Compatible Solutes and novel Biocatalysts in (hyper)thermophilic Archaea: From Identification and Physiological Significance to Application in Biotechnology

Dissertation

zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

- Dr. rer. nat. -

Arbeitskreis für Molekulare Enzymtechnologie und Biochemie

Abteilung für Umweltmikrobiologie und Biotechnologie

Fachbereich Chemie

der

Universität Duisburg-Essen

vorgelegt von

Christina Stracke

November 2020

DOI: 10.17185/duepublico/73873
URN: urn:nbn:de:hbz:465-20250107-144906-6

Die vorliegende Arbeit wurde im Zeitraum von Mai 2014 bis September 2020 im Labor von Frau 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: 14.01.2021

Gutachter:	Prof. Dr. Bettina Siebers
	Prof. Dr. Peter Bayer
Vorsitzender:	Prof. Dr. Christian Mayer

Für meine verstorbenen Eltern,

in Gedenken und tiefer Verbundenheit.

"Science knows no country,

because knowledge belongs to humanity,

and is the torch which illuminates the world…"

Louis Pasteur

Table of content

1 Introduction

1.1 Archaea

The Archaea was recognized as third domain of life 40 years ago and the deep phylogenetic topology of the tree of life is in center of long-time research. Molecular 16S/18S rRNA-based phylogeny divided the Bacteria, Archaea, and Eukarya into three district domains [1]. This Woeseian three-domain tree hypothesis, with the Eukarya evolving as a sister clade to the Archaea, has been adopted as the paradigm for the universal tree of life in textbooks for years (Figure 1A). According to this taxonomy, Archaea encompassed two Phyla, Euryarchaeota and Crenarchaeota. The phylum of Euryarchaeota includes diverse methanogens, extremely halophiles, thermoacidophiles, and a few (hyper)thermophiles [2]. In contrast, the phylum of Crenarchaeota includes mainly (hyper)thermophiles [2]. Commonly, these representatives thrive in anaerobic niches, salt lakes, and marine hydrothermal systems or solfataras. Many of these habitats represent extreme environments concerning temperature, osmotic pressure, and pH and remind on the conditions of the early earth. However, Archaea are ubiquitous, also present in many mesophilic habitats, ranging from freshwater lakes and soils to the human oral- or intestinal-tract [3, 4].

In the last decades, new insights accompany the advances of phylogenetic reconstruction approaches. Increasingly available methods in the field of genomics and metagenomics have been used to identify new archaeal lineages and demonstrate the ancient divergence process and the deep phylogenetic relationship. Not least by these methods, several sister phyla of the Crenarchaeota have been identified, established a monophyletic superphylum termed as TACK (including **T**haum-, **A**ig-, **C**ren-, and **K**orarchaeota, and currently expanding with Bathy-, Verstraete, and Geoarchaeota) [5]. The genomic contents reconstructed from TACK lineages by advanced phylogenomic approaches have just brought new visions to the understanding of the tree of life [6-8]. Recently, two novel archaeal superphyla were discovered, the DPANN superphylum (including **D**iapherotrites, **P**arv-, **A**enigm-, **N**ano-, and **N**anohaloarchaeota), and the Asgard superphylum (including Loki-, Thor- Odin-, and Heimdallarchaeota), which is phylogenomically closest to Eukaryotes [5, 9-11]. Culture-independent genomics of Asgard Archaea revealed several genes formerly believed to be eukaryotic specific. For instance, the reported eukaryal signature proteins (ESPs), offer valuable information about early events during eukaryogenesis [7, 8]. These findings strongly support a new two-domain taxonomy, in which the Eukarya branches as a sister group of the Asgard by sharing a common archaeal ancestor, while Bacteria were classified as an independent domain of life (Figure 1B).

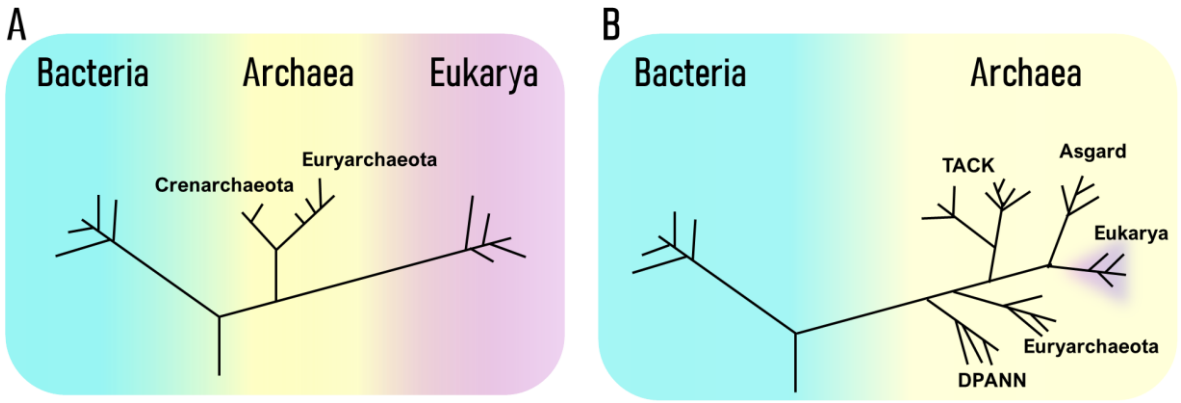


Figure 1: Overview of the three-domain and two-domain tree topology. The three-domain tree of life is based on phylogenetic 16S/18S rRNA sequence comparison (A)**.** The new two-domain tree of life is based on phylogenomic approaches (B).

Besides expanding knowledge of the diversity of life by genome-resolving approaches, the evolutionary history of Archaea revealed that they exhibit a “mosaic” of bacterial and eukaryal features. The unicellular lifestyle, lack of organelles, cell size, and shape as well as the DNA structure (e.g., one circular chromosome, operon structures, plasmids) resembles that of Bacteria. However, most mechanisms involved in information processing (e.g., replication, transcription, repair, and translation) are generally regarded as fewer complex versions of that found in Eukaryotes [12-14]. For instance, the archaeal transcription machinery consists of a multi-subunit RNA-Polymerase (RNAP) and general transcription factors (GTFs) which are homologs of the eukaryotic RNA-Polymerase II, TATA-binding protein (TBP), and the transcription factor TFIIB (TFB) [12, 15]. Therefore, in comparison to Bacteria, Archaea and Eukaryotes share a more complex transcription machinery [12, 15]. Also, in respect to protein modification systems (e.g., protein phosphorylation, protein glycosylation) Archaea are more similar to Eukaryotes [12, 13, 16].

However, Archaea possess some unique features, especially the chemical and structural composition of archaeal membranes differ from those found in Bacteria and Eukarya in several ways. Thus, murein the typical peptidoglycan of the bacterial cell walls is absent in Archaea. Instead, other features are observed such as pseudomurein (e.g., in *Methanosphaera*, *Methanothermus*; N-acetylglucosamine and N-acetyltalosaminuronic acid) [17, 18]. Also, many Archaea has a surface layer (S-layer) (e.g., *Thermopro*teales, *Sulfolobales*), which is composed of a monomolecular paracrystalline array of proteins and glycoproteins, and in some cases, they have no cell envelops at all (e.g., in *Thermoplasmatales*) [17, 19]. Additionally, the membrane lipids of Archaea are unique. In bacterial and eukaryal counterparts the hydrophobic moieties of the membrane lipids are straight chains of saturated or unsaturated fatty acids, which are ester-linked to *sn*-glycerol-3-phosphate. In contrast, archaeal membrane lipids are isoprenoid chains with occasional cyclopentane/cyclohexane rings; these chains are

ether-linked to *sn*-glycerol-1-phosphate [20, 21]. Furthermore, Archaea possess several unique metabolic features, many of which differ from the classical bacterial or eukaryal metabolism [22]. One individual type of pathway is methanogenesis. In this distinctive energy-generating process, methanogens use H_2 + CO_2, formate, methylated C1 compounds, or acetate as energy and carbon sources to produce methane as a primary end product of their metabolism [23, 24]. Notably, also central carbohydrate metabolism (CCM) of Archaea revealed great metabolic diversity. For instance, Archaea employ modified versions of the Embden-Meyerhof-Parnas (EMP) and Entner-Doudoroff (ED) pathway for hexose degradation [22, 25]. In comparison, the archaeal pathways can be distinguished by its numerous unique enzymes [22, 26], and in contrast to Bacteria, the archaeal ED pathway is branched into a semiphosphorylative (sp) and a non-phosphorylative (np) branch [25]. Noteworthy examples for individual archaeal enzymes are, e.g., ADP or PP_i dependent glucokinases and phosphofructokinases or the non-phosphorylating glyceraldehyde-3-phosphate (GAP) -dehydrogenase (GAPN) and the GAP: ferredoxin oxidoreductase (GAPOR), for the conversion of GAP to 3-phosphoglycerate (3PG) without ATP formation [27-32]. In addition, the classical pentose phosphate pathway is only particularly present (if at all) and most Archaea utilize the reversed ribulose monophosphate pathway for pentose generation [33]. Also, pentose degradation significantly differs from that known from bacterial model organisms [34, 35].

Besides this great metabolic diversity, especially in-depth study of (hyper)thermophilic Archaea gained particular scientific interest due to their fascinating extreme lifestyle and very stable enzymes. Their natural environments resemble the conditions used in many specific biotechnologies. Therefore, extremophiles present a potentially valuable resource for the exploitation of novel products and/or unique enzymes for the development of new industrial bioprocesses [26, 36-38]. Notably, extremophilic proteases, esterases, and lipases found applications in food-, textile-, detergent-, and fine chemical industry [39-41]. Mainly, this biotechnological potential as well as the high demand for novel efficient biocatalysts or bioproducts played a significant role in the research of the present thesis. Additionally, the outstanding adaption strategies of hyperthermophilic Archaea, as well as their ability to accumulate compatible solutes as stress response, were addressed in this work. The application of compatible solutes and the transfer of their exceptional protecting properties on biomolecules seems to be beneficial in many fields and open versatile applications for the cosmetic-, pharmaceutical-, medical-, and food industry. Also, this study provides a detailed analysis of involved factors, metabolic pathways, and respective enzymes in archaeal stress response in one of the most popular model organisms of the Crenarchaeota, i.e., *Sulfolobus acidocaldarius*.

1.2 *Sulfolobus acidocaldarius*

The Crenarchaeon *S. acidocaldarius* was firstly isolated from acidic solfatara fields in the Norris Geyser Basin of the Yellowstone National Park (Wyoming, USA) in the early seventies [42]. It is an obligate aerobe and grows naturally in sulfur-containing springs with a pH of 2-3 and temperatures between 70-80°C; hence it is defined as thermoacidophile. *Sulfolobus* species gain their energy chemolithoautotrophically by oxidizing Sulphur granules around the solfatara, thermal soils, or hot springs, generating sulphuric acid and thereby lowering the pH. However, chemoheterotrophic growth, using sulfur to oxidize simply reduced carbon compounds was only observed in the presence of oxygen. Nowadays, several strains have been reported in geographically distant thermal areas worldwide (Figure 2). For instance, *S. acidocaldarius* was reported in the Jigokudani thermal field on the island of Hokkaido, Japan, and in several hot springs in the Rotorua, Wai-O-Tapu region, New Zealand [43-46]. Also, in Germany, it has been reported in thermoacidic soils of self-heating mining wastes near Ronneburg [45, 47].

Figure 2: Hot vent in the geothermal area Wai-O-Tapu of New Zeeland. The inlay shows a collection of *S. acidocaldarius* cell, which naturally occur there.

The cells of *S. acidocaldarius* are highly irregular in shape, often lobed, but occasionally spherical. As mentioned above, the membrane composition of archaea differs, and membrane lipids are unique. Thus, cell envelop of *S. acidocaldarius* is composed of glycoprotein subunits and the cell membrane consists of isopranyl ether lipids, 95% tetraethers with cyclopentane rings (glycerol-dialkylglycerol tetraethers, GDGT and glycerol-dialkyl-nonitol tetraethers, GDNT), and 5% diethers [20, 21, 48]. Furthermore, flagella and pilus-like structures form the so-called archaellum [49, 50]. Also, unique quinones were found, e.g., caldariellaquinone and sulfoquinone [48]. Accordingly, *Sulfolobus* spp. possess an unbending and rigid membrane, which is assumed to be an essential feature for life under low pH and high temperature [51, 52].

So far, *S. acidocaldarius* and *S. solfataricus*, are the two best-described members of this genus [42, 53-55]. Due to the fact, that both can be grown under laboratory conditions and even systems for genetic manipulation have been developed, makes them to popular extremophilic model organisms in several research fields [56-61]. Detailed studies in case of genomics, proteomics, and metabolomics could be addressed so far [49, 50, 62-64]. Metabolic studies revealed *S. solfataricus* as most diverse species among them. Therefore, under laboratory conditions *S. solfataricus* has a high metabolic potential and is able to utilize a broad number of growth substrates including monosaccharides (D- glucose/galactose, D- and L- fucose, D-fructose, D/L-arabinose and D-xylose), disaccharides (cellobiose, maltose, sucrose, trehalose and lactose), poly- and oligo-saccharides (cellulose, starch and dextrin), amino acids (glutamate), peptides (tryptone), and alcohols (ethanol, phenol) [65-67]. Besides, this diversity is represented by high numbers of mobile genetic elements in the genome (insertion sequence and miniature inverted-repeat transposable elements, shortened IS- and MITE-like elements) [68]. This leads to consequential genome instability in *S. solfataricus* [68]. In contrast, the *S. acidocaldarius* genome is different, it is smaller in size (~ 25%) and lacks the genes for several sugar transporters, resulting in a limited growth on a smaller number of carbon sources, e.g., D-glucose, L-arabinose, D-xylose, sucrose, maltotriose, dextrin, and starch, but is well adapted to grow on peptides and amino acids [60, 69-71]. However, *S. acidocaldarius* is less diverse in metabolic function, but due to its significant smaller genome size and low mutational rates, it is much more stable. Thus, *S. acidocaldarius* is useful for genetic manipulations and modifications, which played a crucial role in the present work, to obtain detailed insights into the adaptation mechanisms and stress response in this archaeal model organism. However, it is known that the ATP-dependent hexokinase provides sugar phosphates for the generation of pentoses (RuMP, reverse ribulose monophosphate pathway), and glucose-6-phosphate (G6P) serves as building block for synthesis of the carbon storage compound glycogen. Glycogen can be further utilized for trehalose formation [72-74], which is regarded to be accumulated as stress response. So far, trehalose is the only reported compatible solute in *Sulfolobus* spp [75-77]. In comparison to Bacteria and Eukarya, in Archaea only a limited number of related research on the trehalose metabolism and its function as compatible solute is available. Therefore, this thesis provides detailed results on the involved pathways and enzymes in trehalose biosynthesis in *S. acidocaldarius*.

1.3 Trehalose biosynthesis

Microorganisms, including those thriving under hostile conditions like (hyper)thermophiles, thermoacidophiles, or extreme halophiles, require efficient adaptation strategies and need to respond to environmental changes and various stresses. Besides known intrinsic factors for the protection of cellular components, e.g., structural and mechanistic adaptation of DNA, proteins, and membranes, also extrinsic factors like compatible solutes are essential for the adaptation and stress response. Compatible solutes are low molecular weight, highly water-soluble compounds, that do not interfere with the cell metabolism and can, therefore, be accumulated in high amounts [78, 79]. Either the accumulation occurs via up-take from the environment or via *de-novo* synthesis by the cell. The diversity of compatible solutes is large but falls into a few different chemical categories such as (i) amino acids (e.g., α-glutamate, β-glutamate) and amino acid derivatives (e.g., glycine-betaine, ectoine, hydroxyectoine) as well as (ii) sugars (e.g., trehalose, mannosylglycerate), polyols (e.g., glycerol) and derivatives [80]. Some of these small organic molecules are widespread in nature, while others are specific and restricted to only a small number of microorganisms. The nonreducing disaccharide trehalose (α-D-glucopyranosyl-1,1-α-D-glucopyranoside) is widespread and essential in a large number of microorganisms. It is widely distributed in bacteria and eukaryotes where it functions as compatible solute in general stress protection against heat shock, osmotic stress, desiccation, and cessation of growth [81]. Furthermore, trehalose is used as reserve carbohydrate in some fungi (spore germination) and insects (flight muscles) and some prokaryotes utilize trehalose as carbon and energy source [82]. The multifunctional role of trehalose particularly in eukaryotes (fungi and plants) even extends to signaling and regulation as well as to infection mechanisms in pathogenic fungi and bacteria [81, 83]. So far, there are at least six pathways for trehalose biosynthesis and degradation described: **(i)** the trehalose-6-phosphate synthase (TPS)/trehalose-6-phosphate phosphatase (TPP) pathway (TPS/TPP or otsA/otsB), **(ii)** the maltooligosyltrehalose synthase (TreY) / trehalohydrolase (TreZ) route, **(iii)** the trehalose synthase (TreS) pathway, **(iv)** the synthesis via the trehalose glycosyltransferring synthase (TreT), and **(v)** the trehalose phosphorylase (TreP). Degradation of trehalose is catalyzed via **(vi)** trehalases (TreH) (Figure 3).

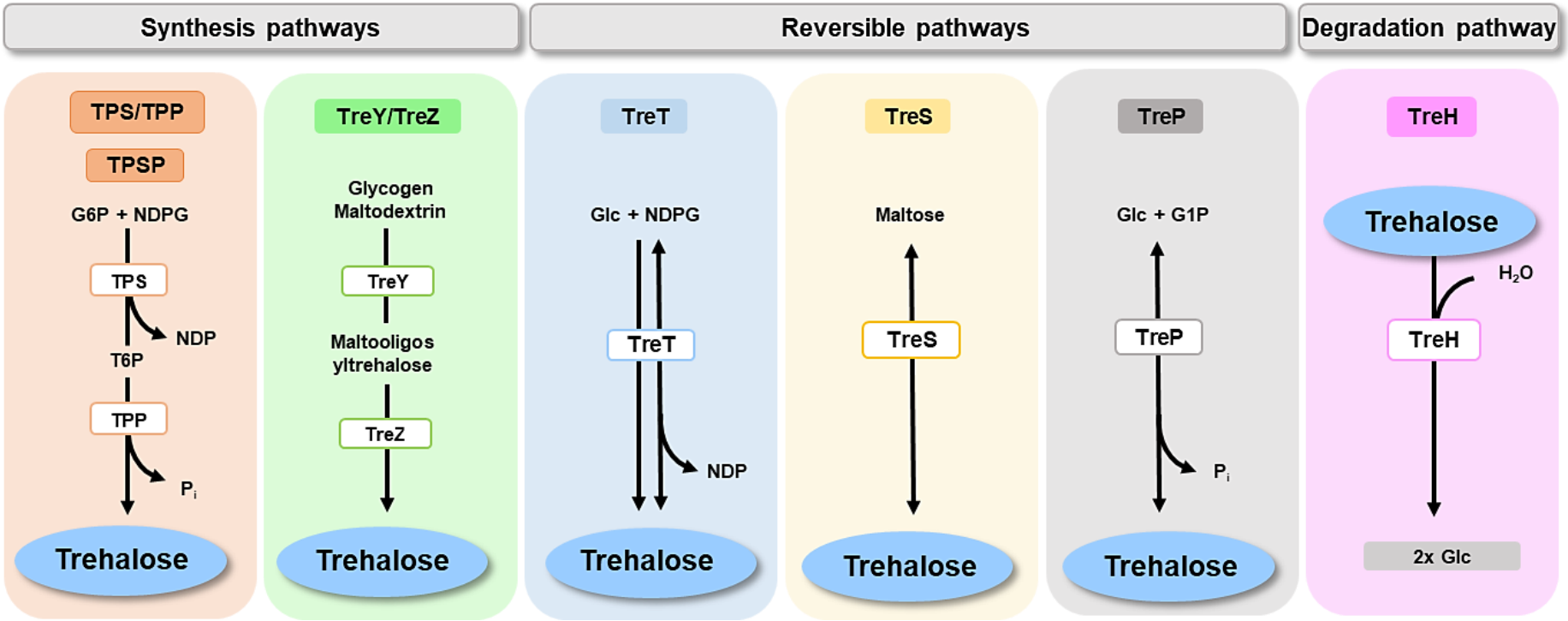


Figure 3: Trehalose metabolizing pathways including enzymes, substrates and intermediates. Abbreviations; TPS, trehalose-6-phosphate synthase; TPP, trehalose-6-phosphate phosphatase; TPSP, bifunctional fusion protein with trehalose-6-phosphate synthase/trehalose-6-phosphate phosphatase domain; TreY, maltooligosyltrehalose synthase; TreZ, maltooligosyltrehalose trehalohydrolase; TreT, trehalose glycosyltransferring synthase; TreS, trehalose synthase; TreP, trehalose phosphorylase; TreH, trehalase; Glc, glucose; G6P, glucose 6-phosphate; G1P, glucose 1-phosphate, NDP, nucleoside diphosphate; NDPG , nucleoside diphosphate glucose, T6P, trehalose 6-phosphate.

In Eukarya, trehalose is mainly synthesized by the TPS/TPP pathway. Whereas, only some fungi and algae were reported to utilize the trehalose phosphorylase (TreP) [83-85]. Conversely, in Bacteria in addition to the widespread TPS/TPP route, also the TreT, TreY/TreZ, TreS, and TreP enzymes have been described [85, 86]. In numerous cases, one bacterial species harbors more than one pathway in parallel, which contributes to trehalose synthesis as shown by several mutational approaches [87, 88]. This led to the assumption that eukaryotes use a single pathway for trehalose synthesis whereas Bacteria employ multiple pathways [84-86]. Besides Eukarya and Bacteria, trehalose biosynthesis was also shown to be present in several Archaea, mainly in Crenarchaeota including hyperthermophiles of the genera *Thermoproteus* and *Pyrobaculum* as well as the thermoacidophiles *Sulfolobus*, *Acidianus,* and *Metallosphaera* [89]. Unusually, trehalose was found in some Euryarchaeota like *Thermoplasma* and also in some extremely halophilic Archaea, where additionally the derivative sulfotrehalose was described (e.g., in *Natronococcus occultus*) [90, 91]. Comparably to Bacteria, also in Archaea several pathways for trehalose synthesis have been described, which often occur in one organism in parallel [85, 92]. Therefore, prokaryotes showed the greatest distribution of trehalose synthesis pathways. The presence of several biosynthetic routes in the same organism may be due to the strict requirement to accumulate trehalose under changeable environmental conditions and highlighting the availability of each pathway. Therefore, further details on the individual trehalose biosynthesis pathways are summarized in the following sections of this thesis.

1.3.1 TPS/TPP pathway

The most widely distributed pathway for trehalose biosynthesis is the TPS/TPP pathway. The TPS (trehalose-6-phosphate synthase) (EC 2.4.1.15) catalyzes the transfer of one glucose unit from an activated sugar (e.g., UDPG, ADPG, GDPG) onto glucose-6-phosphate (G6P) to produce the intermediate trehalose-6-phosphate (T6P), while uridine diphosphate (UDP or ADP, GDP respectively) is released. The TPP (trehalose-6-phosphate phosphatase) (EC 3.1.3.12) dephosphorylates T6P, yielding trehalose and inorganic phosphate. However, TPS enzymes belong to the glycosyltransferase family 20 (GT-20) and TPP enzymes belong to the HAD (L-2-haloacid dehalogenase) superfamily of magnesium-dependent phosphatases and phosphotransferases, which are distributed in both, prokaryotes and eukaryotes [93]. In Bacteria, TPS and TPP enzymes clustered together (e.g., *E. coli* TPS (OTSA) and TPP (OTSB)) [85, 94, 95], whereas in Eukarya, TPS proteins are mainly fused to the TPP domain and trehalose biosynthesis seems to occur preferred via this pathway [85]. Also in many eukaryotes, several TPS copies were reported [85, 96] e.g., in *Saccharomyces cerevisiae* four proteins with TPS and TPP domains were described, forming a holoenzyme complex, while the TPS1 (TPS) and TPS2 (TPP) subunits possess a catalytic function the TPS3 and TSL1

proteins play a regulatory role [97, 98]. However, mutational approaches demonstrated that all subunits are required for optimal enzymatic activity in yeast [98]. In Archaea, comparative genomics revealed several *tps* and *tpp* homologs, either as single or as fused TPS and TPP domains (e.g., in *Methanothermobacter thermoautotrophicus* (single domain), *Thermoplasma volcanium* (single domain) *Thermoproteus uzoniensis* (single domain), *Pyrolobacum aerophilum* (fused domains) or *Methanoculleus marisnigri* (fused domains)) [85, 96, 99, 100]. However, the single TPS/TPP proteins from *Thermoplasma acidophilum* have been characterized in detail [101]. Furthermore, complex-forming bifunctional TPSP, a fusion protein with TPS and TPP domain likewise to eukaryotic counterparts, was characterized from *Thermoproteus tenax* [100]. The genes coding for *T. tenax* TPSP are organized in an operon together with a glycosyltransferase and a putative mechanosensitive channel [100]. The TPSP requires an activation, catalyzed by the glycosyltransferase to exhibit the bifunctional synthase and phosphatase activities, otherwise, only high TPP activity was observed [100]. It has been demonstrated that enzyme fusions and complexes showed either a preferred TPS or TPP catalytic activity without activation of other regulatory and structural elements or enzymes [96, 98, 100, 102].

1.3.2 TreY/TreZ pathway

The conversion of maltodextrins such as starch or glycogen is catalyzed by the TreY/TreZ pathway. Trehalose is synthesized via the maltooligosyltrehalose synthase (TreY) (EC 5.4.99.15), which promotes the transglycosylation of the last glucose moiety at the reduced end of maltodextrins (α,α -1,4 to a α,α-1,1 glycosidic bond), yielding maltooligosyltrehalose. The final hydrolytic release of trehalose is catalyzed via the maltooligosyl trehalose trehalohydrolase (TreZ) (EC 3.2.1.141) [76, 103, 104]. This pathway is utilized for irreversible anabolic trehalose formation [86] and both enzymes belong to the α-amylase family [105]. However, in Archaea and especially in the order of *Sulfolobales*, this pathway plays a major role in trehalose biosynthesis. Functional enzymes were analyzed in detail from *S. acidocaldarius*, *S. solfataricus*, *S. tokodaii*, *S. shibatae* and *Metallosphaera hakonensis* [76, 77, 105-110]. In the genomes of *S. acidocaldarius* and *S. solfataricus* the *treY* and *treZ* genes are clustered together with treX, coding for a glycogen debranching enzyme [76, 111]. The TreX enzyme is assumed to influence the glycogen and trehalose breakdown via transglycosylation reactions [112]. Although homologs of both genes were reported in many bacterial genomes [85], a functional TreY/TreZ pathway has been only reported in a limited number of bacteria, e.g., in *Anthrobacter* sp. [113], *Corynebacterium glutamicum* [88, 114], *Bradyrhizobium japonicum* [115] *and Mycobacterium smegmatis* [116]. In Eukarya, the TreY/TreZ pathway is absent, no homologs have been predicted so far [85].

1.3.3 TreT pathway

The TreT pathway involves the trehalose glycosyltransferring synthase (TreT) (EC 2.4.1.245), which catalyzes the formation of trehalose from glucose and activated sugars as donor substrates. TreT enzymes were exclusively found in hyperthermophilic Archaea. One single bacterial homolog has been predicted in *Thermotoga maritima,* which inherited a substantial part of its genome from Archaea by lateral gene transfer [83, 85, 117]. TreT enzymes were analyzed in detail from *Thermococcus litoralis*, *Pyrococcus horikoshii* and *T. tenax* [92, 118-120]. The proteins reported in *T. litoralis* and *P. horikoshii* catalyzes a reversible formation of trehalose from glucose with differential preferences on several activated sugars as donor substrates. In contrast, the TreT described in *T. tenax* catalyzes the unidirectional synthesis of trehalose with a preference for UDPG as donor substrate [92]. However, in *T. tenax* the encoding gene overlaps 4 bases with an upstream gene, coding for *orfY*, a hypothetical protein of unknown function [92, 121]. In *T. litoralis*, *P. horikoshii,* and *P. furiosus* the gene is organized in a cluster together with a trehalose/maltose ABC transporter. Comparative genomics revealed evidence of horizontal gene transfer for the encoding 16 kb ABC transporter fragment [122]. Furthermore, previous studies showed a preferred trehalose uptake rather than an *de novo* synthesis in *T. litoralis* [123, 124]. Therefore, in *Thermococcales* a physiological role of TreT rather in trehalose degradation than in synthesis was discussed [118]. Both types of TreTs (uni- or bidirectional) show a distinct distribution, the single *tre*T-trehalose/maltose ABC transporter gene cluster exclusively occurring in Euryarchaeota, whereas the putative operon of *tre*T together with *orfY* was only predicted in *Thermoproteales* and *Sulfolobales*. A similar organized cluster was identified in *S. acidocaldarius* (*saci*_1827 & *saci*_1826) [92] and further investigation of the *treT* homolog was a major part of this thesis.

1.3.4 TreS pathway

The trehalose synthase (TreS) (EC 5.4.99.16) catalyzes a reversible conversion of maltose to trehalose via isomerization (α,α-1,4 bond in maltose to α,α-1,1 bond yielding trehalose) [81, 125]. In Bacteria, comparative genomics predicted a wide distribution of the TreS pathway, whereas no homologs were found in Eukarya. In Bacteria, the role of TreS in synthesis and degradation was discussed and seems not to be fully clarified [85]. Several bacterial TreS enzymes were characterized in detail from *Pimelobacter* sp. R48, *Pseudomonas syringae,* and *Thermus caldophilus* [125-127]. However, mutational approaches indicating rather involvement in trehalose degradation than in synthesis [88]. In *T. caldophilus* the TreS additionally catalyze the formation of small amounts of α-β-trehalose *in vitro* (3.5% from 10 mM maltose), which has never been reported for any TreS before [126]. So far, only one archaeal TreS was analyzed in detail from *Picrophilus torridus*, a thermophilic, hyperacidophile

that grows optimally at 60°C and pH 0.7 [128, 129]. Therefore, this TreS was suggested as a candidate in process optimization for industrial trehalose production from maltose [130].

1.3.5 TreP pathway

The trehalose phosphorylase (TreP) catalyzes the reversible conversion of glucose-1-phosphate (G1P) and glucose to trehalose. TreP enzymes have been reported in bacteria, fungi, and protists [83, 85]. Although comparative genomics revealed sequence identities of TreP enzymes in archaeal genomes, e.g., homologs were predicted in *Pyrococcus horikoshii, Thermofilum pendens,* and *Thermococcus barophilu*m, no functional enzymes have been identified so far [85, 131]. Instead, bacterial thermophilic TrePs have been characterized in detail from *Thermoanaerobacter tengcongensis, Thermoanaerobacter brockii* and *Caldanaerobacter subterraneus* [132-134]. In *T. tengcongensis in vitro* studies reported rather preferred trehalose biosynthesis than degradation [135]. Furthermore, upregulation of the *treP* gene has been demonstrated in response to salt stress and a down-regulation was observed when trehalose was supplied [135]. In contrast to this report, TreP in *T. brockii* and *Catellatospora ferruginea* [133, 136] catalyzed rather a trehalose degradation than synthesis. Comparable results were observed for eukaryotic TreP enzymes, likewise enzymes from mushrooms *Schizophyllum commune* and *Agaricus biosporus* catalyzing preferred the degradation of trehalose [137-139]. Interestingly, there existing two types of TreP enzymes, retaining and inverting reaction mechanisms were reported (EC 2.4.1.231 & EC 2.4.1.64) [140]. Retaining enzymes were confirmed in *Schizophyllum commune* and *Agaricus biosporus* [137-139], catalyzing the formation of α-D-glucose and α-D-glucose-1-phosphate from α,α-trehalose, and inorganic phosphate. These enzymes belong to glycosyltransferase family 4 (GT4). In contrast, inverting enzymes producing β-D-glucose and β-D-glucose-1-phosphat and belongs to glycosidhydrolase family 65 (GH65). For instance, GH65 members were described in *Euglena gracilis* [141], *Bradyrhizobium japonicum* [142], *Micrococcus varians* [143], or *Caldanaerobacter subterraneus* [134]. Due to the high energy content of a glycosyl phosphate, phosphorolysis is reversible and thus TreP enzymes can either catalyze anabolic or catabolic reactions [144]. Nevertheless, from all studies on inverting and retaining TrePs, not much about the *in vivo* function of these enzymes could be concluded and it reveals still unclear if these enzymes participate in trehalose synthesis or degradation.

1.3.6 TreH pathway

Trehalose degradation is catalyzed by trehalases (TreH) (EC 3.2.1.28), which hydrolyze trehalose into two glucose moieties. TreHs are widely distributed in Bacteria and Eukarya and recently also a rare number of TreHs were described in Archaea [147-160]. Although vertebrates have no functional pathways for trehalose biosynthesis, degradation plays an important role in various tissues [145, 146]. However, TreH enzymes belong to different

glycoside hydrolase families, members of GH15, GH37, and GH65 have been reported so far [147-150]. Therefore, two different types of TreH enzymes were described, based on their pH optimum, acidic and neutral TreHs were confirmed. Both forms can occur in one organism like it was demonstrated for yeast [151]. However, the two types differ in structural motifs and ion dependency, which allows the enzyme a certain mode of action e.g., neutral TreHs are strictly ion-dependent and can be regulated via phosphorylation [151-153].
Furthermore, it was demonstrated that acidophilic enzymes, like *Neurospora crassa* TreH, are involved in assimilation of extracellular trehalose as carbon source [154-156]. In addition, in *E. coli*, a periplasmatic TreH was reported, which may function together with a part of an uptake system, supplying glucose that can be further processed in glycolysis [157].
For a long time, no classical TreH could be identified in Archaea and trehalose degradation was rather assumed via reversible TreT (see above) [118]. Recently, novel findings confirmed first archaeal TreHs in *Thermoplasma volcanium* and *Thermoplasma acidophilum* [150]. Interestingly, comparative genomics revealed TreH-related sequences only in acidophilic Archaea [150]. Therefore, TreH may be the major enzyme involved in trehalose degradation in acidophilic Archaea [150]. Furthermore, trehalose degrading activities were demonstrated in crude extracts of *S. acidocaldarius* [158] and two functional TreHs enzymes (*Saci*_1816 & *Saci*_1250) were identified [159]. Both enzymes belong to the glycoside hydrolase family 15 (GH15) [150, 158, 159, 160] and the hydrolytic activities were analyzed in detail [159,160]. However, trehalose concentrations and buffer conditions differ significantly in both studies (90 mM trehalose / 5 mM trehalose; 50 mM acetate buffer (pH 4.0), 50 mM Gly-HCl (pH 3.7) / 20 mM sodium citrate buffer (pH 4.0)) [159, 160]. We assume that different assay conditions may influence the measured *in vitro* activities and due to the importance of all trehalose converting enzymes for this study I further investigated the enzymatic activity of the *S. acidocaldarius* TreH enzymes (IV Appendix S. 135).

However, most of here summarized information available on archaeal trehalose metabolism is based on comparative genomics and related sequence information as well as analyses of recombinant proteins. Thus, in contrast to Eukarya and Bacteria where numerous mutational analyses have been performed [88, 115, 161, 162], the general function of trehalose in Archaea is unknown so far, and the complexity of the trehalose metabolism including the *in vivo* importance of the different synthesis pathways in is not yet understood. Therefore, this present study deepens the understanding of archaeal trehalose metabolism and provides further insights in stress response of *S. acidocaldarius.*

1.4 Trehalose metabolism in *S. acidocaldarius*

S. acidocaldarius has been described to synthesize trehalose via the maltooligosyltrehalose synthase (TreY) and maltooligosyltrehalose trehalohydrolase (TreZ) pathway [76, 103, 104], and the trehalose glycosyltransferring synthase (TreT) pathway has been predicted by comparative genomics [92] (Figure 4). Although previous studies have assumed that trehalose may functions as compatible solute under temperature stress rather than under osmotic stress in *Sulfolobales*, no consolidated studies have been published so far [89, 163]. However, the function of trehalose and the significance of the two synthesis routes have not been investigated and were one major aim of this work.

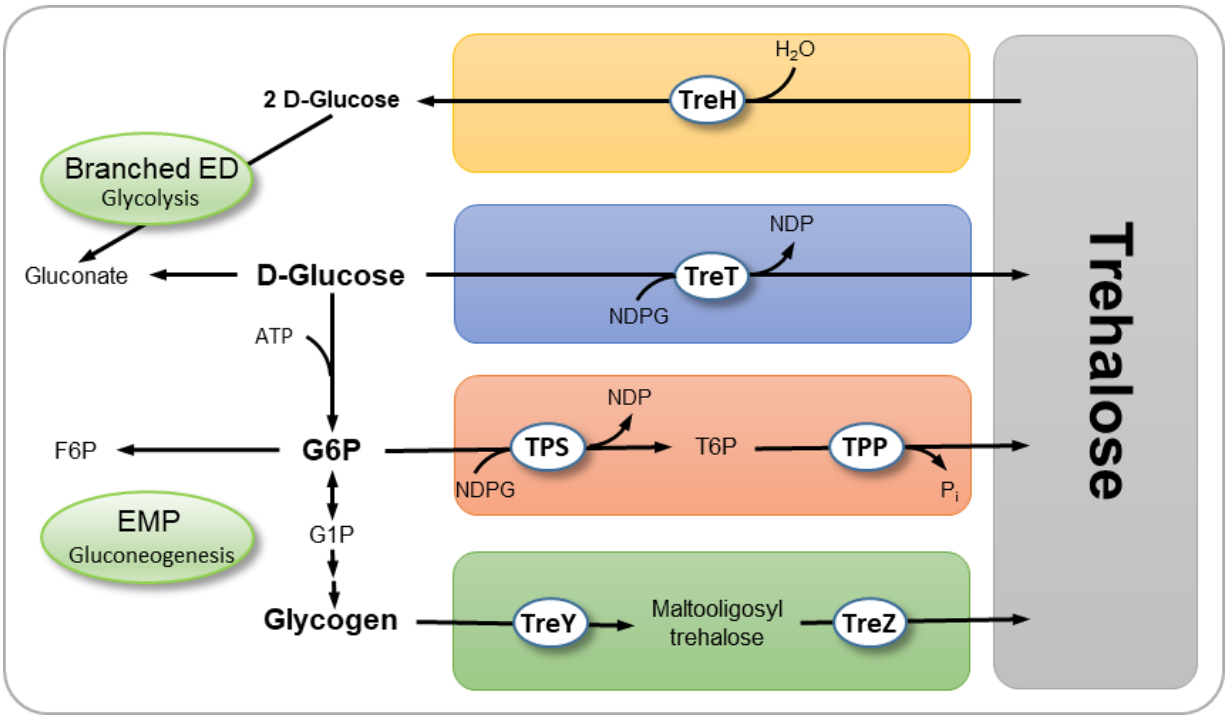


Figure 4: Overview of identified pathways for trehalose synthesis and degradation in *S. acidocaldarius*. *S. acidocaldarius* degrades D-glucose via a modified branched Entner-Doudoroff (ED) pathway, whereas the Embden-Meyerhof-Parnas (EMP) pathway is used only for gluconeogenesis [22]. Glycogen has been reported as a central carbon storage compound in *S. acidocaldarius* [164], and hexokinase activity (broad substrate specificity) has been reported for *S. tokodaii* [73]. The TreY/TreZ [77] and TreT [92] trehalose synthesis pathways and the novel TPS/TPP pathway as well as the trehalases for trehalose degradation [158] are depicted. Abbreviations: TreY, maltooligosyltrehalose synthase; TreZ, maltooligosyltrehalose trehalohydrolase; TPS, trehalose-6-phosphate synthase; TPP, trehalose-6-phosphate phosphatase; TreT, trehalose glycosyltransferring synthase; TreH, trehalase; UDPG, UDP glucose. Enzymes: TreY, maltooligosyltrehalose synthase; TreZ, maltooligosyltrehalose trehalohydrolase; TPS, trehalose-6-phosphate synthase; TPP, trehalose-6-phosphate phosphatase; TreT, trehalose glycosyltransferring synthase; TreH, trehalase. NDPG, nucleotide-diphosphate glucose.

Thus, single and double gene deletion strains of the two pathways (Δ*treY* or Δ*treT* and Δ*treY*/Δ*treT*) were constructed, but even in the double mutant trehalose was still detectable, indicating a further trehalose synthesis route in *S. acidocaldarius*. Using bioinformatics, a novel trehalose-6-phophate (T6P) synthase/T6P phosphatase (TPS/TPP; encoded by *saci*_1249,

*saci*_0016) pathway was identified (Figure 4). On sequence level, the enzymes show no similarity to classical TPSs as well as TPPs. The coding function of the genes was confirmed by characterization of the recombinant enzymes, which catalyzes the *in vitro* synthesis of trehalose via glucose 6-phosphate (Glc-6P) and uridine-diphosphate-glucose (UDPG). The final triple mutant strain (Δ*treY*/Δ*treT*/Δ*tps*) revealed no phenotype under standard growth conditions but was deficient in trehalose formation. However, growth of the triple mutant was abolished under salt stress (250 mM NaCl). Accordingly, salt stressed wild type cells showed a 12-fold increased trehalose accumulation, whereas no effect was observed upon temperature stress (heat and cold shock). Quantitative real-time PCR also demonstrated the upregulation of *treT*, *treY* and *tps* genes upon salt shock, confirming the role of trehalose as compatible solute and the importance of all three synthesis pathways under salt stress (chapter 3.1). The relevance and function of trehalose, demonstrated in this study, is not only beneficial for *S. acidocaldarius*, but have already ensured that trehalose is one of the most important stabilizers utilized in many commercial applications.

1.5 Applications of trehalose

The nonreducing disaccharide trehalose (α-D-glucopyranosyl-1,1-α-D-glucopyranoside) consists of two glucose units connected via an α,α-1,1-glycosidic linkage (low bond energy (>1 kcal/mol)), which makes the sugar very stable in comparison to sucrose [86, 165]. Trehalose appears as a white, odorless powder with a relative sweetness of 45% compared to that of sucrose. Trehalose has several important chemical characteristics, among these are its high hydrophilicity, high thermostability (>99% (120°C/90min)), high pH-stability (pH 3-10, at 100°C, 24 h), and the resistance to Maillard and Amadori reactions (no reducing ends) [86, 166]. These properties are useful for microorganisms to survive under extreme conditions and cope with environmental stress, but also makes the sugar highly applicable in a broad range of commercial and biotechnological fields. In food industry, trehalose serves as important additive, used as sweetener and stabilizer in many frozen products. The cosmetic industry uses trehalose as stabilizer and moisturizer in body lotions, crèmes, and in deo sprays (trehalose masks displeasing odors) [82, 166]. Since it was reported that trehalose also stabilizes the structure of DNA, proteins, and membranes, it has been used for several biotechnological purposes, e.g., to protect DNA-polymerases, restriction enzymes, DNA ligases, and antibodies [167, 168]. Additionally, in pharmaceutical industry trehalose containing solutions support cryopreservation of different cell materials (sperm and stem cells) and can be used as additive in higher reliable organ preservation solutions [165, 169, 170]. Currently, there is a growing interest in medicine and pharmaceutical industry, where trehalose is tested as a relevant therapeutic molecule in research of neurodegenerative diseases (Alzheimer, Huntington, and Parkinson disease) [171-173]. Oxidative stress, aggregation, and proteasomal dysfunction have been considered key mechanisms associated with

neurodegenerative disorders. Trehalose is assumed to act as chemical chaperone, preserving protein structure stability, protein folding as well as reducing aggregation of pathologically misfolded proteins [174]. In addition, there exists tremendous interest in modifying the structure of trehalose to generate analogues. Therefore, chemical as well as chemoenzymatic synthesis of non-degradable trehalose analogues have been applied to provide several novel advantages as bioprotectants, therapeutic molecules, or alternative food additives [175]. Numerous enzymes of the previously described synthesis pathways have been used for this purpose. Recently, the TreT of *T. tenax* gained special interest due to the excellent thermostability and unidirectional synthesis of trehalose [75-177]. The TreT has now been applied to semi-preparative synthesis of numerous analogues bearing a variety of modifications (e.g., fluoro-, deoxy-, azido-, thio-, stereochemical, and isotopic modifications) [175]. These analogues are already useful imaging probes and inhibitors and have future value in diagnosis and treatment of pathogenic microorganisms in which trehalose metabolism plays a major role in virulence, like *Mycobacterium tuberculosis* [175, 178].

However, the here described advantages of trehalose exemplify how high the application and economic potential of compatible solutes are. Therefore, besides trehalose also other compatible solutes with similar properties and function come into biotechnological focus. Especially much rarer compatible solutes, that are particularly synthesized by extremophilic microorganisms have gained special attention due to their unique characteristics.

1.6 Extremolytes – Compatible solutes from extremophiles

Compatible solutes that are exclusively found and synthesized in extremophiles (e.g., halophiles or (hyper)thermophiles) are so-called “extremolytes”. The application of extremolytes seems to be beneficial in many fields, their exceptional properties can be translated into health care promoting and therapeutic activities, which open versatile application for the cosmetic, pharmaceutical, medical, and food industries. In addition to the mesophilic compatible solute trehalose, which was described in the preceding chapter, the commercially most relevant products so far are the extremolytes ectoine and hydroxyectoine (Figure 5). The protective properties of these compounds as well as their synthetic pathways have been well characterized [179-181]. For instance, ectoine and hydroxyectoine maintain the fluidity of membranes and were demonstrated to protect epithelial cell layers from several allergens, UV light, fine dust particles as well as heat or drought [182, 183]. Also, inflammatory reactions caused by these stressors on skin or mucous membranes can be reduced (Bitop AG: “Extremolytes” https://www.bitop.de/de/extremolytes (Date: 21.09.2020). Thus, ectoine became an important ingredient of many high value-added products in human health care and well-being. The production process of ectoine and hydroxyectoine relies on fermentation of the natural producer strain (non-GMO) the halophilic bacterium *Halomonas elongata*, and involves an approach designated as “bacterial milking” [182, 184]. This is basically an osmotic downshock from 15% to 3% NaCl, which forces the cells to release the solute to the medium via mechano-sensitive channels [185, 186]. While ectoine and hydroxyectoine are particularly found in halophilic or halotolerant bacteria, in hyperthermophiles other extremolytes were described (Figure 5), which often are heterosides such as glucosyl-glycerol (GG, glycoin) [187], glucosyl-glycerate (GGA) [188], mannosylglycerate (MG, Firoin) [189, 190], and the very rare mannosylglyceramide (MGA, Firoin-A) [182, 184, 191]. Also, phosphorylated compounds such as di-*myo*-1,1’-inositol-phosphate (DIP) [192, 193], α-diglycerol phosphate (DGP) [194], and cyclic 2,3-diphosphoglycerate (cDPG) have been described for hyperthermophiles [195-198].

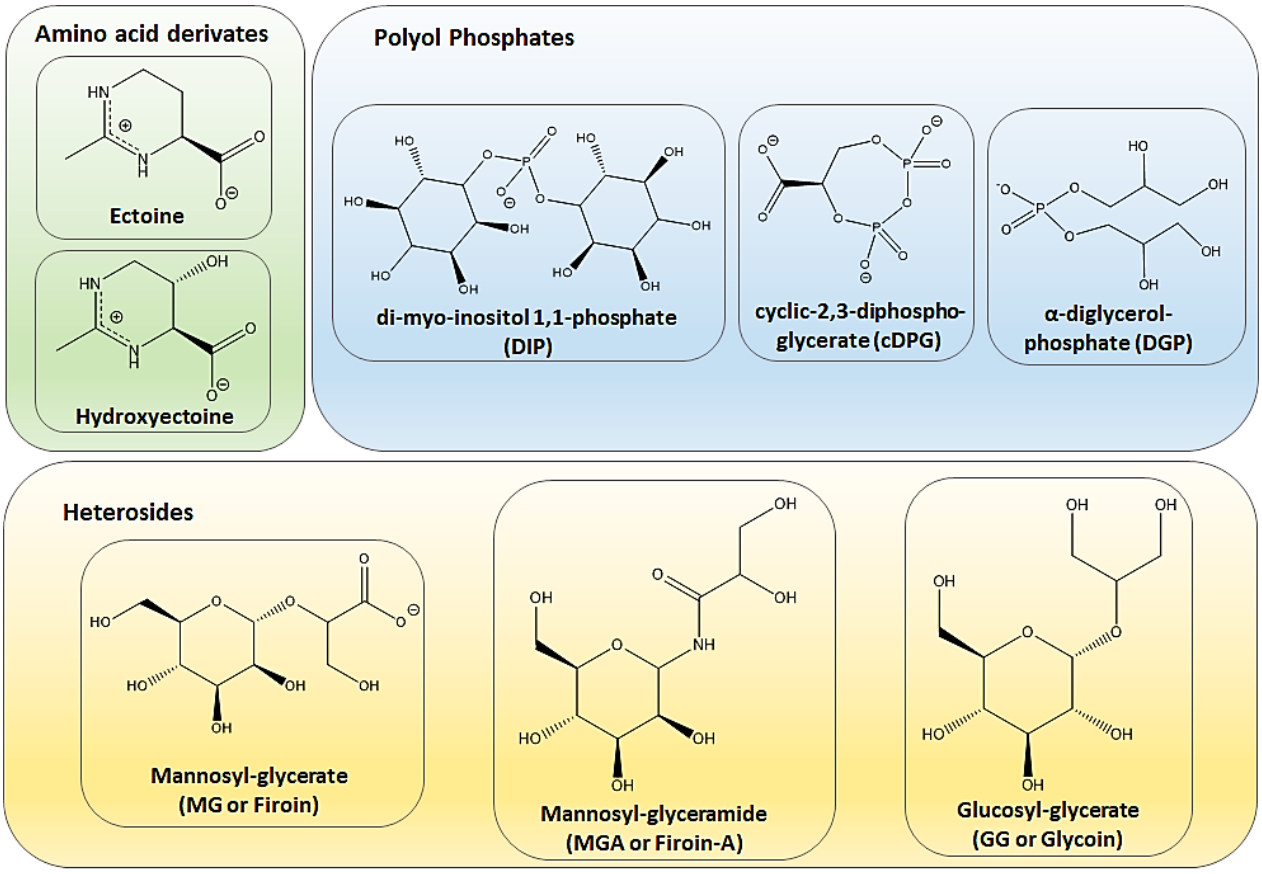


Figure 5: Overview and diversity of extremolytes. Amino acid derivates of pyrimidine as ectoine ((S)-2-methyl-3,4,5,6-tetrahydropyrimidine-4-carboxylic acid) and hydroxyectoine ((S,S)-2-methyl-5-hydroxy-1,4,5,6-tetrahydropyrimidine-4-carboxylic acid) [182]. Phosphorylated extremolytes as di-myo-inositol 1,1-phosphate (DIP) [192, 193], cyclic-2,3-diphosphoglycerate (cDPG) [197], and α-diglycerol-phosphate (DPG) [194]. Sugar derivates such as mannosyl-glycerate (MG) also referred to as Firoin [189, 190] and mannosyl-glyceramide (MGA) also referred to as Firoin-A [182, 184] as well as glycosyl-glycerate (GG) that is also known as Glycoin [187].

Many attempts have been performed to characterize and synthesize extremolytes, recently, also a nice summary with biotechnological focus on microbial production of extremolytes has been reviewed [199]. However, the potential for industrial applications of extremolytes remains largely unexploited, mainly because applicable production hosts and efficient synthesis pathways to make their production economically viable are missing. Consequently, it is of great interest to evolve suitable alternatives to synthesize extremolytes in analytical and preparative scale necessary for industrial purposes. Therefore, the synthesis of the extremolyte cyclic-2,3-diphosphoglycerate (cDPG) was addressed in this study.

1.6.1 Cyclic-2,3-diphosphoglycerate (cDPG) synthesis

Among the previously described extremolytes, accumulation of cDPG has exclusively been confirmed in hyperthermophilic methanogens (Archaea) such as *Methanothermus fervidus, Methanopyrus kandleri,* and *Methanothermobacter thermoautotrophicus* in concentrations from 0.3 - 1.1 M [196, 200-202]. It appears to play a role in thermoadaption of proteins and increased thermostabilities on different model enzymes have been reported [201, 203, 204]. Furthermore, it protects plasmid DNA against oxidative damage by hydroxyl radicals and also functions as a superoxide scavenger [182].

The enzymatic synthesis of cDPG was shown to proceed via two catalytic steps (Figure 6). In the first step, 2-phosphoglycerate kinase (2PGK) synthesizes the phosphate ester 2,3-diphosphoglycerate (2,3DPG) from 2-phosphoglycerate (2PG). In the second step, cyclic-2,3-diphosphoglycerate synthetase (cDPGS) catalyze the formation of an intramolecular phosphoanhydride bond resulting in cDPG.

Although the biotechnological application of cDPG seems very reasonable, since its protective properties and relatively simple biosynthetic pathway are known, no effective production procedure has been established so far. This is due to the low cell yield and challenging cultivation conditions for methanogens under strict anaerobiosis and very high temperature hampering a whole-cell *in vivo* approach like for ectoine. Furthermore, the enzymes for cDPG synthesis characterized from hyperthermophilic methanogens appeared unstable although no detailed information on the enzyme stability has been reported [197, 205]. Also, the conversion efficiency of the enzymes in terms of substrate to product ratios has not been analyzed so far. Therefore, an *in vitro* enzymatic approach provides striking advantages over whole-cell systems. For instance, enzyme cascades are free from challenging cultivation conditions and low cell yields, free from material erosion by competing metabolic pathways, kinetic restrictions, and toxicity problems [206]. However, for an enzymatic *in vitro* approach sufficient amounts of enzyme are required and thus optimized expression and purification procedures are necessary.

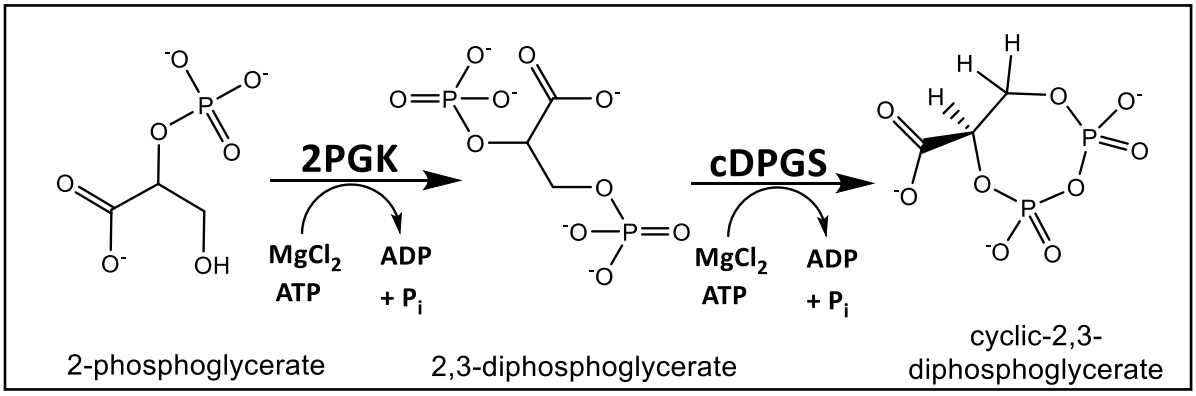


Figure 6: cDPG synthesis. Conversion of 2PG to 2,3DPG by 2PGK and formation of cDPG by cDPGS. Both catalytic steps are ATP-dependent and require additional magnesium ions. Abbreviations: 2PG, 2-phosphoglycerate; 2,3DPG, 2,3-diphosphoglycerate; cDPG, cyclic di-2,3-phosphoglycerate; and cDPGS, cyclic-diphosphoglycerate synthetase.

Therefore, as one chapter of this thesis, I here present first results towards the establishment of an enzymatic *in vitro* approach to produce cDPG for application. The heterologous enzyme production of cDPGS in *E. coli* was optimized with respect to codon adaption, expression conditions, and purification procedures. The improved expression isolation procedure from this study provided a more than two-fold higher yield of recombinant cDPGS compared to previously published results [202]. Furthermore, addition of 400 mM KCl, 10 mM DTT, and 2.5 mM Mg^{2+} increased the short-term stability to at least several days and further supplementation of the buffer with 25% glycerol enabled the long-term storage at -80°C for months without loss of activity. The cDPGS showed a V_{max} of 21.7 U mg^{-1} and complete substrate to product conversion was achieved, which allows for an easy one-step approach for the biocatalytic synthesis of cDPG. Thus, this work paved the way for the cost-effective production of cDPG as added-value product for use in biotechnological applications.

1.7 Biocatalysts from extreme environments

The biocatalytic synthesis of extremolytes described in the preceding chapter exemplifies how important biocatalysts are to enable the synthesis of novel products but also how novel biocatalysts contribute to the development and optimization of new approaches. Therefore, the biocatalytic sector has a continuously high demand for novel enzymes for the development of innovative solutions and applications. The natural microbial biodiversity has an inestimable metabolic variability and hence represents an enormous reservoir for novel enzymes. However, since the vast majority of the total microbial biodiversity is still not cultivable, this enormous potential remains inaccessible. Using metagenomic approaches this hidden diversity becomes more and more accessible. This young, vibrant field provides appropriated tools used for standard genomics and applied them to the study of entire microbial

communities, without the need to isolate and culture the individual microbial species [207-209]. Currently, metagenomic analyses are based on the direct isolation of genomic DNA from environmental samples, and they are either sequence-based (DNA level) or functional-based (protein level) [207-209] (Figure 7). The sequence-based metagenomic screening provides sequencing data from the obtained environmental DNA samples. Recently, next-generation sequencing techniques were developed, enabling the fast mapping of DNA related information without cloning or library construction. The further bioinformatic evaluation provides a high throughput annotation of genes and information about taxonomic and phylogenetic relations. However, sequence-based metagenomics only relies on comparison to already annotated sequences and databases, which only allows the detection of putative enzymes based on their conserved sequences. Accordingly, the information outcome for biotechnology is limited, since mainly new variants of already characterized genes or genes of totally unknown function can be discovered [40, 210, 211]. In contrast, functional metagenomics can be applied for the identification of novel enzymes that possess a desired function or activity on a substrate of interest. This approach specifically refers to the construction of metagenomic libraries that comprises (i) generation of DNA fragments of appropriate size, and the ligation of the fragments into appropriate cloning vectors (e.g., plasmids, cosmids, fosmids, or bacterial artificial chromosome vectors), (ii) introduction of the recombinant vectors into a suitable expression host, and (iii) screening of clones harboring particular activities or containing specific sequences and functions [212] (Figure 7). Therefore, several powerful functional activity screening approaches have been developed in the past, enabling a high-throughput annotation of functional genes and active enzymes [213-218]. These functional screenings revealed a few hydrolytic enzymes, with high potential for industrial applications, e.g., amylases, esterases, and xylanases [219-222]. The functional screenings are carried out mostly in standard expression hosts like the gram-negative bacterium *E. coli* and only a rare number of reports are available using different screening hosts such as *Pseudomonas putida, Caulobacter vibrioides, or Thermus thermophilus* [216, 223]. To underline the role of phylogenetic and physiological distance of the gene donor and the employed expression host, it is a prerequisite that genes can be functionally expressed. Nevertheless, the applicability as host for functional expression appears to be limited due to various reasons: (i) transcriptional signals might not be detected, since functional expression mainly relies on native promoter structures [224], (ii) translation might be weak due to missing adequate ribosomal binding sites, (iii) poor mRNA stability, (iv) a different codon usage, and (v) incorrect folding because of missing chaperons and protein modifying enzymes, which leads to the formation of protein aggregates and inclusion bodies [225]. Consequently, proteins of organisms with different transcriptional and (post)translational mechanisms, and especially many extremophilic or

hyperthermophilic enzymes, which require specific adaption, are therefore frequently excluded from detection.

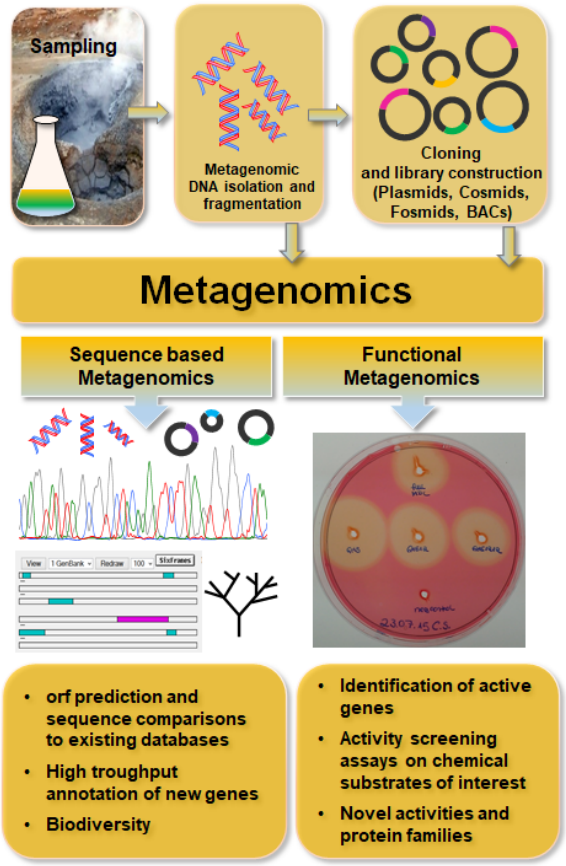


Figure 7: Overview of sequence-based and functional metagenomics. Analysis of environmental samples or enrichment cultures via sequence-based or functional screening methods.

However, for the screening of novel genes and enzymes and thus for the development of novel approaches and applications, especially hyperthermophilic enzymes gain increasing attention because they are typically folded into very stable conformations able to withstand high temperatures, and even require high temperatures for optimal activity. The temperature stability is often associated with high resistance to chemical denaturants commonly used in many industries [226, 227]. Notably, (hyper)thermostable and thermoactive hydrolases degrading recalcitrant polysaccharides like cellulose and xylan into their monomeric constituents, and are important for several industrial applications e.g., in starch-, pulp and paper-, food-, and biofuel industry [226, 227]. All of these applications require enzymes active at high temperatures, thus allowing better substrate solubility, easier mixing, and lowered risk of contamination.

