Sulfolobus acidocaldarius & *Sulfolobus solfataricus*: Exploitation of thermoacidophilic Archaea for biotechnological applications

Dissertation
zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften
- Dr. rer. nat. -

vorgelegt von

JULIA CHRISTIN VERHEYEN

geboren am 09.08.1982 in Dresden

Biofilm Centre, Molekulare Enzymtechnologie und Biochemie
Fachbereich Chemie
der
Universität Duisburg-Essen

2015

Die vorliegende Arbeit wurde im Zeitraum von Oktober 2009 bis April 2015 im Arbeitskreis von Prof. Dr. Bettina Siebers in der Fakultät Chemie, Biofilm Centre, Molekulare Enzymtechnologie und Biochemie der Universität Duisburg-Essen durchgeführt.

Tag der Disputation: 24.06.2015

Gutachter:	Prof. Dr. Bettina Siebers
	Prof. Dr. Peter Bayer
Vorsitzender:	Prof. Dr. Thomas Schrader

Meinen Kindern –
die schönste Bereicherung meines Lebens

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

INTRODUCTION

CHAPTER 1 INTRODUCTION

1.1 (HYPER)THERMOPHILES, THEIR ENZYMES AND THE BIOTECHNOLOGICAL RELEVANCE

(Hyper)thermophiles are prokaryotic organisms with optimal growth at high temperatures thriving in habitats like e.g. terrestrial solfataric fields, hot sediments as well as abyssal hydrothermal vents (black smokers) and are defined as organisms with optimal growth above 80°C (Stetter, 1999). Comparison of 16S and 18S ribosomal RNA (rRNA) sequences revealed that hyperthermophiles, found in the prokaryotic domains Bacteria and Archaea, represent the deepest branches close to the root of the universal tree of life, indicating that these organisms may be the most ancient life-forms evolved on earth (Woese *et al.*, 1990; Stetter, 1996; Barns *et al.*, 1996). (Hyper)thermophiles are a subgroup of extremophiles, which in general are organisms thriving at low (down to -20°C) (Collins *et al.*, 2010) or high temperatures (up to 113°C) (Stetter, 1999), at extreme alkaline (pH 11) or acidic pH (pH 0) (Schleper *et al.*, 1995) or at high salt concentrations (close to saturation) (Fendrihan *et al.*, 2006) or combinations thereof. These organisms developed remarkable features to cope with these harsh environmental conditions. The enzymes of (hyper)thermophiles (thermozymes) for instance show high thermostability and optimal activity in the respective temperature range, some even at 110°C and more (Vieille *et al.*, 1996).

Especially at such high temperatures, but also at extreme pH, pressure or high salt levels, proteins can undergo some serious and irreversible damages, such as ß-elimination, deamidation, hydrolysis and oxidation, leading to a decrease or loss of enzyme function. Additionally, hydrophobic interactions are significantly weakened at temperatures above 110°C resulting in lower stability (Daniel *et al.*, 1996; Jaenicke, 1998). To overcome these problems enzymes and other cellular constituents of (hyper)thermophilic organisms have to be adapted in several ways to remain stable and functional under these harsh conditions. These adaptation strategies can include intrinsic (structural) or extrinsic factors. Generally, a combination of different strategies promotes the thermostability (Vieille *et al.*, 1996).

Intrinsic factors include increased hydrogen bonding as well as hydrophobic or ion-pair interactions, altered amino acid compositions, increased disulfide bridges, anchoring of loose ends, and oligomerization. Hydrophobic interactions are assumed to be a main force in protein stability by increasing the stability by average 1.3 kcal/mol for each additional methyl group being buried in the protein (Pace, 1992; Vieille and Zeikus, 2001). The comparison of amino acid composition of thermophilic versus mesophilic proteins revealed that the charged amino acids glutamate, lysine and arginine (Glu, Lys and Arg, respectively) are more abundant, suggesting the presence of more salt bridges and ionic interactions in thermophilic proteins, whereas the uncharged polar amino acids glutamine, asparagine, threonine, cysteine and serine (Gln, Asn, Thr, Cys and Ser, respectively) are decreased (Haney *et al.*, 1999; Cambillau and Claverie, 2000; Das and Gerstein, 2000). Furthermore, shortened loops, due to the extension or creation of secondary structures, or loop anchoring to the rest

of the protein through hydrogen bonds, ion pairs or hydrophobic interaction, enhances the compactness of a protein, hence, further increasing its thermostability (Vieille and Zeikus, 2001).

Apart from the intrinsic factors, several extrinsic factors, such as salts, coenzymes and substrates, were demonstrated to be involved in the stabilization of proteins (Vieille and Zeikus, 2001). So-called compatible solutes are probably the most prominent stabilizing factors. Compatible solutes are low molecular weight compounds synthesized either *de novo* or taken up by the cells (da Costa *et al.*, 1998) and are accumulated in high concentrations within the cytosol, but do not interfere with the metabolism (Empadinhas and da Costa, 2006). They optimize the physical characteristics of the cytosol, stabilize the conformation of biomolecules against stress factors and function as antagonists of osmotic pressure (Andersen *et al.*, 2011). The two different types of compatible solutes include amino acids and derivatives (proline, aspartate, glutamate, ectoine and glycine betaine) and sugars and polyols and their derivatives (e.g. di-*myo*inositol-phosphate (DIP), cyclic 2,3-diphosphoglycerate (cDPG) sucrose, trehalose, glycerol).

Not only the proteins of (hyper)thermophilic organisms exhibit several thermostabilizing modifications, also their deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) need to be protected at elevated temperatures. For potassium DIP and potassium cDPG and polycationic polyamines like spermidine and norspermidine e.g. a stabilizing function on the secondary structure of nucleic acids has been suggested (Daniel and Cowan, 2000). For the stabilization of structural RNA, particularly tRNA, branched polyamines are utilized, such as N^4-aminopropylspermine (Oshima *et al.*, 2011; Imanaka, 2011). Also, for rRNA a correlation between the higher guanine-cytosine (G-C) content and high growth temperatures was observed, especially for 16S rRNA (Wang *et al.*, 2006), but at the level of DNA this correlation does not seem to exist (Hickey and Singer, 2004). However, the presence of a very specific hyperthermophilic enzyme, the reverse gyrase (topoisomerase-like), gives strong evidence that supercoiling is an important factor for stabilization of DNA (Atomi *et al.*, 2004; Imanaka, 2011). Additionally, the unusual presence of histones in some (hyper)thermophilic prokaryotes compacting their DNA by bending and nucleosome formation, seems also to contribute to the stabilization of DNA (Daniel and Cowan, 2000).

The thermophilic adaptation of thermozymes is often accompanied by higher resistance towards harsh conditions like e.g. non-natural solvents, high temperature and pressure (Unsworth *et al.*, 2007). Such harsh conditions are frequently employed in industrial processes which makes thermozymes favorable for applications. Generally, in e.g. synthetic chemistry, biocatalytic enzyme mediated processes are often advantageous over chemical syntheses due to their high specificity and enantioselectivity and they often employ much less reaction steps. Therefore, many synthetic chemistry processes have already been replaced by more sustainable, environmentally friendly and eco-efficient (mesophilic) enzymatic methods. The use of (hyper)thermophilic enzymes in biotechnological applications instead of their mesophilic counterparts is often beneficial and hence of growing interest for different reasons: (i) They show higher resistance towards chemical denaturants due to their (thermo)stability (as mentioned above). (ii) Thermozymes can easily be purified via heat treatment after expression in mesophilic hosts. (iii) Under thermophilic conditions higher substrate concentrations are possible due to better solubility at higher temperatures. (iv) At high temperatures viscosity is usually lower and (v) the risk of microbial contamination is lower. (vi) Reaction rates are

often higher at higher temperatures and (vii) the thermostability offers the possibility to abolish cooling processes (Vieille and Zeikus, 2001). Furthermore, enzymes from (hyper)thermophiles represent a suitable and eco-efficient alternative since efforts of enhancing the (thermo)stability of mesophilic enzymes via rational design (like site-directed mutagenesis) or directed evolution approaches are time and cost intensive (Unsworth *et al.*, 2007).

The most famous example of application of a (hyper)thermophilic enzyme is the DNA polymerase of *Thermus aquaticus*, which has revolutionized the PCR technology (Saiki *et al.*, 1988). Since then many new (hyper)thermophilic DNA polymerases have been characterized and are now commercially available. There are many other industrial applications for the use of (hyper)thermophilic enzymes. Proteolytic enzymes are utilized in the detergent industry for laundering, in the leather industry for soaking, and for peptide synthesis (Antranikian *et al.*, 2005). Alcohol dehydrogenases are important for the synthesis of optically active alcohols in chemical industry and of co-factors, such as NAD^+ and $NADP^+$ (Radianingtyas and Wright, 2003). Cellulolytic enzymes find use in alcohol production, improvement of juice yields, color extraction from juices, in detergent production for color brightening, softening and soil removal (Allen, 1976), as well as in pretreatments of cellulosic biomass, saccharification of agricultural and industrial wastes and production of fine chemicals (Schülein, 2000). Xylanases are used as additives in food and feed industry, and in pulp and paper bleaching processes for lignin removal (Oksanen *et al.*, 2000). Starch-processing enzymes are employed in production processes for sugar syrup and other valuable products, such as starch-based substances with gelatin-like characteristics and defined linear dextrins. These products find use as fat substitutes, texturizers, aroma stabilizers and prebiotics (Gupta *et al.*, 2003). Also in trehalose production, which stabilizes enzymes, antibodies, vaccines and hormones, starch-processing enzymes are utilized (Schiraldi *et al.*, 2002). Cyclodextrin glycosyltransferases (CGTase) are utilized in cyclodextrin production from starch for gelling, thickening or stabilizing in jelly desserts, dressing, confectionery, dairy and meat products and for improving solubility of hydrophobic compounds important in pharmaceutical and cosmetic industries (Biwer *et al.*, 2002). Several novel (hyper)thermophilic starch and (ligno)cellulose degrading enzymes are known, which bear potential for further optimizing industrial applications (e.g. Bertoldo and Antranikian, 2002; Bhalla *et al.*, 2013; Blumer-Schuette *et al.*, 2014).

The most frequently utilized enzymes in biotechnological applications are, however, lipolytic enzymes, due to their high versatility. Many lipolytic enzymes are active in organic solvents, do not need any co-factors and often show extraordinary stereo-, chemo- and regioselectivities (Jäger and Reetz, 1998). Lipases and esterases are applied in several production processes e.g. of pharmaceuticals and detergents. Moreover, the food industry bears a wide range of application possibilities, for example in flavor development and improvement of cheese. Biodiesel production and synthesis of polymers and fine chemicals are further biotechnological processes employing lipolytic enzymes (Hasan *et al.*, 2006; Salameh and Wiegel, 2007; Levisson *et al.*, 2009).

Lipolytic enzymes are ubiquitous and can be found in all three domains of life. In the presence of water lipolytic enzymes catalyze the reversible hydrolysis of ester bonds into fatty acids and alcohols, whereas in organic solvents a transesterification reaction, where an alkoxy group of an ester is exchanged with another alkoxy group from an alcohol, is catalyzed. In the enzyme commission system this enzyme class of hydrolases acting on ester bonds (EC 3.1) is subdivided amongst others

into the group of carboxylic-ester hydrolases (EC 3.1.1), which in turn harbors 97 sub-classes, including the enzyme classes carboxylesterases (EC 3.1.1.1) and triacylglycerol lipases (EC 3.1.1.3). The two groups differ in their substrate preference. While esterases cleave preferably short-chain acylesters (≤ 10 C-atoms), "true" lipases act preferably on long-chain acylesters (>10 C-atoms) (Levisson *et al.*, 2009). Further, the members of these two enzyme classes (EC 3.1.1.1 and EC 3.1.1.3) were initially categorized into eight families based on amino acid sequence homologies and physiological functions (Arpigny and Jäger, 1999; Hausmann and Jäger, 2010). While the Family I comprises all true lipases within 6 subfamilies, the esterases were categorized into Family II-VIII. The increasing number of new carboxylesterases and lipases, which could not be assigned to any existing family, led to the proposal of Family IX-XIV (Rao *et al.*, 2013). However, due to the high structural diversity among lipolytic enzymes a categorization is often difficult (Chow *et al.*, 2012). Most lipolytic enzymes studied so far were from bacterial sources, including the (hyper)thermophilic esterases from *Thermotoga maritima* and *Thermus thermophilus*. Only few have been characterized from Archaea, amongst them the (hyper)thermophilic carboxylic-ester hydrolases from *Aeropyrum penix*, *Picrophilus torridus*, *Sulfolobus spp.*, *Pyrococcus spp.* and others. These (hyper)thermophilic carboxylic-ester hydrolases mainly belong to the class of esterases (Atomi and Imanaka, 2004; Levisson *et al.*, 2009).

However, in industrial processes most lipolytic enzymes applied so far originate from mesophilic organisms. Nevertheless, (hyper)thermophilic candidates are much more favorable, since most biotechnological applications are performed under or include harsh conditions requiring high enzyme stability at high temperature and high organic solvent concentrations.

1.2 MINING FOR NOVEL ENZYMES - METAGENOMICS

As outlined above there is a high and still growing demand in novel biocatalysts and, therefore, efficient mining strategies for new and unique enzymes are strongly required. In the pre(meta)genomics era, the identification of new enzymes and their encoding sequences relied solely on the culturability of the respective organisms. However, it is assumed, that about 99% of all microbes are not culturable and, thus, the tremendous diversity of living organisms and their biocatalysts remained undetected (Amann *et al.*, 1995). The metagenomics approach, in contrast, circumvents this dependence on culturability by directly investigating the genomic DNA of environmental samples via PCR, metagenomic DNA libraries and more recently also via next generation sequencing techniques.

The term metagenomics was first introduced in 1998 as a culture-independent way for analyzing microbial niches (Handelsman *et al.*, 1998). Since then great progress has been made in this field. With this method not only the phylogenetic biodiversity of a habitat can be analyzed, also functional information is gained. Metagenome analysis is initiated by the isolation of total DNA from an environmental sample, which has been optimized for the different sample types over the past years and even commercial kits are available. The quality of the preparation of the environmental DNA is a prerequisite for successful further analyses, which can either be performed using sequence based or function based approaches. Both strategies have their drawbacks but also strong advantages, implicating a combination of both will bear the most beneficial outcome (Leis *et al.*, 2013).

Sequence based metagenomic analyses include PCR-dependent techniques for e.g. phylogenetic marker genes and hybridization techniques (e.g. FISH), which both rely on at least partially known sequences (like conserved sequence stretches) of the respective molecules of interest. In contrast, the other approaches, like next-generation sequencing and the construction of metagenomic libraries, employed in sequence-based metagenome analyses provide sequencing data without prior information required. Library construction – depending on the fragment size obtained after DNA isolation and fragmentation - either involve the cloning of small-size fragments (< 15 kb) in standard sequencing vectors or of large inserts (up to 40 kb) into fosmids/cosmids and up to 200 kb in bacterial artificial chromosomes (BACs) (Rondon *et al.*, 2000; Entcheva *et al.*, 2001). While small-insert libraries are well suited for the functional identification of single genes (see below), large-insert libraries are utilized for the detection of large gene clusters or operons, thus, retrieving more information per clone. After cloning of metagenomic fragments or PCR products, respectively, subsequent transformation, sub-culturing and re-isolation, the respective fragments are sequenced. In next-generation sequencing approaches the metagenomic DNA fragments are directly sequenced.

All these sequence-based metagenomic methods finally provide sequencing data, which are then further processed by bioinformatic procedures yielding extensive information about e.g. the taxonomic/phylogenetic diversity and the gene content of a microbial habitat as well as the metabolic potential of a given microbial community, although low-abundant species might not be detected. The bioinformatic analyses of the sequencing data, however, are based on comparison with known sequences from databases, thus, resulting mainly in the discovery of new variants of already-known genes or of genes of unknown function when no related sequences are found in databases (e.g. Park *et al.*, 2010; Varaljay *et al.*, 2010).

This drawback of the sequence-based metagenomic analyses is circumvented by the functional screening approach, which in contrast allows for the identification of entirely novel genes and enzymes, since it is based on the activity/function conferred by the cloned metagenomic fragment rather than on sequence comparison. The functional screening approach relies on the construction of metagenomic libraries, which can either be screened by (i) classical activity assays, (ii) heterologous complementation of host strains/mutants, or (iii) induced gene expression (Simon and Daniel, 2011). Unique enzymes with potentially remarkable characteristics with no sequence similarity to known enzymes may be found with this function-based screening procedure (Streit and Schmitz, 2004). A simple and fast phenotypical characterization of clones can be performed by classical plate-screening methods based on e.g. activity via substrate(-analog) supplementation to the growth medium. This can involve chemical dyes, chromophore-containing substrate derivatives or simple halo formation on indicator plates. In contrast, *in vitro* activity assays represent a more time-consuming but also a more sensitive approach, for which a preceeding preparation of crude extracts is required (Simon and Daniel, 2011; Leis *et al.*, 2013). In heterologous complementation studies the utilized host lacks a (metabolic) gene important for growth or survival. An activity derived by a metagenomic fragment, which compensates this deficiency, results in growth of the host in selective medium. This approach represents a fast and selective screening technique with almost no false positive results (Simon and Daniel, 2011). Different variants of the induced gene expression strategy have recently been introduced: SIGEX (substrate-induced gene expression), METREX (metabolite-regulated expression) and PIGEX (product-induced gene expression), are all based on the expression of a reporter-gene,

mostly *gfp* (encoding for the green fluorescent protein). Positive clones are identified via fluorescence and can be separated via FACS analysis (fluorescence-activated cell sorting) (Leis *et al.*, 2013).

The most frequently used host in functional screenings - although few other expression hosts are known in metagenomics - is the gram-negative bacterium *Escherichia coli*. Nevertheless, the applicability as host for functional expression appears to be limited due to various reasons: the transcriptional apparatus of *E. coli* might not detect heterologous transcription signals, since functional expression mainly relies on native promoter structures (Warren *et al.*, 2008). Also, translation might be weak due to missing adequate ribosomal binding sites, little mRNA stability and different codon usage (Sørensen and Mortensen, 2005). Incorrect folding because of missing proper chaperons and protein modifying enzymes may lead to the formation of protein aggregates within the cytoplasm, so called inclusion bodies. In addition, low or no detection of foreign secretion signals by the host may result in incorrect or no secretion of proteins. Finally, missing co-factors may also reduce the detection of activity. All these factors are assumed to influence heterologous expression. Consequently, proteins of organisms with different transcriptional and (post)translational mechanisms or generally a different e.g. extremophilic life style, which requires specific adaptations as described above, are frequently missed in such metagenomic screenings.

Therefore, in addition to e.g. the improvement of expression within a given host or the development of more sensitive screening assays, the establishment of alternative hosts and the usage of several different hosts in parallel (Figure 1) is one of the most promising strategies to overcome the above mentioned drawbacks and improve the outcome of metagenomic analyses (Liebl *et al.*, 2014). This is particularly true for Archaea, which often inhabit extreme environments and harbor unique metabolic pathways and enzymes (see below), which are often not well expressed in the standard expression hosts, such as *E. coli*. The development of an archaeal expression host is, therefore, highly desirable to enable the identification of new extremozymes and the exploitation of their great potential for industrial applications, which would otherwise remain undetected. To date, no archaeal expression host is routinely applied in metagenomics, especially in function driven screenings of metagenomic libraries. The BMBF funded national project ExpresSys (GenoMik-Transfer) with seven academic and five industrial institutions aimed at the comparative screening of metagenomic libraries in phylogenetically diverse host organisms, i.e. the Bacteria *Escherichia coli*, *Pseudomonas putida*, *Rhodobacter capsulatus*, *Pseudomonas antarctica* and *Thermus thermophilus* and the Archaeon *S. acidocaldarius*. As part of this BMBF project the scope of this work was to establish the (hyper)thermoacidophilic Archaeon *S. acidocaldarius* as an alternative expression host for metagenomic analyses.

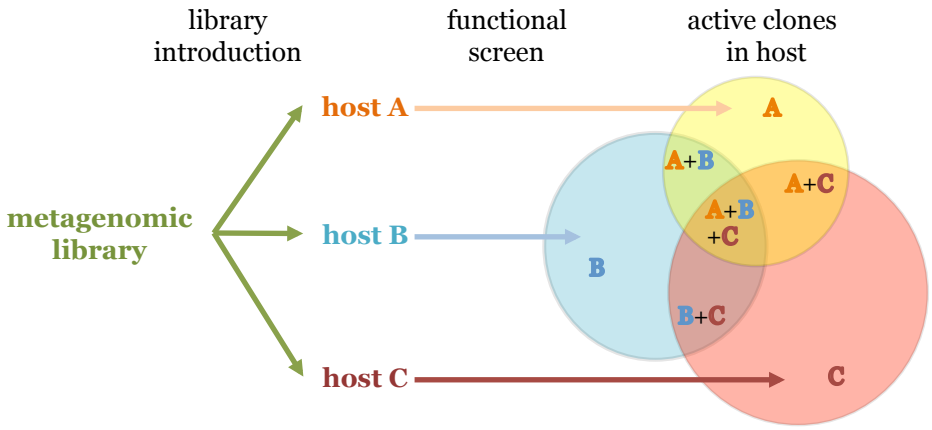


Figure 1. Advantageous use of different expression hosts in parallel for improving metagenomic analyses. Adopted from the BMBF project proposal ExpresSys (GenoMik-Transfer).

1.3 Archaea As Alternative Expression Hosts for Metagenomics

Archaea have been proposed as third domain of life – in addition to Eukarya and Bacteria - in 1977 based on 16S rRNA analysis (Woese and Fox, 1977). To date, this domain is classified into six phyla: the Euryarchaeota, represented by the methanogens, extreme halophiles and some thermoacidophiles and (hyper)thermophiles (Woese *et al.*, 1990); the Crenarchaeota, represented mainly by (hyper)thermophiles (Woese *et al.*, 1990); the Korarchaeota, so far only represented by non-cultivable, anaerobic Archaea and the first proposed member is "*Candidatus* Korarchaeum cryptofilum" (Auchtung *et al.*, 2006; Elkins *et al.*, 2008); the Nanoarchaeota, to date only represented by the two species *Nanoarchaeum equitans*, which can only grow in co-culture with *Ignicoccus hospitalis*, and Nst1, which presumably grows with a member of the *Sulfolobales* (Acd1) (Morris *et al.*, 2013; Podar *et al.*, 2013); the Thaumarchaeota, represented by thermophilic and mesophilic chemolithoautotrophic ammonia oxidizers (e.g. *Nitrososphaera gargensis*, *Nitrosopumilus maritimus*, and *Cenarchaeum symbiosum*) from quite distinct environmental habitats (e.g. freshwater, soil, ocean) (Brochier-Armanet *et al.*, 2008; Schleper and Nicol, 2010; Pester *et al.*, 2011); the Aigarchaeota, with *Caldiarchaeum subterraneum* as its sole representative reported so far, have recently been proposed as sixth archaeal phylum (Nunoura *et al.*, 2011). However, whether it indeed represents a new archaeal branch or if it is just a deep-branching thaumarchaeal lineage is still under debate (Brochier-Armanet *et al.*, 2011). Archaea were initially thought to populate exclusively extreme habitats. For this reason the term extremophiles was often applied for Archaea. Even though Archaea indeed populate a wide range of extreme environments, recent studies revealed that these organisms are ubiquitous also in mesophilic environments, like marine and freshwater sediments, terrestrial soil or even within the human mucosa (DeLong, 1998; Chaban *et al.*, 2006; Matarazzo *et al.*, 2012). Despite their wide distribution no archaeal pathogen has been identified so far.

Archaea combine bacterial, eukaryal as well as unique archaeal features in a mosaic character. Bacterial characteristics are a prokaryotic cell structure without a nucleus, a circular chromosome, additional plasmids and gene organization in operon structures. Mechanisms on the level of transcription and translation, however, are more eukaryal-like, although less complex (Jarrell *et al.*, 2011). Typical archaeal features are for example the archaeal cell membrane and cell wall structure. In

the cytoplasmic membrane polyisoprenyl groups are linked to polar head groups of a *sn*-glycerol-1-phosphate backbone via ether bonds instead of fatty acids linked to a *sn*-glycerol-3-phosphate backbone via ester-bonds as found in Bacteria and Eukarya (Albers and Meyer, 2011). Archaeal cell walls lack peptidoglucan commonly found in bacterial cell walls. Instead, some Archaea have pseudomurein (*Methanobacterium thermoautotrophicum* (König *et al.*, 1982, 1983)), but most Archaea exhibit S-layer proteins on the cell surface (e.g. *Thermoproteales*, *Sulfolobales*) (Albers and Meyer, 2011). Few Archaea do not possess a cell wall at all (*Ferroplasma acidophilum* (Golyshina and Timmis, 2005)). Motile Archaea can produce pili as well as so called archaella (archaeal flagella), which are distinct from bacterial type flagella (Jarrell, 2012; Lassak *et al.*, 2012). Beside these phenotypic characteristics, Archaea also harbor unique metabolic pathways and/or modified versions of classical pathways with unique enzymes (see below) (Siebers and Schönheit, 2005; van der Oost and Siebers, 2006; Zaparty *et al.*, 2008; Sato and Atomi, 2011; Zaparty and Siebers, 2011; Bräsen *et al.*, 2014).

This combination of bacterial, eukaryal and typical archaeal features may attribute to the above mentioned limitations in functional expression of archaeal (hyper)thermophilic proteins within the standard mesophilic hosts used in metagenomics, presumably due to different transcription signals and apparatus, the missing archaeal post-translational machinery, the misfolding of proteins at lower temperatures in mesophilic hosts, leading to reduced or even no enzyme activity, as well as a distinctive different codon usages (Eichler and Adams, 2005; Kim and Lee, 2006b). Therefore, the development of an archaeal, hyperthermophilic expression host for functional metagenomics is highly desirable, for which the availability of genetic tools is an important prerequisite. For several archaeal members, including the halophilic *Haloferax volcanii* and *Halobacterium salinarium*, the methanogenic *Methanococcus maripaludis* and *Methanosarcina spp.*, and the (hyper)thermophilic *Thermococcus kodakarensis*, *Pyrococcus furiosus* and *Sulfolobus spp.* (Leigh *et al.*, 2011; Atomi *et al.*, 2012), genetic toolboxes have been established. Particularly, *Sulfolobus spp.*, as thermoacidophilic organisms growing at temperatures around 80°C and a pH of 2-3, represent very promising expression host candidates in metagenomics, because they qualify not only for the identification of thermophilic enzymes, but also for enzymes exhibiting additionally acidophilic properties (= thermoacidophilic enzymes). Such enzymes are of great interest for many industrial applications.

The family *Sulfolobaceae* belongs to the phylum of Crenarchaeota and comprises six genera, one of them represented by the genus *Sulfolobus*. Genomes of several *Sulfolobus spp.* have been sequenced (Kawarabayasi *et al.*, 2001; She *et al.*, 2001; Chen *et al.*, 2005; Reno *et al.*, 2009; Guo *et al.*, 2011; Jaubert *et al.*, 2013) and for some species genetic tools have been established (e.g. *Sulfolobus solfataricus* P2, *Sulfolobus acidocaldarius* and *Sulfolobus islandicus*) (Jonuscheit *et al.*, 2003; Worthington *et al.*, 2003; Deng *et al.*, 2009; She *et al.*, 2009; Wagner *et al.*, 2009; Leigh *et al.*, 2011; Wagner *et al.*, 2012; Zhang and Whitaker, 2012; Wagner *et al.*, 2014). The membranes of *Sulfolobus spp.* are composed of tetraether lipids in a monolayer membrane, which are highly proton impermeable, allowing to keep the intracellular pH at 6.5 in an extreme acidic surrounding (van de Vossenberg *et al.*, 1995). The first isolated organism of the order of *Sulfolobales* was *S. acidocaldarius* from Locomotive Spring in Yellowstone National Park (Brock *et al.*, 1972). The organism grows heterotrophically at optimal temperatures of 78°C and at a pH of 2-3 under aerobic conditions and

utilizes monosaccharides (e.g. D-glucose, D-xylose), polysaccharides (e.g. dextrin, starch or maltotriose), yeast extract, tryptone and casaminoacids as energy and carbon sources (Grogan, 1989). The cells are irregularly lobed coccoid and possess Ups- and Aap-pili, archaella and threads (Figure 2) (Albers and Meyer, 2011). The genome comprises 2.2 Mbp, which is predicted to harbor about 2,329 putative ORFs (Chen *et al.*, 2005).

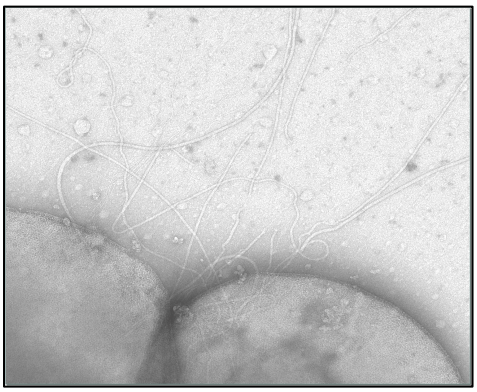


Figure 2. Electron micrograph of *S. acidocaldarius* DSM 639. Kindly provided by Sonja-Verena Albers, University of Freiburg.

A well-engineered genetic tool box is available for *S. acidocaldarius* and offers vector systems for the construction of gene deletion mutants and for heterologous gene expression (Berkner and Lipps, 2008; Berkner *et al.*, 2010; Wagner *et al.*, 2009, 2012, 2014).

The most widely applied selection strategy in *Sulfolobus* transformations so far is the use of auxotrophic mutants. The strain *S. acidocaldarius* MW001 contains a 322 bp deletion within the *pyrE* gene. The genes *pyrE* and *pyrF* encode for the orotate-phosphoribosyltransferase and the orotidin-5-mono-phosphate decarboxylase, respectively. PyrEF catalyzes the last two steps within the *de novo* synthesis of uridine-monophosphate and is, thus, essential for growth (Grogan and Gunsalus, 1993). Therefore, this deletion leads to an uracil auxotrophic strain, facilitating the use of *pyrEF* as a selectable marker (Wagner *et al.*, 2012). Transformation of *S. acidocaldarius* cells is achieved by electroporation.

Since *S. acidocaldarius* contains the restriction endonuclease *Sua*I, foreign DNA needs to be protected by specific N4-methylation of the inner cytosine residues of its GGCC recognition site via the methylase M.*Esa*BC4I (Grogan, 2003). The *E. coli* strain ER1821 contains an additional plasmid harboring this methylase. Thus, transformation of this strain with the desired plasmid and subsequent plasmid isolation yields correctly methylated DNA, which can be used for *S. acidocaldarius* transformation.

The shuttle vector pSVA1450 is used for heterologous expression in *S. acidocaldarius* (Figure 3). The vector is based on the cryptic plasmid pRN1, which is the best studied archaeal plasmid (e.g. Keeling *et al.*, 1996; Lipps, Stegert, *et al.*, 2001; Lipps, Ibanez, *et al.*, 2001; Berkner *et al.*, 2007). Next to pRN2 it occurs naturally in *S. islandicus* REN1H1. pSVA1450 is a derivative of pCmalLacS (Berkner *et al.*, 2010), originating from the shuttle vector pC (Berkner *et al.*, 2007) and utilizes the maltose-inducible promoter of MalE, a maltose-binding protein (Saci_1165) encoded in the maltose transport operon, for heterologous gene expression (Berkner *et al.*, 2010). In addition to the pRN1-based backbone, the vector contains also the *pyrEF* cassette for selection and *lacS* from *S. solfataricus* as a reporter gene for promoter studies in *S. acidocaldarius* MW001. For selection and replication in

E. coli the vector harbors the ampicillin resistance cassette (*bla*, ß-lactamase) and the origin of replication (*ori*, ColE1 replication origin), respectively. The reporter gene *lacS* can be replaced by any gene of interest. For C-terminal tagging (e.g. with a Strep-10xHis-tag) of the gene of interest in the expression vector pSVA1450 a pre-cloning step of the gene of interest via the vector pMZ1 is required. Compared to the shuttle vector pCmalLacS, pSVA1450 contains a copy of the regulatory protein of the maltose transport operon (*malR*, Saci_1161) and a 2 bp deletion within the *malE* promoter region (upstream of the *malE* transcription start site) improving the heterologous expression efficiency (Wagner *et al.*, 2014).

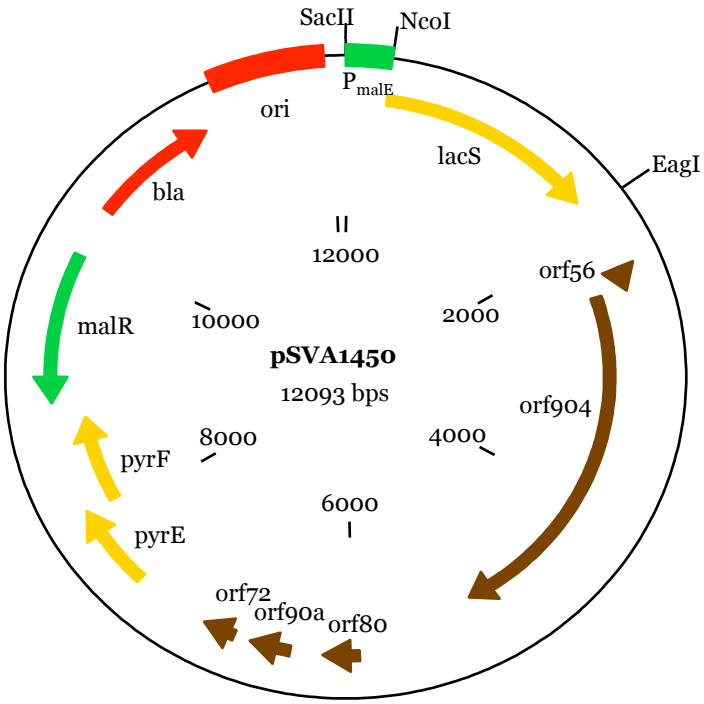


Figure 3. Graphic illustration of the expression vector pSVA1450. Backbone components based on pRN1 are shown in brown. The selection marker cassette *pyrEF* and the reporter gene *lacS* originating from *S. solfataricus* are depicted in yellow. The bacterial origin of replication *ori* and ampicillin resistance cassette *bla* originating from *E. coli* are shown in red. The promoter sequence of the maltose-binding protein MalE (P_{malE}) and the gene *malR* encoding the regulatory protein of the maltose transport operon originating from *S. acidocaldarius* are depicted in green.

Besides the possibility of heterologous expression in *S. acidocaldarius*, the genetic tool box also offers methods for the construction of markerless deletion strains via single- and double-crossover recombination events (Wagner *et al.*, 2009, 2012). As described above, in addition to many other benefits, this is a highly valuable tool for functional metagenomic screenings enabling heterologous complementation studies by constructing mutants depleted in a certain function, e.g. a specific enzyme activity. A metagenomic clone library is then screened for the ability to complement for this respective function. Furthermore, depending on the enzyme class of interest sometimes mutants depleted in the respective enzyme activity are required for successful screening results. Within the ExpreSys project the comparative screening in the different hosts focused on the expression of metagenome derived esterases, for which an esterase deficient *S. acidocaldarius* strain was required.

1.4 GLYCOLYSIS IN *SULFOLOBUS SOLFATARICUS* – THE ROLE OF GDH

The mosaic character of Archaea, outlined above, includes typical archaeal metabolic features. Although the archaeal metabolism is comparably complex like that in Bacteria and lower Eukaryotes, it harbors unique pathways and/or modified versions of classical pathways involving unique, unusual enzymes particularly in the central carbohydrate metabolism (Bräsen *et al.*, 2014). Many classical pathways present in Bacteria and Eukarya are absent in Archaea. Classical sugar degradation routes like Embden-Meyerhof-Parnas (EMP) and Entner-Doudoroff (ED) pathway only exist in modified versions, the oxidative pentose phosphate pathway is replaced either by an (incomplete) non-oxidative pentose phosphate pathway and/or most likely a reversed ribulose monophosphate pathway (RuMP) (van der Oost and Siebers, 2006; Bräsen *et al.*, 2014). Very recently, however, an oxidative pentose phosphate pathway has been described to be present in *Haloferax volcanii* (Pickl and Schönheit, 2015). Also, pentose degradation differs from the routes in bacterial model organisms. For sugar degradation anaerobic, hyperthermophilic Archaea utilize a modified version of the classical EMP pathway with unique enzymes, such as ADP- or PP_i-dependent kinases (Siebers *et al.*, 1998; Tuininga *et al.*, 1999; Verhees *et al.*, 2001), the non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase (GAPN (e.g. Lorentzen *et al.*, 2004; Ahmed *et al.*, 2005; Ettema *et al.*, 2008)) or the ferredoxin-dependent GAP-oxidoreductase (GAPOR (e.g. Reher *et al.*, 2007)) to form pyruvate from glucose. In contrast, modified versions of the ED pathway, i.e. a semi-phosphorylative and a non-phosphorylative version, have been described for most aerobic halophilic and thermoacidophilic Archaea, respectively. As indicated by more detailed studies on *S. solfataricus* and *T. tenax* and subsequent bioinformatic analyses, the most abundant ED variant in Archaea, comprises both, an sp-ED branch and np-ED branch, referred to as the branched ED-pathway (Ahmed *et al.*, 2005). Most of the analyses on the branched ED pathway were carried out with *S. solfataricus* P2, which was first isolated from a solfataric field (Pisciarelli) near Naples, Italy (Zillig *et al.*, 1980). It exhibits optimal growth at 80°C and a pH of 2-4 under aerobic conditions. In contrast to the close relative *S. acidocaldarius* (see above) it utilizes a wider range of carbon and energy sources: pentoses (like D-arabinose, L-arabinose, D-xylose), hexoses (like D-glucose, D-galactose, D-mannose), disaccharides (like maltose, sucrose), polysaccharides (like starch, dextrin), tryptone, peptides and amino acids (Grogan, 1989). The cells are also irregularly lobed coccoid and the organism possesses Ups-pili and archaella (Figure 4). The genome with 3 Mbp in size is the largest among the *Sulfolobus* species with approximately 3,033 predicted ORFs (She *et al.*, 2001).

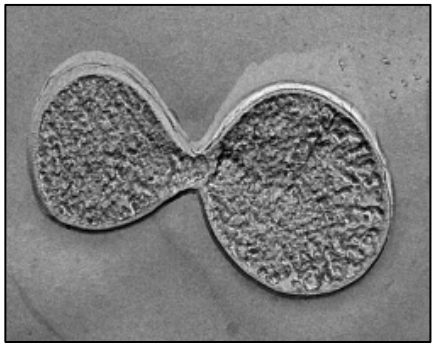


Figure 4. Electron micrograph of *S. solfataricus* P2. Kindly provided by Sonja-Verena Albers, University of Freiburg.

In the first two reactions in the branched ED pathway of *Sulfolobus spp.* common to both, the spED and npED branches, glucose is oxidized to gluconate by Glucose-1-dehydrogenase (GDH), which is then further dehydrated to the characteristic intermediate 2-keto-3-deoxy-gluconate (KDG) by gluconate dehydratase (GAD) (Lamble *et al.*, 2004; Kim and Lee, 2005; Ahmed *et al.*, 2005).

In the npED branch the KDG is first cleaved into glyceraldehyde (GA) and pyruvate by a bifunctional 2-keto-3-deoxy(-6-phospho)-gluconate aldolase (KD(P)GA, active on non-phosphorylated (KDG, GA) and phosphorylated (KDPG, GAP) substrates) (Lamble *et al.*, 2003; Ahmed *et al.*, 2005). A ferredoxin dependent GA oxidoreductase (GAOR) and subsequent glycerate kinase (GK) converts GA via glycerate to 2-phosphoglycerate (2-PG), respectively (Kardinahl *et al.*, 1999; Kouril, Wieloch, *et al.*, 2013). There is no indication of glyceraldehyde dehydrogenase (GADH) activity in *S. solfataricus* like it was shown for *Picrophilus torridus* and *Thermoplasma acidophilum* (Bräsen *et al.*, 2014). 2-PG is then directed into the lower shunt of the EMP pathway, where a second molecule of pyruvate is finally formed via phosphoenolpyruvate (PEP) by the enzymatic action of enolase (ENO) and pyruvate kinase (PK).

