred into a 96-well plate (BRANDplates®, BRAND, Germany). The release of reducing sugars was followed at 575 nm using a Tecan infinite M200 plate reader (Tecan Trading AG, Switzerland) and determined using a calibration curve based on D-glucose. Background absorption due to abiotic substrate degradation and the addition of protein solutions was determined and subtracted from the absorption of the samples. Glutaminase activity was tested using L-gamma-glutamyl-pNP as described before⁸⁹ or by monitoring the formation of glutamate from glutamine by glutamate dehydrogenase (Merck, Germany) via the reduction of NAD⁺. To check for the formation of glucose or xylose during polysaccharide degradation, samples were incubated as described above for 4 h. Afterward, samples were centrifuged at 10,000 \times g, 4 °C for 30 min, and 200 μL of supernatant was incubated with either 2 U of glucose dehydrogenase (Sigma Aldrich, USA) or xylose dehydrogenase (Megazyme, Ireland) and 5 mM of NAD⁺ in 50 mM TRIS pH 8 and the formation of glucose and xylose was then measured by determining the reduction of NAD⁺ to NADH. All enzyme assays were performed in triplicate.

Analysis of hydrolysis products by thin-layer chromatography (TLC). Enzymatic reactions with 0.5 % (w/v) lichenan or beech wood xylan were performed under the same conditions as used for the DNSA assay with 10 μg of Phchr2|2915237. Lichenan and xylan were incubated with Phchr2|2915237 as described above, and samples were removed after 1, 2, and 4 h. Afterward, 2.5 μL of the hydrolysis products and control solutions were applied to aluminum sheet silica gel 60/kieselguhr F254 plates (20 cm \times 20 cm, Merck, Germany) and separated at room temperature with ethyl acetate, methanol and H₂O (68:23:9, v/v/v) as solvent. Plates were dried and products were visualized by treating the plates with a KMnO₄ staining solution (1.5 g KMnO₄, 10 g K₂CO₃, and 1.25 mL 10% aq. NaOH in 200 mL H₂O) and incubation at room temperature for 30 min.

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

Data availability

The mass spectrometry proteomics data for the digestions have been deposited to the ProteomeXchange Consortium via the PRIDE⁹⁰ partner repository (<https://www.ebi.ac.uk/pride/archive/>) with the dataset identifier PXD030618. All data generated during this study are included in this published article (and its supplementary information files). All source data underlying the graphs and charts can be found in Supplementary Data 5. Raw datasets generated during the current study are available upon request.

Received: 21 February 2022; Accepted: 20 October 2022;

Published online: 16 November 2022

References

- Cravatt, B. F., Wright, A. T. & Kozarich, J. W. Activity-based protein profiling: from enzyme chemistry to proteomic chemistry. *Annu. Rev. Biochem.* **77**, 383–414 (2008).
- Sanman, L. E. & Bogyo, M. Activity-based profiling of proteases. *Annu. Rev. Biochem.* **83**, 249–273 (2014).
- Benns, H. J., Wincott, C. J., Tate, E. W. & Child, M. A. Activity- and reactivity-based proteomics: Recent technological advances and applications in drug discovery. *Curr. Opin. Chem. Biol.* **60**, 20–29 (2021).
- Fang, H. et al. Recent advances in activity-based probes (ABPs) and affinity-based probes (ABPs) for profiling of enzymes. *Chem. Sci.* **12**, 8288–8310 (2021).
- Li, N., Overkleeft, H. S. & Florea, B. I. Activity-based protein profiling: an enabling technology in chemical biology research. *Curr. Opin. Chem. Biol.* **16**, 227–233 (2012).
- Verhelst, S. H. L., Bongers, K. M. & Willems, L. I. Bioorthogonal reactions in activity-based protein profiling. *Molecules* **25**, 5994 (2020).
- Niphakis, M. J. & Cravatt, B. F. Enzyme inhibitor discovery by activity-based protein profiling. *Annu. Rev. Biochem.* **83**, 341–377 (2014).
- Deng, H., Lei, Q., Wu, Y. P., He, Y. & Li, W. M. Activity-based protein profiling: recent advances in medicinal chemistry. *Eur. J. Med. Chem.* **191**, 112151 (2020).
- Nomura, D. K., Dix, M. M. & Cravatt, B. F. Activity-based protein profiling for biochemical pathway discovery in cancer. *Nat. Rev. Cancer* **10**, 630–638 (2010).
- Kok, B. P. et al. Discovery of small-molecule enzyme activators by activity-based protein profiling. *Nat. Chem. Biol.* **16**, 997–1005 (2020).
- Parker, C. G. et al. Ligand and target discovery by fragment-based screening in human cells. *Cell* **168**, 527–541 (2017).
- Wright, M. H. & Sieber, S. A. Chemical proteomics approaches for identifying the cellular targets of natural products. *Nat. Prod. Rep.* **33**, 731–733 (2016).
- Morimoto, K. & van der Hoorn, R. A. The increasing impact of activity-based protein profiling in plant science. *Plant Cell Physiol.* **57**, 446–461 (2016).
- Keller, L. J., Babin, B. M., Lakemeyer, M. & Bogyo, M. Activity-based protein profiling in bacteria: applications for identification of therapeutic targets and characterization of microbial communities. *Curr. Opin. Chem. Biol.* **54**, 45–53 (2020).
- Lentz, C. S. What you see is what you get: activity-based probes in single-cell analysis of enzymatic activities. *Biol. Chem.* **401**, 233–248 (2020).
- Krysiak, J. & Sieber, S. A. Activity-based protein profiling in bacteria. *Methods Mol. Biol.* **1491**, 57–74 (2017).
- Heal, W. P. & Tate, E. W. Application of activity-based protein profiling to the study of microbial pathogenesis. *Top. Curr. Chem.* **324**, 115–135 (2012).
- Steuten, K. et al. Challenges for targeting SARS-CoV-2 proteases as a therapeutic strategy for COVID-19. *ACS Infect. Dis.* **7**, 1457–1468 (2021).
- 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).
- Pagar, A. D. et al. Recent advances in biocatalysis with chemical modification and expanded amino acid alphabet. *Chem. Rev.* **121**, 6173–6245 (2021).
- Fuerst, R. & Breinbauer, R. Activity-based protein profiling (ABPP) of oxidoreductases. *ChemBiochem* **22**, 630–638 (2021).
- Rosnow, J. J., Anderson, L. N., Nair, R. N., Baker, E. S. & Wright, A. T. Profiling microbial lignocellulose degradation and utilization by emergent omics technologies. *Crit. Rev. Biotechnol.* **37**, 626–640 (2017).
- Liu, Y. et al. Advancing understanding of microbial bioenergy conversion processes by activity-based protein profiling. *Biotechnol. Biofuels* **8**, 156 (2015).
- Schroder, S. P. et al. Dynamic and functional profiling of xylan-degrading enzymes in aspergillus secretomes using activity-based probes. *ACS Cent. Sci.* **5**, 1067–1078 (2019).
- Rodriguez Benitez, A. & Narayan, A. R. H. Frontiers in biocatalysis: profiling function across sequence space. *ACS Cent. Sci.* **5**, 1747–1749 (2019).
- Barglow, K. T. & Cravatt, B. F. Activity-based protein profiling for the functional annotation of enzymes. *Nat. Methods* **4**, 822–827 (2007).
- Hunerdosse, D. & Nomura, D. K. Activity-based proteomic and metabolomic approaches for understanding metabolism. *Curr. Opin. Biotechnol.* **28**, 116–126 (2014).
- Zhang, T., Liu, H., Lv, B. & Li, C. Regulating strategies for producing carbohydrate active enzymes by filamentous fungal cell factories. *Front. Bioeng. Biotechnol.* **8**, 691 (2020).