Therefore, Archaea come into focus because they are abundant in extreme habitats, especially in hyperthermophilic environments. Archaea exhibit unique metabolic features and modified metabolic pathways characterized by unusual enzymes, not homologs to their bacterial counterparts [22]. Accordingly, archaeal enzymes offer great potential as novel biocatalysts for

industry, which has continuously high demand for innovative solutions and novel processes. However, the majority of so far cultured hyperthermophilic organotrophic Archaea of the two well-represented phyla, Euryarchaeota and Crenarchaeota are anaerobes growing by fermentation of various complex peptides and proteinaceous substrates, such as peptone, tryptone, beef, and yeast extract [228]. In contrast, only a few of these organisms were found to be able to grow on polysaccharides and possess a set of hydrolytic enzymes. According to literature, only three reports are showing weak growth of hyperthermophilic Archaea on cellulose and its derivatives like for *Desulfurococcus fermentans* [229], *Thermococcus sibiricus* [230] and a consortium of three species with predominance of an *Ignisphaera* representative [231]. The latter was also shown to possess cellulase activity. Growth on xylan or heat treated xylan (121°C, 20 min) was demonstrated only for members of the Crenarchaeota such as *Thermosphaera aggregans* [232], *Sulfolobus solfataricus* [233], and *Acidilobus saccharovorans* [234]. In contrast to these scarce reports for growth of hyperthermophilic Archaea on polysaccharides, genomes of many of these organisms harbor genes encoding glycoside hydrolases (GHs) (for additional information see also chapter 3.3, supplementary table 1), and several cellulases and xylanases were isolated from archaeal strains [39, 41]. However, most of these strains were either unable to grow on crystalline cellulose or xylan or were not analyzed for the ability to grow on these substrates [233, 235, 236]. Thus, the function and efficiency of these enzymes for *in vivo* polymer degradation is still unclear. For the discovery of efficient biocatalysts, an alternative and (in case) advantageous strategy to the application of metagenomic approaches is the isolation of novel strains with the desired properties, like the ability to cleave and to grow on cellulose or xylan. Accordingly, improved cultivation approaches have to be applied such as providing the most environmentally close conditions for cultivation [237], utilization of novel substrates and/or electron acceptors, presence or absence of growth factors, as well as the inhibition of cultured fast-growing microorganisms.

1.7.1 Novel archaeal glycoside hydrolases

In the paper presented in chapter three, a multi-layered approach to identify new thermostable polymer degrading hydrolases with potential for biotechnological application is described involving (i) an *in situ* enrichment strategy for organisms that are capable of polymer degradation, (ii) genomics, (iii) comparative genomics as well as (iv) cloning and biochemical characterization of enzymes of biotechnological interest. For the *in situ* enrichment strategy, Hungate tubes (18 mL) containing 200-300 mg of CMC, chitin, agarose, or birchwood xylan, with or without amorphous Fe(III) oxide (ferrihydrite) as external electron acceptor were incubated in a hot vent located in the tidal zone near Goryachiy cape of Kunashir Island (South Kurils, Russian Far East region) (Figure 8).

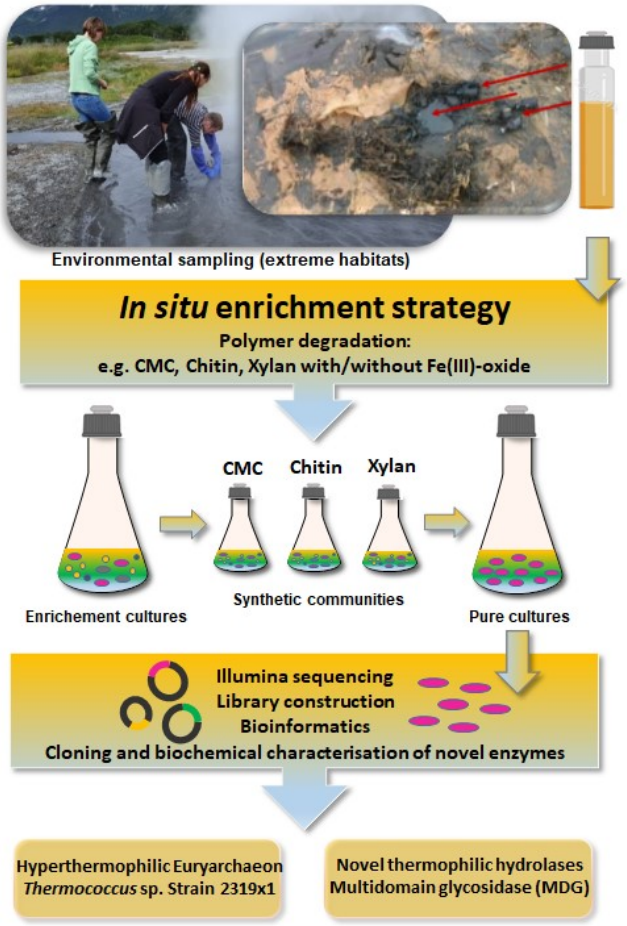


Figure 8: Overview of applied *in situ* enrichment strategy using different polymeric substrates and/or electron acceptors for the screening of novel organisms and biocatalysts of biotechnological relevance.

The enrichment setups were incubated for 6 days, and afterwards sealed and transferred to the lab at ambient temperature. Enrichment cultures were diluted and further incubated at temperatures close to the values at the respective sampling sites. A modified Pfennig medium with ferrihydrite [238] or elemental sulfur (10 g L^{-1}) was used for the isolation and metabolic characterization of new organisms (detailed information chapter 3.3, material and methods). The use of this technique revealed a novel hyperthermophilic Euryarchaeon, designated *Thermococcus* sp. strain 2319x1. This strain was isolated from an enrichment with birchwood xylan as energy and carbon source, and Fe(III) (in the form of insoluble ferrihydrite) as the electron acceptor. The initial inoculum consisted of black sand (enriched with mixed valence Fe mineral magnetite) and hot water from the original hot vent, which has a fluctuating temperature and pH in the range of 76 – 99 °C and 5.0 - 7.0 (Figure 8, inlay with red arrows). Strain 2319x1 grows optimally at 85°C and pH 7.0 on a variety of natural sugar and polysaccharides (e.g., yeast extract, peptone, gelatin, D-xylose, D-glucose, lactose, maltose, sucrose, dextrin, dextran, pullulan, starch, xylan, xyloglucan, AMCH, chitosan, AMC, CMC, alginate, lichenan, and bamboo leaves). The protein fraction extracted from the cells' surface

with Tween 80 exhibited endoxylanase, endoglucanase, and xyloglucanase activities. The genome of the Thermococcus sp. strain 2319x1 was sequenced and assembled into one circular chromosome. Within the newly sequenced genome, in total 18 genes, encoding glycosyl hydrolases (GHs) and carbohydrate esterases (CEs) of different CAZy families [239] (http://www.cazy.org) were identified which hydrolyze the glycosidic bond between two or more sugars or between a carbohydrate and a non-carbohydrate moiety. The majority of these glycosidases were annotated to be involved in hydrolysis of α-linked poly- and oligo-saccharides (starch, dextrin, dextran, pullulan, etc.), while only five enzymes were predicted to catalyze the hydrolysis of β-glycosides. Among these β-glycosidases, one exceptional novel glycosidase (143 kDa) enzyme was identified, which showed a unique multidomain architecture. It consists of three glycoside hydrolase (GH) domains and two carbohydrate-binding modules (CBM) with the domain order GH5-12-12-CBM2-2 (N- to C-terminal direction) (Figure 9A).

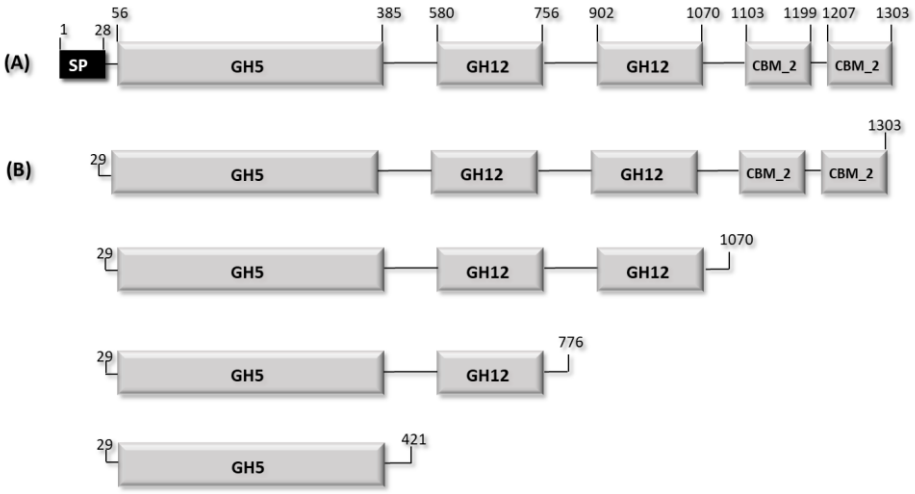


Figure 9: The multi domain glycosidase (MDG) from *Thermococcus* sp. strain2319x1. The predicted sequence signatures and domain-architecture **(A)** and generated recombinant versions of the protein [full length MDG (GH5-12-12-CBM2-2), GH5-12-12, GH5-12 and single GH5] **(B)** are shown. SP. Signal peptide; GH5, glycoside-hydrolase family No.5; GH12, glycoside-hydrolase family No.12; CBM2, cellulose binding-module family 2.

However, only a very rare number of archaeal hyperthermophilic multidomain hydrolases have been identified [231]. Mostly, only single domain GH5 or GH12 enzymes have been reported as for *P. horikoshii* (GH5) and *P. furiosus* (GH12) [235, 240, 241]. So far, no other archaeal multi domain enzyme has been described where GH5 occurs together with GH12. Additionally, in Archaea CBM2 domains are not common and seems to be restricted to the order of *Thermococcales* (18 CAZy entries). Also, the assembly of three catalytic GH domains together seems to be rather unique [242]. Only a very rare number of enzymes with several GH domains have been reported, so far. For instance, the *Caldicellulosiruptor bescii* CelA (GH9-CBM3c-3b-3b-GH48, N-to C-terminal) has a dual catalytic GH function [243]. Usually, structural

signatures of one single catalytic GH-domains together with different CBMs were reported in a large number of mesophilic bacteria and recently, several thermophilic bacteria of the genus, *Thermotoga*; *Caldicellulosiruptor, Caldibacillus* or *Clostridium* have been described to possess these characteristics (1-5 CBMs per enzyme) [242, 244-249]. For instance, the extracellular endo-β-1,4-xylanase (Xyn10B) from *Caldicellulosiruptor lactoaceticus* showed a multidomain domain order CBM22a-22b-22c-GH10-CBM9a-9b (N-to-C-terminal direction) [250]. It was demonstrated that CBM22 domains significantly promotes the thermostability of glycoside hydrolases [250]. In addition, using site-directed mutagenesis, five conserved residues crucial for ligand binding were identified in the C-terminal region of CBM22b domains [251]. The CBM9 modules have been proven to bind crystalline and amorphous cellulose, xylan, and xylooligosaccharides [252]. Therefore, the detailed biochemical analysis of the newly reported archaeal glycosidase with its unique multidomain architecture (GH5-12-12-CBM2-2) offers an interesting research subject, and was the contribution of this thesis to chapter three.

The full length multi domain glycosidase (MDG) and several truncated versions (Figure 9B) were heterologously expressed in *E. coli* and their respective properties were analyzed in detail. The full length MDG was able to hydrolyze various polysaccharides, with the highest activity (16 U mg^{-1}) for barley β-glucan (mixed β-1,3/1,4-glucoside), followed by that for CMC (β-1,4-glucoside), cello oligosaccharides, and galactomannan (3 - 7 U mg^{-1}). The here reported results indicate that the modular MDG structure with multiple glycosidase and carbohydrate-binding domains not only extends the substrate spectrum to a broad range but also seems to allow the degradation of partially soluble and insoluble polymers in a processive manner (Avicel 1 - 5 U mg^{-1}). Therefore, the MDG as well as the here presented truncated versions provide a new benchmark for microbial polymer degradation in extreme environments, which deepens our understanding of the function of archaeal multidomain hydrolases, and represents an innovative source for industrial application.

2 Scope of the thesis

In the present work, the introduced research topics and the contribution I have made to the respective manuscripts are explained in the following chapters below.

In chapter 3.1, the manuscript with the title **"Salt Stress Response of *Sulfolobus acidocaldarius* Involves Complex Trehalose Metabolism Utilizing a Novel Trehalose-6-Phosphate Synthase (TPS)/Trehalose-6-Phosphate Phosphatase (TPP) Pathway"**, the trehalose biosynthesis pathways as well as the stress response in *S. acidocaldarius* was intensively studied through gene deletion studies, bioinformatics, and enzymatic activity determination. In this work, I performed the experiments for the construction of the different double and triple gene deletion strains, to measure and compare the intracellular trehalose concentrations under optimal versus stress conditions. Additionally, I determined respective pathway activities in crude extracts of *S. acidocaldarius* and established the method to analyze the pathway activities on transcriptional level via qRT-PCR. Furthermore, I characterized the novel identified TPS/TPP enzymes in detail.

In chapter 3.2, the manuscript with the title **"Enzymatic Synthesis of the Extremolyte Cyclic-2,3-diphosphoglycerate by the Cyclic-2,3-diphosphoglycerate Synthetase from *Methanothermus fervidus*"**, the biosynthetic production of cDPG was closely examined. Regarding the inaccessibility of cDPG for biotechnical application, first results towards the establishment of an enzymatic *in vitro* approach to produce cDPG were presented. For this purpose, I optimized the heterologous enzyme production with respect to codon adaption, expression conditions, and purification procedures. Furthermore, I sufficiently stabilized the recombinant protein for use and storage and finally demonstrated a complete substrate to product conversion in a small laboratory scale.

In chapter 3.3, the manuscript with the title **"Isolation and Characterization of the First Xylanolytic Hyperthermophilic Euryarchaeon *Thermococcus* sp. Strain 2319x1 and Its Unusual Multidomain Glycosidase"**, the screening for novel thermostable polymer degrading biocatalysts from extreme environments was performed. A new xylan degrading hyperthermophilic *Thermococcus* sp. was identified harboring a unique multidomain glycosidase. In this study, I cloned the full-length glycosidase gene as well as several truncated versions, purified them, analyzed the substrate spectrum, and characterized the respective activities in detail. Additionally, I established the methods for fast screening and visualization of hydrolytic activities on substrate agar plates and zymogram gels at high temperatures.

3 Manuscripts

Chapter 3.1

Salt Stress Response of *Sulfolobus acidocaldarius* Involves Complex Trehalose Metabolism Utilizing a Novel Trehalose-6-Phosphate Synthase (TPS)/Trehalose-6-Phosphate Phosphatase (TPP) Pathway

AMERICAN SOCIETY FOR MICROBIOLOGY | Applied and Environmental Microbiology®

PHYSIOLOGY

Salt Stress Response of *Sulfolobus acidocaldarius* Involves Complex Trehalose Metabolism Utilizing a Novel Trehalose-6-Phosphate Synthase (TPS)/Trehalose-6-Phosphate Phosphatase (TPP) Pathway

Christina Stracke,[a] Benjamin H. Meyer,[a,b] Anna Hagemann,[a,c] Eunhye Jo,[d] Areum Lee,[d] Sonja-Verena Albers,[b] Jaeho Cha,[d] Christopher Bräsen,[a] Bettina Siebers[a]

[a]Molecular Enzyme Technology and Biochemistry, Environmental Microbiology and Biotechnology, Centre for Water and Environmental Research, Department of Chemistry, University of Duisburg-Essen, Essen, Germany
[b]Molecular Biology of Archaea, Institute of Biology II, Albert-Ludwig's University of Freiburg, Freiburg, Germany
[c]Institute of Pharmacology and Toxicology, Department of Health, School of Medicine, Witten/Herdecke University, Witten, Germany
[d]Department of Microbiology, Pusan National University, Busan, Republic of Korea

ABSTRACT The crenarchaeon *Sulfolobus acidocaldarius* has been described to synthesize trehalose via the maltooligosyltrehalose synthase (TreY) and maltooligosyltrehalose trehalohydrolase (TreZ) pathway, and the trehalose glycosyltransferring synthase (TreT) pathway has been predicted. Deletion mutant analysis of strains with single and double deletions of Δ*treY* and Δ*treT* in *S. acidocaldarius* revealed that in addition to these two pathways, a third, novel trehalose biosynthesis pathway is operative *in vivo*: the trehalose-6-phosphate (T6P) synthase/T6P phosphatase (TPS/TPP) pathway. In contrast to known TPS proteins, which belong to the GT20 family, the *S. acidocaldarius* TPS belongs to the GT4 family, establishing a new function within this group of enzymes. This novel GT4-like TPS was found to be present mainly in the *Sulfolobales*. The Δ*treY* Δ*treT* Δ*tps* triple mutant of *S. acidocaldarius*, which lacks the ability to synthesize trehalose, showed no altered phenotype under standard conditions or heat stress but was unable to grow under salt stress. Accordingly, in the wild-type strain, a significant increase of intracellular trehalose formation was observed under salt stress. Quantitative real-time PCR showed a salt stress-mediated induction of all three trehalose-synthesizing pathways. This demonstrates that in *Archaea*, trehalose plays an essential role for growth under high-salt conditions.

IMPORTANCE The metabolism and function of trehalose as a compatible solute in *Archaea* was not well understood. This combined genetic and enzymatic approach at the interface of microbiology, physiology, and microbial ecology gives important insights into survival under stress, adaptation to extreme environments, and the role of compati- ble solutes in *Archaea*. Here, we unraveled the complexity of trehalose metabolism, and we present a comprehensive study on trehalose function in stress response in *S. acido- caldarius*. This sheds light on the general microbiology and the fascinating metabolic repertoire of *Archaea*, involving many novel biocatalysts, such as glycosyltransferases, with great potential in biotechnology.

KEYWORDS osmoadaptation, *Archaea*, compatible solutes, hyperthermophiles, *Sulfolobus acidocaldarius*, TPS/TPP pathway, TreT pathway, trehalose glycosyltransferring synthase, trehalose metabolism, trehalose-6-phosphate phosphatase, trehalose-6-phosphate synthase

Microorganisms, including those thriving under hostile conditions, such as (hyper)thermophiles, thermoacidophiles, and extreme halophiles, require efficient adaptation strategies and need to respond to environmental changes and

Citation Stracke C, Meyer BH, Hagemann A, Jo E, Lee A, Albers S-V, Cha J, Bräsen C, Siebers B. 2020. Salt stress response of *Sulfolobus acidocaldarius* involves complex trehalose metabolism utilizing a novel trehalose-6-phosphate synthase (TPS)/trehalose-6-phosphate phosphatase (TPP) pathway. Appl Environ Microbiol 86:e01565-20. https://doi.org/10.1128/AEM.01565-20.

Editor Haruyuki Atomi, Kyoto University

Address correspondence to Bettina Siebers, bettina.siebers@uni-due.de.

Received 29 June 2020

Accepted 17 September 2020

Accepted manuscript posted online 2 October 2020

various stresses. Beside known intrinsic factors for the protection of cellular com- ponents, such as the structural and mechanistic adaptation of DNA, proteins, and membranes, extrinsic factors like compatible solutes are also essential for the adaptation and stress response. Compatible solutes are low-molecular-weight com- pounds that do not interfere with metabolism and can therefore be accumulated in large amounts (1). A variety of different compounds serve as compatible solutes, such as polyols (e.g., mannitol and sorbitol), amino acids and their derivates (e.g., glutamic acid, proline, ectoine, and hydroxyectoine), quaternary ammonium salts (e.g., glycine betaine), and disaccharides (e.g., trehalose, sucrose) (1).

The nonreducing disaccharide trehalose (a-$_D$-glucopyranosyl-1,1-a-$_D$-glucopyranoside) is widely distributed in bacteria and eukaryotes, where it functions in general stress protection against heat shock, osmotic stress, desiccation, and cessation of growth (2). Furthermore, trehalose is used as a reserve carbohydrate in some fungi (spore germination) and insects (flight muscles), and some prokaryotes utilize trehalose as a carbon and energy source (3). The function of trehalose, particularly in eukaryotes (fungi and plants), even extends to signaling and regulation as well as to infection mechanisms in pathogenic fungi and bacteria (2, 4–13). In *Eukarya*, trehalose is mainly synthesized by the trehalose-6-phosphate synthase (TPS)/trehalose-6-phosphate phosphatase (TPP) pathway. Only some fungi and algae were reported to utilize the trehalose phosphorylase (TreP) (for trehalose synthesis pathways, see Fig. S1) (10, 12, 14). Conversely, in Bacteria, four pathways in addition to the widespread TPS/TPP route have been described: (i) the trehalose glycosyltransferring synthase (TreT) pathway, (ii) the maltooligosyltrehalose synthase (TreY)/maltooligosyltrehalose trehalohydrolase (TreZ) pathway, (iii) the trehalose synthase (TreS) pathway, and (iv) the TreP pathway (6, 14). In numerous cases, a bacterial species harbors more than one pathway in parallel, which all contribute to trehalose synthesis, as shown by mutational approaches (15– 24). This led to the assumption that eukaryotes use a single pathway for trehalose synthesis, whereas bacteria employ multiple pathways (6, 10, 14).

Trehalose was also shown to be present in several *Archaea*, mainly in *Crenarchaeota*, including hyperthermophiles of the genera *Thermoproteus* and *Pyrobaculum* and of the order *Sulfolobales* (25). In the thermoacidophiles *Sulfolobus*, *Acidianus*, and *Metallospha- era*, trehalose has been detected as the sole compatible solute, and these organisms lack other common archaeal protective compounds (25). In addition, trehalose was found in some *Euryarchaeota*, such as *Thermoplasma*, and also in some extremely halophilic *Archaea* in which the derivative sulfotrehalose was also detected (25–27). As in *Bacteria*, several pathways for trehalose synthesis have also been detected in *Archaea*, and they often occur in one organism in parallel (14, 28). Biochemical data have been reported for the TPS/TPP pathway in *Thermoplasma acidophilum* (29) and in halophiles (27), and an unusual complex forming bifunctional TPSP, a fusion protein containing a TPS and TPP domain, was characterized from *Thermoproteus tenax* (30). TreT enzymes from *Thermococcus litoralis* (31) and *T. tenax* (28) and the TreY and TreZ enzymes from *Sulfolobus* and *Metallosphaera* spp. (32–43) have been analyzed in detail. TreS has so far been found only in *Picrophilus torridus* (42, 43). For trehalose degradation, TreH enzymes have also recently been identified (44–48). However, most of the information available on archaeal trehalose metabolism is based on sequence infor- mation and analyses of recombinant proteins. Thus, in contrast to *Eukarya* and *Bacteria*, where numerous mutational analyses have been performed (see above), the general function of trehalose in *Archaea* is so far unknown, and the complexity of trehalose metabolism, including the *in vivo* importance of the different synthesis pathways, is not yet understood.

The (hyper)thermoacidophilic *Crenarchaeon Sulfolobus acidocaldarius*, which has optimal growth at pH 2 to 3.5 and temperatures of 75 to 80°C, was shown to accumulate trehalose intracellularly (47). The TreY/TreZ pathway for synthesis has been biochemically characterized (33–35, 39), whereas the TreT pathway was only proposed from sequence data (28) (Fig. 1). Here, we used a combined genetic and biochemical approach and found that both pathways are operative *in vivo* but do not account for trehalose synthesis alone. Instead, a novel, unusual TPS/TPP pathway was identified and characterized and its *in vivo* relevance was confirmed. Furthermore, the accumulation of trehalose by the concerted action of all three synthesis pathways in response to salt stress was demonstrated.

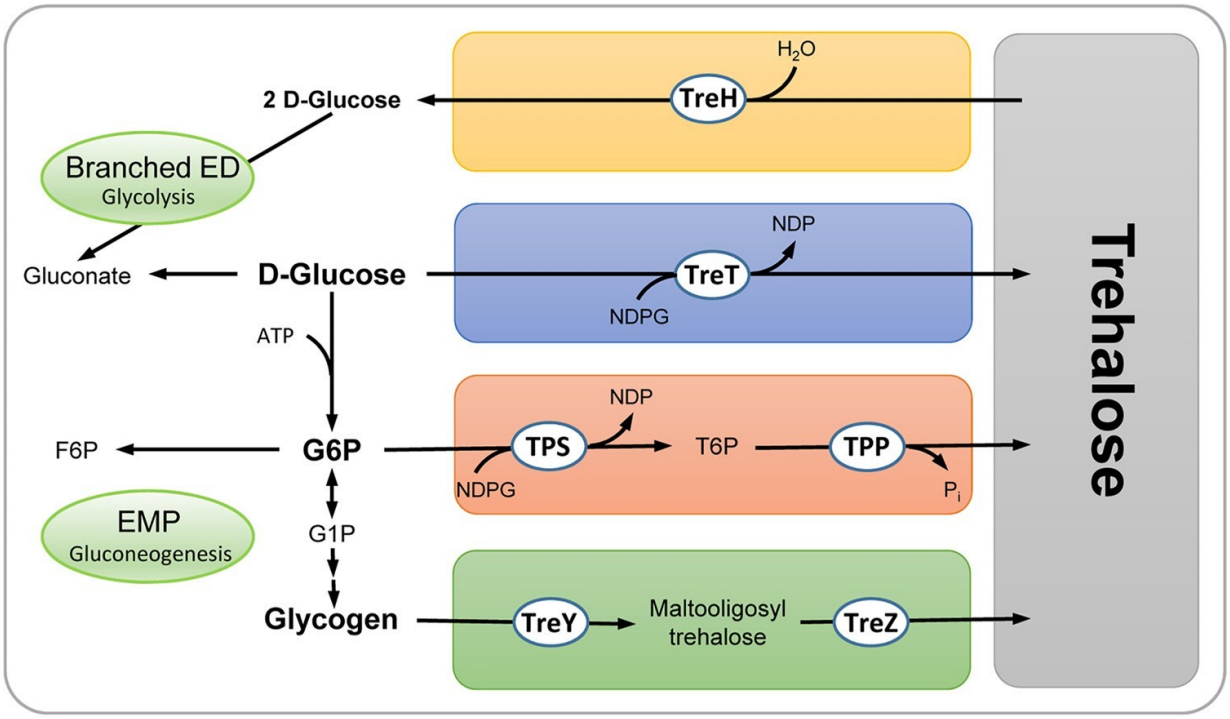


FIG 1 Overview of identified pathways for trehalose synthesis and degradation in *S. acidocaldarius*. *S. acidocaldarius* degrades D-glucose via a modified branched Entner-Doudoroff (ED) pathway, whereas the Embden-Meyerhof- Parnas (EMP) pathway is used only for gluconeogenesis (66). Glycogen has been reported as a central carbon storage compound in *S. acidocaldarius* (61), and hexokinase activity (broad substrate specificity) has been reported for *S. tokodaii* (67). The TreY/TreZ (34, 35) and TreT (28) trehalose synthesis pathways and the novel TPS/TPP pathway as well as the trehalases for trehalose degradation (47) are depicted. Abbreviations: TreY, maltooligosyl- trehalose synthase; TreZ, maltooligosyltrehalose trehalohydrolase; TPS, trehalose-6-phosphate synthase; TPP, trehalose-6-phosphate phosphatase; TreT, trehalose glycosyltransferring synthase; TreH, trehalase; UDPG, UDP glucose.

RESULTS

Complexity of the trehalose metabolism in *S. acidocaldarius*—a novel TPS/TPP pathway. To elucidate the function of the established TreY/TreZ as well as the predicted TreT pathway in trehalose metabolism of *S. acidocaldarius*, single- and double-gene-deletion strains (Δ*treY* [*saci_1436*], Δ*treT* [*saci_1827*], and Δ*treY* Δ*treT*) were constructed. The *treY* gene, which generates the terminal α 1-1 linkage, rather than *treZ*, which releases trehalose from glycogen, was deleted to exclude interference with glycogen metabolism. Compared to the reference strain MW001, the Δ*treY* and Δ*treT* single-deletion strains as well as the double-deletion strain (Δ*treY* Δ*treT*) showed no altered phenotype under standard growth conditions (Fig. 2A). The intracellular treh- alose concentration of 0.096 μmol mg protein^{-1} in the parental strain was only slightly reduced to 0.040 μmol mg protein^{-1} and 0.032 μmol mg protein^{-1} in the Δ*treY* and Δ*treT* strains, respectively (Fig. 2B). In contrast, the trehalose concentration in the Δ*treY* Δ*treT* double mutant was more significantly reduced, to 0.014 μmol mg protein^{-1} (Fig. 2B). These results showed the significance of both pathways for trehalose synthesis and confirmed that the TreT pathway, which had only been predicted so far, is operative *in vivo*. However, the residual amounts of trehalose in the Δ*treY* Δ*treT* mutant (roughly 10 to 15% that of the reference strain) indicated that *S. acidocaldarius* harbors an addi- tional, third, yet-unknown trehalose synthesis pathway.

For the identification of the additional trehalose synthesis pathways in *S. acidocal- darius*, BLAST searches (49) with the *P. horikoshii* TreT revealed an additional TreT homologue with less similarity and sequence coverage (Saci_1249; E values, 3e^{-15} [*P. horikoshii*]) than the previously predicted TreT (Saci_1827; E values, 7e^{-117} [*P. horikoshii*]). The gene (1,335 bp) encodes a protein (444 amino acids) with a calculated molecular mass of 50.3 kDa.

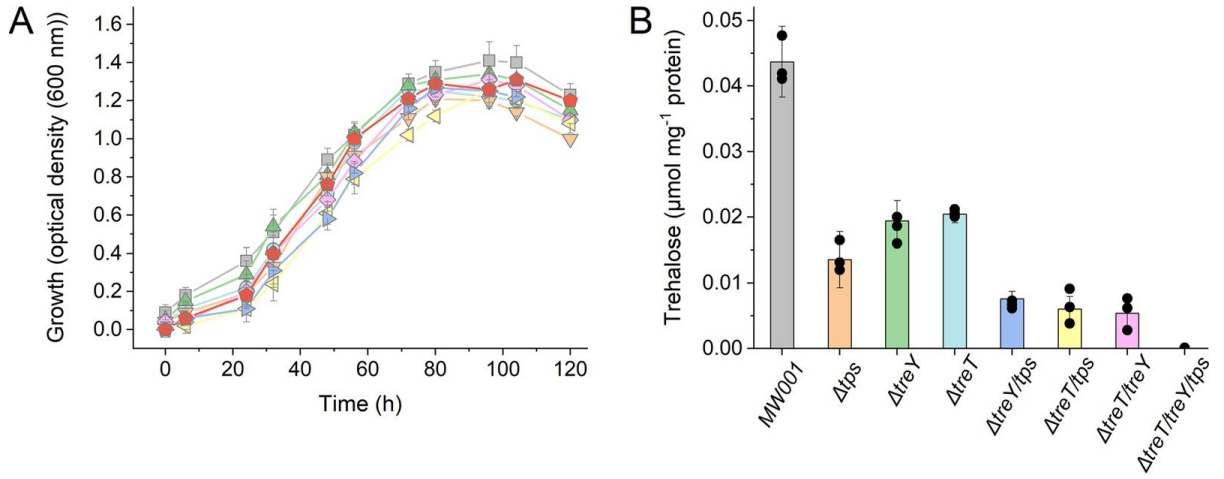


FIG 2 Growth and intracellular trehalose concentrations in *S. acidocaldarius* MW001 and mutants with deletions of the different trehalose synthesis pathways under standard growth conditions. (A) Growth curves of the parental strain MW001 (gray squares) and mutants with the deletions Δ*treT* (blue circles), Δ*treY* (green triangles), Δ*tps* (orange inverted triangles), Δ*treT* Δ*treY* (pink rectangle), Δ*treT* Δ*tps* (yellow left triangles), Δ*treY* Δ*tps* (purple right triangles), and Δ*treT* Δ*treY* Δ*tps* (red circles) under standard growth conditions. (B) Intracellular trehalose concentrations (in micromoles per milligram of protein) in the parental strain MW001 and of the different deletion mutants in the exponential growth phase. The means and standard deviations for three biological replicates (n = 3) are shown.

The protein is annotated as a GT-B-fold superfamily glycosyltransferase of the GT4 family (www.cazy.org [50]). Homology detection and structure prediction tools (HHPred and Phyre [51, 52]) revealed that Saci_1249 is homologous to both TreT and TPS enzymes.

TPS (Saci_1249) was expressed and purified from *Escherichia coli* (Fig. S2A). Under denaturing conditions, the enzyme showed a molecular mass of 50 kDa, and under native conditions via gel filtration, it showed a molecular mass of 100 kDa, indicating a homodimeric structure. Notably, despite the similarity to TreT, no TreT activity with glucose and UDPG/ADPG could be determined. Instead, TPS activity in the presence of Mg^{2+} with glucose-6-phosphate (G6P) as the glycosyl acceptor and UDPG or ADPG as the glycosyl donor forming trehalose 6-phosphate was confirmed (Fig. S3). With UDPG as the glycosyl donor, the enzyme showed a V_{max} of 6 U mg^{-1} and K_m values of 3.2 mM (G6P) and 3.8 mM (UDPG), corresponding to a k_{cat} of 4.9 s^{-1}. For ADPG, the V_{max} was 4.1 U mg^{-1} and the K_m was 1.8 mM. The *S. acidocaldarius* TPS was specific for G6P. Glucose 1-phosphate, fructose 1,6-bisphosphate, fructose 1-phosphate, and ribose 5-phosphate could not serve as acceptor substrates (Fig. S3D). The pH optimum of the enzyme was pH 7.0 (Fig. S3E). Also, the thermostability was analyzed; after 2 h at 70°C or 80°C, 68% or 30% residual activity was observed, respectively (Fig. S3F).

For a functional TPS/TPP pathway, a TPP catalyzing the dephosphorylation of trehalose 6-phosphate to trehalose is required. However, no homologues of characterized TPPs were identified in the *S. acidocaldarius* genome, but 12 members of the HAD superfamily to which canonical TPPs belong were found. Further bioinformatics analyses reduced the number to five promising candidates (Saci_0016, GenBank no. AAY79446; Saci_0094, GenBank no. AAY79521; Saci_0239, GenBank no. AAY79656; Saci_0930, GenBank no. AAY80292; and Saci_1518, GenBank no. AAY80837). We cloned and expressed these genes and confirmed TPP activity, i.e., trehalose formation from trehalose-6-phosphate (T6P), for the purified monomeric Saci_0016 (27 kDa) via thin-layer chromatography (TLC) (Fig. S2B, S4, and S5). In addition, other sugar phosphates, i.e., fructose 6-phosphate, mannose 6-phosphate, and glucose 6-phosphate, were converted, indicating a broader substrate spectrum (Fig. S5A). The activity of the purified enzyme with T6P as the substrate was 8.7 U mg^{-1}. With *p*-nitrophenyl phosphate (pNPP) as the artificial phosphatase substrate, the specific phosphatase activity was 0.06 U mg^{-1}. Furthermore, the combined trehalose-forming activity from UDPG and G6P via both recombinant enzymes (Saci_1249 and Saci_0016) was observed (Fig. S5B).

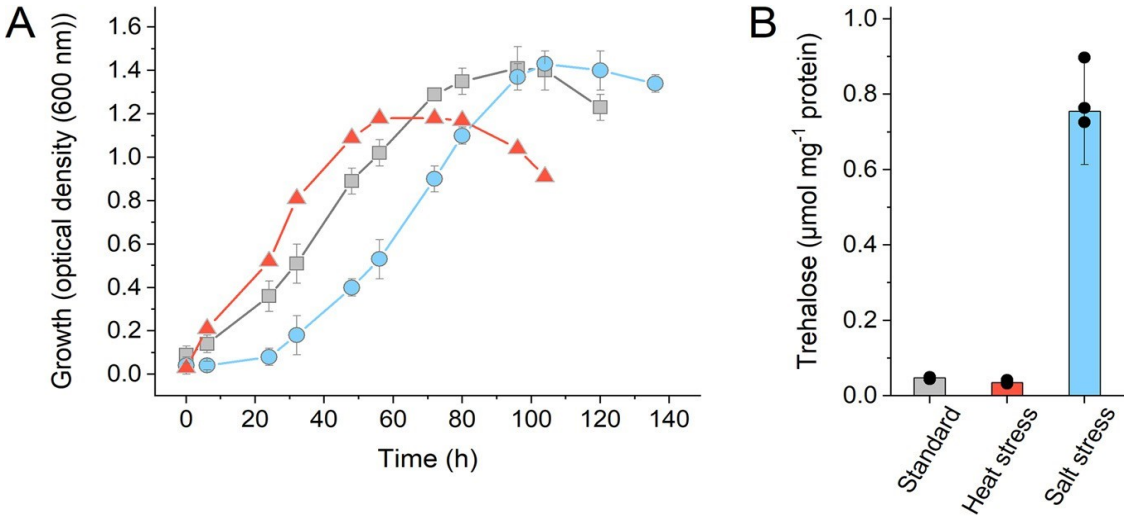


FIG 3 Growth and intracellular trehalose concentration of *S. acidocaldarius* MW001 under standard growth conditions and under heat and salt stress conditions. (A) Growth curves of the parental strain MW001 under optimal growth conditions (gray squares), under heat stress at 83°C (red triangles), and under salt stress in the presence of 250 mM NaCl (blue circles). Cells were adapted to the respective stress conditions for three passages. (B) Intracellular trehalose concentrations (in micromoles per milligram of protein) in cells from the exponential growth phase grown under standard conditions (gray), under heat stress (red), and under salt stress (blue). The means and standard deviations for three biological replicates (n = 3) are shown.

To confirm the function of the TPS/TPP pathway in trehalose metabolism of *S. acidocaldarius*, we constructed the single Δ*tps* mutant (Δ*saci_1249*), the double mutants (Δ*treY*Δ*tps* and Δ*treT*Δ*tps*), and the triple mutant (Δ*treY*Δ*treT*Δ*tps*), in which all three assumed pathways for trehalose formation were deleted. As described for the *treY* and *treT* single and double mutants, no altered growth phenotypes under standard condi- tions for the *tps* single, double, and triple mutants were observed (Fig. 2). The trehalose content compared to that of the parental strain MW001 was decreased in the Δ*tps* single mutant to 0.032 µmol mg protein^{-1}. This decrease was even more pronounced in the double mutants (Δ*treY* Δ*tps* and Δ*treT*Δ*tps*) (0.019 to 0.016 µmol mg protein^{-1}), and no trehalose was detected in the triple mutant. Hence, all three pathways con- tribute to the overall trehalose content of the cells. However, since no altered phenotype was observed, trehalose is dispensable under standard conditions, and thus the role of trehalose was not clarified by this set of experiments.

Physiological role of trehalose in *S. acidocaldarius*. Therefore, we analyzed the impact of the *de novo* trehalose formation under different stress conditions. Under heat stress at 83°C (the optimum growth temperature is 75°C), the growth rate of adapted *S. acidocaldarius* MW001 cells was not severely affected, but the final optical density at 600 nm (OD_{600}) of approximately 1.0 was reduced compared to that at 75°C (OD_{600}, 1.4). However, no trehalose accumulation was observed under heat stress in MW001 (0.087 µmol mg protein^{-1} at 83°C compared to 0.097 µmol mg protein^{-1} at 75°C) (Fig. 3).

For cold shock, *S. acidocaldarius* MW001 was pregrown at 75°C and then transferred to 65°C, resulting in significant growth retardation. However, the cold shock did not enhance the trehalose content of the cells (0.050 µmol mg protein^{-1}) (Fig. S6). These results indicate that trehalose is not involved in thermoprotection in *S. acidocaldarius*. We therefore analyzed the influence of trehalose formation under elevated salinities.

All mutant strains constructed were adapted to high salinity (250 mM NaCl, three passages), and the growth phenotype and the intracellular trehalose concentrations were determined in comparison to those of the parental strain, MW001 (Fig. 3 and 4). In the presence of 250 mM NaCl, the reference strain MW001 showed a 12-fold-higher trehalose accumulation (1.12 µmol mg protein^{-1}) than under standard growth condi- tions (0.097 µmol mg protein^{-1}) (Fig. 3), although the growth was only slightly delayed. Furthermore, the single-gene-deletion strains (Δ*treY* and Δ*treT*) and the double-deletion strain (Δ*treT*Δ*treY*) showed no altered growth phenotype under salt stress conditions compared to

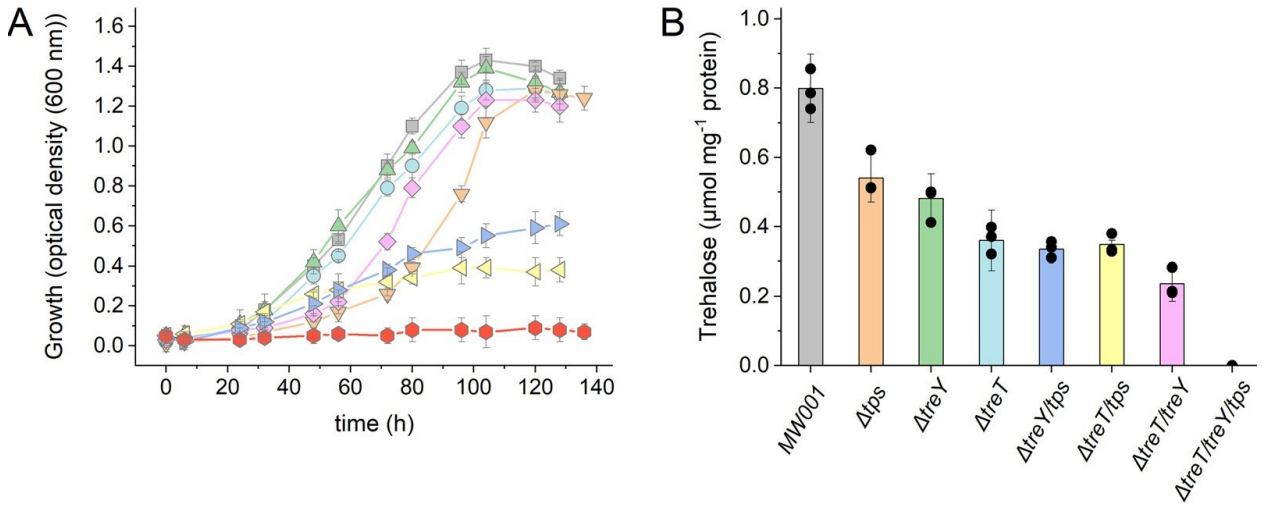


FIG 4 Growth and intracellular trehalose concentration of *S. acidocaldarius* MW001 and deletion mutants of the different trehalose synthesis pathways under salt stress in the presence of 250 mM sodium chloride. (A) Growth curves of parental strain MW001 (gray squares) and mutants with the deletions Δ*treT* (blue circles), Δ*treY* (green triangles), Δ*tps* (orange inverted triangles), Δ*treT* Δ*treY* (pink diamonds), Δ*treT* Δ*tps* (yellow left triangles), Δ*treY* Δ*tps* (purple right triangles), and Δ*treT* Δ*treY* Δ*tps* (red circles) under salt stress. Cells were adapted to salt stress conditions for three passages. (B) Intracellular trehalose concentrations (in micromoles per milligram of protein) in the parental strain MW001 and the different deletion mutants in the exponential growth phase. The means and standard deviations for three biological replicates (n = 3) are shown.

MW001 (Fig. 4A). Conversely, the Δ*tps* single mutant and the double mutants with the *tps* deletion (i.e., Δ*treT* Δ*tps* and Δ*treY* Δ*tps* strains) exhibited a reduced-growth phenotype (Fig. 4A). Growth of the Δ*tps* strain was significantly delayed, although the growth yield was nearly unchanged (Fig. 2A and 4A). However, the Δ*treT* Δ*tps* and Δ*treY* Δ*tps* mutants grew much more slowly, with clearly reduced growth yields (Fig. 4A).

The Δ*treY* Δ*treT* Δ*tps* strain did not contain any measurable trehalose under standard conditions (Fig. 2B). In the presence of 250 mM NaCl, the triple mutant showed no growth and also no intracellular trehalose (Fig. 4). In the single mutants (Δ*treY*, Δ*treT*, and Δ*tps*), the intracellular trehalose concentrations were around 0.499 to 0.714 µmol mg protein^{-1} and thus were reduced by nearly 50% compared to that in the parental strain (Fig. 4B). The double deletions (Δ*treT* Δ*treY*, Δ*treT* Δ*tps*, and Δ*treY* Δ*tps*) resulted in even less trehalose (0.473 to 0.337 µmol mg protein^{-1}) (Fig. 4B). These studies highlight the essential role of trehalose in salt stress response in *S. acidocaldarius*. The complementation of the triple mutant with either of the pathways partially restored growth under salt stress and trehalose formation (Fig. S7). The complementation of the triple mutant with pSVAmz-SH10_tps was expected to result in a growth phenotype resembling that of the Δ*treT* Δ*treY* double mutant, which shows only a slightly reduced growth at high salinities (Fig. 4). The observed growth defect of the complemented strain (Fig. S7) is most likely due to different transcription levels of *tps*, between the native and the constitutively low-expressed promoter of the complementation plasmid. However, the addition of trehalose to the medium did not restore the growth of the triple mutant. Also, *S. acidocaldarius* is not able to utilize trehalose as the sole source of carbon and energy, although a functional trehalase is present, and the addition of trehalose under standard conditions did not result in an altered intracellular trehalose level (data not shown). These results strongly suggest that extracellular trehalose cannot be taken up from the medium and thus that *S. acidocaldarius* relies on the *de novo* synthesis of trehalose for salt stress response.

To gain further insights into the response of *S. acidocaldarius* to elevated salinities, the induction of gene expression was studied by quantitative reverse transcriptase PCR (qRT-PCR) in the parental strain MW001 for the transcription of *treY*, *treT*, and *tps* genes upon NaCl addition to growing cultures (salt shock). Samples were taken from the control and salt-induced cultures over a period of 8 h for transcription analysis via qRT-PCR and determination of intracellular trehalose concentration (Fig. 5). After salt stress induction, cell growth was nearly completely arrested and was recovered only after about 60 h of further incubation (Fig. 5A).

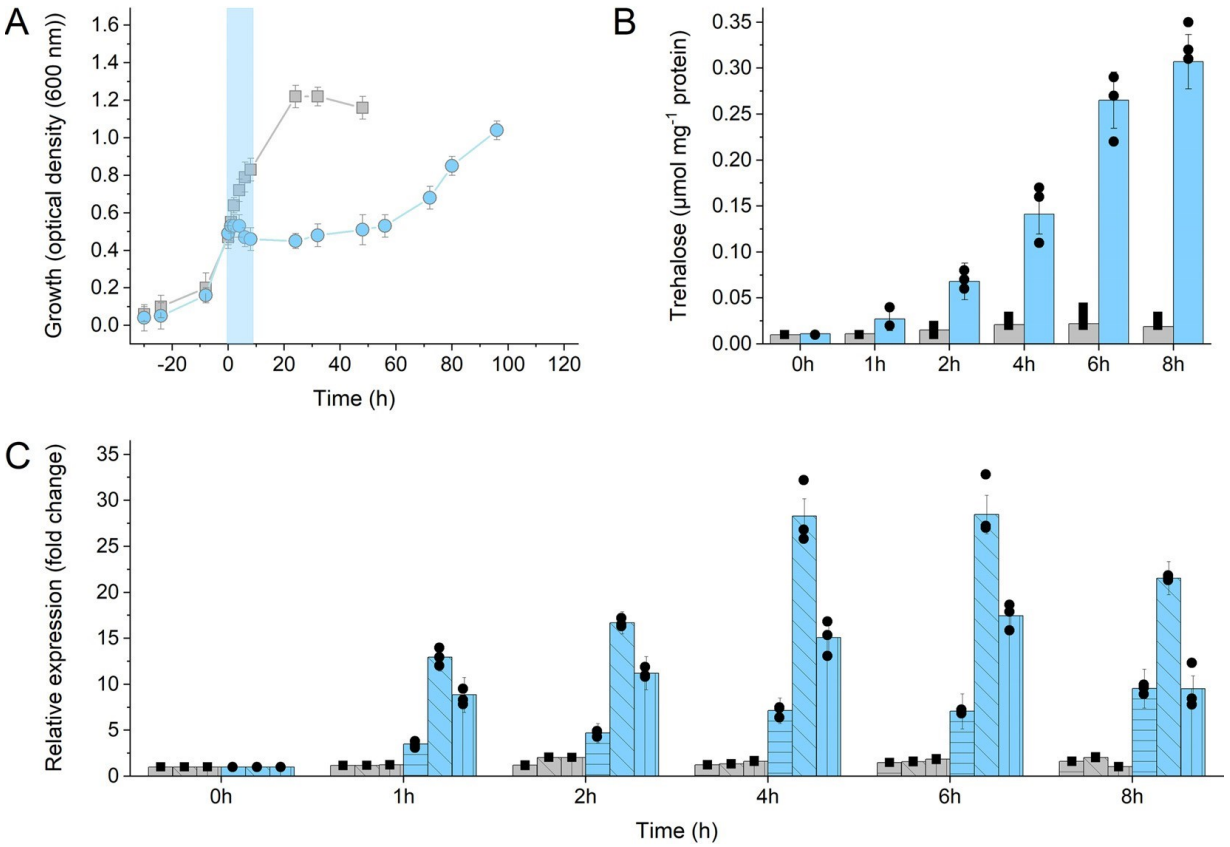


FIG 5 Growth, intracellular trehalose concentrations, and relative expression levels of the trehalose synthesis genes in *S. acidocaldarius* MW001 under optimal growth conditions and under salt shock conditions. (A) Growth curves of *S. acidocaldarius* parental strain MW001 under standard growth conditions (gray squares) and under salt shock conditions (blue circles). The cultures were grown under standard growth conditions until exponential growth phase (OD_{600} 0.5–0.6); then, either preheated Brock medium (control) or preheated Brock medium plus NaCl was added (250 mM final concentration). At the times indicated, 10-ml samples were taken for trehalose determination and total RNA isolation for cDNA synthesis. (B) Intracellular trehalose concentration (in micromoles per milligram of protein) in *S. acidocaldarius* MW001 grown under standard (gray bars) and salt shock (blue bars) conditions. (C) Differential expression of *treY* (horizontally hatched bars), *treT* (diagonally hatched bars), and *tps* (vertically hatched bars) in *S. acidocaldarius* MW001 under standard conditions (gray bars) and upon salt shock (blue bars). After RNA isolation and cDNA synthesis, qRT-PCR analysis was performed using specific primers for the *treY*, *treT*, and *tps* genes. Relative transcript expression levels of each gene were normalized to that of the internal control gene *secY*. For the growth curves (A) and intracellular trehalose determination (B), the means and standard deviations for three biological replicates (n=3) are shown, and for qPCR (C), means and standard deviations for three technical replicates (n=3) are shown.

However, upon salt stress, the trehalose content of the cells continuously increased over the observed 8-h period. As soon as 1 h after induction, a 1.7-fold-higher trehalose concentration was observed in the salt- treated samples, and after 8 h, an 11-fold-higher concentration was observed (Fig. 5B). Accordingly, transcription of all three trehalose genes was induced (Fig. 5C). The qRT-PCR experiments revealed that in salt-stressed cultures, *treT* transcription showed the fastest response (25-fold upregulation after 6 h), followed by *treY* transcription (17-fold upregulated after 6 h), whereas *tps* transcription increased much more slowly but continuously over time (10-fold upregulation after 8 h) (Fig. 5C). This salt-mediated induction of all three trehalose-synthesizing pathways accompanied by trehalose for- mation confirms the function of trehalose in salt stress response in *S. acidocaldarius*.

Furthermore, the induction of the trehalose synthesis pathways was confirmed by pathway activities in crude extracts of salt (250 mM NaCl)-stressed cells of *S. acidocaldarius* MW001. For the TreY/TreZ, TreT, and TPS/TPP pathways, crude extract activities of 3.1 U mg^{-1}, 0.071 U mg^{-1}, and 0.037 U mg^{-1}, respectively, were measured (Fig. S8). In crude extracts of cells grown under standard growth conditions, the respective pathway activities were not detectable using the same assay conditions. In the single- deletion strains (Δ*treT*, Δ*treY*, and Δ*tps*), the corresponding activities were not detected, thus confirming the successful deletion of the respective genes.

DISCUSSION

A novel GT4-like TPS in *S. acidocaldarius*. Unexpectedly, *S. acidocaldarius* could still synthesize trehalose after deletion of the *treY* and *treT* genes, indicating the presence of a third trehalose synthesis pathway. BLAST searches identified Saci_1249 as a distant TreT homologue in *S. acidocaldarius*, which, however, also showed structural homology to TPS enzymes. This is in accordance with the CAZy classification of Saci_1249 as a GT-B-folded protein retaining glycosyltransferase activity as well as the homodimeric structure (100 kDa), which applies to both TreTs and TPSs (www.cazy.org [50]). In addition, the alignment of Saci_1249 with TreT and TPS sequences shows — according to the GT-B affiliation — some conservation, especially in amino acid residues involved in donor substrate binding (Fig. S9). However, the characterization of Saci_1249 clearly revealed TPS but no TreT activity. Deletion of the *tps* gene resulted in the loss of TPS activity in crude extracts of salt-stressed cells. The triple deletion Δ*treY* Δ*treT* Δ*tps* abolished trehalose formation, and the resulting mutant was not able to grow under salt stress, clearly demonstrating the function of Saci_1249 as a TPS in trehalose biosynthesis. Notably, like TreT enzymes, Saci_1249 is annotated as a member of the GT4 family and thus seems distinct from classical TPSs, which are classified as GT20 family members (www.cazy.org [50]).

Close homologues of the newly identified TPS (Saci_1249) could be identified only within the TACK superphylum, mainly in the *Sulfolobales*, with only very few candidates scattered in the *Euryarchaeota*. Thus, Saci_1249 represents a novel type of GT4-like TPS with a narrow distribution, mainly in the *Sulfolobales*. To our knowledge, this is the first report of TPS of the GT4 family, which is one of the largest GT families, with ~150,000 sequences comprising numerous enzyme classes, also including the TreTs (CAZy data- base).

Moreover, no homologue of canonical TPPs was identified, indicating not only a novel type of TPS in *S. acidocaldarius* but also a novel TPP. Thus, a completely unusual TPS/TPP pathway is operative in *S. acidocaldarius*, which is likely restricted to *Sulfolo- bales*. For the HAD (haloacid dehydrogenase) superfamily member Saci_0016, TPP activity was confirmed, and hence, the enzyme represents a potential TPP candidate. Interestingly, from BLAST searches, it appeared that the distributions of TPS and TPP homologues are roughly similar, and this cooccurrence might support the functional correlation. However, future experiments, including those using gene deletions, are needed to confirm its role in the trehalose metabolism.

Function of trehalose in *S. acidocaldarius*. The presence of trehalose and the biochemistry of some of the synthesis pathways are well established in *Archaea*, and recently, unusual trehalases, presumably for trehalose degradation, were also detected in *S. acidocaldarius* (for reviews, see references 30 and 47). The intracellular trehalose concentration determined in this work for *S. acidocaldarius* is well within the range described previously for *Sulfolobales* under standard conditions (25, 47). Here, we show that trehalose is accumulated and essential for growth of *S. acidocaldarius* under salt stress rather than under temperature stress, to increase the osmolarity of the cytoplasm and thus prevent water efflux from the cells (1).

In other *Archaea*, trehalose accumulation in response to enhanced salinity was also shown for *T. litoralis* (53). However, in this organism, alternative compatible solutes were detected (i.e., mannosylglycerate [MG] and di-myo-inositol-1,1-phosphate [DIP]), and the significance of trehalose remains unclear. The main compatible solute formed upon salt stress in *Thermococcales* seems to be MG. Furthermore, it is not yet clear whether trehalose is synthesized *de novo* or whether it is taken up from the growth medium (53, 54). Also, in halophilic archaea, trehalose (and sulfotrehalose) was de- tected with the TPS/TPP pathway as the major synthesis route. However, haloarchaea mainly utilize the “salt-in” strategy and accumulate potassium salts/chloride to coun- terbalance variable and at least temporarily high environmental salt concentration. In these organisms, trehalose seems to be accumulated in response to decreasing salin- ities in order to allow survival under low-salt conditions (27). In accordance with the findings for *S. acidocaldarius*, trehalose accumulation under unfavorable temperature conditions has so far not been reported in archaea. In *T. litoralis* and *Pyrococcus furiosus*, upon temperature stress, mainly DIP is formed as the key solute (53, 55).

Also in *Bacteria*, if trehalose is used as a compatible solute, it plays a predom- inant role in salt and osmotic stress adaptation (15–17). However, it was also shown to be involved in heat stress response, for example, in *E. coli*, *Mycobacterium smegmatis* (18), and *Rhizobium etli* (19), as well as to perform additional functions in desiccation tolerance in a couple of rhizobial species, where it is also involved in symbiosis/nodulation (20–23). Although thermal stress response in *S. acidocaldarius* was shown to be independent from trehalose accumulation, further functions of trehalose in addition to salinity stress, e.g., in desiccation tolerance or under unfavorable pH or growth conditions, cannot be ruled out and will be addressed in future studies. There are already indications that trehalose is formed in the station- ary growth phase (56) and that the TreY/TreZ pathway is upregulated under starvation (57). The recently reported presence of trehalases indicates that the organism is at least able to degrade trehalose (47), and the close relative *Saccha- rolobus solfataricus* has already been shown to utilize trehalose as the sole carbon and energy source (58).

Function of multiple trehalose synthesis pathways. For many *Euryarchaeaota*, like *Thermococcales* and methanogens, several compatible solutes, such as DIP, MG, and cyclic 2,3-diphosphoglycerate (cDPG), have been described, and trehalose seems to play only a minor or supplementary role, if any. These organisms often harbor only one trehalose synthesis pathway. In contrast, many *Archaea*, which utilize trehalose as a main or exclusive compatible solute, have the tendency to possess multiple (i.e., at least two) synthesis pathways. In general, the presence of multiple synthesis pathways in single organisms is thought to reflect the importance of this compatible solute (12). For *T. tenax*, the unidirectional TreT and the TPSP pathway have been described (28, 30), which from genome analyses is also deduced for other *Thermoproteales*, like *Pyrobaculum* spp. Also, in *S. acidocaldarius*, trehalose appears to be the exclusive compatible solute. The previously reported functional TreY/TreZ pathway was confirmed (33), and the proposed TreT pathway (encoded by *saci_1827*) (28) and the novel TPS/TPP pathway were demonstrated in this study to be operative *in vivo*.

In *S. acidocaldarius*, all three pathways were found to operate in a concerted manner and contribute to the overall trehalose content of the cells. Notably, complementation of the triple mutant with either one of the pathways restored growth under salt stress and trehalose formation. Interestingly, the Δ*treT* Δ*tps* and Δ*treY* Δ*tps* mutants exhibited a more severely altered growth phenotype than the Δ*treT* Δ*treY* mutant but no difference in the trehalose concentration. This might indicate additional functions of the TPS/TPP pathway enzymes in cellular metabolism, e.g., through substrate promis- cuity. Also, a role of the pathway intermediate T6P as a signaling molecule was described especially for fungi and plants (6, 8, 12).

In addition, numerous *Bacteria* employ at least two pathways, TPS/TPP and TreY/TreZ, for trehalose synthesis, although the majority appear to rely solely on the TPS/TPP pathway, which is also the sole biosynthetic route in eukaryotes (1, 14). In *Bacteria*, a couple of mutational approaches have been reported. On the one hand, the results indicate that both pathways contribute to trehalose biosynthesis (15, 20, 22, 59), suggesting a common strategy, with multiple synthesis pathways being responsive to a given stressor like high salinity. On the other hand, e.g., in *Corynebacterium glutami- cum*, the TreY/TreZ pathway predominates under hyperosmotic conditions over the TPS/TPP route (60). Additionally, the TreY/TreZ pathway relies on glycogen/starch as a precursor, which is formed only when sufficient amounts of carbon and energy are available. Thus, the physiological conditions also might trigger the pathway signifi- cance, and for *C. glutamicum*, the interconnection of the TreY/TreZ pathway with the glycogen/starch metabolism has been shown (24). Also, several *Archaea* have been described to utilize glycogen as a storage compound, among them *Sulfolobus* and *Thermoproteus* spp. (61), which also contain the TreY/TreZ pathway. Thus, one might argue that the different pathways starting from different precursors permit the meta- bolic flexibility to ensure sufficient synthesis of the compatible solute under (rapidly) changing physiological conditions. Future experiments will aim to elucidate the influ- ence of the growth phase and physiological state of the cells as well as of growth substrates on the utilization of the different trehalose synthesis routes.

In conclusion, we unraveled the complexity of trehalose metabolism in *S. acidocaldarius*. Three pathways are involved in trehalose synthesis, the previously reported TreY/TreZ pathway, the TreT pathway, and a novel TPS/TPP pathway. The latter is characterized by novel enzymes that show no similarity to classical counterparts, comprising a TPS of the GT4 enzyme family as well as a predicted TPP of the HAD enzyme family. Furthermore, we demonstrate that trehalose is formed in response to salt stress in *S. acidocaldarius*.