The KDG intermediate, however, can also be phosphorylated by KDG kinase (KDGK) (Lamble *et al.*, 2005; Kouril, Wieloch, *et al.*, 2013), thus, entering the spED branch. The resulting 2-keto-3-deoxy-6-phospho-gluconate (KDPG) is then cleaved by KD(P)GA into pyruvate and glyceraldehyde-3-phosphate (GAP). GAP is directly oxidized to 3-phosphoglycerate (3-PG) by an irreversible nonphosphorylating glyceraldehyde-3-phosphate dehydrogenase (GAPN) (Ettema *et al.*, 2008). This is a characteristic enzyme of (hyper)thermophilic Archaea. The usual enzyme couple for this conversion, glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and phosphoglycerate kinase (PGK), acts only gluconeogenically (Kouril, Esser, *et al.*, 2013). 3-PG is converted in the lower common shunt of the EMP pathway to 2-PG by phosphoglycerate mutase (PGAM), and 2-PG is then converted to PEP and finally to pyruvate via ENO and PK, respectively.

As shown particularly for *Sulfolobus spp.*, the branched ED pathway is promiscuous for the breakdown of glucose as well as galactose (Figure 5) (Ahmed *et al.*, 2005; Lamble *et al.*, 2005, 2003). For most of the enzymes detailed biochemical analyses are available. GDH, GAD, KD(P)GA and KDGK were all shown to utilize the respective glucose and galactose derived intermediates as substrates (Lamble *et al.*, 2003, 2004; Ahmed *et al.*, 2005; Lamble *et al.*, 2005; Kim and Lee, 2006a; Milburn *et al.*, 2006). At least some of the hexose degrading enzymes, i.e. GDH and KD(P)GA are also involved in the catabolism of pentoses like D-xylose and L-arabinose (Nunn *et al.*, 2010).

Although most of the enzymes required for a functional EMP pathway are present and employed in the lower shunt common to both EMP and ED pathway, for a functional catabolic EMP pathway a phosphofructokinase (PFK) is missing in *Sulfolobus spp.*. The key enzyme of gluconeogenesis in Archaea (and deeply branching Bacteria), the fructose-1,6-bisphosphate aldolase/phosphatase (FBPA/ase) (Say and Fuchs, 2010), however is present (Kouril, Esser, *et al.*, 2013), indicating that the EMP pathway including the classical GAPDH/PGK couple is exclusively utilized for gluconeogenic purposes (Zaparty and Siebers, 2011).

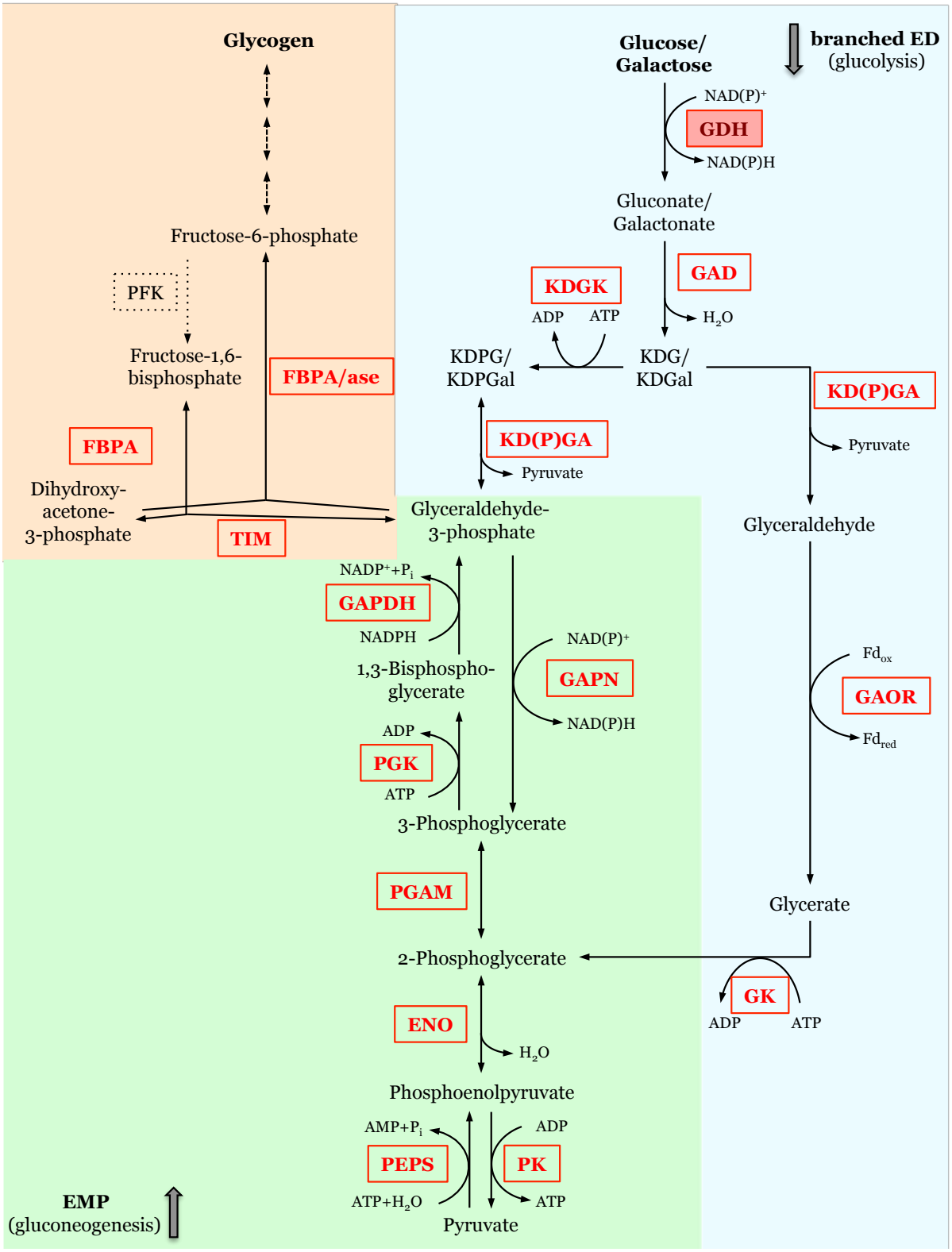


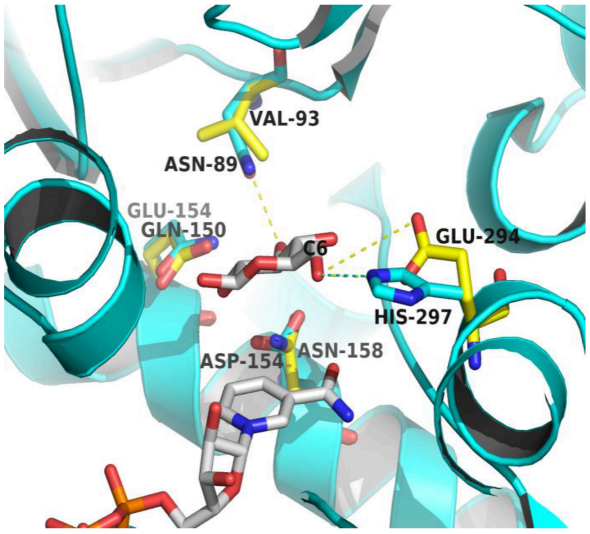
Figure 5. Glycolysis and gluconeogenesis in *S. solfataricus*. Branched ED pathway is boxed in blue, EMP pathway is boxed in orange. Shared reactions of ED and EMP pathway are boxed in green. Enzymes are boxed and given in red. GDH: glucose dehydrogenase, GAD: gluconate dehydratase, KD(P)GA: 2-keto-3-deoxy(-6-phospho)gluconate aldolase, KDGK: 2-keto-3-deoxy(-6-phospho)gluconate kinase, GAOR: glyceraldehyde oxidoreductase, GK: glycerate kinase, GAPN: non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase, GAPDH: glyceraldehyde-3-phosphate dehydrogenase, PGK: phosphoglycerate kinase, PGAM: phosphoglycerate mutase, ENO: enolase, PEPS: phosphoenolpyruvate synthetase, PK: pyruvate kinase, TIM: triosephosphate isomerase, FBPA: fructose-1,6-bisphosphate aldolase, FBPA/ase: fructose-1,6-aldolase/bisphosphatase, PFK: phosphofructokinase (Pathway map modified from (Zaparty and Siebers, 2011)).

As mentioned above, the first step of the modified branched ED pathway is the oxidation of glucose/galactose to gluconate/galactonate catalyzed by GDH-1 (SSO3003) in *S. solfataricus*. The enzyme was purified from glucose grown cells and characterized in detail. It has been shown that GDH-1 displays a broad substrate spectrum, oxidizing D-glucose and additionally also D-galactose, D-xylose, L-arabinose, D-fucose and 6-deoxy-D-glucose (Lamble *et al.*, 2003). GDH-1 is a member of the medium-chain alcohol/polyol dehydrogenase/reductase (MDR) branch of the pyridine nucleotide dependent alcohol/polyol/sugar dehydrogenase superfamily (EC 1.1.1.47) (Riveros-Rosas *et al.*, 2003). GDHs can be found in all three domains of life and several archaeal GDHs have been characterized as either homotetramers or homodimers composed of ~40 kDa subunits (Bhaumik and Sonawat, 1999; Lamble *et al.*, 2003; Ohshima *et al.*, 2003; Angelov *et al.*, 2005; Haferkamp *et al.*, 2011).

The GDH-1 enzyme can use both co-substrates NAD^+ and $NADP^+$ with approximately the same efficiency (Milburn *et al.*, 2006). Its crystal structure has been solved and shows the typical two-domain topology of the MDR superfamily members: the central nucleotide binding domain, displaying the characteristic nucleotide binding motif GXGXXG, is flanked by protein regions at the N and C termini, which together comprise the catalytic domain. The dual co-substrate specificity is attributed to the positively charged arginine residue at position 213, interacting with the phosphate group of $NADP^+$, which is absent in NAD^+ specific MDR dehydrogenases (Milburn *et al.*, 2006). The catalytic domain further contains the four conserved structural zinc-binding residues (Cys^{93}, Cys^{96}, Cys^{99} and Cys^{107}) as well as the catalytic zinc-binding residues (Cys^{39}, His^{66}, Glu^{67}) (Milburn *et al.*, 2006).

Several MDR paralogs have been identified in the genome of *S. solfataricus*, three of which have been characterized as alcohol dehydrogenases (SSO2536, SSO0764, SSO1646) (Chong *et al.*, 2007). More recently, one further MDR paralog with 41% sequence identity to the Sso-GDH-1 was recombinantly expressed, purified and characterized as glucose dehydrogenase, referred to as GDH-2 (SSO3204). In contrast to GDH-1, GDH-2 showed a clear preference for D-glucose as substrate and $NADP^+$ as co-substrate, indicating high substrate specificity (Haferkamp *et al.*, 2011). Based on these different substrate specificities it has been suggested that GDH-2 is the main enzyme for glucose degradation, whereas GDH-1 might function in different pentose and hexose degradation pathways and, therefore, serves as a standby enzyme for quick adaptation to different carbon sources (Haferkamp *et al.*, 2011). All residues shown to participate in nucleotide, co-substrate and zinc binding in GDH-1 are also highly conserved in GDH-2. However, the structural determinants responsible for the different substrate specificity between these very closely related enzymes are not known to date. Consistent with the high sequence identity structural modeling revealed a very similar overall structure and also a very similar active site of GDH-2 compared to GDH-1, however, some striking differences in the substrate binding residues have been observed (Figure 6). In GDH-1 the residues Asn^{89}, Glu^{114}, Gln^{150}, Asp^{154}, Asn^{307} and His^{297} are involved in sugar substrate binding (Milburn *et al.*, 2006). The main differences observed in GDH-2 are the substitutions of the polar Asn^{89} involved in complexing the C3-hydroxy group of the sugar in GDH-1, to the hydrophobic Val^{93} and of the partially positively charged His^{297}, interacting with the C6-hydroxyl of hexoses, to the negatively charged Glu^{294}. These substitutions have been proposed to be mainly responsible for the higher specificity of GDH-2 for glucose (Haferkamp *et al.*, 2011), whereas the other substitutions Asp^{154} and Gln^{150} in GDH-1 to Asn^{158} and Glu^{154} in GDH-2, respectively, are more similar with respect to length and charge and were

hence argued to be less relevant for substrate specificity. However, the significance of these active site differences for substrate specificity has not been experimentally confirmed so far.



GDH-1 SSO3003	GDH-2 SSO3204
C39	C40
A41	T42
N89*	**V93**
R90	R94
E114*	E118
I117	I121
Q150*	E154
D154*	N158
F279	-
V306	V303
N307*	N304
G308	A305
H297*	**E294**

Figure 6. Structural modeling of *S. solfataricus* GDH-2 on the basis of the solved crystal structure of *S. solfataricus* GDH-1 (Haferkamp *et al.*, 2011)**.** The glucose (centre, grey/red) binding pocket of GDH-1 is displayed in cyan, the according GDH-2 model is shown in yellow. The table lists the amino acids of GDH-1 within 0.5 nm of the glucose molecule (left column) and of GDH-2 at the corresponding positions (right column). Amino acids involved in glucose interaction according to Milburn *et al.*, 2006 are marked with an asterisk. The residues investigated in this thesis are marked in bold.

1.5 Aim Of This Work

Owing to their unique metabolism comprising many unusual pathways and enzymes and due to the (hyper)thermophilic life style of many representatives, Archaea constitute a highly valuable source for novel biocatalysts exhibiting high stability under harsh conditions and novel catalytic properties advantageous for many industrial applications. The identification of such enzymes by functional screenings of e.g. metagenomic libraries, however, is often hampered due to the lack of suitable screening hosts, which recognize the special features of archaeal, (hyper)thermophilic genes derived from environmental samples.

Thus, one major part of the present work aimed to establish the genetically tractable (hyper)thermoacidophilic Crenarchaeon *S. acidocaldarius* as an archaeal, hyperthermophilic expression host for functional screenings of metagenomic libraries. Therefore, (i) the promoter selectivity of *S. acidocaldarius* was examined using PFK promoters from different (hyper)thermophilic cren- and euryarchaeal as well as bacterial sources. The expression plasmids were constructed by cloning the respective promoters together with their downstream located natively controlled PFK encoding genes. For comparison, the PFK encoding genes were also cloned under the control of the inducible *Sulfolobus* specific, vector encoded *malE* promoter and expression was comparatively analyzed by PFK activity measurements; (ii) the expression capabilities of *S. acidocaldarius* were investigated using several esterase genes isolated from different metagenomes, which were cloned into the *S. acidocaldarius* expression vector pSVA2301 under control of the inducible *malE* promoter. For the expression of the genes an esterase deficient *S. acidocaldarius* host strain was constructed to

eliminate the host specific esterase background activity and expression efficiency was determined by esterase enzyme measurements.

The second major topic of the work was to elucidate the structural features determining the different substrate specificities of highly homologous sugar dehydrogenases, key enzymes of sugar degradation via the modified branched ED pathway in *S. solfataricus*. Mutant enzymes of the glucose-specific GDH-2 were constructed using site directed mutagenesis. As revealed by structure modeling, the amino acid residues found in corresponding positions to the substrate binding residues in GDH-1, exhibiting a broad substrate spectrum, were accordingly substituted in GDH-2 and the mutant enzymes were expressed, purified and characterized with respect to substrate specificity and kinetic properties in comparison to wild-type GDH-2 and to GDH-1.

Chapter 2

MATERIALS & METHODS

CHAPTER 2 MATERIALS & METHODS

2.1 CHEMICALS, VECTORS AND COMMERCIAL KITS

The following (co)substrates and auxiliary enzymes were purchased from Sigma-Aldrich: 2-deoxy-D-glucose, 6-deoxy-D-glucose, adenosine diphosphate (ADP, monopotassium salt), adenosine triphosphate (ATP, disodium salt), D-(-)-arabinose; dextrin, D-fructose 6-phosphate (F6P, disodium salt hydrate), D-fructose 1,6-bisphosphate (FBP, trisodium salt hydrate), D-(+)-fucose, D-(+)-galactose, D-(+)-glucosamine (hydrochloride), D-(+)-glucose, D-lactose (monohydrate), D-(-)-lyxose, D-(+)-maltose (monohydrate), D-(+)-mannose, D-(-)-ribose, D-(+)-xylose, fructose-1,6-bisphosphate aldolase (FBPA, rabbit muscle, EC 4.1.2.13), L-(+)-arabinose, L-(-)-xylose, NZ-amine, ρ-nitrophenyl butyrate (C4), ρ-nitrophenyl decanoate (C10), ρ-nitrophenyl dodecanoate (C12), ρ-nitrophenyl octanoate (C8), sodium pyrophosphate (PP_i, tetrabasic decahydrate), sucrose, triosephosphate isomerase (TIM, rabbit muscle, EC 5.3.1.1).

The following co-substrates were purchased from Gerbu Biotechnik GmbH: nicotinamide-adenine-dinucleotide (oxidized, NAD^+, hydrate), nicotinamide-adenine-dinucleotide-phosphate (oxidized, $NADP^+$, disodium salt).

All other chemicals and enzymes were purchased from Amersham Pharmacia Biotech Europe GmbH, Bio-Rad Laboratories GmbH, Difco Laboratories, Gerbu Biotechnik GmbH, Life Technologies, Merck Millipore, Qiagen, Roche Diagnostics GmbH, Roth GmbH, Serva Electrophoresis GmbH, Sigma-Aldrich, Thermo Scientific, and VWR International in analytical grade.

For heterologous expression in *Escherichia coli* the pET vector system (pET-11c (amp^r) and pET-24a (kan^r) (Merck Millipore)) was used. For heterologous expression in *Sulfolobus acidocaldarius* the vectors pMZ1 (amp^r, kindly provided by Dr. Melanie Zaparty), pSVA1450 (amp^r), pSVA2300 (amp^r) and pSVA2301 (amp^r) (kindly provided by Dr. Sonja-Verena Albers) were used.

The following commercial kits were used: DNeasy Blood and Tissue Kit (Qiagen), GeneJET Plasmid Miniprep Kit (Thermo Scientific), GoScript™ Reverse Transcription System (Promega), Protino® Ni-TED (Machery-Nagel), QIAfilter Plasmid Midi Kit (Qiagen), QIAquick Nucleotide Removal Kit (Qiagen), Wizard® SV Gel and PCR Clean-Up System (Promega).

2.2 INSTRUMENTS

Table 1. List of instruments

Instrument	Description	Manufacturer
Agarose gel electrophoresis system and power supply	B1A EasyCast™	Owl Separation Systems
	Consort E835	MS Laborgeräte
Analytical scale	EW 4200-2NM TE124S TE601	Kern & Sohn GmbH
Autoclaves	H+P Varioklav, 25T H+P Varioklav, 75S	Federgari Autoklav, Integra Bioscience
Centrifuges	Sorvall Centrifuge RC6 (rotors: Sorvall SLC-3000, SLA-1500 SuperLite)	Kendro Laboratory Products GmbH
	5415R, 5810R (rotors: F-34-6-38, F-45-30-11), miniSpin®Plus, Concentrator Plus	Eppendorf AG
Chromatography columns	HiLoad 26/60 Superdex 200 prep grade Resource Q HIC column (butyl sepharose media, self packed)	GE-Healthcare Life Sciences
Clean Bench	HERAsafe®KSP Class II Bio-safety Cabine	Kendro Laboratory Products GmbH
Fast Flow Liquid Chromatography (FPLC)	Äkta Purifier	GE-Healthcare Life Sciences
Electroporation device	Gene Pulser Xcell™ Microbial System	Bio-Rad Laboratories
Gel-documentation	GelDoc Gel Documentation System	Bio-Rad Laboratories GmbH
Heat Block	Thermoblock v4.6	Hardware&Service
	TSC ThermoShaker	Analytik Jena
Heater/Stirring device	Powertherm Variomag®	H+P Labortechnik AG
Incubators	HT Multitron (37°C) HT Thermotron (65-80°C)	Infors
	Heraus B6 (37°C) Heraus T20 (78°C)	Kendro Laboratory Products GmbH

Membrane-vacuum pump	Laborport Typ:N816.3KN.18	KNF Neuberger GmbH
Microwave	HF1612	Siemens
pH-Meter	WTW Series inoLab pH 720; pH-Electrode: SenTIX 81 pH0-14/0-100°C/3mol/KCl	WTW GmbH
Photometer	BioPhotometer Plus	Eppendorf AG
	Specord®200	Analytik Jena
	Specord®210 compact regulator: JUMO dTRON 308	
Plate Reader	Infinite Pro M200	Tecan
SDS PAGE system	Mini-Protean 3 System	Bio-Rad Laboratories GmbH
		MS Laborgeräte
	Power supply: Consort E835	
Sonicator	UP 200s	Hielscher Ultrasonics GmbH
Thermocycler	Mastercycler personal	Eppendorf AG
	Thermocycler C1000	Bio-Rad Laboratories GmbH
Vacuum centrifuge	RVC 2-25	Christ
VersaDoc imaging model	VersaDoc Model 4000 system	Bio-Rad Laboratories GmbH
Wet Tank blotting system	Mini Trans-Blot®	Bio-Rad Laboratories GmbH

2.3 MICROBIOLOGICAL METHODS

2.3.1 Strains of *Escherichia coli* & growth conditions

For cloning the *Escherichia coli* K-12 DH5α strain (DSMZ 6897) was used and for expression of recombinant proteins the *E. coli* Rosetta (DE3) strain (Agilent Technologies), harboring the plasmid pRARE (cam^r), was used.

Cultures of different *E. coli* strains were grown aerobically in 5 – 2000 mL Luria Bertani Broth (LB) medium in either reaction tubes or Erlenmeyer flasks, respectively, at 37°C with aeration by gyratory shaking (180 rpm). Growth was monitored photometrically at 600 nm. For colony growth solid medium plates were prepared. For this 1.5% (w/v) agar-agar was added to the medium. The appropriate antibiotics were added according to the encoded plasmid resistance at the following concentrations: ampicillin 100 mg/L (pET11c, pMZ1, pSVA1450, pSVA2301), kanamycin 50 mg/L (pET24a, pM.EsaBC4I), chloramphenicol 34 mg/L (pRARE).

2.3.2 Strains of *Sulfolobus acidocaldarius* & growth conditions

The strains *S. acidocaldarius* MW001, MW001Δ1105, MW001Δ1116 and MW001Δ1105Δ1116 were routinely cultivated at 78°C in minimal Brock medium (Table 2) (Brock *et al.*, 1972) containing 0.1% (w/v) NZ-amine, 0.2% (w/v) dextrin and 20 mg/L uracil, if not stated otherwise. If the strains harbored an expression plasmid carrying the *pyrEF* cassette, uracil was omitted. For induction of

expression 0.4% (w/v) maltose was added to the medium. Liquid cultivation was done aerobically by gyratory shaking at 180 rpm for at least 1-2 d. Growth was monitored photometrically at 600 nm. For colony growth 3 mM $CaCl_2$, 10 mM $MgCl_2$ and 6 g/L gellan gum (Gelzan™, Sigma-Aldrich) was added to the Brock medium. Plates were incubated at 78°C under steady conditions for at least 3-4 d.

Table 2. Composition of modified minimal Brock medium according to Brock *et al.* (1972).

	Component	Amount per L stock solution	Amount per L final Brock medium
Brock I			
	$CaCl_2$ x 2 H_2O	70 g	0.07 g
Brock II			
	$(NH_4)_2SO_4$	130 g	1.3 g
	$MgSO_4$ x 7 H_2O	25 g	0.25 g
	H_2SO_4 (50% (v/v))	1.5 mL	
Brock III			
	KH_2PO_4	56 g	0.028 g
	H_2SO_4 (50% (v/v))	1.5 mL	
	Trace elements		
	$MnCl_2$ x 4 H_2O	36 mg	1.8 mg
	Na_2MoO_4 x 2 H_2O	44 mg	0.22 mg
	$ZnSO_4$ x 7 H_2O	10 mg	0.05 mg
	$CuCl_2$ x 2 H_2O	6 mg	0.03 mg
	$NaMoO_4$ x 2 H_2O	2 mg	0.01 mg
	$VOSO_4$ x 2 H_2O	0.9 mg	4.5 mg
	$CoSO_4$ x 7 H_2O	6 mg	0.03 mg
1 L final Brock			
	Brock I		1 mL
	Brock II		10 mL
	Brock III		5 mL
	$FeCl_3$ x 6 H_2O (20% (w/v))		1 mL
	H_2SO_4 (50% (v/v))		Adjust to pH 3.5

Table 3. Composition of glycerol-Brock solution.

Component	Amount per 50 mL glycerol Brock solution
Brock I	50 µL
Brock II	500 µL
Brock III	250 µL
glycerol (50% (w/v))	50 mL

Brock I, II, and III (Table 2) were autoclaved and stored at room temperature (RT). Uracil (250 x stock (5 mg/mL)) was filter sterilized and stored at -20°C. $FeCl_3$ x 6 H_2O (20% (w/v)), NZ-amine (20% (w/v)), dextrin (20% (w/v)), sucrose (20% (w/v)) and maltose (20% (w/v)) were filter sterilized and stored at RT.

Glycerol stocks were prepared by resuspending a cell pellet of a 5 mL culture in 200 µL glycerol-Brock solution (Table 3) and stored at -80°C.

2.4. MOLECULAR BIOLOGICAL METHODS

2.4.1 Preparation of genomic DNA

For the cloning of genes from *E. coli* and *S. acidocaldarius* and for the analysis of *S. acidocaldarius* strains harboring respective expression plasmids a cell pellet of a 5 mL culture (OD_{600} of ~ 0.8) was used for the isolation of genomic DNA, including plasmid DNA, if present. For this, the DNeasy Blood & Tissue Kit (Qiagen) was applied following the manufacturer's protocol for purification of total DNA from Gram-negative bacteria. The incubation at 56°C was done for 1-1.5 h and the DNA was finally eluted in 100 µL RNase/DNase-free water (Gibco, Invitrogen). Genomic DNA was stored at -20°C.

Genomic DNA from *Thermoproteus tenax* was kindly provided by Dr. Theresa Kouril, genomic DNA from *Pyrococcus furiosus* was kindly provided by the group of Prof. Dr. Thomm (University of Regensburg) and genomic DNA from *Thermotoga maritima* was kindly provided by Dr. Verena Kallnik.

2.4.2 Isolation of plasmid DNA from *E. coli*

Plasmid DNA was prepared by using the GeneJET Plasmid Miniprep Kit (Thermo Scientific) following the manufacturer's instructions. RNase/DNase-free water (Gibco, Invitrogen) was used as eluent. Plasmid DNA was stored at -20°C.

Plasmid DNA of the metagenomic esterase clones was kindly provided by Dr. Jennifer Chow (work group of Prof. Streit, University of Hamburg, ExpresSys project partner) and stored at -20°C.

2.4.3 Quantification of DNA

DNA concentrations were measured by ultraviolet absorbance spectrophotometry (BioPhotometer, Eppendorf AG) at λ=260 nm (A_{260}). The purity of DNA can be analyzed by the ratio of the absorbance at 260 nm and 280 nm (A_{260}/A_{280}). The ratio of pure samples should be 1.8 – 2.0 (Sambrook and Russell, 2001). Ratios beyond this frame indicate a contamination of the sample by either proteins or phenol.

2.4.4 Agarose gel electrophoresis

Agarose gel electrophoresis was applied for the analysis and extraction of plasmid DNA, genomic DNA, restriction fragments and PCR products. In general, 1% (w/v) agarose gels were prepared in TAE-buffer (40 mM Tris-acetate, 1 mM EDTA, pH 8.5). DNA samples were mixed with

loading buffer (6 x DNA Loading Dye, Thermo Scientific) and applied to the gel. Depending on the designated separation of DNA bands and the chosen size of the gel chamber (distance from the electrodes) electrophoresis was done at 80-100 V for ~ 50 min at RT. For the determination of the size of the DNA samples a DNA size marker (GeneRuler 1 kb Ladder (Thermo Scientific)) was run in parallel. To visualize fluorescing DNA the gels were incubated in an ethidium bromide bath (500 µg/L) for 15-20 min and destained in H_2O for 5 min at RT. The results were documented under UV light in the GelDoc - Gel Documentation System and the Quantity One Software Package (Bio-Rad Laboratories).

2.4.5 Purification of DNA fragments

The purification of DNA fragments from agarose gels and PCR fragments was achieved by using Wizard® SV Gel and PCR Clean-up System (Promega). For the purification of restriction reactions the QIAquick Nucleotide Removal Kit (Qiagen) was used. All kits were used according to the manufacturer's protocol. RNase/DNase-free water (Gibco, Invitrogen) was used as eluent.

2.4.6 Amplification of DNA by Polymerase Chain Reaction (PCR)

PCR amplification was performed in a reaction mix with a total volume of 50 µL containing 50–100 ng genomic or plasmid DNA as template, 0.5 µM of each forward and reverse primer, 0.2 mM dNTPs (PeqLab), 1 x Phusion® HF buffer (Thermo Scientific), 3% (v/v) DMSO and 1 U of DNA polymerase (Phusion High-Fidelity DNA polymerase, Thermo Scientific). Oligonucleotide primers were purchased from Biolegio. Primers used for cloning are listed in Table 4 and primers used for sequencing in Table 5.

Standard amplification reactions included the following cycling steps, whereas the steps 2-4 were repeated for 24-30 cycles:

1.	initial denaturation	30 sec	98°C
2.	denaturation	30 sec	98°C
3.	annealing	30 sec	52-65 C (depending on primer set)
4.	elongation	30 sec/1000 bp	72°C
5.	final elongation	10 min	72°C

Table 4. List of primers and restriction enzymes used for cloning. Primer names, primer sequences, restriction sites, target vectors and PCR product sizes [bp] are given.

Construct name	Primer name	Sequence of primer (5' → 3')	Restriction sites	vector	Size [bp]
(x) P_{MalE}-TTX_1277-78	TTX_1277-78_*Nco*I_FOR	CCATGGCAAACCTCACCGAGAAA	*Nco*I	pMZ1	1812
	TTX_1277-78_*BamH*I_REV	TTAAAGGATCCGCCCGCTAGGTCCGGCCAATAGG	*BamH*I		
(†) P_{FBPA}-TTX_1277-78	TTX_1277-78_prom_*Bln*I_*Sac*II_FOR	TTAAACCTAGGCCGCGGGCTCCGCCTCAATGG	*Bln*I *Sac*II	pMZ1	1995
	TTX_1277-78_*BamH*I_REV	TTAAAGGATCCGCCCGCTAGGTCCGGCCAATAGG	*BamH*I		
(x) P_{MalE}-TTX_1277	TTX_1277_*Nco*I_FOR	CTGAACCCATGGCGATAGGAGTTCTGAC	*Nco*I	pMZ1	1030
	TTX_1277-78_*BamH*I_REV	TTAAAGGATCCGCCCGCTAGGTCCGGCCAATAGG	*BamH*I		
(x) P_{MalE}-PF1784	PF1784_*Nco*I_FOR	GGTATCCATGGCAGATGAAGTCAGAGAGCTCG	*Nco*I	pMZ1	1380
	PF1784_prom_*BamH*I_REV	CTTAAGGATCCCTGATGCCTTCTTAGGAGGGAG	*BamH*I		
(§) P_{PFK}-PF1784	PF1784_prom-*Kpn*I_*Sac*II_FOR	AAGTTTGGTACCCCGCGGTGAATGTGCATTTCTATTGG	*Kpn*I *Sac*II	pMZ1	1637
	PF1784_prom_*BamH*I_REV	CTTAAGGATCCCTGATGCCTTCTTAGGAGGGAG	*BamH*I		
(x) P_{MalE}-TM0209	TM0209-*Nco*I-FOR	ATCGCCATGGCGAAGATAGCAGTACTTA	*Nco*I	pMZ1	972
	TM0209-*BamH*I-REV	ACTGGATCCTGAAAGCATATGTGCTATTTCG	*BamH*I		
(†) P_{PFK}-TM0209	TM0209-prom-*Bln*I-*Sac*II-FOR	TTAAACCTAGGCCGCGGCTTTACCTCCAGTTCACTTC	*Bln*I *Sac*II	pMZ1	1183
	TM0209-*BamH*I-REV	ACTGGATCCTGAAAGCATATGTGCTATTTCG	*BamH*I		
pSVA2301::*lipS*	*lipS* -*Nhe*I_FOR	TAAGCTAGCATGAGCCCGAAAAGCAGGA	*Nhe*I	pSVA2301	863
	lipS -*Not*I_Stop_REV	TAAGCGGCCGCTCAGCTGTGCTTCCGGATGAACGC	*Not*I		
pSVA2301::*lipT*	*lipT* -*Nhe*I_FOR	AATGCTAGCATGCGGCGGTTACTAGCC	*Nhe*I	pSVA2301	1007
	lipT -short-*Not*I_Stop_REV	AATGCGGCCGCTCAGCGCACCCTAGGCGCCGCCT	*Not*I		
pSVA2301::*est5E5*	*est5E5-Nhe*I-FOR	AAATGCTAGCATGGTCGCTAGGGCGCAGGT	*Nhe*I	pSVA2301	1407
	*est5E5-Not*I_Stop_REV	ATAGCGGCCGCTCACTTCACGATGATGTCGAAGG	*Not*I		
pSVA2301::*lipA8b2*	*lipA8b2-Nhe*I-FOR	ACATGCTAGCATGAAAGGCTGTGGGCTGAT	*Nhe*I	pSVA2301	1273
	*lipA8b2-Not*I_Stop_REV	ATTAGCGGCCGCTTAAGGCCGCAAGCTCGCCAGTT	*Not*I		

pSVA2301::*estA3*	*estA3-Nhe*I-FOR	AAATGCTAGCATGAGCGCCGAAGAACTAGG	*Nhe*I	pSVA2301	1213
	*estA3-Not*I_Stop_REV	ATTAGCGGCCGCTTAGGCGGCGAGCGCGCTGTAG	*Not*I		
pSVA2301::*lipC5*	*lipC5-Nhe*I-For	AATGCTAGCATGGCGTCACTTGCTGCA	*Nhe*I	pSVA2301	980
	*lipC5-Not*I_Stop_REV	AATGCGGCCGCTCATGCAACAGCCTCCTGCGTGGT	*Not*I		
pSVA2301::*lacS*	*lacS*-SSO-*Nhe*I-FOR	AAAGCTAGCATGTACTCATTTCCAAATAG	*Nhe*I	pSVA2301	1490
	lacS-SSO-Stop-*Not*I-REV	AAAGCGGCCGCTTAGTGCCTTAATGGCTTT	*Not*I		
pET24a::Saci_1105 tagless	Saci_1105-*Nhe*I-For	AAAGCTAGCATGCCGTTGGATCCTGAAG	*Nhe*I	pET24a	939
	Saci_1105-Stop-*Xho*I-Rev	AAACTCGAGTTACTTAAAAATGTCCTTAAGTAG	*Xho*I		
pET24a::Saci_1105	Saci_1105-*Nhe*I-For	AAAGCTAGCATGCCGTTGGATCCTGAAG	*Nhe*I	pET24a	936
	Saci_1105-*Xho*I-Rev	AAACTCGAGCTTAAAAATGTCCTTAAGTAGTG	*Xho*I		
pET24a::Saci_1116 tagless	Saci_1116_*Nhe*I-For	AAAGCTAGCATGCCTTTAGATCCAACCATA	*Nhe*I	pET24a	936
	Saci_1116_Stop_*Xho*I_Rev	AAACTCGAGTCAAAAACTGTTGTAGAATACT	*Xho*I		
pET24a::Saci_1116	Saci_1116_*Nhe*I-For	AAAGCTAGCATGCCTTTAGATCCAACCATA	*Nhe*I	pET24a	933
	Saci_1116_*Xho*I_Rev	AAACTCGAGAAAACTGTTGTAGAATACTCG	*Xho*I		

(x) These PFK constructs were first cloned into the pre-vector pMZ1 via *Nco*I/*BamH*I, then subcloned into the expression vector pSVA1450 via *Nco*I/*Eag*I.
(†) These PFK constructs were first cloned into the pre-vector pMZ1 via *Bln*I/*BamH*I, then subcloned into the expression vector pSVA1450 via *Sac*II/*Eag*I.
(§) This PFK construct was first cloned into the pre-vector pMZ1 via *Kpn*I/*BamH*I, then subcloned into the expression vector pSVA1450 via *Sac*II/*Eag*I.

Table 5. List of primers used for sequencing.

Primer name	Sequence of primer (5' → 3')
pSVA1450 Seq F	CGGAGGTGTCCTTAAGTTTAG
pSVA1450 Seq R	CGGGCGTGATAAAGTCTGTCTC
pMZ1-seq-prom-FOR	GCATCCGAGGAGTCTACTTC
pSVA-MZ_seq_FOR_MZ	AACAAAACGTCTTTTACGGAAATAT
pMZ1-seq-his-REV_new	CCGGCAATCTAATGAAAATG
TTX1277-78mprom-seq-FOR	GGCTTTGAGTGGAAGATGTTTG
TM0209mprom-seq-FOR	GAAGTGATCGGAGTGAGAAG
PF1784 seq primer_FOR	CCAGTAGTCGAGGAAGACAAAC
pMZ1-seq-tag-R	AAGCGAGAAGAATCATAATGGG

2.4.7 Restriction of DNA

Plasmid DNA and PCR products were restricted by incubating the samples with the appropriate restriction enzymes (Table 4) at 37°C in an appropriate buffer following the manufacturer's instructions. Reaction volumes and amount of enzymes (FastDigest® enzymes (Thermo Scientific)) were dependent on DNA template and downstream applications.

The construct P_{MalE}-PF1784 contained an internal *Nco*I restriction site, which also had to be used for cloning. For the construct P_{MalE} – PF1784 this meant, that the PCR product (for cloning into the pre-vector pMZ1 via *Nco*I/*BamH*I) and later on the plasmid pMZ1::P_{MalE}-PF1784 (for cloning into the expression vector pSVA1450 via *Nco*I/*Eag*I) had to be digested following a partial digest protocol. For this, conventional restriction enzymes and an appropriate buffer were used (Thermo Scientific). The DNA template was first incubated in three separate reaction tubes with *BamH*I for 30 min at 37°C. Afterwards the samples were transferred to RT and the restriction enzyme *Nco*I was added. The restriction reaction was continued for further 1, 2 and 3 min, respectively. The reactions were analyzed via agarose gel (see 2.4.4) electrophoresis and the DNA bands with the target size were purified from the gel (see 2.4.5).

The construct P_{FBPA}-TTX_1277-78 contained an internal *Sac*II restriction site, which had to be used for cloning into the expression vector pSVA1450. Since *Sac*II is only available as conventional enzyme and *Eag*I only as FastDigest® enzyme, a sequential partial digestion of the plasmid pMZ1::P_{FBPA}-TTX_1277-78 was performed. Plasmid DNA was first digested with the enzyme *Eag*I and purified by using the QIAquick Nucleotide Removal Kit (Qiagen) (see 2.4.5). This purified DNA template was then applied to a partial digest reaction with *Sac*II as described above. The reactions were analyzed via agarose gel electrophoresis (see 2.4.4)and the DNA bands with the target size were purified from the gel (see 2.4.5).

2.4.8 Ligation

DNA ligation was performed by incubating the restricted DNA fragments with the accordingly restricted linearized vector in the presence of T4 DNA ligase (Thermo Scientific). The DNA ligase catalyzes the formation of phosphodiester bonds between a 3'-hydroxyl group and a 5'-phosphate group. Vector DNA (50-75 ng) and insert DNA were mixed in a molar ratio of 1:3 in a total volume of

8 µL and incubated at 45°C for 5 min to break possible secondary structures. Subsequently 1 U T4 DNA ligase and 1 µL 10 x reaction buffer (Thermo Scientific) were added. The ligation reaction was carried out overnight at 16°C and then stopped at 65°C for 10 min. The complete ligation mix was immediately used for transformation.