29. Chen, L. et al. Specificity of O-glycosylation in enhancing the stability and cellulose binding affinity of Family 1 carbohydrate-binding modules. *Proc. Natl Acad. Sci. USA* **111**, 7612–7617 (2014).
30. Chauvine-Hines, L. M. et al. Suite of activity-based probes for cellulose-degrading enzymes. *J. Am. Chem. Soc.* **134**, 20521–20532 (2012).
31. Jain, N., Tamura, K., Dejean, G., Van Petegem, F. & Brumer, H. Orthogonal active-site labels for mixed-linkage endo-beta-glucanases. *ACS Chem. Biol.* **16**, 1968–1984 (2021).
32. Chen, Y. et al. Activity-based protein profiling of retaining alpha-amylases in complex biological samples. *J. Am. Chem. Soc.* **143**, 2423–2432 (2021).
33. de Boer, C. et al. Glycosylated cyclophellitol-derived activity-based probes and inhibitors for cellulases. *RSC Chem. Biol.* **1**, 148–155 (2020).
34. McGregor, N. G. S. et al. Rational design of mechanism-based inhibitors and activity-based probes for the identification of retaining alpha-l-arabinofuranosidases. *J. Am. Chem. Soc.* **142**, 4648–4662 (2020).
35. McGregor, N. G. S. et al. Activity-based protein profiling reveals dynamic substrate-specific cellulase secretion by saprotrophic basidiomycetes. *Biotechnol. Biofuels Bioprod.* **15**, 6 (2022).
36. Isikgor, F. H. & Becer, C. R. Lignocellulosic biomass: a sustainable platform for the production of bio-based chemicals and polymers. *Polym. Chem.* **6**, 4497–4559 (2015).
37. Himmel, M. E. et al. Biomass recalcitrance: engineering plants and enzymes for biofuels production. *Science* **315**, 804–807 (2007).
38. Himmel, M. E. & Bayer, E. A. Lignocellulose conversion to biofuels: current challenges, global perspectives. *Curr. Opin. Biotechnol.* **20**, 316–317 (2009).
39. Houfani, A. A., Anders, N., Spiess, A. C., Baldrian, P. & Benallaoua, S. Insights from enzymatic degradation of cellulose and hemicellulose to fermentable sugars – a review. *Biomass. Bioenerg.* **134**, 105481 (2020).
40. van den Brink, J. & de Vries, R. P. Fungal enzyme sets for plant polysaccharide degradation. *Appl. Microbiol. Biotechnol.* **91**, 1477–1492 (2011).
41. Andlar, M. et al. Lignocellulose degradation: an overview of fungi and fungal enzymes involved in lignocellulose degradation. *Eng. Life. Sci.* **18**, 768–778 (2018).
42. Ohm, R. A. et al. Genomics of wood-degrading fungi. *Fungal Genet. Biol.* **72**, 82–90 (2014).
43. da Silva, R. R., Peduzzi, R. & Souto, T. B. Exploring the bioprospecting and biotechnological potential of white-rot and anaerobic Neocallimastigomycota fungi: peptidases, esterases, and lignocellulolytic enzymes. *Appl. Microbiol. Biotechnol.* **101**, 3089–3101 (2017).
44. Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M. & Henrissat, B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* **42**, D490–D495 (2014).
45. Martinez, D. et al. Genome sequence of the lignocellulose degrading fungus *Phanerochaete chrysosporium* strain RP78. *Nat. Biotechnol.* **22**, 695–700 (2004).
46. Xie, C. et al. Comparative secretome of white-rot fungi reveals co-regulated carbohydrate-active enzymes associated with selective ligninolysis of ramie stalks. *Microb. Biotechnol.* **14**, 911–922 (2021).
47. Simon, G. M. & Cravatt, B. F. Activity-based proteomics of enzyme superfamilies: serine hydrolases as a case study. *J. Biol. Chem.* **285**, 11051–11055 (2010).
48. Liu, Y. S., Patricelli, M. P. & Cravatt, B. F. Activity-based protein profiling: the serine hydrolases. *Proc. Natl Acad. Sci. USA* **96**, 14694–14699 (1999).
49. Kallemeijn, W. W. et al. Novel activity-based probes for broad-spectrum profiling of retaining beta-exoglucosidases in situ and in vivo. *Angew. Chem. Int. Ed. Engl.* **51**, 12529–12533 (2012).
50. Wu, L. et al. An overview of activity-based probes for glycosidases. *Curr. Opin. Chem. Biol.* **53**, 25–36 (2019).
51. Bachovchin, D. A. et al. Superfamily-wide portrait of serine hydrolase inhibition achieved by library-versus-library screening. *Proc. Natl Acad. Sci. USA* **107**, 20941–20946 (2010).
52. Chandrasekar, B. et al. Broad-range glycosidase activity profiling. *Mol. Cell Proteom.* **13**, 2787–2800 (2014).
53. Husaini, A. M. et al. Multiplex fluorescent, activity-based protein profiling identifies active alpha-glycosidases and other hydrolases in plants. *Plant Physiol.* **177**, 24–37 (2018).
54. Sista Kameshwar, A. K. & Qin, W. Understanding the structural and functional properties of carbohydrate esterases with a special focus on hemicellulose deacetylating acetyl xylan esterases. *Mycology* **9**, 273–295 (2018).
55. Machado, A. S. et al. The secretome of *Phanerochaete chrysosporium* and *trametes versicolor* grown in microcrystalline cellulose and use of the enzymes for hydrolysis of lignocellulosic materials. *Front. Bioeng. Biotechnol.* **8**, 826 (2020).
56. Huy, N. D., Thiyagarajan, S., Kim, D. H. & Park, S. M. Cloning and characterization of a novel bifunctional acetyl xylan esterase with carbohydrate binding module from *Phanerochaete chrysosporium*. *J. Biosci. Bioeng.* **115**, 507–513 (2013).
57. Mistry, J. et al. Pfam: the protein families database in 2021. *Nucleic Acids Res.* **49**, D412–D419 (2021).
58. Jones, P. et al. InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**, 1236–1240 (2014).
59. Gabler, F. et al. Protein sequence analysis using the MPI bioinformatics toolkit. *Curr. Protoc. Bioinforma.* **72**, e108 (2020).
60. Artola, M. et al. 1,6-Cyclophellitol cycloulfates: a new class of irreversible glycosidase inhibitor. *ACS Cent. Sci.* **3**, 784–793 (2017).
61. Espina, G. et al. A novel beta-xylosidase structure from *Geobacillus thermoglucosidasius*: the first crystal structure of a glycoside hydrolase family GH52 enzyme reveals unpredicted similarity to other glycoside hydrolase folds. *Acta Crystallogr. D. Biol. Crystallogr.* **70**, 1366–1374 (2014).
62. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
63. Nimlos, M. R. et al. Binding preferences, surface attachment, diffusivity, and orientation of a family 1 carbohydrate-binding module on cellulose. *J. Biol. Chem.* **287**, 20603–20612 (2012).
64. Eriksson, K.-E. & Petterson, B. Purification and partial characterization of two acidic proteases from the white-rot fungus *Sporotrichum pulverulentum*. *Eur. J. Biochem* **124**, 635–642 (1982).
65. Taylor, S. C., Ferguson, A. D., Bergeron, J. J. & Thomas, D. Y. The ER protein folding sensor UDP-glucose glycoprotein-glucosyltransferase modifies substrates distant to local changes in glycoprotein conformation. *Nat. Struct. Mol. Biol.* **11**, 128–134 (2004).
66. Gloster, T. M. et al. Characterization and three-dimensional structures of two distinct bacterial xyloglucanases from families GH5 and GH12. *J. Biol. Chem.* **282**, 19177–19189 (2007).
67. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
68. Nakatani, Y., Larsen, D. S., Cutfield, S. M. & Cutfield, J. F. Major change in regiospecificity for the exo-1,3-beta-glucanase from *Candida albicans* following its conversion to a glycosynthase. *Biochemistry* **53**, 3318–3326 (2014).
69. Arntzen, M. O. et al. Quantitative comparison of the biomass-degrading enzyme repertoires of five filamentous fungi. *Sci. Rep.* **10**, 20267 (2020).
70. Chen, C.-C., Dai, L., Ma, L. & Guo, R.-T. Enzymatic degradation of plant biomass and synthetic polymers. *Nat. Rev. Chem.* **4**, 114–126 (2020).
71. Kirk, T. K., Schultz, E., Connors, W., Lorenz, L. & Zeikus, J. Influence of culture parameters on lignin metabolism by *Phanerochaete chrysosporium*. *Arch. Microbiol.* **117**, 277–285 (1978).
72. Ding, C., Wang, X. & Li, M. Evaluation of six white-rot fungal pretreatments on corn stover for the production of cellulolytic and ligninolytic enzymes, reducing sugars, and ethanol. *Appl. Microbiol. Biotechnol.* **103**, 5641–5652 (2019).
73. Komiya, D. et al. Crystal structure and substrate specificity modification of acetyl xylan esterase from *Aspergillus luchuensis*. *Appl. Environ. Microbiol.* **83**, e01251–17 (2017).
74. Chambers, R. S. et al. An exo-beta-(1,3)-glucanase of *Candida albicans*: purification of the enzyme and molecular cloning of the gene. *J. Gen. Microbiol.* **139**, 325–334 (1993).
75. Nebreda, A. R., Vazquez, C. R., Villa, T. G., Villanueva, J. R. & del Rey, F. Heterogeneous glycosylation of the EXG1 gene product accounts for the two extracellular exo-beta-glucanases of *Saccharomyces cerevisiae*. *FEBS Lett.* **220**, 27–30 (1987).
76. Tamano, K. et al. The beta-1,3-exoglucanase gene *exgA* (*exg1*) of *Aspergillus oryzae* is required to catabolize extracellular glucan, and is induced in growth on a solid surface. *Biosci. Biotechnol. Biochem.* **71**, 926–934 (2007).
77. Bauer, S., Vasu, P., Persson, S., Mort, A. J. & Somerville, C. R. Development and application of a suite of polysaccharide-degrading enzymes for analyzing plant cell walls. *Proc. Natl Acad. Sci. USA* **103**, 11417–11422 (2006).
78. Wymelenberg, A. V. et al. The *Phanerochaete chrysosporium* secretome: database predictions and initial mass spectrometry peptide identifications in cellulose-grown medium. *J. Biotechnol.* **118**, 17–34 (2005).
79. Franco-Cirigliano, M. N. et al. *Streptomyces misionensis* PESB-25 produces a thermoacidophilic endoglucanase using sugarcane bagasse and corn steep liquor as the sole organic substrates. *Biomed. Res. Int.* **2013**, 584207 (2013).
80. Wessel, D. & Flüggel, U. I. A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal. Biochem.* **138**, 141–143 (1983).
81. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2**, 1896–1906 (2007).
82. Olsen, J. V. et al. Parts per million mass accuracy on an orbitrap mass spectrometer via lock mass injection into a C-trap. *Mol. Cell. Proteom.* **4**, 2010–2021 (2005).
83. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).