MATERIALS AND METHODS

Strains and growth conditions. The *E. coli* strains were grown in Luria-Bertani medium. *S. acidocaldarius* MW001 (uracil auxotroph mutant) (62) was grown in Brock's basal medium at 75°C and pH 3.5 (63) supplemented with 0.2% (wt/vol) N-Z-Amine, 0.1% (wt/vol) dextrin, and 10 μg/ml of uracil. Salt stress experiments were performed in the presence of 250 mM sodium chloride, and temperature stress was tested at 83°C. To this end, cells were adapted for three passages to grow in the presence of elevated salt concentrations or temperature. For the salt shock experiment, cells were grown under standard conditions to an OD_{600} of 0.6 to 0.7 and either prewarmed Brock medium (control) or prewarmed Brock medium with NaCl (250 mM final concentration) was added. For the cold shock experiment, the cells were grown at 75°C (control) and at an OD_{600} of 0.6, the cultures were transferred to 65°C.

Cloning, heterologous expression, and purification of the recombinant proteins. The *tps* (*saci_1249*, GenBank no. AAY80595.1) and *tpp* (*saci_0016*, GenBank no. AAY79446.1) genes were ampli- fied via PCR from genomic DNA of *S. acidocaldarius* as the template (for primers, see Table 1). PCR fragments were cloned into the vector pET15b (Novagen, USA). Genes were expressed in *E. coli* BL21(DE3)-CodonPlus induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG). Harvested cells (6,000 × *g*, 20 min, 4°C) were resuspended (1 g cells in 5 ml of 50 mM NaH_2PO_4–300 mM NaCl, pH 8) and disrupted with a French press (20.000 lb/in^2). After centrifugation (4°C, 25,000 × *g*, 45 min), the super- natant was subjected to heat precipitation (75°C, 30 min) and centrifuged. His tagged proteins were purified from the supernatant via Protino Ni-TED (Tris-carboxymethyl ethylene diamine) (Macherey-Nagel, Düren, Germany) following the manufacturers' instructions. Purified proteins were dialyzed (30 mM Tris–300 mM NaCl, pH 7) and concentrated via ultrafiltration (Vivaspin concentrator; molecular weight [MW] cutoffs, 30 kDa [TPS] and 10 kDa [TPP]; Sartorius, Göttingen, Germany). The proteins were further purified via size exclusion chromatography (Superdex 200 HiLoad 26/60 columns [GE Healthcare Life Sciences, Freiburg, Germany]; buffer, 30 mM Tris–300 mM NaCl, pH 7) and concentrated via ultrafil- tration. The pure proteins were used for enzymatic characterization. Purification and molecular mass were monitored by SDS-polyacrylamide gel electrophoresis.

Generation of deletion mutants of *S. acidocaldarius*. For in-frame markerless deletion mutations in *S. acidocaldarius* MW001, plasmids were constructed essentially as described in reference 62. Briefly, the ~500-bp upstream and downstream regions of each of the genes (*tps* [*saci_1249*, GenBank no. AAY80595.1], *treT* [*saci_1827*, GenBank no. AAY81133.1], and *treY* [*saci_1436*, GenBank no. NC_007181.1, 1225387 to 1227549]) were amplified by PCR, and the respective up- and downstream PCR products were fused via overlap PCR. The resulting products were cloned into pSVA407 or pSVA406 (for plasmids and strains, see Table 1). Correct cloning was confirmed by sequencing. To prevent plasmid degradation in the recipient strain, all plasmids were methylated by transformation into *E. coli* ER1821 and after reisolation were transformed into *S. acidocaldarius* MW001. First selection was performed via uracil auxotrophy and second selection via 5-FOA (5-fluoroorotic acid). Successful gene knockouts were confirmed by PCR and sequencing. For the construction of the complementation plasmids, PCR products of *tps*, *treY*, and *treT* were amplified from genomic DNA from *S. acidocaldarius* and cloned into the pSVAmzSH10 expression vector (for primers and plasmids, see Table 1). The respective methylated plasmids were transformed into the MW001 triple deletion strain (Δ*treY* Δ*treT* Δ*tps*).

Enzyme assays. (i) TreT. TreT activity was measured discontinuously in 50 mM HEPES-KOH (pH 7) with 10 mM $MgCl_2$, as glucose (10 mM)- and ADP glucose (ADPG, 10 mM)- or UDP glucose (UDPG, 10 mM)-dependent formation of ADP or UDP at 70°C. The formation of UDP or ADP was detected via pyruvate kinase (PK) (rabbit muscle; 8 U) (Merck, Darmstadt, Germany) and L-lactate dehydrogenase (L-LDH) (rabbit muscle; 4 U) (Merck, Darmstadt, Germany) at 55°C in 50 mM HEPES-KOH (pH 7.5), 10 mM $MgCl_2$, 2 mM phosphoenolpyruvate, 0.5 mM NADH. In the reverse direction, the glucose formation from trehalose (10 mM) and ADP or UDP (10 mM) was assayed discontinuously in 50 mM HEPES-KOH (pH 7) with 10 mM $MgCl_2$, at 70°C. Glucose was determined enzymatically using the glucose dehydrogenase (GDH) from *Pseudomonas* sp. (5 U) (Merck, Darmstadt, Germany) at 37°C in 100 mM HEPES-KOH (pH 7) with 5 mM $MgCl_2$–5 mM $NADP^+$ in a total volume of 500 μl. The reduction of $NADP^+$ was measured at 340 nm using a Specord UV/visible-light (Vis) spectrometer (Analytic Jena AG, Jena, Germany).

(ii) **TreH.** Glucose formation from trehalose, was determined in McIlvaine buffer (0.2 M Na_2HPO_4, titrated with 0.1 M citric acid to pH 7 at 75°C). The resulting D-glucose was quantified by addition of the reagent DNSA (3,5-dinitrosalicylic acid), incubation for 15 min at 100°C, and absorbance measurements at 575 nm with D-glucose as the positive control (64).

TABLE 1 Primers, plasmids, and strains used in this work

Primer, plasmid, or strain	Sequence (5′–3′) or description[a]	Source or reference
Primers		
*saci*_1249 *tps* fwd *Nde*I	CCAGGCATATGTTCTCAGTATCTATCAGTA	
*saci*_1249 *tps* rvs XhoI	CGCGCTCGAGCTAAAGGTTAACATTTTTA	
*saci*_0016 *tpp* fwd XhoI	GACGCTCGAGATGAATAAGGACATA	
*saci*_0016 *tpp* rvs BamHI	GCGCGGATCCTTATTTTATAAGAAT	
*saci*_1249 (Δ*tps*) fwd upstream NdeI	GTGAGCATATGCTGTTCTCGTTTACTG	
*saci*_1249 (Δ*tps*) rvs downstream NcoI	GCATCCATGGTGTAATGCGGTAAG	
*saci*_1249 (Δ*tps*) fwd overlapping region	CAATTAGACCTACTAAAGGTTTGAGAACATAAACTAATTTACTCTTTCAG	
*saci*_1249 (Δ*tps*) rvs overlapping region	GTAAATTAGTTTATGTTCTCAAACCTTTAGTAGGTCTAATTGTTATAG	
*saci*_1436 (Δ*treY*) fwd upstream BamHI	GAGGGATCCAGGCTAATAAACTGAACAATG	
*saci*_1436 (Δ*treY*) rvs downstream NcoI	AGCCATGGACTTGCGGAGTTAATAAATG	
*saci*_1436 (Δ*treY*) fwd overlapping region	GGTGGTTGTTGAAATAACGAGGATAGAATTGGTAGC	
*saci*_1436 (Δ*treY*) rvs overlapping region	TTACATTCTAACTAGGGTTGCTGATATCACTGGAACTCTATCC	
*saci*_1827 (Δ*treT*) fwd upstream ApaI	CCCACTGGGCCCCCTCATACCATAAGTGCTAATGGAAC	
*saci*_1827 (Δ*treT*) rvs downstream PstI	CGCCGACTGCAGGCATTAATAGTTTCTGTAGTAGCCTCTGC	
*saci*_1827 (Δ*treT*) fwd overlapping region	CTACCAATTCTATCCTCGTTACTCATATTTCTCTATCATTTCAACAACCACCTCTC	
*saci*_1827 (Δ*treT*) rvs overlapping region	ATGATAGAGAAATATGAGTAACGAGGATAGAATTGGTAGCAGAACTAAGGGTTGG	
*saci*_1827 (*treT*) fwd NdeI	GCACTGGCATATGATAGAGAAATATGAGAAATTTATTG	
*saci*_1827 (*treT*) rvs XhoI	GTCAGACTCGAGTTATACACTATTCCTCTC	
*saci*_1436 (*treY*) fwd NdeI	GTAGAGTTCATATGATATCAGCAACCTACAG	
*saci*_1436 (*treY*) rvs BamHI	GTCACATCGGATCCCTCTATTTTCATATTCTTATTTG	
*saci*_1249 (*tps*) fwd NdeI	CGGCTACATATGTTCTCAGTATCTATCAGT	
*saci*_1249 (*tps*) rvs XhoI	GCTGACTCGAGAAGGTTAACATTTTTACCATAGT	
qRT-PCR primers		
*saci*_0574 (*secY*) fwd	CCTGCAACATCTATCCATAACATACCGA	
*saci*_0574 (*secY*) rvs	CCTCATAGTGTATATGCTTTAGTAGTAG	
*saci*_1827 (*treT*) fwd	GGATATCGGAAGATAAACCGTTAGTAACCC	
*saci*_1827 (*treT*) rvs	AGATCTACGTGCCTCTTAGCTAACTTG	
*saci*_1436 (*treY*) fwd	CAATATGATCAATTCTATAGCCATCAACATC	
*saci*_1436 (*treY*) rvs	ATCCTCCTAGCTATAGACGATTCTTCG	
*saci*_1249 (*tps*) fwd	GCCTCTGATCACCATATTTAACCTGAGG	
*saci*_1249 (*tps*) rvs	CTGTAGGCGGAGTACCAAAGATGATG	
Plasmids		
pET15b	*E. coli* expression plasmid carrying an N-terminal His tag	Novagen, USA
pSVA407	Gene targeting plasmid, pGEM-T Easy backbone, *pyrEF* cassette of *S. solfataricus*	(1)
pSVA406	Gene targeting plasmid, pGEM-T Easy backbone, *pyrEF* cassette of *S. solfataricus*	(1)
pSVAmZ-SH10	Expression plasmid with minimal replicon of pRN1 consisting of the region surrounding orf56 and orf904	Sonja V. Albers
pBS-0539	*E. coli* expression plasmid of *tps* (*saci_1249*) cloned into pET15b	This work
pBS-0586	*E. coli* expression plasmid of *tpp* (*saci_0016*) cloned into pET15b	This work
pSVA1332	In-frame deletion of *treT* (*saci_1827*) cloned into pSVA407 with ApaI and PstI	This work
pBS-0285	In-frame deletion of *treY* (*saci_1436*) cloned into pSVA406 with BamHI and NcoI	This work
pBS-0584	In-frame deletion of *tps* (*saci_1249*) cloned into pSVA407 with NdeI and NcoI	This work
pBS-0627	Expression plasmid of *treT* (*saci_1827*) cloned into pSVAmZ-SH10 with NdeI and XhoI	This work
pBS-0806	Expression plasmid of *treY* (*saci_1436*) cloned into pSVAmZ-SH10 with NdeI and BamHI	This work
pBS-0626	Expression plasmid of *tps* (*saci_1249*) cloned into pSVAmZ-SH10 with NdeI and XhoI	This work
Strains		
E. coli DH5α		Hanahan, USA
E. coli BL21(DE3)		Stratagene, USA
E. coli BL21(DE3) CodonPlus pRil		Stratagene, USA
E. coli ER1821		New England Biolabs, Germany
S. acidocaldarius MW001		(62)
S. acidocaldarius MW001 Δ*treY*		This work
S. acidocaldarius MW001 Δ*treT*		This work
S. acidocaldarius MW001 Δ*tps*		This work
S. acidocaldarius MW001 Δ*treY* Δ*treT*		This work
S. acidocaldarius MW001 Δ*treY* Δ*tps*		This work
S. acidocaldarius MW001 Δ*treT* Δ*tps*		This work
S. acidocaldarius MW001 Δ*treY* Δ*treT* Δ*tps*		This work

(iii) **TPS.** T6P formation was determined continuously at 340 nm as UDP or ADP release from UDPG or ADPG (10 mM) and glucose 6-phosphate (G6P, 10 mM) at 55°C (500-µl total volume) by coupling the reaction to the NADH oxidation via PK (8 units) and LDH (4 units) in 50 mM HEPES-KOH (pH 7) containing 10 mM $MgCl_2$, 2 mM phosphoenolpyruvate, and 0.5 mM NADH.

(iv) **TPP.** T6P phosphatase activity was determined continuously in 50 mM HEPES/KOH pH 7 with 10 mM $MgCl_2$ by following the *para*-nitrophenol release from *para*-nitrophenyl phosphate at 340 nm at 55°C using 4 µg of purified protein.

For preparation of cell extracts, *S. acidocaldarius* cells were grown under standard conditions or in the presence of 250 mM NaCl until mid-exponential phase (OD_{600} = 0.7 to 0.9). Cells were harvested (8.000 × *g*, 10 min, 20°C), resuspended in buffer, and disrupted (Precellys homogenizer; 3 times for 15 s at 6,500 rpm). After centrifugation, the cell extracts were dialyzed overnight at 8°C (cutoff, 10 kDa). Protein concentrations were determined with a modified Bradford method using bovine serum albumin as a standard (Bio-Rad, Feldkirchen, Germany). All three pathway activities were assayed in 50 mM MES (morpholineethanesulfonic acid)-KOH (pH 6.5), 10 mM $MgCl_2$ at 70°C by incubating the respective substrates in the presence of crude extract (100 µg). Samples were taken over time, and trehalose formation was determined with a trehalose assay kit (Megazyme, Bray, Ireland).

Trehalose determination. Trehalose concentrations (intracellular) were measured using a trehalose assay kit (Megazyme, Bray, Ireland), according to the manufacturer's instructions. From 400-ml cultures, 50-ml samples were taken over time; cells were harvested (8,000 × *g*, 10 min, 20°C) and disrupted using a Precellys homogenizer (3 times for 15 s each, 6,500 rpm). After acetone precipitation (80% [vol/vol], 2 h at 4°C), the trehalose concentrations of the samples were determined.

RNA isolation and RT-qPCR. Total RNA was isolated from growing cells (5-ml samples, under salt stress and standard conditions) using the FastGene RNA Premium kit (Nippon Genetics, Düren, Germany) according to the manufacturer's instructions with a prolonged DNase I incubation time (30 min). Salt-stressed cells were washed in Brock basal medium before RNA isolation. Absence of DNA in the RNA samples was confirmed by PCR using *secY* (*saci_0574*, GenBank no. AAY79966) primers (primers used in qRT-PCR experiments are listed in Table 1). The quality of RNA samples was analyzed via 1% (wt/vol) agarose gels. First-strand cDNA templates were synthesized using the FastGene Scriptase II synthesis kit (Nippon Genetics, Düren, Germany). The synthesis was performed with random hexamer primers and 1 µg of RNA as the template. Quantitative PCR (qPCR) was performed using qPCRBIO SyGreen Mix Lo-ROX (PCR Biosystems, London, UK) and a MIC (magnetic induction cycler) PCR machine (Bio Molecular Systems, Coomera, Australia). C_q values (quantification cycles) were automatically determined after 50 cycles using *secY* (*saci_0574*) as the reference gene (65). qPCRs with DNA-free RNA samples were performed as controls. At least three biological replicates and two technical replicates were performed.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, PDF file, 1 MB.

ACKNOWLEDGMENTS

C.S. acknowledges funding by an EVONIK Industries Ph.D. scholarship. C.S., B.H.M., C.B., and B.S. acknowledge funding by the German Federal Ministry of Education and Research (BMBF) grant HotSolute, 031B0612A, within the ERA CoBioTech funding initiative. HotSolute has received funding from the European Union's Horizon 2020 research and innovations program under grant agreement no. 722361. J.C. acknowl- edges funding by the Ministry of Education (2019R1I1A2A01062787) through the Basic Science Research Program by the National Research Foundation of Korea (NRF).

We thank Katharina Dietzel and Marten Dahlhaus for extremely valuable experimen- tal support.

Conceptualization, B.S., C.B., and J.C.; Investigation, C.S., B.H.M., A.H., E.J., and A.L.; Writing - Original Draft, C.S. and B.H.M.; Writing - Review & Editing, C.S., B.H.M., C.B., S.-V.A., J.C., and B.S.; Funding Acquisition, B.S. and J.C.; Supervision, C.B., J.C., S.-V.A., and B.S.

We declare no competing interests.

REFERENCES

1. Empadinhas N, da Costa MS. 2008. Osmoadaptation mechanisms inprokaryotes: distribution of compatible solutes. Int Microbiol 11: 151–161.
2. Elbein AD, Pan YT, Pastuszak I, Carroll D. 2003. New insights on trehalose: a multifunctional molecule. Glycobiology 13:17R–27R. https://doi.org/10.1093/glycob/cwg047.
3. Arguelles JC. 2000. Physiological roles of trehalose in bacteria and yeasts: a comparative analysis. Arch Microbiol 174:217–224. https://doi.org/10.1007/s002030000192.
4. Thammahong A, Puttikamonkul S, Perfect JR, Brennan RG, Cramer RA.2017. Central role of the trehalose biosynthesis pathway in the pathogenesis of human fungal infections: opportunities and challenges for therapeutic development. Microbiol Mol Biol Rev 81:e00053-16. https://doi.org/10.1128/MMBR.00053-16.
5. Lunn JE, Delorge I, Figueroa CM, Van Dijck P, Stitt M. 2014. Trehalosemetabolism in plants. Plant J 79:544–567. https://doi.org/10.1111/tpj .12509.
6. Iturriaga G, Suárez R, Nova-Franco B. 2009. Trehalose metabolism: fromosmoprotection to signaling. Int J Mol Sci 10:3793–3810. https://doi.org/ 10.3390/ijms10093793.
7. Zeidler S, Hubloher J, Schabacker K, Lamosa P, Santos H, Müller V. 2017. Trehalose, a temperature- and salt-induced solute with implications in pathobiology of *Acinetobacter baumannii*. Environ Microbiol 19: 5088–5099. https://doi.org/10.1111/1462-2920.13987.
8. Eleutherio E, Panek A, De Mesquita JF, Trevisol E, Magalhães R. 2015. Revisiting yeast trehalose metabolism. Curr Genet 61:263–274. https://doi.org/10.1007/s00294-014-0450-1.
9. Tournu H, Fiori A, Van Dijck P. 2013. Relevance of trehalose inpathogenicity: some general rules, yet many exceptions. PLoS Pathog 9:e1003447. https://doi.org/10.1371/journal.ppat.1003447.
10. Arguelles JC. 2014. Why can't vertebrates synthesize trehalose? J MolEvol 79:111–116. https://doi.org/10.1007/s00239-014-9645-9.
11. Mihaela I, Ryozo I. 2011. Trehalose and abiotic stress in biologicalsystems, p 215–234. *In* Shanker A, Venkateswarlu B (ed), Abiotic stress in plants. Mechanisms and adaptations. InTechOpen, Rijeka, Croatia.
12. Paul MJ, Primavesi LF, Jhurreea D, Zhang Y. 2008. Trehalose metabolismand signaling. Annu Rev Plant Biol 59:417–441. https://doi.org/10.1146/ annurev.arplant.59.032607.092945.
13. MacIntyre AM, Barth JX, Pellitteri Hahn MC, Scarlett CO, Genin S, Allen C.2020. Trehalose synthesis contributes to osmotic stress tolerance and virulence of the bacterial wilt pathogen *Ralstonia solanacearum*. Mol Plant Microbe Interact 33:462–473. https://doi.org/10.1094/MPMI-08-19 - 0218-R.
14. Avonce N, Mendoza-Vargas A, Morett E, Iturriaga G. 2006. Insights on theevolution of trehalose biosynthesis. BMC Evol Biol 6:109. https://doi.org/ 10.1186/1471-2148-6-109.
15. Makihara F, Tsuzuki M, Sato K, Masuda S, Nagashima KVP, Abo M, OkuboA. 2005. Role of trehalose synthesis pathways in salt tolerance mechanism of *Rhodobacter sphaeroides* f. sp. *denitrificans* IL106. Arch Microbiol 184:56–65. https://doi.org/10.1007/s00203-005-0012-5.
16. Silva Z, Alarico S, da Costa MS. 2005. Trehalose biosynthesis in *Thermus thermophilus* RQ-1: biochemical properties of the trehalose-6-phosphate synthase and trehalose-6-phosphate phosphatase. Extremophiles 9:29–36. https://doi.org/10.1007/s00792-004-0421-4.
17. Silva Z, Alarico S, Nobre A, Horlacher R, Marugg J, Boos W, Mingote AI,da Costa MS. 2003. Osmotic adaptation of *Thermus thermophilus* RQ-1: lesson from a mutant deficient in synthesis of trehalose. J Bacteriol 185:5943–5952. https://doi.org/10.1128/jb.185.20.5943-5952.2003.
18. Woodruff PJ, Carlson BL, Siridechadilok B, Pratt MR, Senaratne RH,Mougous JD, Riley LW, Williams SJ, Bertozzi CR. 2004. Trehalose is required for growth of *Mycobacterium smegmatis*. J Biol Chem 279: 28835–28843. https://doi.org/10.1074/jbc.M313103200.
19. Strom AR, Kaasen I. 1993. Trehalose metabolism in *Escherichia coli*: stress protection and stress regulation of gene expression. Mol Microbiol 8:205–210. https://doi.org/10.1111/j.1365-2958.1993.tb01564.x.
20. Sugawara M, Cytryn EJ, Sadowsky MJ. 2010. Functional role of *Bradyrhizobium japonicum* trehalose biosynthesis and metabolism genes during physiological stress and nodulation. Appl Environ Microbiol 76: 1071–1081. https://doi.org/10.1128/AEM.02483-09.
21. Reina-Bueno M, Argandoña M, Nieto JJ, Hidalgo-García A, Iglesias-GuerraF, Delgado MJ, Vargas C. 2012. Role of trehalose in heat and desiccation tolerance in the soil bacterium *Rhizobium etli*. BMC Microbiol 12:207. https://doi.org/10.1186/1471-2180-12-207.
22. McIntyre HJ, Davies H, Hore TA, Miller SH, Dufour J-P, Ronson CW. 2007. Trehalose biosynthesis in *Rhizobium leguminosarum bv. trifolii* and its role in desiccation tolerance. Appl Environ Microbiol 73:3984–3992. https://doi.org/10.1128/AEM.00412-07.
23. Domínguez-Ferreras A, Soto MJ, Pérez-Arnedo R, Olivares J, Sanjuán J.2009. Importance of trehalose biosynthesis for *Sinorhizobium meliloti* osmotolerance and nodulation of alfalfa roots. J Bacteriol 191: 7490–7499. https://doi.org/10.1128/JB.00725-09.
24. Tzvetkov M, Klopprogge C, Zelder O, Liebl W. 2003. Genetic dissection oftrehalose biosynthesis in *Corynebacterium glutamicum*: inactivation of trehalose production leads to impaired growth and an altered cell wall lipid composition. Microbiology (Reading) 149:1659–1673. https://doi.org/10.1099/mic.0.26205-0.
25. Martins LO, Huber R, Huber H, Stetter KO, Da Costa MS, Santos H. 1997. Organic solutes in hyperthermophilic archaea. Appl Environ Microbiol 63:896–902. https://doi.org/10.1128/AEM.63.3.896-902.1997.
26. Nicolaus B, Gambacorta A, Basso AL, Riccio R, De Rosa M, Grant WD.1988. Trehalose in Archaebacteria. Syst Appl Microbiol 10:215–217. https://doi.org/10.1016/S0723-2020(88)80003-1.
27. Youssef NH, Savage-Ashlock KN, McCully AL, Luedtke B, Shaw EI, HoffWD, Elshahed MS. 2014. Trehalose/2-sulfotrehalose biosynthesis and glycine-betaine uptake are widely spread mechanisms for osmoadaptation in the *Halobacteriales*. ISME J 8:636–649. https://doi.org/10.1038/ismej.2013.165.
28. Kouril T, Zaparty M, Marrero J, Brinkmann H, Siebers B. 2008. A noveltrehalose synthesizing pathway in the hyperthermophilic Crenarchaeon *Thermoproteus tenax*: the unidirectional TreT pathway. Arch Microbiol 190:355–369. https://doi.org/10.1007/s00203-008-0377-3.
29. Gao Y, Jiang Y, Liu Q, Wang R, Liu X, Liu B. 2014. Enzymatic andregulatory properties of the trehalose-6-phosphate synthase from the thermoacidophilic archaeon *Thermoplasma acidophilum*. Biochimie 101: 215–220. https://doi.org/10.1016/j.biochi.2014.01.018.
30. Zaparty M, Hagemann A, Bräsen C, Hensel R, Lupas AN, Brinkmann H,Siebers B. 2013. The first prokaryotic trehalose synthase complex identified in the hyperthermophilic crenarchaeon *Thermoproteus tenax*. PLoS One 8:e61354. https://doi.org/10.1371/journal.pone.0061354.
31. Qu Q, Lee SJ, Boos W. 2004. TreT, a novel trehalose glycosyltransferringsynthase of the hyperthermophilic archaeon *Thermococcus litoralis*. J Biol Chem 279:47890–47897. https://doi.org/10.1074/jbc.M404955200.
32. Zhao M, Xu X, Yang S, Liu T, Liu B. 2018. Cloning, expression andcharacterization of the maltooligosyl trehalose synthase from the archaeon *Sulfolobus tokodaii*. Pak J Pharm Sci 31:599–601.
33. Maruta K, Mitsuzumi H, Nakada T, Kubota M, Chaen H, Fukuda S,Sugimoto T, Kurimoto M. 1996. Cloning and sequencing of a cluster of genes encoding novel enzymes of trehalose biosynthesis from thermophilic archaebacterium *Sulfolobus acidocaldarius*. Biochim Biophys Acta 1291:177–181. https://doi.org/10.1016/S0304-4165(96)00082-7.
34. Nakada T, Ikegami S, Chaen H, Kubota M, Fukuda S, Sugimoto T, Kurimoto M, Tsujisaka Y. 1996. Purification and characterization of thermostable maltooligosyl trehalose trehalohydrolase from the thermoacidophilic archaebacterium *Sulfolobus acidocaldarius*. Biosci Biotechnol Biochem 60:267–270. https://doi.org/10.1271/bbb.60.267.
35. Nakada T, Ikegami S, Chaen H, Kubota M, Fukuda S, Sugimoto T, Kurimoto M, Tsujisaka Y. 1996. Purification and characterization of thermostable maltooligosyl trehalose synthase from the thermoacidophilic archaebacterium *Sulfolobus acidocaldarius*. Biosci Biotechnol Biochem 60: 263–266. https://doi.org/10.1271/bbb.60.263.

36. Di Lernia I, Morana A, Ottombrino A, Fusco S, Rossi M, De Rosa M. 1998. Enzymes from *Sulfolobus shibatae* for the production of trehalose and glucose from starch. Extremophiles 2:409–416. https://doi.org/10.1007/s007920050086.
37. Fang TY, Tseng WC, Chung YT, Pan CH. 2006. Mutations on aromaticresidues of the active site to alter selectivity of the *Sulfolobus solfataricus* maltooligosyltrehalose synthase. J Agric Food Chem 54:3585–3590. https://doi.org/10.1021/jf060152z.
38. Fang TY, Tseng WC, Guo MS, Shih TY, Hung XG. 2006. Expression, purification, and characterization of the maltooligosyltrehalose trehalohydrolase from the thermophilic archaeon *Sulfolobus solfataricus* ATCC 35092. J Agric Food Chem 54:7105–7112. https://doi.org/10.1021/jf061318z.
39. Gueguen Y, Rolland JL, Schroeck S, Flament D, Defretin S, Saniez MH,Dietrich J. 2001. Characterization of the maltooligosyl trehalose synthase from the thermophilic archaeon *Sulfolobus acidocaldarius*. FEMS Microbiol Lett 194:201–206. https://doi.org/10.1111/j.1574-6968.2001.tb09470.x.
40. Seo J-S, An JH, Baik M-Y, Park CS, Cheong J-J, Moon TW, Park KH, ChoiYD, Kim CH. 2007. Molecular cloning and characterization of trehalose biosynthesis genes from hyperthermophilic archaebacterium *Metallosphaera hakonesis*. J Microbiol Biotechnol 17:123–129.
41. Seo JS, An JH, Cheong JJ, Choi YD, Kim CH. 2008. Bifunctional recombinant fusion enzyme between maltooligosyltrehalose synthase and maltooligosyltrehalose trehalohydrolase of thermophilic microorganism *Metallosphaera hakonensis*. J Microbiol Biotechnol 18:1544–1549.
42. Chen YS, Lee GC, Shaw JF. 2006. Gene cloning, expression, and biochemical characterization of a recombinant trehalose synthase from *Picrophilus torridus* in *Escherichia coli*. J Agric Food Chem 54:7098–7104. https://doi.org/10.1021/jf060828q.
43. Chou H-H, Chang S-W, Lee G-C, Chen Y-S, Yeh T, Akoh CC, Shaw J-F.2010. Site-directed mutagenesis improves the thermostability of a recombinant *Picrophilus torridus* trehalose synthase and efficiency for the production of trehalose from sweet potato starch. Food Chem 119: 1017–1022. https://doi.org/10.1016/j.foodchem.2009.08.010.
44. Sakaguchi M, Shimodaira S, Ishida S-N, Amemiya M, Honda S, Sugahara Y, Oyama F, Kawakita M. 2015. Identification of GH15 family thermophilic archaeal trehalases that function within a narrow acidic-pH range. Appl Environ Microbiol 81:4920–4931. https://doi .org/10.1128/AEM.00956-15.
45. Yuasa M, Okamura T, Kimura M, Honda S, Shin Y, Kawakita M, Oyama F,Sakaguchi M. 2018. Two trehalose-hydrolyzing enzymes from Crenarchaeon *Sulfolobus acidocaldarius* exhibit distinct activities and affinities toward trehalose. Appl Microbiol Biotechnol 102:4445–4455. https://doi .org/10.1007/s00253-018-8915-7.
46. Lee J, Lee A, Moon K, Choi KH, Cha J. 2018. Saci_1816: a trehalose thatcatalyzes trehalose degradation in the thermoacidophilic crenarchaeon *Sulfolobus acidocaldarius*. J Microbiol Biotechnol 28:909–916. https://doi .org/10.4014/jmb.1802.02038.
47. Moon JH, Lee W, Park J, Choi KH, Cha J. 2016. Characterization of atrehalose-degrading enzyme from the hyperthermophilic archaeon *Sulfolobus acidocaldarius*. J Biosci Bioeng 122:47–51. https://doi.org/10.1016/j.jbiosc.2015.12.011.
48. Sakaguchi M. 2020. Diverse and common features of trehalases and theircontributions to microbial trehalose metabolism. Appl Microbiol Biotechnol 104:1837–1847. https://doi.org/10.1007/s00253-019-10339-7.
49. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic localalignment search tool. J Mol Biol 215:403–410. https://doi.org/10.1016/ S0022-2836(05)80360-2.
50. Coutinho PM, Deleury E, Davies GJ, Henrissat B. 2003. An evolvinghierarchical family classification for glycosyltransferases. J Mol Biol 328: 307–317. https://doi.org/10.1016/S0022-2836(03)00307-3.
51. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. 2015. ThePhyre2 web portal for protein modeling, prediction and analysis. Nat Protoc 10:845–858. https://doi.org/10.1038/nprot.2015.053.
52. Zimmermann L, Stephens A, Nam S-Z, Rau D, Kübler J, Lozajic M, GablerF, Söding J, Lupas AN, Alva V. 2018. A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. J Mol Biol 430:2237–2243. https://doi.org/10.1016/j.jmb.2017.12.007.
53. Lamosa P, Martins LO, Da Costa MS, Santos H. 1998. Effects of temperature, salinity, and medium composition on compatible solute accumulation by *Thermococcus* spp. Appl Environ Microbiol 64:3591–3598. https://doi.org/10.1128/AEM.64.10.3591-3598.1998.
54. Neves C, da Costa MS, Santos H. 2005. Compatible solutes of thehyperthermophile *Palaeococcus ferrophilus:* osmoadaptation and thermoadaptation in the order *Thermococcales*. Appl Environ Microbiol 71:
8091–8098. https://doi.org/10.1128/AEM.71.12.8091-8098.2005.
55. Martins LO, Santos H. 1995. Accumulation of mannosylglycerate anddi-myo-inositol-phosphate by *Pyrococcus furiosus* in response to salinity and temperature. Appl Environ Microbiol 61:3299–3303. https://doi.org/10.1128/AEM.61.9.3299-3303.1995.
56. Cheng Q, Gao H, Hu N. 2016. A trehalase from *Zunongwangia* sp.: characterization and improving catalytic efficiency by directed evolution. BMC Biotechnol 16:9. https://doi.org/10.1186/s12896-016-0239-z.
57. Bischof LF, Haurat MF, Hoffmann L, Albersmeier A, Wolf J, Neu A, et al.2019. Early response of *Sulfolobus acidocaldarius* to nutrient limitation. Front Microbiol 9:3201. https://doi.org/10.3389/fmicb.2018.03201.
58. Grogan DW. 1989. Phenotypic characterization of the archaebacterialgenus *Sulfolobus*: comparison of five wild-type strains. J Bacteriol 171: 6710–6719. https://doi.org/10.1128/jb.171.12.6710-6719.1989.
59. Empadinhas N, da Costa MS. 2006. Diversity and biosynthesis of compatible solutes in hyper/thermophiles. Int Microbiol 9:199–206.
60. Wolf A, Krämer R, Morbach S. 2003. Three pathways for trehalosemetabolism in *Corynebacterium glutamicum* ATCC13032 and their significance in response to osmotic stress. Mol Microbiol 49:1119–1134. https://doi.org/10.1046/j.1365-2958.2003.03625.x.
61. König H, Skorko R, Zillig W, Reiter W-D. 1982. Glycogen in thermoacidophilic archaebacteria of the genera *Sulfolobus*. Arch Microbiol 132: 297–303. https://doi.org/10.1007/BF00413378.
62. Wagner M, van Wolferen M, Wagner A, Lassak K, Meyer BH, Reimann J,Albers S-V. 2012. Versatile genetic tool box for the crenarchaeote *Sulfolobus acidocaldarius*. Front Microbiol 3:214. https://doi.org/10.3389/fmicb.2012.00214.
63. Brock T, Brock K, Belly R, Weiss R. 1972. *Sulfolobus*: a new genus of sulfur-oxidizing bacteria living at low pH and high temperature. Arch Mikrobiol 84:54–68. https://doi.org/10.1007/BF00408082.
64. Miller GL. 1959. Use of dinitrosalicylic acid reagent for determinationof reducing sugar. Anal Chem 31:426–428. https://doi.org/10.1021/ac60147a030.
65. van der Sluis EO, Nouwen N, Koch J, de Keyzer J, van der Does C, TampéR, Driessen AJM. 2006. Identification of two interaction sites in SecY that are important for the functional interaction with SecA. J Mol Biol 361: 839–849. https://doi.org/10.1016/j.jmb.2006.07.017.
66. Bräsen C, Esser D, Rauch B, Siebers B. 2014. Carbohydrate metabolism inArchaea: current insights into unusual enzymes and pathways and their regulation. Microbiol Mol Biol Rev 78:89–175. https://doi.org/10.1128/MMBR.00041-13.
67. Nishimasu H, Fushinobu S, Shoun H, Wakagi T. 2006. Identificationand characterization of an ATP-dependent hexokinase with broad substrate specificity from the hyperthermophilic archaeon *Sulfolobus tokodaii*. J Bacteriol 188:2014–2019. https://doi.org/10.1128/JB.188.5 .2014-2019.2006.

Supplementary information on

Salt Stress Response of *Sulfolobus acidocaldarius* Involves Complex Trehalose Metabolism Utilizing a Novel Trehalose-6-Phosphate Synthase (TPS)/Trehalose-6-Phosphate Phosphatase (TPP) Pathway

Authors

Christina Stracke[a], Benjamin H. Meyer[a,b], Anna Hagemann[a,c], Eunhye Jo[d], Areum Lee[d], Sonja-Verena Albers[b], Jaeho Cha[d], Christopher Bräsen[a], Bettina Siebers[a,#]

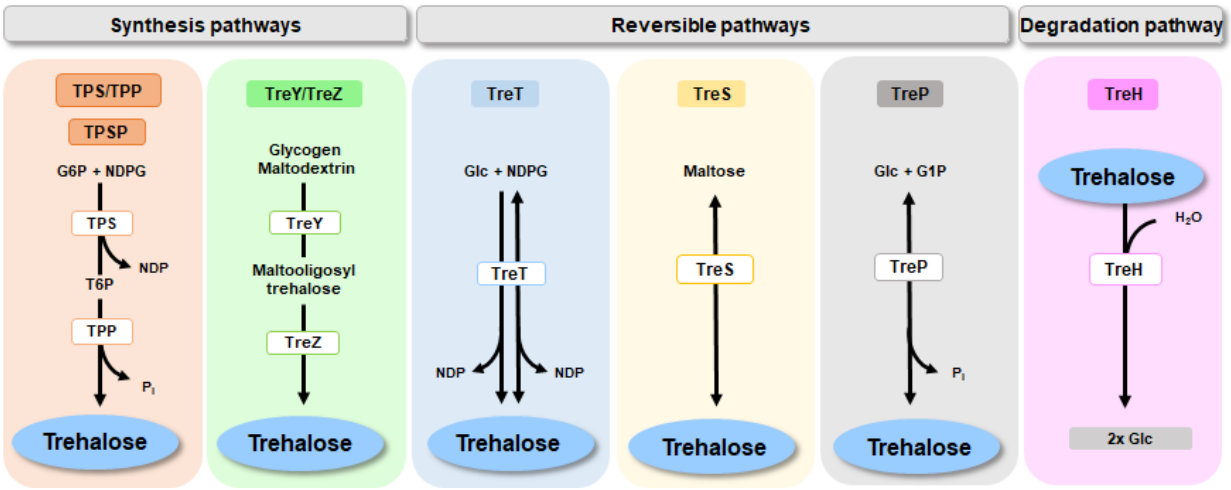
[a]Molecular Enzyme Technology and Biochemistry (MEB), Environmental Microbiology and Biotechnology (EMB), Centre for Water and Environmental Research (CWE), Department of Chemistry, University of Duisburg-Essen, Universitätsstrasse 5, 45141 Essen, Germany

[b]Molecular Biology of Archaea, Institute of Biology II, Albert-Ludwig's University of Freiburg, Schänzlestr. 1, 79104 Freiburg, Germany

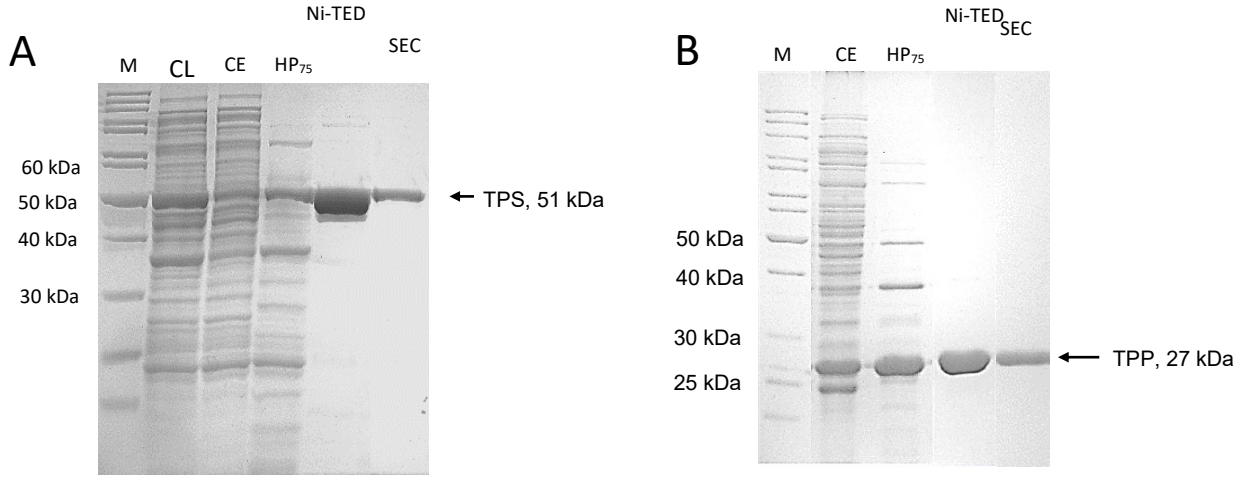
[c]Institute of Pharmacology and Toxicology, Witten/Herdecke University, Department of Health, School of Medicine, Alfred-Herrhausen-Straße 50, 58448 Witten, Germany

[d]Department of Microbiology, Pusan National University, Busan 46241, Republic of Korea

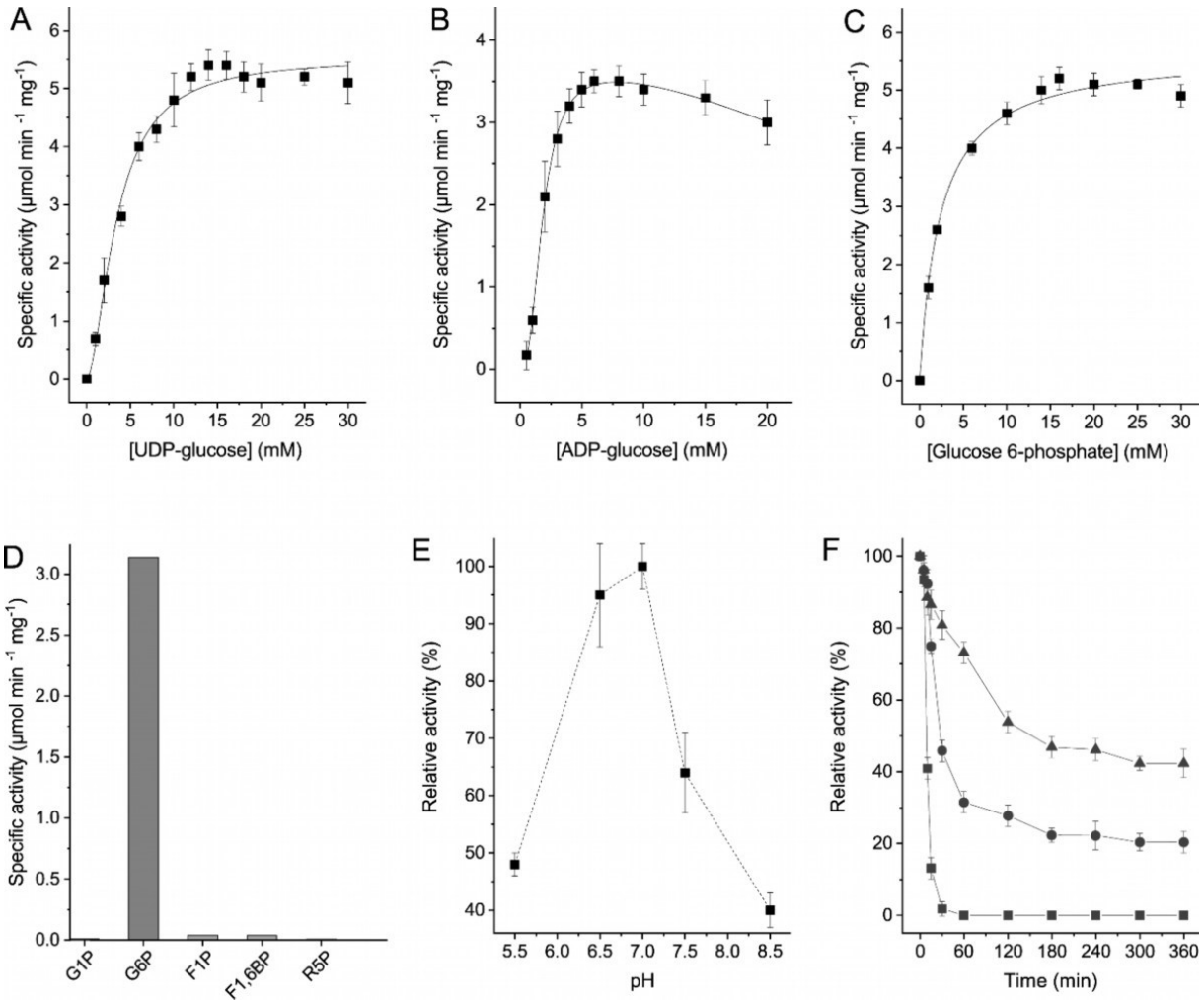
[#]Address correspondence to: Bettina Siebers, bettina.siebers@uni-due.de



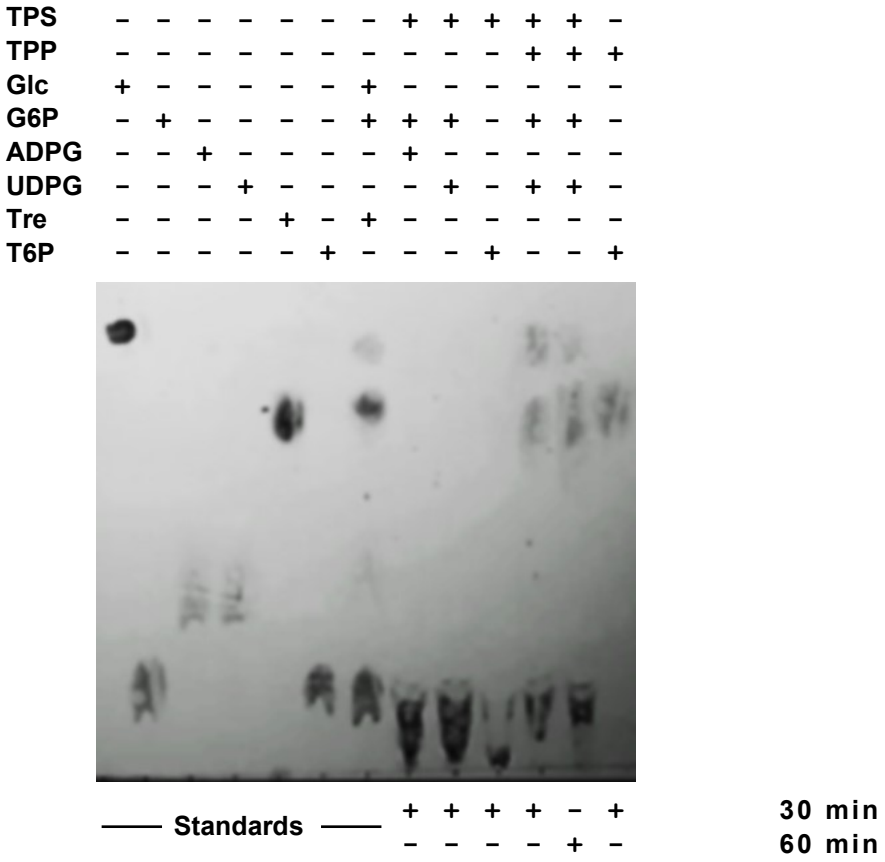
Supplementary figure 1: Trehalose metabolizing pathways including enzymes, substrates and intermediates. Abbreviations; TPS, trehalose-6-phosphate synthase; TPP, trehalose-6-phosphate phosphatase; TPSP, bifunctional fusion protein with trehalose-6-phosphate synthase/trehalose-6-phosphate phosphatase domain; TreY, maltooligosyltrehalose synthase; TreZ, maltooligosyltrehalose trehalohydrolase; TreT, trehalose glycosyltransferring synthase; TreS, trehalose synthase; TreP, trehalose phosphorylase; TreH, trehalase; Glc, glucose; G6P, glucose 6-phosphate; G1P, glucose 1-phosphate, NDP, nucleoside diphosphate; NDPG , nucleoside diphosphate glucose, T6P, trehalose 6-phosphate.



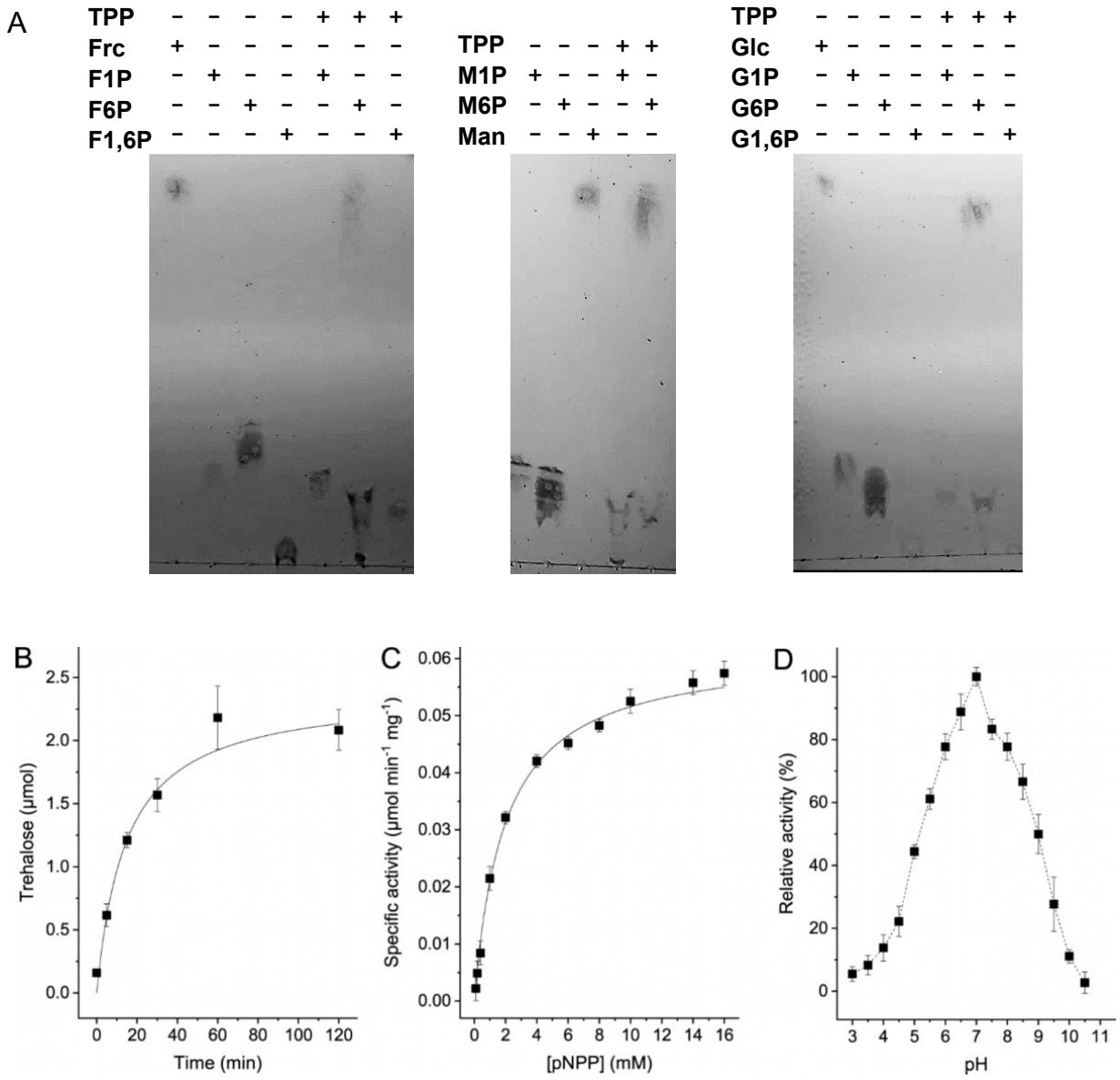
Supplementary figure 2: Purification of the recombinant TPS (A) and TPP (B) from *S. acidocaldarius*. The *tps* gene *saci_1249* and *tpp* gene *saci_0016* were cloned into pET15b and expressed in *E. coli* BL21(DE3)-Codon-Plus with N-terminal His-tag. SDS-PAGE (12.5%, Coomassie staining) of protein fractions (3-10 µg protein) of the respective purification steps are shown. Abbreviations: CL, crude lysate; CE, crude extract; HP_{75}, soluble fraction after heat precipitation at 75°C for 30 min; Ni-TED, protein fraction after Ni-TED affinity chromatography; SEC, protein fraction after size exclusion chromatography; M, protein marker (unstained protein ladder, Thermo Fisher Scientific, Schwerte, Germany).



Supplementary figure 3: Characterization of the recombinant TPS from *S. acidocaldarius* in respect to kinetic properties (A-C), substrate specificity (D), pH dependence (E) and thermostability (F). The enzymatic activity **(A-C)** was determined at 55°C (340 nm) using 5 µg of purified enzyme by coupling the formation of UPD or ADP from UDPG or ADPG to the oxidation of NADH with pyruvate kinase (PK) and (lactic acid dehydrogenase (L-LDH) from rabbit muscle as auxiliary enzymes. The specific activity of purified TPS was determined with G6P concentrations from 0-30 mM and 10 mM UDPG **(A)**, UDPG concentrations from 0 - 30 mM and 10 mM G6P **(B)** or ADPG concentrations from 0-20 mM with 10 mM G6P **(C)**. For determination of the substrate specificity **(D)** UDPG was used as glycosyl donor and 10 mM of each glucose 1-phosphate (G1P), glucose 6-phosphate (G6P), fructose 1-phosphate (F1P), fructose 1,6-bisphosphate (FBP) and ribose 5-phosphate (R5P) were used as acceptor substrate. The pH dependence (**E**) of TPS was analyzed in 50 mM TRIS/HCl, MES/KOH and HEPES/KOH. The activity was determined as described above with G6P and UDPG as substrate. The thermostability (**F**) was analyzed at 90°C (triangles), 80°C (circles) and 70°C (squares) over time. Therefore, purified TPS was incubated in 200 mM potassium phosphate buffer pH 7 at the respective temperatures in a thermocycler. At the time indicated samples were removed and the reaction was stopped by incubation for 10-15 min on ice, followed by determination of the residual activity as described). The maximal activity of TPS 5.1 U/mg was defined as 100% activity. Three independent measurements (n=3) were performed and error bars indicate the standard error of the mean (SEM).

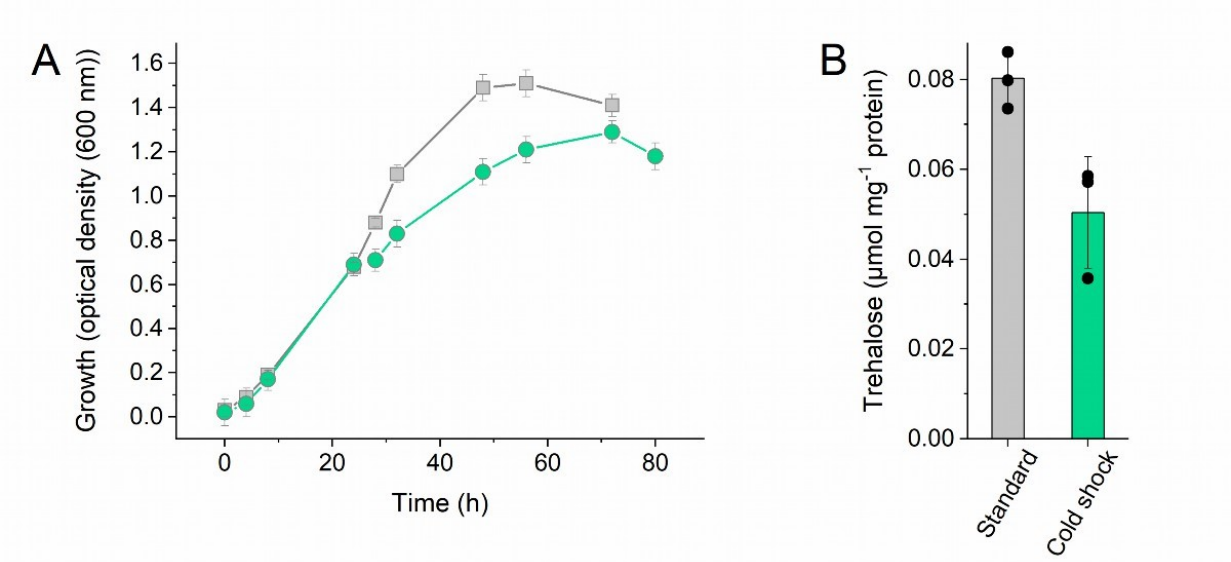


Supplementary figure 4: Analysis of TPS and TPP activity via thin layer chromatography. The TPS and TPP enzyme assays were performed with 10 µg of the purified enzymes in 50 mM HEPES/KOH (pH 7 at 75 °C) with 20 mM $MgCl_2$. The reaction mixtures were supplemented with either 10 mM G6P and 10 mM UDPG or ADPG (TPS) or T6P (TPP reaction) as indicated. After the respective incubation time (30 or 60 min at 55 °C) samples were supplemented with 80% (v/v) acetone and frozen at -20°C. Acetone was removed at 60°C using a Speed Vac concentrator (Eppendorf, Hamburg, Germany). 10 µL of the respective samples as well as standards (3 µl of 10 mM Glc, G6P, ADPG, UDPG, Tre, T6P and a mixture of Glc, G6P and Tre (10 mM each)) were applied and analyzed on Kieselguhr coated TLC aluminium sheets (fluorescent indicator F254, 20 x 20 cm, Merck, Darmstadt, Germany) using 1-butanol, ethanol and bidistilled water (5:3:2) as mobile phase. The samples were separated for 5 hours at room temperature and plates were dried and developed with 20 % (v/v) sulphuric acid followed by incubation at 100 °C for 20 min. Abbreviations; TPS, trehalose-6-phosphate synthase; TPP, trehalose-6-phosphate phosphatase; Glc, glucose; G6P, glucose 6-phosphate; ADPG, adenosine diphosphate glucose; UDPG, uridine diphosphate glucose; Tre, trehalose; T6P, trehalose 6-phosphate.

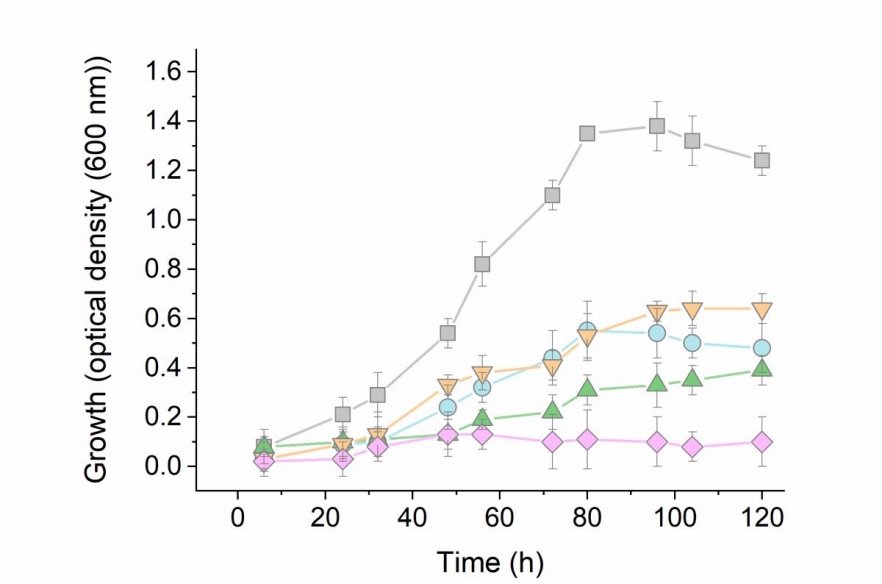


Supplementary figure 5: Characterization of the recombinant TPP from *S. acidocaldarius* in respect to substrate specificity (A), kinetic properties (B, C) and pH dependence (D). TPP substrate specificity for different sugar phosphates was analyzed by TLC **(A).** 3 µl of 10 mM standard solutions of Frc, F1P, F6P, F1,6P, Man, M1P, M6P, Glc, G1P, G6P, G1,6P were applied. TPP reactions were performed with 10 µg of purified enzyme in 50 mM HEPES/KOH (pH 7 at 75 °C) containing 20 mM $MgCl_2$. TLC and plate development was performed as described for supplementary figure 4. TPP activity **(B)** was assayed in a combined assay with the recombinant TPS. Therefore, T6P was first formed by incubation of 10 µg of purified TPS with 10 mM G6P, 10 mM UDPG, 5 mM $MgCl_2$ in 50 mM HEPES/KOH (pH 7.0 at 75°C) for 30 min. Afterwards, 10 µg of purified TPP was added and trehalose formation over time (0 - 120 min) was determined via the Megazyme trehalose assay kit. The initial velocity corresponded to a specific TPP activity of 8.7 U mg^{-1} protein. The general phosphatase activity **(C)** and the pH

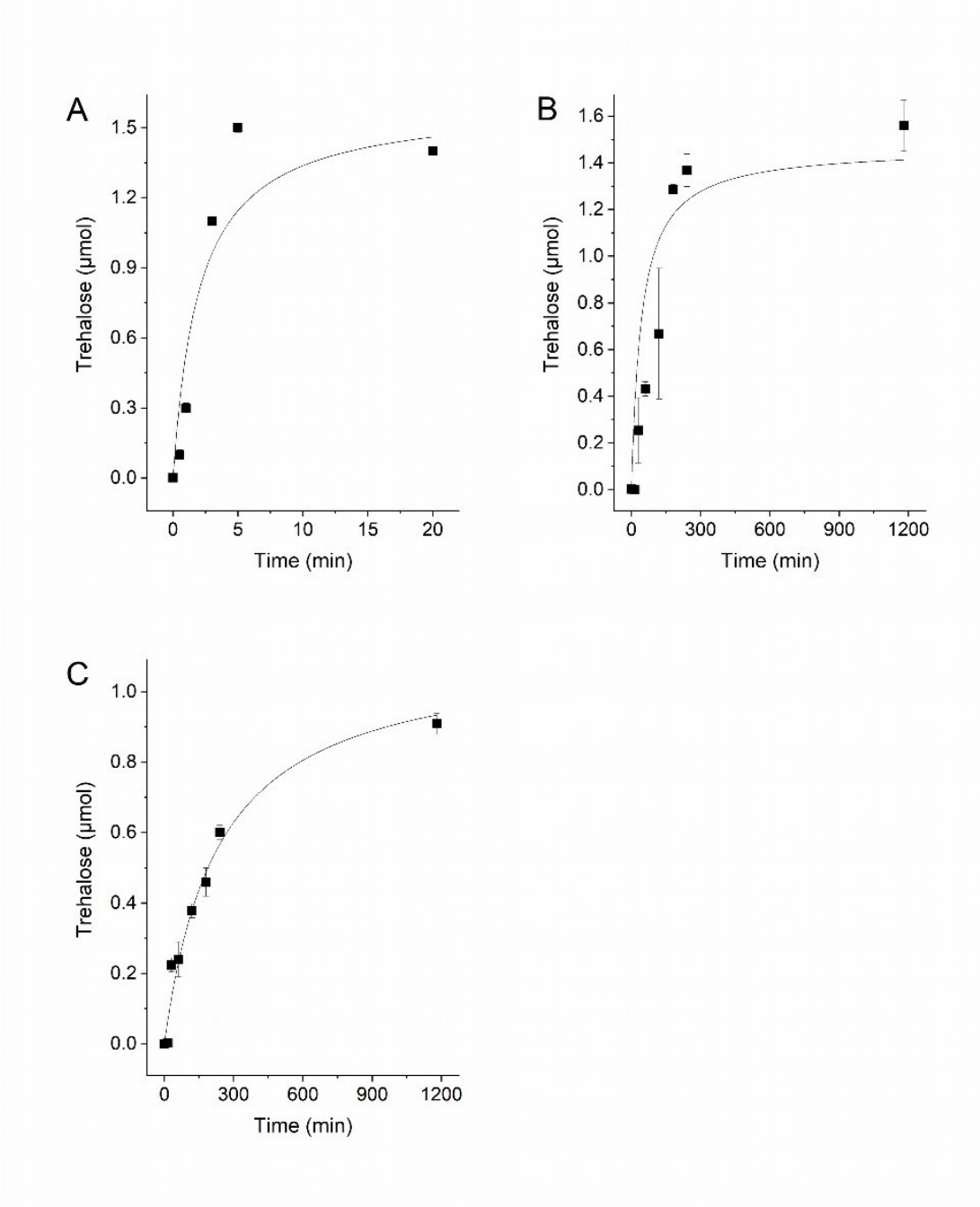
dependence **(D)** was followed using para-nitrophenylphosphate (pNPP) as substrate. TPP activity was continuously determined at 55°C by monitoring the formation of para-nitrophenol (pNP) from pNPP at 405 nm, using 4 µg of the purified enzyme and substrate concentrations from 0 - 16 mM (0.5 mM for the pH dependence) in 50 mM HEPES/KOH pH 7 with 10 mM $MgCl_2$. For pH dependency the buffers as described for supplementary figure 3 was used. Three independent measurements (n=3) were performed and error bars indicate the standard error of the mean (SEM). Abbreviations: Frc, fructose, F1P, fructose 1-phosphate; F6P, fructose 6-phosphate; F1,6P, fructose 1,6-bisphosphate; Man, mannose; M1P, mannose 1-phosphate; M6P, mannose 6-phosphate; G1,6P, glucose 1,6-bisphosphate; For additional abbreviations see supplementary figure 4.