2.4.9 Preparation of chemically competent *E. coli* cells

Competent cells from *E. coli* DH5α, ER1821 and Rosetta (DE3) strains were prepared by using the modified calcium chloride method. Shortly, 200 mL LB-medium (supplemented with appropriate antibiotic if necessary) was inoculated with 1% (v/v) of an *E. coli* overnight culture and incubated at 37°C (180 rpm) until an OD_{600} of 0.4 was reached. The cell suspension was cooled down on ice for 10 min and then centrifuged at 4,000 x g for 7 min at 4°C. All subsequent steps were strictly carried out on ice. The cell pellet was gently resuspended in 10 mL ice cold buffer (60 mM $CaCl_2$, 10 mM PIPES, 15% (v/v) glycerol, pH 7.0). After a second centrifugation step (4,000 x g, 5 min, 4°C), the cell pellet was resuspended in 2 mL of the same buffer. Aliquots of 80 µL were quick-frozen in liquid nitrogen and stored at -80°C.

2.4.10 Transformation of competent *E. coli* cells

Recombinant DNA (0.5-1 µg plasmid DNA or 10 µL ligation mix) was incubated with competent *E. coli* cells on ice for 30 min. Transformation was achieved by a heat shock treatment at 42°C for 80 sec. Cells were transferred immediately on ice, 500 µL LB-medium was added and cell suspension was incubated at 37°C for 50 min (180 rpm). 100 µL of transformed cells were spread on a LB-agar-plate containing the respective antibiotics. The remaining sample was centrifuged (6,000 x g, 2 min, RT), the cell pellet resuspended in 100 µL LB-medium and spread on a second LB-agar-plate. After incubation over night at 37°C, single colonies were used for inoculation of 5 mL of liquid LB medium, supplemented with the appropriate antibiotic, for plasmid isolation in order to identify positive clones via test restrictions. If no single colonies were needed (e.g. transformation of *E. coli* ER1821), the LB-medium, supplemented with the appropriate antibiotics, was directly inoculated with 20-50 µl of the transformation mixture. The incubation was performed at 37°C overnight under aerobic conditions.

2.4.11 Identification of correctly ligated recombinant plasmid DNA

E. coli clones harboring the putative recombinant plasmid were identified by performing a test restriction. For this, plasmid DNA was isolated from the clones (5 mL culture, see 2.4.2) and applied in a 10 µL restriction reaction (see 2.4.7) by using ~ 400 ng plasmid DNA and 0.5 µL FastDigest® enzyme. The restriction pattern was analyzed via agarose gel electrophoresis (see 2.4.4). If the pattern was as expected the respective clone was sequenced (see 2.4.12).

2.4.12 DNA sequencing

Automated DNA sequencing was performed by LGC Genomics. Per sequencing reaction 12 µL of plasmid DNA (100 ng/µL) and 2 µL of specific primer (10 µM) or, if appropriate primer were available at LGC Genomics, 15 µL of plasmid DNA (100 ng/µL) was used.

2.4.13 Preparation of competent *S. acidocaldarius* cells

A pre-culture of a *S. acidocaldarius* strain was used for inoculation of 50 mL Brock medium (Table 2), supplemented with 0.1% (w/v) NZ-amine, 0.2% (w/v) dextrin and 20 mg/L uracil. The cells were then grown overnight at 78°C to an OD_{600} of 0.1-0.3, incubated on ice for 20 min and then centrifuged at 2,000 x g for 20 min at 4°C. All subsequent steps were strictly carried out on ice. The cell pellet was washed three times in ice cold, autoclaved 20 mM sucrose solution (50 mL, 25 mL, 1 mL). After a final centrifugation step (2,000 x g for 20 min at 4°C) the cells were resuspended in sucrose solution to a theoretical OD_{600} of 10 and stored in 50 µL aliquots at -80°C.

2.4.14 Transformation of competent *S. acidocaldarius* cells via electroporation

Before the transformation of competent *S. acidocaldarius* cells with recombinant plasmid DNA via electroporation, the DNA needs to be methylated to escape the restriction by *Sua*I from *S. acidocaldarius*. For this, the respective plasmid DNA was used for transformation of *E. coli* ER1821 cells (see 2.4.10). This strain harbors the plasmid pM.EsaBC4I (kan^r) encoding a methylase, which is responsible for the correct methylation of plasmid DNA. The methylated plasmid DNA was subsequently isolated from the *E. coli* ER1821 cells and used for transformation of competent *S. acidocaldarius* cells. 50 µL competent cells were gently thawed on ice, mixed with 400 ng methylated plasmid DNA, incubated on ice for 5 min and transferred to a pre-cooled electroporation cuvette (0.1 cm, Bio-Rad Laboratories). Electroporation was done at 1500 V, 600 Ω and 25 µF by using the Gene Pulser Xcell™ Microbial System (Bio-Rad Laboratories). Cells were immediately mixed with 50 µl pre-warmed (RT) 2 x regeneration buffer (1% (w/v) sucrose, 20 mM β-alanine-malate buffer (20 mM β-alanine, 20 mM malate-solution, pH adjusted to 4.5 using malate-solution), 10 mM $MgSO_4$ x 7 H_2O) and incubated at 75°C for 30 min at 600 rpm. Cells were finally spread on gellan gum solidified Brock medium plates (supplemented with 0.1% (w/v) NZ-amine and 0.2% (w/v) dextrin) and incubated at 78°C for 3-5 d.

2.4.15 Identification of *S. acidocaldarius* clones harboring recombinant expression plasmid

S. acidocaldarius clones harboring the recombinant expression plasmid were identified by either performing a test restriction or colony PCR. For a test restriction, the plasmid DNA was isolated by using the DNeasy Blood & Tissue Kit (Qiagen) (see 2.4.1) and applied in a restriction reaction as described in 2.4.11. For a colony PCR, the *S. acidocaldarius* clones were transferred to a new gellan gum solidified Brock medium plate by using a pipette tip and subsequently lysed in 30 µL 0.2 M NaOH for 10 min at RT. The lysed cells were neutralized with 70 µL Tris-HCl-buffer (260 mM, pH 7.8) and

1 µL were applied to a PCR using a vector-specific primer and a gene-specific primer. PCR was performed as described in 2.4.6.

2.4.16 RNA isolation from *S. acidocaldarius* cells and DNaseI digestion

For RNA isolation from *S. acidocaldarius* cells from 75 mL culture were homogenized in 1 mL TRIzol (Invitrogen) by rigorous pipetting and incubated at 65°C for 10 min. This suspension was supplemented with 200 µL chloroform, incubated at RT for 5 min while shaking by hand and centrifuged at 16,000 x g for 15 min. Subsequently the aqueous phase was transferred to a fresh reaction tube, 500 µL iso-propanol was added and mixed gently by inverting. For RNA precipitation, the sample was incubated at -80°C overnight. After centrifugation at 16,000 x g for 30 min at 4°C the RNA pellet was washed with 1 mL cold 70% (v/v) ethanol, followed by another centrifugation step 16,000 x g for 5 min at 4°C. Ethanol was removed with a pipette and the pellet was air-dried at RT for 5 min before resuspending in 50 µL RNase-free water (Invitrogen). Finally, 20 U Recombinant RNasin® Ribonuclease Inhibitor (Promega) was added to the sample and incubated at 45°C for 5 min in a thermoshaker for homogenization. RNA was stored at -80°C.

For DNaseI digestion, 2 µg RNA was mixed with 1 µL 10x reaction-buffer (without $MnCl_2$, Thermo Scientific), 0.25 µL 100 mM $MnCl_2$ (Thermo Scientific) and 1 µL DNaseI (Thermo Scientific) in a total volume of 10 µL. The reaction was incubated at 37°C for 1 h before adding another 1 µL DNaseI. Again, the sample was incubated at 37°C for 1 h. Finally, the reaction was stopped by adding 1 µL 25 mM EDTA (pH 8) and incubation at 65°C for 10 min. RNA was stored at -80°C. Complete digestion was verified by PCR as described in 2.4.17.

2.4.17 cDNA synthesis and PCR

cDNA synthesis was carried out by using the GoScript™ Reverse Transcription System (Promega). For this, 5 µg RNA was mixed with 1 µl random primer in a final volume of 13 µL, incubated at 70°C for 10 min and transferred immediately on ice. Subsequently, 4 µl GoScript™ 5 x Reaction Buffer, 1 µl 25 mM $MgCl_2$, 1 µl dNTPs and 1 µl GoScript™ Reverse Transcriptase were added in the given order to the sample. For annealing, the samples were incubated at RT for 10 min, then transferred to 42°C and incubated for 1 h for elongation. The reverse transcriptase was inactivated by incubation at 70°C for 10 min. Synthesized cDNA was either used directly in a PCR or stored at -20°C.

Basically, this PCR is performed as described in 2.4.6 with few adjustments. 40 ng cDNA or 80 ng RNA was applied to a 30 µl PCR mix containing 0.3 µM of each forward and reverse primer (Table 6), 0.2 mM dNTPs (PeqLab), 1 x Phusion® HF buffer (Thermo Scientific) and 0.2 U of DNA polymerase (Phusion High-Fidelity DNA polymerase, Thermo Scientific). The cycling program was as followed (steps 2-4 were repeated for 30 cycles):

1.	initial denaturation	20 sec	98°C
2.	denaturation	10 sec	98°C
3.	annealing	30 sec	50 or 59°C
4.	elongation	1 min	72°C
5.	final elongation	10 min	72°C

For esterase specific primer sets an annealing temperature of 50°C was chosen, whereas for the control *secY* primer set an annealing temperature of 59°C was applied. PCR products were analyzed via agarose gel electrophoresis (see 2.4.4).

Table 6. List of primers used for PCR with cDNA-templates. Primer names, primer sequences, annealing temperature used [°C] and PCR product sizes [bp] are given.

Primer name	Sequence of primer (5' → 3')	Size [bp]
qRT_*lipS*_FOR	AAAGGAGTGGTTCCGGTATG	133
qRT_*lipS*_REV	AAGCGCATCGACTGCGGTGAAC	
qRT_*lipT*_FOR	GAGGTTCCAGGTGGCGTATG	175
qRT_*lipT*_REV	GAAGAAGCTCCTGCACATCC	
qRT_*est5E5*_FOR	TCGTGGCGCTCAACTACTAC	187
qRT_*est5E5*_REV	AGCAGGAAGGTCTCGCTGTC	
qRT_*lipA8b2*_FOR	TGGTTTGTGTTCGGCCTATC	136
qRT_*lipA8b2*_REV	CCCAGTACTTGAACCCAAAC	
qRT_*estA3*_FOR	ATGACCAAACCGGTCACCTC	175
qRT_*estA3*_REV	CAATCCGCATCGGCATTGTC	
qRT_*lipC5*_FOR	AAGCGCGACATTCAGGACCC	149
qRT_*lipC5*_REV	TTGACGGTGCTCATCCACTG	
Sas*ecY*_qRT_ F	TGGCGTCAATTCCTTTATATGG	101
Sas*ecY*_qRT_ R	CGCAAGTGTACCCTGTGATG	

RNA isolation, cDNA synthesis and PCR experiments were carried out by Marcel Blum.

2.5 BIOCHEMICAL METHODS

2.5.1 Heterologous expression of recombinant proteins in *E. coli* Rosetta (DE3) and preparation of cell-free extracts

Protein expression was performed in erlenmeyer flasks containing 2 L LB-medium and the respective antibiotics inoculated with 1-2% (v/v) of an overnight culture (see 2.3.1) was used. Protein expression was induced at an OD_{600} of 0.5-0.6 by adding 1 mM isopropyl-beta-D-thiogalactopyranoside (IPTG) and growth was continued for 3-4 hours. Cells were chilled on ice and harvested by centrifugation (6,000 x g, 15 min, 4°C). The obtained cell pellet was stored at -80°C.

Cell pellets were resuspended in respective buffers (Table 7) at a ratio of 3 mL/g cells. Cells were disrupted by either passing three times through a French pressure cell at 20,000 psi or by sonication (3 x 5 min, amplitude 50%, time cycle 0.5) while the cell suspension was chilled on ice. Finally, the samples were centrifuged (16,000 x g, 45 min, 4°C) to separate soluble cell-free extracts from the insoluble fraction.

2.5.2 Purification of recombinant target proteins from *E. coli* cells

The cell-free extracts were prepared as described in section 2.5.1 and subsequently diluted 1:1 with lysis buffer (see Table 7 for the respective buffers). The protein solution was incubated at high

temperature (see Table 7 for respective temperature and time settings) in order to denature the majority of thermolabile *E. coli* host proteins. The denatured proteins were removed by centrifugation (16,000 x g, 30 min, 4°C).

Protein samples were either directly applied to a chromatography step (esterases Saci_1105, Saci_1116) or dialyzed (GDH-2 SSO3204, SSO3204_V93N, SSO3204_E294H, SSO3204_V93N_E294H) using dialysis tubes (Spectra/Por® Dialyse Membrane, Molecular Weight Cut off (MWCO): 3.5 kDa, Spectrum Laboratories). Dialysis was performed in 5 L IC-buffer (Table 7) overnight at 4°C, in order to remove salts and low molecular weight substances, interfering with subsequent purification steps or activity assays and in order to change buffer conditions.

Table 7. Purification of recombinant proteins heterologously expressed in *E. coli* Rosetta (DE3). Buffer conditions for cell lysis (lysis) and chromatography steps (Ni-TED affinity chromatography (Ni-TED). ion exchange chromatography (IC), hydrophobic interaction chromatography (HIC), gel filtration (GF)) are given, as well as temperature and time settings used for heat precipitation (HP).

enzyme	**lysis**	**HP**	**Ni-TED**	**IC**	**HIC**	**GF**
Saci_1105	1 x LEW (Protino® Ni-TED Kit)	80°C, 30 min	1 x LEW, 1 x elution buffer	---	---	---
Saci_1116	1 x LEW (Protino® Ni-TED Kit)	80°C, 30 min	---	---	---	---
SSO3204	100 mM HEPES/KOH (pH 7, RT)	70°C, 30 min	---	20 mM HEPES/KOH (pH 7, RT)	20 mM HEPES/KOH (pH 7, RT), 3 M NaCl	50 mM HEPES/KOH (pH 7, RT), 300 mM NaCl
SSO3204 _V93N	100 mM HEPES/KOH (pH 7, RT)	70°C, 30 min	---	20 mM HEPES/KOH (pH 7, RT)	---	50 mM HEPES/KOH (pH 7, RT), 300 mM NaCl

SSO3204 _E294H	100 mM HEPES/KOH (pH 7, RT)	70°C, 30 min	---	20 mM HEPES/KOH (pH 7, RT)	---	50 mM HEPES/KOH (pH 7, RT), 300 mM NaCl
SSO3204 _V93N _E294H	100 mM HEPES/KOH (pH 7, RT)	70°C, 30 min	---	20 mM HEPES/KOH (pH 7, RT)	---	50 mM HEPES/KOH (pH 7, RT), 300 mM NaCl

2.5.2.1 Ion exchange chromatography (IC)

The respective GDH-2 enzyme variants (wild-type and mutants) were applied onto ion exchange chromatography via Resource Q (GE-Healthcare Life Sciences; column bed volume (CV): 6 mL, maximal protein loading: 270 mg, flow rate: 2 mL/min, column dimension: 16x 30 mm, maximal operating pressure: 87 psi). The column was equilibrated with 60 mL IC-buffer (Table 7; 10 x CV) prior to sample application. Once the sample was applied, the column was washed with 2 x CV IC-buffer and proteins were separated by a linear salt gradient from 0-1 M NaCl using 20 x CV of IC-buffer. Fractions containing the enriched recombinant enzyme (verified by SDS-PAGE and enzyme activity (see 2.5.7 and 2.6.3)) were pooled and dialyzed, if applied to a HIC column, or concentrated via centrifugal concentrators (Vivaspin6, Sartorius AG), if applied to a GF column.

2.5.2.2 Hydrophobic interaction chromatography (HIC)

The protein sample obtained from the IC column was dialyzed against 5 L HIC-buffer (Table 7) overnight at 4°C. The dialyzed sample was subsequently applied onto hydrophobic interaction chromatography via butyl sepharose media (self packed; CV 7.5 mL, flow rate: 4 mL/min, maximal operating pressure: 200 psi). The column was equilibrated with 10 x CV HIC-buffer, sample was applied to the column and the column was washed again with 2 x CV HIC-buffer. The target protein was eluted by a decreasing linear salt gradient from 3-0 M NaCl using 20 x CV of HIC buffer. Fractions containing the enriched protein (verified by SDS-PAGE and enzyme activity (see 2.5.7 and 2.6.3)) were pooled and concentrated via centrifugal concentrators (Vivaspin6, Sartorius AG) before application to a GF column.

2.5.2.3 Gel filtration (GF)

Protein samples obtained after IC or HIC, respectively, were finally subjected to a gel filtration column. For this, the HiLoad 26/60 Superdex 200 prep grade column (GE-Healthcare Life Sciences; column volume: 330 mL, flow rate: 2 mL/min, separation range of proteins: 10-600 kDa, maximal operating pressure: 50 psi) was washed with 1.2 x CV filtered distilled H_2O and equilibrated in 1.2 x CV GF-buffer (Table 7). The protein samples were dialyzed against 5 L GF-buffer overnight at 4°C before application to the column via a sample loop and separated using 1.5 x CV GF-buffer. Fractions

containing the enriched protein (verified by SDS-PAGE and enzyme activity (see 2.5.7 and 2.6.3)) were pooled and stored at 4°C until subjected to enzyme characterization.

2.5.2.4 Ni-TED affinity chromatography

Protein sample (*Saci*_1105) retrieved after heat precipitation was applied to Ni-TED affinity chromatography according to the manufacturer's protocol (Protino® Ni-TED Kit, Machery-Nagel). Shortly, the Ni-TED column was equilibrated with 1 x LEW buffer (50 mM NaH_2PO_4, 300 mM NaCl, pH adjusted with NaOH to 8) prior to sample application. The column was then washed twice with 1 x LEW buffer before the protein was eluted successively in 3 steps with 1 x elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imidazole, pH adjusted with NaOH to 8). Fractions were analyzed via SDS-PAGE and enzyme activity assays (see 2.5.7 and 2.6.2) and stored at 4°C.

2.5.3 Expression and purification of the thermophilic auxiliary enzyme GAPN

To determine PFK enzyme activity at high temperature in a continuous assay the production of the thermophilic auxiliary enzyme GAPN was necessary, which had to be provided in excess in the assay, so the velocity of the target enzyme PFK was not limited. For this, it was sufficient to enrich the auxiliary enzyme via heat-precipitation. The expression strain (pET24d::*SSO*3194 in *E. coli* Rosetta (DE3)) was kindly provided by Dr. Theresa Kouril.

Expression of the *Sso*-GAPN was carried out in a 2 L Erlenmeyer flask containing 1 L LB medium at 37°C and 180 rpm. The cells were cultivated, harvested and disrupted in a buffer containing 100 mM HEPES/KOH (pH 6-5, 70C) and 5 mM DTT as described above (see 2.5.1). The cell-free extract was then applied to heat-precipitation at 80°C for 20 min and denatured proteins were removed by centrifugation (16,000 x g, 30 min, 4°C). The retrieved supernatant was dialyzed using dialysis tubes (Spectra/Por® Dialyse Membrane (Molecular Weight Cut off (MWCO): 3.5 kDa, Spectrum Laboratories, Inc.)), in order to remove salts and low molecular weight substances, interfering with subsequent enzyme assays. *Sso*-GAPN was stored at -80°C supplemented with glycerol (20% (v/v).

2.5.4 Heterologous expression in *S. acidocaldarius*, preparation of cell-free extracts and purification of recombinant target proteins

For heterologous expression in *S. acidocaldarius* 400 mL-1 L Brock medium (Table 2), supplemented with 0.1% (w/v) NZ-amine, 0.2% (w/v) sucrose and 0.4% (w/v) maltose, was inoculated with a pre-culture (exponential phase, OD_{600} of ~ 0.6) to an start OD_{600} of ~ 0.05. Expression cultures were grown in Erlenmeyer flasks and harvested when an OD_{600} of ~ 0.8-0.9 was reached. For this, the cultures were chilled on ice for 20 min and then centrifuged at 6,000 x g for 15 min at 4°C. The cell pellets were stored at -80°C.

The cell pellets were resuspended in 1 x LEW buffer, containing 1 x protease inhibitor cocktail (7 x stock: 1 tablet cOmplete Mini, EDTA free (Roche) in 1.5 mL 100 mM HEPES, pH 7.0 (storage at -20°C)), at a ratio of 3 mL/g cells. Cells were disrupted by sonication (3 x 5 min, amplitude 50%, time cycle 0.5) while the cell suspension was chilled on ice. Finally, the samples were centrifuged (16,000 x g, 45 min, 4°C) to separate soluble cell-free extracts from the insoluble fraction. Crude

extracts were immediately subjected to purification via Ni-TED affinity chromatography as described in 2.5.2.4.

2.5.5 Protein quantification

Protein concentrations were determined using the Bio-Rad Protein assay based on the method of Bradford (Bradford, 1976), following the manufacturer's protocol. Bovine serum albumin (BSA) was used as standard (1-25 g/L). For the determination of the target protein, samples were diluted with water to a final volume of 600 µL, 400 µL Bradford-reagent were added and incubated for 5 min in the dark (RT). The absorbance was determined at λ=595 nm (Biophotometer, Eppendorf).

2.5.6 Protein precipitation

The protein sample was supplemented with 30% (w/v) trichloroacetic acid at a ratio of 2:1 (final TCA concentration 10% (w/v)) and centrifuged at 16,000 x g for 10 min at 4°C. The pellet was resuspended in Aqua bidest and the pH was neutralized with 2-4 µL 1 M Tris/HCl, pH 9, if necessary.

2.5.7 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

Proteins were analyzed via SDS-PAGE according to Laemmli (Laemmli, 1970). By adding sodium dodecylsulfate (SDS) polypeptides denature into their primary structure. The concentration of polyacrylamide (PAA) was adjusted to the molecular mass of the proteins of interest, whereas for larger proteins a lower concentration of PAA is needed. In general, 12.5% (v/v) SDS-polyacrylamide gels were used. The stacking gel consisted of 4.0% (v/v) acrylamide-bisacrylamide (30%), 125 mM Tris (pH 6.8, RT), 0.1% (w/v) SDS, 0.045% (w/v) ammoniumpersulfate (APS; 100 g/L) and 0.1% (v/v) N,N,N´,N´-Tetramethylethylenediamine (TEMED). The separation gel consisted of 12.5% (v/v) acrylamide-bisacrylamide (30%), 375 mM Tris (pH 8.8, RT), 0.1% (w/v) SDS, 0.067% (w/v) APS (100 g/L), 0.05% (v/v) TEMED. The separation gel was poured into the gel casting chamber, overlaid with iso-propanol (100%). Polymerization was achieved after 20 min, iso-propanol was removed and the stacking gel was poured on top. The comb was inserted and after 20 min of polymerization the gel was stored at 4°C or used directly. Electrophoresis was performed at constant 25 mA using a Mini-Protean 3-System (Bio-Rad Laboratories). The electrophoresis buffer was composed of 25 mM Tris-HCl, 190 mM glycerol and 0.1% (v/v) SDS (pH 8.3). Before application, the protein samples were mixed with loading buffer to a final concentration of 62.5 mM Tris-HCl, pH 6.8, 10% (v/v) glycerol, 2% (w/v) SDS, 5% (w/v) β-mercaptoethanol, and 0.005% (w/v) bromophenol blue and then denatured by incubation at 95°C for 5 min. The samples were then applied onto the SDS gel. For size determination of the target proteins 5 µL of a protein size marker (PageRuler™ Unstained Protein Ladder, 14.2-66.0 kDa, Thermo Scientific). To visualize the proteins, gels were stained with Coomassie Brilliant Blue (0.05% (w/v) Coomassie Brilliant Blue G-250, 40% (v/v) ethanol and 10% (v/v) acetic acid) for 45 min at RT and destained overnight in A. bidest. Gel analysis and documentation was performed using the Gel Doc System (Bio-Rad Laboratories) and the Quantity One Software Package (Bio-Rad Laboratories).

2.5.8 Immuno-blotting

Analysis of proteins fused to a 10xhistidine/Strep-tactin-tag was done via immuno-blotting with anti-Strep-tactin-antibody (*Strep*-Tactin® AP conjugate, iba). For this, a tank blot system (Bio-Rad Laboratories) was used. The samples were first applied to a SDS-PAGE (see 2.5.7) using additionally the PageRuler™ prestained Protein Ladder (Thermo Scientific) as protein size marker. After electrophoresis, the SDS-gel was equilibrated in transfer buffer (25mM Tris, 190 mM glycerol, 20% (v/v) methanol) for 30 min at RT. For blotting a Roti®-PVDF membrane (Carl Roth GmbH) was used, which was washed in 100% (v/v) methanol for 15 min, then washed in H_2O for 2 min and finally equilibrated in transfer buffer for 5 min. The transfer was performed at 20 C at 4°C overnight.

For detection the membrane was handled as followed:

fixation	5 min in 20% (v/v) methanol
washing 1	3 x 5 min in TBST
blocking	2 h in TBST + 0.2% (w/v) I-Block (Tropix, Applied Biosystems)
washing 2	3 x 5 min in TBST + 0.1% (w/v) I-Block
hybridization	1 h in TBST + 0.1% (w/v) I-Block + 1:4000 *Strep*-Tactin® AP conjugate (iba)
washing 3	3 x 5 min in TBST + 0.1% (w/v) I-Block
washing 4	2 x 5 min in TBS

TBST: TBS (500 mM Tris-HCl, 1.5 M NaCl, pH 7.6) + 0.1% (v/v) Tween20 for washing 1 step. For all other steps the TBST buffer contained 0.3% (v/v) Tween20.
All steps were performed at RT on a rocking platform.

Finally, the membrane was incubated with 1 mL CDP-Star® (Sigma-Aldrich) for ~ 5 min and chemiluminescence was detected with the VersaDoc Imaging System (Bio-Rad Laboratories).

2.6 Enzyme Assays

2.6.1 Phosphofructokinase (PFK)

fructose 6-phosphate + ATP/ ADP/ PP_i → fructose 1,6-bisphosphate + ADP/AMP/P_i

The PFK activity was determined in a continuous assay at 70°C in 100 mM Tris-HCl buffer, pH 6.5 (70°C) with ~ 10 µg protein (E1 fraction after Ni-affinity chromatography). The reaction mixture (500 µL) contained the following:

Table 8. Assay composition for the determination of PFK activity.

	ATP-PFK (*Tma*)	**ADP-PFK (*Pfu*)**	**PP_i-PFK (*Ttx*)**
F6P	10 mM	10 mM	10 mM
ATP/ADP/PP*i*	2.5 mM	2.5 mM	5 mM
$MgCl_2$	2.5 mM	10 mM	1 mM
FBPA (rabbit muscle)	5 U	5 U	5 U
TIM (rabbit muscle)	10 U	10 U	10 U
GAPN (*Sso*, HP, see 2.5.3)	20 µL	20 µL	20 µL
Tris-HCl buffer (pH 6.5, 70°C)	100 mM	100 mM	100 mM
$NADP^+$	2 mM	2 mM	2 mM
protein	~ 10 µg	~ 10 µg	~ 10 µg

The reaction was started by adding the substrate F6P and was followed by an increase in absorbance at 340 nm at a Specord 210 photometer (Analytik Jena).

2.6.2 Esterase

Esterase activity was determined with different ρ-nitrophenyl esters (ρNP-butyrate (C4), ρNP-octanoate (C8) and ρNP-decanoate (C10); ρNP-dodecanoate (C12)) by measuring the amount of ρ-nitrophenol released photometrically at 410 nm. The assay, if not stated otherwise, was performed in 100 mM acetate buffer (pH 6.5, 70°C) in a reaction volume of 500 µL containing 1 mM ρ-nitrophenyl substrate and protein (5 µg crude extract or 2 µg purified protein after Ni-TED affinity chromatography). Buffer and protein were pre-incubated at the reaction temperature of 70°C for 2 min and the reaction was initiated by adding the substrate. The reaction was incubated for 5 min before stopping by chilling on ice for 2 min. Final centrifugation at 16,000 x g for 5 min at 4°C lead to a clear supernatant, from which 200 µL were applied to a photometric measurement at 410 nm in a plate reader. The effect of auto-hydrolysis of substrates was measured in parallel and subtracted from the enzymatic measurements. The extinction coefficient of ρ-nitrophenol was determined for the buffer condition prior to the measurements by mapping a standard curve with ρNP-butyrate as substrate.

2.6.3 Glucose dehydrogenase (GDH)

$$\text{glucose} + NAD(P)^+ \rightarrow \text{gluconate} + NAD(P)H + H^+$$

The GDH activity was determined in a continuous assay at 70°C by following an increase in absorbance at 340 nm at a Specord 210 photometer (Analytik Jena). The standard screening reaction mixture (500 µl) contained 100 mM HEPES buffer (pH 6.5, 70°C), 0.3 µg protein (after gel filtration), 10 mM $NAD(P)^+$, 10 mM $MgCl_2$ and 10 mM substrate. The reaction was started by the addition of the substrate.

2.6.4 β-galactosidase

β-galactosidase activity was measured by the conversion of β-X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) to galactose and 5-bromo-4-chloro-3-hydroxyindole, which dimerizes and is oxidized into 5,5'-dibromo-4,4'-dichloro-indigo. The latter is a distinct blue product. 50 µL of *S. acidocaldarius* crude extract was supplemented with 25 µL of 50 mg/mL β-X-gal solution and the reaction was incubated for 10 min at 75°C. The β-X-gal solution was prepared in dimethyl sulfoxide (DMSO).

2.7 KINETIC PARAMETERS

Enzyme activity was measured photometrically by determining the linear change in absorption per min (ΔE/min). From this the specific activity [U/mg] was calculated as followed:

$$spec.act.[U/mg] = \frac{\Delta E/\min \cdot V_{Total}}{\varepsilon \cdot d \cdot c \cdot V_{Protein}}$$

ΔE/min	extinction change per minute
V_{Total}	total volume of the assay (mL)
$V_{Protein}$	volume of the protein solution applied in the assay (mL)
d	thickness of cuvette (cm)
c	protein concentration of protein solution applied in the assay (mg protein/mL)
ε	extinction coefficient: ε_{NADH}= 5.8 mM^{-1} cm^{-1} (70°C); ε_{NADPH}= 5.71 $mM^{-1}cm^{-1}$ (70°C); ε_{pNP}= 1.2 mM^{-1} cm^{-1} (70°C)

One unit of enzyme activity was defined as the enzyme activity catalyzing the conversion of 1 µmol substrate/min.

Kinetic analyses of *Sso*-GDH-2 and its variants were performed using the classical Michaelis-Menten equation and the accordingly linear Hanes Woolf plot.

V_{max} was defined as the maximum reaction velocity rate achieved at maximum (saturating) substrate concentrations and was determined by iterative curve-fitting following the Michaelis-Menten equation.

K_m represented the substrate concentration at which the reaction rate was half of V_{max}, if the enzyme follows classical Michaelis-Menten kinetics, and was determined via a Hanes Woolf plot ([S]/v versus [S]). The intercept in Hanes plot is K_m/ V_{max} and the slope is 1/V_{max}.

The turnover number k_{cat} was defined as the maximum number of substrate molecules an enzyme converts to product per catalytic site per unit of time. k_{cat} was calculated as followed:

$$k_{cat}[s^{-1}] = \frac{V_{max} \cdot MW}{60}$$

V_{max}	maximum reaction velocity rate [µmol/mg·min]
MW	molecular weight of protein applied in assay [µmol/mg]

The catalytic efficiency of an enzyme was calculated as k_{cat}/K_m.

2.8 SOFTWARE, INTERNET DATABASES AND TOOLS

- BPROM was used for the prediction of promoter elements in bacterial promoter regions (http://linux1.softberry.com/berry.phtml?topic=bprom).
- BRENDA was used to retrieve enzyme functional data and information about appropriate literature (http://www.brenda-enzymes.org/).
- CHROMAS was used for the visualization of sequence chromatograms
- CloneManager 7.0 was used to find appropriate vectors and restriction sites for endonucleases, for primer design and to analyze the results of sequencing and to rule out mutations in the nucleotide sequence.
- ExPASy was used for calculation of theoretical molecular weight of proteins (http://web.expasy.org/compute_pi/)
- KEGG was used as a Gene and Genome database (http://www.genome.jp/kegg/).
- NCBI BLAST was used for investigation of sequence similarity (http://blast.ncbi.nlm.nih.gov/Blast.cgi).
- NCBI Protein database was used to retrieve protein sequences (http://www.ncbi.nlm.nih.gov/protein)
- Origin 9 was used to plot curves (Microcal Software Inc.).
- Quantity One was used for gel documentation.
- SignalP was used for the prediction of signal peptides (http://www.cbs.dtu.dk/services/SignalP/).
- TMHMM was used for transmembrane helices prediction (http://www.cbs.dtu.dk/services/TMHMM-2.0/).
- UniProt was used as a resource of protein sequence and functional information (http://www.uniprot.org/).
- WINASPECT (Analytik Jena) was used for enzyme assay performance with Specord 210 photometer (Analytik Jena).

Chapter 3

RESULTS

CHAPTER 3 RESULTS

In course of the present work as part of the BMBF funded ExpresSys project, investigations towards the establishment of *S. acidocaldarius* as expression host particularly for functional metagenome library screenings were performed. For this, and to make the system more variable for different applications, alternative promoters, controlling recombinant gene expression in *S. acidocaldarius*, were comparatively analyzed to the *malE* promoter. Different archaeal and bacterial promoters controlling PFK encoding genes were cloned in the pSVA1450 expression vector available for *S. acidocaldarius* and expression efficiency was determined for the different promoters using their own natively controlled PFK encoding genes as reporters. Furthermore, a suitable *S. acidocaldarius* strain had to be constructed for functional metagenome screenings for esterase activity. Wild-type and esterase deficient *S. acidocaldarius* strains were characterized for their ability to degrade acyl esters of different chain length and the deleted *S. acidocaldarius* esterase genes were heterologously expressed and the recombinant esterases were characterized. The esterase deficient strain was used for functional expression of metagenome derived esterase genes from different sources to investigate its suitability for functional library screenings.

A further aim of the work was to investigate the structural determinants of the different substrate specificities of the sugar dehydrogenase paralogs SSO3003 and SSO3204 in *S. solfataricus*. Mutant proteins were constructed, expressed and comparatively characterized, in which differing putative substrate binding residues in SSO3204, identified by comparative structural modeling, were exchanged for the amino acids occupying the respective positions in SSO3003, for which the involvement in substrate binding has been shown by crystal structure analyses.

3.1 ESTABLISHMENT OF AN ARCHAEAL EXPRESSION HOST FOR BIOTECHNOLOGICAL APPLICATIONS

3.1.1 Promoter selectivity – Comparison of bacterial and archaeal promoters for heterologous expression in *Sulfolobus acidocaldarius*

So far the heterologous protein expression in *S. acidocaldarius* is exclusively based on the maltose-inducible promoter of the maltose-binding protein encoding gene *malE* (Saci_1165) (Wagner *et al.*, 2014). To elucidate the applicability of the vector for functional screenings of metagenomic libraries, analyses of the promoter selectivity were performed. For this purpose, the promoters of different archaeal and bacterial PFK encoding genes were analyzed using their natively controlled downstream encoded PFK genes as reporters. Since *S. acidocaldarius* is naturally PFK deficient this allows for easy activity based analyses of expression efficiency. The different archaeal and bacterial phosphofructokinases (PFK) genes, i.e. the crenarchaeal PP_i-dependent PFK from *Thermoproteus tenax* (TTX_1277, *pfp*), the euryarchaeal ADP-dependent PFK from *Pyrococcus furiosus* (PF1784) and the bacterial ATP-dependent PFK from *Thermotoga maritima* (TM0209, *pfkA*) (Siebers *et al.*, 1998; Tuininga *et al.*, 1999; Hansen *et al.*, 2002), were cloned together

with their upstream promoter regions (169 bp upstream of transcription start site of *T. tenax fba-pfp*-operon, 245 bp upstream of transcription start site of *P. furiosus* PF1784 gene, and 199 bp upstream of transcription start site of *T. maritima pfkA* gene, see Table 4 for primer sequences) into the *S. acidocaldarius* expression vector pSVA1450 using the *Sac*II and *Eag*I restriction sites thereby substituting the *malE* promoter implemented in the vector sequence. For comparison, only the coding region of the respective PFK encoding genes were cloned under the control, i.e. downstream, of the vector based *malE* promoter using the *Nco*I and *Eag*I restriction sites (see vector map Figure 3). In *T. tenax* the *pfp* gene (TTX_1277) and the upstream located *fba* gene, encoding the PFK and a fructose-1,6-bisphosphate aldolase (FBPA, TTX_1278), respectively, are organized within one operon. Both genes overlap by 4 bp with TG of the stop codon TGA of *fba* being the last two nucleotides of the start codon ATG of *pfp*. Thus, this whole operon is under the control of the P_{FBPA} promoter upstream of the *fba* gene and the two genes have been shown to be mainly co-transcribed (Siebers *et al.*, 2001). Therefore, the whole *fba-pfp* operon was cloned under the control of both, its own as well as the *malE* promoter respectively. Additionally, for comparison the single *pfp* gene was cloned under the control of the *malE* promoter (the different constructs are depicted in Table 9).

Table 9. List of constructed plasmids for heterologous expression of archaeal and bacterial phosphofructokinases under the control of different promoters in *S. acidocaldarius*.

construct	gene sequence	promoter
P_{MalE} - TTX_1277-78	*fba-pfp* operon (*T. tenax*)	vector based maltose promoter
P_{FBPA} – TTX_1277-78	*fba-pfp* operon (*T. tenax*)	natural operon promoter
P_{MalE} - TTX_1277	*pfp* (*T. tenax*)	vector based maltose promoter
P_{MalE} – PF1784	PF1784 (*P. furiosus*)	vector based maltose promoter
P_{PFK} – PF1784	PF1784 (*P. furiosus*)	natural PFK promoter
P_{MalE} – TM0209	*pfkA* (*T. maritima*)	vector based maltose promoter
P_{PFK} – TM0209	*pfkA* (*T. maritima*)	natural PFK promoter

Briefly, the respective sequences were amplified by PCR and cloned into the pre-vector pMZ1 for introducing a C-terminal Strep-10xHis-tag (primer sets and restriction sites see Table 4, section 2.4.6). After sequencing, the gene fragments (including native promoter sequence and/or C-terminal tag) were transferred into the expression vector pSVA1450. Plasmid DNA harboring the respective genes under the control of the different promoters was transformed into *S. acidocaldarius* MW001 cells (see 2.4.14). For positive *S. acidocaldarius* clones, harboring the recombinant plasmid DNA, was screened (see 2.4.15) and expression was carried out in a Brock medium (Brock *et al.*, 1972) containing 0.1% (w/v) NZ-amine/0.2% (w/v) sucrose/0.4% (w/v) maltose (see 2.3.2). The cells were harvested by centrifugation and disrupted by sonication (see 2.5.4). The recombinant proteins were partially purified by Ni-TED-affinity chromatography (see 2.5.2.4) and tested for PFK activity (see 2.6.1). Furthermore, these samples were also analyzed by SDS-PAGE and Immuno-blotting with anti-Strep-antibody (see 2.5.7 and 2.5.8).

The analysis of the partially purified elution fractions after Ni-TED-affinity chromatography via SDS-PAGE yielded protein bands of the expected size (37 kDa *Ttx*-PFK, 52 kDa *Pfu*-PFK and 34 kDa *Tma*-PFK) (Figure 7A). The Immuno blot using anti-Strep-tag-antibody (Figure 7B) showed distinct signals of expected size for TTX1277 under P_{MalE} control, for PF1748 under control of both its

native and the *malE* promoter as well as for Tm0209 under P_{MalE} control. Consistent with this, PFK activity (1-2.5 U/mg) could be detected in the same samples, i.e. PPi-dependent PFK activity in the P_{MalE}-TTX_1277 sample, ADP-PFK activity in both MalE and native-promoter controlled PF1784 samples and ATP-PFK activity in the P_{MalE}-TM0209 sample (Figure 7C). These results indicate that (i) heterologous expression in the *S. acidocaldarius* system is generally possible, since PFKs of all three organisms were expressed successfully under the control of the vector based maltose-inducible promoter, and (ii) the Crenarchaeon *S. acidocaldarius* is capable of recognizing the foreign promoter structures from the euryarchaeal *P. furiosus* PFK. In this case, the expression efficiency under the pyrococcal PFK promoter in terms of specific ADP-PFK activity was comparable to that under the *malE* promoter. However, the crenarchaeal promoter from *T. tenax* or the bacterial promoter from *T. maritima* appeared not to be recognized by the host.