84. Cox, J. et al. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell. Proteom.* **13**, 2513–2526 (2014).
85. Tyanova, S. et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **13**, 731–740 (2016).
86. Borch, K. et al. Cellobiohydrolase variants and polynucleotides encoding same. US patent 10,036,050 (2014).
87. Alhifhi, A. & Williams, S. J. Unimolecular, Bimolecular, and Intramolecular Hydrolysis Mechanisms of 4-Nitrophenyl β -d-Glucopyranoside. *The Journal of Organic Chemistry*, **86**, 9530–9539 (2021).
88. McIlvaine, T. C. A buffer solution for colorimetric comparison. *J. Biol. Chem.* **49**, 183–186 (1921).
89. Koibuchi, K., Nagasaki, H., Yuasa, A., Kataoka, J. & Kitamoto, K. Molecular cloning and characterization of a gene encoding glutaminase from *Aspergillus oryzae*. *Appl. Microbiol. Biotechnol.* **54**, 59–68 (2000).
90. Vizcaino, J. A. et al. 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res.* **44**, D447–D456 (2016).

Acknowledgements

Funding by the Mercur Mercator Research Center Ruhr (Pr-2017-0020, to D.B., B.S., and M.K.) as well as the DFG (INST 20876/322-1, to M.K. and F.K.) is greatly acknowledged.

Author contributions

C.B., D.B., B.S., and M.K. conceived and designed the study; C.S., L.S., G.H., K.J., and F.W. performed and analyzed all chemical biology experiments. C.S. and K.J. conducted growth studies and protein expression. C.S. performed enzyme characterizations. H.S.O. provided the activity-based probes. L.S., C.S., G.H., and F.K. performed mass spectrometry and related data analysis. C.S., L.S., B.S., and M.K. wrote the manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s42003-022-04141-x>.

Correspondence and requests for materials should be addressed to Bettina Siebers or Markus Kaiser.

Peer review information *Communications Biology* thanks Jean-Guy Berrin and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Calvin Henard and Christina Karlsson Rosenthal. Peer reviewer reports are available.

Reprints and permission information is available at <http://www.nature.com/reprints>

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022, corrected publication 2022

Supplementary Information

Identification of fungal lignocellulose-degrading biocatalysts secreted by *Phanerochaete chrysosporium* via activity-based protein profiling

Christian Schmerling, Leonard Sewald, Geronimo Heilmann, Frederick Witfeld, Dominik Begerow, Kenneth Jensen, Christopher Bräsen, Farnusch Kaschani, Hermen S. Overkleeft, Bettina Siebers, Markus Kaiser