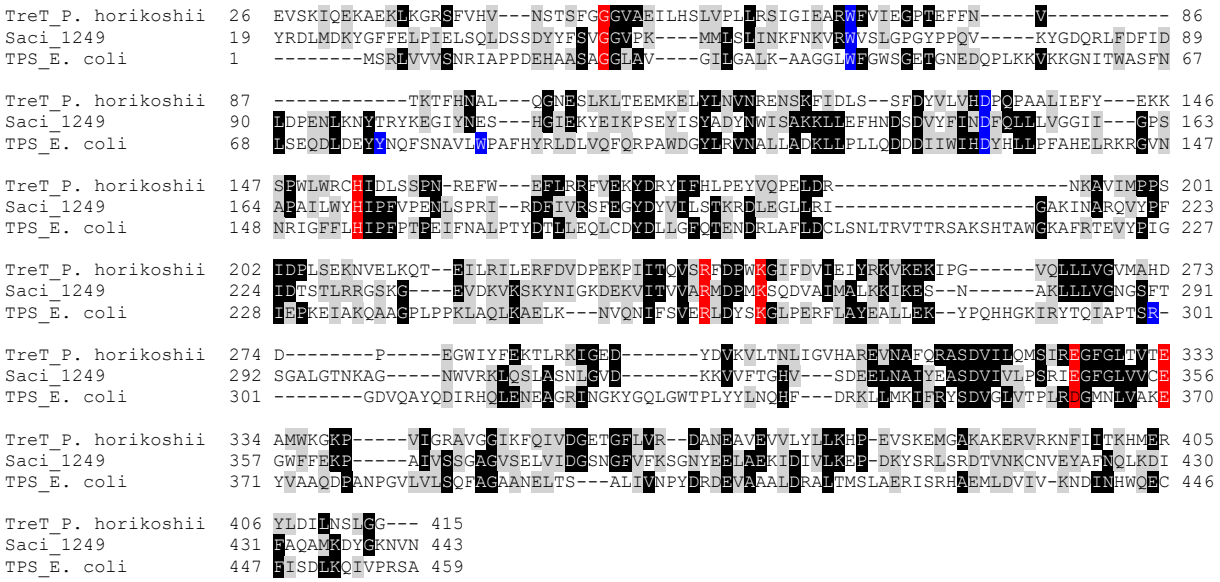
Supplementary figure 6: Growth and intracellular trehalose concentration of *S. acidocaldarius* MW001 under standard growth conditions and under cold shock conditions. Growth curves *S. acidocaldarius* MW001 **(A)** under standard conditions (grey squares) and under cold shock conditions at 65°C (green circles). The cultures were grown under optimal conditions at 75°C until log phase (OD_{600} 0.6), for cold shock conditions the cultures were transferred to 65°C. Intracellular trehalose concentration **(B)** in µmol mg^{-1} protein was determined from cells of late exponential phase (OD 1.0 - 1.1). Three independent measurements (n=3) were performed and error bars indicate the standard error of the mean (SEM



Supplementary figure 7: Growth of the trehalose deficient *S acidocaldarius* MW001 triple mutant strain *ΔtreT/ΔtreY/Δtps* under salt stress conditions complemented with either one of the deleted trehalose synthesis genes. Growth curves of the parental strain MW001 (grey squares) as well as the different (complemented) mutant strains in the presence of 250 mM NaCl is shown. The triple mutant strain *ΔtreT/ΔtreY/Δtps* was transformed with empty plasmid pSVAmz-SH10 as control (pink diamonds), with pSVAmz-SH10_*treT* (blue circles), with pSVAmz-SH10_*treY* (green up triangles), and with pSVAmz-SH10_*tps* (orange down triangles) and was grown on Brock medium supplemented with 250 mM NaCl. Three independent measurements (n=3) were performed and error bars indicate the standard error of the mean (SEM).



Supplementary figure 8: Activity of the different trehalose synthesizing pathways in crude extracts of salt stressed *S. acidocaldarius* MW001 cells. The enzyme assays were performed with 100 µg of crude extract in 50 mM HEPES/KOH (pH 7.0 at 75°C) supplemented with 5 mM $MgCl_2$. The activity of the TreY/TreZ pathway **(A)** was determined) with 10 mM maltopentaose, of the TreT pathway **(B)** with 10 mM glucose and 10 mM UDPG, and of the TPS/TPP pathway **(C)** with 10 mM G6P and 10 mM UDPG. Samples were taken at the time indicated and trehalose formation was determined via the Megazyme trehalose assay kit. Three independent measurements (n=3) were performed and error bars indicate the standard error of the mean (SEM).



Supplementary figure 9: Sequence alignment of GT4-like TPS from *S. acidocaldarius* with the TreT from *P. horikoshii* and the classical GT20 TPS from *Escherichia coli*. The alignment was retrieved from HHPred [2] results after analyses with Saci_1249 as template. The amino acid residues shown to participate in acceptor substrate (i.e. G6P) binding in the *E. coli* TPS are highlighted in red those involved in donor substrate (UDPG) binding in blue [3, 4].

References

1. Wagner M, van Wolferen M, Wagner A, Lassak K, Meyer BH, Reimann J, et al. Versatile genetic tool box for the crenarchaeote *Sulfolobus acidocaldarius*. Front Microbiol. 2012;3:214.

2. Zimmermann L, Stephens A, Nam S-Z, Rau D, Kübler J, Lozajic M, et al. A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. J Mol Biol. 2018;430(15):2237-43.

3. Gibson R, P Turkenburg J, Charnock S, Lloyd R, J Davies G. Insights into trehalose synthesis provided by the structure of the retaining glucosyltransferase OtsA. Chem Biol. 2002;9(12):1337-46.

4. Gibson RP, Tarling CA, Roberts S, Withers SG, Davies GJ. The donor subsite of trehalose-6-phosphate synthase: binary complexes with UDP-glucose and UDP-2-deoxy-2-fluoro-glucose at 2 A resolution. J Biol Chem. 2004;279(3):1950-5.

Chapter 3.2

Enzymatic Synthesis of the Extremolyte Cyclic-2,3-diphosphoglycerate by the Cyclic-2,3-diphosphoglycerate Synthetase from *Methanothermus fervidus*

Enzymatic Synthesis of the Extremolyte Cyclic-2,3-diphosphoglycerate by the Cyclic-2,3-diphosphoglycerate Synthetase from *Methanothermus fervidus*

(preliminary manuscript for submission)

Keywords

Hyperthermophiles – Archaea – Stress response – Compatible solutes – Extremolytes – Thermoadaption – Thermoprotection – 2PGK/cDPGS – 2-phosphoglycerate-kinase – cyclic-2,3-diphosphoglycerate

Abbreviations

(cDPGS) cyclic-2,3-diphosphoglycerate synthetase – (2,3DPG) 2,3-diphosphoglycerate – (cDPG) cyclic 2,3 diphosphoglycerate – (2PG) 2-phosphoglycerate – (2PGK) 2-phosphoglycerate kinase

Abstract

Extremolytes are compatible solutes exclusively formed by extremophiles. They are able to protect biological structures like membranes, proteins, and DNA under adverse conditions including temperature stress and osmotic stress. With these protective properties, extremolytes have an outstanding potential in various fields of application ranging from pharmaceutical, health care and well-being applications up to life science approaches. However, only few of these compounds particularly ectione and hydroxyectoine are established on the market. This is mainly due to missing efficient production strategies for other extremolytes. The extremolyte cDPG found in hyperthermophilic methanogenic Archaea and is synthesized by these organisms from 2PG via two enzymes 2-phosphoglycerate kinase (2PGK) and cyclic-2,3-diphosphoglycerate synthetase (cDPGS).

Here, we present first results towards the establishment of a one-step enzymatic *in vitro* approach to produce cDPG from 2,3 DPG as raw material using the cyclic 2,3-diphosphoglycerate synthetase from *Methanothermus fervidus*. The heterologous production of the cDPGS in *Escherichia coli* was optimized with respect to codon usage and expression conditions as well as an easy two-step purification including heat precipitation and size exclusion chromatography was established. The recombinant protein showed a V_{max} of 21 U mg^{-1} and K_m values of 1.4 mM and 1.1 mM for 2,3 DPG and ATP, respectively. As a prerequisite for application, the recombinant cDPGS could sufficiently be stabilized for use and storage by addition of KCl (400 mM) and reducing agent DTT (10 mM), and additionally for long term storage at -80°C by addition of glycerol (25% v/v). Finally, a complete conversion of 0.76 mg substrate to product was achieved using 7.4 µg of enzyme within 60 min. These findings will now enable the development of the larger scale enzymatic production of cDPG for application.

Introduction

Compatible solutes are ubiquitous and are found in all three domains of life. They accumulate to high intracellular concentrations (up to 2 M) in response to diverse environmental stresses (e.g., heat, cold, osmotic stress, and desiccation), are compatible with cellular metabolism and thus allow their hosts to survive harsh environmental conditions by stabilizing and protecting biomolecular structures as well as the cellular conformation [124, 163]. Compatible solutes can be either taken up from the environment by the cell or are formed via *de-novo* synthesis. The diversity of compatible solutes is large but falls into a few different chemical categories such as (i) amino acids (e.g., α-glutamate, β-glutamate) and amino acid derivatives (e.g., glycine-betaine, ectoine, hydroxyectoine) as well as (ii) sugars (e.g. trehalose, mannosylglycerate), polyols (e.g., glycerol) and derivatives [80]. Their exceptional protective effect on biological structures such as enzymes, DNA, membranes, and whole cells, promotes the use of compatible solutes in many commercial applications in different industrial fields such as food, health and consumer care, and cosmetics. Compatible solutes that are exclusively found and synthesized in extremophiles (e.g., halophiles or (hyper)thermophiles), so-called “extremolytes”, have gained particular industrial interest. In addition to the mesophilic compatible solute trehalose, the commercially most relevant products so far are the extremolytes ectoine and hydroxyectoine. The protective properties of these compounds as well as their synthetic pathways have been well characterized [179, 253, 254]. For instance, ectoine and hydroxyectoine maintain the fluidity of membranes and were demonstrated to protect epithelial cell layers from several allergens, UV light, fine dust particles as well as heat or drought [182, 183]. Also, inflammatory reactions caused by these stressors on skin or mucous membranes can be reduced (Bitop AG: “Extremolytes” https://www.bitop.de/de/extremolytes (Date: 21.01.2020). Thus, ectoine became an important ingredient of many high value-added products in human health care and well-being. The production process of ectoine and hydroxyectoine relies on fermentation of the natural producer strain (non-GMO) of the halophilic bacterium *Halomonas elongata* and involves an approach designated as “bacterial milking” [184]. This is basically an osmotic downshock from 15% to 3% NaCl, which forces the cells to release the solute to the medium via mechano-sensitive channels [185, 186]. While ectoine and hydroxyectoine are particularly found in halophilic or halotolerant bacteria, in hyperthermophiles other extremolytes were mainly described, which often are heterosides such as glucosyl-glycerol (GG, glycoin) [187], glucosyl-glycerate (GGA) [188], mannosylglycerate (MG, Firoin) [189, 190], and the very rare mannosylglyceramide (MGA, Firoin-A) [182, 191]. Also, phosphorylated compounds such as di-*myo*-1,1’-inositol-phosphate (DIP) [192, 193], α-diglycerol phosphate (DGP) [194], and cyclic 2,3-diphosphoglycerate (cDPG), have been described for hyperthermophiles [195-197]. However, their potential for biotechnological and industrial applications remains largely

unexploited, mainly because applicable production hosts and efficient synthesis pathways to make their production economically viable are missing.

The extremolyte cDPG has been exclusively found in the hyperthermophilic methanogens (Archaea) like *Methanothermus fervidus, Methanopyrus kandleri,* and *Methanothermobacter thermoautotrophicus* in concentrations from 0.3 - 1.1 M [196, 200-202]. It appears to play a role in the thermoprotection of proteins and increased thermostabilities on different model enzymes have been reported [201, 203, 204]. Furthermore, it protects plasmid DNA against oxidative damage by hydroxyl radicals and also functions as a superoxide scavenger [182].

The enzymatic synthesis of cDPG was shown to proceed via two catalytic steps (Figure 1). In the first step 2-phosphoglycerate kinase (2PGK) synthesizes the phosphate ester 2,3-diphosphoglycerate (2,3DPG) from 2-phosphoglycerate (2PG). In the second step cyclic-2,3-diphosphoglycerate synthetase (cDPGS) catalyze the formation of an intramolecular phosphoanhydride bond resulting in cDPG. Although cDPG with its known protective properties has a high potential for industrial applications and the synthesis pathways are well investigated, so far, no effective production procedure has been developed. This is due to the low cell yield and challenging cultivation conditions for methanogens under strict anaerobiosis and very high temperature hampering a whole-cell *in vivo* approach like for ectoine. Furthermore, the enzymes characterized from the hyperthermophilic methanogens appeared unstable although no detailed information on the enzyme stability has been reported. Also, the conversion efficiency of the enzymes in terms of substrate to product ratios has not been analyzed so far.

Here, we present a first results towards the establishment of an enzymatic *in vitro* approach to produce cDPG. The heterologous enzyme production was optimized with respect to codon optimization, expression conditions, and purification procedures. Furthermore. the recombinant protein could sufficiently be stabilized for use and storage and finally, the complete substrate to product conversion was achieved. These findings will now enable the development of the larger scale enzymatic production of cDPG for application.

Material and Methods

CLONING, HETEROLOGOUS EXPRESSION, AND PURIFICATION OF THE RECOMBINANT PROTEINS

Cloning. The cyclic 2,3-diphosphoglycerate synthetase gene (*cdpgs,* 1383 bp) from *M. fervidus* (Mfer_0077, CAA70986) was amplified via PCR from chromosomal DNA of *M. fervidus* (primer sets for cloning listed in supplementary table 1). The PCR products were purified using the Wizard® SV Gel and PCR Clean-Up System (Promega, Fitchburg, USA) according to the instructions of the manufacturer. Restricted DNA fragments were cloned into the expression vectors pET24a (Kan^R) and pET15b (Amp^R) (Novagen, Merck, Darmstadt, Germany) with C- or N-terminal His-tag and were transformed into *E. coli* DH5α (Life Technologies, Carlsbad, USA). Also, the *cdpgs* gene was codon-optimized (Eurofins, Ebersberg, Germany) for expression in *E. coli* and subcloned in pET15b (Amp^R) without affinity tag (primer sets and plasmids listed in supplementary table 1). All sequences and plasmids were verified by sequencing (LGC genomics, Berlin, Germany).

Heterologous expression. Competent cells of *E. coli* BL21(DE3)-CodonPlus-pRIL (Cam^R) (Merck, Darmstadt, Germany) were transformed with the constructed plasmids and grown in 200 mL – 400 mL lysogeny broth medium (LB broth) (Carl Roth GmbH + Co. KG, Karlsruhe, Germany). Gene expression was induced by the addition of 1 mM isopropyl-β-d-thiogalactopyranoside (IPTG) at an OD_{600} between 0.6 - 0.8. Cultures were incubated over night at 30°C, and for improved expression results at 16°C. Cells were harvested (6.000 x g, 20 min, 4°C) and per 1 g of the wet cell mass 5 mL of buffer were used for resuspension (50 mM MES (2-(N-morpholino)ethanesulfonic acid), 5 mM DTT, 2 mM $MgCl_2$, 300-500 mM KCl, pH 6.5). For cell lysis, cell suspensions were passed three times through a French pressure cell at 20.000 psi. The cell lysate was centrifuged (25.000 x g, 45 min, 4°C) to remove all cell debris. The supernatant containing the recombinant thermophilic cDPGS was subjected to a heat precipitation step at 75°C for 30 min (Eppendorf Thermomixer at 700 rpm) followed by centrifugation (25.000 x g, 45 min, 4°C) to remove denatured mesophilic host proteins.

Purification. Protein fractions obtained after heat treatment were further purified. For affinity chromatography, the supernatant was transferred onto a Protino® Ni-TED column (tris-carboxymethyl ethylene diamine 1000 or 2000) (Macherey-Nagel GmbH & Co. KG, Düren, Germany) and purified following the manufacturers' instructions using 50 mM Tris/HCl or 50 mM potassium phosphate buffer pH 8 (50 mM Tris or 50 mM K_2HPO_4/KH_2PO_4, 5 mM DTT, 300 mM KCl or 300 mM NaCl, 250 mM imidazole). Protein dialysis was performed via the Spectra/Por® dialysis membrane standard RC tubing (Repligen, Waltham, USA) with a molecular weight cutoff of 30 kDa. For the protein that was produced using the codon-optimized gene without affinity tag, purification was obtained via size exclusion chromatography (SuperdexTM 200, HiLoadTM 26/60 column, 320 ml volume, GE Healthcare

Life science, Freiburg, Germany) equilibrated in 50 mM MES/KOH, 300 mM KCl, pH 6.5 (flow rate 2 mL/min). The expression, purification, and subunit molecular mass were monitored by denaturing SDS-polyacrylamide gel electrophoresis. Protein concentrations were determined with a modified Bradford method using bovine serum albumin as a standard (Bio-Rad, Feldkirchen, Germany).

DETERMINATION OF cDPGS ACTIVITY IN A CONTINUOUS ASSAY, CONSUMPTION OF 2,3 DPG AND cDPG FORMATION IN A DISCONTINUOUS ASSAY SYSTEM

Enzyme Properties. The enzymatic activity was determined by coupling the ADP formation from ATP to the oxidation of NADH via the pyruvate kinase (PK) and L-lactate dehydrogenase (LDH) (rabbit muscle, Merck, Darmstadt, Germany). The assay mixtures (0.5 mL) contained 50 mM MES/KOH pH 6.5 with 2.5 mM $MgCl_2$, 400 mM KCl, 2.5 mM ATP, 0.5 mM NADH, 3.7 µg purified cDPGS, 8 U PK, 4 U LDH, and 2 mM phosphoenolpyruvate. After 1 min of preincubation at 55°C the reactions were started by the addition of 2,3 di-phosphoglycerate (2,3 DPG, MERCK, Darmstadt, Germany) and the oxidation of NADH was followed in a Specord 210 UV/VIS spectrometer (Analytic Jena AG, Jena, Germany). All measurements were performed in triplicates.

Effect of salt and pH. Stimulation of enzyme activity was analyzed in presence of KCl. Therefore, the enzymatic activity was analyzed as described above but with different salt concentrations of up to 500 mM and 16.8 µg partial purified cDPGS after heat treatment. In addition, the pH optimum was determined in a range from pH 5.5 – 8.0 using 50 mM MES/KOH pH 5.5 and 6.5, 50 mM TES/KOH pH 6.8, 50 mM potassium phosphate buffer pH 6.8 and HEPES/KOH pH 7.0 and 8.0 supplemented with 400 mM KCl.

Consumption of 2,3 DPG and formation of cDPG. The cDPG synthesis was measured in a discontinuous assay (1 mL) in 50 mM MES/KOH pH 6.5, 2.5 mM $MgCl_2$, 2.5 mM ATP, 1 mM 2,3 DPG, 400 mM KCl at 55°C using 7.4 µg of the purified cDPGS. In regular intervals 100 µL samples were removed into precooled reaction tubes and directly stored at -80°C to stop the reaction. The amount of consumed 2,3 DPG in 10 µL of these samples was determined via the 2,3 DPG test kit from ROCHE (MERCK, Darmstadt, Germany). The assay was performed according to the manufacturer's instructions but modified for 96-well plates in a total volume of 200 µl using a NADH standard curve obtained via the Tecan Microplate Reader at 340 nm (Tecan group Ltd., Lifesciences, Männedorf, Switzerland). Respectively the amount of resulting ADP from ATP was determined from 50 µL of these samples using the PK and LDH assay conditions as it was described above.

Results

EXPRESSION AND PURIFICATION OF THE *M. fervidus* cDPGS

The heterologous expression of the cyclic-2,3-diphosphoglycerate synthetase (*cdpgs*) from *M. fervidus* in *E. coli* yielded only low amounts of soluble protein (supplementary Figure 1). Since the codon adaption index for the cDPGS expression in *E. coli* was low (CAI = 0.54), the gene was codon optimized (Eurofins, Ebersberg, Germany) for the expression in *E. coli*. This led to higher amounts of soluble protein (supplementary Figure 2). However, the His-tagged cDPGS obtained from these expressions was rather instable in solution. We therefore expressed the codon-optimized gene without any affinity tag in *E. coli* BL21(DE3)-CodonPlus-pRIL (Cam^R) (MERCK, Darmstadt, Germany) and additionally lowered the expression temperatures to 16°C. After overnight expression, the enzyme was purified to apparent homogeneity via an easy heat precipitation step (75°C, 30 min) followed by size exclusion chromatography. From 400 mL expression culture corresponding to 1.7 g cell wet weight, 3.5 mg of apparently pure protein was obtained. The molecular weight of the cDPGS was 52 kDa under denaturating conditions, determined by SDS-PAGE (Figure 2) and 119 kDa under native conditions determined by size exclusion chromatography indicating a homodimeric structure of the protein.

CHARACTERIZATION OF THE *M. fervidus* cDPGS

The cDPGS catalyzes the ATP depended formation of an intramolecular phosphoanhydride bond in 2,3 DPG resulting in cyclic 2,3-diphosphoglycerate (cDPG). The activity was determined by following the ADP-formation via pyruvate kinase (PK) and L-lactate dehydrogenase (LDH) (rabbit muscle, Merck, Darmstadt, Germany). However, cDPGS activity was proven to be highly salt dependent: A 5.6-fold activation was observed in presence of 500 mM KCl (Figure 3A). The pH optimum of the cDPGS was pH 6.5 and the optimal buffer for characterization was 50 mM MES/KOH supplemented with 400 mM KCl, 2.5 mM $MgCl_2$, and 10 mM DTT (Figure 3B).

The rate dependence of the pure cDPGS followed classical Michaelis-Menten kinetics with a V_{max} of 21.7 U mg^{-1} and K_m values of 1.4 mM and 1.1 mM for 2,3 DPG and ATP, respectively (Figure 4A & B).

The addition of 400 mM KCl, 2.5 mM $MgCl_2$,and 10 mM DTT as a reducing agent was required during all purification steps. The stability of the cDPGS could be significantly improved under these conditions. For storage at -80°C the addition of 25% (v/v) glycerol to the same buffer was required to maintain enzyme activity. After 1 week of storage at -80°C the enzyme remained 100% active, and even after 1.5 month nearly full activity (95%) was retained (Figure 5).

cDPG PRODUCTION

The *in vitro* synthesis of cDPG from 2,3 DPG was analyzed at 55°C in a small-scale. Therefore, 1 mM 2,3 DPG (0.76 mg in 1 mL) was incubated with 2.5 mM ATP in 50 mM MES/KOH pH 6.5 containing 400 mM KCl, 2.5 mM $MgCl_2$, 10 mM DTT and 7.4 µg of the purified cDPGS (corresponding to 0.11 U). The consumption of 2,3 DPG was followed in a discontinuous assay via the modified assay procedure of the 2,3 DPG test Kit from Roche (Merck, Darmstadt, Germany) as described in the Material and Methods section. After 10 min of incubation, 56% of the 2,3 DPG were converted by cDPGS, after 30 min 92%, and after 60 min the 2,3 DPG was completely consumed (Figure 6). In addition, the corresponding amount of ADP was formed as determined via the pyruvate kinase (PK) and L-lactate dehydrogenase (LDH) (rabbit muscle, Merck, Darmstadt, Germany) (Figure 6). As a control, the non-enzymatic hydrolysis of ATP at 55°C was determined under the same condition but without the recombinant cDPGS. It turned out that only 4.5% of the ATP was lost after 60 min due to non-enzymatic hydrolysis. Together, these results indicate a complete conversion of 2,3, DPG to cDPG by the recombinant cDPGS from *M. fervidus*.

Discussion

Extremolytes have a high application and economic potential and thus represents valuable products for the cosmetic, pharmaceutical, medical, and food industry, where the protection of biomolecular structures is an essential key to success. However, only a few extremolytes have been established on the market so far particularly ectoine and hydroxyectoine, which are derived and produced by the halophilic bacterium *Halomonas elongata* ([182, 184] Bitop AG: https://www.bitop.de/de/extremolytes)). Among extremolytes of hyperthermophiles, cyclic 2,3-diphosphoglycerate (cDPG) was exclusively reported in methanogenic Archaea and the stabilizing potential on biomolecular structures have already been demonstrated [182, 184]. In *Methanothermus fervidus* an increase of intracellular cDPG concentration with increasing growth temperature was observed [196,197]. Furthermore, in hyperthermophilic methanogens intracellular cDPG concentrations tended to increase with the optimal growth temperature of the organism [255]. Thus, an important role of cDPG in thermoadaption was suggested [196, 197, 255]. *In vitro*, cDPG was shown to enhance the thermostability of some model proteins including the lactate-dehydrogenase from rabbit muscle, the glyceraldehyde-3-phosphate dehydrogenase and the 3-phosphoglycerate kinase from *M. fervidus,* as well as the methyl-coenzyme M reductase and cyclohydrolase from *M. kandleri* [201, 203, 204]. Furthermore, prevention of protein misfolding and aggregate formation in presence of extremolytes have been shown in a few studies, which is of great importance in the development of new drugs for the treatment of neurodegenerative diseases [256, 257]. On DNA level, cDPG functions as superoxide scavenger and protects plasmid DNA against oxidative damage [182].

Despite this broad range of beneficial properties, a strategy of cDPG production has not yet been developed. An *in vivo* production by whole-cell biocatalysis appears not feasible due to challenging cultivation conditions posed by the strictly anaerobic and hyperthermophilic lifestyle of the natural methanogenic producer strains, which results in only low cell yields and therefore likewise low product yields. *In vitro* enzymatic approaches can circumvent these cultivation related problems and additionally have some further advantages over whole-cell systems including circumvention of regulation issues, of competing metabolic pathways, kinetic restrictions, and toxicity problems [206]. Furthermore, for cDPG the biosynthetic pathway is relatively simple and starts from the central metabolic intermediate 2-phosphoglycerate (2PG) involving only two enzymes [197]. However, such enzymatic *in vitro* approaches require sufficient amounts of enzyme, and thus optimized expression and purification procedures. And also, sufficient enzyme stability for long term usage and storage is an important property to render such approaches efficient.

To meet these challenges and to keep the system as simple as possible, we first decided to go for a one-step approach using 2,3DPG as a raw material since it is less costly than 2PG, and only the cDPGS is required as biocatalyst. However, expression levels and protein yields of former heterologous production efforts in *E. coli* were low (0.9 mg g^{-1} cells wet weight) [202] and in initial experiments, this problem was maintained. Therefore, the differences in gene structures and codon usage may explain these experimental difficulties sometimes encountered when methanogenic genes are heterologously expressed in non-methanogens [258]. Codons, such as AUA, AGA, and AGG, are frequently found in methanogens but rarely in *E. coli* resulting in insufficient expression and protein yields [259, 260]. In addition, the high level of As and Ts in the third position of the codon, often found in methanogens, may also pose a problem for proper expression in other organisms [258]. Improved expression of archaeal methanogenic proteins in *E. coli* were described in serval cases when expressed in the presence of rare tRNAs (e.g., argU, ileY, ileX, and leuW) [260, 261]. We therefore decided to codon optimize the cDPGS encoding gene for expression in *E. coli* and then also lowered the expression temperature which represents another common strategy to solve expression issues [262].

The recombinant cDPGS from *M. fervidus* showed a remarkable instability which is also in accordance with previous findings for cDPGSs [202, 263] and was also described for other proteins from methanogens [201, 203, 264-266]. Furthermore, the cDPGS was shown to rely on potassium ions (up 1 M) as well as magnesium ions (2.5 mM) for full activity [197, 205, 263] and was also indicated to be stabilized by potassium ions as well as by reducing agents like DTT to minimize oxygen susceptibility as described for other proteins from methanogens e.g., *Methanpyrus kandleri* or *Methanothermobacter thermoautotrophicus* [264-266].

Herein, it was shown that the addition of 400 mM KCl, 10 mM DTT, and 2.5 mM Mg^{2+} increased the short-term stability to at least several days. And we could also show that the additional supplementation of the buffer with 25% glycerol enabled the long-term storage at -80°C for months without loss of activity. However, the His-tag we initially introduced to the recombinant enzyme to streamline the purification appeared to even increase the instability. And this might therefore be another example that protein tags in some cases interfere with protein function, folding, activity, and stability [267-269]. The recombinant protein was therefore expressed as the native non-tagged version and purification was carried out in the presence of 400 mM KCl, 2.5 mM $MgCl_2$, and 10 mM DTT. For purification, a simple two-step procedure was established involving heat precipitation for 30 min at 75°C and size exclusion chromatography. This finally resulted in a more than two-fold increased protein yield (2 mg g^{-1} cell wet weight) compared to that reported previously (0.9 mg g^{-1} cells wet weight) [202]. The advantage of this procedure appears that the KCl concentration can be kept high during the whole purification procedure of the recombinant cDPGS. Conversely, the method used by Matussek and coworkers [202] involved an affinity chromatography on hydroxyapatite which required low salt conditions. This likely caused a tremendous loss in active protein (see above). The cDPGS produced by this newly developed procedure showed a much higher V_{max} (21.7 U mg^{-1} at 55°C) than reported previously (17 U mg^{-1} at 83°C) and also the Km values tended to be lower. These differences are likely due to the different purification and assay systems used in our study. The homodimeric structure determined however confirmed the previous findings [202, 205, 263].

An important result of this work was that cDPGS completely convert 2,3DPG to cDPG. In a mL range, 1 µmol of substrate was consumed within 60 min by 7.4 µg of purified enzyme. The conversion efficiency has so far not been analyzed in detail. The detailed kinetic analyses of the *M. fervidus* cDPGS in both direction of the reaction revealed that the catalytic efficiency is roughly 100-fold higher in the cDPG synthesis direction. This already shows a strong preference for the synthesis direction [202] although the exact thermodynamics of the reaction is not known. Overall, a phosphoanhydride bond is formed in cDPG at the expense of a phosphoanhydride bond in ATP, which might suggest full reversibility. However, the $\Delta G^{0'}$ of hydrolysis of the anhydride bond in pyrophosphate, and cDPG is classified as trianionic pyrophosphate [182], with ~16 kJ mol^{-1}, whereas that of ATP to ADP and Pi is -26 kJ mol^{-1} under standard conditions, pH 7.0, and ionic strength of 0.1 M (eQuilibrator) [270]. This suggests that the anhydride bond formation in 2,3DPG driven by ATP hydrolysis is exergonic by -10 kJ mol^{-1} and would thus explain the complete conversion.

Taken together we developed an efficient cDPGS production and purification procedure enabling sufficient stabilization of the protein. We further showed that the conversion is complete, which allows for an easy one-step approach for the biocatalytic synthesis of cDPG

and significantly simplifies the downstream purification of the product. The obtained results now allow the up-scaling of the process as well as to extend the process to a more complex enzyme cascade e.g., involving an ATP regenerating system or to start synthesis from even more inexpensive substrates like glycerate. Thus, our study paved the way for the cost-effective production of cDPG as added-value product for use in biotechnological applications.

Figures

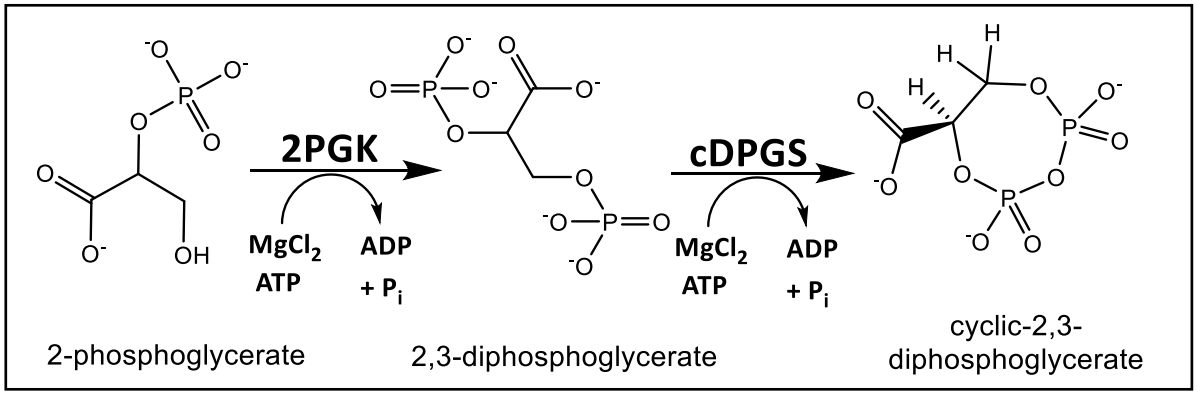


Figure 1: cDPG synthesis. Conversion of 2PG to 2,3DPG by 2PGK and formation of cDPG by cDPGS. Both catalytic steps are ATP-dependent and require additional magnesium ions. Abbreviations: 2PG, 2-phosphoglycerate; 2,3DPG, 2,3-diphosphoglycerate; cDPG, cyclic di-2,3-phosphoglycerate; and cDPGS, cyclic-diphosphoglycerate synthetase.

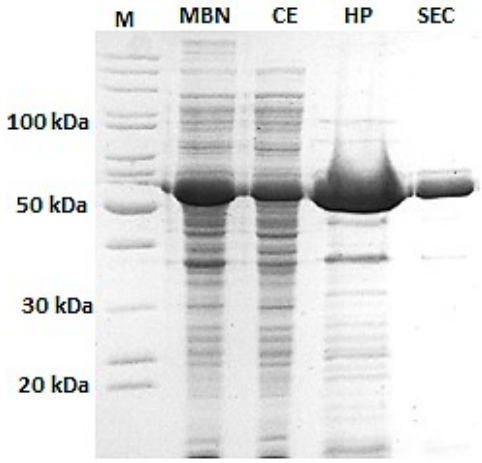


Figure 2: Purification of recombinant cDPGS from *M. fervidus* heterologously expressed in *E. coli*. The codon optimized *cdpgs* was cloned in pET15b without affinity tag. The expression was optimized in *E. coli* BL21(DE3)-Codon-Plus and the protein was purified by heat treatment (HP) and size exclusion chromatography (SEC). Protein fractions (5-10 µg) after successive purification steps were analyzed via SDS-PAGE (12.5 %) and stained with Coomassie Brilliant Blue. Mbn: Membrane fraction, CE: Crude extract, HP: Soluble fraction after heat precipitation (75°C, 30 min), SEC: Fraction after size exclusion chromatography. M: Protein marker, unstained protein ladder (Thermo Fisher Scientific, Carlsbad, USA).

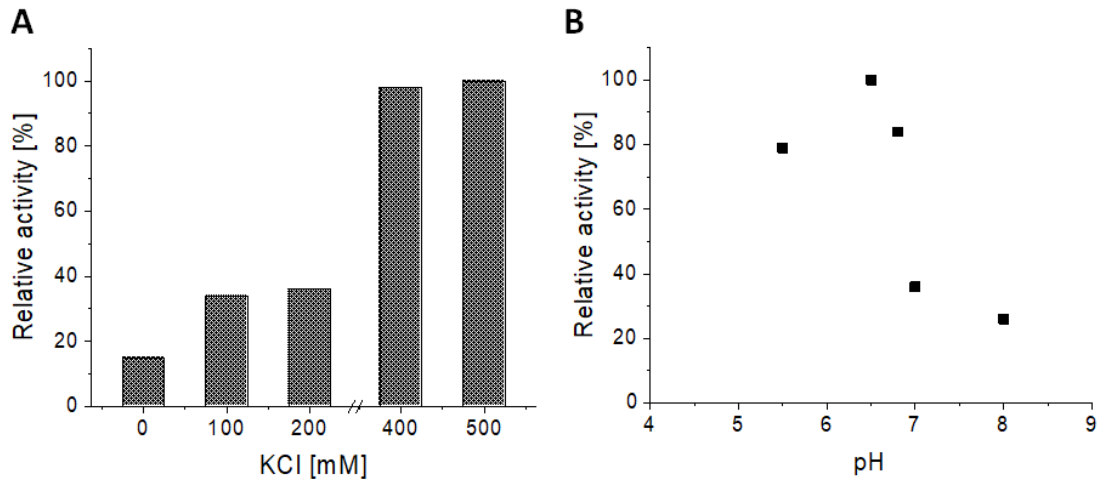


Figure 3: Effect of salt and pH on the activity of recombinant cDPGS. Effect of KCl (0 - 500 mM) **(A)** and of pH (5.5 - 8) **(B)** on the activity of partial purified cDPGS (16.8 µg) after heat treatment (75°C, 30 min). The formation of ADP from ATP was coupled to the oxidation of NADH via the pyruvate kinase (PK) and lactic acid dehydrogenase (L-LDH) from rabbit muscle. The activity was determined as decrease in absorbance at 340 nm at 55°C. The pH optimum was determined in the buffer systems as described in the material and method section. The respective buffers was supplemented with 400 mM KCl and 2.5 mM $MgCl_2$. 100 % of relative activity correspond to a specific activity of 1.8 U mg^{-1}.

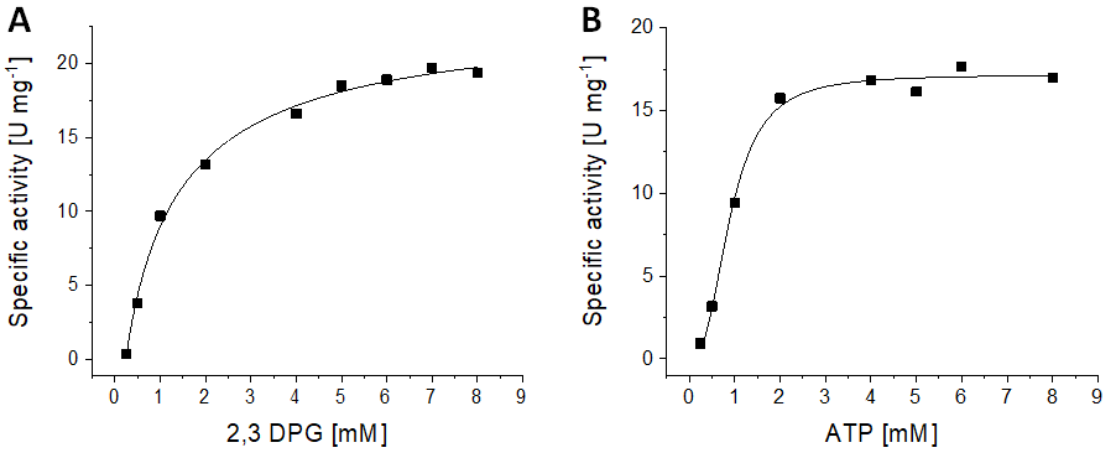


Figure 4: Enzymatic properties of the recombinant cDPGS. The enzymatic activity **(A & B)** was determined at 55°C (340 nm) using 3.7 µg of purified enzyme by coupling the formation of ADP from ATP to the oxidation of NADH with pyruvate kinase (PK) and (lactic acid dehydrogenase (L-LDH) from rabbit muscle as auxiliary enzymes. The specific activity of purified cDPGS was determined with 2,3 DPG concentrations from 0-8 mM and 2.5 mM ATP **(A)** and ATP concentrations from 0-8 mM and 2 mM 2,3 DPG **(B)**.

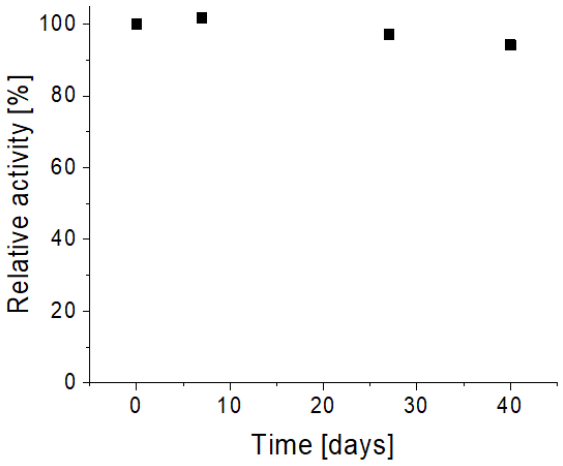


Figure 5: Long term stability of the recombinant cDPGS. The enzyme was stored at -80°C in a buffer composed of MES/KOH pH 6.5, 25% (v/v) glycerol, 400 mM KCl, 5 mM $MgCl_2$ and 10 mM DTT. The residual activity was determined at the time points indicated using the continuous assay as previously described in the material and methods section. 100 % of relative activity correspond to a specific activity of 13.5 U mg^{-1}.

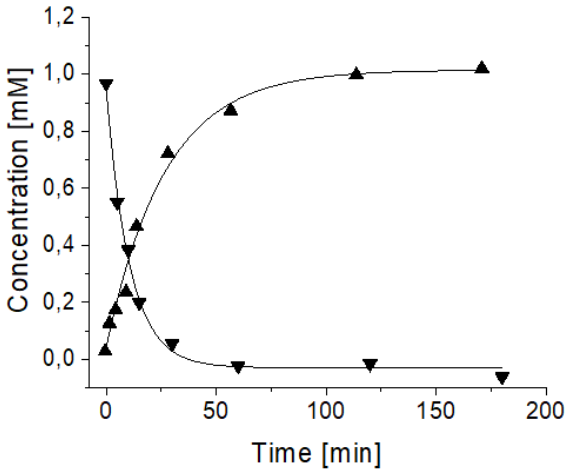


Figure 6: Kinetics of the cDPGS mediated 2,3 DPG conversion to cDPG. The ATP and 2,3 DPG dependent production of cDPG was determined in a lab scale conversion assay. The time dependent consumption of 2,3 DPG (inverted triangles) was analyzed via the 2,3-DPG assay kit from Roche (Mannheim, Germany) according to the manufacturer's instructions but modified for 96-well plates. The corresponding amount of ADP (triangles) produced from ATP was calculated from the respective reactions coupled to the oxidation of NADH using the pyruvate kinase (PK) and lactic acid dehydrogenase (L-LDH) from rabbit muscle as auxiliary enzymes as previously described in the material and method section.

Tab. 1: Purification of recombinant 2PGK and cDPGS from 1.7 g E. coli cells

Purification step	Protein (mg)	Total activity (U)	Spec. act. (U/mg)	Purification (-fold)
Crude extract	22.5	2.9	0.13	-
Heat treatment	5.2	24.9	4.8	37
Size exclusion chromatography	3.5	75.9	21.7	167

Supplementary information on

Enzymatic Synthesis of the Extremolyte Cyclic-2,3-diphosphoglycerate by the Cyclic-2,3-diphosphoglycerate Synthetase from *Methanothermus fervidus*

Supplementary table 1: List of primers for cloning of *M. fervidus cdpgs*.

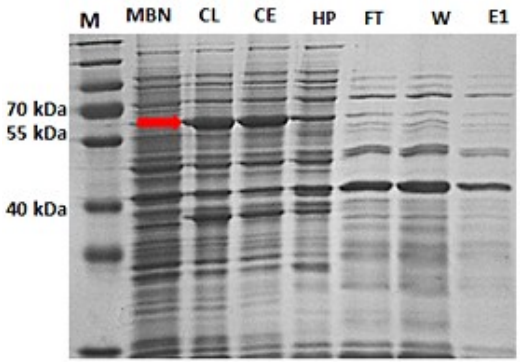
Primer	Sequence (5'-3')
f-*Nco*I_*cdpgs*	tctaccatgggggcgagacaaaaa
r-*Bam*HI_ *cdpgs*	catcggatccttagcggttgttctt
f-p24-*cdpgs-Hind*III	gacgaagcttatgggtgaaactaaaaaaatg
r-p24- *cdpgs-Xho*I	gcacctcgagcctattgtttttaaaatcatctattgcc
f-*EcoR*I-*cdpgs*	tcgtgaattcatgaccgccgtgaagaggatac
r-*Hind*III-*cdpgs*	taataagcttaccctcgaaccgatcgatcgcca
Plasmids	**Source**
pEX_K4_cdpgs_Mfer_opt_bac	Eurofins genomics
pET24a_cdpgs_orig:His-tag	This work
pET15b_ cdpgs_opt:His-tag	This work
pET15b_cdpgs_opt_no-tag	This work

>*Mfer_cdpgs* original

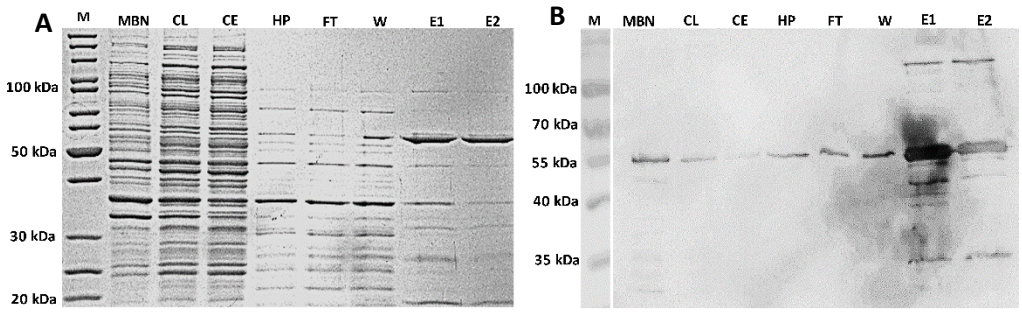
atgggtgaaactaaaaaaatgatatgcttagtagatggtgaacattattttcctgtagttaaagattctattgaaatattggatgatcttgaacatatagacgttgtagct
gtagtattcattgggggaactgaaaaactccaaattgaagatcctaaagaatattctgaaaaattaggaaaaccagttttttttgggcctgatccaaaaaaaatac
catatgatgtaataaaaaaaatgtgtaaaaaaatataatgctgatatagttatggatttaagtgatgaacctgttgtagattatacaaaaagatttagaatagcatcca
tagttctaaaagaaggcgctgtatatcagggcgctgattttaaatttgaaccacttacagaatatgatgtgttagaaaagccatcgattaaaattataggaactgga
aaaagaataggaaaaactgcggtatctgcatatgctgcaagagtcattcataaacataaatacaatccatgtgtagttgcgatgggtcgtggaggaccacgag
aacctgaaatagtagaaggaaataaaatagaaataacagctgaatatttattagaacaagctgataaaggtgtacatgcagcatcagatcattgggaagatgc
attaatgagtagaattcttacagtcggatgtagaagatgtggcggtggaatgcttggtgatacattcataacaaatgtgaaaagaggtgctgaaattgccaataaa
ttagattctgactttgttataatggaaggtagtggagcagcaatacctcctgttaaaacaaataggcaaatagttacagttggtgcaaatcagccaatgataaatatt
aacaatttctttggaccatttaggataggattagcagatcttgtcataataacaatgtgtgaagaacctatggcaactacagaaaaaattaaaaaggtagaaaaa
tttataaaagaaataaatccatcagcaaatgttattcctactgtatttagacctaaaccagtaggcaatgttgaaggtaaaaaagtgttatttgcaactacagcacc
aaaagttgttgtagggaaattagttaattatctagaaagtaaatatggatgcgatgtggtaggtgtcacaccacatttatcaaatcgtccattattacgtagagattta
aaaaaatatataaacaaggcagatttaatgttgacggaattaaaggctgcagctgttgacgttgcaactagggtagctatagaagctggtctagatgttgtatattg
tgataatattcctgtagttatagatgagagttatggaaacatcgatgatgcaattattgaagttgtagaaatggcaatagatgattttaaaaacaataggtga

>*cdpgs* codon optimized for expression in *E. coli*

ATGGGCGAGACAAAAAAGATGATTTGCCTGGTAGATGGGGAACACTATTTTCCTGTTGTTAAAGACAGCATTG
AAATCCTCGATGATCTGGAGCATATCGACGTAGTGGCTGTGGTATTCATCGGCGGAACCGAGAAACTGCAGA
TTGAAGATCCGAAAGAATATTCGGAAAAACTGGGCAAACCTGTGTTCTTTGGACCCGATCCGAAGAAAATTCC
GTATGACGTTATCAAGAAATGCGTCAAGAAATACAATGCGGATATTGTGATGGATCTTTCTGACGAACCAGTA
GTGGACTACACCAAACGGTTTCGCATCGCCTCCATTGTGCTGAAAGAGGGCGCAGTTTATCAAGGGGCCGA
TTTTAAATTTGAACCGCTGACTGAATACGATGTTTTGGAGAAACCGTCTATCAAAATTATTGGTACCGGGAAAC
GCATTGGTAAGACAGCGGTGAGTGCGTATGCAGCCCGTGTGATTCACAAGCATAAATACAATCCCTGTGTAG
TTGCAATGGGCCGTGGTGGACCACGTGAACCGGAGATTGTGGAGGGCAACAAAATCGAAATCACCGCCGAA
TATCTGCTTGAGCAAGCGGATAAAGGCGTTCATGCAGCCAGCGATCATTGGGAAGATGCCCTGATGAGTCG
CATTCTGACGGTTGGATGTCGTCGTTGTGGTGGTGGGATGCTGGGCGACACGTTCATTACCAACGTCAAACG
TGGTGCAGAGATTGCGAACAAACTGGACTCAGATTTTGTCATTATGGAAGGTTCAGGTGCGGCAATTCCGCC
GGTGAAAACGAATCGGCAGATTGTCACTGTTGGCGCCAATCAGCCGATGATCAACATCAATAACTTCTTTGG
CCCGTTTCGCATTGGCTTAGCCGATTTGGTCATCATTACCATGTGTGAAGAACCGATGGCGACCACCGAAAA
GATCAAGAAAGTTGAGAAATTCATTAAAGAGATCAATCCCAGCGCTAATGTGATTCCGACGGTTTTCCGCCCA
AAACCTGTGGGTAACGTCGAAGGTAAAAAAGTGTTGTTTGCGACCACGGCCCCGAAAGTTGTGGTAGGGAAA
CTCGTGAATTACCTGGAATCGAAATATGGCTGCGATGTAGTGGGTGTTACGCCACACCTGAGCAATCGCCCT
CTGTTACGTCGCGATTTAAAGAAATACATTAACAAAGCGGATCTTATGCTCACTGAACTGAAAGCGGCTGCTG
TGGATGTCGCGACACGCGTAGCTATTGAAGCGGGCTTAGATGTCGTGTATTGCGACAACATCCCAGTCGTCA
TCGACGAATCCTATGGCAACATTGACGATGCAATCATCGAAGTGGTCGAAATGGCTATCGACGACTTCAAGA
ACAACCGCTAA

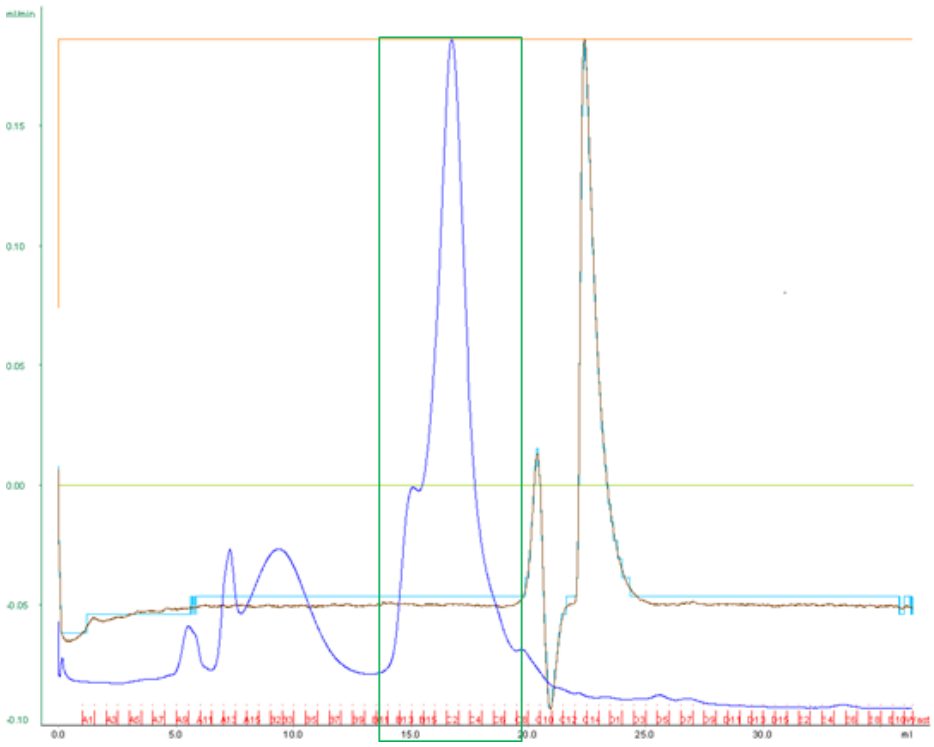


Supplementary figure 1: Expression and purification of recombinant cDPGS from *M. fervidus* heterologously expressed in *E. coli*. The *cdpgs* gene was cloned in pET24a with C-terminal His-tag and expressed in *E. coli* BL21(DE3)-Codon-Plus. The protein fractions (5-10 µg) after successive purification steps were analyzed via SDS-PAGE (12.5 %) and stained with Coomassie Brilliant Blue. Mbn: Membrane fraction, CE: Crude extract, CL: crude lysate HP: Soluble fraction after heat precipitation (75°C, 30 min), FT: Flow through fraction-, W: Wash fraction-, E1: Elution fraction 1 from Ni-TED affinity chromatography. M: Protein marker, prestained protein ladder (Thermo Fisher Scientific, Carlsbad, USA). High amounts of the recombinant protein were detected in the insoluble membrane fraction, marked by the red arrow.



Supplementary figure 2: Expression and purification of recombinant codon optimized cDPGS from *M. fervidus* and immunodetection via anti-His-tag antibody. The codon optimized *cdpgs* gene was cloned in pET15b with N-terminal His-tag and expressed in *E. coli* BL21(DE3)-Codon-Plus. The protein fractions (5-10 µg) after successive purification steps were analyzed via SDS-PAGE (12.5 %) **(A)** and stained with Coomassie Brilliant Blue. For immunodetection of the His-tagged cDPGS **(B)** anti-His-tag antibody was used (Thermo Fisher Scientific, Carlsbad, USA). Mbn: Membrane fraction, CE: Crude extract, CL: crude lysate HP: Soluble fraction after heat precipitation (75°C, 30 min), FT: Flow through fraction-, W: Wash fraction-, E1-2: Elution fraction 1-2 - from Ni-TED affinity chromatography. M: Protein marker, (A) unstained and (B) prestained protein ladder (Thermo Fisher Scientific, Carlsbad, USA).

Purification



Supplementary figure 3: Purification chromatogram from size exclusion chromatography (SEC). The filtered soluble fraction of cDPGS after heat precipitation at 75°C (5.6 mg) was loaded onto Supderdex200 size exclusion column. The peak containing the recombinant cDPGS is boxed in green. Fractions containing the pure enzyme were pooled and concentrated.

Chapter 3.3

Isolation and Characterization of the First Xylanolytic Hyperthermophilic Euryarchaeon *Thermococcus* sp. Strain 2319x1 and Its Unusual Multidomain Glycosidase

frontiers in Microbiology

ORIGINALRESEARC H
published:03May201 6
doi:10.3389/fmicb.2016.0055 2

Isolation and Characterization of the First Xylanolytic Hyperthermophilic Euryarchaeon *Thermococcus* sp. Strain 2319x1 and Its Unusual Multidomain Glycosidase

Sergey N. Gavrilov[1†], Christina Stracke[2†], Kenneth Jensen[3], Peter Menzel[4], Verena Kallnik[2], Alexei Slesarev[1,5], Tatyana Sokolova[1], Kseniya Zayulina[1], Christopher Bräsen[2], Elizaveta A. Bonch-Osmolovskaya[1], Xu Peng[4], Ilya V. Kublanov[1] and Bettina Siebers[2]**

[1] Winogradsky Institute of Microbiology, Research Center of Biotechnology, Russian Academy of Sciences, Moscow, Russia,

[2] Molecular Enzyme Technology and Biochemistry, Biofilm Centre, Centre for Water and Environmental Research, University Duisburg-Essen, Essen Germany

[3] Novozymes A/S, Bagsværd, Denmark, [4] Department of Biology, University of Copenhagen, Copenhagen, Denmark

[5] Fidelity Systems, Inc., Gaithersburg, MD, USA

Edited by:
Belinda Ferrari,
University of New South Wales,
Australia

Reviewed by:
Jun-Jie Zhang,
Indiana University School of Medicine,
USA
Nick Sirijovski,
Lund University, Sweden

***Correspondence:**
Bettina Siebers bettina.siebers@uni-due.de;
Ilya V. Kublanov kublanov.ilya@gmail.com

†These authors have contributed equally to this work.

Specialty section:
This article was submitted to Microbiotechnology, Ecotoxicology and Bioremediation, a section of the journal
Frontiers in Microbiology

Received: 26 December 2015
Accepted: 04 April 2016
Published: 03 May 2016

Citation:
Gavrilov SN, Stracke C Jensen K Menzel P, Kallnik V, Slesarev A, Sokolova T, Zayulina K, Bräsen C, Bonch-Osmolovskaya EA, Peng X, Kublanov IV and Siebers B (2016) Isolation and Characterization of the First Xylanolytic Hyperthermophilic Euryarchaeon Thermococcus sp. Strain 2319x1 and Its Unusual Multidomain Glycosidase. Front. Microbiol. 7:552. doi: 10.3389/fmicb.2016.00552

Enzymes from (hyper)thermophiles "Thermozymes" offer a great potential for biotechnological applications. Thermophilic adaptation does not only provide stability toward high temperature but is also often accompanied by a higher resistance to other harsh physicochemical conditions, which are also frequently employed in industrial processes, such as the presence of, e.g., denaturing agents as well as low or high pH of the medium. In order to find new thermostable, xylan degrading hydrolases with potential for biotechnological application we used an *in situ* enrichment strategy incubating Hungate tubes with xylan as the energy substrate in a hot vent located in the tidal zone of Kunashir Island (Kuril archipelago). Using this approach a hyperthermophilic euryarchaeon, designated *Thermococcus* sp. strain 2319x1, growing on xylan as sole energy and carbon source was isolated. The organism grows optimally at 85°C and pH 7.0 on a variety of natural polysaccharides including xylan, carboxymethyl cellulose (CMC), amorphous cellulose (AMC), xyloglucan, and chitin. The protein fraction extracted from the cells surface with Tween 80 exhibited endoxylanase, endoglucanase and xyloglucanase activities. The genome of *Thermococcus* sp. strain 2319x1 was sequenced and assembled into one circular chromosome. Within the newly sequenced genome, a gene, encoding a novel type of glycosidase (143 kDa) with a unique five-domain structure, was identified. It consists of three glycoside hydrolase (GH) domains and two carbohydrate-binding modules

Escherichia coli and their activity was analyzed. The full length multidomain glycosidase (MDG) was able to hydrolyze various polysaccharides, with the highest activity for barley β-glucan (β1,3/1,4-glucoside), followed by that for CMC (β-1,4-glucoside), cellooligosaccharides and galactomannan. The results reported here indicate that the modular MDG structure with multiple glycosidase and carbohydrate-binding domains not only extends thesubstrate spectrum, but also seems to allow the degradation of partially soluble and insoluble polymers in a processive manner. This report highlights the great potential in a multi-pronged approach consisting of a combined *in situ* enrichment, (comparative) genomics, and biochemistry strategy for the screening for novel enzymes of biotechnological relevance.