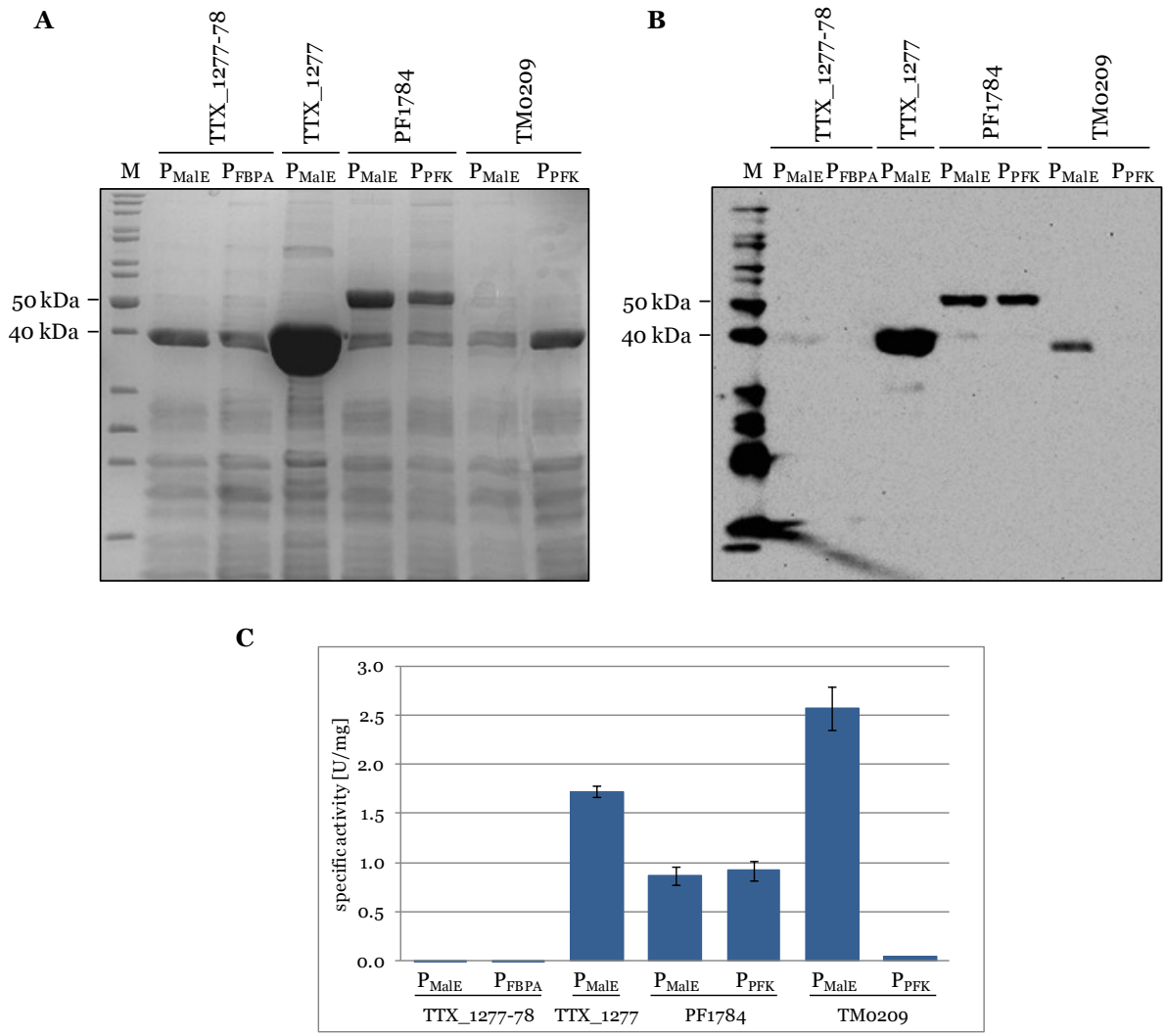


Figure 7. Heterologous expression of archaeal and bacterial phosphofructokinases in *Sulfolobus acidocaldarius* MW001 under the control of different promoters (P_{FBPA}, P_{PFK} or P_{MalE}). Analysis of functional expression via SDS-PAGE (**A,** 12.5% SDS, Coomassie stained), immuno-blotting (**B,** anti-Strep-antibody) and PFK activity assays (**C,** results are the mean of two measurements, standard deviation is given). For comparison same volumes (23 µL) of Ni-TED-affinity elution fractions (E1) were applied to the SDS-PAGE and immuno-blot-analysis and ~ 10 µg protein (E1 fraction) were used for activity assays. Expression was performed at 78°C either under the control of the natural gene promoter (P_{PFK}) or the pSVA1450 based maltose inducible promoter (P_{MalE}). (M) PageRuler™ Unstained Protein Ladder.

3.1.2 Functional Metagenomics – Establishment of *Sulfolobus acidocaldarius* as screening host

To date most functional metagenomic screenings rely on commonly used hosts such as *Escherichia coli* and yeast. However, the applicability appears to be limited, especially for (archaeal) (hyper)thermophilic proteins, e.g. due to different transcription signals, the missing archaeal post-translational machinery, the misfolding of proteins at lower temperatures in mesophilic hosts as well as different codon usages (Kim and Lee, 2006b). To establish the archaeal model organism *S. acidocaldarius* as archaeal, (hyper)thermophilic expression host a suitable strain for the screening of metagenomic libraries for esterase activity, enzymes of high biotechnological importance, was developed and characterized.

3.1.2.1 Examination of esterase activity in *Sulfolobus acidocaldarius*

First, the *S. acidocaldarius* MW001 host strain for genetic manipulations was analyzed for native esterase activity using different ρ-nitrophenyl esters as substrates under standard growth conditions. Strain MW001 was grown in Brock medium (Brock *et al.*, 1972) containing 0.1% (w/v) NZ-amine, 0.2% (w/v) dextrin and 20 mg/L uracil (see 2.3.2), cells were harvested by centrifugation and disrupted by sonication (see 2.5.4). The crude extract after centrifugation was immediately tested for esterase activity at 70°C in 100 mM acetate buffer pH 6.5 containing 1 mM ρNP-substrate. After 5 min incubation p-nitrophenyl formation was photometrically measured at 410 nm (see 2.6.2).

Crude extracts of the reference strain MW001 showed esterase activity at 70°C for all tested ρ-nitrophenyl substrates (17 U/mg for ρNP-butyrate (C4), 18 U/mg for ρNP-octanoate (C8); 15 U/mg for ρNP-decanoate (C10); 8 U/mg for ρNP-dodecanoate (C12)) (Figure 8). These substrates differ in the length of their acyl ester chains. A distinct preference for short-chain acyl esters (shorter than 8-10 carbon atoms) could be observed, thus, the activity seen is most likely due to esterases rather than lipases (Bornscheuer, 2002).

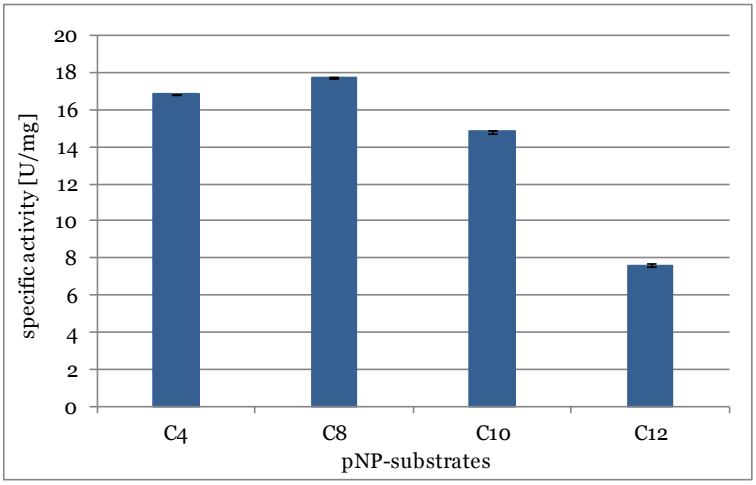


Figure 8. Specific esterase activity in *Sulfolobus acidocaldarius* strain MW001 crude extract. The activity assay was performed at 70°C with 5 µg of MW001 crude extract and 1 mM of ρ-nitrophenyl substrates by measuring the release of ρ-nitrophenol at A_{410nm}. The results are the mean of two measurements and the standard deviation is given.

For the use of *S. acidocaldarius* as screening host of metagenomic libraries for esterase activity, this background activity had to be eliminated. Two esterase homologs Saci_1105 and Saci_1116 were identified in *S. acidocaldarius* by genome sequence analyses (group of Prof. Jäger, ExpresSys project partner, Institute of Molecular Enzyme Technology (IMET), research centre Jülich (FZJ)). Saci_1105 is annotated as an esterase/lipase and shows 28% amino acid identity to the heroin esterase from *Rhodococcus sp.* (PDB code 1lzl_A; (Zhu *et al.*, 2003)), whereas Saci_1116 is annotated as an acetyl esterase and shows 44% amino acid identity to the esterase from *Pyrobaculum calidifontis* (PDB code 3zwq_A; (Palm *et al.*, 2011)). Both gene products are, thus, likely to be responsible for the observed esterase activity in *S. acidocaldarius* MW001 crude extract.

Therefore, *S. acidocaldarius* markerless single and double knockout strains based on the reference strain MW001 were constructed by Alexander Wagner (ExpresSys project partner, MPI Marburg, work group of Dr. Sonja-Verena Albers): MW001Δ1105, MW001Δ1116 and MW001Δ1105Δ1116. All knockouts were markerless and, thus, still PyrEF deficient, allowing for selection required for the transformation with (expression) plasmids for metagenomic esterase screenings. These three knockout strains, which did not show any growth differences on standard medium compared to the MW001 parental strain (data not shown), were tested for esterase activity and compared to the MW001 reference strain (Figure 9). The specific esterase activity is reduced in the single knockout strains MW001Δ1105 and MW001Δ1116 for the substrates ρNP-butyrate (C4), ρNP-octanoate (C8) and ρNP-decanoate (C10). MW001Δ1105 exhibits 40%-50% less activity towards ρNP-butyrate and ρNP-decanoate and 28% less activity towards ρNP-octanoate compared to MW001. The reduction of esterase activity in MW001Δ1116 was less pronounced for ρNP-butyrate and ρNP-octanoate (24% and 6% less activity, respectively), but MW001Δ1116 lost 60% of its activity towards ρNP-decanoate compared to MW001. However, in the double knockout strain MW001Δ1105Δ1116 the esterase activity for the short-chain substrates ρNP-butyrate (C4, 0.83 U/mg, 94% loss) and ρNP-octanoate (C8, 2 U/mg, 88% loss) was nearly completely abolished, whereas for ρNP-decanoate the esterase activity was merely reduced by 43% (C10). For longer chain ρNP-acyl esters (C12) no significant activity alterations were observed.

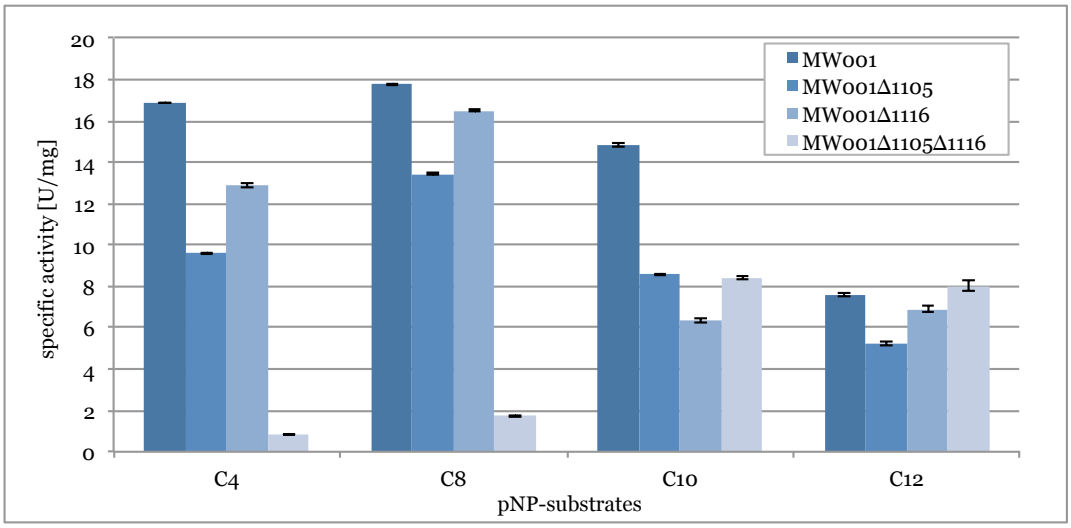


Figure 9. Specific esterase activity in *Sulfolobus acidocaldarius* crude extract from different deletion strains (MW001Δ1105, MW001Δ1116 and MW001Δ1105Δ1116) compared to the reference strain MW001. For the measurements 5 µg of crude extract from the different deletion strains and 1 mM of ρ-nitrophenyl substrates were used and the release of ρ-nitrophenol was followed at A_{410nm} at 70°C. The results are the mean of two measurements and the standard deviation is given.

These results show that the double knockout strain MW001Δ1105Δ1116, due to the abolished background activity, is suitable to screen metagenomic clones for esterase activity towards short-chain acyl ester substrates.

3.1.2.2 Recombinant expression of Saci_1105 and Saci_1116 in *Escherichia coli* and protein characterization

For further investigations the gene sequences Saci_1105 and Saci_1116 were cloned into the pET24a vector and expressed in *E. coli* Rosetta (DE 3) (see 2.5.1). The recombinant esterase Saci_1105 was expressed in the soluble fraction, however, in low amounts not detectable as distinct protein band in the crude extract. The protein was purified by heat precipitation and Ni-TED-affinity chromatography (see 2.5.2.4) to ~70% homogeneity as determined by SDS-PAGE, which showed distinct protein bands of the expected size (37 kDa) (Figure 10A). From 5 g recombinant cells (wet weight) ~ 0.3 mg recombinant Saci_1105 protein was obtained. The esterase Saci_1116 (~ 37 kDa) was sufficiently expressed in *E. coli*, however, within the insoluble fraction (Figure 10B).

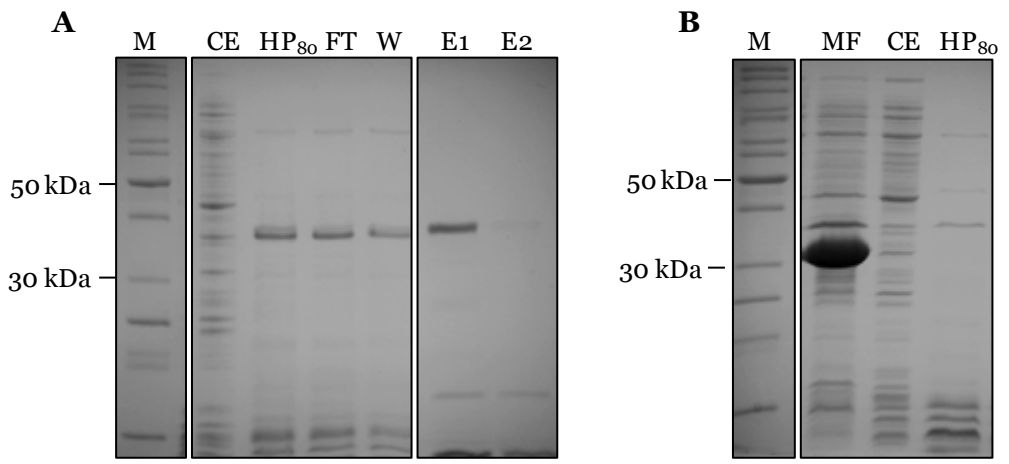


Figure 10. Recombinant expression of the esterases Saci_1105 and Saci_1116 in *Escherichia coli* Rosetta (DE3). Analysis of respective purification fractions of recombinantly expressed **(A)** Saci_1105 and **(B)** Saci_1116 via SDS-PAGE (12.5% SDS, Coomassie staining). **A.** 4 µg crude extract (CE) and4 µg heat precipitation (80°C, 30 min, HP_{80}); 3 µg flow through (FT), 2 µg wash fraction (W), 2.5 µg elution fraction (E1) and 1 µg elution fraction (E2) after Ni-TED-affinity chromatography were applied. **B.** 5 µg membrane fraction (MF), 5 µg crude extract (CE) and 5 µg heat precipitation (80°C, 30 min, HP_{80}) were applied. 5 µL PageRuler™ Unstained Protein Ladder (M).

The esterase Saci_1105 (E1 fraction, Figure 10A) showed a pH optimum at pH 6.5 corresponding well to the internal pH reported for *S. acidocaldarius* (Figure 11A). The substrate dependency of the esterase Saci_1105 followed Michaelis-Menten kinetics (Figure 11B) and a V_{max} of 354 U/mg and a K_m value of 0.6 mM were determined corresponding to a catalytic efficiency of k_{cat}/K_m of 365 $s^{-1} \cdot mM^{-1}$.

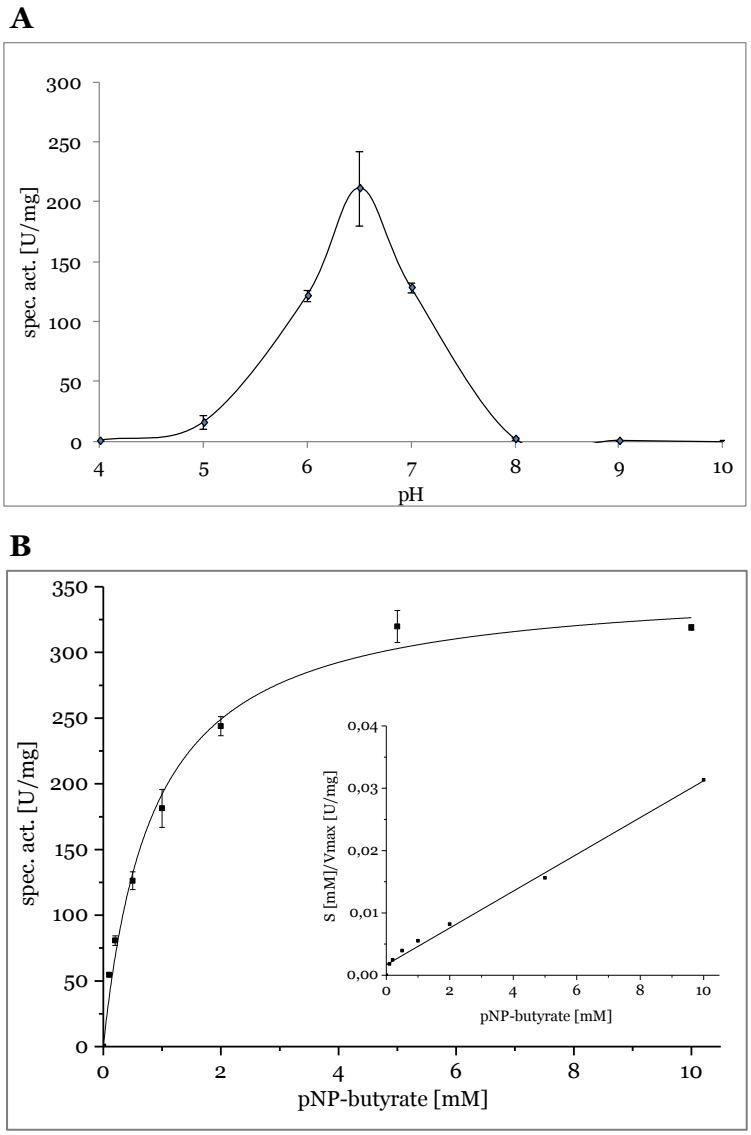


Figure 11. Determination of the pH optimum and kinetic parameters of the *S. acidocaldarius* esterase Saci_1105 recombinantly expressed in *E. coli* Rosetta (DE3). A. Esterase activity was tested at different pH (4-10) at 70°C with 0.2 µg elution fraction E1 after Ni-TED-affinity chromatography and 1 mM ρNP-butyrate (pH 4-6.5 100 mM acetate buffer, pH 7-9 100 mM Tris/HCl buffer, pH 10 carbonate buffer). The results are the mean of three measurements. **B.** Esterase activity was tested with different ρNP-butyrate concentrations (0-10 mM) at 70°C in acetate buffer pH6.5. The results are the mean of three measurements and the standard deviation is given. Insert shows the according Hanes-Woolf plot.

3.1.2.3 Heterologous expression of selected metagenomic clones in *Sulfolobus acidocaldarius* MW001Δ1105Δ1116

Six different metagenomic clones with esterase activity upon expression in *E. coli* were selected for comparative expression studies (LipS, LipT, Est5E5, LipA8b2, EstA3 and LipC5, obtained from Prof. Streit, University of Hamburg, ExpresSys project partner). The clones were derived from different metagenomic libraries: LipS – enrichment of soil and water samples of a botanical garden (65°C), LipT – enrichment of heating water samples (75°C), Est5E5 – biofilm of ethylene-propylene-dien-monomer coated gate valves, LipA8b2 - enrichment of soil and water samples (*Geobacillus* GBHH01), EstA3 – biofilm of gate valve of drinking water pipeline, LipC5 - biofilm of gate valve of drinking water pipeline. The corresponding gene sequences were subcloned into the pSVA1450

derived expression vector pSVA2301 containing an optimized multiple cloning site (N-terminal 10x-histidine- and Strep-tactin-tag) via the restriction sites *Nhe*I and *Not*I (see Table 4, section 2.4.6). *S. acidocaldarius* MW001Δ1105Δ1116 was transformed with the constructed recombinant plasmid constructs (see 2.4.14). As positive control for the expression systems, the β-galactosidase from *S. solfataricus* (*Sso*-LacS), which has already been successfully expressed in *S. acidocaldarius* (Wagner *et al.*, 2014), was also cloned into the vector pSVA2301. For expression *S. acidocaldarius* MW001Δ1105Δ1116 cells transformed with the respective plasmid constructs were grown at 78°C or 65°C in Brock medium (Brock *et al.*, 1972) containing 0.1% (w/v) NZ-amine, 0.2% (w/v) sucrose and 0.4% (w/v) maltose (see 2.5.4). After cell disruption (see 2.5.4) the recombinant proteins were purified from crude extracts via Ni-TED-affinity chromatography (see 2.5.2.4). The respective samples were tested immediately for esterase activity (see 2.6.2) and analyzed by SDS-PAGE and immuno-blotting with anti-Strep-antibody (see 2.5.7 and 2.5.8).

The heterologous expression of the esterases in *S. acidocaldarius* MW001Δ1105Δ1116 at standard conditions (78°C, Brock medium containing 0.1% (w/v) NZ-amine, 0.2% (w/v) sucrose and 0.4% (w/v) maltose) did neither show any distinct protein bands on SDS-PAGE or the immuno-blot corresponding to the estimated sizes of chosen proteins, nor any measurable esterase activity (data not shown).

Since the majority of the six selected esterase candidates have their activity optimum at lower temperatures than the optimal growth temperature of *S. acidocaldarius* (78°C) (Table 11) the expression temperature was lowered to 65°C. At this temperature *S. acidocaldarius* is still able to grow even though with decreased growth rates. However, also under lower temperatures no expression could be detected for the esterases. The SDS-PAGE and immuno-blot-analysis with anti-Strep-antibody of crude extracts and Ni-TED-affinity elution fractions of the respective esterase expression strains (in *S. acidocaldarius* MW001Δ1105Δ1116) in Figure 12 did not show any distinct protein band of the expected size.

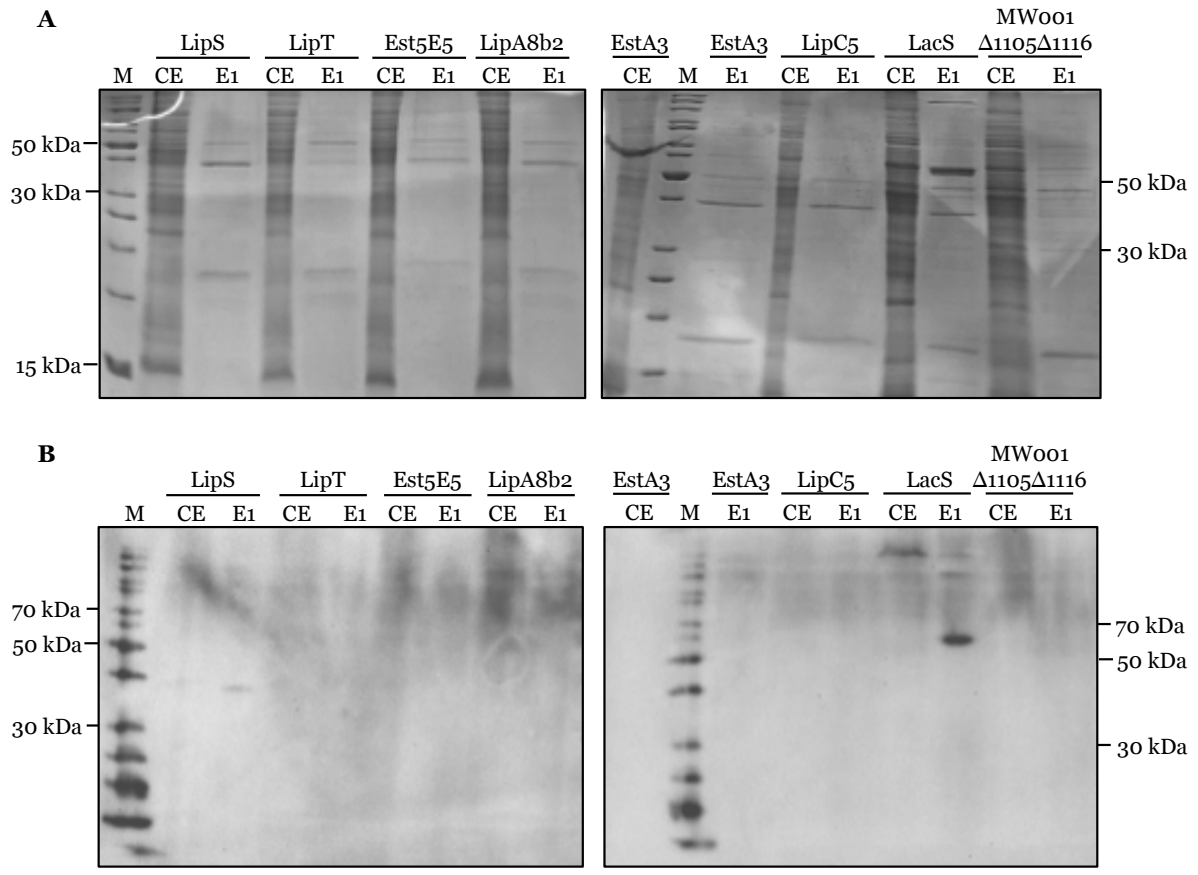
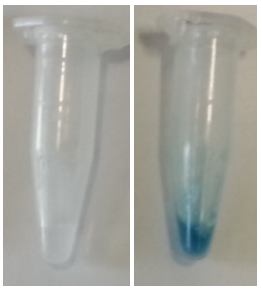


Figure 12. Heterologous expression of metagenomic esterases via the expression vector pSVA2301 in *Sulfolobus acidocaldarius* MW001Δ1105Δ1116 at 65°C. Expression strains were grown to an OD_{600} of 0.8-0.9 at 65°C in Brock medium supplemented with 0.1% (w/v) NZ-amine, 0.2% (w/v) sucrose and 0.4% (w/v) maltose. Analysis of expression patterns via **(A)** SDS-PAGE (12.5% SDS, Coomassie stained) and **(B)** immuno-blotting (anti-Strep-antibody). Samples of 20 µL crude extract (CE) and 4 µL Ni-TED-affinity elution fraction (E1) were applied. pSVA2301::*lacS* was expressed in *S. acidocaldarius* MW001 cells and served as positive control. MW001Δ1105Δ1116 cells without any expression plasmid served as negative control. M: PageRuler™ Unstained Protein Ladder.

However, the positive control strain MW001 pSVA2301::LacS revealed a slight overexpression of LacS indicated by a distinct protein band of approximately 60 kDa (estimated size 62 kDa) in SDS-PAGE (Figure 12) and immuno-blot confirmed by X-gal staining (Figure 13) (under both temperature conditions). pSVA2301 based LacS expression also functioned in combination with the esterase deficient strain MW001Δ1105Δ1116 indicated by comparable results (data not shown).



blank pSVA2301:: *lacS*

Figure 13. Analysis of LacS expression of *Sulfolobus acidocaldarius* MW001 harboring the expression plasmid pSVA2301::*lacS* via β-X-gal staining. Expression strain was grown to an OD_{600} of 0.8-0.9 at 78°C in Brock medium supplemented with 0.1% (w/v) NZ-amine, 0.2% (w/v) sucrose and 0.4% (w/v) maltose. 50 µL crude extract were incubated with 1250 µg β-X-gal (see 2.6.4) for 10 min at 75°C. 1 x LEW buffer (Protino® Ni-TED Kit) served as blank.

In summary, expression with the vector system pSVA2301 was achieved in both *S. acidocaldarius* strains (MW001 and MW001Δ1105Δ1116) for the positive control *Sso*-LacS, although at lower rates compared to pSVA1450 (see above, personal correspondence group of Prof. Dr. Sonja-Verena Albers). The metagenome derived esterase proteins, however, were not sufficiently overexpressed in *S. acidocaldarius* strain MW001Δ1105Δ1116.

3.1.2.4 Transcription analysis of selected metagenomic clones in *Sulfolobus acidocaldarius* MW001Δ1105Δ1116

To investigate whether the missing expression of the esterases with the pSVA2301 vector system in *S. acidocaldarius* MW001Δ1105Δ1116 cells is based on failures at the transcriptional or at the translational level, the mRNA transcripts of the esterase genes were analyzed. The expression strains were grown at 65°C in Brock medium (Brock *et al.*, 1972) containing 0.1% (w/v) NZ-amine, 0.2% (w/v) sucrose and 0.4% (w/v) maltose (see 2.5.4) and the cells were harvested by centrifugation (see 2.5.4). The cells were then used for RNA preparation using TRIzol reagent, a monophasic solution of phenol and guanidinium isothiocyanate, followed by organic extraction and alcohol precipitation of the RNA (see 2.4.16). The prepared RNA was subsequently subjected to DNaseI digestion (see 2.4.16) and the DNA-free RNA was reversely transcribed into cDNA (see 2.4.17) using reverse transcriptase, which was then used in PCR analysis with primer sets specific for the respective esterase genes. Each of the primer sets were constructed so that a fragment of the 5' end of the respective cloned gene was amplified (*lipS* – 133 bp; *lipT* – 175 bp; *est5E5* – 187 bp; *lipA8b2* – 136 bp; *estA3* – 175 bp; *lipC5* – 149 bp) (see Table 6, see 2.4.17). Plasmid DNA (pDNA) with the respective genes was used as positive control, RNA preparation was used to confirm that DNA was successfully removed. As a positive control for the whole RT-PCR procedure *secY* was used as housekeeping gene (Saci_0574, expected fragment size 101 bp) (van der Sluis *et al.*, 2006). As a negative control cDNA preparation from *S. acidocaldarius* MW001Δ1105Δ1116 harboring the empty vector pSVA2301 was prepared (data not shown) (RNA isolation, cDNA synthesis, reverse transcription-PCR and analysis via agarose gel electrophoresis were carried out by Marcel Blum).

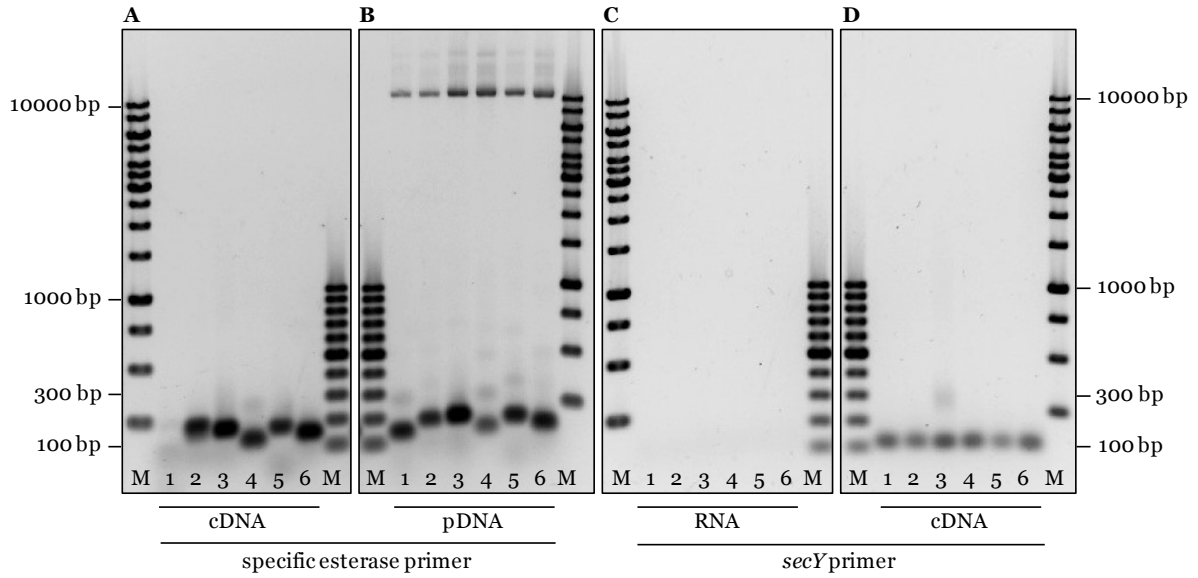


Figure 14. Analysis of RNA transcripts of the respective metagenomic esterase genes heterologously expressed in *S. acidocaldarius* MW001Δ1105Δ1116. RNA transcripts of the respective metagenomic esterase genes were analyzed via agarose gel electrophoresis after heterologous expression in *S. acidocaldarius* MW001Δ1105Δ1116 at 65°C. From each expression strain whole RNA was isolated and complementary DNA (cDNA) was synthesized. PCR amplification was done with cDNA and for the esterase genes specific primer sets **(A)**. Different controls were incorporated **(B-D)**. 1. *lipS* 2. *lipT* 3. *est5E5* 4. *lipA8b2* 5. *estA3* 6. lipC5 M. 5 µL GeneRuler™ 1kb DNA ladder or 5 µL GeneRuler™ 100bp DNA ladder, respectively.

As shown in the agarose gel analyses, transcripts of *lipT*, *est5E5*, *lipA8b2*, *estA3* and *lipC5* could be detected in the respective cDNA preparations (Figure 14A), which showed bands of the expected sizes also observed in the positive control with plasmid DNA as template (Figure 14B). For lipS a transcript of the expected size (133 bp) could not clearly be demonstrated, since a distinct band like in the pDNA control was not observed. However, transcripts of the housekeeping *secY* gene were detected in all six preparations (expected size of the PCR product 101 bp, Figure 14D) and DNA-free RNA preparation was confirmed (Figure 14C). Thus, the genes *lipT*, *est5E5*, *lipA8b2*, *estA3* and *lipC5* are actively transcribed into mRNA in the respective *S. acidocaldarius* strains indicating that missing functional protein expression is caused on the translational or post-translational level.

3.2 Enzyme Modification Of *S. solfataricus* GDH-2 – Influence Of Mutations In Active Site On Substrate Specificity

The first step of the branched Entner-Doudoroff (ED) pathway is the oxidation of glucose to gluconate catalyzed by a glucose dehydrogenase (GDH). The enzyme purified from glucose grown *S. solfataricus* cells, designated as GDH-1 (SSO3003), was shown to be promiscuous for the oxidation of different C5 and C6 sugars (D-galactose, D-xylose, L-arabinose, 6-deoxy-D-glucose and D-fucose) in addition to D-glucose (Lamble *et al.*, 2003). One further paralog for SSO3003 has been recently characterized in more detail (GDH-2 (SSO3204)), revealing that the GDH-2 appears to be specific for D-glucose (Haferkamp *et al.*, 2011). GDH-2 shows a preference for $NADP^+$ over NAD^+ as co-substrate, while GDH-1 shows not a pronounced difference in the use of $NADP^+$ or NAD^+, respectively. The structural determinants for the different sugar substrate specificities of both enzymes have so far not been analyzed. The comparison of the amino acid sequences of both enzymes revealed differences

within their active sites. Most strikingly, based on comparative structural modeling, two amino acid residues, involved in sugar binding, have been suggested to cause this difference in substrate specificity. Instead of Asn^{89} and His^{297} in GDH-1 the enzyme GDH-2 exhibits Val^{93} and Glu^{294}, respectively (Haferkamp *et al.*, 2011). Via a site directed mutagenesis approach GDH-2 mutants have been constructed carrying the aforementioned substitutions (the single mutants GDH-2_V93N and GDH-2_E294H as well as the double mutant GDH-2_V93N_E294H; latter two constructs were cloned by Dr. Eleni Theodosiou). These mutated enzymes were purified and analyzed with respect to substrate specificity in comparison to the wild-type and to GDH-1.

3.2.1 Heterologous expression of wild-type GDH-2 and its mutants in *E. coli* Rosetta (DE3) and purification of the proteins

The wild-type GDH-2 and the respective mutants were recombinantly expressed in *E. coli* Rosetta (DE3) (see 2.5.1) and purified from the soluble fraction by a four step procedure including heat precipitation, ion exchange chromatography (IC), hydrophobic interaction chromatography (HIC) and gel filtration (GF) (see 2.5.2, 2.5.2.1-2.5.2.3).

After size exclusion chromatography GDH-2 was essentially pure as shown by a distinct band at ~ 40 kDa corresponding well to the calculated mass of 39.97 kDa (Figure 15). From 7.7 g recombinant cells (wet weight) ~ 1 mg GDH-2 (SSO3204) were obtained. The purification of the respective mutants yielded similar results (data not shown).

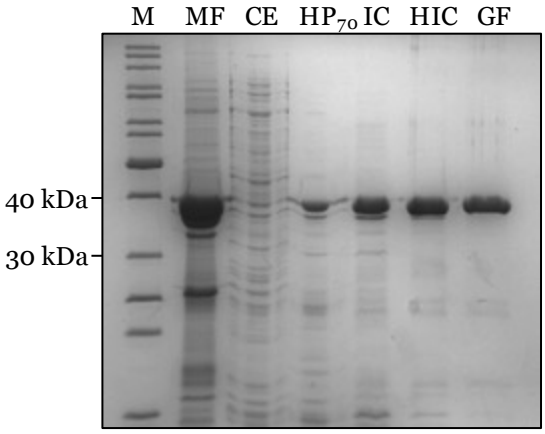


Figure 15. Purification of the wild-type GDH-2 (SSO3204) recombinantly expressed in *Escherichia coli* Rosetta (DE3). Analysis of purification fractions via SDS-PAGE (12.5% SDS, Coomassie stained). 2 µg of each fraction were applied to this analysis. M: 5 µL PageRuler™ Unstained Protein Ladder. MF: membrane fraction. CE: crude extract. HP_{70}: heat precipitation (70°C, 30 min). IC: ion exchange chromatography (Resource Q 6 mL). HIC: hydrophobic interaction chromatography (butyl sepharose 7.5 mL). GF: gel filtration (HiLoad 26/60 Superdex 200 prep grade).

3.2.2 Biochemical characterization of wild-type and mutant GDH-2 with special respect to substrate specificity

With the purified recombinant GDH-2 (wild-type and mutants) an initial substrate screening was performed with 16 different sugar substrates in the presence of either NAD^+ or $NADP^+$ as co-substrate, respectively. For comparison, this substrate screening was performed with the same substrates as reported for GDH-1 (Lamble *et al.*, 2003). The substrates tested included D-glucose,

D-galactose, D-mannose, D-xylose, L-arabinose, D-ribose, D-lyxose, D-arabinose, L-xylose, D-glucosamine, 2-deoxy-D-glucose, 6-deoxy-D-glucose, D-fucose, maltose, D-lactose, sucrose (Figure 16, Figure 17). The activities were measured with 0.3 µg protein in the presence of 10 mM substrate, 10 mM co-substrate (NAD^+/$NADP^+$) and 10 mM $MgCl_2$ in 100 mM HEPES buffer (pH 6.5 (70°C)) (2.6.3).

Under these conditions with NAD^+ as co-substrate wild-type GDH-2 showed a clear preference for D-glucose as substrate. Only minor activities in the range of 10% or less were observed for the substrates D-xylose, D-glucosamine, 2-deoxy-D-glucose and 6-deoxy-D-glucose. All other substrates tested were not converted by GDH-2. (Figure 16).