Content

Supplementary Figures	3 – 8
Supplementary References	9

a



b

```

3002168          MHSSFSLQDCPIRSLHPVTWTATPFNPASVPLAVRSPYLSAWLNQGGTALNAI
7dks.1.A beta-glucosidase -----
5npf.1.A Glucosylceramidase -----
4rhh.1.A Beta-xylosidase -----

3002168          TSTQSTFVMTSGPIDLTVNFI SPVE ----- PTDIVKQSLPFSYVILTATG
7dks.1.A beta-glucosidase  PNSWYTYTNKDLPVQLAVKQFSPIL ----- PYNKYKETSYPVAVFKWTAYN
5npf.1.A Glucosylceramidase  PNSWYTYTNKDLPVQLAVKQFSPIL ----- PYNKYKETSYPVAVFKWTAYN
4rhh.1.A Beta-xylosidase    EFRVATDTWKAG--DLTLTIINSPVKAVDPETASEEELKLLALVPAVIVEMTIDN

3002168          -VHOVQLOSOTPFGEVSDHIQGSAFYATLAHS--GTTYQ-----TGQDITVVR
7dks.1.A beta-glucosidase  EIVAAVMGNI---S-NDNEEWNGEYSIGVKKVPGVDISYKAKFVTTGDGSDLWH
5npf.1.A Glucosylceramidase  EIVAAVMGNI---S-NDNEEWNGEYSIGVKKVPGVDISYKAKFVTTGDGSDLWH
4rhh.1.A Beta-xylosidase    ---MRRI-DDTCPOLRGVCGRILGIASKD-----EGVRSAL

3002168          GHVRDPAVEYIVAGGKTQSRSLFEWSQF---SSVCAAISSFLGDYSNALSRAK
7dks.1.A beta-glucosidase  SW-DLPIMKFGGGD--K--WYKMYTKYFGKNGKNSFAILKEALNNYQWEKMI
5npf.1.A Glucosylceramidase  SW-DLPIMKFGGGD--K--WYKMYTKYFGKNGKNSFAILKEALNNYQWEKMI
4rhh.1.A Beta-xylosidase    CFYRGGCVT-AGM----DASFEYTRFF---HNIEEVGLYALEQAEVLKEQAE

3002168          SINTSDVLMFMKEIISSDGNVNTVDVIFPSWPIFELTNPNLCKYLLPLLEYO
7dks.1.A beta-glucosidase  DKRTNNMFGL-LACFDYNYEYETLDVRFYGSFPLVMLWPDIEKQVMRQFADTIN
5npf.1.A Glucosylceramidase  DKRTNNMFGL-LECFDYNYYETLDVRFYGSFPLVMLWPDIEKQVMRQFADTIN
4rhh.1.A Beta-xylosidase    -----KPIWVVNGG-EYRMMNT-DLITVDQLFEELKMNFWTVKNVLDIFYVERYS-

3002168          -ALGH-----NDGNDEAMP---VEESGNMLIMTLSYVQKTG--DRSLINSY
7dks.1.A beta-glucosidase  -IKIN-----AYD--WQNPNIWKDLNSKYVLLVYRDYVLTGKTDKEFLKYTW
5npf.1.A Glucosylceramidase  -IKIN-----AYD--WQNPNIWKDLNSKYVLLVYRDYVLTGKTDKEFLKYTW
4rhh.1.A Beta-xylosidase    SRPHYSSYELYGISGCFSEMT---HEQLVNWVLCAAVYIEQTK--DNAWRDRRI

3002168          TDDF----AGALANQTNLAIKGIVGIKAMSQIASLAGKSSVAANYSSIAASYVI
7dks.1.A beta-glucosidase  TYDT---WSMKGTSAYCGSLWLAALKAAQEIGKVLKDNEAYIKYNEWYKIAQC
5npf.1.A Glucosylceramidase  TYDT---WSMKGTSAYCGSLWLAALKAAQEIGKVLKDNEAYIKYNEWYKIAQC
4rhh.1.A Beta-xylosidase    TYDSL DVSIQARNNLYLAKCWAAYVALEKLEFRDVGKEELAAALAREQAEKCAF

3002168          -N-LFPQSVYEMHSGVPLDRHTYTKSDWSIWTAATTTAVRDLFISAVHSY
7dks.1.A beta-glucosidase  GD-ILPKDH-----
5npf.1.A Glucosylceramidase  GD-ILPKDH-----
4rhh.1.A Beta-xylosidase    -HEALRED-----

3002168          VRG
7dks.1.A beta-glucosidase  ---
5npf.1.A Glucosylceramidase  ---
4rhh.1.A Beta-xylosidase    ---

```

C

3002168 PKNMFFNAHHSVPVGAFAFASFTLGFP-----GKSGGLDLELGRPPRQNVY 67
 4C1P PFNPASVPLAVR--SPYLSAWLNQGSALTALNADWPR----FWTGSILGWA 45

3002168 IGVASLSQPGMYEVLPPFFEAGDDESKRYDIENPDPNPEKPKQILVPPFNEM 117
 4C1P GFIKVD---GTAYNFLGAPS-----MPCVTFQKSV---Q 73

3002168 IQREFHVSTDTWKAG----DLTFTIYSPVKSVPNPDTAKEEDLKFALVPA 163
 4C1P KSMTFTSTQSTFVMTSGPIDLTVNFLSPVEPTD-----LVKQSLPF 114

3002168 VIAELTIDNTK-GTSPRRAFFGFEGND-----PYTSMRRIDDTCPPLRG 206
 4C1P SYVTLTATSTDGKAHSVQLYTDISAEWVTGDNSLLANWSTAV--SSSTLV 162

3002168 VGQG-----RITAIIVSKHSDVRSALHFS--LEDILTTP 237
 4C1P HQVQLQSQTPFGEVSDHIQQGSAFYATLAHSGTTYQTGQDTVVRAQFVNS 212

3002168 LEENWTFGL-----GKVGALIMDTPAGM-KRTYQFAVCFYRSGYATA 278
 4C1P GTLPNTQDTRFRAVQDAWPVFAFAHDLGSVTGTSLPVVIAGHVDPAVE 212

3002168 -----GLDTSYFYTRFFKNIEEVGKYALDHIEALKERAFQSNQLIER- 320
 4C1P YIVAGGKTQSRSLFFWSQFSSVGAAISSFLGDYSNALSRAKTFDAKVQGD 312

3002168 -DWLSDDQKFMMAHAIRSYYGNTQLLEQEG-----KPIWVVNEGEY- 360
 4C1P ASKISADYASLVALSVRQAFGATEITASRNGDGSINTSDVLMFMKEISSD 362

3002168 RMMNTFDLTVDQLFFELKMNPNWTVKNVLDLYVERY---YYDRVRFPGEE 407
 4C1P GNVNTVDVIFPSWPIFLYTNPNLKGKYLPLFEYQATGQYPN----- 404

3002168 KEYPGGISFTHDMGVANTFSRPHYSAYELYGIDGCFSHMTHEQLVNWVLC 457
 4C1P -----KWSVHDMGAHYQPALGHNDGNDE-----AMPVEESGNMLIM 440

3002168 AAVYIEQTKDWAWRQEKLPILEQCLES MVNRDHPDPEKRNQVMGLDSTRT 507
 4C1P TLSYVQKTGDKSLINSYYNLLDQWTQFLITDS-----LVPA----- 476

3002168 MGGAEITTYDSL DVSLGQARNNLYLAGKCWAAYVALEKIFRDTGKEALAA 557
 4C1P ---NQISTDDFA----GALANQTNLAIKGIKAMSQIASLAGKSSVAA 519