Keywords: archaea, hyperthermophiles, *Thermococcus*, xylan, multidomain glycosidase, endoglucanase

INTRODUCTION

The use of molecular ecology approaches revealed that the vast majority of the total microbial biodiversity was so far not cultivated (Barns et al., 1996; Hugenholtz et al., 1998) highlighting that most of the natural diversity with an inestimable metabolic variability and potential is still hidden. This offers a great reservoir for novel biocatalysts with significant potential in biotechnological applications and process optimization. Enzymes from hyperthermophiles are typically folded into very stable conformations able to withstand high temperatures. The temperature stability is often associated with a high resistance to chemical denaturants commonly used in many industrial applications. The high stability combined with an optimal activity at high temperatures has led to a strong interest in using enzymes from hyperthermophiles in a wide range of commercial applications, most well-known, e.g., in starch and the cellulosic ethanol industries. Both of these applications require enzymes active at high temperatures, thus allowing better substrate solubility, easier mixing and lowered risk of contamination. Archaea are abundant in extreme habitats, especially in hyperthermophilic environments. They exhibit unique metabolic features and modified metabolic pathways characterized by unusual enzymes, not homologous to their bacterial counterparts. Therefore, archaeal enzymes offer great potential as novel biocatalysts for industry, which has continuously high demand for innovative solutions in the biocatalytic sector. Especially, (hyper)thermostable and thermoactive hydrolases degrading recalcitrant polysaccharides like cellulose and xylan into their monomeric constituents are of great importance for, e.g., pulp, biofuel and food industries (Egorova and Antranikian, 2005; Unsworth et al., 2007).

The majority of so far cultured hyperthermophilic organotrophic archaea of the two well-represented phyla, Euryarchaeota and Crenarchaeota are anaerobes growing by fermentation of various complex peptides and proteinaceous substrates, such as peptone, tryptone, beef, and yeast extract (Amend and Shock, 2001). In contrast, only few of these organisms were found to be able to grow on polysaccharides (Supplementary Figure S1). According to our knowledge, there are only three reports showing weak growth of hyperthermophilic archaea on cellulose and its derivatives, i.e., for *Desulfurococcus fermentans* (Perevalova et al., 2005), *Thermococcus sibiricus* (Mardanov et al., 2009) and a consortium of three species with predominance of an *Ignisphaera* representative (Graham et al., 2011). The latter was also shown to possess cellulase activity. Growth on xylan or heat treated xylan (121°C, 20 min) was demonstrated only for members of the *Crenarchaeota*, i.e., *Thermosphaera aggregans* (Huber et al., 1998), *Sulfolobus solfataricus* (Cannio et al., 2004), and *Acidilobus saccharovorans* (Prokofeva et al., 2009). In contrast to these scarce reports for growth of hyperthermophilic archaea on polysaccharides, genomes of many of these organisms harbor genes encoding glycoside hydrolases (GHs[1], Supplementary Table S1), and several cellulases and xylanases were isolated from archaeal strains. However, most of these strains were either unable to grow on crystalline cellulose or xylan or were not analyzed for the ability to grow on these substrates (Ando et al., 2002; Cannio et al., 2004; Maurelli et al., 2008). Therefore, the function and efficiency of these enzymes for *in vivo* polymer degradation is still unclear.

For the identification of novel enzymes two main approaches are currently applied: They can be obtained either directly from the environment using high-throughput techniques such as (functional) metagenomics (Ferrer et al., 2015), or through the enrichment and isolation of novel microorganisms. In order to discover efficient biocatalysts, the isolation of novel strains with the desired properties, like the ability to cleave and to grow on cellulose or xylan is advantageous. Therefore, improved cultivation approaches have to be applied, such as providing the most environmentally close conditions for cultivation (Kublanov et al., 2009), utilization of novel substrates and/or electron

[1]http://www.cazy.org/

acceptors, presence or absence of growth factors, as well as the inhibition of cultured fast-growing microorganisms.

Here we describe a multilayered approach for the isolation of novel biocatalysts for biotechnological applications using (i) an *in situ* enrichment strategy for organisms, that are capable of polymer degradation, (ii) genomics, (iii) comparative genomics as well as (iv) cloning and biochemical characterization of enzymes of interest. Using this *in situ* enrichment technique on mineral medium with xylan as the sole carbon source, we isolated a new representative of the *Euryarchaeota, Thermococcus* sp. strain 2319x1. The strain was able to grow on xylan, xyloglucan, alginate, amorphous and CMCs, starch and its derivatives, as well as on mono- and disaccharides. The complete genome of the novel strain was sequenced and revealed the presence of genes encoding diverse hydrolytic enzymes. One of these hydrolases, constitutes a novel multidomain enzyme with a unique three catalytic glycosidase and two carbohydrate-binding domain organization, called multidomain glycosidase (MDG). Upon expression in *Escherichia coli*, the purified recombinant enzyme (full-sized and truncated versions) exhibited endoglucanase as well as numerous additional activities. The broad activity spectrum appears to be facilitated by the modular multidomain architecture also allowing to processively degrade partially soluble and insoluble polymers.

MATERIALS AND METHODS

Chemicals (for Growth Experiments and Enzyme Assays)

Starch was purchased from Merck (Germany). Yeast extract, peptone, gelatin, mono- and disaccharides, barley glucan, birchwood and beechwood xylan, carboxymethyl cellulose (CMC), microcrystalline cellulose (MCC) Avicel, inulin, cellobiose, dextrin, dextran, pullulan, laminarin, lichenan, pectin, and alginate were purchased from Sigma Aldrich (Taufkirchen, Germany), or kindly provided by Dr. R. Wohlgemuth. Agarose (agarose MP) was purchased from Boehringer (Mannheim, Germany) and chitin (crab chitin) from Bioprogress (Russia). Chitin and chitosan were kindly provided by Dr. S. Lopatin from the Centre of Bioengineering, Research Center of Biotechnology, RAS, Moscow, Russia. Amorphous chitin (AMCH) and amorphous cellulose (AMC) for growth experiments and native activity measurements were prepared according to Sorokin et al. (2015). Other polysaccharides, such as glucomannan, galactomannan, arabinoxylan, and curdlan, were purchased from Megazyme (Ireland). Bamboo leaves collected near the sampling site were dried at room temperature and used as the growth substrate.

Sampling, Enrichment, Isolation and Cultivation

Sampling and primary enrichment procedures were performed during the expedition to Kunashir Island (Russia) in July 2011. Hungate tubes (18 mL) containing 200-300 mg of CMC, chitin, agarose, or birchwood xylan, with or without amorphous Fe(III) oxide (ferrihydrite) as external electron acceptor were prepared. Tubes were filled with thermal water and sediments from the sampling site (coordinates: $44^{\circ}11.08333^{0}$, $145^{\circ}50.46667^{0}$), sealed with rubber stoppers and screw caps. Two 1.2-mm syringe needles were inserted into the rubber stoppers to allow the fluid exchange with the environment without a significant loss of insoluble substrates. The enrichment setups were placed into the hot spring, incubated for 6 days, and afterwards sealed and transferred to the lab at ambient temperature. A modified Pfennig medium with ferrihydrite (Gavrilov et al., 2007) or elemental sulfur (10 g L^{-1}) was used for the isolation and metabolic characterization of new organisms. The medium contained 0.12 g L^{-1} Na_2Sx9H_2O, 9.0 g L^{-1} NaCl, 2.0 g L^{-1}, $MgCl_2x6H_2O$, and 0.05 g L^{-1} of yeast extract; the pH was adjusted to 6.0– 7.0. Ten mililiter portions of the medium were dispensed in 18 mL Hungate tubes with 10 mg birchwood xylan or with other substrates as carbon and energy source as indicated. Tubes were inoculated with 1/10 (v/v) of primary enrichments and incubated at temperatures close to values at respective sampling sites. Pure cultures were obtained by the serial dilution-to-extinction technique in the same liquid media.

DNA Isolation, Genome Sequencing and Annotation

The culture was centrifuged at 4500 × *g* for 10 min, and the cell pellet was resuspended in 1 mL TNE buffer, pH 7.4 (20 mM TRIS HCl, 15 mM NaCl, 20 mM EDTA). Three repeated freezing and thawing cycles followed by 30 min incubation at room temperature with the addition of lysozyme (200 µg mL^{-1}) and RNase (DNAse-free, 5 µg mL^{-1}) were performed to disrupt the cells. Subsequently, proteinase K (10 µg mL^{-1}) and SDS [0.5% (w/v)] were added, and the mixture was incubated for 30 min at 54°C. DNA was extracted using phenol:chlorophorm:isoamyl alcohol (50:50:1), pelleted by centrifugation (17000 × *g*, 4°C), washed (70% ice-cold ethanol) and dissolved in TE buffer (10 mM TRIS, 1 mM EDTA, pH 7.4). DNA sequencing was performed by BGI at Shenzhen, China[2]. A 500 bp paired-end library was constructed and sequenced with Illumina HiSeq using the standard protocol of the company.

The Ilumina sequencing resulted in ∼6.1 million read pairs of 2 × 90 bp length, which equals to a ∼560-fold coverage given the final genome size of 1,961,221 bp. An assembly using Velvet (Zerbino and Birney, 2008) resulted in 20 contigs. This initial assembly was used to identify repeat regions that were

[2] http://www.genomics.cn/

subsequently removed. Because of the presence of multiple repeat regions longer than 500 bp, a 2 kb mate-paired Illumina library was constructed and sequenced, and the obtained paired end information was used to arrange multiple screened contigs into a single scaffold using the Phred/Phrap/Consed (Gordon et al., 1998) software package. This package was used for further sequence assembly and quality assessment in the subsequent finishing process. Sequence gaps between contigs that represented repeats were filled with Dupfinisher (Han and Chain, 2006), and a single scaffold was created and verified using available paired end information. Together, the combination of reads from the two Illumina libraries provided a 700-fold coverage of the genome. The origin of replication was predicted using the Orifinder2 tool (Luo et al., 2014). The initial *de novo* gene prediction and annotation was done using the RAST web server (Aziz et al., 2008), which employs the Glimmer-3 gene caller (Delcher et al., 2007). Additionally, non-coding RNAs were predicted by Infernal 1.1 (Nawrocki and Eddy, 2013) using Rfam 12.0 (Nawrocki et al., 2015) as reference. A further manual curation of the RAST functional annotation was accomplished according to Toshchakov et al. (2015).

Native Enzyme Activity Measurements

Native enzyme activities were determined in culture broth as well as in protein fractions that were extracted from cell surfaces. The extraction was performed in 0.05 M MOPS, $pH^{25°C}$ 8.3 with 0.5% (v/v) Tween-80, which appeared to be the most efficient among various tested compounds [i.e., 1 M NaCl, 3 M urea, 0.18 mM SDS, 0.5% (v/v) Triton X-100 or 0.5% (v/v) Tween80] for solubilization of proteins anchored to the cell surface (as determined by liberation of proteins into the solution after 1 h at 25°C). The resulted protein solution was diluted 10-fold with0.05 M MOPS $pH^{25°C}$ 8.3. Glycosidase activities were determined using the dinitrosalicylic acid (DNS) assay (Miller, 1959). Optical density values were converted into sugar concentrations using a D-glucose calibration curve. The calibration curve with D-xylose showed a similar slope as with D-glucose. To ensure that β– glucan does not interfere with the assay, a β–glucan calibration curve was performed, which revealed no effect of β-glucan below a concentration of 0.05 g L^{-1}, both at 25 and at 85°C. Protein concentrations were measured using the bicinchoninic method (Smith et al., 1985).

Cloning of the *mdg* Gene and its Truncated Versions for Protein Expression in *E. coli*

The signal sequence of the *mdg* gene encoding the MDG (ADU37_CDS22600) was identified using the SignalP tool[3]. The cloning of the *mdg* gene (3912 bp) and its truncated versions was performed using the In-Fusion ® HD Cloning Kit (Takara Bio Company). For PCR amplification primers with 15 bp extensions at the 5^0 ends corresponding to the sequence of the *Not*I (forward) and *Eco*RI (reverse) linearized expression vector pET24a were designed for the *mdg* gene and all gene truncations (Supplementary Table S2).

PCR products were purified with the Wizard ® SV Gel and PCR Clean-Up System (Promega). Plasmid DNA was isolated with the GeneJET Plasmid Miniprep Kit (Fermentas). The In-Fusion ® cloning reaction was performed for 15 min at 50°C and the reaction mixtures were directly transformed into Stellar competent cells (Takara) following the instructions of the manufacturer. The presence of the *mdg* full length and truncated inserts was confirmed by digestion of the respective plasmid DNA with gene/fragment specific restriction enzymes (i.e., *Xba*I, *Hind*III, *Xho*I and *Eco*RI), followed by sequencing (LGC Genomics, Berlin).

Protein Expression and Purification

Expression of the recombinant MDG (GH5-12-12-CBM2-2) and the different truncated proteins (GH5-12-12, GH5-12 or single GH5) was performed in *E. coli* BL21 (DE3)–CodonPluspRIL (Novagen). For large scale protein expression 2 L fresh LB medium containing 50 μg mL^{-1} kanamycin and 50 μg mL^{-1} chloramphenicol was inoculated with 20 mL of an overnight culture. The cells were grown at 37°C with constant shaking at 180 rpm. Protein expression was induced at OD(600 nm) 0.6–0.9 by addition of 1 mM IPTG (isopropylβ-D-thiogalactopyranosid) and the cells were grown at 30°C overnight with constant shaking (180 rpm). The cells (∼12 g wet weight) were harvested by centrifugation (6000 × *g* for 15 min at 4°C) and either stored at −80°C or directly resuspended in 40 mL buffer (50 mM TRIS HCl $pH^{25°C}$ 8.0) for further purification. Cell lysis was performed by passing the cells three times through a French pressure cell at 1200 psi in the presence of protease inhibitors (cOmplete ULTRA Tablets, Mini, EDTA-free, EASYpack, Roche). Cell debris and unbroken cells were removed by centrifugation (12000 × *g* for 30 min at 4°C) and the resulting crude extract was diluted 1:1 and subjected to a heat precipitation for 30 min at 60°C. After heat precipitation, the samples were cleared by centrifugation (12000 × *g* for 30 min at 4°C) and dialyzed overnight against 50 mM TRIS HCl $pH^{25°C}$ 7.0. The partial purified protein fractions were either used immediately for enzyme assays (MDG, GH5-12-12 protein) or further purified (GH5 and GH5-12 proteins). Since an affinity chromatography by Ni-TED column revealed no binding of the His-tagged GH5 and GH5-12 proteins, an alternative purification protocol via fractionated ammonium sulfate precipitation, ion exchange chromatography and size exclusion chromatography was established (for detailed description see Supplementary section Purification of the GH5 and GH5-12 Protein).

[3] http://www.cbs.dtu.dk/services/SignalP/

Visualization of Cellulase Activity on Substrate Agar Plates and Zymogram Gels

For the first qualitative analysis, cellulolytic activity was followed using substrate agar plates and zymogram gels. Substrate agar plates contained 1.5% (w/v) agar-agar and 0.2% (w/v) CMC. The protein samples (5–15 µg, after heat precipitation) were transferred into punched holes and the plates were incubated at 60°C for 4 h or overnight. For zymography 0.2% (w/v) CMC was embedded in standard 7.5% SDS-PAGE gels (10 × 7.5 cm). Enzyme aliquotes were added to sample buffer (100 mM TRIS HCl, $pH^{25°C}$ 6.8, 10% (w/v) SDS, 5% (v/v) 2-mercaptoethanol, 1% (w/v) bromphenol blue) but were not heated before loading. In each lane 2–3 µg of recombinant protein was applied and as protein standard the PageRuler Prestained Protein Ladder (Thermo Scientific) was used. After electrophoreses (BioRAD Mini-Protean System) gels were incubated twice for 20 min in renaturation buffer (100 mL, 50 mM potassium phosphate buffer pH 6.5, 2.5% (v/v) Triton X-100) at room temperature followed by 30 min incubation in developing buffer (100 mL volume; 20 mM MOPS $pH^{25°C}$ 6.5, 100 mM NaCl, 2 mM $CaCl_2$). Finally, in order to visualize endoglucanase activity gels and plates were stained with 0.2% (w/v) Congo red for 30 min (5 mL for plates, 10 mL for gels) and destained with 1 M NaCl three times over 15 min at room temperature. Crude extract of *E. coli* BL21 (DE3)-CodonPlus-pRIL with empty vector was used as control for substrate agar plates and zymogram gels.

Analysis of Hydrolysis Products by Thin Layer Chromatography (TLC)

Enzymatic reactions with 1% (w/v) oligomeric substrates (Megazyme) were performed under the same conditions as used for the DNS assay. Glucose, cellobiose, cellotriose, cellotetraose, cellopentaose, and cellohexaose were incubated with the MDG and the three truncated enzymes (25 µg) at 60°C for 120 min. Controls (stored on ice) as well as hydrolysis products were separated on aluminum sheet (20 × 20 cm) silica gel 60/kieselguhr F_{254} plates (Merck) with ethyl acetate, methanol and H_2O (68:23:9, v/v/v) as a solvent and visualized with $KMnO_4$ solution (1.5 g $KMnO_4$, 10 g K_2CO_3 and 1.25 mL 10% aq. NaOH in 200 mL H_2O).

Activity Measurements of Recombinant Enzymes

Enzyme activities were determined in the presence of soluble polysaccharide substrates, i.e., β-D-glucan from barley, CMC, hydroxyethyl cellulose (HEC), birchwood xylan and other β-linked polysaccharides. Standard assay mixtures contained 100 µL McIlvaine buffer (0.2 M Na_2HPO_4 titrated with 0.1 M citric acid $pH^{25°C}$ 6.0) and 250 µL of the respective 1% (w/v) substrate solution. After pre-equilibration for 5 min at 60°C, the reaction was started by the addition of enzyme. Following a 30 min incubation at 60°C, the reaction mixture was transferred to ice for 5 min, mixed with 750 µL DNS reagent (Bernfeld, 1955), and boiled for 10 min at 100°C. The color shift from yellow to brown depends on the amount of reducing groups and was determined at 575 nm. One unit of enzyme activity was defined as the amount of enzyme required to release 1 µmol glucose equivalents per minute. The pH and temperature optimum for the full length MDG (GH5-12-12-CBM2-2) and the GH5 protein were determined with partially purified enzyme (44 µg) and purified protein (28 µg), respectively, using β-D-glucan [1% (w/v)] from barley as the substrate. AMC was produced by phosphoric acid treatment of Avicel as described previously by Zhang et al. (2006).

RESULTS

Enrichment, Isolation and General Features

Strain 2319x1 was isolated from an *in situ* enrichment with birchwood xylan as the energy and carbon source, and Fe(III) (in the form of insoluble ferrihydrite) as the electron acceptor. The initial inoculum consisted of black sand (enriched with mixed valence Fe mineral magnetite) and hot water from a hot spring, located in the tidal zone near Goryachiy cape of Kunashir Island (South Kurils, Russian Far East region) (Supplementary Figure S2). The enrichment culture was incubated in the same spring, with temperature and pH fluctuating in the range of 76–99°C and 5.0–7.0, respectively. After 6 days of incubation the culture contained two major morphotypes: (i) short rods later on identified as *Pyrobaculum arsenaticum* (manuscript in preparation) and (ii) small irregular cocci. Reduction of ferrihydrite to a magnetic mineral was observed. After the next three subsequent transfers at 90°C on a modified Pfennig medium (pH 6.0–6.2) containing xylan and ferrihydrite but devoid of yeast extract, the coccoid cells became dominating, while the rods disappeared and ferrihydrite reduction ceased. When ferrihydrite (90 mM) was replaced with elemental sulfur (5 g L^{-1}), hydrogen sulfide was formed, but no detectable stimulation of growth was observed. After 10-fold serial (up to 10^{-8}) dilutions on the initial medium with xylan and ferrihydrite, a pure culture was obtained designated as strain 2319x1.

TABLE 1 | Growth substrates of *Thermococcus* sp. strain 2319x1.

Utilized substrates	Non-utilized substrates
Yeast extract	D-mannose
Peptone	D-fructose
Gelatin	Cellobiose
D-xylose	Trehalose
D-glucose	Raffinose
Lactose	β-glucan
Maltose	Avicel
Sucrose	arabinoxylan
Dextrin	arabinan
Dextran	Arabinogalactan
Pullulan	Galactan
Starch	Galactomannan
Xylan	Glucomannan
Xyloglucan	Mannan
AMCH Chitosan	Laminarin
Amorphous cellulose (AMC)	Curdlan
CMC	Inulin
Alginate	Agarose
Lichenan Bamboo leaves	Pectin

16S rRNA gene-based phylogenetic analysis revealed *T. alcaliphlus*, *T. sibiricus,* and *T. litoralis* as the nearest validly published relatives (99.7, 98.4, and 99.5% of 16S rRNA gene identity, respectively; Supplementary Figure S3). All substrates supporting growth of strain 2319x1 in addition to yeast extract and cellulose as well as those which did not serve as growth substrates are depicted in **Table 1**. During the growth on xylan the generation time was ca. 1 h and the final growth yield was 2.5×10^7 cells mL^{-1}.

Genome Sequence and General Genome Features

The final, assembled circular chromosome comprises 1,961,221 bp with an average GC content of 44.6%. The RAST gene caller identified 2,294 protein coding genes. Eight hundred and sixty-four of these protein coding genes were designated as "hypothetical," since no assigned database matches could be found. Both, RAST and Infernal identified the same set of the four rRNAs and 46 tRNAs. The rRNA genes are not located in a single operon: 16S and 23S rRNAs are separated by a single tRNA gene, while two non-identical (two nucleotide substitutions per molecule) 5S rRNAs are distantly located.

Additionally, the Infernal annotation revealed the presence of SRP RNA (signal recognition particle RNA) and RNAse P RNA, several riboswitches and the SscA RNA (secondary structure conserved A RNA), a putative non-coding RNA conserved in hyperthermophiles, which is so far exclusively found in the genera *Pyrococcus* and *Thermococcus*. The average nucleotide

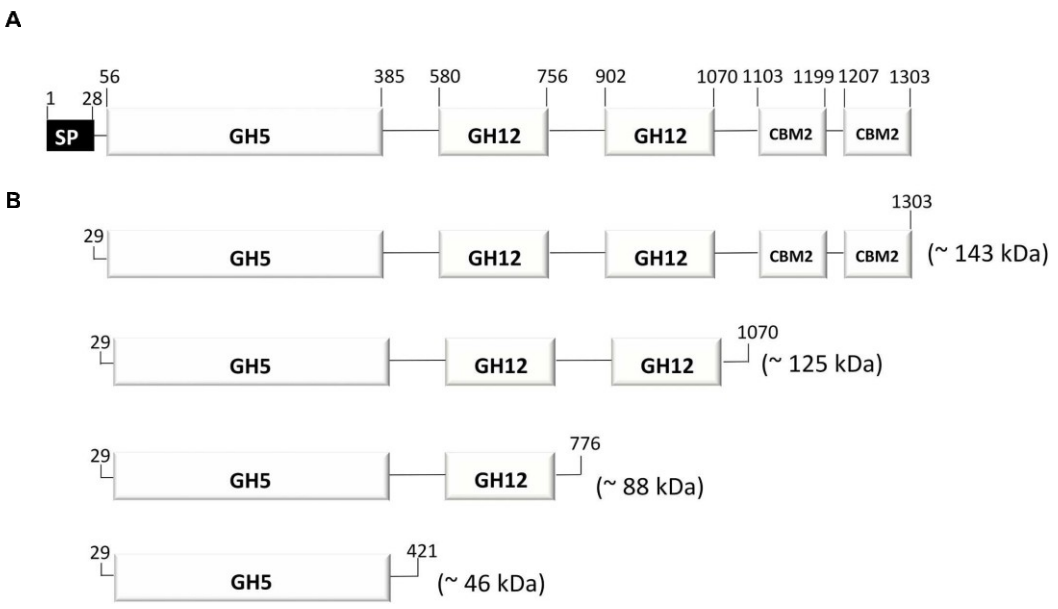


FIGURE 1 | The multidomain glycosidase (MDG) from *Thermococcus* sp. strain 2319x1. The predicted sequence signatures and domain-architecture **(A)** and generated recombinant versions of the protein [full length MDG (GH5-12-12-CBM2-2), GH5-12-12, GH5-12 and single GH5] **(B)** are shown. SP, signal peptide; GH5, glycoside-hydrolase family No. 5; GH12, glycoside-hydrolase family No. 12; CBM2, cellulose binding-module family 2.

identity (ANI) between the genomes of strain 2319x1 and *T. litoralis* (90.4%) or *T. sibiricus* (77.7%) is well below the species level of 95% (Goris et al., 2007). Together with 16S rRNA-based phylogenetic analysis, these results suggest
that either strain 2319x1 belongs to the *T. alcaliphilus* species (so far no genome sequence is available) or represents a novel species of this genus.

The *Thermococcus* sp. strain 2319x1 genome is deposited in GenBank under the accession number CP012200.

Genome-Scale Metabolic Reconstruction of Polysaccharide, D-Glucose and D-Xylose Degradation

A total of 18 genes, encoding GHs and carbohydrate esterases (CEs) of different CAZy families (Lombard et al., 2014) were identified in the genome of strain 2319x1 (Supplementary Table S3). Four of them were predicted to be extracellular. Among those, only one protein had a C-terminal transmembrane region (ADU37_CDS18940, protein sequence region 1075–1097) presumably involved in anchoring the enzyme to the cell membrane. In agreement with the list of substrates supporting growth, the majority of glycosidases were annotated to be involved in hydrolysis of alpha-linked poly- and oligosaccharides (starch, dextrin, dextran, pullulan etc.), while only five enzymes were predicted to catalyze the hydrolysis of β-glycosides. The only exception appeared to be encoded by ADU37_CDS22600 which was predicted to be an endoglucanase/endoxylanase. The protein shows an unique domain organization and is composed of three GH domains (one GH family 5 and two GH family 12), and two family 2 carbohydrate-binding modules (CBM2), with the N- to C-terminal domain order GH5-12-12-CBM2-2 (**Figure 1A**). The analysis of signal peptide cleavage sites using SignalP (Petersen et al., 2011) and transmembrane helices using TMHMM (Möller et al., 2001) revealed an N-terminal signal peptide (amino acid residues 1–28).

Besides GHs and CEs, the genome encoded about 20 glycosyl transferases (GTs) of different families (mainly GT2 and GT4, according to CAZy). The majority is presumably involved in glycosylation of *S*-layer proteins, exopolysaccharide (biofilm) formation or biosynthesis of intracellular oligo- and/or polysaccharides. Few of them belong to families with a retaining mechanism, which theoretically allows either hydrolysis or phosphorolysis of glycosidic bonds. One of these proteins, encoded by ADU37_CDS05640 belongs to the GT35 family containing glycogen phosphorylases and enzymes with related activities.

Since strain 2319x1 was growing on various poly- and monosaccharides, it should possess an efficient uptake system for diverse mono-, di- and oligosaccharides. Indeed, more than 100 genes coding for transporters of the ATP-binding Cassette (ABC) Superfamily (3.A.1) were identified in the genome, along with over 20 genes coding for the Major Facilitator Superfamily (MFS, 2.A.1) transporters and two genes coding for the Solute:Sodium Symporter (SSS) family (2.A.21, according to TCDB database) (Saier et al., 2014).

The majority of substrates supporting the growth of the strain contained either D-glucose (D-glucose itself, maltose, starch and derivatives, cellulose and derivatives, xyloglucan) or D-xylose (D-xylose itself, xylan and xyloglucan). Accordingly, all genes encoding enzymes operative in a *Thermococcales* like modified Embden-Meyerhof-Parnas (EMP) pathway were identified (Siebers and Schönheit, 2005; Bräsen et al., 2014) (e.g., archaeal ADP glucokinase ADU37_CDS0192; ADP phosphofructokinase ADU37_CDS0945; glyceraldehyde-3-phosphate ferredoxin: oxidoreductase (GAPOR) ADU37_CDS1714; non-phosphorylating $NAD(P)^+$- dependent glyceraldehyde-3-phosphate dehydrogenase (GAPN) ADU37_CDS1986). In contrast, the pathway for D-xylose degradation remained unclear. The only characterized archaeal D-xylose degradation pathway, described for *Haloferax volcanii* (Johnsen et al., 2009) and *S. solfataricus* (Nunn et al., 2010) seems to be absent since not a single protein of this pathway was found in the *in silico* translated proteome of strain 2319x1. Moreover, BLAST searches of key enzymes of this pathway did not reveal any homologs in other archaea, except haloarchaea and few members of the *Thermoprotei*. In bacteria, D-xylose can be utilized similarly to the *H. volcanii* pathway, as demonstrated, e.g., for *Caulobacter crescentus* (Stephens et al., 2007). Alternatively, and more commonly, D-xylose degradation is initiated by xylose isomerase and kinase yielding xylulose-5-phosphate, which further enters the metabolic network via the pentosephosphate pathway (Bräsen et al., 2014). However, no obvious homologs (e.g., xylose isomerase, xylulose kinase) involved in the established bacterial route for D-xylose degradation were identified. Therefore, the pathway for D-xylose degradation still remains unclear, the genome contains numerous genes encoding $NAD(P)^+$-dependent oxidoreductases, sugar kinases as well as aldose-ketose isomerases with unknown or uncertain function, possibly involved in D-xylose degradation.

The modified Entner–Doudoroff pathway (Bräsen et al., 2014) as well as the complete TCA cycle is absent, as it was shown for other members of *Thermococcales* (Atomi et al., 2005). Homologs for pyruvate:ferredoxin oxidoreductaseand several ADP-forming ac(et)yl-CoA synthetases were identified in the genome of strain 2319x1, indicating a similar energy metabolism as described for *T. kodakarensis* and *P. furiosus* (Siebers and Schönheit, 2005; Bräsen et al., 2014).

Strain 2319x1, as most of other archaea, possesses no classical pentose-phosphate pathway (PPP). The genes, encoding proteins of the oxidative part of the PPP were absent, and only an incomplete set of proteins of the non-oxidative part was identified, i.e., a transketolase, which is split in an N- and C-terminal part (ADU37_CDS1012 and ADU37_CDS1013) and a ribose-5-phosphate isomerase (ADU37_CDS0756) (Bräsen et al.,

2014). D-ribulose-5-phosphate is formed from D-fructose 6-phosphate via the reversed Ribulose Monophosphate Pathway – a two-step reaction, catalyzed by D-arabino-3-hexulose-6-phosphate formaldehyde-lyase and phosphohexuloisomerase (Kato et al., 2006; Bräsen et al., 2014). In *T. kodakaraensis* the genes, encoding these enzymes, are fused (Orita et al., 2006), and a highly similar homolog was found in the genome of strain 2319x1 (ADU37_CDS03570).

Identification of Extracellular Polysaccharolytic Activities

During the growth of strain 2319x1 on AMC, xylan and xyloglucan extracellular, cell-bound endoglucanases with broad specificity were produced (**Figure 2**). The cell yields were quite similar (1–1.5 × 10^8 cells mL^{-1}) for all three substrates. The concentrations of total cell-associated proteins were also of the same order of magnitude (49.6, 84.6, and 66.7 µg mL^{-1} for AMC, xylan or extracellular cell surface-bound protein fractions were not clearly linked to the different growth substrates utilized (**Figure 2**). This might be explained by the co-occurrence of cellulose and hemicelluloses in plant cell walls, resulting in the expression of the same enzyme or enzyme sets during the growth on both types of polysaccharides. It is important to mention that cells grown on gelatin did not produce detectable amounts of extracellular polysaccharide degrading enzymes (data not shown) further supporting their inducible state.

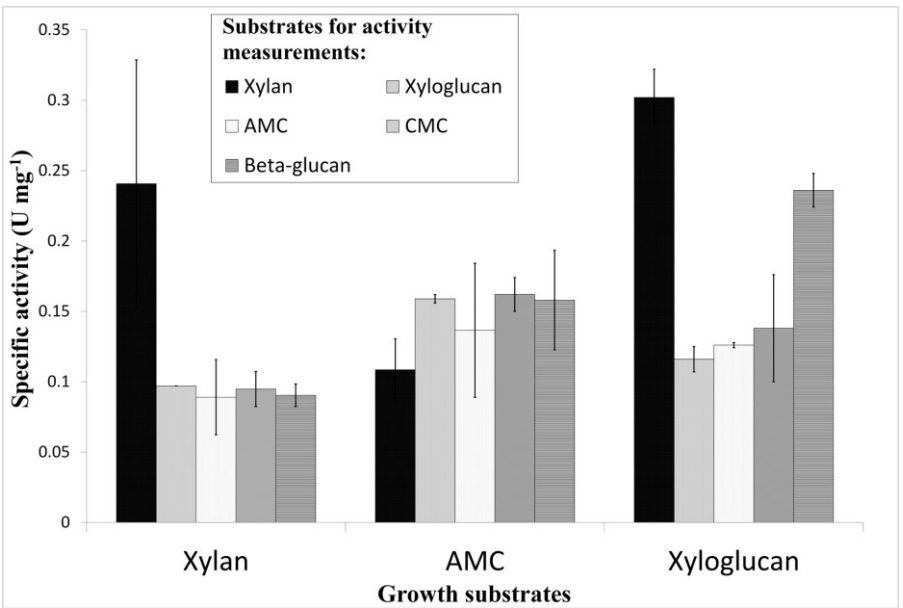


FIGURE 2 | Hydrolytic activities of surface protein fractions from strain 2319x1 cells, grown on xylan, AMC and xyloglucan. The formation of reducing sugars using the DNS method per mg of protein (bicinhoninic assay) is shown. The different growth substrates (2 g^{-1} L (*x*-axis) as well as the substrates for activity measurements (0.01 g L^{-1}) (different gray shades of columns shown in the right part of the figure) are depicted. Incubations were performed at 85 °C for 20 h.

Cloning, Expression and Purification of the Full Length MDG and Truncated MDG Proteins

Among the identified hydrolases, ADU37_CDS22600 was chosen for further analysis due to its unusual domain composition, GH512-12-CBM2-2. The *mdg* gene (3912 bp) encodes for a protein with a predicted molecular mass of 143 kDa. In order to unravel the enzymatic activity of the MDG and to elucidate the function associated with the different domains, the full length protein (GH5-12-12-CBM2-2, without signal peptide) along with the three C-terminal truncated versions of the MDG were cloned (via In-Fusion cloning) and expressed in *E. coli* using the pET expression system (pET24a, C-terminal His_{6x}-tag). The truncated versions were: GH5-12-12 (missing the two CBM2s), GH5-12 (missing two CBM2s and the second GH12) and single GH5(missing all other domains) (**Figure 1**).

All four proteins were recombinantly expressed in *E. coli* BL21 (DE3)-CodonPlus-pRIL. The truncated versions GH5 and GH5-12 were formed as fully soluble proteins, whereas only a small portion of the full length MDG (GH5-12-12-CBM2-2) and GH5-12-12 was obtained in the soluble protein fraction as demonstrated by SDS-PAGE and immunoblotting with His-tag specific antibodies (Supplementary Figure S4). Additional efforts to improve the soluble expression of the MDG and GH5-12-12 using different expression strains (e.g., *E. coli* Rosetta and *E. coli* SoluBL21) or expression conditions, e.g., lower temperature and

different culture media, were not successful. Also, purification from the membrane fraction using denaturation (e.g., urea, triton X-100) and renaturation approaches did not significantly improve the yield and purity of either protein. Due to the low yield of soluble protein, the MDG and GH5-12-12 were only partially purified and the enzymatic assays were performed with the soluble fraction after heat precipitation. From 8 g (wet weight) of recombinant *E. coli* cells, 12.4 mg and 9.6 mg of the MDG and the GH5-GH12-GH12 protein were obtained. For the soluble proteins GH5 and GH5-12, a purification protocol via heat precipitation (60°C, 30 min), ammonium sulfate fractionation [GH5, 2.2 M $(NH_4)_2SO_4$; GH5-12, 2.6 M $(NH_4)_2SO_4$], followed by ion exchange and size exclusion chromatography was established (Supplementary Figure S5). Notably, although all four proteins were cloned with a C-terminal His-tag, none of the proteins bound to Ni-TED columns. The purification procedure yielded 10.4 and 6.9 mg of essentially pure GH5 and GH5–GH12 protein, from 15 g (wet weight) of recombinant *E. coli* cells, respectively. For storage of protein, samples were supplemented with glycerol 20% (v/v) and frozen at −70°C.

Cellulolytic Activities of the Full Length and Truncated MDG Versions

For all four proteins, activity with 0.2% (w/v) CMC was detected on substrate agar plates and in zymogram gels (**Figure 3**). In the zymogram gels, the enzyme activity correlated to protein bands of the expected size for the full length and truncated MDG proteins (**Figure 3B**).

Substrate Specificity of the Full Length and Truncated MDG Proteins

The substrate specificity of the partially purified full length MDG and GH5-12-12 protein as well as the pure GH5-12 and GH5 versions was investigated using the DNS assay (Bernfeld1955) under standard conditions (60°C, pH 6.0) (**Figure 4**, Supplementary Table S4).

For all four proteins, the highest activity was observed on mixed β-1,3/1,4-glucans, i.e., barley β-glucan (4-16 U mg^{-1}) and lichenan from icelandic moss (7 U mg^{-1}). Endo-β-1,4glucanase activity was also high on β-1,4-linked glucans such as CMC (3–7 U mg^{-1}), hydroxyethyl cellulose (2–5 U mg^{-1}), phosphoric acid treated (PAT) Avicel (2–5 U mg^{-1}) and Avicel (1 U mg^{-1}). In addition, all four proteins exhibited significant activity on cellooligosaccharides (cellohexaose (2–6 U mg^{-1}) but only minor activity with cellobiose (0.5–1 U mg^{-1}). TLC analysis of cellobiose and oligosaccharides (3–6 glucose units, G3 – G6) hydrolysis products revealed cellobiose as a main product for all four proteins. In addition, some glucose was detected for all proteins and cellotriose for full length MDG and GH5-12-12 (Supplementary Figure S6). Very low activity (<1 U mg^{-1}) was observed for xylan (birchwood, β-1,4 linked D-xylose units), xyloglucan (β-1,4 linked D-glucose backbone with α-1,6 linked D-xylose side chains) and curdlan (β-1,3-glucan, 0.4 – 0.5 U mg^{-1}). Notably high activity (2–9 U mg^{-1}) was also observed on galactomannan, i.e., locust bean gum (LBG, β-1,4 linked Dmannose backbone with α-1,6 linked D-galactose side chains) suggesting mannan endo-1,4-β-mannosidase or β-mannosidase activity.

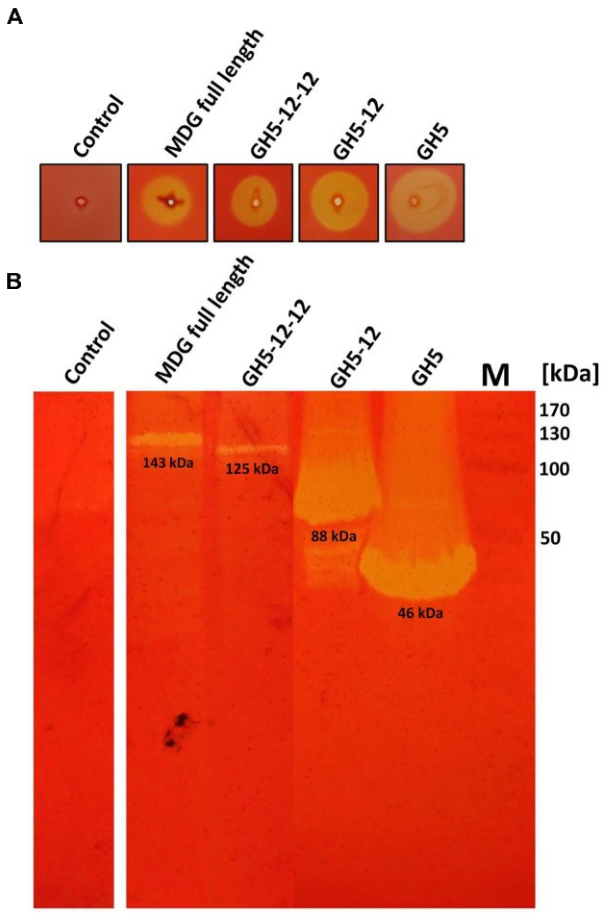


FIGURE 3 | Cellulolytic activities of full length and truncated MDG proteins. Hydrolytic activities were analyzed on CMC-agar-plates **(A)** and in CMC-zymogram gels **(B)**. For CMC screening plates 10 μL (5–10 μg of total protein after heat precipitation at 60° C) and for zymogram gels 20μL (3μg of protein after heat precipitation at 60° C (GH5 and GH5-12) or 2 μg of solubilized membrane fraction (GH5-12-12 and MDG full length)) were applied. As negative control, crude extract of *E. coli* BL21 (DE3) pRIL with vector pET24a without insert was used. Marker, M **(B)**:prestained protein standard (Fermentas), GH5 (~46kDa) and GH 5-12 (~ 88kDa), GH5-12-12 (~125 kDa) and MDG full length (~143 kDa).

Characterization of the Full Length MDG and the GH5 Protein

For industrial applications the temperature and pH optima of enzymes as well as their stability toward detergents are of high interest. The respective enzymatic properties were analyzed for the partially purified full length MDG as well as the purified GH5 protein with barley β-glucan as substrate using the DNS assay. The enzymatic activities were determined in a temperature range from 40°C to 90°C and a pH range from 4.5 to 9.5 (**Figure 5**).

In order to establish a purification protocol, the MDG was solubilized from the *E. coli* membrane fraction by treatment with 8 M urea. Interestingly, the enzyme was active even after 2 days of incubation in the presence of 8 M urea (data not shown). These results indicate a high stability of the MDG toward denaturing agents.

DISCUSSION

Archaea of the genus *Thermococcus* are widespread in shallow water and deep-sea marine environments, as well as in deepsubsurface habitats. At present, this genus comprises more than 30 species with a G + C content ranging from 40.2 to 56.1%.

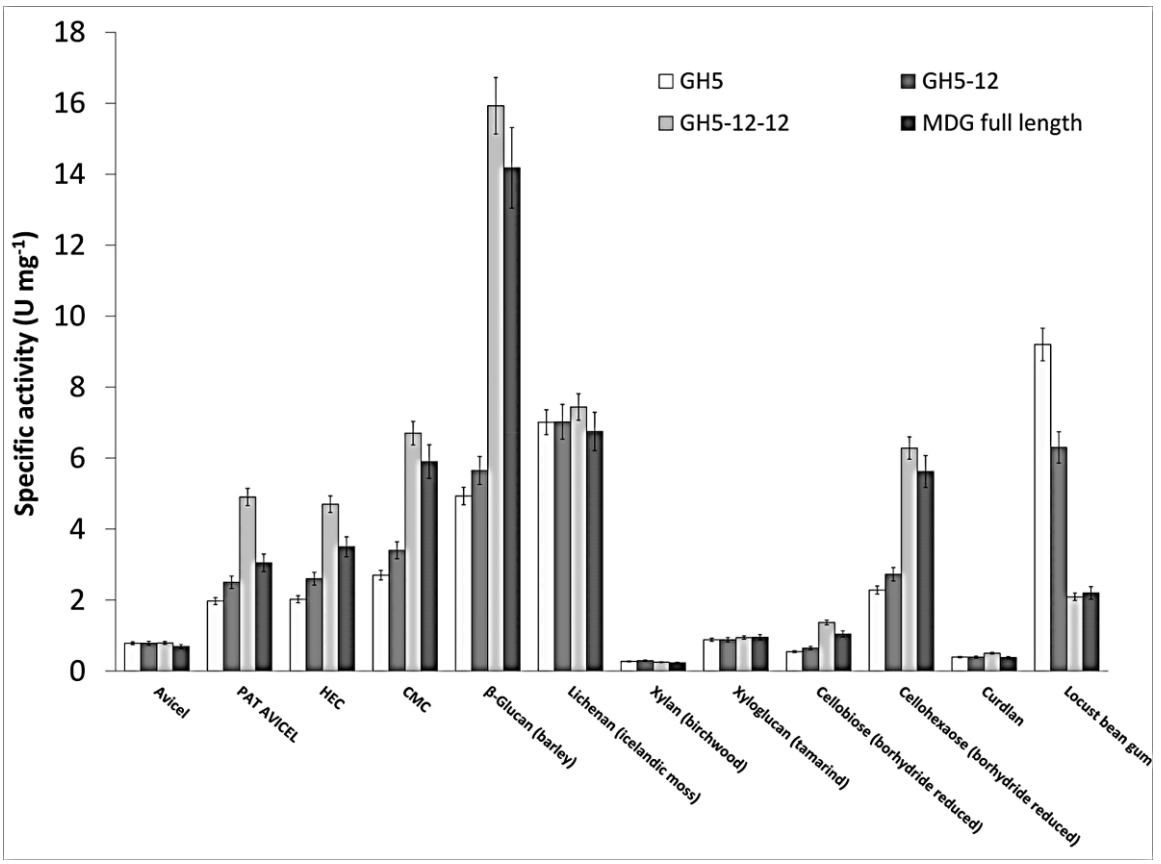


FIGURE 4 | Substrate specificity and specific activity of the full length MDG and truncated MDG proteins. Enzme activities were determined using the DNS assay. Substrate solutions [1% (w/v)] were incubated for 30 min with 33 μg of enriched protein fractions at 60°C. For GH5 and GH5-12 purified enzymes (after storage at -70°C)and for the full length MDG and GH5-12-12 enriched protein fractions after heat precipitation were used. microcrystalline cellulose (MCC) Avicel phosphoric acid treated (PAT) Avicel, hydroxyethyl cellulose (HEC), carboxymethyl cellulose (CMC).

The full length MDG showed the highest activity (4 U mg^{-1}) at a temperature of 60°C and possesses a pH-optimum from 7.5 to 9.5 with a maximum activity of 8 U mg^{-1} pH 8.5. In contrast the single GH5- protein showed the highest activity at 90°C and exhibited a broad pH optimum (pH 4.5–9.5) with highest activity at pH 5.5. Therefore, the truncation seems to broaden the temperature range to higher temperatures as well as the pH range to more acidic pH-values.

Phenotypically *Thermococcus* species are very similar: they all are obligate anaerobes degrading peptides or, less frequently, polysaccharides, and use elemental sulfur as a terminal electron acceptor (Schut et al., 2014). Among *Thermococcales*, the ability to grow with C1 compounds (CO and formate) coupled with hydrogen production seems to be restricted to deepsea representatives possessing corresponding enzyme clusters

(Sokolova et al., 2009). Strain 2319x1 was isolated from the hot vent, located in the tidal zone of Kunashir Island, Southern Kurils. Organic matter of plants, alga and animals is brought to the vent with tidal water providing a constant supply of substrates for hyperthermophilic microorganisms inhabiting sand and water of the vent. In addition, to proteinaceous substrates, which are common energy sources of all *Thermococcales,* strain 2319x1 degrades a wide spectrum of sugars and polysaccharides, including highly recalcitrant higher plant polysaccharides, like cellulose and xylan. These substrates are more abundant in terrestrial hot springs predominantly inhabited by organotrophic *Crenarchaeota*. Accordingly, the ability to degrade cellulose and xylan has been identified so far mainly in this archaeal phylum (Supplementary Figure S1).

Genome-Scale Metabolic Reconstruction of Central Carbohydrate Metabolism

Thermococcus sp. strain 2319x1 shares several metabolic features with well-characterized members of the order *Thermococcales,* i.e., *P. furiosus* and *T. kodakarensis* (Bräsen et al., 2014), but is unique in regard to its broad sugar substrate spectrum, especially in its capability to grow on xylan, xyloglucan, cellulose as well as on monosaccharides, such as D-glucose and D-xylose. The spectrum of sugars utilized by *P. furiosus* and *T. kodakarensis* is restricted to maltose (*P. furiosus* only) and to higher maltooligosaccharides, whereas glucose cannot be used. According to the broad substrate specificity and physiological versatility of *Thermococcus* sp. strain 2319x1,

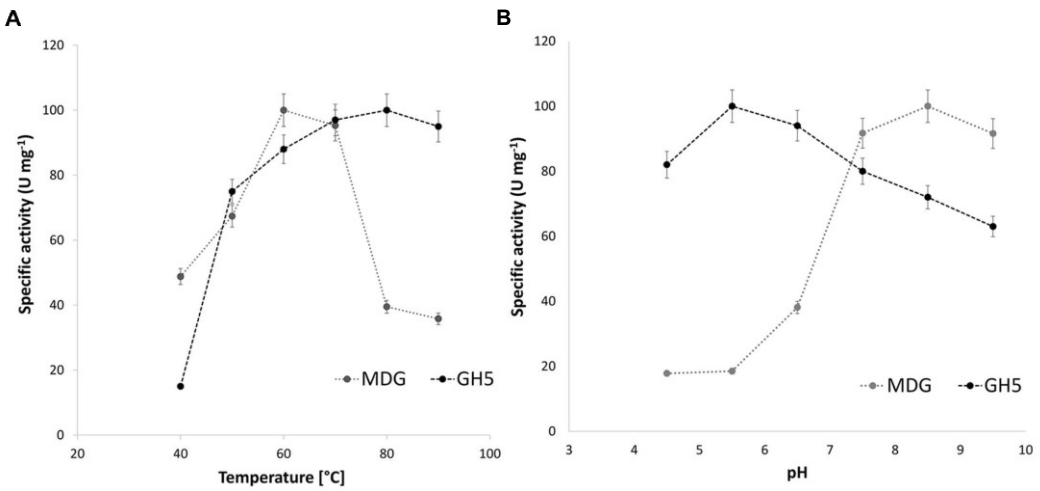


FIGURE 5 | Effect of temperature (A) and pH (B) on the specific activity of the full length MDG and GH5 protein. Temperature and pH dependend activity was determined with barley β-glucan [1% (w/v)] as substrate using the DNS assay. Substrate solutions were incubated for 30 min with 44 μg of enriched protein fractions after heat precipitation (MDG) or 28 μg of purified protein (GH5) at the indicated pH value (at 60°C) or temperature (at pH 7) in McIlvaine buffer (0.2 M Na_2HPO_4 titrated with 0.1 M citric acid).

Thermococcus sp. strain 2319x1 isolated from the organics-rich tidal zone is, to our knowledge, the first representative of the phylum *Euryarchaeota* and the first marine hyperthermophilic isolate capable to grow efficiently on cellulose or xylan. The analysis of cell-bound exoenzyme activities, of cells grown on different polysaccharides, compared to cells grown on gelatin revealed induction of exoenzyme expression in response to polysaccharides although the observed polysaccharolytic activity is not always clearly linked to the growth substrate (**Figure 2**). In its natural habitat, this exoenzyme repertoire will allow the strain 2319x1 to utilize allochthonous organic matter originating from plants or algae. The growth yield of AMC grown cells of strain 2319x1 is the highest among known hyperthermophilic cellulolytic archaea (Supplementary Figure S7).

numerous genes encoding GHs, GEs and GTs of different enzyme families were identified in its genome (Supplementary Table S3). Due to its broad substrate specificity (**Table 1**), the presence of sugar transporters is extremely important for strain 2319x1. In *T. kodakarensis*, only one oligosaccharide ABC transporter has been shown to be essential for carbohydrate uptake (Matsumi et al., 2007), whereas in *P. furious* three ABC transporters for trehalose/maltose, maltodextrin and cellobiose have been characterized (Koning et al., 2001; Lee et al., 2003). In the genome of *Thermococcus* sp. strain 2319x1 a number of genes encoding ABC transporters were identified. For three of them, the transport of either D-glucose/D-xylose (ADU37_CDS01060-01090), cellobiose / cellooligosaccharides (ADU37_CDS07290-07330) or maltodextrins (ADU37_CDS18 970-18930), respectively, is annotated based on sequence similarities to the characterized homologs (Xavier et al., 1996; Koning et al., 2001; Erbeznik et al., 2004). Notably, one of two SSS transporter genes found in the strain 2319x1 genome (ADU37_CDS10740), exhibits

high similarity to the characterized human myoinositol:Na^+ symporter SMIT2, which efficiently transports both D-glucose and D-xylose, and might therefore represent a glucose:Na^+ symporter (2.A.21.3.2) (Coady et al., 2002).

Like in other anaerobic *Thermococcales*, D-glucose is degraded via a modified EMP pathway, and acetate, CO_2 and H_2 are the fermentation products (Siebers and Schönheit, 2005; Bräsen et al., 2014). Pentoses are formed from fructose-6-phosphate via the reversed ribulose monophosphate pathway. The formation is catalyzed by a fused D-arabino-3-hexulose-6phosphate formaldehyde-lyase and phosphohexuloisomerase as described for *T. kodakaraensis*, and an incomplete nonoxidative pentose phosphate pathway (Kato et al., 2006; Orita et al., 2006; Bräsen et al., 2014). However, the central metabolic pathway for D-xylose degradation still remains to be elucidated. Homologs were neither found for the reported archaeal nor for the bacterial degradation pathways.

Enzymatic Properties of the MDG of *Thermococcus* sp. Strain 2319x1

The analysis of hydrolases capable of degrading β-linked polysaccharides revealed two candidates in the genome sequence of strain 2319x1, i.e., the GT35 family representative ADU37_CDS05640 and ADU37_CDS22600. The latter is a MDG with so far unknown, unique domain architecture (GH512-12-CBM2-2), which has not been observed in any of the publicly available database entries. Homologs of the few known archaeal cellulases or xylanases could not be identified in strain 2319x1. Also, related polysaccharide hydrolyzing enzymes of other enzyme families were absent.

The GT35 enzyme, purified from *Themococcus zilligii*, was previously characterized as a xylanase (Uhl and Daniel, 1999). However, addition of xylan to the medium neither supported growth of *T. zilligii* nor enhanced its xylanase activity. In addition, the enzyme sequence was highly similar to maltodextrin phosphorylases, e.g., from *T. litoralis*, raising questions about its enzymatic activity (Rolland et al., 2002). The GT35 sequence of strain 2319x1 (ADU37_CDS05640) is highly similar to the characterized maltodextrin phosphorylase from *T. litoralis* (96.7% identity) (Xavier et al., 1996). Less pronounced similarity was observed to the putative xylanase from *T. zilligii* (75.8% amino acid sequence identity), suggesting maltodextrin phosphorylase activity for this protein.

Thus, the only remaining predicted enzyme, capable of hydrolyzing β-linked polysaccharides was ADU37_CDS22600. In order to unravel the enzymatic function of the MDG as well as its different protein domains, the encoding gene as well as three truncated versions encoding proteins with reduced domain complexity were cloned and expressed in *E. coli*. The full length protein MDG as well as the GH5-12-12 protein were only partially soluble and despite many efforts, only a partial purification by heat precipitation was achieved (Supplementary Figure S4). For the soluble GH5-12 and GH5 proteins a successful purification protocol was established (Supplementary Figure S5). The activities of all recombinant proteins were confirmed by substrate agar plates, as well as by zymography with CMC as substrate (**Figure 3**).

All four proteins were highly active against the mixed β1,3/1,4-glucans, barley β-glucan and lichenan. So far, only few archaeal endoglucanases have been reported to hydrolyze such substrates (Supplementary Table S4; **Figure 4**). The activity against curdlan was low, indicating a high specificity of all catalytic domains of ADU37_CDS22600 to β-1,4-glucosidic bonds. The activity against microcrystalline cellulose (MCC) Avicel, phosphoric acid treated (PAT) Avicel, HEC and especially CMC suggests that all four proteins act as endoglucanases (Béguin and Aubert, 1994). Furthermore, the specific activity increased with increasing domain complexity for all substrates containing more than one consecutive β-1,4-glycosidic bond. For substrates containing none or no more than one consecutive β-1,4-glycosidic bond (i.e., curdlan, xylan, galactomannan or lichenan, respectively), no effect was observed. Also, α1,6-linked side chained substrates (xyloglucan) or insoluble substrates (Avicel) were not converted more efficiently by the higher complexity variants MDG and GH5-12-12 of ADU37_CDS22600. This suggests that the addition of the second GH12 domain has a positive influence on enzyme activity and that this second GH12 domain probably acts as exoglucanase and β-glucosidase, enabling a more efficient hydrolysis of substrates. Hence, the full-length MDG seems to act in a processive mode. Since for both proteins MDG and GH5-12-12 only partially purified enzyme fractions were used, compared to the purified GH5 and GH5-12 proteins, the specific activity is even underestimated. However, we have no information if possible structural changes in the multidomain enzyme or additional enzyme activities (e.g., exoglucanase activity) provided by the second GH12 domain are responsible for the enhanced enzyme activity. Therefore, for a final evaluation, future studies with the two single GH12 domain enzymes have to be awaited.

In the presence of cellooligosaccharides (i.e., cellotriose to cellohexaose), the formation of cellobiose as the main product and of some glucose was shown for all four proteins. Cellotriose was identified as additional product only for the full length MDG and GH5-12-12. Notably, additional activity with galactomannan (LBG) was observed. In contrast to the increase in glucanase activity with increasing domain complexity, the additional β-(endo)-mannosidase activity seems to be linked to GH5 domain. Hence, the combination of different enzyme domains in MDG extends the substrate specificity of the enzyme and seems to allow the processive conversion of various polysaccharides differing in size and structure.

The substrate specificity of the MDG protein was rather unexpected, since *Thermococcus* sp. strain 2319x1 showed no growth on barley β-glucan as well as on galactomannan, and the enrichment was performed on xylan, which is, if at all, a minor substrate of the enzyme. However, significant growth was observed on CMC and AMC suggesting that the enzyme has a major role in cellulose degradation. This assumption is in

agreement with determined exoenzyme activities with β-glucan, AMC, CMC, xylan and xyloglucan as substrates of xylan-, AMCand xyoglucan-grown cells.

The detailed characterization of the full length MDG revealed a slightly alkaline pH optimum (7.5–9.5) and a rather low temperature optimum (60–70°C), compared to the optimal growth temperature and pH of the organism (75–85°C, pH 7.0). In contrast, the single GH5 protein exhibits a slightly acidic, broad pH optimum (4.5–9.5) and high temperature (90°C) optimum, consistent with the respective growth requirements of the organism. However, the expression of the MDG in *E. coli* is challenging and only a very small fraction of the MDG and GH5-12-12 protein was soluble, probably due to the size of the protein and the complex domain structure. Also, for several CBMs, the need of glycosylation has been reported (Boraston et al., 2003; Chen et al., 2014) and the archaeal system generating these kinds of modification is absent in the bacterial host. Preliminary studies on the binding to mircrocrystalline cellulose (MCC) Avicel of the full length recombinant MDG compared to its single GH5 domain revealed no obvious difference in binding affinity (data not shown). Therefore, the different optima of MDG and GH5, the different solubility, and the similar binding patterns to MCC (unexpected because of the presence and absence of the CBM domains in MDG and GH5, respectively) might be caused by the heterologous expression in *E. coli* which does not allow for the required posttranslational modifications.

Phylogenetic and Functional Implications for the Single MDG Domains

The presence of the MDG with an unusual domain complexity raises questions about the evolutionary origin and the relationship of the different protein domains to available functionally characterized proteins. BLASTp searches revealed the highest identity of the GH5 domain (position 56–385) to the characterized endo-1,4-β-glucanase from the close relative *P. horikoshii* (PH1171, EGPh, 86% identity). However, in contrast to the artificial GH5 domain from the strain 2319x1 MDG, the *P. horikoshii* enzyme possesses the highest activity with CMC (100%), but much lower activity with lichenan (43.9%) and very low activity with β-glucose oligomers (cellobiose to cellopentaose, <2%) (Ando et al., 2002; Kashima et al., 2005). PH1171 showed a pH optimum between 5.4 and 6.0, a temperature optimum at 97°C, and a specific activity of 8 U mg^{-1} [0.5% (w/v) CMC, pH 5.6, 85°C] (Ando et al., 2002). For PH1171 processive hydrolytic activity with direct formation of cellobiose by random cleavage was reported (Kim and Ishikawa, 2010). Also, structural prediction tools (HHpred[4]; Söding et al., 2005) revealed the highest structural similarity of the GH5 domain to the crystal structure of PH1171, which exhibits a typical triosephosphate isomerase (TIM) $(\alpha\beta)_8$ barrel fold (Kim and Ishikawa, 2010, 2011). The eight residues found in the catalytic cleft of PH1171 as well as in other family 5 enzymes (*Pho*: Arg102, His155, Asn200, Glu201, His297, Tyr299, Glu342 and Trp377) as well as the four cysteines (*Pho*: Cys106-Cys159, Cys372-Cys412) involved in disulfide bond formation, are conserved in the GH5 domain (Kim and Ishikawa, 2010, 2011) (Supplementary Figure S8).