With $NADP^+$ as co-substrate the preference for glucose was less pronounced. The highest activity again was observed with D-glucose, although reduced (~40%) compared to that obtained with NAD^+ as co-substrate. However, the activities towards D-xylose, 2-deoxy-D-glucose and 6-deoxy-D-glucose were significantly increased (40%-90% of the activity seen with glucose/$NADP^+$). Additional activity could be observed with D-galactose (~15%), whereas that with D-glucosamine appeared slightly decreased compared to NAD^+ as co-substrate.

All GDH-2 mutants (V93N, E294H and the double mutant V93N_E294H) were analyzed with the same substrates under the same conditions. The V93N single mutation and the double mutation V93N_E294H had a serious effect on the enzyme activity. With both co-substrates only very low (if at all) residual activities were observed with all substrates tested. The effect of the E294H mutation was much less significant: especially with $NADP^+$ as co-substrate nearly no decrease in activity with D-glucose, D-xylose, 6-deoxy-D-glucose and D-glucosamine was observed, whereas activities with D-galactose and 2-deoxy-D-glucose were also nearly completely abolished. However - towards the tested substrates - none of the mutations caused a significantly broader substrate specificity compared to the wild-type.

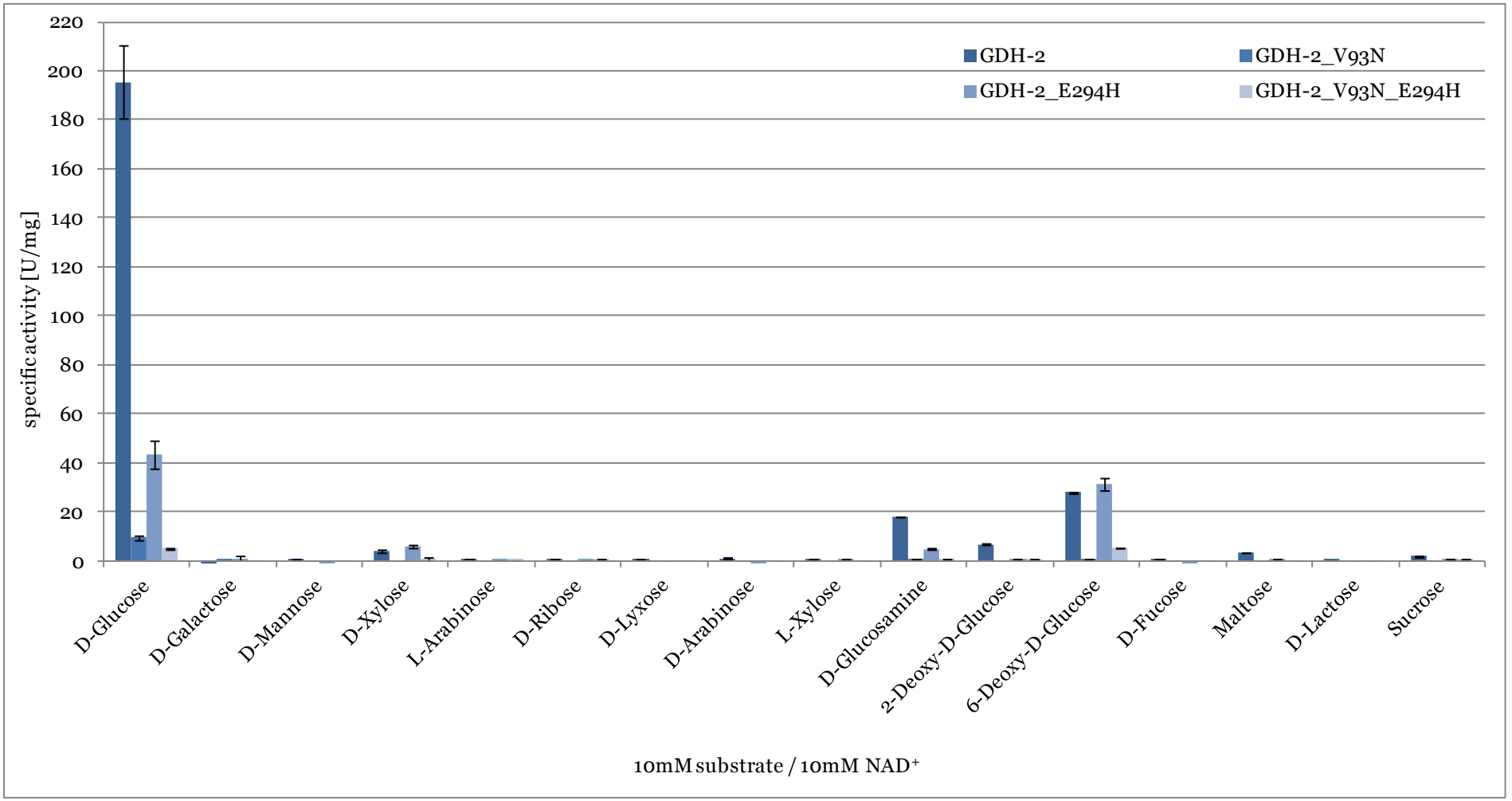


Figure 16. Enzyme activities of wild-type GDH-2 (*SSO*3204), GDH-2_V93N, GDH-2_E294H and GDH-2_V93N_E294H in the presence of various substrates and NAD⁺ as co-substrate. The activity was determined with 0.3 µg protein at 70°C in 100 mM HEPES buffer (pH 6.5 (70°C)) in the presence of 10 mM $MgCl_2$, 10 mM co-substrate and 10 mM substrate by following the formation of NADH at 340 nm. The results are the mean of two measurements and the standard deviation is given.

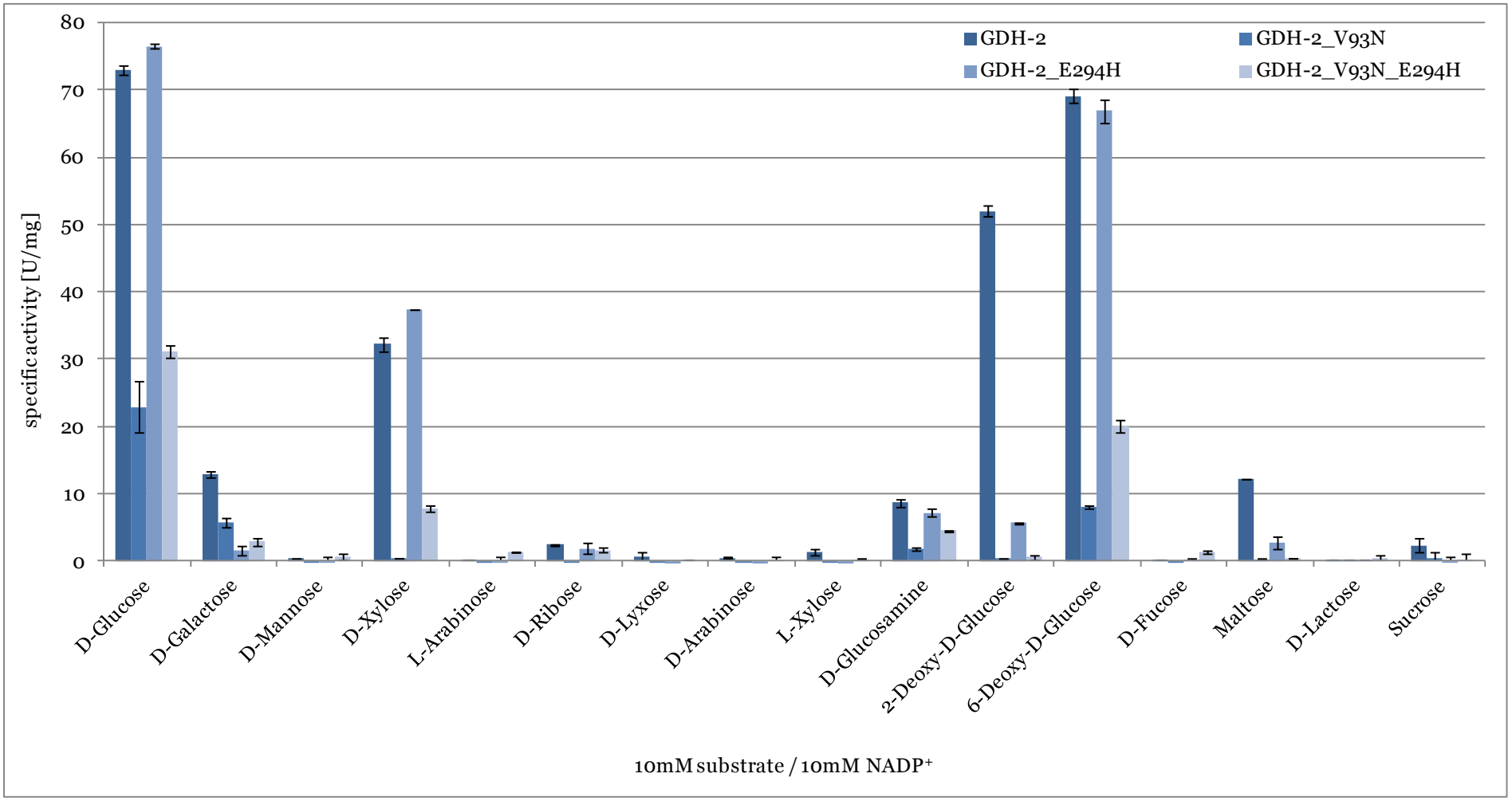


Figure 17. Enzyme activities of wild-type GDH-2 (*SSO*3204), GDH-2_V93N, GDH-2_E294H and GDH-2_V93N_E294H in the presence of various substrates and NADP⁺ as co-substrate. The activity was determined with 0.3 µg protein at 70°C in 100 mM HEPES buffer (pH 6.5 (70°C)) in the presence of 10 mM $MgCl_2$, 10 mM co-substrate and 10 mM substrate by following the formation of NADPH at 340 nm. The results are the mean of two measurements and the standard deviation is given.

Wilde-type and mutant GDH-2 were further characterized in detail with respect to kinetic parameters for all six substrates, for which activity could be detected in the substrate screenings described above. In all cases rate dependence followed classical Michaelis-Menten kinetics. The activities were measured under comparable conditions in the presence of 0.3 µg protein, 10 mM co-substrate (NAD^+/$NADP^+$) and 10 mM $MgCl_2$ in 100 mM HEPES buffer (pH 6.5 (70°C)) and sugar substrate concentrations between 0 and 100 mM. The specific activities [U/mg] were plotted as a function of the substrate concentration [mM] (Figure 18, Figure 19) and the kinetic parameters were determined using the origin software package.

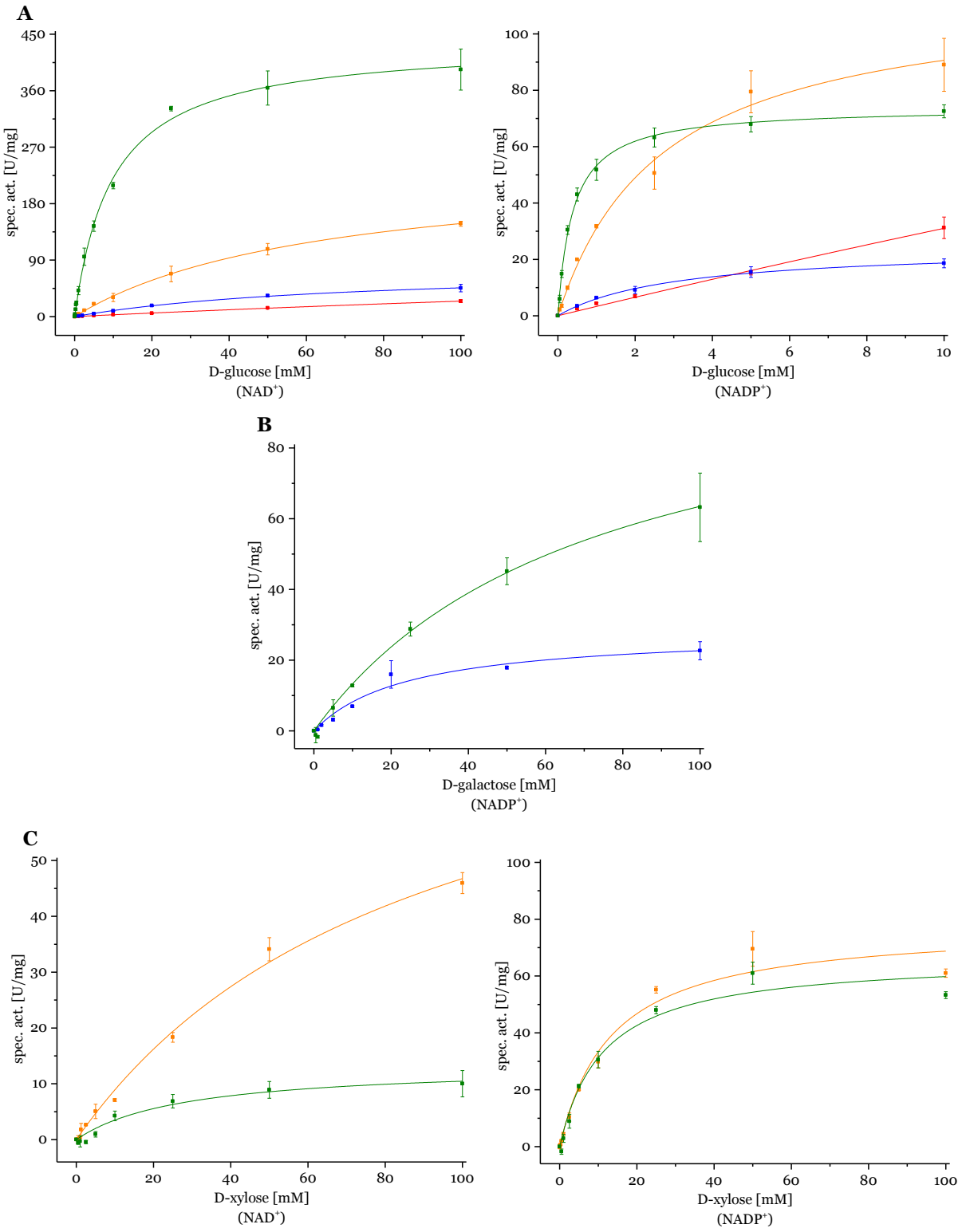


Figure 18. GDH activity of wild-type GDH-2 (■) and the various mutants GDH-2_V93N (■), GDH-2_E294H (■) and GDH-2_V93N_E294H (■) in the presence of (A) D-glucose, (B) D-galactose or (C) D-xylose, respectively, and $NAD^+/NADP^+$ as co-substrate. GDH activity was measured with 0.3 µg protein at 70°C in 100 mM HEPES buffer (pH 6.5 (70°C)) in the presence of 10 mM $MgCl_2$, 10 mM co-substrate and 0-100 mM substrate by following the reduction of $NAD(P)^+$ to NAD(P)H + H^+ at 340 nm. The specific activity [U/mg] was plotted as a function of substrate concentration [mM]. The results are the mean of three measurements and the standard deviation is given.

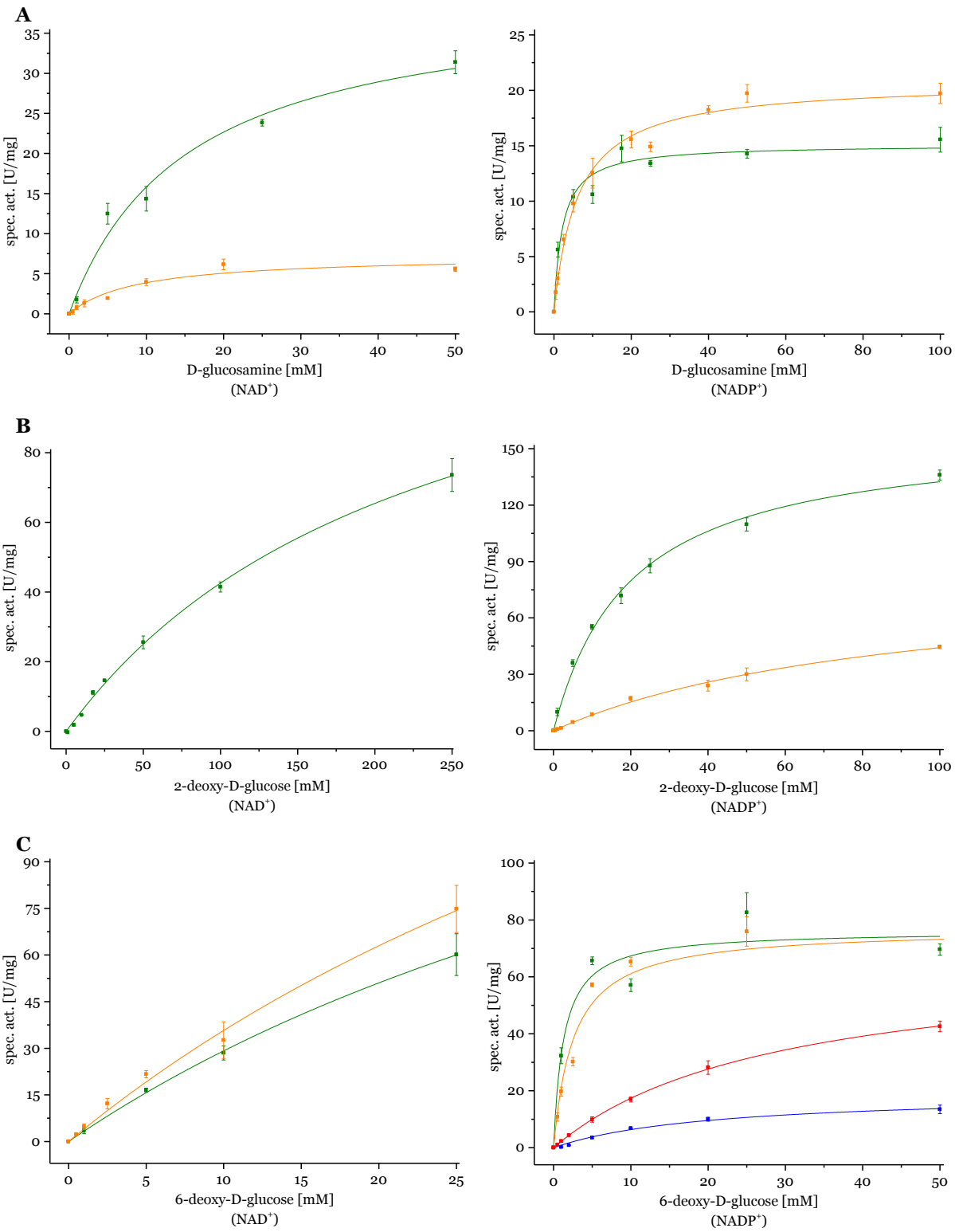


Figure 19. GDH activity of wild-type GDH-2 (■) and the various mutants GDH-2_V93N (■), GDH-2_E294H (■) and GDH-2_V93N_E294H (■) in the presence of (A) D-glucosamine, (B) 2-deoxy-glucose or (C) 6-deoxy-glucose, respectively, and NAD^+/$NADP^+$ as co-substrate. GDH activity was measured with 0.3 µg protein at 70°C in 100 mM HEPES buffer (pH 6.5 (70°C)) in the presence of 10 mM $MgCl_2$, 10 mM co-substrate and 0-100 mM substrate by following the reduction of $NAD(P)^+$ to $NAD(P)H + H^+$ at 340 nm. The specific activity [U/mg] was plotted as a function of substrate concentration [mM]. The results are the mean of three measurements and the standard deviation.

The kinetic parameters for the sugar substrates (V_{max}, K_m, k_{cat} and k_{cat}/K_m) for GDH-2 wild-type and mutant enzymes under fixed co-substrate concentrations of 10 mM are summarized in Table 10. For all sugar substrates GDH-2 wild-type and variants showed considerably lower K_m values and higher catalytic efficiencies with $NADP^+$ as co-substrate, indicating a clear preference for the phosphorylated co-substrate. Also, the pronounced preference of the enzyme for D-glucose as sugar substrate could be confirmed in terms of catalytic efficiencies (wild-type GDH-2: D-glucose/$NADP^+$ 117.7 $s^{-1}\cdot mM^{-1}$, D-galactose/$NADP^+$ 0.9 $s^{-1}\cdot mM^{-1}$, D-xylose/$NADP^+$ 3.7 $s^{-1}\cdot mM^{-1}$, D-glucosamine/$NADP^+$ 4.3 $s^{-1}\cdot mM^{-1}$, 2-deoxy-D-glucose/$NADP^+$ 5.0 $s^{-1}\cdot mM^{-1}$, 6-deoxy-D-glucose/$NADP^+$ 36.0 $s^{-1}\cdot mM^{-1}$).

The detailed kinetic analyses of the mutant enzymes also support the findings that the E294H mutation is less severe (D-glucose/$NADP^+$: 26.3 $s^{-1}\cdot mM^{-1}$) compared to the V93N (D-glucose/$NADP^+$: 4.9 $s^{-1}\cdot mM^{-1}$) and V93N_E294H mutation (D-glucose/$NADP^+$: 2.1 $s^{-1}\cdot mM^{-1}$). This trend was also observed for the other substrates. Only for D-galactose/$NADP^+$, merely converted by the wild-type enzyme (0.9 $s^{-1}\cdot mM^{-1}$), the V93N mutant showed unchanged catalytic efficiency (0.7 $s^{-1}\cdot mM^{-1}$), whereas the E294H mutant did not utilize this substrate at all.

Generally, the E294H mutation induces primarily a decrease in substrate affinity (increased K_m values) for the C6 sugars (D-glucose, D-glucosamine and 2-deoxy-D-glucose) whereas the V_{max} values are not significantly affected (< 2-fold decrease). For D-xylose/$NADP^+$ the kinetic parameters remained nearly unchanged and for 6-deoxy-D-glucose/$NADP^+$ only a twofold increase in K_m value was observed. This indicates a function of this residue in the interaction with the C6-hydroxyl group of the substrates, which is absent in D-xylose and 6-deoxy-D-glucose explaining the minor effects of this mutation on the kinetic parameters towards these substrates. Although for D-xylose/NAD^+ the affinity decreased 3-fold, the catalytic efficiency however remained unchanged (due to a 7-fold increased V_{max}). Thus, confirming the above described trend.

In contrast, the V93N mutation resulted in both, an increase of K_m with a simultaneous decrease of V_{max} for all sugar substrates (despite D-galactose, see above) and accordingly in a significantly decreased catalytic efficiency (Table 10). The results, therefore, indicate that this residue is important not only for substrate binding, but might also be involved in active site architecture and/or dynamics during catalysis.

In accordance with the effects the single mutants caused on substrate specificity, the kinetic parameters for the double mutation V93N_E294H reflect a combination thereof. For this mutant residual activity with dramatically decreased catalytic efficiency could only be observed for the two substrates D-glucose and 6-deoxy-D-glucose. For all other substrates the activity was completely abolished.

Table 10. Kinetic parameters of GDH-2 (SSO3204) and the mutants (GDH-2_V93N, GDH-2_E294H and GDH-2_V93N_E294H) for various substrates with NAD^+/$NADP^+$ as co-substrate. The GDH activities were determined with 0.3 µg protein at 70°C in 100 mM HEPES buffer (pH 6.5 (70°C)) in the presence of 10 mM $MgCl_2$, 10 mM co-substrate and 0-250 mM substrate by following the reduction of $NAD(P)^+$ to NAD(P)H + H^+ at 340 nm.

	GDH-2		**GDH-2_V93N**		**GDH-2_E294H**		**GDH-2_V93N_E294H**	
D-Glucose								
co-substrate	**NAD^+**	**$NADP^+$**	**NAD^+**	**$NADP^+$**	**NAD^+**	**$NADP^+$**	**NAD^+**	**$NADP^+$**
K_m [mM]	9.8	0.4	72.1	3.1	64.2	2.7	534.0	141.9
V_{max} [U/mg]	437.9	73.9	78.9	24.5	243.6	115.2	156.1	471.5
k_{cat} [s^{-1}]	270.3	45.6	48.7	15.1	150.4	71.1	96.4	291.1
k_{cat}/K_m [$s^{-1}\cdot mM^{-1}$]	27.6	117.7	0.7	4.9	2.3	26.3	0.2	2.1
D-Galactose								
co-substrate	**NAD^+**	**$NADP^+$**	**NAD^+**	**$NADP^+$**	**NAD^+**	**$NADP^+$**	**NAD^+**	**$NADP^+$**
K_m [mM]	--	72.0	--	24.4	--	--	--	--
V_{max} [U/mg]	--	109.2	--	28.2	--	--	--	--
k_{cat} [s^{-1}]	--	67.4	--	17.4	--	--	--	--
k_{cat}/K_m [$s^{-1}\cdot mM^{-1}$]	--	0.9	--	0.7	--	--	--	--
D-Xylose								
co-substrate	**NAD^+**	**$NADP^+$**	**NAD^+**	**$NADP^+$**	**NAD^+**	**$NADP^+$**	**NAD^+**	**$NADP^+$**
K_m [mM]	28.4	11.2	--	--	89.1	13.3	--	--
V_{max} [U/mg]	13.4	66.5	--	--	88.4	77.9	--	--
k_{cat} [s^{-1}]	8.3	41.1	--	--	54.6	48.1	--	--
k_{cat}/K_m [$s^{-1}\cdot mM^{-1}$]	0.3	3.7	--	--	0.6	3.6	--	--
D-Glucosamine								
co-substrate	**NAD^+**	**$NADP^+$**	**NAD^+**	**$NADP^+$**	**NAD^+**	**$NADP^+$**	**NAD^+**	**$NADP^+$**
K_m [mM]	15.3	2.2	--	--	8.8	6.2	--	--
V_{max} [U/mg]	40.0	15.1	--	--	7.3	20.8	--	--
k_{cat} [s^{-1}]	24.7	9.3	--	--	4.5	12.8	--	--
k_{cat}/K_m [$s^{-1}\cdot mM^{-1}$]	1.6	4.3	--	--	0.5	2.1	--	--
2-Deoxy-D-Glucose								
co-substrate	**NAD^+**	**$NADP^+$**	**NAD^+**	**$NADP^+$**	**NAD^+**	**$NADP^+$**	**NAD^+**	**$NADP^+$**
K_m [mM]	232.7	19.8	--	--	--	89.9	--	--
V_{max} [U/mg]	141.6	158.6	--	--	--	83.9	--	--
k_{cat} [s^{-1}]	87.4	97.9	--	--	--	51.8	--	--
k_{cat}/K_m [$s^{-1}\cdot mM^{-1}$]	0.4	5.0	--	--	--	0.6	--	--
6-Deoxy-D-Glucose								
co-substrate	**NAD^+**	**$NADP^+$**	**NAD^+**	**$NADP^+$**	**NAD^+**	**$NADP^+$**	**NAD^+**	**$NADP^+$**
K_m [mM]	59.9	1.3	--	20.4	64.1	2.6	--	28.8
V_{max} [U/mg]	203.8	76.2	--	19.3	264.9	77.0	--	67.4
k_{cat} [s^{-1}]	125.8	47.0	--	11.9	163.5	47.5	--	41.6
k_{cat}/K_m [$s^{-1}\cdot mM^{-1}$]	2.1	36.0	--	0.6	2.6	18.3	--	1.4

Chapter 4

DISCUSSION

CHAPTER 4 DISCUSSION

In the course of this work the genetically tractable (hyper)thermophilic Crenarchaeon *S. acidocaldarius* was examined for serving as an archaeal (hyper)thermophilic expression host for functional screenings of metagenomic libraries. For this, the promoter selectivity of *S. acidocaldarius* was analyzed via PFK promoters from different (hyper)thermophilic cren- and euryarchaeal as well as bacterial sources using their natively controlled downstream encoded PFK genes as reporters. Heterologous promoter recognition by *S. acidocaldarius* was successfully shown for the euryarchaeal *Pyrococcus furiosus* PFK promoter sequence. Furthermore, *S. acidocaldarius* effectively expressed all three heterologous PFKs under the control of the *Sulfolobus malE* promoter. Moreover, the expression capabilities of *S. acidocaldarius* under control of the inducible *malE* promoter were investigated using several esterase genes isolated from different metagenomes. Since the *S. acidocaldarius* expression strain MW001 exhibited esterase activity towards the short-chain acyl esters butyrate (C4) and octanoate (C8), an esterase deficient host strain was successfully constructed, by deleting the identified esterase genes Saci_1105 and Saci_1116, for serving as an esterase screening host. Functional expression of the metagenome derived esterase genes from the utilized vector could not be shown under various expression conditions, the transcription of the genes, however, was proven by mRNA analysis. Besides, the esterase genes Saci_1105 and Saci_1116 were recombinantly expressed in *E. coli* and initial biochemical characterization studies showed a pH optimum at pH 6.5 and a catalytic efficiency of 365 $s^{-1} \cdot mM^{-1}$ for the esterase Saci_1105, which followed classical Michaelis-Menten kinetics.

The second major topic of the work was to elucidate the structural determinants conferring the different substrate specificities of the highly homologous sugar dehydrogenases, key enzymes of sugar degradation via the modified branched ED pathway in *S. solfataricus*. Based on structure modeling mutant enzymes of the glucose-specific GDH-2 were constructed using site directed mutagenesis, substituting amino acids found in GDH-2 in corresponding positions to the substrate binding residues in GDH-1: GDH-2_V93N, GDH-2_E294H and the double mutant GDH-2_V93N_E294H. The mutant enzymes were expressed, purified and characterized with respect to substrate specificity and kinetic properties. Comparisons to wild-type GDH-2 and to GDH-1 did not show a shift in substrate specificity of the GDH-2 mutants towards the GDH-1 substrate specificity, but revealed the V93N substitution having a high impact on catalytic efficiency, whereas the E294H substitution does not seem to play a major role.

4.1 A NEW ARCHAEAL EXPRESSION HOST - *SULFOLOBUS ACIDOCALDARIUS*

4.1.1 Heterologous expression and promoter studies

An important requisite for an archaeal expression host suitable for e.g. functional metagenome screenings is the ability to initiate transcription using foreign promoters. The promoter selectivity of *S. acidocaldarius* towards different archaeal and bacterial promoters of phosphofructokinases (PFK)

was investigated via the expression of their natively controlled PFK encoding genes, since the lack of a *S. acidocaldarius* PFK allowed easy activity detection of the respective enzymes: the crenarchaeal PP_i-dependent PFK from *Thermoproteus tenax* (*Ttx*, TTX_1277), the euryarchaeal ADP-dependent PFK from *Pyrococcus furiosus* (*Pfu*, PF1784) and the bacterial ATP-dependent PFK from *Thermotoga maritima* (*Tma*, TM0209). For comparison, the same PFKs were expressed under the control of the maltose-inducible *malE* promoter (Figure 7).

Heterologous expression initiated from the *malE* promoter. Under the control of the established *malE* promoter all three PFKs were successfully expressed in *S. acidocaldarius* as shown via SDS-PAGE, Immuno blot with anti-Strep-antibody and enzyme activity assays (Figure 7).

The *malE* promoter shows low basal activity and is induced upon maltose supplementation (Berkner *et al.*, 2010) and was therefore used as inducible promoter for the construction of the pSVA1450 plasmid for recombinant overexpression of proteins in *S. acidocaldarius* MW001. The *malE* gene (Saci_1165) encodes the sugar binding protein MalE of an ABC transporter and is the first gene of the recently identified maltose ABC transporter operon containing the classical genes for archaeal ABC transporters (Albers *et al.*, 2004) and genes encoding carbohydrate degradation enzymes: the *malE* gene is followed downstream by two permease encoding genes (*malG*, Saci_1164 and *malF*, Saci_1163) and a membrane-bound α-amylase (AmyA, Saci_1162). The operon is flanked downstream by a α-glucosidase (MalA, Saci_1160) and a predicted transcriptional regulator (MalR, Saci_1161) and upstream by an ABC transporter ATPase (MalK, Saci_1166) (Figure 20) (Choi *et al.*, 2013). MalR has been shown to act as a transcriptional activator of the maltose regulon by binding to the *malE* promoter. Amino acid sequence comparison with other archaeal regulators, such as *P. furiosus* TrmB (PF1743) or TrmBL1 (PF0124), revealed only low levels of identity (in a range of 22.7-24.1%) and in a phylogenetic analysis MalR forms a separate group within the TrmB family, which generally act as repressors on their corresponding transport operons. The mechanism of activating regulation remains to be elucidated (Wagner *et al.*, 2014). The expression efficiency of pSVA1450 was, therefore, optimized by the introduction of the *malR* gene into the vector, in addition to the genomic copy. Additionally, also the distance between the TATA box and the start codon has been shown to be of importance (Koerdt *et al.*, 2011). Thus, in accordance to other strong promoters, a 2 bp deletion upstream of the translation start codon of *malE* also contributes to a higher expression efficiency of pSVA1450 (Wagner *et al.*, 2014).

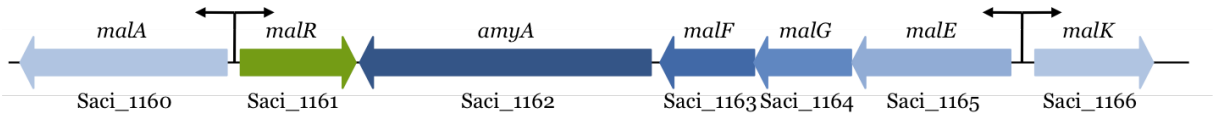


Figure 20. Gene organization within the maltose transport operon of *S. acidocaldarius*. Adopted from Wagner *et al.*, 2014.

In several studies the successful use of pSVA1450 for homologous expression of various *S. acidocaldarius* proteins has been described (e.g. Banerjee *et al.*, 2012; Reimann *et al.*, 2012; Meyer *et al.*, 2013; Orell *et al.*, 2013; Reindl *et al.*, 2013; Liu *et al.*, 2014). To date, two studies reported on the heterologous expression in this *S. acidocaldarius* system of (i) the protein PVAP (protein forming virus-associated pyramids) encoded from the *Sulfolobus islandicus* rod-shaped virus 2 (SIRV2) from

pSVA1450 and (ii) the *S. solfataricus* ABCE1 protein (ATP binding cassette (ABC) protein with two [4Fe-4S]$^{2+}$ clusters, SSO0287) from the pSVA1450 precursor vector pCmalLacS (Berkner *et al.*, 2010; Daum *et al.*, 2014). However, both proteins originated from close relatives of *S. acidocaldarius*. So far, only one study has described a successful heterologous protein expression from a distantly related organism in *Sulfolobales*, the sulfur oxygenase reductase (SOR) from the Crenarchaeote *Acidianus ambivalens* in *S. solfataricus* under the control of the inducible *araS* promoter (SSO3066) (Albers *et al.*, 2006). Thus, the findings presented in this work are one of the first reports on controlled heterologous expression of proteins in *Sulfolobales* originating from very distantly related organisms (Wagner *et al.*, 2014). Very recently, another study reported on the heterologous expression in *S. acidocaldarius* (amylomaltase from *Pyrococcus horikoshii* and α-glucosidase from *Thermoplasma acidophilum*), however, this was based on the constitutive *gdhA* promoter (Saci_0155) (Hwang *et al.*, 2015).

Within the genus of *Sulfolobus* genetic toolboxes have been established for two other species, i.e. *S. solfataricus* and *S. islandicus*. The standard genetic manipulation methods developed for *S. islandicus* are comparable to those for *S. acidocaldarius*. Regulated gene expression, however, is based on the heterologous *araS* promoter (SSO3066) from *S. solfataricus* (Peng *et al.*, 2009). The genetic toolbox of *S. solfataricus,* in contrast, differs with regard to shuttle vector systems (Jonuscheit *et al.*, 2003; Berkner *et al.*, 2007) and selection (Cannio *et al.*, 2001; Jonuscheit *et al.*, 2003; Worthington *et al.*, 2003). Similar to the expression in *S. islandicus* the regulated gene expression relies on the *araS* promoter (SSO3066) (Albers *et al.*, 2006). To broaden the range of applicable promoters, several other promoters have been tested *in vivo* in reporter gene assays for expression efficiency in *S. solfataricus* and *S. acidocaldarius*, including the heat shock inducible *tf55α* promoter (α-subunit of thermosome) and the constitutive *sac7d* promoter (abundant DNA binding protein). Both promoters initiated transcription with high efficiency (Berkner *et al.*, 2010) but also present drawbacks for certain applications: (i) (Over)expression studies rely on regulated induction of protein expression, thus, the *sac7d* promoter is not suitable owing to its constitutive nature; (ii) The *tf55α* promoter is inducible via heat shock, but shows high basal activity (Jonuscheit *et al.*, 2003). Homologous and heterologous expression has been performed under the control of this promoter in *S. solfataricus* for the expression of ATPase FlaI (SSO2316) and SOR from *Acidianus ambivalens* (Albers *et al.*, 2006). The study revealed that the required induction via heat shock also induces a stress response by the host, negatively influencing transcription, translation and replication (Berkner *et al.*, 2010) and the maintenance of the plasmid with this promoter was reduced in *S. acidocaldarius* M31 cells (Berkner *et al.*, 2007).

Hence, up to date the *malE* and *araS* promoters are the best suited promoters available and are routinely applied in inducible *Sulfolobus* expression systems. E.g. (i) in *S. solfataricus* homologous expression of ABCE1 protein (SSO0287) and a membrane-associated ATPase (SSO2680) was successful using the *araS* promoter (SSO3066) (Albers *et al.*, 2006). (ii) In *S. islandicus* the esterase EstA (SiRe_0290) was homologously and the *S. solfataricus* LacS (SSO3019) heterologously expressed using the *araS* promoter (SSO3066) (Peng *et al.*, 2012). And (iii) in *S. acidocaldarius* heterologous *S. solfataricus* ABCE1 protein expression (SSO0287) using the *malE* promoter (Saci_1165) succeeded (Berkner *et al.*, 2010) next to several other proteins (see above).

Next to the Sulfolobales, also members of *Thermococcales*, methanogens and halophiles are genetically tractable (reviewed in Leigh *et al.*, 2011): (i) Genetic tools have been established for the *Thermococcales* members *Thermococcus kodakarensis*, *Pyrococcus abyssi* and *Pyrococcus furiosus*. Protein overexpression is driven by the use of the *gdh* promoter (glutamate dehydrogenase), the nutritionally regulated *fbp* promoter (FBPA/ase) and the constitutive *csg* promoter (cell surface glycoprotein) in *Thermococcus kodakarensis*. The latter promoter was successfully utilized for the heterologous expression of *S. solfataricus* α-1,4-glucan phosphorylase (SSO2538) (Mueller *et al.*, 2009). Also, in *P. furiosus* overexpression is based on either the *gdh* promoter (glutamate dehydrogenase) or the nutritionally regulated *fbp* promoter (FBPase). (ii) Controlled expression in methanogens is based on either tetracycline-responsive promoters (*Methanosarcina spp.*) or the *nif* promoter (nitrogen fixation) (*Methanococcus spp.*), both being strong promoters. (iii) Halophilic expression systems utilize the constitutive *bop* promoter (bacterioopsin) and *fdx* promoter (ferredoxin protein) in *Halobacterium salinarum*, thus, regulated overexpression is not possible, whereas for *Haloferax volcanii* expression is controlled via a tryptophan-inducible p.*tnaA* promoter (tryptophanase), which shows no basal activity and fast induction upon tryptophan presence.

The successful P_{MalE} controlled heterologous expression of a PFK itself also pioneers first steps towards metabolic engineering of *S. acidocaldarius*. *S. acidocaldarius* utilizes the EMP-pathway only for gluconeogenesis in anabolic direction, whereas for glycolysis the branched ED-pathway is employed (Figure 5) (Zaparty and Siebers, 2011). In order to direct the EMP-pathway into a glycolytic direction a PFK homolog for the conversion of F6P to FBP is missing in the *S. acidocaldarius* genome. Besides, the presence of a functional FBPA, the enzyme responsible for the subsequent conversion of FBP to DHAP/GAP, has not yet been confirmed in *S. acidocaldarius*.

S. solfataricus harbors a homolog of a FBPA class Ia (SSO3226), but the enzymatic function still needs to be revealed. *S. acidocaldarius* contains an ORF homolog with 74% amino acid identity, but the encoding gene is not yet as ORF defined and annotated. Besides the putative FBPA, *S. acidocaldarius* also possesses a putative FBPA/ase (Saci_0671), the key enzyme of gluconeogenesis in hyperthermophilic Archaea and deeply branching Bacteria (Say and Fuchs, 2010), but in *S. solfataricus* this bifunctional enzyme (SSO0286) is only active in anabolic direction (Kouril, Esser, *et al.*, 2013).