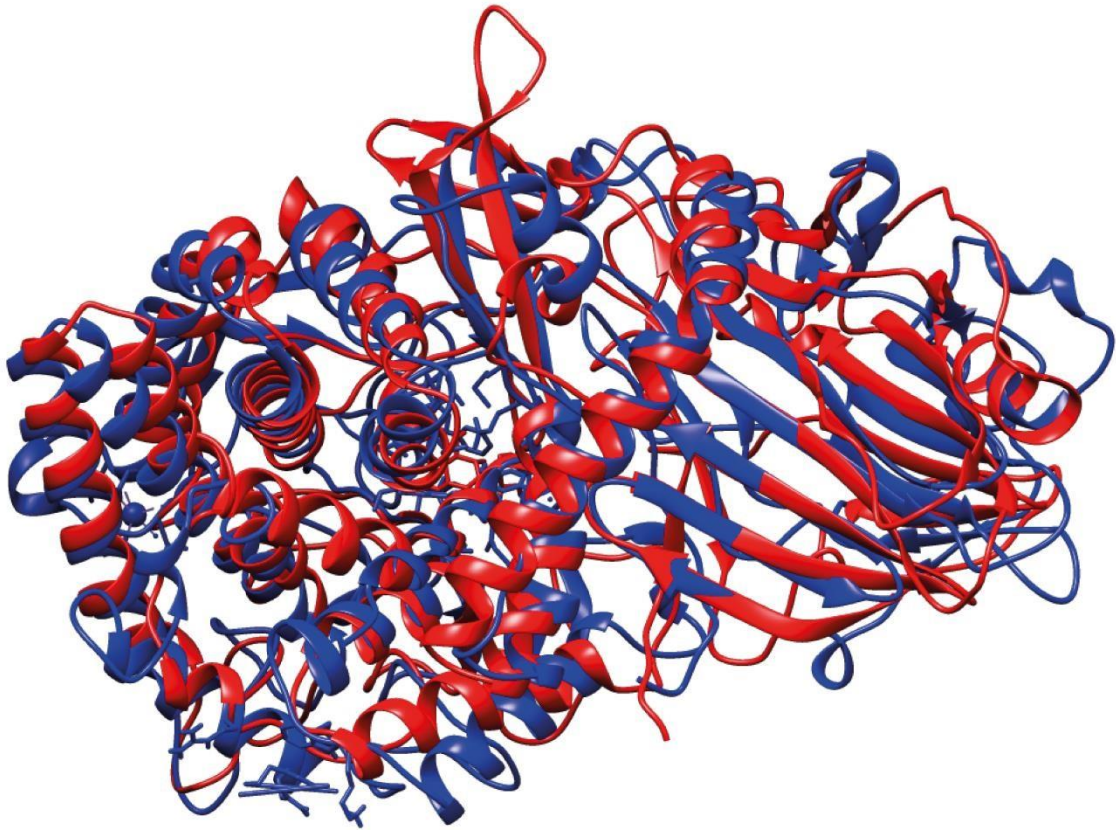
3002168 LAGEQAEEKCAATIVSY--VTEQGYIPAVMGE GND SKIIPAIEGLVFPYFT 605
 4C1P NYSSIAASYVTQWQGFATSKTGAHLTLSYG-NDASWGLA--YNLYGDKLL 566

3002168 NCHEALDPHGRFGEYIRALRKHLQYVLTEGICLFPDGGWKISSTSNNSWL 655
 4C1P GLNL-----FPQSVYEMH--SFG-----VPLDTRHTYTKS 594

3002168 SKIYLCQFIARRILGWKWDEAGAKADAAHVAVLTHPTLSVWSWSDQIIA- 704
 4C1P DWSIWTAIA----TTTA--VRDLFISAVHSYAADGKSAQ-PLGDWYETT 637