The second GH12-2 domain, comprising the amino acid residues 902–1070 of the full length MDG, is highly similar to the well characterized endo-1,4-β-glucanase encoded by ORF PF0854 from *P. furiosus* (EGPf, 85% idenitity). Although this ORF appears also as closest homolog of the first GH12-1 domain (position 580–756 of full length MDG) in

BLAST searches, the amino acid identity is much lower (34% identity). Like the MDG, PF0854 cleaves β-1,4-bonds in mixed linkage glucans (β-1,3/1,4-bonds) and in cellulose (β-1,4-bonds). Barley-glucan, lichenan, and to a lesser extend

CMC and Avicel PH101 have been reported as substrates (Bauer et al., 1999). However, the highest specific activity has been reported toward β-1,4-linked cellulose oligosaccharides, i.e., cellotetraose, cellopentaose and cellohexaose. Additionally, a function in attacking the β-1,4-amorphous insoluble regions within the insoluble cellulose was suggested (Bauer et al., 1999). Moreover, two characterized GH12 enzymes from the hyperthermophilic bacterium *Thermotoga maritima* were also among the most similar characterized homologs. *Thermotoga maritima* enzymes were characterized as intracellular endoβ-1,4-glucanase TM1524 and extracellular exo-β-1,4-glucanase TM1525. Both enzymes exhibit similar substrate specificity on a variety of β-linked biopolymers (CMC, β-glucan, and *p*-nitrophenyl-β-D-cellobioside) (Bronnenmeier et al., 1995; Liebl et al., 1996; Vanfossen et al., 2008), but differ in their pH optima, long-term thermostability, as well as in stability in the presence of salts (NaCl). Both genes are in juxtaposition on the genome and a recent gene duplication event has been proposed (Liebl et al., 1996). The crystal structure of TM1524/Cel12 from *T. maritima* was solved (Cheng et al., 2011).

Another characterized homolog is the GH12 xylanase/ cellulase from S. solfataricus (SSO1354), which exhibits similar activity on various xylans, CMC and arabinan (Maurelli et al., 2008). Although this protein was not active against Avicel, mannan and xyloglucan, its activity spectrum indicates the possibility of a single GH12 domain being active against various polysaccharides including xylan that differ in size and structure.

HHpred analyses (Söding et al., 2005) revealed that both, the GH12-1 and the GH12-2 domain, of the 2319x1 full length MDG show a high structural similarity to the solved PF0854 structure, which adopts a compact β-jellyroll fold with a calciumbinding motif (DxDxDG) (Kim et al., 2012). The catalytic residues, the nucleophile PF0854-E197 and the proton donor PF0854-E290, conserved in family 12 GHs, are found in both MDG-GH12 domains. In contrast the calcium-binding motif, important for activity, thermostability and protein folding, as well as the residues involved in metal ion coordination (PF0854: Asp68,

[4] http://toolkit.tuebingen.mpg.de/hhpred

Asp70, Asp72, Asn74, Glu76, Asp142), are only present in GH12-2 domain. Notably the metal ion coordinating residues, except Asp142, were found in the linker region in front of the GH12-2 domain, which shows only low sequence similarity (Supplementary Figure S9).

Carbohydrate-binding modules enable enzymes to attach to polymer surfaces and, thus, to increase the local substrate concentration and to enhance the catalytic efficiency (Simpson et al., 2000). CBMs are not common for archaeal glycosidases. Among all 1953 CBM2 accessions currently available in the CAZy database, only six are from Archaea. Both the *P. horikoshii* GH5 and the *P. furiosus* GH12, do not possess cellulose binding domains. The CBM2-1, comprising residues 1103– 1199 of the full length MDG, shows the highest similarity (BLASTP) to the CBM of the endo-1,4-β-endoglucanase Cel12E from an unidentified prokaryotic organism (37% identity) and to the CBM of the chitinase PF1233 from *P. furiosus* (35% identity) (Supplementary Table S5). Cel12E has recently been reported to be of archaeal origin and also exhibits an unusual though less complex N- to C-terminal GH12-CBM2CBM2 domain architecture. All three domains of Cel12E show significant similarity to the thermococcal homologous (Leis et al., 2015). However, the CBM2-2, i.e., residues 1208– 1303 of the full length MDG, shows statistically reliable hits only to bacterial proteins with the highest similarity to the CBM of the chitinase from *Mycobacterium kansasii* (35% identity).

CBM2 represents one of the largest CBM families in CAZy and comprises members that can bind to cellulose, xylan and chitin. CBM2 domains are β-sandwich fold domains, typically comprised of two four stranded sheets that contain a planar face that interacts with a ligand via a hydrophobic strip of aromatic residues (Xu et al., 1995). Structural searches (HHpred) predict 25% (CBM2-1) and 26% (CBM2-2) identity to the cellulosebinding domain of the exo-β-1,4-glucanase from *Cellulomonas fimi* (PDB 1exg_A), and 29% (CBM2-1) and 19% (CBM2-2) identity to the chitinase-binding domain of the *P. furiosus* chitinase (Xu et al., 1995; Simpson et al., 2000; Nakamura et al., 2008) (Supplementary Figure S10).

Additional phylogenetic implications come from genomic context analysis in *Thermococcus* sp. strain 2319x1. The *mdg* gene is localized in the genome together with three other genes, encoding one putative GT ADU37_CDS22610 and two hypothetical proteins, which are flanked by two transposases encoding genes (ADU37_CDS22570-22590 and ADU37_CDS22640-22650) (Supplementary Figure S11), supporting lateral gene transfer to the strain 2319x1 or its closest ancestor. Moreover, as outlined above, individual domains of the full length MDG (ADU37_CDS22600) have different sets of nearest relatives (BLAST hits), indicating their different origin. While the GH5 and the GH12-2 domains have highly similar homologs, the GH12-1 and the CBM2-2 domains have less similar homologs, all from the bacterial domain. The CBM2-1 domain has less than ten statistically significant hits, mainly from thermococcal/pyrococcal chitinases (Supplementary Table S5). Phylogenetic analysis of the individual MDG glycosidase domains thus supports their different derivation (Supplementary Figures S12 and S13). The GH5 domain was, most probably, horizontally transferred from the ancestors of *Xanthomonadales,* and the GH12-2 domain from *Thermotogales*. The GH12-1 domain of strain 2319x1 MDG was possibly also acquired from *Thermotogales,* but presumably much earlier as indicated by the reduced similarity. Our findings are in accordance with the general view on evolution of GHs-encoding genes in *Archaea*, that regards horizontal gene transfer as the determining factor behind archaeal GH repertoires (Henrissat et al., 2009).

In general, the identification of new thermostable cellulases with appropriate performance and/or novel functionalities offers great potential for biotechnological applications. Thus for the endo-1,4-β-glucanase from *P. horikoshii* (PH1171, GH5 family) in combination with β-glucosidase from *P. furiosus* (PF0073, Bauer et al., 1996), a complete saccharification of cellulose [phosphoric acid treated (PTA) Avicel] to glucose (incubation for 3 days at 85°C) was demonstrated (Kim and Ishikawa, 2010). In contrast the endo-1,4-β-glucanase from *P. furiosus* (PF0854, GH12 family, forming G2, G3, and G4 oligosaccharides) and PF0073, from the same organism, allowed for complete saccharification of barley glucan and lichenan to glucose (Kataoka and Ishikawa, 2014).

CONCLUSION

Here we describe a novel multidisciplinary approach for the screening of hydrolases of biotechnological interest, based on *in situ* enrichments using selected substrates, (comparative) genomics and biochemical characterization of enzymes. *Thermococcus* sp. strain 2319x1 was isolated from the tidal hydrothermal spring, rich in organic matter. It is the first archaeon growing on xyloglucan, one of few hyperthermophilic archaea growing on cellulose and xylan, and the first euryarchaeon growing on xylan. While the D-glucose utilization pathway by strain 2319x1 seems to proceed as described for other *Thermococcales*, the Dxylose utilization pathway probably involves novel steps and requires further experimental analysis. Genomic analysis revealed a single gene, annotated as a glycosidase, capable of degrading β-glycosides. This protein (MDG) possesses a unique multidomain organization and acts mainly as endoglucanase and β-(endo)-mannosidase with various (although very low) side activities including xyloglucanase, xylanase, 1,3-glucanase.

The GH5 domain of MDG exhibits mainly endoglucanase and β-(endo)-mannosidase activities, while the GH12 domains presumably give additional multifunctionality to the MDG. The two CBM2 domains of this enzyme probably allow its attachment and binding to insoluble polysaccharides. Altogether, this domain organization seems to allow a processive hydrolysis of various polysaccharides, including insoluble ones. The alkaline pH optimum (pH 8.5) of the MDG, as well as its multifunctionalty, might be of interest for future biotechnological applications,

TABLE 2 | Characterized endoglucanases and endoxylanases from hyperthermophilic Archaea.

Microorganism	Characterized enzyme	Uniprot accession	Domain organization	T_{opt}°C	pH	Substrates	Reference
Thermococcus sp. strain 2319x1	Multifunctional glycosidase		GH5-12-12-CBM2-2	60 (in *E. coli*)	8.5	CMC, amorphous cellulose, CE-cellulose, β-glucan, lichenan, xyloglucan, Avicel, xylan (beech, birch), locust bean gum,	This work
				90 (single GH5)	5.5 (single GH5)	cellooligosaccharides (cellotriose to cellohexaose), cellobiose	
Enrichment culture (*Ignisphaera*)	Endoglucanase	F6M085	GH new GH-A clan + possibly 3 unknown domains	109	5.5	Lichenan, Avicel, CMC, β-glucan	Graham et al., 2011
Sulfolobus solfataricus MT 4 (DSM 5833)	Endoglucanase/ xylanase	Q97YG7	GH12	90	7.0	xylan (oat, beech, birch), CMC, arabinan	Cannio et al., 2004 Maurelli et al., 2008
Pyrococcus furiosus (DSM 3638)	Endoglucanase	Q9V2T0	GH12	100	6.0	Cellotriose, cellopentaose, cellohexaose, β-glucan, lichanan, CMC	Bauer et al., 1999
Pyrococcus furiosus (DSM 3638)	Endo-1,3-β-glucanase	O73951	GH16	100–105	5.5–6.5	β-glucan, laminarin, lichenan	Gueguen et al., 1997
Pyrococcus horikoshii OT3	Endoglucanase	O58925	GH5	>97	5.4–6.0	CMC, Avicel, lichenan	Ando et al., 2002
Thermococcus zilligi AN1 (DSM 2770)	Xylanase*	Q977U0	GT35	85	6.0	Xylan (larch, birch, oat), arabinoxylan (wheat)	Uhl and Daniel, 1999; Rolland et al., 2002
Environmental (*Thermococcus*)	Endolgucanase	A0A0F7YYA5	GH12 + CBM2 + CBM2	92	5.5	CMC, β-glucan, lichenan, AMC	Leis et al., 2015

**Questionable, see "Discussion" for details.*

since most reported thermostable endoglucanases possess more acidic pH range (**Table 2**). Horizontal gene transfer is regarded as the determining factor behind archaeal GH repertoires (Henrissat et al., 2009). The genomic context of the MDG-encoding gene supports its proneness to horizontal gene transfer. Moreover, phylogenetic analysis of the individual protein domains of the MDG approved that the unusual domain architecture was gained by independent HGT events ("tinkering") from different bacterial phyla, i.e., *Proteobacteria*, *Actinobacteria,* and *Thermotogae*. Massive horizontal gene transfer from *Thermotogales* to *Thermococcales* and, more general, from Bacteria to Archaea has been reported previously (Nelson et al., 1999; Nelson-Sathi et al., 2015).

AUTHOR CONTRIBUTIONS

SG, AS, TS, KZ, IK contributed to isolation, growth experiments, extracellular enzyme measurements of the new *Thermococcus* strain. PM, IK, and XP contributed to genome sequenzing and annotation of the genome sequence. CS, KJ, VK, CB contributed to cloning and enzymatic characterization of the multidomain cellulase. EB-O, BS designed the work. CS, IK, EB-O, and BS wrote the manuscript.

ACKNOWLEDGMENTS

CS, VK, KJ, PM, AS were supported by the HotZyme project (Project reference: 265933) within the seventh framework programme for research and technological development (FP7) of the European Union. The work of TS, EB-O was supported by a RAS programme "Molecular and Cell Biology" and the work of SG, KZ, and IK was supported by RFBR grant A_13-04-02157 and RSF grant 16-14-0012

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at:http://journal.frontiersin.org/article/10.3389/fmicb.2016.00552

REFERENCES

Amend, J. P., and Shock, E. L. (2001). Energetics of overall metabolic reactions of thermophilic and hyperthermophilic Archaea and bacteria. *FEMS Microbiol. Rev.* 25, 175–243. doi: 10.1111/j.1574-6976.2001.tb00576.x

Ando, S., Ishida, H., Kosugi, Y., and Ishikawa, K. (2002). Hyperthermostable Endoglucanase from *Pyrococcus horikoshii*. *Appl. Environ. Microbiol.* 68, 430–433. doi: 10.1128/AEM.68.1.430-433.2002

Atomi, H., Fukui, T., Kanai, T., Matsumi, R., Fujiwara, S., and Imanaka, T. (2005). Complete genome sequence of the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1 and comparison with *Pyrococcus* genomes. *Genome Res.* 15, 352–363. doi: 10.1101/gr.3003105

Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., et al. (2008). The RAST server: rapid annotations using subsystems technology. *BMC Genomics* 9:75. doi: 10.1186/1471-2164-9-75

Barns, S. M., Delwiche, C. F., Palmer, J. D., and Pace, N. R. (1996). Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences. *Proc. Natl. Acad. Sci. U.S.A.* 93, 9188–9193. doi: 10.1073/pnas.93.17.9188

Bauer, M. W., Bylina, E. J., Swanson, R. V., and Kelly, R. M. (1996). Comparison of a beta-glucosidase and a beta-mannosidase from the hyperthermophilic archaeon *Pyrococcus furiosus*. Purification, characterization, gene cloning, and sequence analysis. *J. Biol. Chem.* 271, 23749–23755. doi: 10.1074/jbc.271.39.23749

Bauer, M. W., Driskill, L. E., Callen, W., Snead, M. A., Mathur, E. J., and Kelly, R. M. (1999). An endoglucanase, EglA, from the hyperthermophilic archaeon *Pyrococcus furiosus* hydrolyzes beta-1,4 bonds in mixedlinkage (1– >3),(1– > 4)-beta-D-glucans and cellulose. *J. Bacteriol.* 181, 284–290.

Béguin, P., and Aubert, J. P. (1994). The biological degradation of cellulose. *FEMS Microbiol. Rev.* 13, 25–58.

Bernfeld, P. (1955). Amylases, α and β. *Methods Enzymol.* 1, 149–158. doi: 10.1016/0076-6879(55)01021-5

Boraston, A. B., Sandercock, L., Warren, R. A., and Kilburn, D. G. (2003). O-glycosylation of a recombinant carbohydrate-binding module mutant secreted by *Pichia* pastoris. *J. Mol. Microbiol. Biotechnol.* 5, 29–36. doi: 10.1159/000068721

Bräsen, C., Esser, D., Rauch, B., and Siebers, B. (2014). Carbohydrate metabolism in Archaea: current insights into unusual enzymes and pathways and their regulation. *Microbiol. Mol. Biol. Rev.* 78, 89–175. doi: 10.1128/MMBR.00041-13

Bronnenmeier, K., Kern, A., Liebl, W., and Staudenbauer, W. L. (1995). Purification of *Thermotoga maritima* enzymes for the degradation of cellulosic materials. *Appl. Environ. Microbiol.* 61, 1399–1407.

Cannio, R., Di Prizito, N., Rossi, M., and Morana, A. (2004). A xylan-degrading strain of *Sulfolobus solfataricus*: isolation and characterization of the xylanase activity. *Extremophiles* 8, 117–124. doi: 10.1007/s00792-003-0370-3

Chen, L., Drake, M. R., Resch, M. G., Greene, E. R., Himmel, M. E., Chaffey, P. K., et al. (2014). Specificity of O-glycosylation in enhancing the stability and cellulose binding affinity of Family 1 carbohydrate-binding modules. *Proc. Natl. Acad. Sci. U.S.A.* 111, 7612–7617. doi: 10.1073/pnas.1402518111 Cheng, Y. S., Ko, T. P., Wu, T. H., Ma, Y., Huang, C. H., Lai, H. L., et al. (2011). Crystal structure and substrate-binding mode of cellulase 12A from *Thermotoga maritima*. *Proteins* 79, 1193–1204. doi: 10.1002/prot.22953

Coady, M. J., Wallendorff, B., Gagnon, D. G., and Lapointe, J. Y. (2002). Identification of a novel Na+/myo-inositol cotransporter. *J. Biol. Chem.* 277, 35219–35224. doi: 10.1074/jbc.M204321200

Delcher, A. L., Bratke, K. A., Powers, E. C., and Salzberg, S. L. (2007). Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics* 23, 673–679. doi: 10.1093/bioinformatics/btm009

Egorova, K., and Antranikian, G. (2005). Industrial relevance of thermophilic Archaea. *Curr. Opin. Microbiol.* 8, 649–655. doi: 10.1016/j.mib.2005. 10.015

Erbeznik, M., Hudson, S. E., Herrman, A. B., and Strobel, H. J. (2004). Molecular analysis of the xylFGH operon, coding for xylose ABC transport, in *Thermoanaerobacter ethanolicus*. *Curr. Microbiol.* 48, 295–299. doi: 10.1007/s00284-003-4202-6

Ferrer, M., Martinez-Martinez, M., Bargiela, R., Streit, W. R., Golyshina, O. V., and Golyshin, P. N. (2015). Estimating the success of enzyme bioprospecting through metagenomics: current status and future trends. *Microb. Biotechnol.* 9, 22–34. doi: 10.1111/1751-7915.12309

Gavrilov, S. N., Slobodkin, A. I., Robb, F. T., and de Vries, S. (2007). [Characterization of membrane-bound Fe(III)-EDTA reductase activities of the thermophilic gram-positive dissimilatory iron-reducing bacterium *Thermoterrabacterium ferrireducens*]. *Mikrobiologiia* 76, 164–171.

Gordon, D., Abajian, C., and Green, P. (1998). Consed: a graphical tool for sequence finishing. *Genome Res.* 8, 195–202. doi: 10.1101/gr.8.3.195

Goris, J., Konstantinidis, K. T., Klappenbach, J. A., Coenye, T., Vandamme, P., and Tiedje, J. M. (2007). DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int. J. Syst. Evol. Microbiol.* 57, 81–91. doi: 10.1099/ijs.0.64483-0

Graham, J. E., Clark, M. E., Nadler, D. C., Huffer, S., Chokhawala, H. A., Rowland, S. E., et al. (2011). Identification and characterization of a multidomain hyperthermophilic cellulase from an archaeal enrichment. *Nat. Commun.* 2:375. doi: 10.1038/ncomms1373

Gueguen, Y., Voorhorst, W. G., van der Oost, J., and de Vos, W. M. (1997). Molecular and biochemical characterization of an endo-beta-1,3- glucanase of the hyperthermophilic archaeon *Pyrococcus furiosus*. *J. Biol. Chem.* 272, 31258–31264. doi: 10.1074/jbc.272.50.31258

Han, C., and Chain, P. (2006). "Finishing repetitive regions automatically with dupfinisher," in *Proceedings of 2006 International Conference on Bioinformatics & Computational Biology: 2006*, eds H. R. Arabnia and H. Valafar (Las Vegas, NV: CSREA Press), 142–147.

Henrissat, B., Surolia, A., and Stanley, P. (2009). "A genomic view of glycobiology," in *Essentials of Glycobiology*, 2nd Edn, Chap. 7, eds A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, C. R. Bertozzi, et al. (New York, NY: Cold Spring Harbor Laboratory Press).

Huber, R., Dyba, D., Huber, H., Burggraf, S., and Rachel, R. (1998). Sulfur-inhibited *Thermosphaera aggregans* sp. nov., a new genus of hyperthermophilic archaea isolated after its prediction from environmentally derived 16S rRNA sequences. *Int. J. Syst. Bacteriol.* 48(Pt 1), 31–38. doi: 10.1099/00207713-48-1-31

Hugenholtz, P., Pitulle, C., Hershberger, K. L., and Pace, N. R. (1998). Novel division level bacterial diversity in a Yellowstone hot spring. *J. Bacteriol.* 180, 366–376.

Johnsen, U., Dambeck, M., Zaiss, H., Fuhrer, T., Soppa, J., Sauer, U., et al. (2009). D-xylose degradation pathway in the halophilic archaeon *Haloferax volcanii*. *J. Biol. Chem.* 284, 27290–27303. doi: 10.1074/jbc.M109. 003814

Kashima, Y., Mori, K., Fukada, H., and Ishikawa, K. (2005). Analysis of the function of a hyperthermophilic endoglucanase from *Pyrococcus horikoshii* that hydrolyzes crystalline cellulose. *Extremophiles* 9, 37–43. doi: 10.1007/s00792004-0418-z

Kataoka, M., and Ishikawa, K. (2014). Complete saccharification of betaglucan using hyperthermophilic endocellulase and beta-glucosidase from *Pyrococcus furiosus*. *Biosci. Biotechnol. Biochem.* 78, 1537–1541. doi: 10.1080/09168451.2014.923300

Kato, N., Yurimoto, H., and Thauer, R. K. (2006). The physiological role of the ribulose monophosphate pathway in bacteria and archaea. *Biosci. Biotechnol. Biochem.* 70, 10–21. doi: 10.1271/bbb.70.10

Kim, H., and Ishikawa, K. (2011). Functional analysis of hyperthermophilic endocellulase from *Pyrococcus horikoshii* by crystallographic snapshots. *Biochem. J.* 437, 223–230. doi: 10.1042/BJ20110292

Kim, H. W., and Ishikawa, K. (2010). Structure of hyperthermophilic endocellulase from *Pyrococcus horikoshii*. *Proteins* 78, 496–500. doi: 10.1002/prot. 22602

Kim, H. W., Kataoka, M., and Ishikawa, K. (2012). Atomic resolution of the crystal structure of the hyperthermophilic family 12 endocellulase and stabilizing role of the DxDxDG calcium-binding motif in *Pyrococcus furiosus*. *FEBS Lett.* 586, 1009–1013. doi: 10.1016/j.febslet.2012.02.029

Koning, S. M., Elferink, M. G., Konings, W. N., and Driessen, A. J. (2001). Cellobiose uptake in the hyperthermophilic archaeon *Pyrococcus furiosus* is mediated by an inducible, high-affinity ABC transporter. *J. Bacteriol.* 183, 4979–4984. doi: 10.1128/JB.183.17.4979-4984.2001

Kublanov, I. V., Bidjieva, S. K. H., Mardanov, A. V., and Bonch-Osmolovskaya, E. A. (2009). *Desulfurococcus kamchatkensis* sp. nov., a novel hyperthermophilic protein-degrading archaeon isolated from a Kamchatka hot spring. *Int. J. Syst. Evol. Microbiol.* 59, 1743–1747. doi: 10.1099/ijs.0. 006726-0

Lee, S. J., Engelmann, A., Horlacher, R., Qu, Q., Vierke, G., Hebbeln, C., et al. (2003). TrmB, a sugar-specific transcriptional regulator of the trehalose/maltose ABC transporter from the hyperthermophilic archaeon *Thermococcus litoralis*. *J. Biol. Chem.* 278, 983–990. doi: 10.1074/jbc.M210236200

Leis, B., Heinze, S., Angelov, A., Pham, V. T., Thurmer, A., Jebbar, M., et al. (2015). Functional screening of hydrolytic activities reveals an extremely thermostable cellulase from a deep-sea archaeon. *Front. Bioeng. Biotechnol.* 3:95. doi: 10.3389/fbioe.2015.00095

Liebl, W., Ruile, P., Bronnenmeier, K., Riedel, K., Lottspeich, F., and Greif, I. (1996). Analysis of a *Thermotoga maritima* DNA fragment encoding two similar thermostable cellulases, CelA and CelB, and characterization of the recombinant enzymes. *Microbiology* 142(Pt 9), 2533–2542. doi: 10.1099/00221287-142-9-2533

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., and Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 42, D490–D495. doi: 10.1093/nar/gkt1178

Luo, H., Zhang, C. T., and Gao, F. (2014). Ori-Finder 2, an integrated tool to predict replication origins in the archaeal genomes. *Front. Microbiol.* 5:482. doi: 10.3389/fmicb.2014.00482

Mardanov, A. V., Ravin, N. V., Svetlitchnyi, V. A., Beletsky, A. V., Miroshnichenko, M. L., Bonch-Osmolovskaya, E. A., et al. (2009). Metabolic versatility and indigenous origin of the archaeon *Thermococcus sibiricus*, isolated from a siberian oil reservoir, as revealed by genome analysis. *Appl. Environ. Microbiol.* 75, 4580–4588. doi: 10.1128/AEM.00718-09

Matsumi, R., Manabe, K., Fukui, T., Atomi, H., and Imanaka, T. (2007). Disruption of a sugar transporter gene cluster in a hyperthermophilic archaeon using a host-marker system based on antibiotic resistance. *J. Bacteriol.* 189, 2683–2691. doi: 10.1128/JB.01692-06

Maurelli, L., Giovane, A., Esposito, A., Moracci, M., Fiume, I., Rossi, M., et al. (2008). Evidence that the xylanase activity from *Sulfolobus solfataricus* Oalpha is encoded by the endoglucanase precursor gene (sso1354) and characterization of the associated cellulase activity. *Extremophiles* 12, 689–700. doi: 10.1007/s00792-008-0175-5

Miller, G. L. (1959). Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Anal. Chem.* 31, 426–428. doi: 10.1021/ac60147a030

Möller, S., Croning, M. D., and Apweiler, R. (2001). Evaluation of methods for the prediction of membrane spanning regions. *Bioinformatics* 17, 646–653. doi: 10.1093/bioinformatics/17.7.646

Nakamura, T., Mine, S., Hagihara, Y., Ishikawa, K., Ikegami, T., and Uegaki, K. (2008). Tertiary structure and carbohydrate recognition by the chitin-binding domain of a hyperthermophilic chitinase from *Pyrococcus furiosus*. *J. Mol. Biol.* 381, 670–680. doi: 10.1016/j.jmb.2008.06.006

Nawrocki, E. P., Burge, S. W., Bateman, A., Daub, J., Eberhardt, R. Y., Eddy, S. R., et al. (2015). Rfam 12.0*:* updates to the RNA families database. *Nucleic Acids Res.* 43, D130–D137. doi: 10.1093/nar/gku1063

Nawrocki, E. P., and Eddy, S. R. (2013). Infernal 1.1*:* 100-fold faster RNA homology searches. *Bioinformatics* 29, 2933–2935. doi: 10.1093/bioinformatics/btt509

Nelson, K. E., Clayton, R. A., Gill, S. R., Gwinn, M. L., Dodson, R. J., Haft, D. H., et al. (1999). Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of *Thermotoga maritima*. *Nature* 399, 323–329. doi: 10.1038/20601

Nelson-Sathi, S., Sousa, F. L., Roettger, M., Lozada-Chavez, N., Thiergart, T., Janssen, A., et al. (2015). Origins of major archaeal clades correspond to gene acquisitions from bacteria. *Nature* 517, 77–80. doi: 10.1038/nature13805

Nunn, C. E., Johnsen, U., Schönheit, P., Fuhrer, T., Sauer, U., Hough, D. W., et al. (2010). Metabolism of pentose sugars in the hyperthermophilic archaea *Sulfolobus solfataricus* and *Sulfolobus acidocaldarius*. *J. Biol. Chem.* 285, 33701–33709. doi: 10.1074/jbc.M110.146332

Orita, I., Sato, T., Yurimoto, H., Kato, N., Atomi, H., Imanaka, T., et al. (2006). The ribulose monophosphate pathway substitutes for the missing pentose phosphate pathway in the archaeon *Thermococcus kodakaraensis*. *J. Bacteriol.* 188, 4698–4704. doi: 10.1128/JB.00492-06

Perevalova, A. A., Svetlichny, V. A. I, Kublanov, V., Chernyh, N. A., Kostrikina, N. A., Tourova, T. P., et al. (2005). *Desulfurococcus fermentans* sp. nov., a novel hyperthermophilic archaeon from a Kamchatka hot spring, and emended description of the genus *Desulfurococcus*. *Int. J. Syst. Evol. Microbiol.* 55, 995–999. doi: 10.1099/ijs.0.63378-0

Petersen, T. N., Brunak, S., von Heijne, G., and Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* 8, 785–786. doi: 10.1038/nmeth.1701

Prokofeva, M. I., Kostrikina, N. A., Kolganova, T. V., Tourova, T. P., Lysenko, A. M., Lebedinsky, A. V., et al. (2009). Isolation of the anaerobic thermoacidophilic crenarchaeote *Acidilobus saccharovorans* sp. nov. and proposal of Acidilobales ord. nov., including *Acidilobaceae fam*. nov. and *Caldisphaeraceae* fam. nov. *Int. J. Syst. Evol. Microbiol.* 59, 3116–3122. doi: 10.1099/ijs.0.010355-0

Rolland, J. L., Gueguen, Y., Flament, D., Pouliquen, Y., Street, P. F., and Dietrich, J. (2002). Comment on "The first description of an archaeal hemicellulase: the xylanase from *Thermococcus zilligii* strain AN1": evidence that the unique N-terminal sequence proposed comes from a maltodextrin phosphorylase. *Extremophiles* 6, 349–350. doi: 10.1007/s00792-001-0258-z

Saier, M. H. Jr., Reddy, V. S., Tamang, D. G., and Vastermark, A. (2014). The transporter classification database. *Nucleic Acids Res.* 42, D251–D258. doi: 10.1093/nar/gkt1097

Schut, G. J., Lipscomp, G. L., Han, Y., Notey, J. S., Kelly, R. M., and Adams, M. M. W. (2014). *The order Thermococcales and the Family Thermococcaceae, The Prokaryotes: Other Major Lineages of Bacteria and The Archaea*. Berlin: Springer-Verlag, 363–383.

Siebers, B., and Schönheit, P. (2005). Unusual pathways and enzymes of central carbohydrate metabolism in Archaea. *Curr. Opin. Microbiol.* 8, 695–705. doi: 10.1016/j.mib.2005.10.014

Simpson, P. J., Xie, H., Bolam, D. N., Gilbert, H. J., and Williamson, M. P. (2000). The structural basis for the ligand specificity of family 2 carbohydrate-binding modules. *J. Biol. Chem.* 275, 41137–41142. doi: 10.1074/jbc.M006948200

Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., et al. (1985). Measurement of protein using bicinchoninic acid. *Anal. Biochem.* 150, 76–85. doi: 10.1016/0003-2697(85)90442-7

Söding, J., Biegert, A., and Lupas, A. N. (2005). The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* 33, W244–W248. doi: 10.1093/nar/gki408

Sokolova, T. G., Henstra, A. M., Sipma, J., Parshina, S. N., Stams, A. J., and Lebedinsky, A. V. (2009). Diversity and ecophysiological features of thermophilic carboxydotrophic anaerobes. *FEMS Microbiol. Ecol.* 68, 131–141. doi: 10.1111/j.1574-6941.2009.00663.x

Sorokin, D. Y., Toshchakov, S. V., Kolganova, T. V., and Kublanov, I. V. (2015). Halo(natrono)archaea isolated from hypersaline lakes utilize cellulose and chitin as growth substrates. *Front. Microbiol.* 6:942. doi: 10.3389/fmicb.2015.00942

Stephens, C., Christen, B., Fuchs, T., Sundaram, V., Watanabe, K., and Jenal, U. (2007). Genetic analysis of a novel pathway for D-xylose metabolism in *Caulobacter crescentus*. *J. Bacteriol.* 189, 2181–2185. doi: 10.1128/JB.01438-06

Toshchakov, S. V., Kublanov, V. I., Messina, E., Yakimov, M. M., and Golyshin, P. N. (2015). "Genomic analysis of pure cultures and communities," in *Hydrocarbon and Lipid Microbiology Protocols: Cultivation, Springer Protocols Handbooks*, eds T. J. McGenity, K. N. Timmis, and B. Nogales Fernández (Heidelberg: Springer-Verlag).

Uhl, A. M., and Daniel, R. M. (1999). The first description of an archaeal hemicellulase: the xylanase from *Thermococcus zilligii* strain AN1. *Extremophiles* 3, 263–267. doi: 10.1007/s007920050126

Unsworth, L. D., van der Oost, J., and Koutsopoulos, S. (2007). Hyperthermophilic enzymes: stability, activity and implementation strategies for high temperature applications. *FEBS J.* 274, 4044–4056. doi: 10.1111/j.1742-4658.2007. 05954.x

Vanfossen, A. L., Lewis, D. L., Nichols, J. D., and Kelly, R. M. (2008). Polysaccharide degradation and synthesis by extremely thermophilic anaerobes. *Ann. N. Y. Acad. Sci.* 1125, 322–337. doi: 10.1196/annals.1419.017

Xavier, K. B., Martins, L. O., Peist, R., Kossmann, M., Boos, W., and Santos, H. (1996). High-affinity maltose/trehalose transport system in the hyperthermophilic archaeon *Thermococcus litoralis*. *J. Bacteriol.* 178, 4773–4777.

Xu, G. Y., Ong, E., Gilkes, N. R., Kilburn, D. G., Muhandiram, D. R., Harris-Brandts, M., et al. (1995). Solution structure of a cellulose-binding domain from *Cellulomonas* fimi by nuclear magnetic resonance spectroscopy. *Biochemistry* 34, 6993–7009. doi: 10.1021/bi00021a011

Zerbino, D. R., and Birney, E. (2008). Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* 18, 821–829. doi: 10.1101/gr.074492.107

Zhang, Y. H., Cui, J., Lynd, L. R., and Kuang, L. R. (2006). A transition from cellulose swelling to cellulose dissolution by o-phosphoric acid: evidence from enzymatic hydrolysis and supramolecular structure. *Biomacromolecules* 7, 644–648. doi: 10.1021/bm050799c

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Supplemtary information on

Isolation and characterization of the first xylanolytic hyperthermophilic euryarchaeon *Thermococcus* sp. 2319x1 and its unusual multidomain glycosidase

Sergey N. Gavrilov[1#], Christina Stracke[2#], Kenneth Jensen[3], Peter Menzel[4], Verena Kallnik[2], Alexei Slesarev[1,5], Tatyana Sokolova[1], Kseniya Zayulina[1], Christopher Bräsen[2], Elizaveta A. Bonch-Osmolovskaya[1], Xu Peng[4], Ilya V. Kublanov[1*], Bettina Siebers[2*]

[#]Both authors contributed equally
[*]Corresponding authors

[1]Winogradsky Institute of Microbiology, Research Center of Biotechnology, Russian Academy of Sciences, Moscow, Russia
[2]Molecular Enzyme Technology and Biochemistry (MEB), Biofilm Centre, Centre for Water and Environmental Research (CWE), University Duisburg-Essen, Essen, Germany
[3]Novozymes A/S, Bagsværd, Denmark
[4]Department of Biology, University of Copenhagen, Copenhagen, Denmark
[5]Fidelity Systems, Inc., Gaithersburg, MD USA

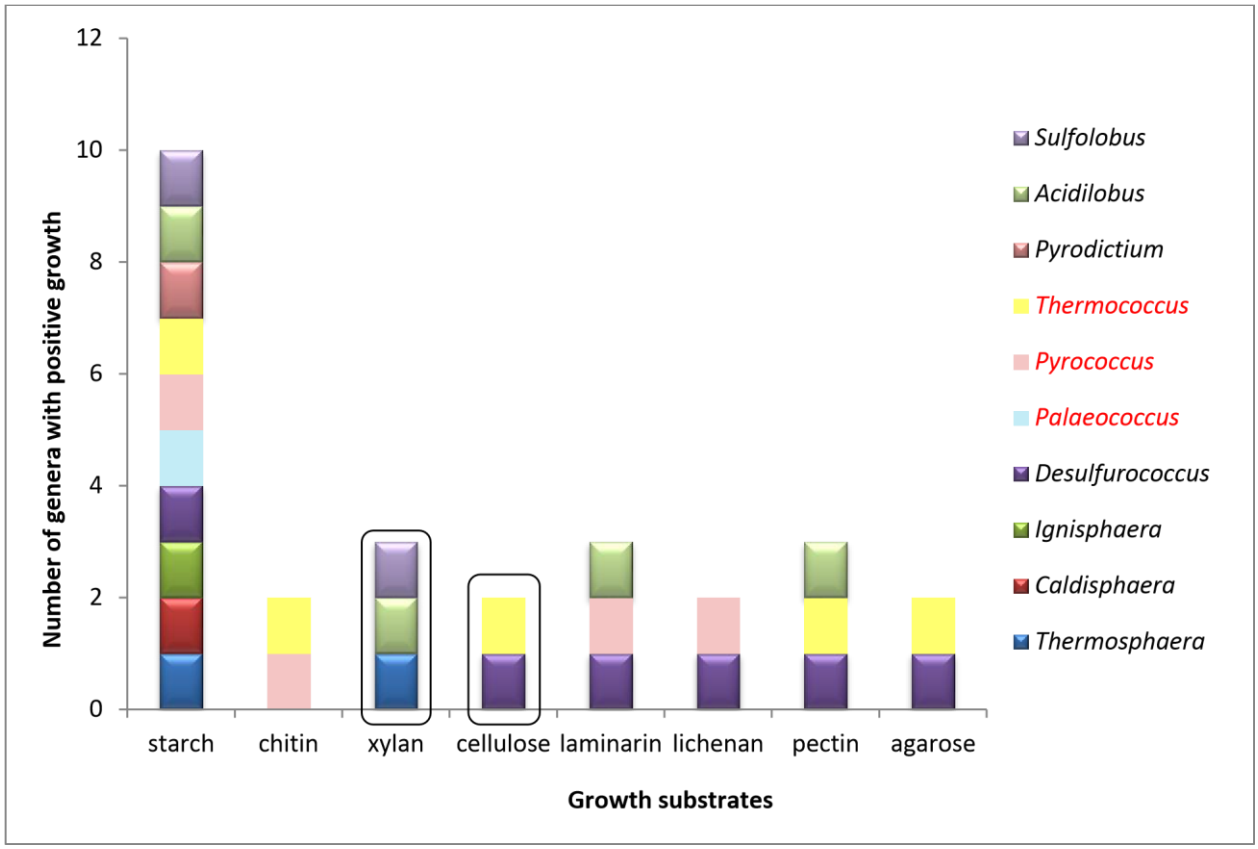
1. Introduction

Supplementary Table 1. Characterized and predicted cellulases and xylanases from hyperthermophlic Archaea.

Family	GH3	GH5	GH9	GH10	GH11	GH12	GH16	GH26	GH43	GH51	GH94
Activities	beta-xylo-gluco	beta-many (mainly endo)	beta-xylo-gluco (mainly endo)	endo-xylo	endo-xylo	beta-xylo-gluco (endo)	beta-many (mainly endo)	beta-many	xylo-arabino mainly	beta-xylo-gluco-arabino	beta-gluco-chito phospho rylases
# proteins from cultivated thermophilic archaea	**83**	**55**	**4**	**14**	**5**	**62**	**18**	**5**	**14**	**18**	**8**
Crenarchaeota											
Acidianus hospitalis W1						**1**					
Acidilobus saccharovorans 345-15	**1**	**1**				**3**					
Caldivirga maquilingensis IC-167	**2**	**1**				**5**					
Ignisphaera aggregans DSM 17230	**3**	**3**				**5**				**1**	**1**
Metallosphaera cuprina Ar-4						**1**					
Sulfolobus acidocaldarius (all strains)						**4**					
Sulfolobus islandicus (all strains)	**10**	**10**				**20**					
Sulfolobus solfataricus (all strains)	**5**	**5**				**12**					
Thermofilum carboxyditrophus 1505	**1**										
Thermofilum pendens Hrk 5	**1**	**1**				**1**					
Thermofilum sp. 1807-2	**1**										
Thermofilum sp. 1910b	**1**										
Thermoproteus uzoniensis 768-20		**1**				**2**					
Thermoproteus tenax Kra 1						**1**					
Thermosphaera aggregans DSM 11486							**1**				
Vulcanisaeta distributa DSM 14429		**1**				**1**					

Euryarchaeota												
Palaeococcus pacificus DY20341						**1**						
Picrophilus torridus DSM 9790	**1**	**1**										
Pyrococcus abyssi GE5		**1**										
Pyrococcus furiosus COM1						**1**	**1**					
Pyrococcus furiosus DSM 3638						**1**	**1**					
Pyrococcus horikoshii OT3		**1**										
Pyrococcus sp. ST04		**1**										
Thermococcus sp. ES1		**1**										
Thermococcus sp. AM4						**1**						
Thermococcus guaymasensis DSM 11113							**1**					
Thermococcus sibiricus MM 739						**3***	**1**					
Thermoplasma volcanium GSS1		**1**										

According to CAZy.
* - very distant, not in CAZy



Supplementary Figure 1. Polysaccharide-degrading hyperthermophilic Archaea. Euryarchaeal genera are in red and their elements are in 2D; Crenarchaeal genera are in black and their elements are in 3D. Genera, growing on xylan or cellulose are boxed. Alpha-linked starch and its derivatives are widely utilized by hyperthermophilic archaea while beta-linked polysaccharides, such as cellulose and xylan are used rarely, mainly by representatives of the Crenarchaeota.

2 Material and Methods

2.1 Cloning, expression and purification of the MDG and truncated versions

Supplementary Table 2. Primers for In-Fusion® cloning of the *mdg* gene and its truncated versions. The 15 bp extensions complementary to the linearized expression vector pET24a are underlined.

Primer ID:	Sequence:
fwd-*mdg* (without signalpeptide)	5'-TCGCGGATCCGAATTATGACGACTGACACAAGCACG-3
rvs- *mdg* (GH5-12-12-CBM2-2)	5'-GTGCTCGAGTGCGGCAAGGATTTCGACTCCAAGGAG-3'
rvs- *mdg-trunc1* (GH5-12-12)	5'-GGTGCTCGAGTGCGGGAGCGAGACGTTGTAGAACC-3'
rvs- *mdg-trunc2* (GH5-12)	5'-GGTGCTCGAGTGCGGCTGGACCTCTCCGATGAGAAC-3'
rvs- *mdg-trunc3* (GH5)	5'-GGTGCTCGAGTGCGGCGGGGCAATGGCGTTTCCAG-3'

2.2 Purification of the GH5 and GH5-12 protein

Since an affinity chromatography by Ni-TED column revealed no binding of the His-tagged GH5 and GH5-12 proteins an alternative purification protocol via fractionated ammonium sulfate (AS) precipitation, ion exchange chromatography and size exclusion chromatography was established. For fractionated AS precipitation the protein fraction after heat precipitation was first incubated (30 min, stirring at 4°C) in the presence of 1 M AS and after centrifugation (26000 x g for 30 min at 4°C) the GH5 protein was precipitated with 2.2 M AS and the GH5-12 protein by addition of 2.6 M AS followed by centrifugation (26000 x g for 30 min at 4°C). After resuspension and dialysis against 20 mM TRIS HCl, $pH^{25°C}$ 7.0 the protein fractions were applied on a Resource-Q column (6 mL column volume, flow rate 2 mL min^{-1}, GE Healthcare) equilibrated in 20 mM TRIS HCl $pH^{25°C}$ 7.0, followed by washing with equilibration buffer (20-fold column volume) and elution with a linear salt gradient (0-1 M NaCl, 20-fold column volume, flow rate 2 mL min^{-1}, Äkta purifier, GE Healthcare). Fractions containing the respective proteins (checked by SDS-PAGE and activity tests) were pooled and concentrated via Vivaspin concentrators (30 kDa MW cut off, Sartorius Stedim Biotech). After dialysis overnight against 50 mM TRIS HCl, 300 mM NaCl, $pH^{25°C}$ 7, both proteins were purified via size exclusion chromatography (Superdex^{TM} 200, HiLoad^{TM} 26/60 column, 320 mL volume, GE Healthcare) equilibrated in dialysis buffer (flow rate 2 mL min^{-1}, Äkta purifier, GE Healthcare). Protein fractions containing the purified protein (checked by SDS-PAGE and activity tests) were pooled and concentrated via Vivaspin concentrators (30 kDa MW cut off, Sartorius Stedim Biotech). The protein solutions were stored at -70°C (in the presence of 20% glycerol) and remained active for at least 5 month.

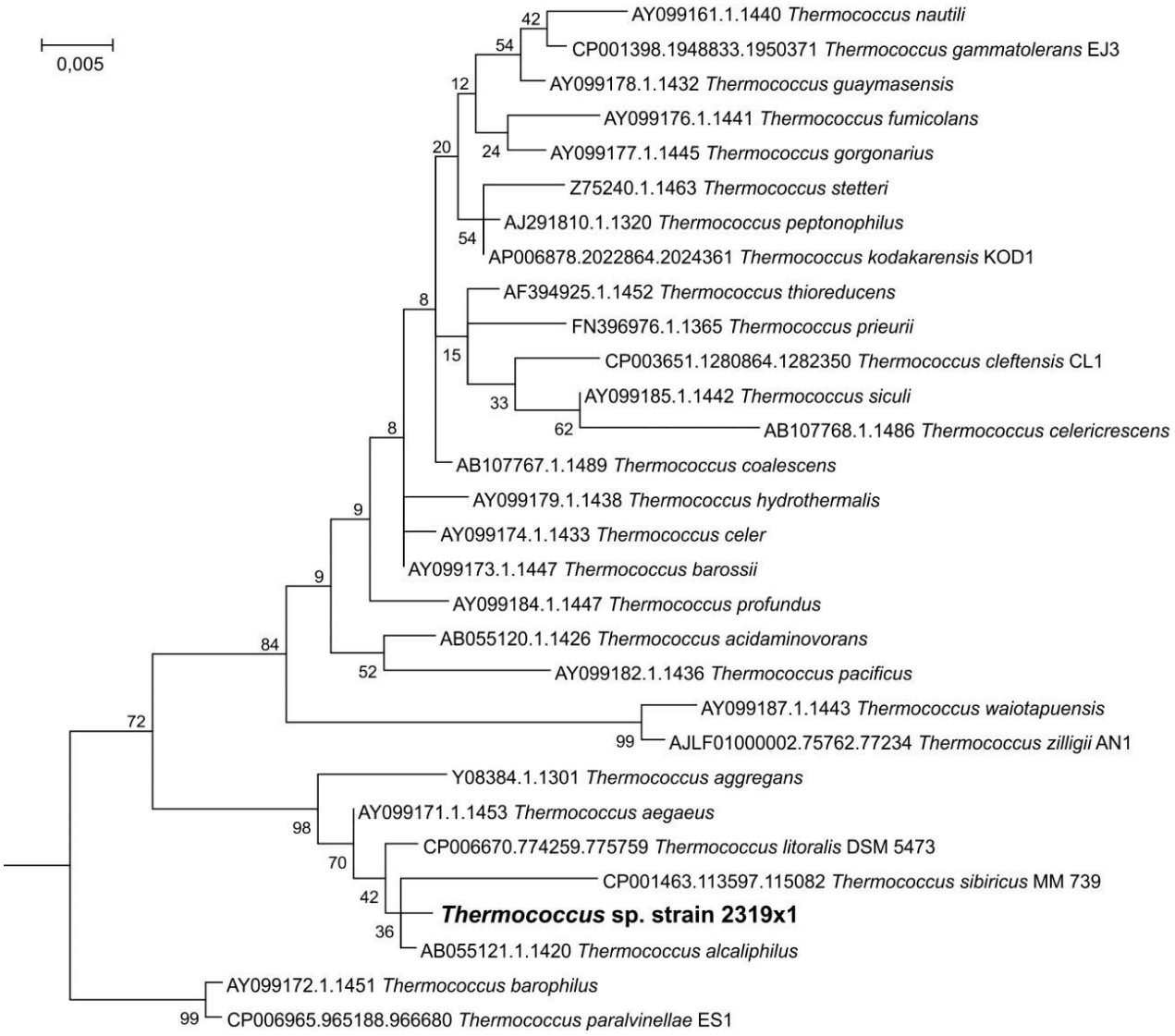
3 Results

Supplementary Table 3. Glycoside hydrolases **(**GHs) and carbohydrate esterases (CEs) encoding sequences, identified in the genome of *Thermococcus* strain 2319x1.

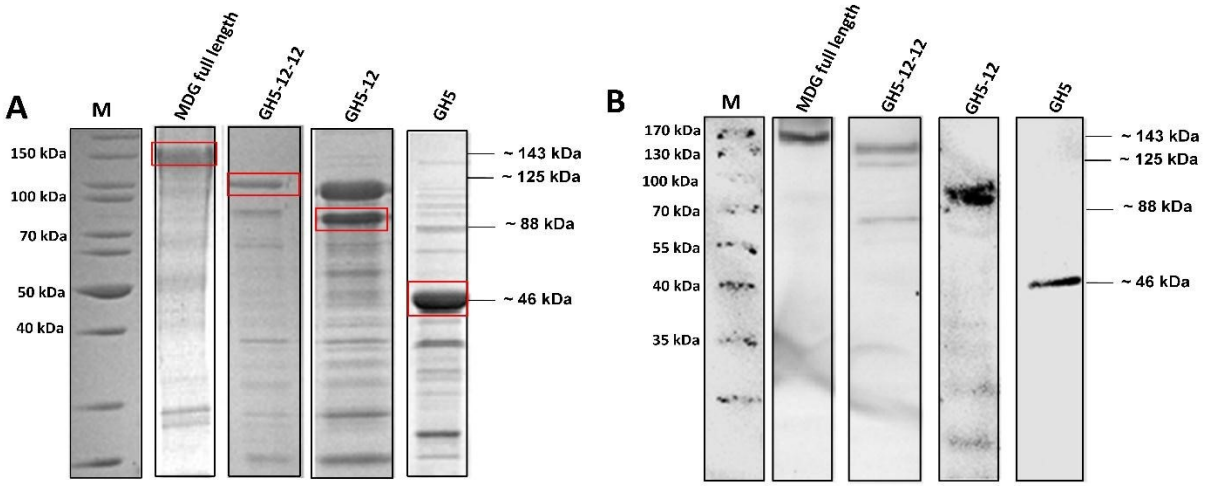
№	GenBank ID	Cazy family	Pred Hel=	Topology=	SignalP (G(+) & eukar)
1.	ADU37_CDS00170	Glycoside hydrolase 13 family enzyme, alpha-amylase	0		N
2.	ADU37_CDS02890	Glycoside hydrolase 63 family enzyme, glycogen debranching enzyme, putative	0		N
3.	ADU37_CDS02900	Glycoside hydrolase 130 family enzyme, beta-1,4 mannooligosaccharide phosphorilase	0		N
4.	ADU37_CDS03850	Glycoside hydrolase 57 family enzyme, alphaamylase/4alpha-gluconotranspherase, putative	0		N
5.	ADU37_CDS05320	Glycoside hydrolase 1 family enzyme, beta-glucosidase, putative	0		N
6.	ADU37_CDS05340	Glycoside hydrolase 57 family enzyme	0		N
7.	ADU37_CDS07270	Glycoside hydrolase 35 family enzyme,exo-beta Dglucosaminidase, putative	0		N
8.	ADU37_CDS07340	Glycoside hydrolase 1 family enzyme, beta-galactosidase, putative	0		N
9.	ADU37_CDS07440	Glycoside hydrolase 1 family enzyme, beta-galactosidase, putative	0		N
10.	ADU37_CDS07750	Glycoside hydrolase 122 family enzyme, alphaglucosidase, putative	0		N
11.	ADU37_CDS12970	Glycoside hydrolase 57 family enzyme, 1,4alphabranching enzyme	0		N
12.	ADU37_CDS13420	Carbohydrate esterase, putative	0		Y
13.	ADU37_CDS18940	Glycoside hydrolase 57 family enzyme, amylopullulanase, putative	1	o10751097i	Y
14.	ADU37_CDS18980	Glycoside hydrolase 13 family enzyme, neopullulanase, putative	0		N
15.	ADU37_CDS20530	Glycoside hydrolase 13 family enzyme	0		N
16.	ADU37_CDS21030	Glycoside hydrolase 13 family enzyme	1	i7-25o	Y
17.	ADU37_CDS21840	Putative oxidoreductase, related to GH109 family	0		N
18.	ADU37_CDS22600	Multidomain glycoside hydrolase protein, endoglucanase/endoxylanase	1	i7-29o	Y

PredHel and Topology= are results of transmembrane regions predictions by the TMHMM server (http://www.cbs.dtu.dk/services/TMHMM). "iXXX-oYYY" indicates the localization and coordinates of the transmembrane (TM) region in the protein with „i“ inside and “o” outside. TM predicitions at the N-terminal end of the protein often indicate the presence of a signal peptide (i.e.i7-25o in ADU37_CDS21030, i7-29o in ADU37_CDS22600). For ADU37_CDS18940 a Cterminal TM is predicted (protein region o1075-1097i). SignalP (G(+) & eukar): results of signal peptide predictions by the SignalP server (for Gram positives and Eukaryotes; http://www.cbs.dtu.dk/services/SignalP/index.php).

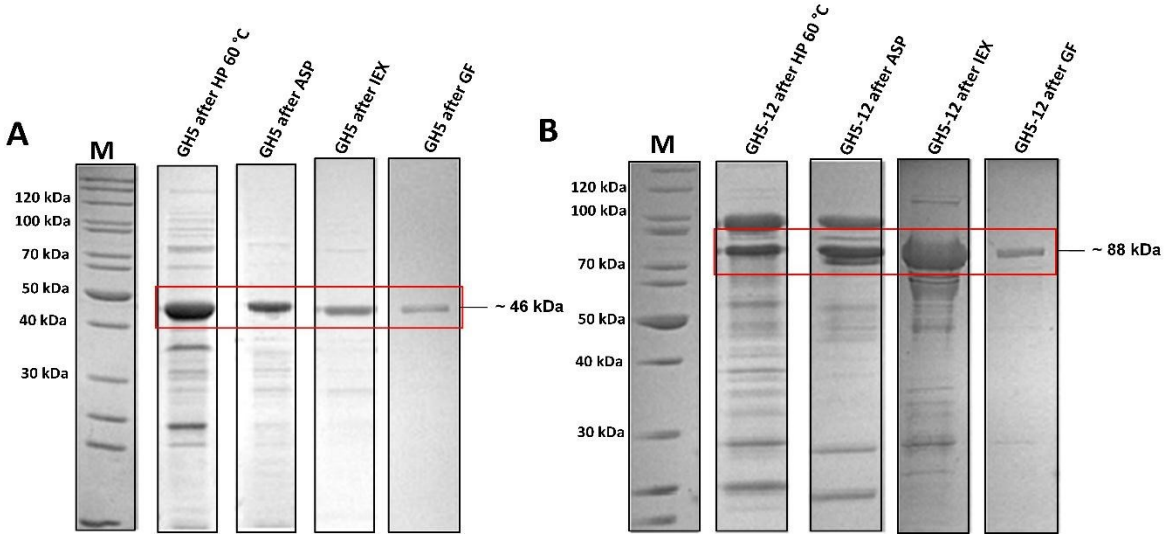
Supplementary Figure 2. Sampling site and *in situ* enrichment cultures. Arrows indicate the 50 mL bottles and 18 mL Hungate tubes incubated in the hot spring, located in the tidal zone near Goryachiy cape of Kunashir Island (South Kurils, Russian Far East region).



Supplementary Figure 3. 16S rRNA gene sequence-based Maximum likelihood phylogenetic tree of *Thermococcus* sp. strain 2319x1 and validly published representatives of the *Thermococcus* genus. The tree with the highest log likelihood (-3041.4022) is shown. The number of trees of totally 100 repetitions (bootstrap analysis) in which the associated taxa clustered together is shown next to the branches. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 31 nucleotide sequences. All positions with less than 95% site coverage were eliminated. There were a total of 1286 positions in the final dataset. Evolutionary analyses were conducted in MEGA6 (Tamura et al., 2013). Bar, 0.50 nucleotides substitutions per 100. *Thermococcus* sp. strain 2319x1 is shown in bold. *Pyrococcus yayanosii* strain CH1 (NR 102853.1) was used as outgroup.



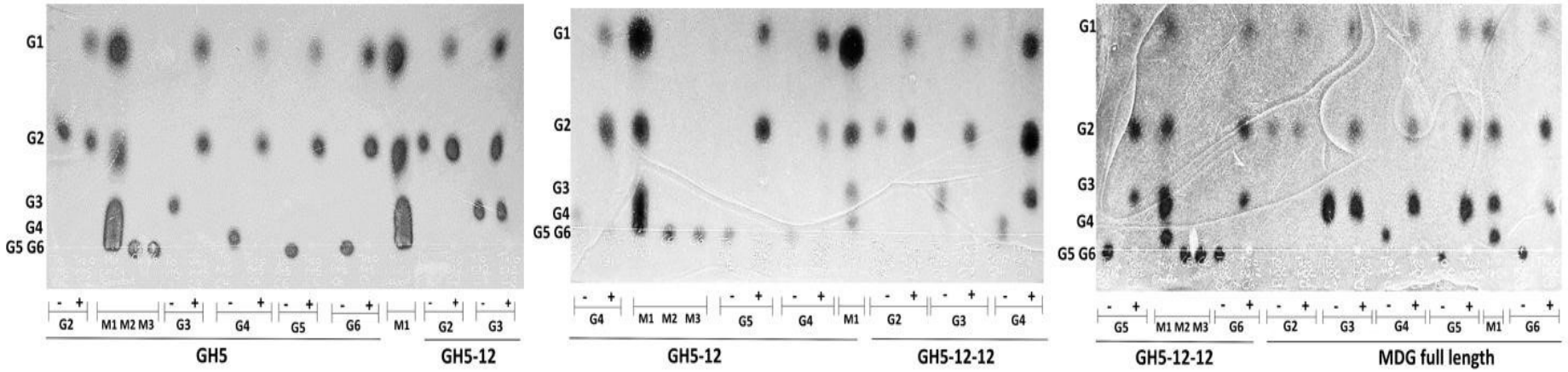
Supplementary Figure 4. Expression of the recombinant MDG and the different truncated versions in *Escherichia coli*. (A) Protein fractions (5 µg protein each lane) after heat precipitation at 60°C (30 min) were separated via SDS-PAGE (12.5% SDS-PAGE, coomassie staining) and (B) expressed proteins were identified by immunodetection via anti His-tag antibodies. Marker (M): unstained (A) and prestained (B) protein standard (Fermentas); MDG full length (~ 143 kDa); GH5-12-12 (~ 125 kDa); GH5-12 (~ 88 kDa); single GH5 (~ 46 kDa)).



Supplementary Figure 5. Purification of the truncated GH5 (A) and GH5-12 (B) protein. Purification was followed by SDS-PAGE (12.5% SDS-PAGE, coomassie staining). Protein fractions (5 µg protein) after heat precipitation (HP), ammonium sulphate precipitation (ASP), ion exchange chromatography (IEX) and protein fractions (2 µg protein) after size exclusion chromatography (SEC) are shown. Marker (M): unstained protein standard (Fermentas).

Supplementary Table 4. Substrate specificity and specific activity of the full length MDG and the truncated MDG proteins.

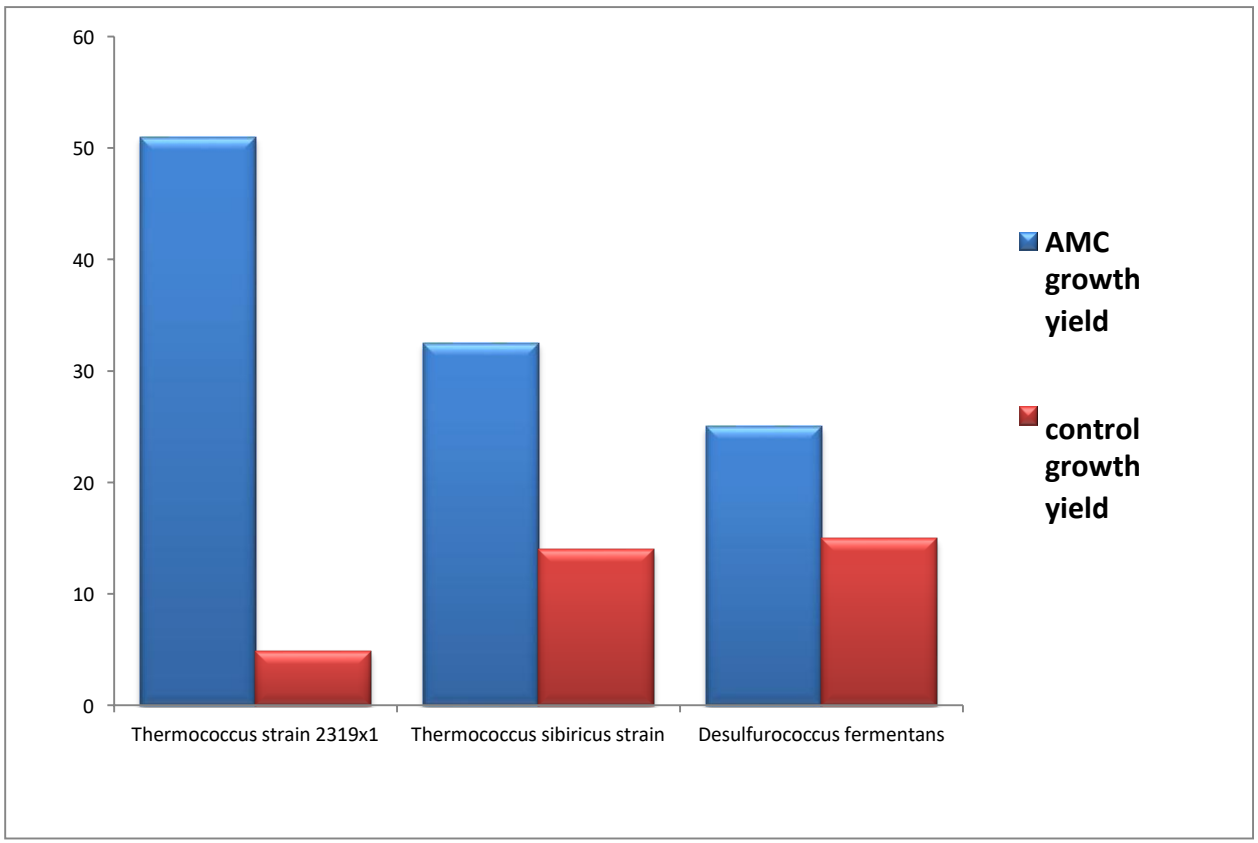
		Specific activities [U mg^{-1}] (K_{cat} [s^{-1}])			
Substrate	**Backbone linkage(s)**	**GH5**	**GH5-12**	**GH5-12-12**	**MDG full length**
Glucan oligosaccharides					
Cellobiose	β-1,4 only	0.5 (1498)	0.6 (3403)	1.4 (10248)	1.0 (8937)
Cellohexaose	β-1,4 only	2.3 (6293)	2.7 (14380)	6.3 (47120)	5.6 (48250)
Glucan polysaccharides					
Barley glucan	β-1,3/4	4.3 (13606)	5.6 (29832)	15.9 (119475)	14.2 (121664)
Lichenan	β-1,3/4	7.01 (19347)	7.02 (37065)	7.44 (39283)	6.75 (57915)
Carboxymethyl cellulose	β-1,4 only	2.7 (7452)	3.4 (17952)	6.7 (50250)	5.9 (50622)
Hydroxyethyl cellulose	β-1,4 only	2.02 (5575)	2.6 (13728)	4.7 (35250)	3.3 (30030)
Xyloglucan	β-1,4/6	0.88 (2428)	0.88 (4646)	0.94 (4963)	0.95 (8151)
MCC Avicel PH101	β-1,4 only	0.8 (2152)	0.8 (4418)	0.8 (5920)	0.7 (5920)
PAT Avicel PH101	β-1,4 only	2.0 (5437)	2.5 (13200)	5.0 (36750)	3.1 (26169)
Curdlan	β-1,3 only	0.39 (1076)	0.39 (2059)	0.5 (3750)	0.38 (3260)
Mannans					
Galactomannan	β-1,4/6	9.2 (25392)	6.3 (33264)	2.09 (15675)	2.2 (18876)
Xylans					
Birchwood xylan	β-1,4 only	0.27 (745)	0.29 (1531)	0.25 (1875)	0.23 (1973)



Supplementary Figure 6. Degradation of cellobiose and cellooligosaccharides by the MDG and its truncated versions.
Substrate solutions (1% (w/v)) were mixed with 25 µg of each protein and either directly transferred to ice and the reaction was stopped by the addition of 80 % (v/v) acetone **(-)**, or incubated for 2 h at 60°C **(+)**. Hydrolysis products were separated on aluminium sheet (20 x 20 cm) silica gel 60/kieselguhr F_{254} plates (Merck) using a mixture of ethyl acetate, methanol and H_2O (68:23:9, v/v/v) as mobile phase. Products were visualized with $KMnO_4$ solution (1.5 g $KMnO_4$, 10 g K_2CO_3 and 1.25 ml 10 % aq. NaOH in 200 ml H_2O). G: Number of glucose units. G1 = glucose; G2 = cellobiose; G3 = cellotriose; G4 = cellotetraose; G5 = cellopentaose; G6 = cellohexaose. Marker: M1 = G1, G2, G3, G4. M2 = G5. M3 = G6.