The missing PFK can now be provided by heterologous expression in *S. acidocaldarius* cells harboring the respective plasmid DNA. Two further requisites are necessary for the introduction of a catabolic EMP in *S. acidocaldarius* in course of metabolic engineering studies: (i) the introduction of a functional FBPA for the conversion of FBP to DHAP/ GAP (e.g. the *Ttx*-FBPA) and (ii) the disruption of the branched ED by deletion of e.g. the GAD encoding gene in *S. acidocaldarius* (Saci_0885).

Promoter comparison. PFK promoters from different (hyper)thermophilic cren- and euryarchaeal and bacterial sources were analyzed with regard to promoter selectivity of *S. acidocaldarius*.

In bacterial promoters two elements are important for RNA polymerase (RNAP) recognition: a -35 (consensus in *E. coli*: TTGACA) and a -10 element (consensus in *E. coli*: TATAAT) (Harley and Reynolds, 1987). The RNAP binding to the DNA is initially unspecific and weak. Upon σ-factor binding

to the −10 and −35 elements the RNAP-DNA interaction is enhanced. Strong promoters may also harbor an AT-rich UP-element, located upstream of the −35 element, which facilitates the binding of the RNAP α subunit (consensus in *E. coli*: NNAAAWWTWTTTTNNAAAANNN, N=any, W=A/T (Estrem *et al.*, 1998; Aiyar *et al.*, 1998)).

Archaeal promoters, in contrast, resemble the eukaryotic RNAPII promoters and exhibit the core elements TATA-box (consensus in Crenarchaeota: YTTTAWA; halophiles: TTTWWW (Soppa, 1999), methanogens: TWTATATA (Li *et al.*, 2008), *Pyrococcus*: TTWWWAW (van de Werken *et al.*, 2006), W=A/T, Y=CT) and BRE site (consensus in Crenarchaeota: RNWAAW, Pyrococcus: VRAAA (van de Werken *et al.*, 2006), halophiles: CGAAA (Brenneis *et al.*, 2007), methanogens: MRCCGAAAAG (Li *et al.*, 2008), V=C/G/A, R=G/A, M=A/C, N=any). The TATA-box is located approximately 24 bp upstream of the transcription start site and the BRE site follows 2-6 bp upstream of the TATA-box and promotes promoter strength.

Analysis of the *S. acidocaldarius malE* promoter revealed, next to the core promoter elements TATA-box (-32 TTTATAAAT -24) and the transcription factor B response element (BRE, -38 ATTAAG -33), an 8 bp repeat upstream of the BRE site (-106 ATAATACT -99 and -139 ATAATACT -132) to be essential for successful transcription (Wagner *et al.*, 2014) (Figure 21). Furthermore, an inverted TATA box motif found upstream of the BRE site (-56 TAAATATTT -48) at a corresponding position as the "ara-box" in the arabinose-inducible promoter of *araS* in *S. solfataricus* (Peng *et al.*, 2009), the so-called *mal* motif, was also shown to be important for successful transcription (Wagner *et al.*, 2014). The *Ttx-pfp* gene is located within an operon downstream of the *fba* gene encoding a fructose-1,6-bisphosphate aldolase (FBPA, TTX_1278). Thus, the promoter upstream of the *fba* gene naturally controls the expression of the operon. For the *T. tenax fba* promoter putative promoter elements have been described: a BRE site (-37 ACAAAA -32) and a TATA box motif (-29 ATATTAAA -22) (Siebers *et al.*, 2001) (Figure 21). The *P. furiosus* PF1784 promoter harbors, next to the core elements BRE site (-36 AAAAA -32) and TATA box (-29 TTTTTAA -23), also a TGM sequence (+7 ATAACACTGAGGGTGGTA +24) downstream of the BRE/TATA sites, which has been shown to be essential for the transcriptional regulation of the glycolytic PFK gene (see below) (van de Werken *et al.*, 2006; Lee *et al.*, 2007, 2008) (Figure 21). To date, the *T. maritima pfkA* promoter has not been described. Therefore, the typical bacterial core promoter elements have been predicted using the software BPROM, based on the known -10 and -35 elements of *E. coli*, which are recognized by the σ70 factor: the -35 element (-35 TTGTTA -30) and the -10 element (-15 TGGTAAATT -7). An UP element, typical for strong promoters, could not be identified manually (Figure 21).

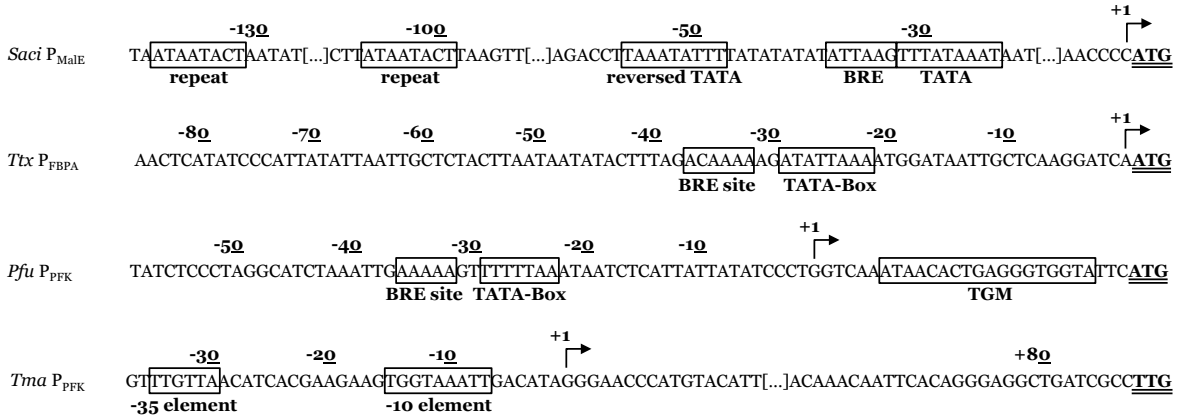


Figure 21. Identification of putative promoter elements of the *T. tenax fba-pfp*-operon promoter (P_{FBPA}), the *P. furiosus* PFK promoter (P_{PFK}) and the *T. maritima* PFK promoter (P_{PFK}) in comparison to the maltose-inducible *malE* promoter of *S. acidocaldarius*. Upstream nucleotide sequences of the different promoters regions are shown and the putative transcription factor B recognition elements (BRE site), the TATA box promoter elements and the Thermococcales-Glycolytic-Motif (TGM) are marked in boxes for the archaeal promoter sequences (Wagner *et al.*, 2014; van de Werken *et al.*, 2006; Siebers *et al.*, 2001). For the bacterial promoter sequence the putative -10 and -35 elements are marked in boxes and have been predicted by the software BPROM, as well as the transcription start site. The putative starting points of transcription are marked by an arrow, and the start codons are bold and underlined.

The general transcriptional machinery in Archaea is of eukaryal nature, whereas the archaeal transcriptional regulators are more bacterial-like. This could explain the lack of expression in *S. acidocaldarius* using the bacterial *T. maritima* promoter. Comparison of the archaeal promoters from *S. acidocaldarius, T. tenax* and *P. furiosus* showed, that only the presence of the core elements BRE site/TATA box appears to be common to both, crenarchaeal and euryarchaeal. The TGM motifs seem to be restricted to the genus *Thermococcales*, since these have not been described for other archaeal members, and thus could not be identified neither within the *S. acidocaldarius malE* promoter nor the *T. tenax fba* promoter. And vice versa, no repeats, as seen in *S. acidocaldarius malE* promoter, have been described for the *P. furiosus* PF1784 promoter or *T. tenax fba* promoter. Hence, from the archaeal promoter architectures no direct conclusions can be drawn with regard to the presence or absence of protein expression initiation.

The *in vivo* transcription regulation of the PFKs in their native hosts is not fully described for all three PFKs. In case of the *Ttx*-PFK transcript analyses of the *fba-pfp*-operon revealed higher transcript amounts under heterotrophic growth conditions (glucose) compared to autotrophic growth conditions (CO_2 and H_2), indicating glucose being the inductor for transcription (Siebers *et al.*, 2001). For the *Pfu*-PFK it has been shown that its transcription is up-regulated when cells were grown on maltose, indicating the presence of a maltose-inducible promoter (Schut *et al.*, 2003). Indeed, due to the presence of the Thermococcales-Glycolytic-Motif (TGM) downstream of the BRE/TATA sites in the promoter region, this glycolytic gene is transcriptionally regulated by the regulator TrmBL1 via the effector maltose/maltotriose (activation at low maltose concentrations, repression at no or high maltose concentrations) (Lee *et al.*, 2007, 2008). No information about the transcription induction of the bacterial *Tma*-PFK is available so far.

Expression from heterologous promoters. The ability of *S. acidocaldarius* to initiate expression from these heterologous promoters was analyzed using the promoter's natively controlled downstream encoded PFK genes as reporters.

From the natural PFK promoter sequences only the *P. furiosus* promoter was recognized by *S. acidocaldarius*. The expression of the *Pfu*-PFK under the control of its natural promoter was very similar to the PFK-expression under the control of the *malE* promoter.

The similar native transcriptional regulation of the *P. furiosus* PF1784 and *S. acidocaldarius malE* promoter, both activated in the presence of maltose, may therefore explain the successful PFK expression from the native euryarchaeal promoter under *S. acidocaldarius* expression conditions. In contrast, the native inductor of the *T. tenax fba-pfp*-operon is glucose (Siebers *et al.*, 2001) and in the presence of maltose no growth of *T. tenax* was observed (Zillig et al., 1981). However, since this operon was also not expressed under the *S. acidocaldarius malE* promoter, the different use of inductors does not seem to be essential for the successful FBPA and PFK expression from its native operon promoter.

The use of heterologous promoters for (over)expression in Archaea has been described before, although in most cases the promoter donor organism was closely related to the expression host (Thomm *et al.*, 1992; Gregor and Pfeifer, 2005; Peng *et al.*, 2009). A few studies, however, did describe successful protein expression controlled by heterologous promoters derived from distantly related organisms. In *Haloferax volcanii* expression of a yeast $tRNA^{ProM}$ gene was driven by *P. furiosus gdh* (glutamate dehydrogenase, constitutive) and *mlrA* (maltose regulated) promoters, respectively (Schreier *et al.*, 1999). Or in *T. kodakarensis* expression was shown under the control of the constitutive promoter of the *M. thermautotrophicus* histone encoding *hmtB* gene (Santangelo *et al.*, 2010).

Pyrococcus, as a representative of the euryarchaeal phylum is only distantly related to the Crenarchaeon *Sulfolobus*. Conversely, although *T. tenax* originates from the same phylum as *S. acidocaldarius* and appears thus more likely to resemble the transcriptional apparatus of *S. acidocaldarius* including its promoter architecture, the thermoproteal operon was not expressed in *S. acidocaldarius*. Within the *T. tenax fba-pfp* operon the genes overlap by 4 bp with TG of the *fba*-stop codon TGA being part of the *pfp*-start codon ATG. Generally, for most archaeal genomes 50-70% of the genes are organized within operon structures (Torarinsson *et al.*, 2005), therefore, the transcriptional and translational apparatus of *S. acidocaldarius* should be able to express operon structures. Previous investigation of this operon revealed a co-transcription of both genes (Siebers *et al.*, 2001). Polycistronic transcripts with overlapping genes, however, have also been described for *S. acidocaldarius* (e.g. *gar*1 (Saci_1340) and *tfb*2 (Saci_1341) (Rauch, 2013)) or *S. solfataricus* (e.g. subunits *porA* and *porD* within the *por* operon, SSO2128-2130 (Peeters *et al.*, 2009)), both showing the exact same overlapping sequence as seen in the *Ttx-fba-pfp* operon. But neither PP_i-PFK nor any FBPA activity was detected in the crude extracts gained from the expression strains harboring the operon constructs, independent of the promoter (P_{FBPA} or P_{MalE}) (data not shown).

Besides, heterologous operon expression has been successfully shown in *P. furiosus* (e.g. Keller *et al.*, 2014). In this study an artificial pathway for butanol production was introduced to *P. furiosus*, by inserting synthetic operons into the host's genome simultaneously deleting the α-subunit of acetyl-CoA synthetase I (converts acetyl-CoA to acetate), thus rerouting the acetyl-CoA into the butanol production and away from acetate formation. The artificial operons consisted of six genes derived from three different bacterial donors with optimal growth temperatures ranging 25-40°C below the

P. furiosus optimum of 100°C. The protein expression from these operons was assured by utilizing the strong and constitutive *P. furiosus por* promoter (γ-subunit of pyruvate oxidoreductase, POR), inserting *P. furiosus* ribosomal binding sites from highly expressed genes between each gene of the operon and including a *P. furiosus* terminator sequence at the end of the operon (Keller *et al.*, 2014). This study, thus, showed host specific translational elements, such as ribosomal binding sites, to have an enhancing effect on heterologous expression.

Limitations may, next to the transcriptional level, also occur on the translational level, which includes e.g. the use of very specific codon usages. This can differ especially in distantly related organisms such as *S. acidocaldarius* and *T. maritima*. The successful expression of all three heterologous PFKs under the control of the *malE* promoter indicates possible divergent codon usages not being a general matter. Comparison of the codons in the *pfkA* gene of *T. maritima* with the codon usage of *S. acidocaldarius* did also not reveal a high overall discrepancy, however, the natural start codon of the *pfkA* is TTG, which is a very rarely used start codon in *S. acidocaldarius* (Vendittis and Bocchini, 1996). During cloning the *pfkA* gene downstream of the *malE* promoter the start codon was changed to ATG. Therefore, the absence of the expression of the bacterial PFK within the archaeal host under the control of its natural promoter sequence may be attributed to the different use of start codons.

Nevertheless, the successful expression from the heterologous *P. furiosus* promoter in *S. acidocaldarius* offers new prospects. On the one hand, heterologous promoter recognition is an important requisite for metagenome analyses of large-insert libraries, harboring large DNA fragments containing foreign promoters. *S. acidocaldarius* did show to be able to initiate expression from a heterologous promoter to a certain extent and, therefore, may be suitable in serving as an expression host in such screenings. On the other hand, it might be beneficial to drive expression from a variety of promoters in metagenomic screenings of small-insert libraries, but also simply for heterologous overexpression. The available expression system may now be further expanded by the application of differently inducible promoters, e.g. cold-inducible promoters as shown for *P. furiosus* (Basen *et al.*, 2012). The induction with a metabolic compound, such as sugars, is not always favorable, due to negative effects on heterologous overexpression by decreasing expression levels over time, which in turn might limit the production of sufficient biomass. Basen *et al.* have demonstrated the use of a cold-inducible promoter (P_{CipA}) in *P. furiosus* for the heterologous expression of a bacterial lactate dehydrogenase (LDH (Cbes_1918), *Caldicellulosiruptor bescii*), being active at almost 30°C below the growth temperature of *P. furiosus* (100°C) (Basen *et al.*, 2012). The successful expression of this bacterial enzyme in an archaeal host as well as the possibility of a metabolic switch towards a different end product in a temperature-mediated manner was reported (hydrogen/acetate towards lactate). BLAST searches revealed a hypothetical protein (SSO1273) in *S. solfataricus* P2 showing 28% amino acid identity with CipA, whereas no homolog was found in *S. acidocaldarius* DSM639. Also, a homolog for CipB (PF1408), a further cold-induced protein, was found in *S. solfataricus* P2 with 24% amino acid identity, annotated as a dipeptide ABC transporter (periplasmic dipeptide binding protein dppA, SSO2619). These results reflect the findings of Weinberg *et al.* (Weinberg *et al.*, 2005). Reporter gene assays with the *S. acidocaldarius* expression vector pSVA1450 would reveal whether the promoter

sequences of these *S. solfataricus* P2 CipA/B homologs exhibit similar cold-inducible characteristics as the *Pfu-cipA* promoter, and, thus, could serve as an alternative expression system.

4.1.2 Screening of metagenomic libraries

The optimization of industrial enzymatic processes often relies on the discovery of novel and unique enzymes via methods like metagenomic analyses. Functional metagenomics are based on the heterologous expression of clone libraries in suitable host organisms with *Escherichia coli* representing thereby the most frequently used host. Unfortunately, not all proteins, especially so-called extremozymes, can easily be functionally expressed in this standard host. This is particularly true for proteins of Archaea, which often inhabit extreme environments and harbor unique metabolic pathways and enzymes. The development of an archaeal expression host is, therefore, highly desirable to enable the identification of new extremozymes and the exploitation of their great potential for industrial applications, which would otherwise remain undetected.

Therefore, in addition to the examination of promoter selectivity, the expression capabilities of *S. acidocaldarius* were exemplarily investigated using several esterase genes, functionally identified from different metagenomic libraries, for a further establishment of *S. acidocaldarius* as an expression host for functional metagenomics. The respective gene sequences were cloned into the *S. acidocaldarius* expression vector pSVA2301 (pSVA1450 derivative) under control of the inducible *malE* promoter.

Lipolytic enzymes show a high versatility, since many are active in organic solvents, do not need any co-factors and often display extraordinary stereo-, chemo- and regioselectivities (Jäger and Reetz, 1998), making them one of the most important biocatalysts for biotechnological applications to date. Lipolytic enzymes are applied in several production processes, e.g. of pharmaceuticals and detergents, in the food industry, e.g. in flavor development and improvement of cheese or in biodiesel production and in synthesis of polymers and fine chemicals (Hasan *et al.*, 2006; Salameh and Wiegel, 2007; Levisson *et al.*, 2009). In industrial processes most lipolytic enzymes applied so far originate from mesophilic organisms. Nevertheless, (hyper)thermophilic candidates are much more favorable, since most biotechnological applications are performed under harsh conditions requiring high enzyme stability at high temperature and high organic solvent concentrations.

Carboxylesterases (EC 3.1.1.1) and triacylglycerol lipases (EC 3.1.1.3) represent subclasses of hydrolase enzymes acting on ester bonds (EC 3.1) and can be found in all three domains of life. Most lipolytic enzymes studied so far were conducted using bacterial sources, including the (hyper)thermophilic esterases from *Thermotoga maritima* and *Thermus thermophilus*. Several archaeal lipolytic enzymes have also been characterized as well, amongst them the (hyper)thermophilic carboxylic-ester hydrolases from *Aeropyrum penix*, *Archaeoglobus fulgidus*, *Picrophilus torridus*, *Pyrobaculum calidifontis*, *Sulfolobus spp.*, and *Pyrococcus spp.*. These (hyper)thermophilic carboxylic-ester hydrolases mainly belong to the class of esterases (Atomi and Imanaka, 2004; Levisson *et al.*, 2009). Furthermore, many esterases and lipases have also been identified from metagenomic screenings (reviewed in Steele *et al.*, 2009), but only few have been characterized in more detail (e.g. Elend *et al.*, 2006; Chow *et al.*, 2012). To date, no true lipases have been reported for Archaea (Salameh and Wiegel, 2007; Levisson *et al.*, 2009).

In this study the six selected metagenomic clones for the comparative screening in the multiple hosts, exhibited lipolytic activity when expressed in *E. coli* and showed a diverse range of characteristics with regard to origin, temperature optima, and substrate specificity (Table 11). The clones were derived from different metagenomic libraries: LipS - enrichment of soil and water samples botanical garden (65°C), LipT - enrichment of heating water samples (75°C), Est5E5 - biofilm of ethylene-propylene-dien-monomer coated gate valves, LipA8b2 - enrichment of soil and water samples (*Geobacillus* GBHH01), EstA3 - biofilm of gate valve of drinking water pipeline, LipC5 - biofilm of gate valve of drinking water pipeline. Three of the lipolytic enzymes have been recently characterized: EstA3, LipS and LipT (Elend *et al.*, 2006; Chow *et al.*, 2012). EstA3 shows esterase activity only in a multimeric form (crystallization studies propose an octameric form) and prefers short-chain acyl esters (ρNP-C4 -C8) (Elend *et al.*, 2006). LipS and LipT both are highly thermo-stable enzymes and exhibit lipase activities with a broad substrate spectrum ranging from medium- to long-chain acyl esters (ρNP-C4 -C14), with LipS preferring ρNP-C8 and LipT ρNP-C10 (Chow *et al.*, 2012). LipS is one of the first metagenomic lipases to be analyzed by crystallography (Chow *et al.*, 2012). Est5E5 and LipC5 are classified as esterases, due to their substrate preference for short- to medium-chain acyl esters. Est5E5 acts preferably on ρNP-C8 and shows only low thermostability at 50°C (Chow, 2012). LipA8b2 is also classified as an esterase, but does accept also acyl esters with a chain length of C18.

Table 11. Characteristics of the metagenomic lipolytic enzymes. Kindly provided by Dr. J. Chow, group Prof. Streit. Protein names, molecular weights [kDa], metagenome origin, substrate utilization, optimal activity temperature [°C], sequence available and classified function are given. TBT: tributyrin, pNP: ρ-nitrophenyl substrates with different chain length, EPDM: ethylene-propylene-dien-monomer.

Name	Mol. weight (kDa)	Metagenome origin	Sub-strate	Temp. Opt. (°C)	Seq.	Func-tion
LipS	32	soil and water samples botanical garden (65°C)	TBT, pNP C4-14	70	yes	esterase/ lipase
LipT	38	heating water samples (75°C)	TBT, pNP C4-15	75	yes	esterase/ lipase
Est5E5	53	biofilm of EPDM coated gate valves	TBT, pNP C4-10	55	yes	Esterase
LipA8b2	48	soil and water samples (*Geobacillus* GBHH01)	TBT, olive oil, pNP C6-18	60	yes	Lipase
EstA3	50	biofilm of gate valve of drinking water pipeline	TBT, pNP C4-10	50	yes	Esterase
LipC5	42	biofilm of gate valve of drinking water pipeline	TBT, pNP C4-10	n.a.	yes	esterase

Successful expression of heterologous proteins in general might be hampered on the transcriptional, translational or post-translational level: non-detection of heterologous transcription signals, missing adequate ribosomal binding sites, little mRNA stability, different codon usage, incorrect folding due to the absence of proper chaperons and protein modifying enzymes, incorrect or no secretion of proteins due to low or no detection of foreign secretion signals, or missing co-factors may reduce the detection of activity (Mergulhão *et al.*, 2005; Sørensen and Mortensen, 2005; Villegas and Kropinski, 2008; Warren *et al.*, 2008). Furthermore, heterologous proteins might be degraded by the hosts increased proteolytic activities, which can be a type of stress response caused by the heterologous expression (Glick, 1995). All these factors are assumed to influence heterologous expression. Supposedly, the closer the phylogenetic relationship between DNA donor organism and expression host, the more likely is a successful functional expression (Liebl *et al.*, 2014).

In order to overcome these limitations, the co-operative project ExpresSys aimed for a comparative screening in multiple expression hosts, divergent in terms of phylogeny and physical requirements: the thermoacidophilic Archaeon *S. acidocaldarius* and the Bacteria *Escherichia coli*, *Pseudomonas putida*, *Rhodobacter capsulatus*, *Pseudomonas antarctica* and *Thermus thermophilus*. Several studies already demonstrated the discovery of novel enzymes in functional screenings in multiple hosts (e.g. Courtois *et al.*, 2003; Craig *et al.*, 2010; Kakirde *et al.*, 2011). However, in most cases the hosts were relatively closely related to each other, thus, they comprise quite similar transcription and translation machineries. Therefore, it is assumed that a more diverse range of hosts will even more increase the expression and activity levels and, thus, yield more novel enzyme discoveries.

In a screening with multiple, but closely related hosts often the same shuttle-vectors are used for the transfer of metagenomic libraries between the organisms. However, the use of the same vector in distantly related organisms often necessitates the implementation of additional origins of replication, promoter structures, diverse selection markers and multiple cloning sites, and possibly different tags, which could increase the size of the vector significantly and decrease transformation efficiency. Thus, the development of organism specific expression vectors is required, multiplying the cloning effort. However, it is expected that this time-consuming fact can be outweighed by the increase in novel enzyme identifications.

For the functional screening of the selected metagenomic esterases, esterase activity by *S. acidocaldarius* itself under standard assay conditions had to be ruled out. In course of this work a background esterase activity of 8-18 U/mg, depending on the ρ-nitrophenyl substrate (ρNP-C4 -C12) used, was detected in *S. acidocaldarius* MW001 crude extracts (Figure 8). Most studies of archaeal esterases report on the characterization of the recombinant proteins, thus data about esterase activity in archaeal crude extracts are rare. A previous study on *S. acidocaldarius* DSM639 esterase activity yielded 0.15-0.18 U/mg in crude extracts towards the ρNP-C2 substrate (Sobek and Görisch, 1988, 1989). In *S. solfataricus* P1 crude extracts an esterase activity of ~7 U/mg towards ρNP C16 or phenyl acetate was reported (Park *et al.*, 2006, 2008). Since this background activity in *S. acidocaldarius* MW001 might interfere with possibly low esterase activities in functional metagenomic screenings, a deletion strain was constructed, depleted in any esterase activity towards the substrates used. Genome

sequence analyses (performed by the group of Prof. Jäger, IMET, Jülich) revealed the presence of two esterase homologs in *S. acidocaldarius*: Saci_1105 (annotated as esterase/lipase) and Saci_1116 (annotated as acetyl esterase). Both genes encode putative esterases, which show 28% and 44% amino acid identities to the structurally characterized heroin esterase from *Rhodococcus sp*. (PDB code 1lzl_A, (Zhu *et al.*, 2003)) and the structural resolved esterase of *Pyrobaculum calidifontis* (PDB code 3zwq_A, (Palm *et al.*, 2011)), respectively, and were assumed to cause the detected esterase activity in the *S. acidocaldarius* MW001 strain. Deletion of the esterase genes Saci_1105 and Saci_1116 resulted in a decreased activity towards all short-chain acyl esters tested, especially ρNP-C4 and ρNP-C10, in the single deletion strains (MW001Δ1105 and MW001Δ1116), and nearly no activity towards ρNP-C4 and ρNP-C8 in the double deletion strain (MW001Δ1105Δ1116). Nearly no difference in activity towards the substrate ρNP-C12 was observed for all deletion strains, confirming that the two genes Saci_1105 and Saci_1116 do encode for esterases instead for lipases (Figure 9).

In *E. coli* as well as *R. capsulatus* and *P. putida*, all representing selected hosts in the ExpresSYS project, no host specific esterase background activity was detected using tributyrin (TBT, C4) and ρNP-C4 -C12 as substrates (personal correspondence with Dr. Anuschka Malach (Malach, 2013)). However, *T. thermophilus* HB27 showed activity with TBT, ρNP-C4 -C16 and 5-bromo-4-chloro-3-indoxyl-C4, -C8, -C16 and 10 esterase/lipase homologs were identified. Deletion of three of the afore mentioned homologs resulted in a 75% decrease of this background activity, thus, the respective triple-mutant *T. thermophilus* strain is now a suitable host for metagenomic esterase screenings (Leis *et al.*, 2014). With the double esterase deletion strain *S. acidocaldarius* MW001Δ1105Δ1116 also an archaeal screening host is now available.

Heterologous expression is generally possible in *S. acidocaldarius* with the pSVA1450 vector, as shown in this work for PFK genes from different sources and elsewhere (see above). However, the use of pSVA1450 for metagenomic screenings is rather limited due to the restricted multiple cloning site and the necessity of a pre-cloning step via pMZ1 for the fusion of a C-terminal Strep-10xHis-tag to the gene of interest. Therefore, a refined system, more suitable for metagenomic screenings, was required and pSVA1450 was further improved yielding the vector pSVA2301. The functional expression of the metagenomic esterases under the control of the inducible *malE* promoter with pSVA2301, however, could not be achieved, neither at the optimal *S. acidocaldarius* growth temperature of 78°C nor at a reduced temperature of 65°C (Figure 12). As mentioned above, a functional expression may be hampered on the transcriptional level. For this, the transcription of the esterase genes was analyzed and mRNA transcripts, and thus transcription, for five of six metagenomic esterases, i.e. *lipT*, *est5E5*, *lipA8b2*, *estA3* and *lipC5*, could be demonstrated (Figure 14). However, the transcription efficiency might be significantly reduced causing decreased expression levels as indicated by the LacS control, which showed notably lower galactosidase activity compared to the LacS expression with pSVA1450. This suggests that different vector properties of pSVA2301 are responsible for suboptimal expression.

The presence of mRNA suggests a failure on the translational level of the *S. acidocaldarius* expression machinery. Limitations on the translational level may include several aspects, as described above, whereas the divergent codon usage probably is most hindering in heterologous protein expression (Ekkers *et al.*, 2012). Codon usage implies one or more codons encoding for the same

amino acid and differences in codon preferences in different organisms. The use of different codons often comes along with a different GC content in the genome (Knight *et al.*, 2001). The more similar the GC content between donor and recipient organism is the more likely is the successful heterologous expression. In metagenomic studies the donor organism is often not known, therefore, the parallel use of different organisms with diverse GC contents may be an advantageous alternative. Regarding the GC content, the alternative bacterial expression hosts utilized within the ExpresSys project (*E. coli* – 50.8%, *P. putida* – 61.5%, *R. capsulatus* – 68% and *T. thermophilus* – 69.4%) share by far more similarities with the metagenomic clones (LipS – 59.6%, LipT – 59.4%, Est5E5 – 61.1%, LipA8b2 – 55.5%, EstA3 – 56.7% and LipC5 – 57.3%) than the archaeal expression host *S. acidocaldarius* (36.7%) presented in this work. This might explain the clearly more successful expression of the selected esterase clones within the bacterial hosts (personal correspondence with the respective work groups).

The esterase encoding genes Saci_1105 and Saci_1116 were also recombinantly expressed in *E. coli* Rosetta (DE3) for further characterization. Whereas the esterase Saci_1105 was expressed at a rather low level as soluble protein, the expression of the esterase Saci_1116 was quite sufficient but apparently as insoluble protein within the membrane fraction (Figure 10). The esterase Saci_1105 was partially characterized. The protein exhibited activity within a pH range from 5 to 8, with a pH optimum of 6.5 (Figure 11A). This correlates with the intracellular pH of *S. acidocaldarius*, indicating Saci_1105 being an intracellular esterase. Its substrate dependency followed Michaelis-Menten kinetics with a catalytic efficiency k_{cat}/K_m of 365 $s^{-1} \cdot mM^{-1}$ and a K_m value of 0.6 mM for the substrate ρNP-C4 (Figure 11B).

The esterase Saci_1116 was actually the first hyperthermophilic esterase to be natively isolated and characterized (Sobek and Görisch, 1988, 1989). This tetrameric esterase was shown to possess four active centres and it exhibits a two-step reaction mechanism with a simultaneous acyl-group transfer and substrate hydrolysis (Sobek and Görisch, 1989). The enzyme exhibits a broad substrate specificity and is inhibited by phenylmethanesulphonyl fluoride and paraoxon, classifying the enzyme as a serine esterase (Sobek and Görisch, 1988). To date, recombinant expression experiments or kinetic parameters are not reported for this protein. The reported preference for short-chain acyl esters of this enzyme is in accordance with the detected phenotype of the deletion strain MW001Δ1116. Another previously characterized esterase from *S. acidocaldarius* (Saci_2213) is a putative extracellular active enzyme due to the similarity of its N-terminal sequence to typical prokaryotic signal peptides. The enzyme also shows the conserved three residues comprising the catalytic centre of lypolytic enzymes. Activity screenings with the recombinantly expressed protein revealed the hydrolysis of tributyrin, ρNP-C4 and 1,2-*O*-dilauryl-*rac*-glycero-3-glutaric acid-resorufin ester (DGGR, chromogenic triglyceride analogue) and heat resistance (Arpigny *et al.*, 1998).

Since then several esterases have been described from other archaeal species, for some the genes are identified and for some catalytic kinetics were characterized (Levisson *et al.*, 2009). From the close relative *S. solfataricus* P2 four esterases have been described (Chung *et al.*, 2000; Kim and Lee, 2004; Mandrich *et al.*, 2007; Wang *et al.*, 2010; Shang *et al.*, 2010), for three of them the genes have been identified: SSO2493, SSO2517 and SSO2518. The esterase SSO2517 shares 72% sequence identity with Saci_1105 and 50% sequence identity with Saci_1116 (Mandrich *et al.*, 2007). Interestingly, a short and a long version of SSO2517 were reported as active enzymes, with the longer

version exhibiting better catalytic features towards short-chain acyl esters and temperature and pH characteristics corresponding better to *S. solfataricus* P2 life style. However, it remains unclear which version is present *in vivo*.

For *S. solfataricus* P1 the presence of an arylesterase has been reported, exhibiting 100% amino acid sequence identity with the *S. acidocaldarius* Saci_1105 esterase (Park *et al.*, 2008). The *S. solfataricus* P1 arylesterase (EC 3.1.1.2) showed a high preference for the organophosphate substrate paraoxon (120 $s^{-1} \cdot \mu M^{-1}$), followed by the aromatic esters phenyl acetate (31 $s^{-1} \cdot \mu M^{-1}$) and α-naphthyl acetate (15 $s^{-1} \cdot \mu M^{-1}$), and the ρNP esters ρNP-caproate (C6) (15 $s^{-1} \cdot \mu M^{-1}$) and ρNP-octanoate (C8) (15 $s^{-1} \cdot \mu M^{-1}$), whereas with ρNP-butyrate (C4) only a catalytic efficiency of 1.3 $s^{-1} \cdot \mu M^{-1}$ was achieved (Park *et al.*, 2008). Furthermore, a carboxylesterase with 97% amino acid sequence identity to the *S. acidocaldarius* esterase Saci_1116 has been characterized from *S. solfataricus* P1 with (Park *et al.*, 2006). The *S. solfataricus* P1 carboxylesterase exhibited a high preference for ρNP-caproate (C6) (179 $s^{-1} \cdot \mu M^{-1}$) and ρNP-octanoate (C8) (207 $s^{-1} \cdot \mu M^{-1}$). The extraordinary similarity on amino acid level between the *S. solfataricus* P1 and *S. acidocaldarius* esterases suggests also a high similarity of enzyme characteristics. The phenotype of MW001Δ1116 agrees well with the reported characteristics of the *S. solfataricus* P1 carboxylesterase. However, the phenotype of MW001Δ1105 implies a preference for ρNP-butyrate (C4) and ρNP-decanoate (C10) over ρNP-octanoate (C8) of the esterase Saci_1105, which is not in accordance with the substrate preference described for the *S. solfataricus* P1 arylesterase.

Based on sequence comparison with the characterized *S. solfataricus* esterases the highly conserved active site consensus GXSXG motif, typical for lipolytic enzymes (Jäger *et al.*, 1994), was also found in the *S. acidocaldarius* esterases presented in this work (Saci_1105: G_{154}-G_{158}, Saci_1116: G_{149}-G_{153}).

In summary, in course of this work *S. acidocaldarius* was further developed towards a suitable screening host for diverse metagenomic analyses and protein expression from distantly related organisms. Heterologous expression has been shown to be successful for phylogenetically different phosphofructokinases (PFK) under the control of the homologous *malE* promoter with the pSVA1450 vector. This enables for future metabolic engineering studies. Furthermore, PFK expression was also effectively initiated using a heterologous euryarchaeal promoter. Heterologous promoter recognition is a pre-requisite for functional metagenomic screenings of especially large-insert libraries. Besides, an esterase deficient *S. acidocaldarius* MW001Δ1105Δ1116 strain has been successfully constructed, suitable for metagenomic screenings for esterase activity. Unfortunately, the further improvement of the vector system towards pSVA2301 did not yield the expected functional expression of metagenomic esterases, although the transcription into mRNA could be proven.

4.2 Modification of GDH-2 Substrate Specificity By Active Site Mutations

The branched ED-pathway is the most abundant ED-pathway variant in Archaea for glucose degradation, comprised of a semiphosphorylative-ED branch and a nonphosphorylative-ED branch (Ahmed *et al.*, 2005). As shown particularly for *Sulfolobus spp.*, the branched ED pathway is promiscuous for the breakdown of glucose as well as galactose (Figure 5) (Ahmed *et al.*, 2005; Lamble

et al., 2005, 2003). The first pathway enzyme reaction, the oxidation of D-glucose to D-gluconate, is catalyzed by two isoenzymes, GDH-1 (SSO3003) and GDH-2 (SSO3204) (Lamble *et al.*, 2003; Haferkamp *et al.*, 2011). GDH-1 exhibits a broad substrate spectrum, oxidizing D-glucose, D-galactose, D-xylose, L-arabinose, D-fucose and 6-deoxy-D-glucose, with a preference for D-xylose, and uses both co-substrates NAD^+ and $NADP^+$ with approximately the same efficiency (Lamble *et al.*, 2003; Milburn *et al.*, 2006). In contrast, the GDH-2 shows high specificity for D-glucose as substrate and $NADP^+$ as co-substrate (Haferkamp *et al.*, 2011), which could be confirmed in the present work (117.7 $s^{-1} \cdot mM^{-1}$ D-glucose/$NADP^+$). Additionally, the enzyme utilizes in the presence of the co-substrate $NADP^+$ 6-deoxy-D-glucose (36.0 $s^{-1} \cdot mM^{-1}$) with moderate activity, and D-xylose (3.7 $s^{-1} \cdot mM^{-1}$), D-galactose (0.9 $s^{-1} \cdot mM^{-1}$), D-glucosamine (4.3 $s^{-1} \cdot mM^{-1}$), 2-deoxy-D-glucose (5.0 $s^{-1} \cdot mM^{-1}$) are utilized with minor activities (Table 10).

To date, several archaeal GDHs are characterized next to the *S. solfataricus* GDH-1 and GDH-2: *Thermoplasma acidophilum* (Ta0897) (Smith *et al.*, 1989), *Haloferax mediterranei* (HFX_1090) (Bonete *et al.*, 1996), *Thermoproteus tenax* (TTX_0329) (Siebers *et al.*, 1997) and *Picrophilus torridus* (PTO1070) (Angelov *et al.*, 2005). The crystal structures have been solved for *S. solfataricus* GDH-1 (Milburn *et al.*, 2006), *H. mediterranei* (Ferrer *et al.*, 2001; Esclapez *et al.*, 2005; Baker *et al.*, 2009) and *T. acidophilum* (Bright *et al.*, 1991; John *et al.*, 1994). For *S. solfataricus* GDH-1 as well as for *P. torridus* GDH and *T. acidophilum* GDH a promiscuity for utilizing D-glucose and at least D-galactose has been reported (Giardina *et al.*, 1986; Smith *et al.*, 1989; Angelov *et al.*, 2005; Lamble *et al.*, 2005), whereas the *S. solfataricus* GDH-2, *H. mediterranei* GDH and *T. tenax* GDH are described to be specific for D-glucose (Siebers *et al.*, 1997; Bonete *et al.*, 1996; Haferkamp *et al.*, 2011). Due to their different substrate preference it is hypothesized that GDH-2 is the main enzyme for glucose degradation in *S. solfataricus*, while GDH-1 might function in different sugar degradation pathways and, therefore, serves more as a standby enzyme for quick adaptation to different carbon sources (Haferkamp *et al.*, 2011).

Glucose dehydrogenases can be found in all three domains of life and the characterized GDHs have been reported to be either homotetramers or homodimers composed of ~40 kDa subunits (Bhaumik and Sonawat, 1999; Lamble *et al.*, 2003; Ohshima *et al.*, 2003; Angelov *et al.*, 2005; Haferkamp *et al.*, 2011). GDHs belong to the medium-chain alcohol/polyol dehydrogenase/reductase (MDR) superfamily, exhibiting the typical two-domain organization: the central nucleotide binding domain is flanked by protein regions at the N and C termini, together representing the catalytic domain (Edwards *et al.*, 1996).