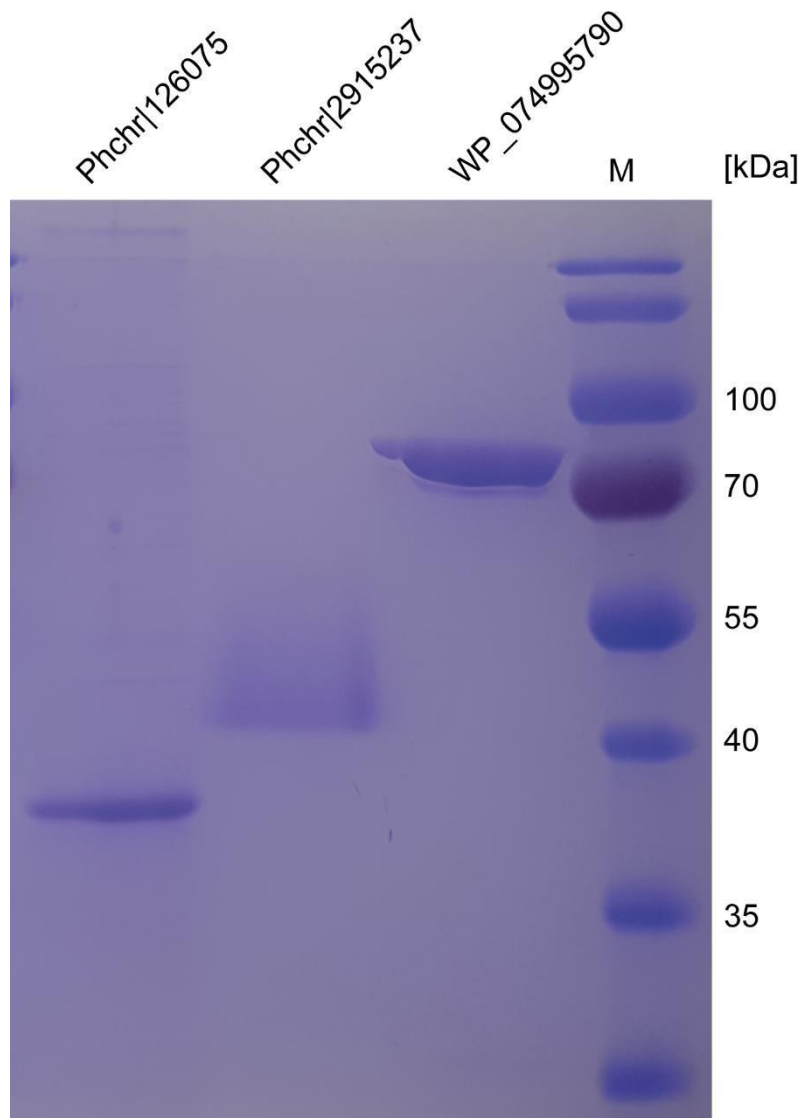
3002168 -GEISGSKYYPRGVT SILWLEEGK 727
 4C1P DGSVEGFRARPVVGGHLALVSLVI 661

d

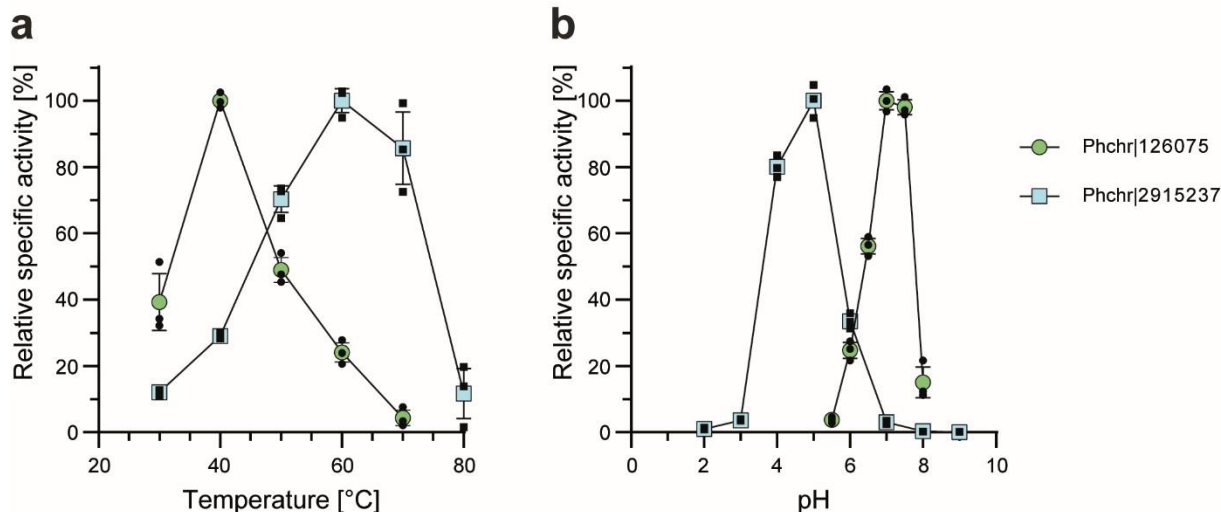


3002168 **SSSS--SSSHSSSSSH-SSSSHSSH----HHH-HHH-HSHH-S-HH**
4C1P **SSSSSHSSSHSSSSSSHSSH-HSSHHSSSHHSSSHHSSSH-**

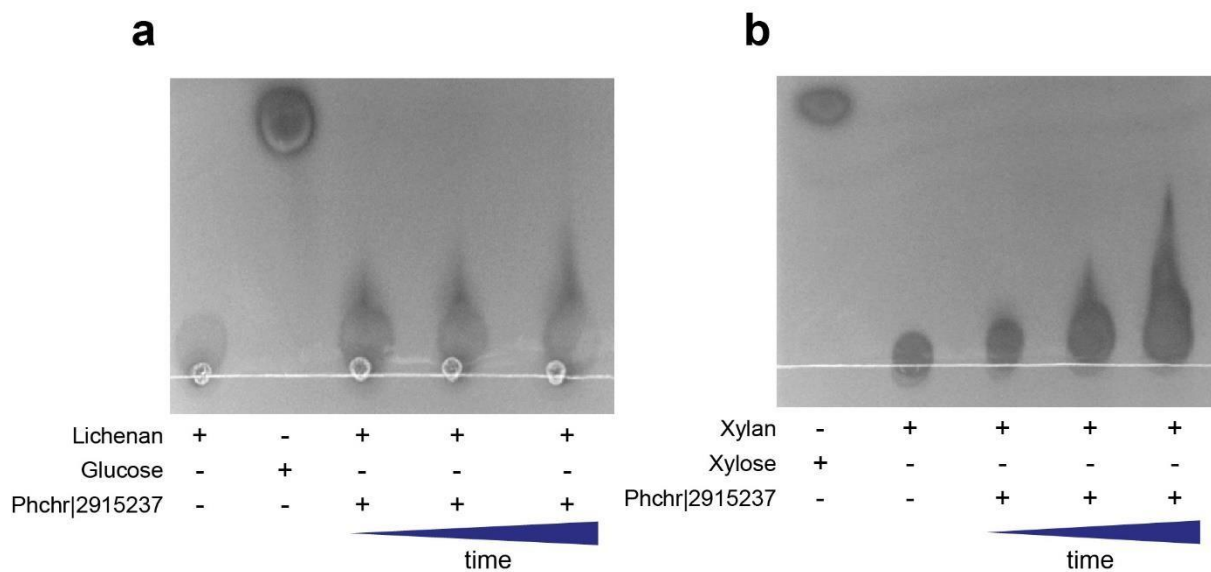
Supplementary Figure 1. Combined bioinformatics and structural analysis of the four domains of unknown function (DUF4964, DUF5127, DUF4965 and DUF1739) protein Phchr2|3002168. (a) The domain structure predicted by PFAM¹ and HMMER² shows the presence of four DUF domains (DUF4964, DUF5127, DUF4965 and DUF1739) in Phchr2|3002168. (b) Alignment of Phchr2|3002168 with different GH family proteins (pdb code 4RHH - *Geobacillus stearothermophilus*, pdb code 5NPF - *Thermoanaerobacterium xylanolyticum* and pdb code 7DKS - *Thermoanaerobacterium xylanolyticum* LX-11) predicted to be structurally similar by HHpred³. Phchr2|3002168 shows low sequence similarity to known GH52 (pdb codes 4RHH and 5NPF, 17.5 % and 15.1 % sequence identity, respectively) and GH116 (pdb code 7DKS, 14.1% sequence identity). Black shaded residue, same amino acid; grey shaded residue, similar amino acid. (c) Pairwise structural alignment by PDBeFold⁴ shows that Phchr2|3002168 and pdb code 4C1P share 70% of secondary structure elements. Yellow shaded residues indicate secondary structures that are the same in the predicted structure of Phchr2|3002168 and pdb code 4C1P. (d) Superimposition of the AlphaFold⁵ predicted structure of Phchr2|3002168 with GH52 (pdb code 4C1P, 16.1 % sequence identity) and a structure-based alignment show a high similarity in predicted tertiary structure. S = betasheet; H = alpha-helix.



Supplementary Figure 2. Expression and purification of Phchr2|126075 (*P. chrysosporium*), Phchr2|2915237 (*P. chrysosporium*) and WP_074995790 (*S. misionensis*, close homologue of Phchr2|3002168). Phchr2|126075 was heterologously expressed in *K. lactis*. Phchr2|2915237 was expressed in *Aspergillus oryzae*. WP_074995790 was heterologously overexpressed in *E. coli* Rosetta. Proteins (2 µg) were separated via SDS-PAGE and stained via Coomassie Blue. Marker (M): prestained PageRuler™ (Thermo Fischer scientific, USA).



Supplementary Figure 3. Temperature and pH optimum for the heterologously expressed proteins Phchr2|126075 and Phchr2|2915237 from *P. chrysosporium*. **(a)** Temperature and **(b)** pH optimum were determined for Phchr2|126075 with *p*NP-acetate and for Phchr2|2915237 with *p*NP-Glc as substrate using a citric acid/phosphate buffer in a discontinuous assay by determining the release of *p*-nitrophenol. All activity measurements were performed in triplicate ($n = 3$), mean values are shown and the error bars indicate the standard deviation (SD).



Supplemental Figure 4: Thin layer chromatography (TLC) of lichenan (**a**) and beech wood xylan (**b**) hydrolysis products formed by Phchr2|2915237. Lichenan (0.5 % (w/v)) or xylan (0.5 % (w/v)) was incubated with 10 μ g of Phchr2|2915237, and the reaction products after 1, 2 and 4 hours were separated by TLC and visualized. Lichenan and glucose (**a**) as well as xylan and xylose (**b**) served as standards. Hydrolysis products and standards were separated on aluminum sheet silica gel 60/kieselguhr F254 plates using ethyl acetate, methanol and H₂O (68:23:9, v/v/v) as solvent and stained using a KMnO₄ solution.

Supplementary References

- 1 Mistry, J. *et al.* Pfam: The protein families database in 2021. *Nucleic Acids Res* **49**, D412-D419, (2021).
- 2 Potter, S. C. *et al.* HMMER web server: 2018 update. *Nucleic Acids Res* **46**, W200W204, (2018).
- 3 Gabler, F. *et al.* Protein Sequence Analysis Using the MPI Bioinformatics Toolkit. *Curr Protoc Bioinformatics* **72**, e108, (2020).
- 4 Krissinel, E. & Henrick, K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr* **60**, 2256-2268, (2004).
- 5 Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589, (2021).

5. Summary

The first part (Part I) dealt with the metabolism in the third domain of life, the Archaea. One characteristic feature of Archaea is their membrane composition that consists of isoprenoid side chains that are ether linked to G1P, in contrast to bacterial and eukaryotic membrane lipids that consist of FA side chains that are ester linked to G3P. 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 FAs in Archaea. Since one of the main functions of FAs as key part of membrane phospholipids and thus of cell structure in Bacteria and Eukaryotes is substituted by isoprenoids in Archaea, the presence and function of FAs in Archaea remains unknown. In this work it was shown that the archaeal model organism *S. acidocaldarius* contains a fully functional largely bacterial-like β oxidation pathway for FA degradation which was biochemically characterized in detail. Furthermore, a potential novel pathway for FA synthesis acting completely independent from β oxidation was identified and characterized. Furthermore, this work elucidated the utilization of glycerol by *S. acidocaldarius* and compares its breakdown with known pathways from Bacteria and Eukaryotes and discusses how thermoacidophilic Archaea such as *S. acidocaldarius* have evolved strategies to mitigate metabolic instabilities associated with high temperatures.