3

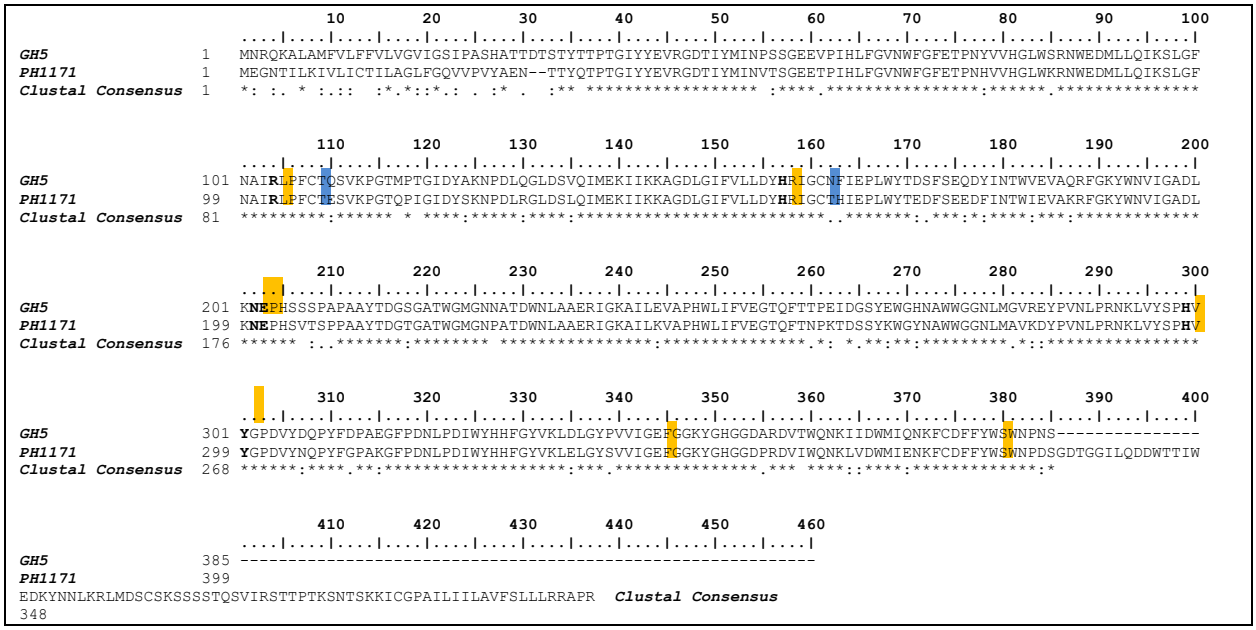
4 Discussion



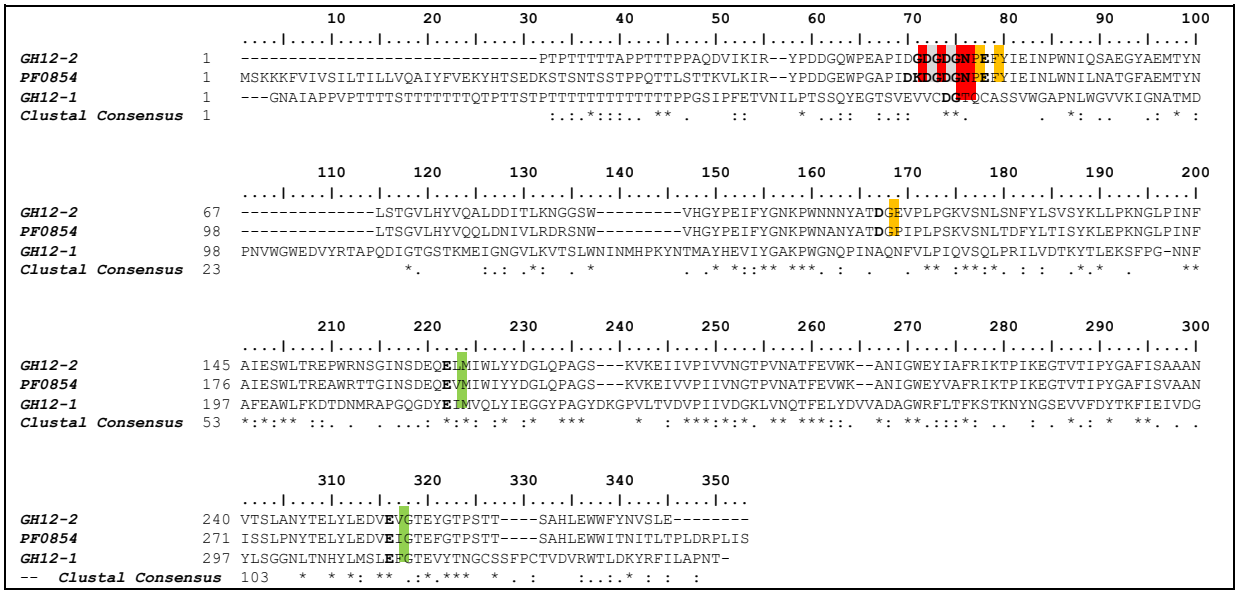
Supplementary Figure 7. Comparative analysis of growth yields of three hyperthermophilic archaeal strains, growing on amorphous cellulose (AMC). All strains were cultivated under their optimal growth conditions. Control means the same medium without AMC substrate.

Supplementary Table 5. Best BLAST hits of CBM2-1 and CBM2-2

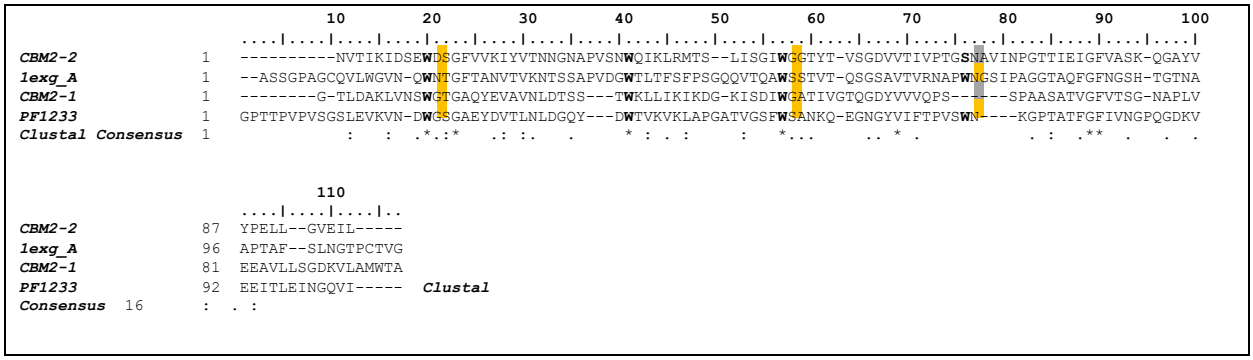
	CBM2-1 (1103-1199)	max	total	% query	% idenity	e-value	acession
1	Cel12E [unidentified prokaryotic organism]	50,4	50,4	0,97	0,37	0,00006	CRL09147.1
2	chitinase [Pyrococcus furiosus]	50,4	50,4	0,97	0,35	0,00007	WP_011012376.1
3	Chain A, Crystal Structure Of Chitin Biding Domain Of Chitinase From Pyrococcus furiosus	46,6	46,6	0,94	0,35	0,0002	2CWR_A
4	chitinase [Thermococcus kodakarensis]	47	47	0,9	0,34	0,0007	WP_048053867.1
5	glycoside hydrolase family 5 [Cyanobacterium aponinum PCC 10605]	47,4	47,4	0,73	0,35	0,0008	AFZ54640.1
6	chitinase [Thermococcus kodakaraensis]	47	47	0,9	0,34	0.001	BAA88380.1
7	hypothetical protein, partial [Thermococcus nautili]	44,7	44,7	0,9	0,29	0.006	WP_042693178.1
8	Chitinase [Thermococcus nautili]	44,7	44,7	0,9	0,29	0.007	AHL22682.1
9	Glycoside Hydrolase Family 5 candidate endoglucanase [Bifidobacterium thermophilum RBL67]	43,9	43,9	0,72	0,35	0.009	AGH41463.1
	CBM2-2 (1207-1303)	**max**	**total**	**% query**	**% idenity**	**e-value**	**acession**
1	chitinase [Mycobacterium kansasii]	68,9	135	0,93	0,36	2,00E-11	WP_036448276.1
2	hypothetical protein [Mycobacterium kansasii]	68,9	135	0,93	0,36	4,00E-11	WP_036391652.1
3	hypothetical protein [Mycobacterium kansasii]	68,6	135	0,93	0,36	4,00E-11	WP_042313822.1
4	hypothetical protein [Mycobacterium gastri]	67,4	131	0,93	0,33	1,00E-10	WP_036410849.1
5	hypothetical protein [Mycobacterium kansasii]	66,6	66,6	0,93	0,33	1,00E-10	WP_036444190.1
6	hypothetical protein [Mycobacterium kansasii]	63,2	63,2	0,93	0,34	1,00E-10	WP_042313524.1
7	hypothetical protein MGAST_01715 [Mycobacterium gastri 'Wayne']	67	131	0,93	0,33	2,00E-10	ETW25608.1
8	glycoside hydrolase [Calothrix sp. 336/3]	66,6	66,6	0,91	0,38	2,00E-10	WP_046815061.1
9	cellulose binding domain protein [Mycobacterium kansasii 824]	61,6	61,6	0,89	0,34	3E-09	ETZ98981.1
10	endoglucanase [Microcoleus sp. PCC 7113]	62,8	62,8	0,95	0,32	4E-09	WP_015184298.1
11	hypothetical protein [Mycobacterium kansasii]	62,8	62,8	0,9	0,34	6E-09	WP_023367572.1
12	hypothetical protein [Mycobacterium kansasii]	62,4	62,4	0,9	0,34	6E-09	WP_041327210.1
13	mannan endo-1,4-beta-mannosidase [Streptomyces xiamenensis]	61,2	61,2	0,9	0,37	1E-08	WP_046725132.1
14	chitinase [Mycobacterium kansasii]	61,2	61,2	0,89	0,34	1E-08	KEP39480.1
15	mannan endo-1,4-beta-mannosidase [Streptomyces sp. NRRL F-2890]	61,2	61,2	0,9	0,37	1E-08	WP_030731241.1
16	hypothetical protein MPTA5024_00600 [Microbispora sp. ATCC PTA-5024]	60,8	60,8	0,85	0,35	1E-08	ETK38093.1
17	chitinase [Mycobacterium kansasii]	57,8	57,8	0,9	0,3	1E-08	WP_036394018.1
18	hypothetical protein [Microbispora sp. ATCC PTA-5024]	60,5	60,5	0,88	0,34	2E-08	WP_036322646.1
19	chitinase [Mycobacterium liflandii]	60,8	60,8	0,9	0,31	2E-08	WP_041300546.1
20	hypothetical protein [Calothrix sp. PCC 7103]	60,5	60,5	0,98	0,33	2E-08	WP_035173330.1
21	PE-PGRS family protein [Mycobacterium marinum M]	60,5	118	0,9	0,31	3E-08	ACC42687.1
22	xyloglucanase [Streptomyces griseoflavus]	60,5	60,5	0,95	0,31	3E-08	WP_004922772.1
23	chitinase [Mycobacterium marinum]	60,5	118	0,93	0,31	3E-08	WP_041324889.1
24	hypothetical protein [Mycobacterium gastri]	60,5	60,5	0,93	0,33	4E-08	WP_036414736.1
25	hypothetical protein [Mycobacterium ulcerans]	60,5	117	0,9	0,31	4E-08	WP_011739252.1



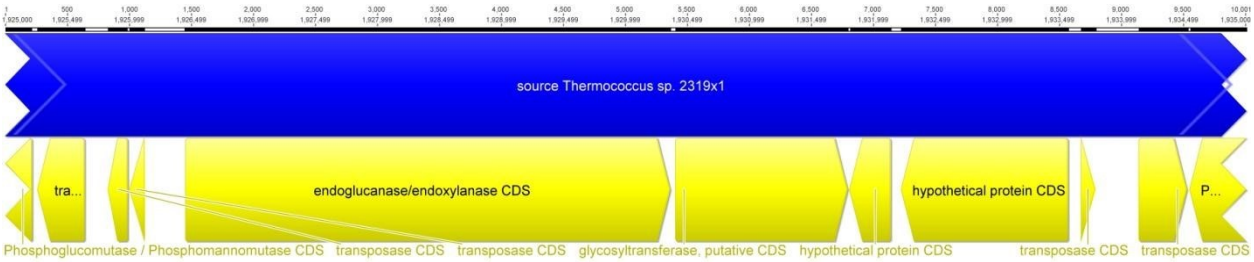
Supplementary Figure 8. Sequence Alignment of the MDG GH5 domain (residues 1 - 385) and the endo-1,4-β-glucanase PH1171 (EGPh) from *Pyrococcus horikoshii* (Ando et al. 2002; Kashima et al. 2005). All identical amino acid residues are marked with a star (*). Conserved residues that are similar to the catalytic cleft of PH1171 (PH1171: Arg102, His155, Asn200, Glu201, His297, Tyr299, Glu342 and Trp377) are highlighted in orange. Cysteins that are involved in the formation of a disulfide bond (PH1171: Cys106 – Cys159) are marked in blue. Alignments were performed with ClustalX and edited with BioEdit.



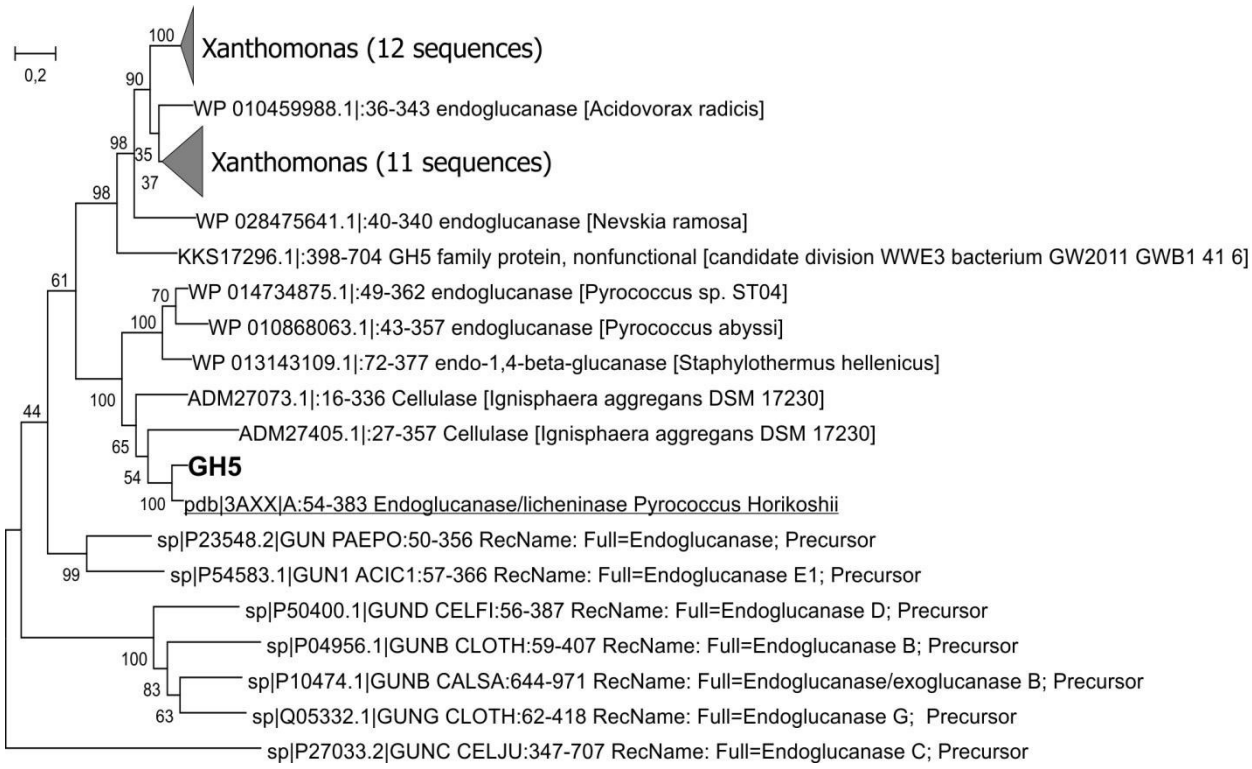
Supplementary Figure 9. Sequence Alignment of the MDG GH12-1 domain (residues 416 - 761), the MDG GH12-2 domain (residues 791 - 1070) (both including upstream linker regions) and the well characterized endo-1,4-β-glucanase PF0854 (EGPf) of *Pyrococcus furiosus* (Bauer et al. 1999). The nucleophile and the proton donor (PF0854: E197, E290, respectively) conserved in family 12 GHs are found in both GH12 domains of the MDG and are highlighted in green. The calcium-binding motif (DxDxDG) as well as the residues involved in metal ion coordination identified in PF0854 (PF0854: Asp68 , Asp70, Asp72, Asn74, Glu76 and Asp142) are found in the upstream linker region of GH12-2 (with low sequence similarity), except Asp142 which is located in the predicted GH12-2 domain. Residues for binding and coordination of Ca^{2+} are highlighted in red and orange, respecively. Alignments were performed with ClustalX and edited with BioEdit.



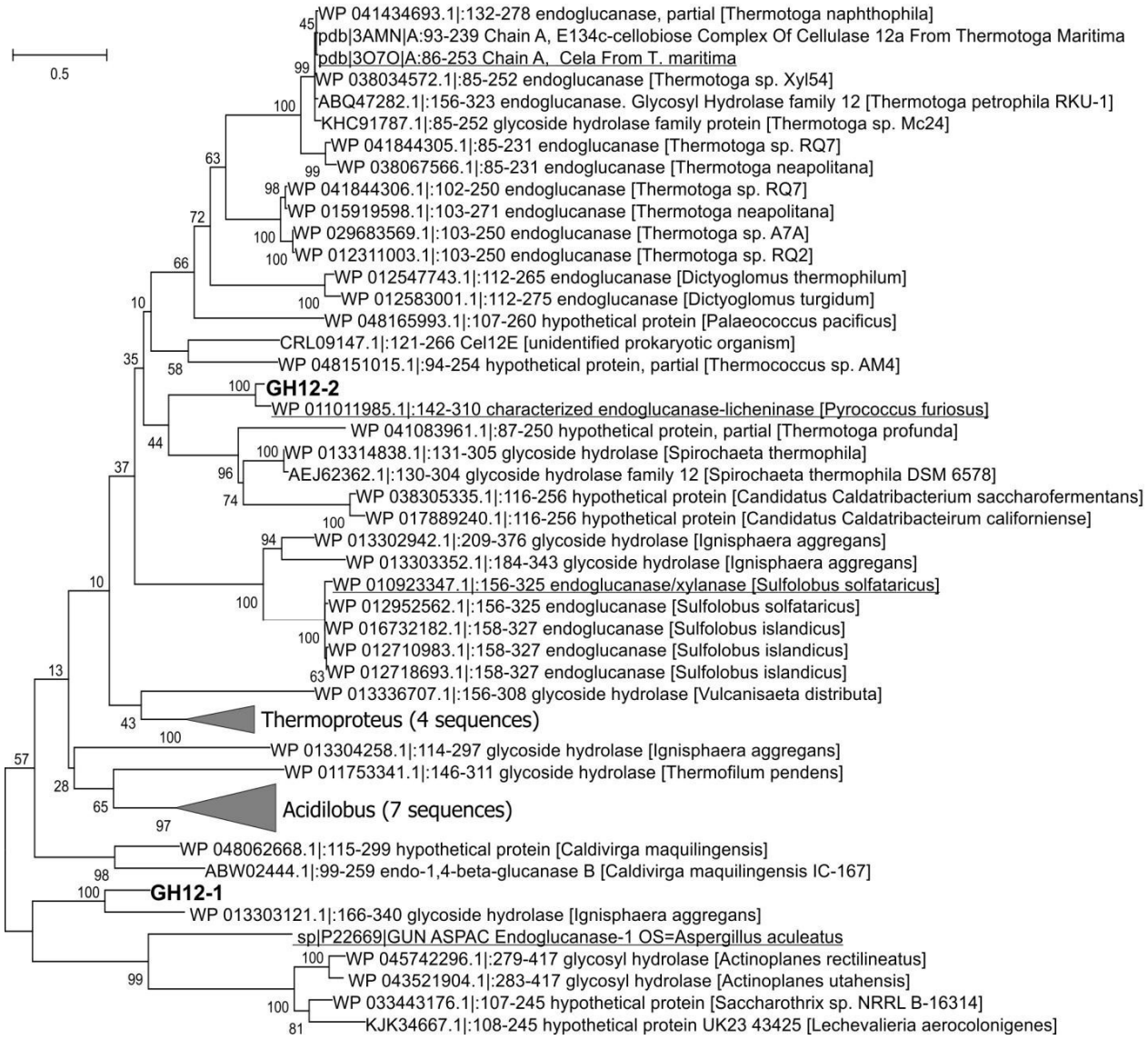
Supplementary Figure 10. Sequence Alignment of the MDG CBM2-1 domain (residues 1103 - 1199), the MDG CBM2-2 domain (residues 1208 - 1303), 1exg_A of *Cellulomonas fimi* and PF1233 of *Pyrococcus furiosus* (Xu et al. 1995; Nakamura et al. 2008; Simpson et al. 2000). Conserved surface-exposed tryptophans in CBM domains are highlighted in orange. The third conserved tryptophan residue is only present in 1exg_A of *Cellulomonas fimi* and PF1233 of *Pyrococcus furiosus*, whereas it is completely missing in the CBM2-1 domain and replaced by a serine in the CBM-2 domain of the MDG (differences are highlighted in grey). Alignments were performed with ClustalX and edited in BioEdit.



Supplementary Figure 11. The *mdg* gene is encoded in a gene cluster together with a putative glycosyltransferse and two hypothetical proteins in the genome of *Thermococcus* sp. strain 2319x1. The gene cluster is flanked by transposase encoding genes.



Supplementary Figure 12. Maximum likelihood phylogenetic tree of the GH5 domain of the novel multidomain glycosidase (MDG, ADU37_CDS22600) and the nearest relatives. The tree with the highest log likelihood (-9647.9039) is shown. The number of trees of totally 100 repetitions (bootstrap analysis) in which the associated taxa clustered together is shown next to the branches. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 40 amino acid sequences. All positions with less than 95% site coverage were eliminated. There were a total of 292 positions in the final dataset. Evolutionary analyses were conducted in MEGA6 (Tamura et al., 2013). Bar, 20 amino acid substitutions per 100. The *Thermococcus* sp. strain 2319x1 MDG GH5 domain is depicted in bold. The nearest characterized realtive is underlined.



Supplementary Figure 13. Maximum likelihood phylogenetic tree of the GH12 domains of the novel multifunctional, mulitdomain glycosidase (MDG, ADU37_CDS22600) and the nearest relatives. The tree with the highest log likelihood (-9585.6354) is shown. The number of trees of totally 100 repetitions (bootstrap analysis) in which the associated taxa clustered together is shown next to the branches. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 54 amino acid sequences. All positions with less than 95% site coverage were eliminated. There were a total of 125 positions in the final dataset. Evolutionary analyses were conducted in MEGA6 (Tamura et al., 2013). Bar, 50 amino acid substitutions per 100. The *Thermococcus* sp. strain 2319x1 MDG GH12-1 and GH12-2 domains are shown in bold. The nearest characterized relatives are underlined.

Supplementary Literature

Tamura K., Stecher G., Peterson D., Filipski A., and Kumar S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution* 30: 2725-2729.

4 Summary

The accumulation of compatible solutes is a prerequisite for the adaptation of microorganisms to respond to various stressors or to survive in extreme environments. The protection mechanisms of compatible solutes on biological macromolecules are still a relevant research field and is far from being thoroughly elucidated. The disaccharide trehalose is one of the most widespread compatible solutes in nature. The respective biosynthetic pathways and involved enzymes were analyzed in a large number of microorganisms. However, often the regulatory mechanisms and physiological functions are unknown so far. Although the thermoacidophilic Crenarchaeon *Sulfolobus acidocaldarius* belongs to the best-studied archaeal species up to date, only less is known about trehalose biosynthesis and its role in this organism.

In chapter 3.1, a comprehensive study on the trehalose metabolism in *S. acidocaldarius* was performed. Therefore, gene deletion mutants, growth studies, bioinformatics, and transcriptional analysis via qRT-PCR were accomplished. Results obtained from single and double gene deletion strains of already reported or predicted trehalose biosynthesis routes (Δ*treY* or Δ*treT* and Δ*treY*/Δ*treT*) still demonstrated trehalose formation and thus indicated a further trehalose biosynthesis pathway. Comprehensive bioinformatics revealed a novel TPS/TPP biosynthesis route for trehalose. The novel enzymes show no similarity to classical counterparts, the pathway comprising a TPS (trehalose-6P-synthase) of the GT4 enzyme family as well as a predicted TPP (trehalose-6P-phosphatase) of the HAD (haloacid dehalogenase) enzyme family. Furthermore, the physiological function of trehalose in salt stress response was confirmed. Accordingly, salt stressed wild type cells (250 mM NaCl) showed a 12-fold increased trehalose accumulation. In contrast, upon cold and heat shock (65°C and 83°C), no effect on trehalose biosynthesis was observed. The induction of trehalose synthesis pathways was confirmed on transcriptional level via qRT-PCR, and pathway activities in crude extracts of salt stressed cells. In conclusion, the results obtained in this study unraveled the complexity of the trehalose metabolism in *S. acidocaldarius* and demonstrate a function of trehalose as compatible solute upon salt stress. However, there is a growing interest in the utilization of compatible solutes for biotechnological and industrial purposes. The transfer of the protecting properties on biomolecules is a favorable approach in medical, cosmetics, and pharmaceutical research. Especially “extremolytes”, compatible solutes from extremophilic microorganisms, offer great potential for versatile applications. The diversity of extremolytes has increased since more (hyper)thermophilic Archaea have been investigated. However, in most cases, the biotechnological potential remains unused so far. In chapter 3.2, the *in vitro* production of the extremolyte cyclic-2,3-diphosphoglycerate (cDPG) was extensively studied. Although cDPG with its known protective properties has a high potential for industrial applications and the synthesis pathway is known, no effective production procedure has been developed so far. This is due to the low cell yield and challenging

cultivation conditions for methanogens under strict anaerobiosis and very high temperature hampering a whole-cell *in vivo* approach. Therefore, for the first time an easy biocatalytic *in vitro* one-step process with a straightforward expression and purification protocol using the codon-optimized cDPGS from *Methanothermus fervidus* was developed. The expression yield of recombinant cDPGS was significantly improved and the enzyme could be sufficiently stabilized for use and storage, finally, the complete substrate to product conversion was achieved. These findings will now enable the development of the larger scale enzymatic production of cDPG for application.

Besides the biotechnological relevance of extremolytes, likewise extremophiles itself as well as their enzymes are beneficial for biotransformations under hostile conditions. Especially the conversion of lignocellulosic biomass and polymeric substrates is of high industrial demand.

In chapter 3.3, a multi-layered approach was used to identify novel organisms and enzymes from extreme environments, that are capable of polymer degradation. Therefore, an *in situ* enrichment strategy, genomics, and comparative genomics, as well as cloning, expression, and biochemical characterization of novel enzymes were applied. Thus, a new xylan degrading hyperthermophilic euryarchaeon has been isolated, *Thermococcus* sp. strain 2319x1. Within the genome, a novel type of glycosidase, with a unique five-domain structure was identified. The multi domain glycosidase (MDG) is assembled of three glycoside hydrolase (GH) domains and two carbohydrate-binding modules (CBM) (GH5-12-12-CBM2-2 N-to C-terminal direction), which has not been reported for any other archaeal enzyme so far. The full-length protein, as well as several truncated versions, were analyzed in detail. The MDG as well as several truncated versions were analyzed in detail. The MDG was able to hydrolyze various polysaccharides, with the highest activity for barley β-glucan (β-1,3/1,4-glucoside) followed by that for CMC (β-1,4-glucoside), cello oligosaccharides, and galactomannan. The obtained results indicate that the modular MDG structure, with multiple GH- and CBM- domains not only extends the substrate spectrum but also seems to allow the degradation of partially soluble and insoluble polymers in a processive manner. The MDG and its truncated versions provides an innovative source for polymer degradation in industrial processes.

5 Zusammenfassung

Die Synthese und Anreicherung kompatibler Solute ist eine erfolgreiche Strategie zur Anpassung von Mikroorganismen, um auf verschiedene Stress- und Umweltfaktoren zu reagieren oder um in extremen Habitaten überleben zu können. Die schützenden Eigenschaften, die kompatible Solute auf biologische Makromoleküle ausüben, sind nach wie vor ein relevantes Forschungsfeld und sind noch nicht vollständig aufgeklärt. Eines der in der Natur am weitesten verbreitetsten kompatiblen Solute ist das Disaccharid Trehalose. Die jeweiligen Biosynthesewege und beteiligten Enzyme wurden in einer Vielzahl von Mikroorganismen untersucht, dennoch sind die Regulationsmechanismen und physiologischen Funktionen häufig nicht bekannt. Das trifft auch für das thermoacidophile Crenarchaeon *Sulfolobus acidocaldarius* zu, obwohl es zu den am besten untersuchten Archaeen gehört, ist über die Trehalose Biosynthese und seine Bedeutung für diesen Organismus nur wenig bekannt. Daher wurde in Kapitel 3.1 eine umfassende Studie zum Trehalose Metabolismus in *S. acidocaldarius* durchgeführt. Dazu wurden Gendeletionsmutanten, Wachstumsstudien, Bioinformatik und Transkriptionsanalyse mittels qRT-PCR durchgeführt. Die Untersuchung von verschiedenen Deletionsmutanten, der bereits bekannten oder vorhergesagten Trehalose Biosynthesewegen (Δ*treY* oder Δ*treT* und Δ*treY* / Δ*treT*) zeigte, dass nach wie vor Trehalose gebildet werden kann. Diese Ergebnisse lieferten die nötigen Hinweise auf einen weiteren bislang noch unbekannten Syntheseweg. Eine umfassende bioinformatische Analyse ergab einen neuartigen TPS/TPP-Biosyntheseweg für Trehalose. Der neu identifizierte Syntheseweg besteht aus einer TPS (Trehalose-6P-Synthase) der GT4-Enzymfamilie, sowie einer putativen TPP (Trehalose-6P-Phosphatase) der HAD-Enzymfamilie (Haloacid-Dehalogenase-Familie), beide Enzyme zeigen keine Ähnlichkeit zu den klassischen bereits charakterisierten Enzymen. Darüber hinaus konnte die physiologische Funktion von Trehalose als Reaktion auf Salzstress gezeigt werden, dementsprechend zeigten salzgestresste Wildtypzellen (250 mM NaCl) eine 12-fach erhöhte Trehaloseakkumulation, wohingegen bei Kälte- und Hitzeschock (65 ° C und 83 ° C) keine Auswirkung auf die Trehalose Biosynthese beobachtet werden konnte. Die Induktion der einzelnen Trehalose Synthesewege wurde auf Transkriptionsebene mittels qRT-PCR analysiert und zusätzlich wurden die Aktivitäten der einzelnen Biosynthesewege im Rohextrakt von salzgestressten Zellen bestätigt. Zusammenfassend tragen die Ergebnisse dieser Studie zur Aufklärung der Komplexität des Trehalose Metabolismus in *S. acidocaldarius* bei und zeigen erstmals die Funktion von Trehalose als kompatibles Solut unter Salzstress. An der biotechnologischen Anwendung von kompatiblen Soluten besteht ein erhebliches industrielles Interesse. Die Übertragung der protektiven Eigenschaften auf andere Biomoleküle ist ein vielversprechender Ansatz in der medizinischen, kosmetischen und pharmazeutischen Forschung. Insbesondere Extremolyte, kompatible Solute aus extremophilen

Mikroorganismen, bieten erhebliches industrielles und biotechnologisches Potenzial. Die Vielfalt der bekannten Extremolyte hat zugenommen, da in der Vergangenheit mehr (hyper)thermophile Archaeen untersucht wurden, dennoch bleibt in den meisten Fällen das Potenzial völlig ungenutzt. In Kapitel 3.2 wurde die *in vitro* Produktion des Extremolyt cDPG (cyclisches 2,3-Diphosphoglycerat) eingehend untersucht. Obwohl der Syntheseweg und die schützenden Eigenschaften von cDPG bereits untersucht wurden und es damit ein hohes Potenzial für industrielle Anwendungen besitzt, konnte weder ein effizienter Produktionsstamm oder Prozess etabliert werden. Dies ist auf die geringe Zellausbeute und die schwierigen Kultivierungsbedingungen von Methanogenen unter Ausschluss von Sauerstoff und extrem hohen Temperaturen zurückzuführen, die eine *in vivo* Ganzzellbiotransformation erheblich behindern und zu niedrigen Produktausbeuten führen. Im Rahmen dieser Arbeit wurde zum ersten Mal ein vielversprechendes *in vitro* Verfahren mit Hilfe eines einzigen Katalyse Schritts durchgeführt. Für dieses Verfahren wurde ein optimiertes Expressions- und Reinigungsprotokoll unter Verwendung der codon-optimierten cDPGS aus *Methanothermus fervidus* entwickelt. Dadurch konnte die Ausbeute der rekombinantem cDPGS signifikant verbessert werden, das Enzym konnte für die Anwendung und Lagerung ausreichend stabilisiert werden und schließlich konnte die vollständige Umwandlung vom Substrat in Produkt erreicht werden. Diese Ergebnisse bilden die Grundlage zur Entwicklung der enzymatischen Produktion von cDPG in größerem Maßstab, um die Anwendung zu ermöglichen. Neben der biotechnologischen Relevanz von Extremolyten bieten auch extremophile Mikroorganismen und ihre besonders stabilen Enzyme eine vorteilhafte Quelle für Biotransformation unter industriellen Bedingungen. Insbesondere die Umwandlung von lignocellulose-haltiger Biomasse und anderen polymeren Substraten ist von erheblicher Bedeutung.

In Kapitel 3.3 wurde ein mehrschichtiger Ansatz verwendet, um neuartige Organismen und Enzyme aus extremen Umgebungen zu identifizieren, die zum Polymerabbau fähig sind. Dazu wurden eine *In-situ*-Anreicherungsstrategie, Genomik und vergleichende Genomik, sowie die Klonierung, Expression und biochemische Charakterisierung neuartiger Enzyme durchgeführt. Mit Hilfe dieser Strategie wurde ein neues Xylan abbauendes hyperthermophiles Euryarchaeon isoliert, *Thermococcus* sp. Stamm 2319x1. Innerhalb des Genoms wurde eine neuartige Glycosidase mit einer einzigartigen Fünf-Domänen-Struktur identifiziert. Diese Multidomänen-Glycosidase (MDG) besteht aus drei Glycosidhydrolase (GH) -Domänen und zwei Kohlenhydratbindungsmodulen (CBM) (GH5-12-12-CBM2-2 N-zu-C-terminus). Das Protein voller Länge, sowie mehrere verkürzte Versionen wurden detailliert analysiert. Die MDG ist in der Lage verschiedene Polysaccharide zu hydrolysieren, die höchsten Aktivität wurden für Gersten-β-Glucan (β-1,3/1,4-Glucosid) und CMC (β-1,4-Glucosid) gemessen, gefolgt von Cellooligosacchariden und Galactomannan. Die Ergebnisse zeigen, dass die

modulare MDG-Struktur mit mehreren GH- und CBM-Domänen nicht nur das Substratspektrum erweitert, sondern auch den prozessiven Abbau von teilweise löslichen und unlöslichen Polymeren ermöglicht. Die MDG inklusive der verkürzten Versionen bietet eine innovative Quelle für den Polymerabbau in industriellen Prozessen.

6 References

1. Woese, C.R. and G.E. Fox, *Phylogenetic structure of the prokaryotic domain: The primary kingdoms.* Proceedings of the National Academy of Sciences, 1977. **74**(11): p. 5088.
2. Woese, C.R., O. Kandler, and M.L. Wheelis, *Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya.* Proceedings of the National Academy of Sciences, 1990. **87**(12): p. 4576.
3. DeLong, E., *Oceans of Archaea.* ASM News, 2003. **69**: p. 503-511.
4. Bang, C. and R.A. Schmitz, *Archaea associated with human surfaces: not to be underestimated.* FEMS Microbiology Reviews, 2015. **39**(5): p. 631-648.
5. Guy, L. and T. Ettema, *Guy L, Ettema TJG.. The archaeal 'TACK' superphylum and the origin of eukaryotes. Trends Microbiol 19: 580-587.* Trends in microbiology, 2011. **19**: p. 580-7.
6. Spang, A., et al., *Complex archaea that bridge the gap between prokaryotes and eukaryotes.* Nature, 2015. **521**(7551): p. 173-179.
7. Eme, L., et al., *Archaea and the origin of eukaryotes.* Nature Reviews Microbiology, 2017. **15**(12): p. 711-723.
8. Zaremba-Niedzwiedzka, K., et al., *Asgard archaea illuminate the origin of eukaryotic cellular complexity.* Nature, 2017. **541**(7637): p. 353-358.
9. Cox, C.J., et al., *The archaebacterial origin of eukaryotes.* Proceedings of the National Academy of Sciences of the United States of America, 2008. **105**(51): p. 20356-20361.
10. Williams, T., et al., *An archaeal origin of eukaryotes supports only two primary domains of life.* Nature, 2013. **504**: p. 231-236.
11. Raymann, K., C. Brochier-Armanet, and S. Gribaldo, *The two-domain tree of life is linked to a new root for the Archaea.* Proceedings of the National Academy of Sciences, 2015. **112**(21): p. 6670.
12. Grohmann, D. and F. Werner, *Recent advances in the understanding of archaeal transcription.* Current Opinion in Microbiology, 2011. **14**(3): p. 328-334.
13. Bell, S. and S. Jackson, *Bell SD, Jackson SP.. Mechanism and regulation of transcription in archaea. Curr Opin Microbiol 4: 208-213.* Current Opinion in Microbiology, 2001. **4**: p. 208-213.
14. Tan, C. and J. Tomkins, *Information Processing Differences Between Archaea and Eukaraya—Implications for Homologs and the Myth of Eukaryogenesis.* Answers Research Journal, 2015. **8**: p. 121–141.
15. Werner, F. and D. Grohmann, *Evolution of multisubunit RNA polymerases in the three domains of life.* Nature reviews. Microbiology, 2011. **9**: p. 85-98.
16. Esser, D., et al., *Protein phosphorylation and its role in archaeal signal transduction.* FEMS Microbiology Reviews, 2016. **40**: p. fuw020.
17. Kandler, O. and H. Konig, *Cell wall polymers in Archaea (Archaebacteria).* Cell Mol Life Sci, 1998. **54**(4): p. 305-8.
18. Sara, M. and U.B. Sleytr, *S-Layer proteins.* J Bacteriol, 2000. **182**(4): p. 859-68.
19. Golyshina, O.V., et al., *The novel extremely acidophilic, cell-wall-deficient archaeon Cuniculiplasma divulgatum gen. nov., sp. nov. represents a new family, Cuniculiplasmataceae fam. nov., of the order Thermoplasmatales.* International journal of systematic and evolutionary microbiology, 2016. **66**(1): p. 332-340.
20. De Rosa, M., A. Gambacorta, and A. Gliozzi, *Structure, biosynthesis, and physicochemical properties of archaebacterial lipids.* Microbiol Rev, 1986. **50**(1): p. 70-80.
21. Damste, J.S., et al., *Crenarchaeol: the characteristic core glycerol dibiphytanyl glycerol tetraether membrane lipid of cosmopolitan pelagic crenarchaeota.* J Lipid Res, 2002. **43**(10): p. 1641-51.
22. 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.

23. Deppenmeier, U., *The unique biochemistry of methanogenesis.* Prog Nucleic Acid Res Mol Biol, 2002. **71**: p. 223-83.
24. Thauer, R.K., et al., *Methanogenic archaea: ecologically relevant differences in energy conservation.* Nature reviews. Microbiology, 2008. **6**(8): p. 579-591.
25. Sato, T. and H. Atomi, *Novel metabolic pathways in Archaea.* Current Opinion in Microbiology, 2011. **14**(3): p. 307-314.
26. Siebers, B. and P. Schönheit, *Unusual pathways and enzymes of central carbohydrate metabolism in Archaea.* Current Opinion in Microbiology, 2005. **8**(6): p. 695-705.
27. Siebers, B., H.P. Klenk, and R. Hensel, *PPi-dependent phosphofructokinase from Thermoproteus tenax, an archaeal descendant of an ancient line in phosphofructokinase evolution.* Journal of bacteriology, 1998. **180**(8): p. 2137-2143.
28. Kengen, S., et al., *ADP-dependent glucokinase and phosphofructokinase from Pyrococcus furiosus.* Methods in enzymology, 2001. **331**: p. 41-53.
29. Verhees, C.H., et al., *ADP-dependent phosphofructokinases in mesophilic and thermophilic methanogenic archaea.* Journal of bacteriology, 2001. **183**(24): p. 7145-7153.
30. Lorentzen, E., et al., *Structural Basis of Allosteric Regulation and Substrate Specificity of the Non-Phosphorylating Glyceraldehyde 3-Phosphate Dehydrogenase from Thermoproteus tenax.* Journal of Molecular Biology, 2004. **341**(3): p. 815-828.
31. Reher, M., S. Gebhard, and P. Schönheit, *Glyceraldehyde-3-phosphate ferredoxin oxidoreductase (GAPOR) and nonphosphorylating glyceraldehyde-3-phosphate dehydrogenase (GAPN), key enzymes of the respective modified Embden–Meyerhof pathways in the hyperthermophilic crenarchaeota Pyrobaculum aerophilum and Aeropyrum pernix.* FEMS Microbiology Letters, 2007. **273**(2): p. 196-205.
32. Ettema, T.J.G., et al., *The non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase (GAPN) of Sulfolobus solfataricus: a key-enzyme of the semi-phosphorylative branch of the Entner–Doudoroff pathway.* Extremophiles, 2008. **12**(1): p. 75-88.
33. Soderberg, T., *Biosynthesis of ribose-5-phosphate and erythrose-4-phosphate in archaea: a phylogenetic analysis of archaeal genomes.* Archaea (Vancouver, B.C.), 2005. **1**(5): p. 347-352.
34. Brouns, S., et al., *Identification of the missing links in prokaryotic pentose oxidation pathways: Evidence for enzyme recruitment.* The Journal of biological chemistry, 2006. **281**: p. 27378-88.
35. Sutter, J.-M., et al., *Pentose degradation in archaea: Halorhabdus species degrade D-xylose, L-arabinose and D-ribose via bacterial-type pathways.* Extremophiles, 2020. **24**(5): p. 759-772.
36. Littlechild, J.A., *Archaeal Enzymes and Applications in Industrial Biocatalysts.* Archaea, 2015. **2015**: p. 147671.
37. Quehenberger, J., et al., *Sulfolobus – A Potential Key Organism in Future Biotechnology.* Frontiers in Microbiology, 2017. **8**: p. 2474.
38. Schocke, L., C. Brasen, and B. Siebers, *Thermoacidophilic Sulfolobus species as source for extremozymes and as novel archaeal platform organisms.* Curr Opin Biotechnol, 2019. **59**: p. 71-77.
39. Suzuki, Y., K. Miyamoto, and H. Ohta, *A novel thermostable esterase from the thermoacidophilic archaeon Sulfolobus tokodaii strain 7.* FEMS Microbiol Lett, 2004. **236**(1): p. 97-102.
40. Park, S.J., et al., *Metagenome microarray for screening of fosmid clones containing specific genes.* FEMS Microbiol Lett, 2008. **284**(1): p. 28-34.
41. Gogliettino, M., et al., *A new pepstatin-insensitive thermopsin-like protease overproduced in peptide-rich cultures of Sulfolobus solfataricus.* Int J Mol Sci, 2014. **15**(2): p. 3204-19.
42. Brock, T.D., et al., *Sulfolobus: a new genus of sulfur-oxidizing bacteria living at low pH and high temperature.* Arch Mikrobiol, 1972. **84**(1): p. 54-68.

43. Bohlool, B.B., *Occurrence of Sulfolobus acidocaldarius, an extremely thermophilic acidophilic bacterium, in New Zealand hot springs.* Archives of Microbiology, 1975. **106**(3): p. 171-174.
44. Kurosawa, N., et al., *Characterization and identification of thermoacidophilic archaebacteria isolated in Japan.* Journal of General and Applied Microbiology - J GEN APPL MICROBIOL TOKYO, 1995. **41**: p. 43-52.
45. Mao, D. and D. Grogan, *Genomic evidence of rapid, global-scale gene flow in a Sulfolobus species.* Isme j, 2012. **6**(8): p. 1613-6.
46. Anderson, R.E., et al., *Structured Populations of Sulfolobus acidocaldarius with Susceptibility to Mobile Genetic Elements.* Genome biology and evolution, 2017. **9**(6): p. 1699-1710.
47. Fuchs, T., et al., *Metallosphaera prunae, sp. nov., a Novel Metal-mobilizing, Thermoacidophilic Archaeum, Isolated from a Uranium Mine in Germany.* Systematic and Applied Microbiology, 1995. **18**(4): p. 560-566.
48. Don J. Brenner, N.R.K.G.M.G.J.T.S.D.R.B.P.V.M.G.F.A.R.K.H.S., et al., *Bergey's Manual® of Systematic Bacteriology*. 2001: Springer.
49. Albers, S.V. and K.F. Jarrell, *The archaellum: how Archaea swim.* Front Microbiol, 2015. **6**: p. 23.
50. Albers, S.V. and K.F. Jarrell, *The Archaellum: An Update on the Unique Archaeal Motility Structure.* Trends Microbiol, 2018. **26**(4): p. 351-362.
51. Gabriel, J.L. and P.L. Chong, *Molecular modeling of archaebacterial bipolar tetraether lipid membranes.* Chem Phys Lipids, 2000. **105**(2): p. 193-200.
52. Bischof, L.F., et al., *Early Response of Sulfolobus acidocaldarius to Nutrient Limitation.* Front Microbiol, 2018. **9**: p. 3201.
53. Zillig, W., et al., *The Sulfolobus-"Caldariella" group: Taxonomy on the basis of the structure of DNA-dependent RNA polymerases.* Archives of Microbiology, 1980. **125**(3): p. 259-269.
54. Sakai, H.D. and N. Kurosawa, *Saccharolobus caldissimus gen. nov., sp. nov., a facultatively anaerobic iron-reducing hyperthermophilic archaeon isolated from an acidic terrestrial hot spring, and reclassification of Sulfolobus solfataricus as Saccharolobus solfataricus comb. nov. and Sulfolobus shibatae as Saccharolobus shibatae comb. nov.* International Journal of Systematic and Evolutionary Microbiology, 2018. **68**(4): p. 1271-1278.
55. Zhang, C., et al., *The essential genome of the crenarchaeal model Sulfolobus islandicus.* Nature Communications, 2018. **9**(1): p. 4908.
56. Albers, S.-V., et al., *Production of recombinant and tagged proteins in the hyperthermophilic archaeon Sulfolobus solfataricus.* Applied and Environmental Microbiology, 2006. **72**(1): p. 102–111.
57. Berkner, S., et al., *Small multicopy, non-integrative shuttle vectors based on the plasmid pRN1 for Sulfolobus acidocaldarius and Sulfolobus solfataricus, model organisms of the (cren-)archaea.* Nucleic Acids Res, 2007. **35**(12): p. e88.
58. Berkner, S., et al., *Inducible and constitutive promoters for genetic systems in Sulfolobus acidocaldarius.* Extremophiles : life under extreme conditions, 2010. **14**(3): p. 249–259.
59. Deng, L., et al., *Unmarked gene deletion and host-vector system for the hyperthermophilic crenarchaeon Sulfolobus islandicus.* Extremophiles : life under extreme conditions, 2009. **13**(4): p. 735–746.
60. Grogan, D.W., *Homologous recombination in Sulfolobus acidocaldarius: genetic assays and functional properties.* Biochemical Society transactions, 2009. **37**(Pt 1): p. 88–91.
61. Wagner, M., et al., *Expanding and understanding the genetic toolbox of the hyperthermophilic genus Sulfolobus.* Biochemical Society transactions, 2009. **37**(Pt 1): p. 97–101.
62. Ulas, T., et al., *Genome-scale reconstruction and analysis of the metabolic network in the hyperthermophilic archaeon Sulfolobus solfataricus.* PLoS One, 2012. **7**(8): p. e43401.

63. Kort, J.C., et al., *A cool tool for hot and sour Archaea: proteomics of Sulfolobus solfataricus.* Proteomics, 2013. **13**(18-19): p. 2831-50.
64. Albers, S.V. and B.H. Meyer, *The archaeal cell envelope.* Nat Rev Microbiol, 2011. **9**(6): p. 414-26.
65. Elferink, M.G., et al., *Sugar transport in Sulfolobus solfataricus is mediated by two families of binding protein-dependent ABC transporters.* Mol Microbiol, 2001. **39**(6): p. 1494-503.
66. Izzo, V., et al., *The thermophilic archaeon Sulfolobus solfataricus is able to grow on phenol.* Res Microbiol, 2005. **156**(5-6): p. 677-89.
67. Stark, H., et al., *Oxidative Stickland reactions in an obligate aerobic organism - amino acid catabolism in the Crenarchaeon Sulfolobus solfataricus.* FEBS J, 2017. **284**(13): p. 2078-2095.
68. Brugger, K., et al., *Mobile elements in archaeal genomes.* FEMS Microbiol Lett, 2002. **206**(2): p. 131-41.
69. Grogan, D.W., *Phenotypic characterization of the archaebacterial genus Sulfolobus: comparison of five wild-type strains.* J Bacteriol, 1989. **171**(12): p. 6710-9.
70. Chen, L., et al., *The genome of Sulfolobus acidocaldarius, a model organism of the Crenarchaeota.* Journal of bacteriology, 2005. **187**(14): p. 4992–4999.
71. Joshua, C.J., et al., *Absence of diauxie during simultaneous utilization of glucose and Xylose by Sulfolobus acidocaldarius.* J Bacteriol, 2011. **193**(6): p. 1293-301.
72. Orita, I., et al., *The ribulose monophosphate pathway substitutes for the missing pentose phosphate pathway in the archaeon Thermococcus kodakaraensis.* Journal of bacteriology, 2006. **188**(13): p. 4698-4704.
73. Nishimasu, H., et al., *Identification and Characterization of an ATP-Dependent Hexokinase with Broad Substrate Specificity from the Hyperthermophilic Archaeon <em>Sulfolobus tokodaii</em&gt.* Journal of Bacteriology, 2006. **188**(5): p. 2014.
74. Kouril, T., et al., *Unraveling the function of the two Entner-Doudoroff branches in the thermoacidophilic Crenarchaeon Sulfolobus solfataricus P2.* FEBS J, 2013. **280**(4): p. 1126-38.
75. Nicolaus, B., et al., *Trehalose in Archaebacteria.* Systematic and Applied Microbiology, 1988. **10**(3): p. 215-217.
76. Maruta, K., et al., *Cloning and sequencing of a cluster of genes encoding novel enzymes of trehalose biosynthesis from thermophilic archaebacterium Sulfolobus acidocaldarius.* Biochim Biophys Acta, 1996. **1291**(3): p. 177-81.
77. Nakada, T., et al., *Purification and characterization of thermostable maltooligosyl trehalose synthase from the thermoacidophilic archaebacterium Sulfolobus acidocaldarius.* Biosci Biotechnol Biochem, 1996. **60**(2): p. 263-6.
78. Brown, A.D., *Compatible solutes and extreme water stress in eukaryotic micro-organisms.* Advances in microbial physiology, 1978. **17**: p. 181–242.
79. Empadinhas, N. and M.S. da Costa, *Osmoadaptation mechanisms in prokaryotes: distribution of compatible solutes.* Int Microbiol, 2008. **11**(3): p. 151-61.
80. Empadinhas, N. and M.S. da Costa, *Diversity and biosynthesis of compatible solutes in hyper/thermophiles.* Int Microbiol, 2006. **9**(3): p. 199-206.
81. Elbein, A.D., et al., *New insights on trehalose: a multifunctional molecule.* Glycobiology, 2003. **13**(4): p. 17r-27r.
82. Arguelles, J.C., *Physiological roles of trehalose in bacteria and yeasts: a comparative analysis.* Arch Microbiol, 2000. **174**(4): p. 217-24.
83. Paul, M.J., et al., *Trehalose metabolism and signaling.* Annu Rev Plant Biol, 2008. **59**: p. 417-41.
84. Arguelles, J.C., *Why can't vertebrates synthesize trehalose?* J Mol Evol, 2014. **79**(3-4): p. 111-6.
85. Avonce, N., et al., *Insights on the evolution of trehalose biosynthesis.* BMC Evol Biol, 2006. **6**: p. 109.
86. Iturriaga, G., R. Suarez, and B. Nova-Franco, *Trehalose metabolism: from osmoprotection to signaling.* Int J Mol Sci, 2009. **10**(9): p. 3793-810.

87. Makihara, F., et al., *Role of trehalose synthesis pathways in salt tolerance mechanism of Rhodobacter sphaeroides f. sp. denitrificans IL106.* Arch Microbiol, 2005. **184**(1): p. 56-65.
88. Tzvetkov, M., et al., *Genetic dissection of trehalose biosynthesis in Corynebacterium glutamicum: inactivation of trehalose production leads to impaired growth and an altered cell wall lipid composition.* Microbiology, 2003. **149**(Pt 7): p. 1659-73.
89. Martins, L.O., et al., *Organic solutes in hyperthermophilic archaea.* Appl Environ Microbiol, 1997. **63**(3): p. 896-902.
90. Desmarais, D., et al., *2-Sulfotrehalose, a novel osmolyte in haloalkaliphilic archaea.* J Bacteriol, 1997. **179**(10): p. 3146-53.
91. Youssef, N.H., et al., *Trehalose/2-sulfotrehalose biosynthesis and glycine-betaine uptake are widely spread mechanisms for osmoadaptation in the Halobacteriales.* Isme j, 2014. **8**(3): p. 636-49.
92. Kouril, T., et al., *A novel trehalose synthesizing pathway in the hyperthermophilic Crenarchaeon Thermoproteus tenax: the unidirectional TreT pathway.* Arch Microbiol, 2008. **190**(3): p. 355-69.
93. Rao, K.N., et al., *Crystal structure of trehalose-6-phosphate phosphatase-related protein: biochemical and biological implications.* Protein Sci, 2006. **15**(7): p. 1735-44.
94. Kaasen, I., J. McDougall, and A.R. Strom, *Analysis of the otsBA operon for osmoregulatory trehalose synthesis in Escherichia coli and homology of the OtsA and OtsB proteins to the yeast trehalose-6-phosphate synthase/phosphatase complex.* Gene, 1994. **145**(1): p. 9-15.
95. Kaasen, I., J. McDougall, and A.R. Strøm, *Analysis of the otsBA operon for osmoregulatory trehalose synthesis in Escherichia coli and homology of the OtsA and OtsB proteins to the yeast trehalose-6-phosphate synthase/phosphatase complex.* Gene, 1994. **145**(1): p. 9-15.
96. Avonce, N., et al., *The Cytophaga hutchinsonii ChTPSP: First Characterized Bifunctional TPS–TPP Protein as Putative Ancestor of All Eukaryotic Trehalose Biosynthesis Proteins.* Molecular Biology and Evolution, 2010. **27**(2): p. 359-369.
97. Strom, A.R. and I. Kaasen, *Trehalose metabolism in Escherichia coli: stress protection and stress regulation of gene expression.* Mol Microbiol, 1993. **8**(2): p. 205-10.
98. Bell, W., et al., *Composition and functional analysis of the Saccharomyces cerevisiae trehalose synthase complex.* J Biol Chem, 1998. **273**(50): p. 33311-9.
99. Zaparty, M., et al., *The central carbohydrate metabolism of the hyperthermophilic crenarchaeote Thermoproteus tenax: Pathways and insights into their regulation.* Archives of Microbiology, 2008. **190**(3): p. 231-245.
100. Zaparty, M., et al., *The first prokaryotic trehalose synthase complex identified in the hyperthermophilic crenarchaeon Thermoproteus tenax.* PLoS One, 2013. **8**(4): p. e61354.
101. Gao, Y., et al., *Enzymatic and regulatory properties of the trehalose-6-phosphate synthase from the thermoacidophilic archaeon Thermoplasma acidophilum.* Biochimie, 2014. **101**: p. 215-20.
102. Blazquez, M.A., et al., *Isolation and molecular characterization of the Arabidopsis TPS1 gene, encoding trehalose-6-phosphate synthase.* Plant J, 1998. **13**(5): p. 685-9.
103. Mukai, K., et al., *Production of Trehalose from Starch by Thermostable Enzymes from Sulfolobus acidocaldarius.* Starch - Stärke, 1997. **49**(1): p. 26-30.
104. Gueguen, Y., et al., *Characterization of the maltooligosyl trehalose synthase from the thermophilic archaeon Sulfolobus acidocaldarius.* FEMS Microbiol Lett, 2001. **194**(2): p. 201-6.
105. Fang, T.Y., et al., *Characterization of the trehalosyl dextrin-forming enzyme from the thermophilic archaeon Sulfolobus solfataricus ATCC 35092.* Extremophiles, 2004. **8**(4): p. 335-43.
106. Nakada, T., et al., *Purification and characterization of thermostable maltooligosyl trehalose trehalohydrolase from the thermoacidophilic archaebacterium Sulfolobus acidocaldarius.* Biosci Biotechnol Biochem, 1996. **60**(2): p. 267-70.