A sequence alignment of the characterized archaeal GDHs revealed a high conservation of residues involved in binding interactions within the active site (Figure 22). The fingerprint motif of the nucleotide binding site (GxGxxG/A) is highly conserved throughout the archaeal GDHs. In *S. solfataricus* GDH-1 the structural zinc is coordinated by four cysteine residues, of which three are conserved in *T. acidophilum* GDH (John *et al.*, 1994; Milburn *et al.*, 2006). In *H. mediterranei* GDH these residues are not conserved at all, indicating the structural zinc not being essential for function, as described for other MDR family members (Persson *et al.*, 1994; Edwards *et al.*, 1996; Pauly *et al.*, 2003; Ying and Ma, 2011). The dual co-substrate specificity is conferred by a negatively charged and a

conserved positively charged residue (Arg in *S. solfataricus* GDH-1 and *H. mediterranei* GDH, His in *T. acidophilum* GDH) (John *et al.*, 1994; Milburn *et al.*, 2006; Baker *et al.*, 2009). The catalytic zinc binding is coordinated by four highly conserved residues (Cys, His, and two Glu), except in *H. mediterranei* GDH the Cys is exchanged for Asp, discussed as a halophilic adaptation, and in *S. solfataricus* GDH-1 the second Glu is exchanged for Gln (John *et al.*, 1994; Pire *et al.*, 2001; Milburn *et al.*, 2006; Baker *et al.*, 2009). For *H. mediterranei* GDH active site dynamics during catalysis, i.e. the movement of the catalytic zinc ion between the residues, has been described in detail, revealing a subsequent rather than a simultaneous interaction and showing the His residue being the only constant binding partner (Baker *et al.*, 2009). In *S. solfataricus* GDH-1 this important His residue is highly conserved and is, together with the residues Glu and Gln, also involved in coordinating the zinc-bound water (Milburn *et al.*, 2006). Interestingly, in *S. solfataricus* GDH-1 the His residue is also hydrogen bonded by the residue Asn, which is involved in substrate binding (Milburn *et al.*, 2006). This position is occupied by a valine in all other archaeal GDHs, for which a hydrogen bond to the zinc binding histidine is not reported, indicating a different active site architecture being present in these GDHs. Thus, a modification within such a complex interaction network in the active site can have severe consequences on catalysis.

In *S. solfataricus* GDH-1 six residues have been described to be involved in substrate binding, four of which are identical or highly conserved in all archaeal GDHs, indicating their essential role for substrate binding (Asn^{89}, Glu^{114}, Gln^{150} and Asn^{307}, *S. solfataricus* GDH-1 numbering) (Milburn *et al.*, 2006). Except in *S. solfataricus* GDH-1, as mentioned above, the first Asn residue, which was subject for site directed mutagenesis in this study, is substituted for a Val residue in all other analyzed GDHs. Notably, for the position 150 in *S. solfataricus* GDH-1 (Gln) as well as in the *H. mediterranei* GDH (Glu) also a catalytic zinc binding involvement has been described (in Figure 22 these residues are marked both, green and red, indicating the involvement in catalytic zinc and substrate binding, respectively). In contrast, the other two residues, discussed to be involved in substrate binding, appear to be more variable (His^{297} and Asp^{154}, *S. solfataricus* GDH-1 numbering).

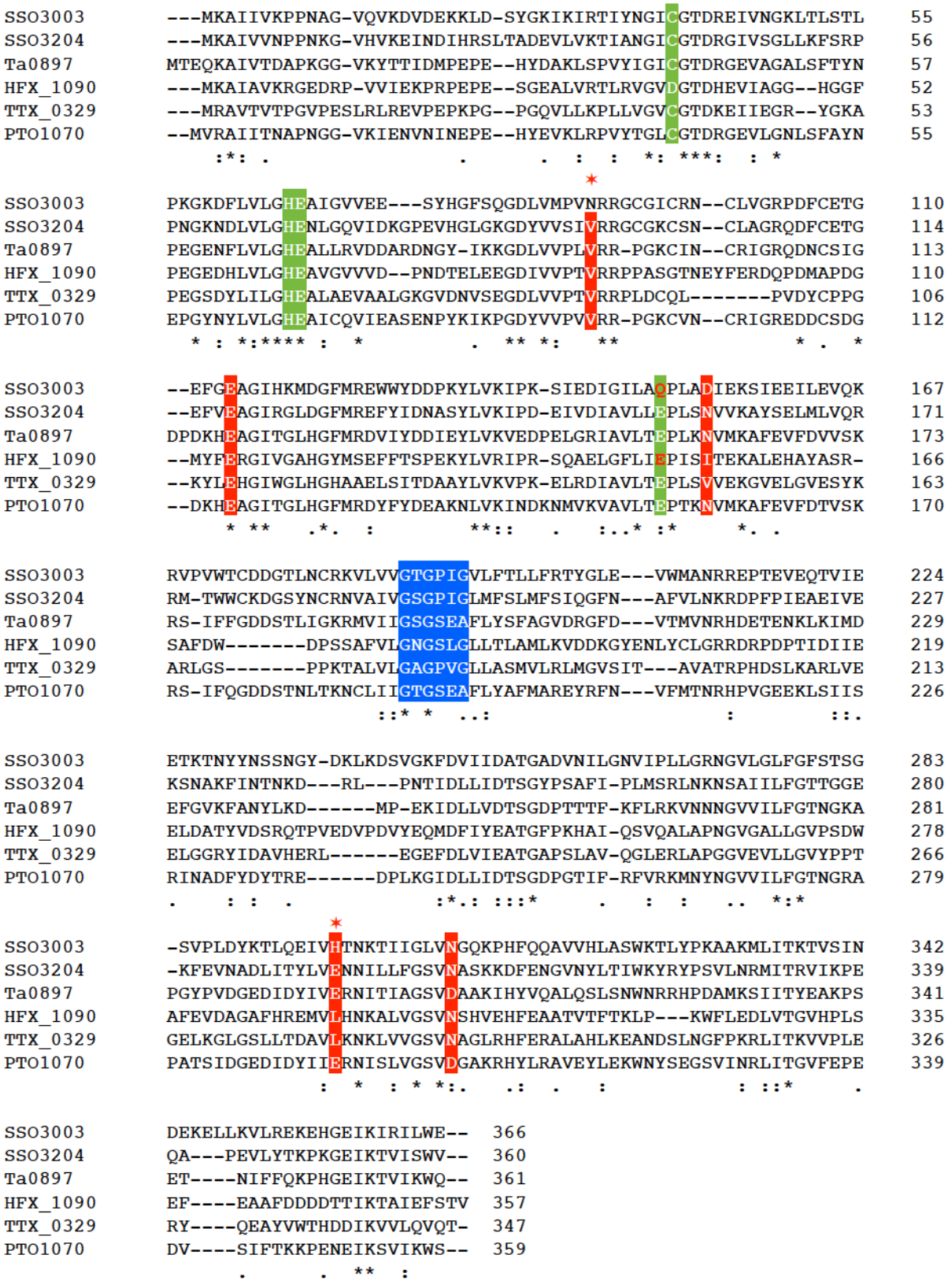


Figure 22. Sequence alignment of archaeal glucose dehydrogenases. Amino acid sequences of characterized archaeal GDH enzymes are shown. *S. solfataricus* (SSO3003 (GDH-1), SSO3204 (GDH-2)), *Thermoplasma acidophilum* (Ta0897), *Haloferax mediterranei* (HFX_1090), *Thermoproteus tenax* (TTX_0329) and *Picrophilus torridus* (PTO1070). The conserved fingerprint motif for nucleotide binding is marked blue. Residues discussed to be involved in diverse binding events are marked as follows: substrate binding (red), catalytic zinc binding (green). Residues subjected to site directed mutagenesis in this study are marked (✶). Alignment symbols: (*) identical amino acids; (:) highly conserved amino acids; (.) weakly conserved amino acids. The program Clustal Omega was used for the alignment (http://www.ebi.ac.uk/Tools/msa/clustalo/).

S. solfataricus GDH-2 shows 41% amino acid sequence identity to GDH-1 and in the range of 29-41% identity to the other archaeal GDHs. GDH-2 contains the conserved residues for nucleotide, co-substrate and zinc binding mentioned above. However, structure modeling of GDH-2, based on the GDH-1 structure and sequence alignment revealed differences in residues within the active site as described above. The main structural differences proposed to explain the higher specificity for D-glucose are the substitutions of Asn^{89} to Val^{93} and His^{297} to Glu^{294} in GDH-2 (Haferkamp *et al.*, 2011). From the structural model it has been suggested that the substitution N89V results in the loss of the hydrogen bond to the substrate's C3-hydroxyl group, thus weakening affinity to D-glucose and to D-xylose. The substitution H297E, in contrast, was assumed to strengthen the hydrogen bonding to the C6-hydroxyl group of the substrate, possibly compensating the weaker interaction to D-glucose caused by the N89V substitution, thus GDH-2 still binds D-glucose (Haferkamp *et al.*, 2011). D-xylose, in contrast, does not have a C6-hydroxyl group, so that the loss of the hydrogen bond due to the N89V substitution cannot be compensated. This may explain the reduced affinity for D-xylose observed for GDH-2. To address the structural basis determining such divergent substrate specificity of so closely related enzymes, which has not been studied for GDH so far, the substitution of the residues Val^{93} to Asn and Glu^{294} to His in GDH-2 was conducted via site directed mutagenesis. The resulting GDH-2 mutants GDH-2_V93N, GDH-2_E294H and GDH-2_V93N_E294H were biochemically characterized in comparison to the wild-type GDH-2 (Table 10).

All three mutations caused a decrease in catalytic efficiency, whereas the E294H single mutation exhibited a less severe effect on catalytic efficiency (D-glucose/$NADP^+$: 26.3 $s^{-1}\cdot mM^{-1}$) compared to the V93N single mutation (D-glucose/$NADP^+$: 4.9 $s^{-1}\cdot mM^{-1}$) and V93N_E294H double mutation (D-glucose/$NADP^+$: 2.1 $s^{-1}\cdot mM^{-1}$). Same tendency was seen also for the other substrates. Primarily, the E294H mutation yields a decreased substrate affinity (K_m) for hexoses, confirming a weaker interaction of the histidine residue towards the substrate's C6-hydroxyl group compared to glutamate, as assumed from structure modeling studies (Milburn *et al.*, 2006; Haferkamp *et al.*, 2011). In accordance with this role in C6-hydroxyl interaction, the kinetic parameters for the substrates 6-deoxy-D-glucose and D-xylose, lacking a C6-hydroxyl group or the complete C6 position, remained nearly the same for this mutant compared to wild-type enzyme.

The V93N mutation, in contrast, was assumed to result in better substrate affinities and consequently better catalysis due to an addition of a hydrogen bond to the hydroxyl group at C3 position of the substrate (Milburn *et al.*, 2006; Haferkamp *et al.*, 2011). The results, however, showed mainly an increase of K_m with a simultaneous decrease of V_{max} or even in an abolished enzyme activity. This may imply that the residue in this position plays an important role not only for substrate binding, but also e.g. for active site architecture and/or dynamics during catalysis.

The effect the double mutant V93N_E294H had on activity reflects a combination of the single mutations. Residual activity with dramatically decreased catalytic efficiency could only be observed for the two substrates D-glucose and 6-deoxy-D-glucose. For all other substrates the activity was completely abolished.

From structural studies on *S. solfataricus* GDH-1 it was concluded, that both C5 and C6 sugars are utilized in a common orientation/position and that mainly the interactions of the enzyme to the substrate's C1-C3-hydroxyl groups have an influence on the substrate binding capacity (Milburn *et al.*, 2006). Structural modeling experiments showed that especially alternative epimers of the C2- and

C3-hydroxyls (compared to glucose) will result in the loss of the majority of hydrogen bonds. This is in accordance with the observed severe effects the V93N mutant (interaction with the substrate's C3-hydroxyl group) had on catalytic efficiencies. In contrast, for substrates with a different orientation of the C4-hydroxyl group (e.g. galactose) it was assumed, that the loss of a hydrogen bond to one of the two interacting enzyme residues (Asn^{307}) may be compensated by the formation of a second hydrogen bond to the other residue (Glu^{114}) in GDH-1. Hence, there is no structural evidence for a difference in affinity between glucose and galactose in GDH-1. The V93N mutant utilized the substrate pair D-galactose/$NADP^+$ with an increased affinity but a decreased V_{max} compared to the wild-type enzyme, whereas the E294H mutant did not utilize this substrate at all. In GDH-1 an interaction only to the C3-hydroxyl group was reported for the residue N89, although the binding mode of galactose has not been investigated. This indicates a possibly different substrate orientation in the active site of GDH-2 compared to GDH-1, where this residue position may, next to the C3-hydroxyl group interaction, also interact with the substrate's C4-hydroxyl group, maybe explaining the simultaneous impact on K_m and V_{max} by the V93N mutation.

In summary, the comparison of the respective GDH-2 mutants with the wild-type GDH-2 revealed that the investigated active site residues do have an impact on substrate binding. However, the structural determinants conferring the different substrate specificities still remain to be elucidated. The structural data available for some archaeal GDHs rely only on crystallization experiments with wild-type enzymes in complex with glucose. Thus, the binding mode of other substrates is not known. Therefore, crystallization studies with the wild-type GDH-2 enzyme for direct comparison would give a more detailed picture on residues involved in substrate binding and possibly in assigning substrate specificity. A sequence alignment of the characterized archaeal GDHs identified the presence of four highly conserved and two variable residues involved in substrate binding.

Summary

SUMMARY

One of the most efficient mining strategies for novel and unique enzymes, meeting the requirements of industrial applications based on enzymatic processes, is a functional metagenomic screening. Especially the successful expression of so-called extremozymes is often limited in standard screening hosts, which is particularly true for archaeal enzymes. However, due to the (hyper)thermophilic life style of many representatives and their unique metabolism encompassing many unusual pathways and enzymes, Archaea represent a highly valuable source for novel biocatalysts with exceptional catalytic properties and high stability under harsh conditions. These characteristics are advantageous for many industrial applications. The development of an archaeal expression host is, therefore, highly desirable to enable the identification of new extremozymes and the exploitation of their great potential for industrial applications, which would otherwise remain undetected.

The establishment of the genetically tractable (hyper)thermophilic Crenarchaeon *S. acidocaldarius* as an expression host for functional metagenomic screenings was a major part of the present work. Heterologous promoter recognition is a prerequisite for functional metagenomic screenings especially of large-insert libraries. Therefore, the promoter selectivity of *S. acidocaldarius* was investigated via phosphofructokinase (PFK) promoters derived from different (hyper)thermophilic cren- and euryarchaeal as well as bacterial sources using their natively controlled PFK genes as reporters: the crenarchaeal PP_i-dependent PFK from *Thermoproteus tenax* (TTX_1277), the euryarchaeal ADP-dependent PFK from *Pyrococcus furiosus* (PF1784) and the bacterial ATP-dependent PFK from *Thermotoga maritima* (TM0209). The euryarchaeal *P. furiosus* PFK promoter recognition by *S. acidocaldarius* as well as the heterologous expression of all three PFKs under the control of the frequently applied *S. acidocaldarius malE* promoter was successfully shown using the vector pSVA1450. These results indicate the suitability of *S. acidocaldarius* as a host for functional screening of large-insert metagenomic libraries as well as heterologous protein (over)expression. Furthermore, *S. acidocaldarius* utilizes a branched Entner-Doudoroff (ED) pathway for glucose degradation and the Embden-Meyerhof-Parnas (EMP) pathway for gluconeogenesis. Hence, the successful heterologous expression of a PFK now enables first steps towards metabolic engineering, i.e. directing the EMP-pathway into a glycolytic direction, since a PFK homolog is missing in the *S. acidocaldarius* genome.

The expression capabilities of *S. acidocaldarius* were further analyzed via esterase genes functionally identified from different metagenomes. The six selected esterases showed a diverse range of characteristics with regard to origin, temperature optima, and substrate specificity. Since the *S. acidocaldarius* MW001 exhibited esterase activity towards short-chain acyl esters (i.e. ρNP-C4 and ρNP-C8), an esterase deficient *S. acidocaldarius* strain (MW001Δ1105Δ1116) was successfully constructed by deleting the identified esterase genes Saci_1105 and Saci_1116 allowing for metagenomic esterase screenings. The applicability of pSVA1450 in metagenomic screenings is rather limited. Therefore, the vector was further improved by inserting additional restriction sites and a N-terminal Strep-10xHis-tag, discarding the pre-cloning step via pMZ1, yielding the vector pSVA2301. Nevertheless, this did not result in the expected functional expression of the metagenomic esterases

under various expression conditions even though the transcription of the genes was proven by mRNA analysis.

The esterases Saci_1105 and Saci_1116 were also subjected to recombinant expression in *E. coli* for further characterization. Initial biochemical studies revealed a pH optimum at pH 6.5 for Saci_1105. This corresponds to the intracellular pH of *S. acidocaldarius*, thus suggesting Saci_1105 being an intracellular esterase. Its substrate dependency followed Michaelis-Menten kinetics with a catalytic efficiency k_{cat}/K_m of 365 $s^{-1} \cdot mM^{-1}$ and a K_m value of 0.6 mM for the substrate ρNP-C4.

For glucose degradation the branched ED-pathway is the most abundant ED-pathway variant in Archaea, which is comprised of a semiphosphorylative-ED branch and a nonphosphorylative-ED branch. In *Sulfolobus spp*. this branched ED pathway is promiscuous for the breakdown of glucose as well as galactose. Two glucose dehydrogenases (GDH), the first pathway enzyme, have been characterized for *S. solfataricus*: GDH-1 (SSO3003) exhibits a broad substrate spectrum whereas GDH-2 (SSO3204) appears to be glucose-specific. Structure modeling revealed differences in residues within the active site, discussed to be involved in substrate binding. The second major part of the present work aimed for the elucidation of the structural determinants conferring the different substrate specificity. For this, GDH-2 mutants were constructed via site-directed mutagenesis, substituting amino acids found in corresponding positions to the substrate binding residues in GDH-1: GDH-2_V93N, GDH-2_E294H and the double mutant GDH-2_V93N_E294H. Biochemical characterization revealed the E294H mutation having a lower impact on substrate utilization for most tested substrates compared to the V93N and V93N_E294H mutation. The V93N mutation showed mainly an increase of K_m with a simultaneous decrease of V_{max} or even an abolished enzyme activity, suggesting the residue in this position is playing an important role not only for substrate binding, but also e.g. for active site architecture and/or dynamics during catalysis. The E294H mutation, in contrast, showed primarily an increasing effect on K_m for the tested hexoses, whereas the V_{max} values were not significantly affected, confirming the proposed interaction with the C6-hydroxyl group of the substrate. In accordance with the effects the single mutants caused on activity, the kinetic parameters for the double mutant V93N_E294H reflect a combination thereof. For this mutant residual activity with dramatically decreased catalytic efficiency was only observed for two substrates, whereas for all other tested substrates the activity was completely abolished. A sequence alignment of six characterized archaeal GDHs revealed a high conservation of residues involved in binding actions within the active site. The presence of four highly conserved and two variable residues involved in substrate binding could be identified.

LITERATURE

LITERATURE

AHMED, H., ETTEMA, T.J.G., TJADEN, B., GEERLING, A.C.M., VAN DER OOST, J., SIEBERS, B. (2005) The semi-phosphorylative Entner-Doudoroff pathway in hyperthermophilic Archaea: a re-evaluation. *Biochem. J.*, **390**, 529–540.

AIYAR, S.E., GOURSE, R.L., ROSS, W. (1998) Upstream A-tracts increase bacterial promoter activity through interactions with the RNA polymerase alpha subunit. *Proc. Natl. Acad. Sci. U. S. A.*, **95**, 14652–7.

ALBERS, S.-V., JONUSCHEIT, M., DINKELAKER, S., URICH, T., KLETZIN, A., TAMPÉ, R., DRIESSEN, A.J.M., SCHLEPER, C. (2006) Production of recombinant and tagged proteins in the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Appl. Environ. Microbiol.*, **72**, 102–111.

ALBERS, S.-V., KONING, S.M., KONINGS, W.N., DRIESSEN, A.J. (2004) Insights into ABC transport in Archaea. *J. Bioenerg. Biomembr.*, **36**, 5–15.

ALBERS, S.-V., MEYER, B.H. (2011) The archaeal cell envelope. *Nat. Rev. Microbiol.*, **9**, 414–426.

ALLEN, W.G. (1976) Potential applications for cellulase enzymes. *Biotechnol. Bioeng. Symp.*, 303–5.

AMANN, R.I., LUDWIG, W., SCHLEIFER, K.H. (1995) Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol. Rev.*, **59**, 143–69.

ANDERSEN, H.D., WANG, C., ARLETH, L., PETERS, G.H., WESTH, P. (2011) Reconciliation of opposing views on membrane-sugar interactions. *Proc. Natl. Acad. Sci. U. S. A.*, **108**, 1874–8.

ANGELOV, A., FÜTTERER, O., VALERIUS, O., BRAUS, G.H., LIEBL, W. (2005) Properties of the recombinant glucose/galactose dehydrogenase from the extreme thermoacidophile, *Picrophilus torridus*. *FEBS J.*, **272**, 1054–1062.

ANTRANIKIAN, G., VORGIAS, C.E., BERTOLDO, C. (2005) Extreme environments as a resource for microorganisms and novel biocatalysts. *Adv. Biochem. Eng. Biotechnol.*, **96**, 219–262.

ARPIGNY, J.L., JÄGER, K.E. (1999) Bacterial lipolytic enzymes: classification and properties. *Biochem. J.*, **343 Pt 1**, 177–83.

ARPIGNY, J.L., JENDROSSEK, D., JÄGER, K.E. (1998) A novel heat-stable lipolytic enzyme from *Sulfolobus acidocaldarius* DSM 639 displaying similarity to polyhydroxyalkanoate depolymerases. *FEMS Microbiol. Lett.*, **167**, 69–73.

ATOMI, H., IMANAKA, T. (2004) Thermostable carboxylesterases from hyperthermophiles. *Tetrahedron: Asymmetry*, **15**, 2729–2735.

ATOMI, H., IMANAKA, T., FUKUI, T. (2012) Overview of the genetic tools in the Archaea. *Front. Microbiol.*, **3**.

ATOMI, H., MATSUMI, R., IMANAKA, T. (2004) Reverse gyrase is not a prerequisite for hyperthermophilic life. *J. Bacteriol.*, **186**, 4829–33.

AUCHTUNG, T.A., TAKACS-VESBACH, C.D., CAVANAUGH, C.M. (2006) 16S rRNA phylogenetic investigation of the candidate division '*Korarchaeota*'. *Appl. Environ. Microbiol.*, **72**, 5077–82.

BAKER, P.J., BRITTON, K.L., FISHER, M., ESCLAPEZ, J., PIRE, C., BONETE, M.J., FERRER, J., RICE, D.W. (2009) Active site dynamics in the zinc-dependent medium chain alcohol dehydrogenase superfamily. *Proc. Natl. Acad. Sci. U. S. A.*, **106**, 779–84.

BANERJEE, A., GHOSH, A., MILLS, D.J., KAHNT, J., VONCK, J., ALBERS, S.-V. (2012) FlaX, a unique component of the crenarchaeal archaellum, forms oligomeric ring-shaped structures and interacts with the motor ATPase FlaI. *J. Biol. Chem.*, **287**, 43322–30.

BARNS, S.M., DELWICHE, C.F., PALMER, J.D., PACE, N.R. (1996) Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences. *Proc. Natl. Acad. Sci. U. S. A.*, **93**, 9188–93.

BASEN, M., SUN, J., ADAMS, M.W.W. (2012) Engineering a Hyperthermophilic Archaeon for Temperature-Dependent Product Formation. *MBio*, **3**.

BERKNER, S., GROGAN, D., ALBERS, S.-V., LIPPS, G. (2007) Small multicopy, non-integrative shuttle vectors based on the plasmid pRN1 for *Sulfolobus acidocaldarius* and *Sulfolobus solfataricus*, model organisms of the (Cren-)Archaea. *Nucl. Acids Res.*, **35**.

BERKNER, S., LIPPS, G. (2008) Genetic tools for *Sulfolobus spp.*: vectors and first applications. *Arch. Microbiol.*, **190**, 217–230.

BERKNER, S., WLODKOWSKI, A., ALBERS, S.-V., LIPPS, G. (2010) Inducible and constitutive promoters for genetic systems in *Sulfolobus acidocaldarius*. *Extrem. Life under Extrem. Cond.*, **14**, 249–259.

BERTOLDO, C., ANTRANIKIAN, G. (2002) Starch-hydrolyzing enzymes from thermophilic Archaea and Bacteria. *Curr. Opin. Chem. Biol.*, **6**, 151–160.

BHALLA, A., BANSAL, N., KUMAR, S., BISCHOFF, K.M., SANI, R.K. (2013) Improved lignocellulose conversion to biofuels with thermophilic Bacteria and thermostable enzymes. *Bioresour. Technol.*, **128**, 751–9.

BHAUMIK, S.R., SONAWAT, H.M. (1999) Kinetic mechanism of glucose dehydrogenase from *Halobacterium salinarum*. *Indian J. Biochem. Biophys.*, **36**, 143–9.

BIWER, A., ANTRANIKIAN, G., HEINZLE, E. (2002) Enzymatic production of cyclodextrins. *Appl. Microbiol. Biotechnol.*, **59**, 609–17.

BLUMER-SCHUETTE, S.E., BROWN, S.D., SANDER, K.B., BAYER, E.A., KATAEVA, I., ZURAWSKI, J. V, CONWAY, J.M., ADAMS, M.W.W., KELLY, R.M. (2014) Thermophilic lignocellulose deconstruction. *FEMS Microbiol. Rev.*, **38**, 393–448.

BONETE, M.J., PIRE, C., LLORCA, F.I., CAMACHO, M.L. (1996) Glucose dehydrogenase from the halophilic archaeon *Haloferax mediterranei*: enzyme purification, characterisation and N-terminal sequence. *FEBS Lett.*, **383**, 227–9.

BORNSCHEUER, U.T. (2002) Microbial carboxyl esterases: classification, properties and application in biocatalysis. *FEMS Microbiol. Rev.*, **26**, 73–81.

BRADFORD, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.*, **72**, 248–54.

BRÄSEN, C., ESSER, D., RAUCH, B., SIEBERS, B. (2014) Carbohydrate metabolism in Archaea: current insights into unusual enzymes and pathways and their regulation. *Microbiol. Mol. Biol. Rev.*, **78**, 89–175.

BRENNEIS, M., HERING, O., LANGE, C., SOPPA, J. (2007) Experimental characterization of Cis-acting elements important for translation and transcription in halophilic Archaea. *PLoS Genet.*, **3**.

BRIGHT, J.R., MACKNESS, R., DANSON, M.J., HOUGH, D.W., TAYLOR, G.L., TOWNER, P., BYROM, D. (1991) Crystallization and preliminary crystallographic study of glucose dehydrogenase from the archaebacterium *Thermoplasma acidophilum*. *J. Mol. Biol.*, **222**, 143–4.

BROCHIER-ARMANET, C., BOUSSAU, B., GRIBALDO, S., FORTERRE, P. (2008) Mesophilic crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. *Nat. Rev. Microbiol.*, **6**, 245–52.

BROCHIER-ARMANET, C., FORTERRE, P., GRIBALDO, S. (2011) Phylogeny and evolution of the Archaea: one hundred genomes later. *Curr. Opin. Microbiol.*, **14**, 274–81.

BROCK, T.D., BROCK, K.M., BELLY, R.T., WEISS, R.L. (1972) *Sulfolobus*: A new genus of sulfur-oxidizing bacteria living at low pH and high temperature. *Arch. Microbiol.*, **84**, 54–68.

CAMBILLAU, C., CLAVERIE, J.M. (2000) Structural and genomic correlates of hyperthermostability. *J. Biol. Chem.*, **275**, 32383–6.

CANNIO, R., CONTURSI, P., ROSSI, M., BARTOLUCCI, S. (2001) Thermoadaptation of a mesophilic hygromycin B phosphotransferase by directed evolution in hyperthermophilic Archaea: selection of a stable genetic marker for DNA transfer into *Sulfolobus solfataricus*. *Extremophiles*, **5**, 153–9.

CHABAN, B., NG, S.Y.M., JARRELL, K.F. (2006) Archaeal habitats--from the extreme to the ordinary. *Can. J. Microbiol.*, **52**, 73–116.

CHEN, L., BRÜGGER, K., SKOVGAARD, M., REDDER, P., SHE, Q., TORARINSSON, E., GREVE, B., AWAYEZ, M., ZIBAT, A., KLENK, H., GARRETT, R.A., BRU, K. (2005) The genome of *Sulfolobus acidocaldarius*, a model organism of the Crenarchaeota. *J. Bacteriol.*, **187**, 4992–4999.

CHOI, K.-H., HWANG, S., CHA, J. (2013) Identification and characterization of MalA in the maltose/maltodextrin operon of *Sulfolobus acidocaldarius* DSM639.

CHONG, P.K., BURJA, A.M., RADIANINGTYAS, H., FAZELI, A., WRIGHT, P.C. (2007) Translational and transcriptional analysis of *Sulfolobus solfataricus* P2 to provide insights into alcohol and ketone utilisation. *Proteomics*, **7**, 424–35.

CHOW, J. (2012) Biochemical and Structural Characterization Of Three Thermostable and Metagenome-Derived Lipolytic Enzymes. *PhD Thesis*, 1–147.

CHOW, J., KOVACIC, F., DALL ANTONIA, Y., KRAUSS, U., FERSINI, F., SCHMEISSER, C., LAUINGER, B., BONGEN, P., PIETRUSZKA, J., SCHMIDT, M., MENYES, I., BORNSCHEUER, U.T., ECKSTEIN, M., THUM, O., LIESE, A., MUELLER-DIECKMANN, J., JÄGER, K.E., STREIT, W.R. (2012) The metagenome-derived enzymes LipS and LipT increase the diversity of known lipases. *PLoS One*, **7**.

CHUNG, Y.M., PARK, C.B., LEE, S.B. (2000) Partial purification and characterization of thermostable esterase from the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Biotechnol. Bioprocess Eng.*, **5**, 53–56.

COLLINS, R.E., ROCAP, G., DEMING, J.W. (2010) Persistence of bacterial and archaeal communities in sea ice through an Arctic winter. *Environ. Microbiol.*, **12**, 1828–41.

DA COSTA, M.S., SANTOS, H., GALINSKI, E.A. (1998) An overview of the role and diversity of compatible solutes in Bacteria and Archaea. *Adv. Biochem. Eng. Biotechnol.*, **61**, 117–53.

COURTOIS, S., CAPPELLANO, C.M., BALL, M., FRANCOU, F.-X., NORMAND, P., HELYNCK, G., MARTINEZ, A., KOLVEK, S.J., HOPKE, J., OSBURNE, M.S., AUGUST, P.R., NALIN, R., GUÉRINEAU, M., JEANNIN, P., SIMONET, P., PERNODET, J.-L. (2003) Recombinant environmental libraries provide access to microbial diversity for drug discovery from natural products. *Appl. Environ. Microbiol.*, **69**, 49–55.

CRAIG, J.W., CHANG, F.-Y., KIM, J.H., OBIAJULU, S.C., BRADY, S.F. (2010) Expanding small-molecule functional metagenomics through parallel screening of broad-host-range cosmid environmental DNA libraries in diverse Proteobacteria. *Appl. Environ. Microbiol.*, **76**, 1633–41.

DANIEL, R.M., COWAN, D.A. (2000) Biomolecular stability and life at high temperatures. *Cell. Mol. Life Sci.*, **57**, 250–64.

DANIEL, R.M., DINES, M., PETACH, H.H. (1996) The denaturation and degradation of stable enzymes at high temperatures. *Biochem. J.*, **317**, 1–11.

DAS, R., GERSTEIN, M. (2000) The stability of thermophilic proteins: a study based on comprehensive genome comparison. *Funct. Integr. Genomics*, **1**, 76–88.

DAUM, B., QUAX, T.E.F., SACHSE, M., MILLS, D.J., REIMANN, J., YILDIZ, Ö., HÄDER, S., SAVEANU, C., FORTERRE, P., ALBERS, S.-V., KÜHLBRANDT, W., PRANGISHVILI, D. (2014) Self-assembly of the general membrane-remodeling protein PVAP into sevenfold virus-associated pyramids. *Proc. Natl. Acad. Sci. U. S. A.*, **111**, 3829–34.

DELONG, E.F. (1998) Everything in moderation: Archaea as 'non-extremophiles'. *Curr. Opin. Genet. Dev.*, **8**, 649–54.

DENG, L., ZHU, H., CHEN, Z., LIANG, Y.X., SHE, Q. (2009) Unmarked gene deletion and host-vector system for the hyperthermophilic crenarchaeon *Sulfolobus islandicus*. *Extremophiles*, **13**, 735–46.

EDWARDS, K.J., BARTON, J.D., ROSSJOHN, J., THORN, J.M., TAYLOR, G.L., OLLIS, D.L. (1996) Structural and sequence comparisons of quinone oxidoreductase, zeta-crystallin, and glucose and alcohol dehydrogenases. *Arch. Biochem. Biophys.*, **328**, 173–83.

EICHLER, J., ADAMS, M.W.W. (2005) Posttranslational Protein Modification in Archaea. *Microbiol. Mol. Biol. Rev.*, **69**, 393–425.

EKKERS, D.M., CRETOIU, M.S., KIELAK, A.M., ELSAS, J.D. VAN (2012) The great screen anomaly--a new frontier in product discovery through functional metagenomics. *Appl. Microbiol. Biotechnol.*, **93**, 1005–20.

ELEND, C., SCHMEISSER, C., LEGGEWIE, C., BABIAK, P., CARBALLEIRA, J.D., STEELE, H.L., REYMOND, J.-L., JÄGER, K.E., STREIT, W.R. (2006) Isolation and biochemical characterization of two novel metagenome-derived esterases. *Appl. Environ. Microbiol.*, **72**, 3637–3645.

ELKINS, J.G. *ET AL.* (2008) A korarchaeal genome reveals insights into the evolution of the Archaea. *Proc. Natl. Acad. Sci. U. S. A.*, **105**, 8102–7.

EMPADINHAS, N., DA COSTA, M.S. (2006) Diversity and biosynthesis of compatible solutes in hyper/thermophiles. *Int. Microbiol.*, **9**, 199–206.

ENTCHEVA, P., LIEBL, W., JOHANN, A., HARTSCH, T., STREIT, W.R. (2001) Direct cloning from enrichment cultures, a reliable strategy for isolation of complete operons and genes from microbial consortia. *Appl. Environ. Microbiol.*, **67**, 89–99.

ESCLAPEZ, J., BRITTON, K.L., BAKER, P.J., FISHER, M., PIRE, C., FERRER, J., BONETE, M.J., RICE, D.W. (2005) Crystallization and preliminary X-ray analysis of binary and ternary complexes of *Haloferax mediterranei* glucose dehydrogenase. *Acta Crystallogr. Sect. F. Struct. Biol. Cryst. Commun.*, **61**, 743–6.

ESTREM, S.T., GAAL, T., ROSS, W., GOURSE, R.L. (1998) Identification of an UP element consensus sequence for bacterial promoters. *Proc. Natl. Acad. Sci. U. S. A.*, **95**, 9761–6.

ETTEMA, T.J.G., AHMED, H., GEERLING, A.C.M., VAN DER OOST, J., SIEBERS, B. (2008) 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*, **12**, 75–88.

FENDRIHAN, S., LEGAT, A., PFAFFENHUEMER, M., GRUBER, C., WEIDLER, G., GERBL, F., STAN-LOTTER, H. (2006) Extremely halophilic Archaea and the issue of long-term microbial survival. *Re/views Environ. Sci. bio/technology*, **5**, 203–218.

FERRER, J., FISHER, M., BURKE, J., SEDELNIKOVA, S.E., BAKER, P.J., GILMOUR, D.J., BONETE, M.J., PIRE, C., ESCLAPEZ, J., RICE, D.W. (2001) Crystallization and preliminary X-ray analysis of glucose dehydrogenase from *Haloferax mediterranei*. *Acta Crystallogr. D. Biol. Crystallogr.*, **57**, 1887–9.

GIARDINA, P., DE BIASI, M.G., DE ROSA, M., GAMBACORTA, A., BUONOCORE, V. (1986) Glucose dehydrogenase from the thermoacidophilic archaebacterium *Sulfolobus solfataricus*. *Biochem. J.*, **239**, 517–22.

GLICK, B.R. (1995) Metabolic load and heterologous gene expression. *Biotechnol. Adv.*, **13**, 247–261.

GOLYSHINA, O. V, TIMMIS, K.N. (2005) *Ferroplasma* and relatives, recently discovered cell wall-lacking archaea making a living in extremely acid, heavy metal-rich environments. *Environ. Microbiol.*, **7**, 1277–88.

GREGOR, D., PFEIFER, F. (2005) *In vivo* analyses of constitutive and regulated promoters in halophilic Archaea. *Microbiology*, **151**, 25–33.

GROGAN, D.W. (2003) Cytosine methylation by the *Sua*I restriction-modification system: implications for genetic fidelity in a hyperthermophilic archaeon. *J. Bacteriol.*, **185**, 4657–4661.

GROGAN, D.W. (1989) Phenotypic characterization of the archaebacterial genus *Sulfolobus*: comparison of five wild-type strains. *J. Bacteriol.*, **171**, 6710–6719.

GROGAN, D.W., GUNSALUS, R.P. (1993) *Sulfolobus acidocaldarius* synthesizes UMP via a standard de novo pathway: results of biochemical-genetic study. *J. Bacteriol.*, **175**, 1500–7.

GUO, L., BRÜGGER, K., LIU, C., SHAH, S.A., ZHENG, H., ZHU, Y., WANG, S., LILLESTØL, R.K., CHEN, L., FRANK, J., PRANGISHVILI, D., PAULIN, L., SHE, Q., HUANG, L., GARRETT, R.A. (2011) Genome analyses of Icelandic strains of *Sulfolobus islandicus*, model organisms for genetic and virus-host interaction studies. *J. Bacteriol.*, **193**, 1672–80.

GUPTA, R., GIGRAS, P., MOHAPATRA, H., GOSWAMI, V.K., CHAUHAN, B. (2003) Microbial a-amylases: a biotechnological perspective. *Process Biochem.*, 1–18.

HAFERKAMP, P., KUTSCHKI, S., TREICHEL, J., HEMEDA, H., SEWCZYK, K., HOFFMANN, D., ZAPARTY, M., SIEBERS, B. (2011) An additional glucose dehydrogenase from *Sulfolobus solfataricus*: fine-tuning of sugar degradation? *Biochem. Soc. Trans.*, **39**, 77–81.

HANDELSMAN, J., RONDON, M.R., BRADY, S.F., CLARDY, J., GOODMAN, R.M. (1998) Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem. Biol.*, **5**, R245–9.

HANEY, P.J., BADGER, J.H., BULDAK, G.L., REICH, C.I., WOESE, C.R., OLSEN, G.J. (1999) Thermal adaptation analyzed by comparison of protein sequences from mesophilic and extremely thermophilic *Methanococcus* species. *Proc. Natl. Acad. Sci. U. S. A.*, **96**, 3578–83.

HANSEN, T., MUSFELDT, M., SCHÖNHEIT, P. (2002) ATP-dependent 6-phosphofructokinase from the hyperthermophilic bacterium *Thermotoga maritima*: characterization of an extremely thermophilic, allosterically regulated enzyme. *Arch. Microbiol.*, **177**, 401–409.

HARLEY, C.B., REYNOLDS, R.P. (1987) Analysis of *E. coli* promoter sequences. *Nucleic Acids Res.*, **15**, 2343–61.

HASAN, F., SHAH, A.A., HAMEED, A. (2006) Industrial applications of microbial lipases. *Enzyme Microb. Technol.*, **39**, 235–251.

HAUSMANN, S., JÄGER, K.E. (2010) Lipolytic Enzymes from Bacteria. In, Timmis,K.N. (ed), *Handbook of Hydrocarbon and Lipid Microbiology*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 1099–1126.

HICKEY, D.A., SINGER, G.A.C. (2004) Genomic and proteomic adaptations to growth at high temperature. *Genome Biol.*, **5**, 117.

HWANG, S., CHOI, K.-H., YOON, N., CHA, J. (2015) Improvement of a *Sulfolobus-E. coli* Shuttle Vector for Heterologous Gene Expression in *Sulfolobus acidocaldarius*. *J. Microbiol. Biotechnol.*, **25**, 196–205.

IMANAKA, T. (2011) Molecular bases of thermophily in hyperthermophiles. *Proc. Jpn. Acad. Ser. B. Phys. Biol. Sci.*, **87**, 587–602.

JAENICKE, R. (1998) What ultrastable globular proteins teach us about protein stabilization. *Biochem. Biokhimiiâ*, **63**, 312–21.