In chapter 3.1 the FA metabolism of *S. acidocaldarius* was elucidated by characterizing homologous enzymes of a bacterial-like β oxidation encoded by gene cluster *saci_1103-1126*. Here the functions of multiple acyl-CoA synthetases (FA activation), as well as acyl-CoA dehydrogenases (ACAD), bifunctional 3(S)-hydroxyacyl-CoA dehydrogenases/enoyl-CoA hydratases (HCDH/ECH) and β -ketothiolases (KT) were confirmed via heterologous production and characterization. All single enzymes put together *in vitro* were able to fully degrade short to medium chain acyl-CoA esters to acetyl-CoA in a functional β oxidation cycle. Thus, it was shown that *S. acidocaldarius* encodes enzymes for fully functional β oxidation spiral degrading acyl-CoAs up to chain lengths of C8 and this pathway shows some unusual features with respect to the ETF, the HCDH/ECH bifunctional enzyme and the “archaeal type” KTs.

In contrast to what was previously suggested this β oxidation cycle was not reversible and is thus not responsible for FA synthesis in *S. acidocaldarius*. Instead, a potential novel pathway was studied, that would run independent of this β oxidation cycle. This putative anabolic pathway for FA synthesis shows a 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). Candidates for all these

enzymes were purified and biochemically characterized and were shown to be active against acyl-CoA intermediates with a chain length up to C8. Even though the function of FAs in Archaea is unknown this work provides a basic understanding of FA metabolism in Archaea and functions as a starting point for understanding the presence and significance of FA in Archaea.

Chapter 3.2 presents the elucidation of a pathway for glycerol degradation in *S. acidocaldarius*. Growth curves, transcriptomic, proteomic and metabolomic data show that *S. acidocaldarius* utilizes glycerol via the concerted activity of GK and G3PDH. Biochemical characterization shows the ATP dependent phosphorylation of glycerol by GK and quinone dependent oxidation of intermediately formed G3P to DHAP. In total, this work demonstrates that *S. acidocaldarius* utilizes glycerol as growth substrate and employs a conserved classical GK (homologous to GlpK) for glycerol phosphorylation. However, G3P oxidation is catalysed by a GlpA-subunit like FAD-dependent G3PDH (Saci_2032) lacking the B and C subunits of the classical, bacterial GlpABC complex. Instead, it shows an unusual type of membrane association facilitated by a small CoxG-related protein and thus, the recruitment of the GlpA-like Saci_2032 represents a novel function of CoxG homologues in Archaea.

Chapter 3.3 discusses how Archaea circumvent metabolic problems associated with growth at high temperatures such as formation and accumulation of toxic by/side products. Adaption strategies range from detoxification reactions that remove toxic by/side-products to the usage of different metabolic routes of common pathways that e.g., skip labile intermediates. In total it seems like (hyper)thermophilic Archaea employ three base strategies for mitigation: The concentration of unstable metabolites is reduced, different pathway topologies allow to circumvent labile intermediates, and damaged metabolites are continuously removed via novel metabolic pathways. This review furthermore suggests a significant influence of thermolabile intermediates as well as of promiscuous enzymes and pathways on the evolution of the metabolic network in (hyper)thermophilic Archaea.

The second part of this work in chapter 3.4 showcases the potential of using activity-based protein profiling (ABPP) biocatalyst screening technology in complex and biotechnologically relevant experimental settings, with the aim to rapidly identify promising lignocellulose degrading enzymes from complex protein samples: the secretome of the white rot fungus *P. chrysosporium*. The newly established workflow allowed for identification of a set of enzymes involved in lignocellulose metabolism while paying special attention to the often-neglected substrate bound proteins. The workflow allowed for enrichment of a wide array of serine hydrolases (SH)

and β -cleaving glycoside hydrolases (GH). Protein production and biochemical characterization confirmed the function of two labeled and previously uncharacterized proteins Phchr2|126075 and Phchr2|2915237 as an acetyl-xylan esterase and endo cleaving β -glucanase, respectively. In addition, the results suggest that an unknown class of enzymes, DUF5127 domain proteins, could belong to a class of GH based on sequence analysis and enzyme assays. In total it was shown that a new established ABPP workflow can successfully be used for prescreening of enzymes and the involvement of these enzymes in fungal lignocellulose degradation was established. In theory this workflow could now be employed using different cultivation methods or different target organisms and could be advanced and refined further by the synthesis of novel ABPs.

6. Zusammenfassung

Der erste Teil dieser Arbeit befasst sich mit dem Stoffwechsel in der dritten Domäne des Lebens, den Archaeen. Ein charakteristisches Merkmal der Archaeen ist ihre Membranzusammensetzung, die aus Isoprenoid-Seitenketten besteht, die mit G1P über Etherbindungen verknüpft sind, im Gegensatz zu bakteriellen und eukaryotischen Membranlipiden, die aus Fettsäure-Seitenketten bestehen, die mit G3P verestert sind. Dieser Unterschied der Membranstrukturen zwischen Archaeen und Bakterien wird als "Lipid Divide" bezeichnet und wirft Fragen bezüglich der Existenz und Funktion von Fettsäuren in Archaeen auf. Da eine der Hauptfunktionen von Fettsäuren als Bestandteil von Membranphospholipiden und damit der Zellstruktur in Bakterien und Eukaryoten in Archaeen durch Isoprenoide ersetzt wird, ist die Funktion von Fettsäuren in Archaeen unbekannt. In dieser Arbeit wurde gezeigt, dass der archaeelle Modellorganismus *S. acidocaldarius* eine voll funktionsfähige, weitgehend bakterienähnliche β -Oxidation für den Abbau von Fettsäuren enthält, die biochemisch im Detail charakterisiert wurde. Darüber hinaus wurde ein potenzieller neuer Weg zur Fettsäure-Synthese identifiziert und charakterisiert, der völlig unabhängig zur β -Oxidation funktioniert. Außerdem wurde in dieser Arbeit die Verwertung von Glycerol durch *S. acidocaldarius* aufgeklärt und dieser Abbau mit bekannten Stoffwechselwegen von Bakterien und Eukaryoten verglichen und erörtert, wie thermoacidophile Archaeen wie *S. acidocaldarius* Strategien entwickelt haben, um metabolische Instabilitäten im Zusammenhang mit hohen Temperaturen abzumildern.

In Kapitel 3.1 wurde der FA-Stoffwechsel von *S. acidocaldarius* durch die Charakterisierung homologer Enzyme einer bakterienähnlichen β -Oxidation aufgeklärt, die von einem Gencluster (*saci_1103-1126*) kodiert werden. Dabei wurden die Funktionen von mehreren Acyl-CoA-Synthetasen sowie Acyl-CoA-Dehydrogenasen (ACAD), bifunktionellen 3(S)-Hydroxyacyl-CoA-Dehydrogenasen/Enoyl-CoA-Hydratasen (HCDH/ECH) und β -Ketothiolasen (KT) durch heterologe Produktion und Charakterisierung bestätigt. Alle Enzyme zusammen waren *in vitro* in der Lage, kurz- bis mittelkettige Acyl-CoA-Ester in einem funktionellen β -Oxidationszyklus vollständig zu Acetyl-CoA abzubauen. Somit wurde gezeigt, dass in *S. acidocaldarius* Enzyme für eine voll funktionsfähige β -Oxidationsspirale kodiert sind, die Acyl-CoAs bis zu einer Kettenlänge von C8 abbauen können. Dieser β -Oxidationszyklus zeigte einige ungewöhnliche Merkmale in Bezug auf das ETF, das bifunktionelle HCDH/ECH-Enzym und auf archaeelle KTs.