107. Fang, T.Y., et al., *Expression, purification, and characterization of the maltooligosyltrehalose trehalohydrolase from the thermophilic archaeon Sulfolobus solfataricus ATCC 35092.* J Agric Food Chem, 2006. **54**(19): p. 7105-12.
108. Di Lernia, I., et al., *Enzymes from Sulfolobus shibatae for the production of trehalose and glucose from starch.* Extremophiles, 1998. **2**(4): p. 409-16.
109. Seo, J.S., et al., *Bifunctional recombinant fusion enzyme between maltooligosyltrehalose synthase and maltooligosyltrehalose trehalohydrolase of thermophilic microorganism Metallosphaera hakonensis.* J Microbiol Biotechnol, 2008. **18**(9): p. 1544-9.
110. Seo, J.S., et al., *Molecular cloning and characterization of trehalose biosynthesis genes from hyperthermophilic archaebacterium Metallosphaera hakonesis.* J Microbiol Biotechnol, 2007. **17**(1): p. 123-9.
111. Woo, E.-J., et al., *Structural insight into the bifunctional mechanism of the glycogen-debranching enzyme TreX from the archaeon Sulfolobus solfataricus.* The Journal of biological chemistry, 2008. **283**(42): p. 28641-28648.
112. Nguyen, D.H., et al., *The Reaction Kinetics and the Effect of Substrate Transglycosylation Catalyzed by TreX of Sulfolobus solfataricus on Glycogen Breakdown.* J Bacteriol, 2014.
113. Maruta, K., et al., *Cloning and sequencing of trehalose biosynthesis genes from Arthrobacter sp. Q36.* Biochimica et Biophysica Acta (BBA) - General Subjects, 1996. **1289**(1): p. 10-13.
114. Carpinelli, J., R. Kramer, and E. Agosin, *Metabolic engineering of Corynebacterium glutamicum for trehalose overproduction: role of the TreYZ trehalose biosynthetic pathway.* Appl Environ Microbiol, 2006. **72**(3): p. 1949-55.
115. Sugawara, M., E.J. Cytryn, and M.J. Sadowsky, *Functional role of Bradyrhizobium japonicum trehalose biosynthesis and metabolism genes during physiological stress and nodulation.* Appl Environ Microbiol, 2010. **76**(4): p. 1071-81.
116. De Smet, K.A.L., et al., *Three pathways for trehalose biosynthesis in mycobacteria.* Microbiology, 2000. **146 (Pt 1)**: p. 199-208.
117. Worning, P., et al., *Structural analysis of DNA sequence: evidence for lateral gene transfer in Thermotoga maritima.* Nucleic Acids Res, 2000. **28**(3): p. 706-9.
118. Qu, Q., S.J. Lee, and W. Boos, *TreT, a novel trehalose glycosyltransferring synthase of the hyperthermophilic archaeon Thermococcus litoralis.* J Biol Chem, 2004. **279**(46): p. 47890-7.
119. Ryu, S.I., et al., *A novel trehalose-synthesizing glycosyltransferase from Pyrococcus horikoshii: molecular cloning and characterization.* Biochem Biophys Res Commun, 2005. **329**(2): p. 429-36.
120. Ryu, S.I., et al., *Catalytic reversibility of Pyrococcus horikoshii trehalose synthase: Efficient synthesis of several nucleoside diphosphate glucoses with enzyme recycling.* Process Biochemistry, 2011. **46**(1): p. 128-134.
121. Siebers, B., et al., *Reconstruction of the central carbohydrate metabolism of Thermoproteus tenax by use of genomic and biochemical data.* J Bacteriol, 2004. **186**(7): p. 2179-94.
122. Diruggiero, J., et al., *Evidence of recent lateral gene transfer among hyperthermophilic archaea.* Mol Microbiol, 2000. **38**(4): p. 684-93.
123. Xavier, K.B., et al., *High-affinity maltose/trehalose transport system in the hyperthermophilic archaeon Thermococcus litoralis.* J Bacteriol, 1996. **178**(16): p. 4773-7.
124. Lamosa, P., et al., *Effects of temperature, salinity, and medium composition on compatible solute accumulation by thermococcus spp.* Appl Environ Microbiol, 1998. **64**(10): p. 3591-8.
125. Nishimoto, T., et al., *Purification and Characterization of a Thermostable Trehalose Synthase from Thermus aquaticus*. Vol. 60. 1996. 835-839.
126. Koh, S.K., et al., *Trehalose synthesis from maltose by a thermostable trehalose synthase from Thermus caldophilus.* Biotechnology Letters, 1998. **20**(8): p. 757-761.

127. Freeman Brian, C., C. Chen, and A. Beattie Gwyn, *Identification of the trehalose biosynthetic loci of Pseudomonas syringae and their contribution to fitness in the phyllosphere.* Environmental Microbiology, 2010. **12**(6): p. 1486-1497.
128. Fütterer, O., et al., *Genome sequence of Picrophilus torridus and its implications for life around pH 0.* Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(24): p. 9091.
129. Chen, Y.S., G.C. Lee, and J.F. Shaw, *Gene cloning, expression, and biochemical characterization of a recombinant trehalose synthase from Picrophilus torridus in Escherichia coli.* J Agric Food Chem, 2006. **54**(19): p. 7098-104.
130. Chou, H.-H., et al., *Site-directed mutagenesis improves the thermostability of a recombinant Picrophilus torridus trehalose synthase and efficiency for the production of trehalose from sweet potato starch.* Food Chemistry, 2010. **119**(3): p. 1017-1022.
131. Hagemann, *Trehalose Metabolismus in (Hyper-) thermophilen Archaea*, in *MEB*. 2013, Universität Duisburg Essen.
132. Ren, Y., et al., *Gene expression and molecular characterization of a thermostable trehalose phosphorylase from Thermoanaerobacter tengcongensis*. Vol. 48. 2005. 221-7.
133. Maruta, K., et al., *Gene Encoding a Trehalose Phosphorylase from Thermoanaerobacter brockii ATCC 35047*. Vol. 66. 2002. 1976-80.
134. Van der Borght, J., et al., *Enzymatic properties and substrate specificity of the trehalose phosphorylase from Caldanaerobacter subterraneus.* Appl Environ Microbiol, 2011. **77**(19): p. 6939-44.
135. Ren, Y., et al., *Gene expression and molecular characterization of a thermostable trehalose phosphorylase from Thermoanaerobacter tengcongensis.* Sci China C Life Sci, 2005. **48**(3): p. 221-7.
136. Aisaka, K.M., Tomomi, *Production of trehalose phosphorylase by Catellatospora ferruginea.* FEMS Microbiology Letters, Volume 131, Issue 1, 1 August 1995, Pages 47–51, https://doi.org/10.1111/j.1574-6968.1995.tb07752.x, 1995.
137. Wannet, W.J., et al., *HPLC detection of soluble carbohydrates involved in mannitol and trehalose metabolism in the edible mushroom Agaricus bisporus.* J Agric Food Chem, 2000. **48**(2): p. 287-91.
138. Wannet, W.J.B., et al., *Purification and characterization of trehalose phosphorylase from the commercial mushroom Agaricus bisporus.* Biochimica et Biophysica Acta (BBA) - General Subjects, 1998. **1425**(1): p. 177-188.
139. Goedl, C., et al., *Structure-function relationships for Schizophyllum commune trehalose phosphorylase and their implications for the catalytic mechanism of family GT-4 glycosyltransferases.* Biochem J, 2006. **397**(3): p. 491-500.
140. Desmet, T. and W. Soetaert, *Enzymatic glycosyl transfer: Mechanisms and applications.* Biocatalysis and Biotransformation, 2011. **29**: p. 1-18.
141. Belocopitow, E. and L.R. Maréchal, *Trehalose phosphorylase from Euglena gracilis.* Biochimica et Biophysica Acta (BBA) - Enzymology, 1970. **198**(1): p. 151-154.
142. Salminen, S.O. and J.G. Streeter, *Enzymes of alpha,alpha-Trehalose Metabolism in Soybean Nodules.* Plant Physiol, 1986. **81**(2): p. 538-41.
143. Kizawa, H., K.-i. Miyagawa, and Y. Sugiyama, *Purification and Characterization of Trehalose Phosphorylase from Micrococcus varians.* Bioscience, Biotechnology, and Biochemistry, 1995. **59**(10): p. 1908-1912.
144. Luley-Goedl, C. and B. Nidetzky, *Carbohydrate synthesis by disaccharide phosphorylases: reactions, catalytic mechanisms and application in the glycosciences.* Biotechnol J, 2010. **5**(12): p. 1324-38.
145. Yoneyama, Y. and J.E. Lever, *Apical trehalase expression associated with cell patterning after inducer treatment of LLC-PK1 monolayers.* J Cell Physiol, 1987. **131**(3): p. 330-41.
146. Dahlqvist, A., *Assay of intestinal disaccharidases.* Scand J Clin Lab Invest, 1984. **44**(2): p. 169-72.
147. Horlacher, R., et al., *Characterization of a cytoplasmic trehalase of Escherichia coli.* J Bacteriol, 1996. **178**(21): p. 6250-7.

148. Mittenbuhler, K. and H. Holzer, *Purification and characterization of acid trehalase from the yeast suc2 mutant.* J Biol Chem, 1988. **263**(17): p. 8537-43.
149. Tourinho-dos-Santos, C.F., et al., *Periplasmic trehalase from Escherichia coli--characterization and immobilization on spherisorb.* Braz J Med Biol Res, 1994. **27**(3): p. 627-36.
150. Sakaguchi, M., et al., *Identification of GH15 Family Thermophilic Archaeal Trehalases That Function within a Narrow Acidic-pH Range.* Appl Environ Microbiol, 2015. **81**(15): p. 4920-31.
151. Maicas, S., J.P. Guirao-Abad, and J.C. Arguelles, *Yeast trehalases: Two enzymes, one catalytic mission.* Biochim Biophys Acta, 2016. **1860**(10): p. 2249-54.
152. Londesborough, J. and K. Varimo, *Characterization of two trehalases in baker's yeast.* Biochem J, 1984. **219**(2): p. 511-8.
153. Thevelein, J.M. and M. Beullens, *Cyclic AMP and the stimulation of trehalase activity in the yeast Saccharomyces cerevisiae by carbon sources, nitrogen sources and inhibitors of protein synthesis.* J Gen Microbiol, 1985. **131**(12): p. 3199-209.
154. Sussman, A.S., et al., *Isolation, mapping, and characterization of trehalaseless mutants of Neurospora crassa.* J Bacteriol, 1971. **108**(1): p. 59-68.
155. Nwaka, S., B. Mechler, and H. Holzer, *Deletion of the ATH1 gene in Saccharomyces cerevisiae prevents growth on trehalose.* FEBS Letters, 1996. **386**(2): p. 235-238.
156. d'Enfert, C. and T. Fontaine, *Molecular characterization of the Aspergillus nidulans treA gene encoding an acid trehalase required for growth on trehalose.* Mol Microbiol, 1997. **24**(1): p. 203-16.
157. Malaisse, W.J., et al., *The stimulus-secretion coupling of glucose-induced insulin release. Insulin release due to glycogenolysis in glucose-deprived islets.* Biochemical Journal, 1977. **164**(2): p. 447-454.
158. Moon, J.H., et al., *Characterization of a trehalose-degrading enzyme from the hyperthermophilic archaeon Sulfolobus acidocaldarius.* J Biosci Bioeng, 2016. **122**(1): p. 47-51.
159. Yuasa, M., et al., *Two trehalose-hydrolyzing enzymes from Crenarchaeon Sulfolobus acidocaldarius exhibit distinct activities and affinities toward trehalose.* Appl Microbiol Biotechnol, 2018. **102**(10): p. 4445-4455.
160. Lee, J., et al., *Saci_1816: A trehalase that catalyzestrehalose degradation in the thermoacidophilic crenarchaeon Sulfolobus acidocaldarius.* J Microbiol Biotechnol, 2018.
161. Silva, Z., et al., *Osmotic adaptation of Thermus thermophilus RQ-1: lesson from a mutant deficient in synthesis of trehalose.* J Bacteriol, 2003. **185**(20): p. 5943-52.
162. Silva, Z., S. Alarico, and M.S. da Costa, *Trehalose biosynthesis in Thermus thermophilus RQ-1: biochemical properties of the trehalose-6-phosphate synthase and trehalose-6-phosphate phosphatase.* Extremophiles, 2005. **9**(1): p. 29-36.
163. da Costa, M.S., H. Santos, and E.A. Galinski, *An overview of the role and diversity of compatible solutes in Bacteria and Archaea*, in *Biotechnology of Extremophiles*, G. Antranikian, Editor. 1998, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 117-153.
164. Konig, H., et al., *Glycogen in thermoacidophilic archaebacteria of the genera Sulfolobus, Thermoproteus, Desulfurococcus and Thermococcus.* Arch Microbiol, 1982. **132**(4): p. 297-303.
165. Paiva, C.L. and A.D. Panek, *Biotechnological applications of the disaccharide trehalose.* Biotechnol Annu Rev, 1996. **2**: p. 293-314.
166. Higashiyama, T., *Higashiyama, T. Novel functions and applications of trehalose. Pure Appl. Chem. 74, 1263-1269*. Vol. 74. 2002.
167. Colaco, C., et al., *Extraordinary Stability of Enzymes Dried in Trehalose: Simplified Molecular Biology*. Vol. 10. 1992. 1007-11.
168. Roser, B., *Trehalose, a new approach to premium dried foods.* 1991. **v. 2**.
169. Jain, N.K. and I. Roy, *Effect of trehalose on protein structure.* Protein Science : A Publication of the Protein Society, 2009. **18**(1): p. 24-36.
170. Jain, N.K. and I. Roy, *Trehalose and Protein Stability*, in *Current Protocols in Protein Science*. 2010, John Wiley & Sons, Inc.

171. He, Q., et al., *Treatment with Trehalose Prevents Behavioral and Neurochemical Deficits Produced in an AAV alpha-Synuclein Rat Model of Parkinson's Disease.* Mol Neurobiol, 2016. **53**(4): p. 2258-68.
172. Fernandez-Estevez, M.A., et al., *Trehalose reverses cell malfunction in fibroblasts from normal and Huntington's disease patients caused by proteosome inhibition.* PLoS One, 2014. **9**(2): p. e90202.
173. Tien, N.T., et al., *Trehalose Alters Subcellular Trafficking and the Metabolism of the Alzheimer-associated Amyloid Precursor Protein.* J Biol Chem, 2016. **291**(20): p. 10528-40.
174. Eleutherio, E., *Trehalose: As Sweet as Effective in Biomedical Research and Biotechnology.* Advances in Biotechnology & Microbiology, 2018. **8**(4).
175. O'Neill, M.K., et al., *Tailoring Trehalose for Biomedical and Biotechnological Applications.* Pure Appl Chem, 2017. **89**(9): p. 1223-1249.
176. Urbanek, B.L., et al., *Chemoenzymatic synthesis of trehalose analogues: rapid access to chemical probes for investigating mycobacteria.* Chembiochem, 2014. **15**(14): p. 2066-70.
177. Pena-Zalbidea, S., et al., *Chemoenzymatic radiosynthesis of 2-deoxy-2-[(18)F]fluoro-d-trehalose ([(18)F]-2-FDTre): A PET radioprobe for in vivo tracing of trehalose metabolism.* Carbohydr Res, 2019. **472**: p. 16-22.
178. Kalscheuer, R. and H. Koliwer-Brandl, *Genetics of Mycobacterial Trehalose Metabolism.* Microbiology Spectrum, 2014. **2**.
179. Lippert, K. and E.A. Galinski, *Enzyme stabilization be ectoine-type compatible solutes: protection against heating, freezing and drying.* Applied Microbiology and Biotechnology, 1992. **37**(1): p. 61-65.
180. Goller, K., A. Ofer, and E.A. Galinski, *Construction and characterization of an NaCl-sensitive mutant of Halomonas elongata impaired in ectoine biosynthesis.* FEMS Microbiology Letters, 1998. **161**(2): p. 293–300.
181. Knapp, S., R. Ladenstein, and E.A. Galinski, *Extrinsic protein stabilization by the naturally occurring osmolytes beta-hydroxyectoine and betaine.* Extremophiles : life under extreme conditions, 1999. **3**(3): p. 191–198.
182. Lentzen, G. and T. Schwarz, *Extremolytes: natural compounds from extremophiles for versatile applications.* Appl Microbiol Biotechnol, 2006. **72**(4): p. 623-634.
183. Harishchandra, R.K., et al., *The effect of compatible solute ectoines on the structural organization of lipid monolayer and bilayer membranes.* Biophys Chem, 2010. **150**(1-3): p. 37-46.
184. Lentzen, G. and T. Schwarz, *Kompatible Solute: Mikrobielle Hersetellung und Anwendung*, in *Angewandte Mikrobiologie*, G. Antranikian, Editor. 2006, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 355-371.
185. Sauer, T. and E.A. Galinski, *Bacterial milking: A novel bioprocess for production of compatible solutes.* Biotechnol Bioeng, 1998. **57**(3): p. 306-13.
186. Kunte, H., G. Lentzen, and E. Galinski, *Industrial Production of the Cell Protectant Ectoine: Protection Mechanisms, Processes, and Products*. Vol. 3. 2014.
187. Klähn, S., et al., *Glucosylglycerate: a secondary compatible solute common to marine cyanobacteria from nitrogen-poor environments.* 2010. **12**(1): p. 83-94.
188. Costa, J., et al., *Characterization of the biosynthetic pathway of glucosylglycerate in the archaeon Methanococcoides burtonii.* Journal of bacteriology, 2006. **188**(3): p. 1022-1030.
189. Esteves, A.M., et al., *Mannosylglycerate and Di-myo-Inositol Phosphate Have Interchangeable Roles during Adaptation of Pyrococcus furiosus to Heat Stress.* Appl Environ Microbiol, 2014. **80**(14): p. 4226-33.
190. Borges, N., et al., *Mannosylglycerate: structural analysis of biosynthesis and evolutionary history.* Extremophiles, 2014. **18**(5): p. 835-852.
191. Kumar, R., et al., *Extremophiles: Sustainable Resource of Natural Compounds-Extremolytes*. 2009. p. 279-294.
192. Scholz, S., et al., *Di-myo-inositol-1, 1'-phosphate: A new inositol phosphate isolated from Pyrococcus woesei.* FEBS Letters, 1992. **306**(2-3): p. 239–242.

193. Martins, L.O. and H. Santos, *Accumulation of Mannosylglycerate and Di-myo-Inositol-Phosphate by Pyrococcus furiosus in Response to Salinity and Temperature.* Applied and Environmental Microbiology, 1995. **61**(9): p. 3299–3303.
194. Borges, N., et al., *Biosynthetic pathways of inositol and glycerol phosphodiesters used by the hyperthermophile Archaeoglobus fulgidus in stress adaptation.* J Bacteriol, 2006. **188**(23): p. 8128-35.
195. Seely, R.J. and D.E. Fahrney, *A novel diphospho-P,P'-diester from Methanobacterium thermoautotrophicum.* J Biol Chem, 1983. **258**(18): p. 10835-8.
196. Hensel, R. and H. König, *Thermoadaptation of methanogenic bacteria by intracellular ion concentration.* FEMS Microbiology Letters, 1988. **49**(1): p. 75-79.
197. Lehmacher, A., A.B. Vogt, and R. Hensel, *Biosynthesis of cyclic 2,3-diphosphoglycerate. Isolation and characterization of 2-phosphoglycerate kinase and cyclic 2,3-diphosphoglycerate synthetase from Methanothermus fervidus.* FEBS Lett, 1990. **272**(1-2): p. 94-8.
198. Seely, R.J. and D.E. Fahrney, *Levels of cyclic-2,3-diphosphoglycerate in Methanobacterium thermoautotrophicum during phosphate limitation.* J Bacteriol, 1984. **160**(1): p. 50-4.
199. Becker, J. and C. Wittmann, *Microbial production of extremolytes — high-value active ingredients for nutrition, health care, and well-being.* Current Opinion in Biotechnology, 2020. **65**: p. 118-128.
200. Ciulla, R., et al., *Halotolerance of Methanobacterium thermoautotrophicum delta H and Marburg.* J Bacteriol, 1994. **176**(11): p. 3177-87.
201. Shima, S., et al., *Activation and thermostabilization effects of cyclic 2, 3-diphosphoglycerate on enzymes from the hyperthermophilic Methanopyrus kandleri.* Arch Microbiol, 1998. **170**(6): p. 469-72.
202. Matussek, K., et al., *Cloning, sequencing, and expression of the gene encoding cyclic 2, 3-diphosphoglycerate synthetase, the key enzyme of cyclic 2, 3-diphosphoglycerate metabolism in Methanothermus fervidus.* J Bacteriol, 1998. **180**(22): p. 5997-6004.
203. Hensel, R. and I. Jakob, *Stability of Glyceraldehyde-3-Phosphate Dehydrogenases from Hyperthermophilic Archaea at High Temperature.* Systematic and Applied Microbiology, 1993. **16**(4): p. 742-745.
204. Borges, N., et al., *Comparative study of the thermostabilizing properties of mannosylglycerate and other compatible solutes on model enzymes.* Extremophiles : life under extreme conditions, 2002. **6**(3): p. 209–216.
205. Lehmacher, A. and R. Hensel, *Cloning, sequencing and expression of the gene encoding 2-phosphoglycerate kinase from Methanothermus fervidus.* Mol Gen Genet, 1994. **242**(2): p. 163-8.
206. Fessner, W.-D., *Systems Biocatalysis: Development and engineering of cell-free "artificial metabolisms" for preparative multi-enzymatic synthesis.* New Biotechnology, 2015. **32**(6): p. 658-664.
207. Streit, W.R. and R.A. Schmitz, *Metagenomics--the key to the uncultured microbes.* Curr Opin Microbiol, 2004. **7**(5): p. 492-8.
208. Carrigg, C., et al., *DNA extraction method affects microbial community profiles from soils and sediment.* Appl Microbiol Biotechnol, 2007. **77**(4): p. 955-64.
209. Ferrer, M., et al., *Estimating the success of enzyme bioprospecting through metagenomics: current status and future trends.* Microbial biotechnology, 2016. **9**(1): p. 22-34.
210. Park, S.J., J.C. Chae, and S.K. Rhee, *Application of DNA microarray for screening metagenome library clones.* Methods Mol Biol, 2010. **668**: p. 313-24.
211. Varaljay, V.A., et al., *Deep sequencing of a dimethylsulfoniopropionate-degrading gene (dmdA) by using PCR primer pairs designed on the basis of marine metagenomic data.* Appl Environ Microbiol, 2010. **76**(2): p. 609-17.
212. Daniel, R., *The metagenomics of soil.* Nature Reviews Microbiology, 2005. **3**(6): p. 470-478.
213. Simon, C. and R. Daniel, *Metagenomic analyses: past and future trends.* Appl Environ Microbiol, 2011. **77**(4): p. 1153-61.

214. Uchiyama, T. and K. Miyazaki, *Product-induced gene expression, a product-responsive reporter assay used to screen metagenomic libraries for enzyme-encoding genes.* Appl Environ Microbiol, 2010. **76**(21): p. 7029-35.
215. Uchiyama, T., et al., *Substrate-induced gene-expression screening of environmental metagenome libraries for isolation of catabolic genes.* Nat Biotechnol, 2005. **23**(1): p. 88-93.
216. Leis, B., A. Angelov, and W. Liebl, *Screening and expression of genes from metagenomes.* Adv Appl Microbiol, 2013. **83**: p. 1-68.
217. Williamson, L.L., et al., *Intracellular screen to identify metagenomic clones that induce or inhibit a quorum-sensing biosensor.* Appl Environ Microbiol, 2005. **71**(10): p. 6335-44.
218. Guan, C., et al., *Signal mimics derived from a metagenomic analysis of the gypsy moth gut microbiota.* Appl Environ Microbiol, 2007. **73**(11): p. 3669-76.
219. Yun, J., et al., *Characterization of a novel amylolytic enzyme encoded by a gene from a soil-derived metagenomic library.* Appl Environ Microbiol, 2004. **70**(12): p. 7229-35.
220. Elend, C., et al., *Isolation and biochemical characterization of two novel metagenome-derived esterases.* Appl Environ Microbiol, 2006. **72**(5): p. 3637-45.
221. Voget, S., H.L. Steele, and W.R. Streit, *Characterization of a metagenome-derived halotolerant cellulase.* J Biotechnol, 2006. **126**(1): p. 26-36.
222. Hu, Y., et al., *Cloning and enzymatic characterization of a xylanase gene from a soil-derived metagenomic library with an efficient approach.* Appl Microbiol Biotechnol, 2008. **80**(5): p. 823-30.
223. Craig, J.W., et al., *Expanding small-molecule functional metagenomics through parallel screening of broad-host-range cosmid environmental DNA libraries in diverse proteobacteria.* Appl Environ Microbiol, 2010. **76**(5): p. 1633-41.
224. Warren, R.L., et al., *Transcription of foreign DNA in Escherichia coli.* Genome Res, 2008. **18**(11): p. 1798-805.
225. Sorensen, H.P. and K.K. Mortensen, *Advanced genetic strategies for recombinant protein expression in Escherichia coli.* J Biotechnol, 2005. **115**(2): p. 113-28.
226. Egorova, K. and G. Antranikian, *Industrial relevance of thermophilic Archaea.* Current Opinion in Microbiology, 2005. **8**(6): p. 649-655.
227. Unsworth, L.D., J. van der Oost, and S. Koutsopoulos, *Hyperthermophilic enzymes − stability, activity and implementation strategies for high temperature applications.* The FEBS Journal, 2007. **274**(16): p. 4044-4056.
228. Amend, J.P. and E.L. Shock, *Energetics of overall metabolic reactions of thermophilic and hyperthermophilic Archaea and Bacteria.* FEMS Microbiology Reviews, 2001. **25**(2): p. 175-243.
229. Perevalova, A.A., et al., *Desulfurococcus fermentans sp. nov., a novel hyperthermophilic archaeon from a Kamchatka hot spring, and emended description of the genus Desulfurococcus.* 2005. **55**(3): p. 995-999.
230. Mardanov, A.V., et al., *Metabolic versatility and indigenous origin of the archaeon Thermococcus sibiricus, isolated from a siberian oil reservoir, as revealed by genome analysis.* Appl Environ Microbiol, 2009. **75**(13): p. 4580-8.
231. Graham, J.E., et al., *Identification and characterization of a multidomain hyperthermophilic cellulase from an archaeal enrichment.* Nature Communications, 2011. **2**(1): p. 375.
232. Huber, R., et al., *Sulfur-inhibited Thermosphaera aggregans sp. nov., a new genus of hyperthermophilic archaea isolated after its prediction from environmentally derived 16S rRNA sequences.* 1998. **48**(1): p. 31-38.
233. Cannio, R., et al., *A xylan-degrading strain of Sulfolobus solfataricus: isolation and characterization of the xylanase activity.* 2003. **8**: p. 117-124.
234. Prokofeva, M.I., et al., *Isolation of the anaerobic thermoacidophilic crenarchaeote Acidilobus saccharovorans sp. nov. and proposal of Acidilobales ord. nov., including Acidilobaceae fam. nov. and Caldisphaeraceae fam. nov.* 2009. **59**(12): p. 3116-3122.
235. Ando, S., et al., *Hyperthermostable endoglucanase from Pyrococcus horikoshii.* Appl Environ Microbiol, 2002. **68**(1): p. 430-3.

236. Maurelli, L., et al., *Evidence that the xylanase activity from Sulfolobus solfataricus Oalpha is encoded by the endoglucanase precursor gene (sso1354) and characterization of the associated cellulase activity.* Extremophiles : life under extreme conditions, 2008. **12**(5): p. 689-700.
237. Kublanov, I.V., et al., *Desulfurococcus kamchatkensis sp. nov., a novel hyperthermophilic protein-degrading archaeon isolated from a Kamchatka hot spring.* 2009. **59**(7): p. 1743-1747.
238. Gavrilov, S.N., et al., *[Characterization of membrane-bound Fe(III)-EDTA reductase activities of the thermophilic gram-positive dissimilatory iron-reducing bacterium Thermoterrabacterium ferrireducens].* Mikrobiologiia, 2007. **76**(2): p. 164-71.
239. Lombard, V., et al., *The carbohydrate-active enzymes database (CAZy) in 2013.* Nucleic acids research, 2014. **42**(Database issue): p. D490-D495.
240. Irwin, D., et al., *Roles of the catalytic domain and two cellulose binding domains of Thermomonospora fusca E4 in cellulose hydrolysis.* Journal of bacteriology, 1998. **180**(7): p. 1709-1714.
241. Bauer, M.W., et al., *An endoglucanase, EglA, from the hyperthermophilic archaeon Pyrococcus furiosus hydrolyzes beta-1,4 bonds in mixed-linkage (1-->3),(1-->4)-beta-D-glucans and cellulose.* J Bacteriol, 1999. **181**(1): p. 284-90.
242. Talamantes, D., et al., *Natural diversity of cellulases, xylanases, and chitinases in bacteria.* Biotechnology for Biofuels, 2016. **9**(1): p. 133.
243. Chu, Y., et al., *The GH10 and GH48 dual-functional catalytic domains from a multimodular glycoside hydrolase synergize in hydrolyzing both cellulose and xylan.* Biotechnology for Biofuels, 2019. **12**(1): p. 279.
244. Jia, X. and Y. Han, *The extracellular endo-β-1,4-xylanase with multidomain from the extreme thermophile Caldicellulosiruptor lactoaceticus is specific for insoluble xylan degradation.* Biotechnology for Biofuels, 2019. **12**(1): p. 143.
245. Sunna, A., M. Gibbs, and P. Bergquist, *A novel thermostable multidomain 1,4-b-xylanase from 'Caldibacillus cellulovorans' and effect of its xylan-binding domain on enzyme activity.* Microbiology (Reading, England), 2000. **146 (Pt 11)**: p. 2947-55.
246. Kurokawa, J., et al., *Clostridium thermocellum cellulase CelT, a family 9 endoglucanase without an Ig-like domain or family 3c carbohydrate-binding module.* Applied Microbiology and Biotechnology, 2002. **59**(4): p. 455-461.
247. Blanco, A., et al., *A multidomain xylanase from a Bacillus sp. with a region homologous to thermostabilizing domains of thermophilic enzymes.* Microbiology (Reading, England), 1999. **145 (Pt 8)**: p. 2163-70.
248. Zverlov, V., et al., *The multidomain xylanase A of the hyperthermophilic bacterium Thermotoga neapolitana is extremely thermoresistant.* Applied Microbiology and Biotechnology, 1996. **45**(1): p. 245-247.
249. Winterhalter, C., et al., *Identification of a novel cellulose-binding domain within the multidomain 120 kDa xylanase XynA of the hyperthermophilic bacterium Thermotoga maritima.* Mol Microbiol, 1995. **15**(3): p. 431-44.
250. Araki, R., et al., *Effect of Family 22 Carbohydrate-Binding Module on the Thermostability of Xyn10B Catalytic Module from Clostridium stercorarium.* Bioscience, Biotechnology, and Biochemistry, 2006. **70**(12): p. 3039-3041.
251. Gilbert, H.J. and G.P. Hazlewood, *Bacterial cellulases and xylanases.* 1993. **139**(2): p. 187-194.
252. Notenboom, V., et al., *Crystal Structures of the Family 9 Carbohydrate-Binding Module from Thermotoga maritima Xylanase 10A in Native and Ligand-Bound Forms.* Biochemistry, 2001. **40**(21): p. 6248-6256.
253. Göller, K. and E. A. Galinski, *Protection of a model enzyme (lactate dehydrogenase) against heat, urea and freeze-thaw treatment by compatible solute additives.* Journal of Molecular Catalysis B: Enzymatic, 1999. **7**(1-4): p. 37–45.
254. Knapp, S., R. Ladenstein, and E.A. Galinski, *Extrinsic protein stabilization by the naturally occurring osmolytes beta-hydroxyectoine and betaine.* Extremophiles, 1999. **3**(3): p. 191-8.

255. Huber, R., et al., *A novel group of abyssal methanogenic archaebacteria (Methanopyrus) growing at 110 °C.* Nature, 1989. **342**(6251): p. 833-834.
256. Faria, C., et al., *Inhibition of formation of alpha-synuclein inclusions by mannosylglycerate in a yeast model of Parkinson's disease.* Biochim Biophys Acta, 2013. **1830**(8): p. 4065-72.
257. Autengruber, A., et al., *Signalling-dependent adverse health effects of carbon nanoparticles are prevented by the compatible solute mannosylglycerate (firoin) in vitro and in vivo.* PLoS One, 2014. **9**(11): p. e111485.
258. Lange, M. and B.K. Ahring, *A comprehensive study into the molecular methodology and molecular biology of methanogenic Archaea.* FEMS Microbiology Reviews, 2001. **25**(5): p. 553-571.
259. Reeve, J.N., *Structure and Organization of Genes*, in *Methanogenesis: Ecology, Physiology, Biochemistry & Genetics*, J.G. Ferry, Editor. 1993, Springer US: Boston, MA. p. 493-527.
260. Rosano, G. and E. Ceccarelli, *Rare codon content affects the solubility of recombinant proteins in a codon bias-adjusted Escherichia coli strain.* Microbial cell factories, 2009. **8**: p. 41.
261. Kim, R., et al., *Overexpression of archaeal proteins in Escherichia coli.* Biotechnology Letters, 1998. **20**(3): p. 207-210.
262. Vera, A., et al., *The conformational quality of insoluble recombinant proteins is enhanced at low growth temperatures.* Biotechnology and Bioengineering, 2007. **96**(6): p. 1101-1106.
263. van Alebeek, G.-J.W.M., et al., *Cyclic 2,3-diphosphoglycerate metabolism in Methanobacterium thermoautotrophicum (strain ΔH): characterization of the synthetase reaction.* Archives of Microbiology, 1994. **162**(3): p. 193-198.
264. Breitung, J., et al., *Salt dependence, kinetic properties and catalytic mechanism of N-formylmethanofuran:tetrahydromethanopterin formyltransferase from the extreme thermophile Methanopyrus kandleri.* European Journal of Biochemistry, 1992. **210**(3): p. 971-981.
265. Liu, Y.-F., et al., *Molecular Mechanism Underlying the Interaction of Typical Sac10b Family Proteins with DNA.* PLOS ONE, 2012. **7**(4): p. e34986.
266. Reed, C.J., et al., *Protein Adaptations in Archaeal Extremophiles.* Archaea, 2013. **2013**: p. 373275.
267. Freydank, A.-C., W. Brandt, and B. Dräger, *Protein structure modeling indicates hexahistidine-tag interference with enzyme activity.* Proteins: Structure, Function, and Bioinformatics, 2008. **72**(1): p. 173-183.
268. Young, C.L., Z.T. Britton, and A.S. Robinson, *Recombinant protein expression and purification: A comprehensive review of affinity tags and microbial applications.* Biotechnology Journal, 2012. **7**(5): p. 620-634.
269. Booth, W.T., et al., *Impact of an N-terminal Polyhistidine Tag on Protein Thermal Stability.* ACS omega, 2018. **3**(1): p. 760-768.
270. Flamholz, A., et al., *eQuilibrator--the biochemical thermodynamics calculator.* Nucleic acids research, 2012. **40**(Database issue): p. D770-D775.

I DANKSAGUNG

Besonderer Dank gilt meiner Doktormutter Frau Prof. Dr. Bettina Siebers, die die Bearbeitung dieser spannenden Forschungsthematiken in ihrem Labor ermöglicht hat. Im Zuge dessen bedanke ich mich für die konstruktiven Anregungen und nützlichen Ratschläge zu experimentellen Abläufen, sowie für die Einräumung des wissenschaftlichen Freiraums zur Weiterentwicklung meiner Fähigkeiten als Wissenschaftlerin. Gleichermaßen bedanke ich mich für ihr entgegen gebrachtes Vertrauen und ihr Verständnis während des gesamten Bearbeitungszeitraums.

Herrn Dr. Christopher Bräsen möchte ich für seine Unterstützung und wissenschaftlichen Ratschläge danken. Vor allem bedanke ich mich für das entgegen gebrachte Vertrauen, die Geduld und Erläuterungen zu wissenschaftlichen Problematiken, sowie das Korrekturlesen aller Manuskripte, sowohl als auch dieser Arbeit.

Herrn Prof. Dr. Peter Bayer möchte ich für die Übernahme des Amtes als Zweitgutachter danken.

Der Evonik Industries AG, besonders Herrn Dr. Felix Müller danke ich für die Vergabe des Doktorandenstipendiums, welches die Bearbeitung eines großen Teils dieser Forschungsthematik finanziell ermöglicht hat.

Ich bedanke mich ausnahmslos bei allen Partnern des HotSolute Projekts für die gute Zusammenarbeit und die inspirierenden und spannenden Projekttreffen.

Frau Dr. Verena Kallnik und Frau Dr. Britta Tjaden danke ich für die wissenschaftlichen Diskussionen und Anregungen bezüglich verschiedener experimenteller Methoden und die Zusammenarbeit in den Anfängen dieser Arbeit.

Ich danke Herrn Prof. Dr. Jochen Niemeyer und Herrn Prof. Dr. Oliver Schmitz für die interne Kooperation zur Nutzung weiterer relevanter analytischer Methoden. An dieser Stelle gilt mein Dank ebenfalls Frau Dr. Martha Kohlhaas für ihre Erläuterungen bezüglich der Datenauswertung und der Erweiterung meiner analytischen Kenntnisse.

Für die technische und administrative Unterstützung, welche ich in allen Abschnitten dieser Promotion erfahren durfte, bedanke ich mich herzlichst bei Thomas Knura, Sabine Dietl und Agathe Materla.

In diesem Zuge möchte ich mich auch bei meinem ehemaligen Bachelor und Master Studenten: Katharina Dietzel, Marten Dahlhaus, Andreas Schielke und Yassin Kaspareit für ihr hervorragendes Engagement und ihre Arbeit im Labor bedanken.

Ein weiteres großes Dankeschön möchte ich den ehemaligen, sowie den aktuellen Mitgliedern der Arbeitsgruppe Siebers: Dr. Verena Kallnik, Dr. Britta Tjaden, Dr. Christopher Bräsen, Thomas Knura, Sabine Dietl, Agathe Materla, Dr. Lu Shen, Xiao Xiao Zhou, Dr. Jens Benninghoff, Katharina Fafenroth, Marcel Blum, Claus-Rüdiger Wallis, Dr. Frank Schult, Larissa Schocke, Astrid Neu, Sabine Krevet, Dr. Benjamin Meyer, Laura Kurschmierz, Alexander Wagner, Christian Schmerling, Thomas Klaus und Svenja Höfmann aussprechen. Die angenehme Arbeitsatmosphäre im Labor ließ so manchen „Forschungsfrust“ schnell verfliegen und schaffte einen freien Kopf für neue Ansätze. Bei Dr. Benjamin Meyer möchte ich mich an dieser Stelle nochmal für die Zusammenarbeit im letzten Drittel dieser Arbeit bedanken, für die Unterstützung und die Anregungen zu den verschiedenen Stressexperimenten.

Sicherlich gilt der größte Dank meinen verstorbenen Eltern Manfred und Gerlinde, für ihre uneingeschränkte Unterstützung in allen Lebenslagen und auch dabei, den Weg in die Promotion überhaupt anzutreten. Leider musstet ihr beide in diesen für mich nicht immer einfachen letzten Jahren aus dem Leben scheiden. Zusätzlich zu all dem Stress der ganzen Trauer und auch manchmal unterschwelligen Wut, weiss ich tief in meinem Herzen, wie sehr ihr euch mit mir über diesen finalen Abschluss gefreut hättet und das lässt mich dieses Kapitel mit einem lachenden und einem weinenden Auge abschließen.

Für die Unterstützung, den Trost, das Verständnis, die offenen Ohren und den Beistand während den Trauerphasen möchte ich mich an dieser Stelle noch einmal gesondert bedanken, besonders bei meiner Oma Annelies, danke Oma, dass du immer zu mir gestanden hast. Auch den anderen Anverwandten und Freunden der Familie möchte ich danken, für das entgegen gebrachte Verständnis in dieser turbulenten Zeit.
Thomas Knura, Sabine Dietl und Agathe Materla danke ich an dieser Stelle ein weiteres Mal, da ohne euch so mancher Arbeitstag für mich in dieser Zeit gar nicht möglich gewesen wäre, danke für eure Freundschaft.

Last but not least, danke ich dir Chris, zwischen all den aufreibenden Ereignissen meines Lebens sind wir uns begegnet und es war dir nicht zu kompliziert an meiner Seite zu verweilen, mit mir zu trauern, mich zu stützen und zu unterstützen, mit mir ins Labor zu fahren, mich abzulenken, mich vergessen zu lassen und zum Lachen zu bringen. Danke, mein „Motivator“, ich liebe dich.

II VERFASSUNGSERKLÄRUNG

Hiermit erkläre ich, dass ich die vorliegende Arbeit:

Compatible Solutes and novel Biocatalysts in (hyper)thermophilic Archaea: From Identification and Physiological Significance to Application in Biotechnology

ohne fremde Hilfe und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus fremden Quellen (einschließlich elektronischer Quellen) direkt oder indirekt übernommenen Gedanken sind ausnahmslos als solche kenntlich gemacht. Ich bestätige, dass die eingereichten Arbeiten meine eigenen sind, es sei denn, die Arbeiten waren ein Teil von gemeinsam verfassten Veröffentlichungen.
Ich versichere außerdem, dass ich die vorliegende Dissertation nur in diesem und keinem anderen Promotionsverfahren eingereicht habe und, dass diesem Promotionsverfahren keine endgültig gescheiterten Promotionsverfahren vorausgegangen sind.

Essen den 06.11.2020 Christina Stracke

Ort, Datum (Christina Stracke)

III APPENDIX

Determination of trehalase activity in *S. acidocaldarius*

As described in section 1.3.6, recent studies revealed two trehalases (TreHs) in *S. acidocaldarius* (*Saci*_1250 (TreH1) & *Saci*_1816 (TreH2) [159, 160]. Both studies used different conditions for the determination of enzymatic activities, therefore, maybe one of the reported TreHs showed no clear activity in the report of Lee and co-workers [159, 160]. The analyzed trehalose concentrations and buffers in both studies differs significantly (90 mM trehalose versus 5 mM trehalose (1.5 µg) and 50 mM acetate buffer (pH 4.0), 50 mM Gly-HCl (pH 3.7) versus 20 mM sodium citrate buffer (pH 4.0)). Concerning the importance of the trehalose metabolism in *S. acidocaldarius* for this thesis, both enzymes were cloned in pET15b, expressed in *E. coli.* BL21 (DE3) Codon-plus, purified, and analyzed for their respective activities (Figure 10).

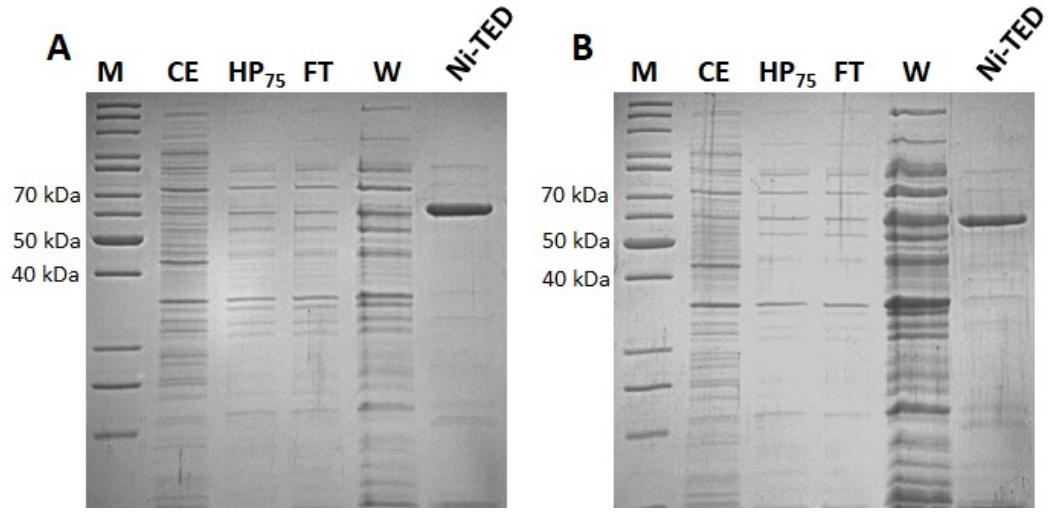


Figure 10: Purification of the recombinant trehalases TreH1 and TreH2 from *S. acidocaldarius*. The genes S*aci*_1250 (*treH1*) and S*aci*_1816 (*treH2*) were cloned into pET15b and expressed in *E. coli* BL21(DE3)-Codon-Plus with N-terminal His-tag. SDS-PAGE (12.5%, Coomassie staining) of protein fractions (2 -10 µg protein) of the respective purification steps are shown for TreH1 (A) or TreH2 (B). Abbreviations: CE, crude extract; HP_{75}, soluble fraction after heat precipitation at 75°C for 30 min; FT, flow through; W, wash fraction; Ni-TED, protein fraction after Ni-TED affinity chromatography; M, protein marker (unstained protein ladder, Thermo Fisher Scientific, Schwerte, Germany).

Since neutral or acidic types of trehalases have been described, the respective activities were briefly analyzed from pH 3.0 - 7.0 (Figure 11). TreH1 (*Saci*_1250) probably is more a neutral trehalase than TreH2 (*Saci*_1816). However, the highest activity was measured at pH 5.0 for both enzymes (TreH1 3.6 U/mg and TreH2 5.7 U/mg). In comparison, TreH2 showed a higher specific activity of 2.1 U/mg at pH 3.0, whereas TreH1 at pH 7.0 (1.7 U/mg). Therefore, in *S. acidocaldarius trehalose* degradation can occur via these two TreH enzymes.

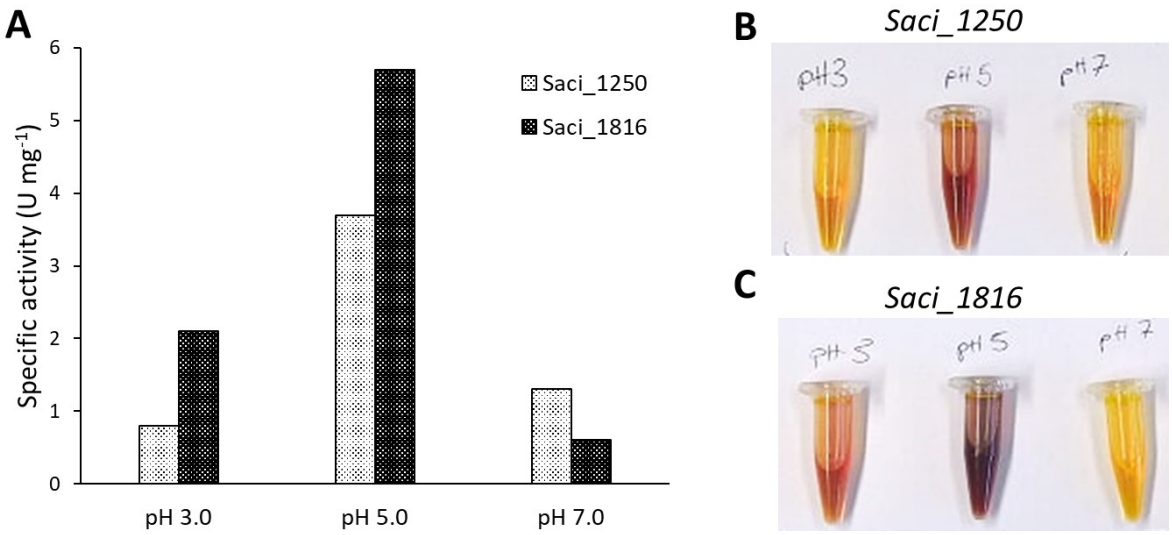


Figure 11: pH depended trehalase activity of recombinant TreH1 and TreH2 from *S. acidocaldarius*. The pH dependence (A) of TreH1 (S*aci*_1250) white bars, and (B) of TreH2 *Saci*_1816) black bars, were analyzed using 5 µg of the purified enzyme in 200 mM McIllvaine Buffer pH 3.0 – 7.0 at 70°C. The activity was determined using the DNSA assays for the detection of reducing ends, which were produced from 10 mM trehalose for TreH1 (B) and TreH2 (C). The method was performed as previously described in chapter 3.3. in the section for activity measurements of recombinant enzymes.

Expression and characterization of the 2-phosphoglycerate kinase (2PGK) from *Methanothermus fervidus*

As described in section 1.5.1 the enzymatic synthesis of cDPG was shown to proceed via two catalytic steps (Figure 6). In the first step, 2-phosphoglycerate kinase (2PGK) synthesizes the phosphate ester 2,3-diphosphoglycerate (2,3DPG) from 2-phosphoglycerate (2PG). In the second step, cyclic-2,3-diphosphoglycerate synthetase (cDPGS) catalyze the formation of an intramolecular phosphoanhydride bond resulting in cDPG. The key enzymatic step via the cDPGS enzyme was analyzed in detail and final results towards the establishment of an enzymatic *in vitro* approach to produce cDPG was presented in chapter 3.2. Concerning the importance of extremolyte biosynthesis, also the first reaction step via the 2-phosphoglycerate kinase (2PGK) was analyzed. Therefore, 2-phosphoglycerate kinase (*2pgk,* Mfer_0247, 915 bp) from *M. fervidus* was purchased as codon optimized gene (Eurofins, Ebersberg, Germany) for expression in *E. coli*. The gen was subcloned into pET15b (Amp^R) (Novagen, Merck, Darmstadt, Germany) and the plasmid was verified by sequencing (LGC genomics, Berlin, Germany). Competent cells of *E. coli* BL21(DE3)-CodonPlus-pRIL (Cam^R) (Merck, Darmstadt, Germany) were transformed with the constructed plasmid and grown in 200 mL – 400 mL lysogeny broth medium (LB broth) (Carl Roth GmbH + Co. KG, Karlsruhe, Germany). The expression was induced by the addition of 1 mM isopropyl-β-d-thiogalactopyranoside (IPTG) at an OD_{600} between 0.6 - 0.8. Cultures were incubated over night at 16°C. Cells were harvested (6.000 x g, 20 min, 4°C) and per 1 g of the wet cell mass 5 mL of buffer was used for resuspension ((KH_2PO_4/K_2HPO_4, 5mM DTT, 300 mM KCl or NaCl, pH 8 or 50 mM Tris 5 mM DTT, 300 mM KCl or NaCl, pH 8). For lysis, cells suspensions were passed three times through a French pressure cell at 20.000 psi. The cell lysate was centrifuged at 4°C, 25,000xg for 45 min to remove all cell debris. The supernatant containing the recombinant 2PGK was subjected to a heat precipitation step at 75°C for 20 min (Eppendorf Thermomixer at 700 rpm) followed by centrifugation at 4°C, 25,000 x g for 45 min to remove denatured mesophilic host proteins. Further purification was obtained via hydroxyapatite affinity chromatography (1 g in 200 mM K_2HPO_4, Tricorn 10/100 Column, protein binding capability ~ 12 mg, gradient 300 – 850 mM KCl) according to the described method by [197]. The expression, purification and subunit molecular mass of 2PGK was monitored by denaturing SDS-polyacrylamide gel electrophoresis (Figure 12). Protein concentrations were determined with a modified Bradford method using bovine serum albumin as a standard (Bio-Rad, Feldkirchen, Germany).

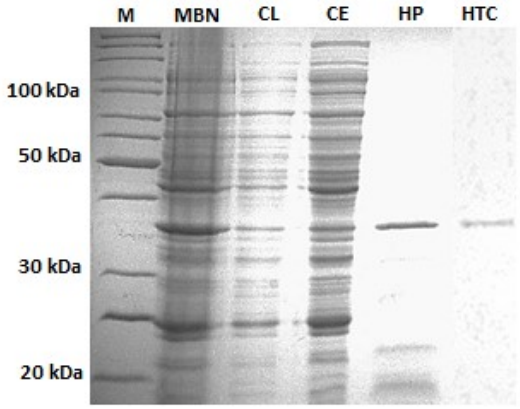


Figure 12: Purification of recombinant 2PGK from *M. fervidus* heterologously expressed in *E. coli*. The codon optimized *2pgk* was cloned in pET15b without affinity tag, the expression was optimized in *E. coli* BL21(DE3)-Codon-Plus, and the protein was purified. Protein fractions (5-10 µg) after successive purification steps were analyzed via SDS-PAGE (12.5 %) and stained with Coomassie Brilliant Blue. Mbn: Membrane fraction, CL: Crude lysate, CE: Crude extract, HP: Soluble fraction after heat precipitation (75°C, 30 min), HTC: Fraction after hydroxylapatite chromatography, M: Protein marker, unstained protein ladder (Thermo Fisher Scientific, Carlsbad, USA).

After purification, recombinant enzyme turned out to be inactive, therefore, the enzymatic properties were determined using partial purified 2PGK after heat treatment. The activity of 2PGK (32 µg) was measured continuously in 50 mM MES/KOH pH 6.5 with 2 mM $MgCl_2$, 5 mM DTT and 300 – 500 mM KCl as 2PG (1 mM) and ATP (2.5 mM) at 55°C in a final volume of 0.5 mL.

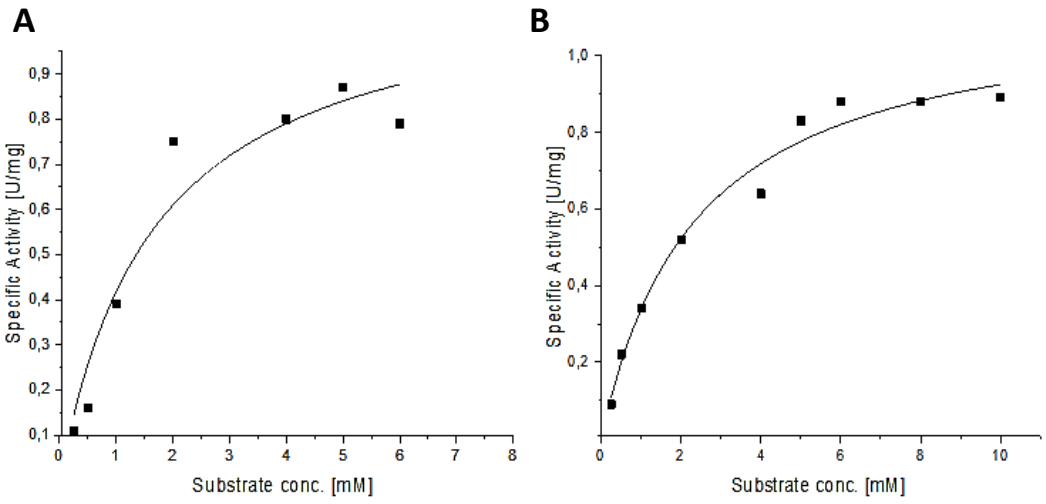


Figure 13: Kinetic properties and substrate specificity of the recombinant 2PGK from *M. fervidus*. The enzymatic activity of 2PGK (32 µg) was determined at 55°C (340 nm) by coupling the formation of ADP from ATP to the oxidation of NADH with pyruvate kinase (PK) and (lactic acid dehydrogenase (L-LDH) from rabbit muscle as auxiliary enzymes. The specific activity was determined with 2PG concentrations from 0 - 6 mM and 2.5 mM ATP (**A**) and ATP concentrations from 0 - 10 mM and 2 mM 2PG (**B**).

Reactions were started after 1 min of preincubation by the addition of the substrate (2PG). The formation of ADP was detected in presence of 2 mM phosphoenolpyruvate and 0.5 mM NADH via the pyruvate kinase (PK, rabbit muscle, 8 U) (Merck, Darmstadt, Germany) and L-lactate dehydrogenase (L-LDH, rabbit muscle, 4 U) (Merck, Darmstadt, Germany) at 340 nm using a Specord UV/VIS spectrometer (Analalytik Jena AG, Jena, Germany). The rate dependence of the 2PGK catalyzed reaction followed classical Michaelis–Menten kinetics with a Vmax of 1.12 U mg^{-1} and a k_m for 2PG of 1.7 mM and 2.3 mM for ATP (Figure 13). The enzymatic activity is 1.6-fold activated in presence of 200 mM KCl and the pH optimum is pH 6.5 (Figure 14).

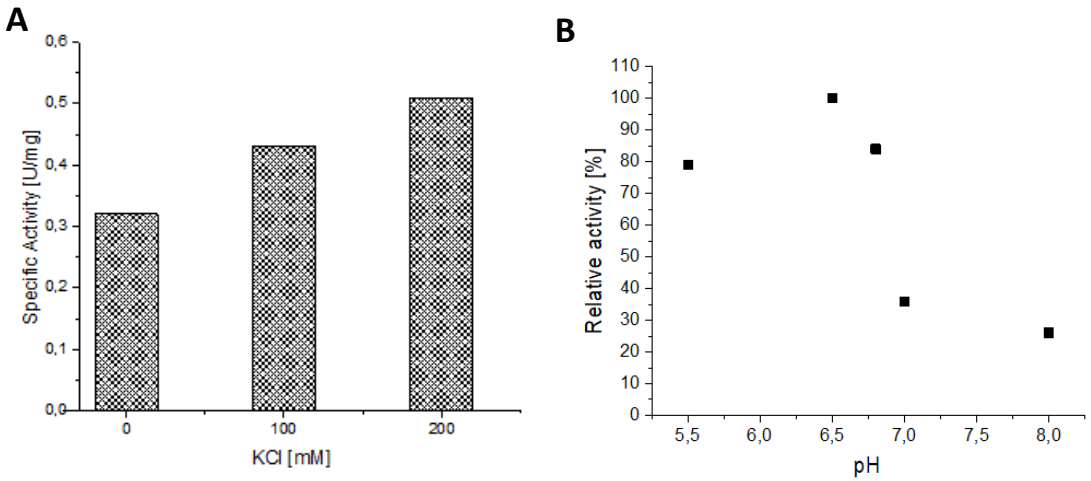


Figure 14: Effect of salt and pH on activity of recombinant 2PGK from *M. fervidus*. The enzymatic activity of 2PGK (32 µg) was determined at 55°C (340 nm) by coupling the formation of ADP from ATP to the oxidation of NADH with pyruvate kinase (PK) and (lactic acid dehydrogenase (L-LDH) from rabbit muscle as auxiliary enzymes. 100 % of relative activity corresponds to a specific activity of 0.4 U mg^{-1}. Measurements were performed in MES/KOH pH 5.5 - 6.5, TES/KOH pH 6.8, HEPES/KOH pH 7.0 - 8.0, with 2 mM 2 PG, 2.5 mM ATP, 3 mM $MgCl_2$ and 0 - 200 mM KCl.

IV LIST OF FIGURES

V LIST OF ABBREVIATIONS AND ACRONYMS

aa	Amino acid
ADP	Adenosine-diphosphate
ADPG	Adenosine-diphosphate-glucose
Amp	Ampicillin
ATP	Adenosine-triphosphate
BLAST	Basic Local Alignment Search Tool
bp	Base pairs
BSA	bovine serum albumin
CE	crude extract, cell free extract
CDP-Star	Disodium 2-chloro-5-(4-methoxyspiro[1,2-dioxetane-3,2′- 5 chlorotricyclo [3.3.1.13.7] decan])-4-yl]-1-phenyl phosphate
cDPG	cyclic-2,3-diphosphoglycerate
Cntrl.	Control
DIP	Di-*myo*-Inositolphosphate
DMSO	Dimethylsulfoxide
DNA	Desoxyribonucleic acid
dNTP	Deoxynucleotide triphosphates
DTT	Dithiothreitol
ED	Entner- Doudoroff- pathway
EDTA	Ethylenediaminetetraacetic acid
e.g.	for example
EMP	Embden-Meyerhof-Parnas-Pathway
g	Gramms
G6P	Glucose-6-phosphate
G6PDH	Glukose 6-phosphat dehydrogenase
GDH	Glukose dehydrogenase
GDP	Guanosine-diphosphate
GDPG	Guanosine-diphosphate-glocose
GER	Germany
GF	Gelfiltration
GT	Glycosyltransferase
GOI	Gen of interest
HEPES	2-(4-(2-Hydroxyethyl)-1-piperazinyl)-Ethansulfonsäure
HK	Hexokinase
HP	Heat precipitation
IEX	Ion exchange chromatography
IPTG	Isopropyl-ß-D-thiogalactopyranoside
Kan	Kanamycin
kDa	Kilodalton
L	Liter
LB	Luria-Bertani
LDH	Lactat dehydrogenase
M	Molarity
Mbp	Megabase pairs
MSC	mechanosensitive channell
min	Minute
MOPS	3-(N-Morpholino)-Propansulfonsäure
MW	molecular weight
NAD^+	Nicotinamide-adenine-dinucleotide (oxidized)
NADH	Nicotinamide-adenine-dinucleotide (reduced)
$NADP^+$	Nicotinamide-adenine-dinucleotide-phosphate (oxidized)
NADPH	Nicotinamide-adenine-dinucleotide-phosphate (reduced)
NCBI	National Center for Biotechnology Information

NDP	Nucleoside-diphosphate
NDPG	Nucleoside-diphosphate-glucose
nt	Nucleotide
OD	Optical density
ORF	open reading frame
OrfY	open reading frame Y
OtsA/OtsB	osmoregulatory trehalose synthese A/B
PAGE	Polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PEP	Phosphoenolpyruvate
pH	negative logarithm of the hydrogen ion (H+) conc.
P_i	inorganic phosphate
PP_i	Pyrophosphate
PK	Pyruvatkinase
Psi	pounds per square inch; 1 psi ~ 69 mbar
PVDF	Polyvinylidenfluoride
RNA	ribonucleic acid
rRNA	ribosomal ribonucleic acid
RT	Room temperature
rpm	revolution per minute
Saci	*Sulfolobus acidocaldarius*
SDS	Sodiumdodecylsulfate
sec	Seconds
Sso	*Sulfolobus solfataricus (Saccharolobus solfataricus)*
T6P	Trehalose-6-phosphate
TLC	Thin layer chromatography
TPP	Trehalose-6-phosphat phosphatase
TPS	Trehalose-6-phosphat synthase
TPSP	Trehalose-6-phosphat phosphatase/synthase
Tre	Trehalose
TreH	Trehalase
TreP	Trehalose phosphorylase
TreS	Trehalose synthase
TreT	Glycosyl transferring synthase
TreX	Glycogen debranching enzyme
TreY	Maltooligosyltrehalose synthase
TreZ	Malto-oligosyltrehalose Trehalohydrolase
Tris	Tris-(hydroxymethyl)-Aminomethane
TTX	*Thermoproteus tenax*
U	unit = Enzymaktivität
UDP	Uridine 5'-triphosphate
UDPG	Uridine 5'-triphosphate-glucose
UV	ultraviolet
V	Volt
Vol	Volume
X-Gal	5-Bromo-4-Chloro-3-Indolyl alpha-D-Glucopyranoside
x g	Earth acceleration / gravity
°C	Degrees Celsius
E	Extinction Coefficient
µ	micro (10^{-6})
λ	Lambda, wavelength
w/v	weight per volume; Gew.-%
v/v	volume per volume
2PG	2-phosphoglycerate
3PG	3-phosphoglycerate