Im Gegensatz zu früheren Annahmen war diese β -Oxidation nicht reversibel und ist daher nicht für die Fettsäure-Synthese in *S. acidocaldarius* verantwortlich. Stattdessen wurde ein möglicher neuer Stoffwechselweg untersucht, der unabhängig von diesem β -Oxidationszyklus abläuft. Dieser mutmaßliche anabole Weg für die Fettsäure-Synthese zeigt einen Mosaikcharakter mit Ähnlichkeiten sowohl zu bakteriellen (fabG, MaoC) als auch zu eukaryontischen (MDR-Reduktase) Merkmalen, gemischt mit einzigartigen archaellen Eigenschaften (DUF35-Domäne/KT-Komplexe, ACP-Unabhängigkeit). Die Kandidaten für alle diese Enzyme wurden gereinigt und biochemisch charakterisiert und ihre Funktion bestätigt.

In Kapitel 3.2 wird die Aufklärung eines Weges für den Glycerolabbau in *S. acidocaldarius* charakterisiert. Wachstumskurven, transkriptomische, proteomische und metabolomische Daten zeigten, dass *S. acidocaldarius* Glycerol über die Aktivität einer GK und G3PDH verwertet. Die biochemische Charakterisierung zeigte die ATP-abhängige Phosphorylierung von Glycerol durch GK und die Chinon-abhängige Oxidation von intermediär gebildetem G3P zu DHAP. Insgesamt zeigt diese Arbeit, dass *S. acidocaldarius* Glycerol abbaut und eine konservierte klassische GK (homolog zu GlpK) für die Glycerolphosphorylierung einsetzt. Die G3P-Oxidation wird jedoch von einer GlpA-Untereinheit katalysiert, die der FAD-abhängigen G3PDH (Saci_2032) ähnelt und der die Untereinheiten B und C des klassischen, bakteriellen GlpABC-Komplexes fehlen. Stattdessen zeigt es eine ungewöhnliche Art der Membranassoziation, die durch ein kleines CoxG-verwandtes Protein vermittelt wird, und somit stellt die Rekrutierung des GlpA-ähnlichen Saci_2032 eine neue Funktion von CoxG-Homologen in Archaeen dar.

In Kapitel 3.3 wird erörtert, wie Archaeen metabolische Probleme umgehen, die mit dem Wachstum bei hohen Temperaturen verbunden sind, wie z. B. die Bildung und Anhäufung von toxischen Nebenprodukten. Die Anpassungsstrategien reichen von Entgiftungsreaktionen, die toxische Nebenprodukte entfernen, bis zur Nutzung verschiedener Stoffwechselwege, die z.B. labile Zwischenprodukte überspringen. Insgesamt scheinen (hyper)thermophile Archaeen drei grundlegende Strategien zur Schadensbegrenzung zu verwenden: Die Konzentration instabiler Metaboliten wird reduziert, verschiedene Stoffwechselwege ermöglichen die Umgehung labiler Zwischenprodukte, und geschädigte Metaboliten werden kontinuierlich über neue Stoffwechselwege entfernt. Diese Übersicht deutet außerdem auf einen signifikanten Einfluss von thermolabilen Zwischenprodukten sowie von promiskuitiven Enzymen und Stoffwechselwegen auf die Evolution des metabolischen Netzwerks in (hyper)thermophilen Archaeen hin.

Der zweite Teil dieser Arbeit in Kapitel 3.4 zeigt das Potenzial von ABPP für das Screening von Biokatalysatoren in einem komplexen und biotechnologisch relevanten experimentellen Umfeld, mit dem Ziel, vielversprechende lignozelluloseabbauende Enzyme aus einer komplexen Proteinprobe zu identifizieren: dem Sekretom des Weißfäulepilzes *P. chrysosporium*. Der neu entwickelte Arbeitsablauf ermöglichte die Identifizierung einer Reihe von Enzymen, die am Lignozellulosestoffwechsel beteiligt sind, wobei den oft vernachlässigten substratgebundenen Proteinen eine besondere Aufmerksamkeit geschenkt wurde. Dieser Arbeitsablauf ermöglichte die Anreicherung einer breiten Palette von Serinhydrolasen (SH) und β -spaltenden Glykosidhydrolasen (GH). Die Proteinproduktion und biochemische Charakterisierung bestätigte die Funktion von zwei zuvor nicht charakterisierten Proteinen Phchr2|126075 und Phchr2|2915237 als Acetyl-Xylan-Esterase bzw. endospaltende β -Glucanase. Darüber hinaus legen die Ergebnisse nahe, dass eine unbekannte Klasse von Enzymen, DUF5127-Domänenproteine, auf der Grundlage von Sequenzanalysen und Enzymtests zu einer Klasse von GH gehören könnten. Insgesamt konnte gezeigt werden, dass ein neu etablierter ABPP-Arbeitsablauf erfolgreich für das Vorscreening von Enzymen eingesetzt werden kann und dass diese Enzyme am Lignocellulose-Abbau durch Pilze beteiligt sind. Theoretisch könnte dieser Arbeitsablauf nun mit anderen Kultivierungsmethoden oder anderen Zielorganismen eingesetzt und durch die Synthese neuartiger ABPs weiterentwickelt und verfeinert werden.

Acknowledgements

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

First, my sincere and heartfelt gratitude goes to Prof. Dr. Bettina Siebers, for offering me the opportunity to conduct my PhD studies in her workgroup and to work under her tutelage. All discussions, her suggestions and constructive criticism were very helpful to me and will shape me for the rest of my scientific life. I would like to greatly thank her for her constant support, trust and understanding.

I would like to express my gratitude to Dr. Christopher Bräsen, who so patiently guided me and taught me so much with his exceptional knowledge. I deeply thank him for helping me whenever I was in need especially for experimental advice and detailed proofreading of the publications and this thesis. Without his instruction, this thesis would not have reached its present form.

I own my sincere gratitude to Prof. Dr. Mark Kaiser 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. Dr. Markus Kaiser, Prof. Dr. Sonja-Verena Albers, Prof. Dr. Jörn Kalinowski, Prof. Dr. Oliver J. Schmitz, Dr. Sven Meckelmann, Dr. Sabrina Ninck, Leonard Sewald, Till Kessenbrock, Dr. Tobias Busche, Jan Bost and Paul Görs. I also thank the Volkswagen Stiftung, the Stiftung Mercator and the BMBF for their financial support of this work.

I would like to extend thanks to Thomas Knura and Sabine Dietl for their technical assistance and I want to especially mention my colleague Xiaoxiao Zhou for her support and amazing teamwork.

And as the story goes and this chapter finally comes to an end, I would like to thank my colleagues and friends Laura Kuschmierz, Carmen Peraglie, Xiaoxiao Zhou, Parppim Boonpranomsri, Simon Kleimann, Hanna Martin and Claudia Szymanski for their warm hearts, for their encouragement and for their kindness. I hope that I will be able to repay what was given to me and I am eternally grateful.

I would also like to thank and acknowledge both my past and present colleagues: Carsten Schroeder, Svenja Höfmann, Ravi Shankar Ojha; Thomas Klaus, Larissa Schocke, Christina Stracke, Benjamin Meyer, Lu Shen. Promise Akua Dziwornu, and Jan-Niklas Tum. My deepest

thanks also goes to all my students in general for their extraordinary hard work, for giving me the opportunity to teach and for extending me the honor of accompanying them in the early stages of their own career.

Finally thank you my friends and Jana for always being there for me. We have always known that there is no help for us but from one another, and that no hand will save us if we do not reach out our hand.

I cannot express my gratitude with words, but they are all I have so I will use them now: Each of you I thank you for your kindness and advice, and for your friendship. I am grateful for the time spent together and I would not have liked to walk with anyone else.

Erklärung der selbstständigen Verfassung der Dissertation

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

Fatty acid and glycerol metabolism in *Sulfolobus acidocaldarius* and establishing new methods for identification of novel biocatalysts in the white-rot fungus *Phanerochaete chrysosporium*

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.

Essen, 09.03.2023

Ort, Datum

C. Schmerling

(Christian Schmerling)