JÄGER, K.E., RANSAC, S., DIJKSTRA, B.W., COLSON, C., VAN HEUVEL, M., MISSET, O. (1994) Bacterial lipases. *FEMS Microbiol. Rev.*, **15**, 29–63.

JÄGER, K.E., REETZ, M.T. (1998) Microbial lipases form versatile tools for biotechnology. *Trends Biotechnol.*, **16**, 396–403.

JARRELL, K.F. (2012) Control of archaellation in *Sulfolobus acidocaldarius*: unravelling of the regulation of surface structure biosynthesis in Archaea begins. *Mol. Microbiol.*, **86**, 1–5.

JARRELL, K.F., WALTERS, A.D., BOCHIWAL, C., BORGIA, J.M., DICKINSON, T., CHONG, J.P.J. (2011) Major players on the microbial stage: why Archaea are important. *Microbiology*, **157**, 919–36.

JAUBERT, C., DANIOUX, C., OBERTO, J., CORTEZ, D., BIZE, A., KRUPOVIC, M., FORTERRE, P., PRANGISHVILI, D., SEZONOV, G., SHE, Q. (2013) Genomics and genetics of *Sulfolobus islandicus* LAL14/1, a model hyperthermophilic archaeon. *Open Biol.*

JOHN, J., CRENNELL, S.J., HOUGH, D.W., DANSON, M.J., TAYLOR, G.L. (1994) The crystal structure of glucose dehydrogenase from *Thermoplasma acidophilum*. *Structure*, **2**, 385–93.

JONUSCHEIT, M., MARTUSEWITSCH, E., STEDMAN, K.M., SCHLEPER, C. (2003) A reporter gene system for the hyperthermophilic archaeon *Sulfolobus solfataricus* based on a selectable and integrative shuttle vector. *Mol. Microbiol.*, **48**, 1241–1252.

KAKIRDE, K.S., WILD, J., GODISKA, R., MEAD, D.A., WIGGINS, A.G., GOODMAN, R.M., SZYBALSKI, W., LILES, M.R. (2011) Gram negative shuttle BAC vector for heterologous expression of metagenomic libraries. *Gene*, **475**, 57–62.

KARDINAHL, S., SCHMIDT, C.L., HANSEN, T., ANEMÜLLER, S., PETERSEN, A., SCHÄFER, G. (1999) The strict molybdate-dependence of glucose-degradation by the thermoacidophile *Sulfolobus acidocaldarius* reveals the first crenarchaeotic molybdenum containing enzyme--an aldehyde oxidoreductase. *Eur. J. Biochem.*, **260**, 540–8.

KAWARABAYASI, Y. *ET AL.* (2001) Complete genome sequence of an aerobic thermoacidophilic crenarchaeon, *Sulfolobus tokodaii* strain7. *DNA Res.*, **8**, 123–140.

KEELING, P.J., KLENK, H.P., SINGH, R.K., FEELEY, O., SCHLEPER, C., ZILLIG, W., DOOLITTLE, W.F., SENSEN, C.W. (1996) Complete Nucleotide Sequence of the *Sulfolobus islandicus* Multicopy Plasmid pRN1. *Plasmid*, **35**, 141–144.

KELLER, M.W., LIPSCOMB, G.L., LODER, A.J., SCHUT, G.J., KELLY, R.M., ADAMS, M.W.W. (2014) A Hybrid synthetic pathway for butanol production by a hyperthermophilic microbe. *Metab. Eng.*, **27**, 101–6.

KIM, S., LEE, S.B. (2006a) Characterization of *Sulfolobus solfataricus* 2-keto-3-deoxy-D-gluconate kinase in the modified Entner-Doudoroff pathway. *Biosci. Biotechnol. Biochem.*, **70**, 1308–16.

KIM, S., LEE, S.B. (2005) Identification and characterization of *Sulfolobus solfataricus* D-gluconate dehydratase: a key enzyme in the non-phosphorylated Entner-Doudoroff pathway. *Biochem. J.*, **387**, 271–80.

KIM, S., LEE, S.B. (2006b) Rare codon clusters at 5'-end influence heterologous expression of archaeal gene in *Escherichia coli*. *Protein Expr. Purif.*, **50**, 49–57.

KIM, S., LEE, S.B. (2004) Thermostable esterase from a thermoacidophilic archaeon: purification and characterization for enzymatic resolution of a chiral compound. *Biosci. Biotechnol. Biochem.*, **68**, 2289–98.

KNIGHT, R.D., FREELAND, S.J., LANDWEBER, L.F. (2001) A simple model based on mutation and selection explains trends in codon and amino-acid usage and GC composition within and across genomes. *Genome Biol.*, **2**.

KOERDT, A., ORELL, A., PHAM, T.K., MUKHERJEE, J., WLODKOWSKI, A., KARUNAKARAN, E., BIGGS, C. A, WRIGHT, P.C., ALBERS, S.-V. (2011) Macromolecular fingerprinting of *Sulfolobus* species in biofilm: a transcriptomic and proteomic approach combined with spectroscopic analysis. *J. Proteome Res.*, **10**, 4105–19.

KÖNIG, H., KANDLER, O., JENSEN, M., RIETSCHEL, E.T. (1983) The primary structure of the glycan moiety of pseudomurein from *Methanobacterium thermoautotrophicum*. *Hoppe. Seylers. Z. Physiol. Chem.*, **364**, 627–36.

KÖNIG, H., KRALIK, R., KANDLER, O. (1982) Structure and Modifications of Pseudomurein in *Methanobacteriales*. *Zentralblatt für Bakteriol. Mikrobiol. und Hyg. I. Abt. Orig. C Allg. Angew. und ökologische Mikrobiol.*, **3**, 191–179.

KOURIL, T., ESSER, D., KORT, J., WESTERHOFF, H. V, SIEBERS, B., SNOEP, J.L. (2013) Intermediate instability at high temperature leads to low pathway efficiency for an *in vitro* reconstituted system of gluconeogenesis in *Sulfolobus solfataricus*. *FEBS J.*, **280**, 4666–4680.

KOURIL, T., WIELOCH, P., REIMANN, J., WAGNER, M., ZAPARTY, M., ALBERS, S.-V., SCHOMBURG, D., RUOFF, P., SIEBERS, B. (2013) Unraveling the function of the two Entner-Doudoroff branches in the thermoacidophilic Crenarchaeon *Sulfolobus solfataricus* P2. *FEBS J.*, **280**, 1126–38.

LAEMMLI, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680–5.

LAMBLE, H., HEYER, N., BULL, S., HOUGH, D., DANSON, M. (2003) Metabolic pathway promiscuity in the archaeon *Sulfolobus solfataricus* revealed by studies on glucose dehydrogenase and 2-keto-3-deoxygluconate aldolase. *J. Biol. Chem.*, **278**, 34066–34072.

LAMBLE, H.J., MILBURN, C.C., TAYLOR, G.L., HOUGH, D.W., DANSON, M.J. (2004) Gluconate dehydratase from the promiscuous Entner-Doudoroff pathway in *Sulfolobus solfataricus*. *FEBS Lett.*, **576**, 133–6.

LAMBLE, H.J., THEODOSSIS, A., MILBURN, C.C., TAYLOR, G.L., BULL, S.D., HOUGH, D.W., DANSON, M.J. (2005) Promiscuity in the part-phosphorylative Entner-Doudoroff pathway of the archaeon *Sulfolobus solfataricus*. *FEBS Lett.*, **579**, 6865–9.

LASSAK, K., NEINER, T., GHOSH, A., KLINGL, A., WIRTH, R., ALBERS, S.-V. (2012) Molecular analysis of the crenarchaeal flagellum. *Mol. Microbiol.*, **83**, 110–124.

LEE, S.-J., SURMA, M., HAUSNER, W., THOMM, M., BOOS, W. (2008) The role of TrmB and TrmB-like transcriptional regulators for sugar transport and metabolism in the hyperthermophilic archaeon *Pyrococcus furiosus*. *Arch. Microbiol.*, **190**, 247–56.

LEE, S.-J., SURMA, M., SEITZ, S., HAUSNER, W., THOMM, M., BOOS, W. (2007) Characterization of the TrmB-like protein, PF0124, a TGM-recognizing global transcriptional regulator of the hyperthermophilic archaeon *Pyrococcus furiosus*. *Mol. Microbiol.*, **65**, 305–18.

LEIGH, J.A., ALBERS, S.-V., ATOMI, H., ALLERS, T. (2011) Model organisms for genetics in the domain Archaea: methanogens, halophiles, *Thermococcales* and *Sulfolobales*. *FEMS Microbiol. Rev.*, **35**, 577–608.

LEIS, B., ANGELOV, A., LI, H., LIEBL, W. (2014) Genetic analysis of lipolytic activities in *Thermus thermophilus* HB27. *J. Biotechnol.*, **191**, 150–157.

LEIS, B., ANGELOV, A., LIEBL, W. (2013) Screening and expression of genes from metagenomes. *Adv. Appl. Microbiol.*, **83**, 1–68.

LEVISSON, M., VAN DER OOST, J., KENGEN, S. (2009) Carboxylic ester hydrolases from hyperthermophiles. *Extremophiles*, **13**, 567–581.

LI, E., REICH, C.I., OLSEN, G.J. (2008) A whole-genome approach to identifying protein binding sites: promoters in *Methanocaldococcus* (Methanococcus) *jannaschii*. *Nucleic Acids Res.*, **36**, 6948–58.

LIEBL, W., ANGELOV, A., JUERGENSEN, J., CHOW, J., LOESCHCKE, A., DREPPER, T., CLASSEN, T., PIETRUZSKA, J., EHRENREICH, A., STREIT, W.R., JÄGER, K.E. (2014) Alternative hosts for functional (meta)genome analysis. *Appl. Microbiol. Biotechnol.*, **98**, 8099–109.

LIPPS, G., IBANEZ, P., STROESSENREUTHER, T., HEKIMIAN, K., KRAUSS, G. (2001) The protein ORF80 from the acidophilic and thermophilic archaeon *Sulfolobus islandicus* binds highly site-specifically to double-stranded DNA and represents a novel type of basic leucine zipper protein. *Nucleic Acids Res.*, **29**, 4973–82.

LIPPS, G., STEGERT, M., KRAUSS, G. (2001) Thermostable and site-specific DNA binding of the gene product ORF56 from the *Sulfolobus islandicus* plasmid pRN1, a putative archaeal plasmid copy control protein. *Nucleic Acids Res.*, **29**, 904–13.

LIU, H., ORELL, A., MAES, D., WOLFEREN, M. VAN, LINDÅS, A.-C., BERNANDER, R., ALBERS, S.-V., CHARLIER, D., PEETERS, E. (2014) BarR, an Lrp-type transcription factor in *Sulfolobus acidocaldarius*, regulates an aminotransferase gene in a β-alanine responsive manner. *Mol. Microbiol.*, **92**, 625–639.

LORENTZEN, E., HENSEL, R., KNURA, T., AHMED, H., POHL, E. (2004) Structural Basis of allosteric regulation and substrate specificity of the non-phosphorylating glyceraldehyde 3-phosphate dehydrogenase from *Thermoproteus tenax*. *J. Mol. Biol.*, **341**, 815–28.

MALACH, A. (2013) *Rhodobacter capsulatus* als Wirtsstamm zum aktivitätsbasierten Screening von Enzymen und zur Synthese von Membranproteinen. *PhD Thesis*, 1–165.

MANDRICH, L., PEZZULLO, M., ROSSI, M., MANCO, G. (2007) SSoNDelta and SSoNDeltalong: two thermostable esterases from the same ORF in the archaeon *Sulfolobus solfataricus*? *Archaea*, **2**, 109–115.

MATARAZZO, F., RIBEIRO, A.C., FAVERI, M., TADDEI, C., MARTINEZ, M.B., MAYER, M.P.A. (2012) The domain Archaea in human mucosal surfaces. *Clin. Microbiol. Infect.*, **18**, 834–40.

MERGULHÃO, F.J.M., SUMMERS, D.K., MONTEIRO, G.A. (2005) Recombinant protein secretion in *Escherichia coli*. *Biotechnol. Adv.*, **23**, 177–202.

MEYER, B.H., PEYFOON, E., DIETRICH, C., HITCHEN, P., PANICO, M., MORRIS, H.R., DELL, A., ALBERS, S.-V. (2013) Agl16, a thermophilic glycosyltransferase mediating the last step of N-Glycan biosynthesis in the thermoacidophilic crenarchaeon *Sulfolobus acidocaldarius*. *J. Bacteriol.*, **195**, 2177–86.

MILBURN, C.C., LAMBLE, H.J., THEODOSSIS, A., BULL, S.D., HOUGH, D.W., DANSON, M.J., TAYLOR, G.L. (2006) The structural basis of substrate promiscuity in glucose dehydrogenase from the hyperthermophilic archaeon *Sulfolobus solfataricus*. *J. Biol. Chem.*, **281**, 14796–804.

MORRIS, B.E.L., HENNEBERGER, R., HUBER, H., MOISSL-EICHINGER, C. (2013) Microbial syntrophy: interaction for the common good. *FEMS Microbiol. Rev.*, **37**, 384–406.

MUELLER, M., TAKEMASA, R., SCHWARZ, A., ATOMI, H., NIDETZKY, B. (2009) 'Short-chain' alpha-1,4-glucan phosphorylase having a truncated N-terminal domain: functional expression and characterization of the enzyme from *Sulfolobus solfataricus*. *Biochim. Biophys. Acta*, **1794**, 1709–14.

NUNN, C.E.M., JOHNSEN, U., SCHÖNHEIT, P., FUHRER, T., SAUER, U., HOUGH, D.W., DANSON, M.J. (2010) Metabolism of pentose sugars in the hyperthermophilic Archaea *Sulfolobus solfataricus* and *Sulfolobus acidocaldarius*. *J. Biol. Chem.*, **285**, 33701–33709.

NUNOURA, T., TAKAKI, Y., KAKUTA, J., NISHI, S., SUGAHARA, J., KAZAMA, H., CHEE, G.-J., HATTORI, M., KANAI, A., ATOMI, H., TAKAI, K., TAKAMI, H. (2011) Insights into the evolution of Archaea and eukaryotic protein modifier systems revealed by the genome of a novel archaeal group. *Nucleic Acids Res.*, **39**, 3204–3223.

OHSHIMA, T., ITO, Y., SAKURABA, H., GODA, S., KAWARABAYASI, Y. (2003) The *Sulfolobus tokodaii* gene ST1704 codes highly thermostable glucose dehydrogenase. *J. Mol. Catal. B Enzym.*, **23**, 281–289.

OKSANEN, T., PERE, J., PAAVILAINEN, L., BUCHERT, J., VIIKARI, L. (2000) Treatment of recycled kraft pulps with *Trichoderma reesei* hemicellulases and cellulases. *J. Biotechnol.*, **78**, 39–48.

VAN DER OOST, J., SIEBERS, B. (2006) The Glycolytic pathways of Archaea: Evolution By Tinkering. In, Garrett,R.A. and Klenk,H.-P. (eds), *Archaea: Evolution, Physiology, and Molecular Biology*. Blackwell Publishing Ltd, Malden, MA, USA, pp. 247–259.

ORELL, A., PEETERS, E., VASSEN, V., JACHLEWSKI, S., SCHALLES, S., SIEBERS, B., ALBERS, S.-V. (2013) Lrs14 transcriptional regulators influence biofilm formation and cell motility of Crenarchaea. *ISME J.*, **7**, 1886–98.

OSHIMA, T., MORIYA, T., TERUI, Y. (2011) Identification, chemical synthesis, and biological functions of unusual polyamines produced by extreme thermophiles. *Methods Mol. Biol.*, **720**, 81–111.

PACE, C.N. (1992) Contribution of the hydrophobic effect to globular protein stability. *J. Mol. Biol.*, **226**, 29–35.

PALM, G.J., FERNÁNDEZ-ÁLVARO, E., BOGDANOVIĆ, X., BARTSCH, S., SCZODROK, J., SINGH, R.K., BÖTTCHER, D., ATOMI, H., BORNSCHEUER, U.T., HINRICHS, W. (2011) The crystal structure of an esterase from the hyperthermophilic microorganism *Pyrobaculum calidifontis* VA1 explains its enantioselectivity. *Appl. Microbiol. Biotechnol.*, **91**, 1061–72.

PARK, S.-J., CHAE, J.-C., RHEE, S.-K. (2010) Application of DNA microarray for screening metagenome library clones. *Methods Mol. Biol.*, **668**, 313–24.

PARK, Y.-J., CHOI, S.Y., LEE, H.-B. (2006) A carboxylesterase from the thermoacidophilic archaeon *Sulfolobus solfataricus* P1: purification, characterization, and expression. *Biochim. Biophys. Acta*, **1760**, 820–8.

PARK, Y.-J., YOON, S.-J., LEE, H.-B. (2008) A novel thermostable arylesterase from the archaeon *Sulfolobus solfataricus* P1: purification, characterization, and expression. *J. Bacteriol.*, **190**, 8086–95.

PAULY, T.A., EKSTROM, J.L., BEEBE, D.A., CHRUNYK, B., CUNNINGHAM, D., GRIFFOR, M., KAMATH, A., LEE, S.E., MADURA, R., MCGUIRE, D., SUBASHI, T., WASILKO, D., WATTS, P., MYLARI, B.L., OATES, P.J., ADAMS, P.D., RATH, V.L. (2003) X-ray crystallographic and kinetic studies of human sorbitol dehydrogenase. *Structure*, **11**, 1071–85.

PEETERS, E., ALBERS, S.-V., VASSART, A., DRIESSEN, A.J.M., CHARLIER, D. (2009) Ss-LrpB, a transcriptional regulator from *Sulfolobus solfataricus*, regulates a gene cluster with a pyruvate ferredoxin oxidoreductase-encoding operon and permease genes. *Mol. Microbiol.*, **71**, 972–88.

PENG, N., DENG, L., MEI, Y., JIANG, D., HU, Y., AWAYEZ, M., LIANG, Y., SHE, Q. (2012) A Synthetic Arabinose-Inducible Promoter Confers High Levels of Recombinant Protein Expression in Hyperthermophilic archaeon *Sulfolobus islandicus*. *Appl. Environ. Microbiol.*, **78**, 5630–5637.

PENG, N., XIA, Q., CHEN, Z., LIANG, Y.X., SHE, Q. (2009) An upstream activation element exerting differential transcriptional activation on an archaeal promoter. *Mol. Microbiol.*, **74**, 928–39.

PERSSON, B., ZIGLER, J.S., JÖRNVALL, H. (1994) A super-family of medium-chain dehydrogenases/reductases (MDR). *Eur. J. Biochem.*, **226**, 15–22.

PESTER, M., SCHLEPER, C., WAGNER, M. (2011) The Thaumarchaeota: an emerging view of their phylogeny and ecophysiology. *Curr. Opin. Microbiol.*, **14**, 300–6.

PICKL, A., SCHÖNHEIT, P. (2015) The oxidative pentose phosphate pathway in the haloarchaeon *Haloferax volcanii* involves a novel type of glucose-6-phosphate dehydrogenase - The archaeal Zwischenferment. *FEBS Lett.*, **589**, 1105–11.

PIRE, C., ESCLAPEZ, J., FERRER, J., BONETE, M.J. (2001) Heterologous overexpression of glucose dehydrogenase from the halophilic archaeon *Haloferax mediterranei*, an enzyme of the medium chain dehydrogenase/reductase family. *FEMS Microbiol. Lett.*, **200**, 221–7.

PODAR, M., MAKAROVA, K.S., GRAHAM, D.E., WOLF, Y.I., KOONIN, E. V, REYSENBACH, A.-L. (2013) Insights into archaeal evolution and symbiosis from the genomes of a nanoarchaeon and its inferred crenarchaeal host from Obsidian Pool, Yellowstone National Park. *Biol. Direct*, **8**, 9.

RADIANINGTYAS, H., WRIGHT, P.C. (2003) Alcohol dehydrogenases from thermophilic and hyperthermophilic Archaea and Bacteria. *FEMS Microbiol. Rev.*, **27**, 593–616.

RAO, L., XUE, Y., ZHENG, Y., LU, J.R., MA, Y. (2013) A novel alkaliphilic bacillus esterase belongs to the 13(th) bacterial lipolytic enzyme family. *PLoS One*, **8**.

RAUCH, B. (2013) Functional analysis of multiple general transcription factors in *Sulfolobus acidocaldarius*. *PhD Thesis*.

REHER, M., GEBHARD, S., SCHÖNHEIT, P. (2007) 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 aero*. *FEMS Microbiol. Lett.*, **273**, 196–205.

REIMANN, J., LASSAK, K., KHADOUMA, S., ETTEMA, T.J.G., YANG, N., DRIESSEN, A.J.M., KLINGL, A., ALBERS, S.-V. (2012) Regulation of archaella expression by the FHA and von Willebrand domain-containing proteins ArnA and ArnB in *Sulfolobus acidocaldarius*. *Mol. Microbiol.*, **86**, 24–36.

REINDL, S., GHOSH, A., WILLIAMS, G.J., LASSAK, K., NEINER, T., HENCHE, A.-L., ALBERS, S.-V., TAINER, J.A. (2013) Insights into FlaI functions in archaeal motor assembly and motility from structures, conformations, and genetics. *Mol. Cell*, **49**, 1069–82.

RENO, M.L., HELD, N.L., FIELDS, C.J., BURKE, P. V, WHITAKER, R.J. (2009) Biogeography of the *Sulfolobus islandicus* pan-genome. *Proc. Natl. Acad. Sci. U. S. A.*, **106**, 8605–8610.

RIVEROS-ROSAS, H., JULIÁN-SÁNCHEZ, A., VILLALOBOS-MOLINA, R., PARDO, J.P., PIÑA, E. (2003) Diversity, taxonomy and evolution of medium-chain dehydrogenase/reductase superfamily. *Eur. J. Biochem.*, **270**, 3309–34.

RONDON, M.R., AUGUST, P.R., BETTERMANN, A.D., BRADY, S.F., GROSSMAN, T.H., LILES, M.R., LOIACONO, K.A., LYNCH, B.A., MACNEIL, I.A., MINOR, C., TIONG, C.L., GILMAN, M., OSBURNE, M.S., CLARDY, J., HANDELSMAN, J., GOODMAN, R.M. (2000) Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. *Appl. Environ. Microbiol.*, **66**, 2541–7.

SAIKI, R.K., GELFAND, D.H., STOFFEL, S., SCHARF, S.J., HIGUCHI, R., HORN, G.T., MULLIS, K.B., ERLICH, H.A. (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*, **239**, 487–91.

SALAMEH, M., WIEGEL, J. (2007) Lipases from extremophiles and potential for industrial applications. *Adv. Appl. Microbiol.*, **61**, 253–83.

SAMBROOK, J., RUSSELL, D.W. (2001) Molecular Cloning: A Laboratory Manual 3rd ed. Cold Spring Harbor Laboratory.

SANTANGELO, T.J., CUBONOVÁ, L., REEVE, J.N. (2010) *Thermococcus kodakarensis* genetics: TK1827-encoded beta-glycosidase, new positive-selection protocol, and targeted and repetitive deletion technology. *Appl. Environ. Microbiol.*, **76**, 1044–52.

SATO, T., ATOMI, H. (2011) Novel metabolic pathways in Archaea. *Curr. Opin. Microbiol.*, **14**, 307–314.

SAY, R.F., FUCHS, G. (2010) Fructose 1,6-bisphosphate aldolase/phosphatase may be an ancestral gluconeogenic enzyme. *Nature*, **464**, 1077–81.

SCHIRALDI, C., GIULIANO, M., DE ROSA, M. (2002) Perspectives on biotechnological applications of Archaea. *Archaea*, **1**, 75–86.

SCHLEPER, C., NICOL, G.W. (2010) Ammonia-oxidising Archaea--physiology, ecology and evolution. *Adv. Microb. Physiol.*, **57**, 1–41.

SCHLEPER, C., PUEHLER, G., HOLZ, I., GAMBACORTA, A., JANEKOVIC, D., SANTARIUS, U., KLENK, H.P., ZILLIG, W. (1995) *Picrophilus* gen. nov., fam. nov.: a novel aerobic, heterotrophic, thermoacidophilic genus and family comprising Archaea capable of growth around pH 0. *J. Bacteriol.*, **177**, 7050–7059.

SCHREIER, H.J., ROBINSON-BIDLE, K.A., ROMASHKO, A.M., PATEL, G. V (1999) Heterologous expression in the Archaea: transcription from *Pyrococcus furiosus* gdh and mlrA promoters in *Haloferax volcanii*. *Extremophiles*, **3**, 11–19.

SCHÜLEIN, M. (2000) Protein engineering of cellulases. *Biochim. Biophys. Acta*, **1543**, 239–252.

SCHUT, G.J., BREHM, S.D., DATTA, S., ADAMS, M.W.W. (2003) Whole-genome DNA microarray analysis of a hyperthermophile and an archaeon: *Pyrococcus furiosus* grown on carbohydrates or peptides. *J. Bacteriol.*, **185**, 3935–47.

SHANG, Y.-S., ZHANG, X.-E., WANG, X.-D., GUO, Y.-C., ZHANG, Z.-P., ZHOU, Y.-F. (2010) Biochemical characterization and mutational improvement of a thermophilic esterase from *Sulfolobus solfataricus* P2. *Biotechnol. Lett.*, **32**, 1151–1157.

SHE, Q. *ET AL.* (2001) The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. *Proc. Natl. Acad. Sci. U. S. A.*, **98**, 7835–7840.

SHE, Q., ZHANG, C., DENG, L., PENG, N., CHEN, Z., LIANG, Y.X. (2009) Genetic analyses in the hyperthermophilic archaeon *Sulfolobus islandicus*. *Biochem. Soc. Trans.*, **37**, 92–6.

SIEBERS, B., BRINKMANN, H., DÖRR, C., TJADEN, B., LILIE, H., VAN DER OOST, J., VERHEES, C.H. (2001) Archaeal fructose-1,6-bisphosphate aldolases constitute a new family of archaeal type class I aldolase. *J. Biol. Chem.*, **276**, 28710–28718.

SIEBERS, B., KLENK, H.P., HENSEL, R. (1998) PPi-dependent phosphofructokinase from *Thermoproteus tenax*, an archaeal descendant of an ancient line in phosphofructokinase evolution. *J. Bacteriol.*, **180**, 2137–2143.

SIEBERS, B., SCHÖNHEIT, P. (2005) Unusual pathways and enzymes of central carbohydrate metabolism in Archaea. *Curr. Opin. Microbiol.*, **8**, 695–705.

SIEBERS, B., WENDISCH, V.F., HENSEL, R. (1997) Carbohydrate metabolism in *Thermoproteus tenax*: *in vivo* utilization of the non-phosphorylative Entner-Doudoroff pathway and characterization of its first enzyme, glucose dehydrogenase. *Arch. Microbiol.*, **168**, 120–7.

SIMON, C., DANIEL, R. (2011) Metagenomic analyses: past and future trends. *Appl. Environ. Microbiol.*, **77**, 1153–1161.

VAN DER SLUIS, E.O., NOUWEN, N., KOCH, J., DE KEYZER, J., VAN DER DOES, C., TAMPÉ, R., DRIESSEN, A.J.M. (2006) Identification of two interaction sites in SecY that are important for the functional interaction with SecA. *J. Mol. Biol.*, **361**, 839–49.

SMITH, L.D., BUDGEN, N., BUNGARD, S.J., DANSON, M.J., HOUGH, D.W. (1989) Purification and characterization of glucose dehydrogenase from the thermoacidophilic archaebacterium *Thermoplasma acidophilum*. *Biochem. J.*, **261**, 973–7.

SOBEK, H., GÖRISCH, H. (1989) Further kinetic and molecular characterization of an extremely heat-stable carboxylesterase from the thermoacidophilic archaebacterium *Sulfolobus acidocaldarius*. *Biochem. J.*, **261**, 993–998.

SOBEK, H., GÖRISCH, H. (1988) Purification and characterization of a heat-stable esterase from the thermoacidophilic archaebacterium Sulfolobus acidocaldarius. *Biochem. J.*, **250**, 453–458.

SOPPA, J. (1999) Normalized nucleotide frequencies allow the definition of archaeal promoter elements for different archaeal groups and reveal base-specific TFB contacts upstream of the TATA box. *Mol. Microbiol.*, **31**, 1589–1592.

SØRENSEN, H.P., MORTENSEN, K.K. (2005) Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. *J. Biotechnol.*, **115**, 113–28.

STEELE, H.L., JÄGER, K.E., DANIEL, R., STREIT, W.R. (2009) Advances in recovery of novel biocatalysts from metagenomes. *J. Mol. Microbiol. Biotechnol.*, **16**, 25–37.

STETTER, K.O. (1999) Extremophiles and their adaptation to hot environments. *FEBS Lett.*, **452**, 22–5.

STETTER, K.O. (1996) Hyperthermophiles in the history of life. *Ciba Found. Symp.*, **202**, 1–10; discussion 11–8.

STREIT, W.R., SCHMITZ, R.A. (2004) Metagenomics--the key to the uncultured microbes. *Curr. Opin. Microbiol.*, **7**, 492–8.

THOMM, M., SANDMAN, K., FREY, G., KOLLER, G., REEVE, J.N. (1992) Transcription *in vivo* and *in vitro* of the histone-encoding gene hmfB from the hyperthermophilic archaeon *Methanothermus fervidus*. *J. Bacteriol.*, **174**, 3508–3513.

TORARINSSON, E., KLENK, H.-P., GARRETT, R.A. (2005) Divergent transcriptional and translational signals in Archaea. *Environ. Microbiol.*, **7**, 47–54.

TUININGA, J., VERHEES, C., VAN DER OOST, J., KENGEN, S., STAMS, A., DE VOS, W. (1999) Molecular and Biochemical Characterization of the ADP-dependent Phosphofructokinase from the Hyperthermophilic Archaeon *Pyrococcus furiosus*. *J. Biol. Chem.*, **274**, 21023–21028.

UNSWORTH, L.D., VAN DER OOST, J., KOUTSOPOULOS, S. (2007) Hyperthermophilic enzymes--stability, activity and implementation strategies for high temperature applications. *FEBS J.*, **274**, 4044–4056.

VARALJAY, V.A., HOWARD, E.C., SUN, S., MORAN, M.A. (2010) Deep sequencing of a dimethylsulfoniopropionate-degrading gene (dmdA) by using PCR primer pairs designed on the basis of marine metagenomic data. *Appl. Environ. Microbiol.*, **76**, 609–17.

VENDITTIS, E. DE, BOCCHINI, V. (1996) Protein-encoding genes in the sulfothermophilic Archaea *Sulfolobus* and *Pryococcus*. *Gene*, **176**, 27–33.

VERHEES, C., TUININGA, J., KENGEN, S., STAMS, A., VAN DER OOST, J., DE VOS, W. (2001) ADP-Dependent Phosphofructokinases in Mesophilic and Thermophilic Methanogenic Archaea. *J. Bacteriol.*, **183**, 7145–7153.

VIEILLE, C., BURDETTE, D.S., ZEIKUS, J.G. (1996) Thermozymes. *Biotechnol. Annu. Rev.*, **2**, 1–83.

VIEILLE, C., ZEIKUS, G.J. (2001) Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. *Microbiol. Mol. Biol. Rev.*, **65**, 1–43.

VILLEGAS, A., KROPINSKI, A.M. (2008) An analysis of initiation codon utilization in the Domain Bacteria - concerns about the quality of bacterial genome annotation. *Microbiology*, **154**, 2559–661.

VAN DE VOSSENBERG, J.L., UBBINK-KOK, T., ELFERINK, M.G., DRIESSEN, A.J., KONINGS, W.N. (1995) Ion permeability of the cytoplasmic membrane limits the maximum growth temperature of Bacteria and Archaea. *Mol. Microbiol.*, **18**, 925–32.

WAGNER, M., BERKNER, S., AJON, M., DRIESSEN, A.J.M., LIPPS, G., ALBERS, S.-V. (2009) Expanding and understanding the genetic toolbox of the hyperthermophilic genus *Sulfolobus*. *Biochem. Soc. Trans.*, **37**, 97–101.

WAGNER, M., WAGNER, A., MA, X., KORT, J.C., GHOSH, A., RAUCH, B., SIEBERS, B., ALBERS, S.-V. (2014) Investigation of the *malE* Promoter and MalR, a Positive Regulator of the Maltose Regulon, for an Improved Expression System in *Sulfolobus acidocaldarius*. *Appl. Environ. Microbiol.*, **80**, 1072–1081.

WAGNER, M., VAN WOLFEREN, M., WAGNER, A., LASSAK, K., MEYER, B.H., REIMANN, J., ALBERS, S.-V. (2012) Versatile Genetic Tool Box for the Crenarchaeote *Sulfolobus acidocaldarius*. *Front. Microbiol.*, **3**.

WANG, H.-C., XIA, X., HICKEY, D. (2006) Thermal adaptation of the small subunit ribosomal RNA gene: a comparative study. *J. Mol. Evol.*, **63**, 120–6.

WANG, J., WANG, J., GONG, X., ZHENG, G. (2010) A novel esterase Sso2518 from *Sulfolobus solfataricus* with a much lower temperature optimum than the growth temperature. *Biotechnol. Lett.*, **32**, 1103–1108.

WARREN, R.L., FREEMAN, J.D., LEVESQUE, R.C., SMAILUS, D.E., FLIBOTTE, S., HOLT, R.A. (2008) Transcription of foreign DNA in *Escherichia coli*. *Genome Res.*, **18**, 1798–805.

WEINBERG, M. V, SCHUT, G.J., BREHM, S., DATTA, S., ADAMS, M.W.W. (2005) Cold shock of a hyperthermophilic archaeon: *Pyrococcus furiosus* exhibits multiple responses to a suboptimal growth temperature with a key role for membrane-bound glycoproteins. *J. Bacteriol.*, **187**, 336–348.

VAN DE WERKEN, H.J.G., VERHEES, C.H., AKERBOOM, J., DE VOS, W.M., VAN DER OOST, J. (2006) Identification of a glycolytic regulon in the Archaea *Pyrococcus* and *Thermococcus*. *FEMS Microbiol. Lett.*, **260**, 69–76.

WOESE, C.R., FOX, G.E. (1977) Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc. Natl. Acad. Sci. U. S. A.*, **74**, 5088–90.

WOESE, C.R., KANDLER, O., WHEELIS, M.L. (1990) Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natl. Acad. Sci. U. S. A.*, **87**, 4576–9.

WORTHINGTON, P., HOANG, V., PEREZ-POMARES, F., BLUM, P. (2003) Targeted disruption of the alpha-amylase gene in the hyperthermophilic archaeon *Sulfolobus solfataricus*. *J. Bacteriol.*, **185**, 482–488.

YING, X., MA, K. (2011) Characterization of a zinc-containing alcohol dehydrogenase with stereoselectivity from the hyperthermophilic archaeon *Thermococcus guaymasensis*. *J. Bacteriol.*, **193**, 3009–19.

ZAPARTY, M., SIEBERS, B. (2011) Reconstruction of the central carbon metabolic network of thermoacidophilic Archaea. In, Horikoshi,K. (ed), *Extremophiles Handbook*. Springer Japan, Tokyo, pp. 601–639.

ZAPARTY, M., ZAIGLER, A., STAMME, C., SOPPA, J., HENSEL, R., SIEBERS, B. (2008) DNA microarray analysis of central carbohydrate metabolism: glycolytic/gluconeogenic carbon switch in the hyperthermophilic crenarchaeum *Thermoproteus tenax*. *J. Bacteriol.*, **190**, 2231–8.

ZHANG, C., WHITAKER, R.J. (2012) A broadly applicable gene knockout system for the thermoacidophilic archaeon *Sulfolobus islandicus* based on simvastatin selection. *Microbiology*, **158**, 1513–22.

ZHU, X., LARSEN, N.A., BASRAN, A., BRUCE, N.C., WILSON, I.A. (2003) Observation of an arsenic adduct in an acetyl esterase crystal structure. *J. Biol. Chem.*, **278**, 2008–14.

ZILLIG, W., STETTER, K.O., SCHÄFER, W., JANEKOVIC, D., WUNDERL, S., HOLZ, I., PALM, P. (1981) *Thermoproteales*: A novel type of extremely thermoacidophilic anaerobic archaebacteria isolated from Icelandic solfataras. *Zentralblatt für Bakteriol. Mikrobiol. und Hyg. I. Abt. Orig. C Allg. Angew. und ökologische Mikrobiol.*, **2**, 205–227.

ZILLIG, W., STETTER, K.O., WUNDERL, S., SCHULZ, W., PRIESS, H., SCHOLZ, I. (1980) The *Sulfolobus*-'Caldariella' group: Taxonomy on the basis of the structure of DNA-dependent RNA polymerases. *Arch. Microbiol.*, **125**, 259–269.

APPENDIX

APPENDIX

LIST OF ABBREVIATIONS

aa	amino acid
A. bidest	aqua bidestillata = two times distilled water
ABCE1	ATP binding cassette (ABC) protein with two $[4Fe\text{-}4S]^{2+}$ clusters
ADP	adenosine diphosphate
AMP	adenosine monophosphate
amp	ampicillin
APS	ammonium persulfate
ATP	adenosine triphosphate
BLAST	*Basic Local Alignment Search Tool*
BRE	transcription factor B recognition element
BSA	bovine serum albumin
cam	chloramphenicol
CCM	central carbohydrate metabolism
cDNA	complementary DNA
CE	cell-free extract fraction
DGGR	1,2-*O*-dilauryl-*rac*-glycero-3-glutaric acid-resorufin ester
DHAP	dihydroxyacetonephosphate
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxy-nucleotide triphosphate
DTT	dithiothreitol
e.g.	for example
ED	Entner-Doudoroff pathway
EDTA	ethylene diamine tetraacetic acid
EMP	Embden-Meyerhof-Parnas pathway
ENO	enolase
et al.	et alii = and the others
FBP	fructose 1,6-bisphosphate
FBPA	fructose -1,6-bisphosphate aldolase
FBPase	fructose-1,6-bisphosphate phosphatase
FBPA/ase	fructose-1,6-bisphosphate aldolase/ phosphatase
F6P	fructose 6-phosphate
GA	glyceraldehyde
GAOR	glyceraldehyde oxidoreductase
GAD	gluconate dehydratase
GADH	glyceraldehyde dehydrogenase
GAP	glyceraldehyde 3-phosphate

GAPDH	glyceraldehyde-3-phosphate dehydrogenase
GAPN	non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase
GAPOR	glyceraldehyde 3-phosphate oxidoreductase
GDH	glucose dehydrogenase
GER	Germany
GF	gel filtration
GK	glycerate kinase
HFX	*Haloferax mediterranei*
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HIC	hydrophobic interaction chromatography
HP	heat precipitation fraction
IC	ion exchange chromatography
i.e.	id est = that is
IPTG	isopropyl-ß-D-thiogalactopyranoside
kan	kanamycin
KDGK	2-keto-3-deoxy gluconate kinase
KD(P)G	2-keto-3-deoxy(-6-phospho)gluconate
KD(P)GA	2-keto-3-deoxy(-6-phospho)gluconate aldolase
K_m	Michaelis constant
LB	Luria-Bertani
LDH	lactate dehydrogenase
MalE	maltose-binding protein
MDR	medium-chain alcohol/polyol dehydrogenase/reductase
MPI	Max Planck Institute
mRNA	messenger RNA
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 of the USA
np	non-phosphorylating
OD	optical density
ORF	open reading frame
2-PG	2-phosphoglycerate
3-PG	3-phosphoglycerate
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PDB	protein Data Bank
pDNA	plasmid DNA
PEP	phosphoenolpyruvate
PEPS	phosphoenolpyruvate synthase

PIPES	piperazine-N,N'-bis(2-ethanesulfonic acid)
PFK	6-phosphofructokinase
PF, *Pfu*	*Pyrococcus furiosus*
PGAM	phosphoglycerate mutase
PGK	phosphoglycerate kinase
PK	pyruvate kinase
ρ-NP	para-nitrophenyl
PP_i	inorganic pyrophosphate
POR	pyruvate oxidoreductase
PTO	*Picrophilus torridus*
PVAP	protein forming virus-associated pyramids
RNAP	RNA polymerase
rRNA	ribosomal ribonucleic acid
RT	room temperature
SDS	sodium dodecyl sulfate
SIRV2	*Sulfolobus islandicus* rod-shaped virus 2
sp	semi-phosphorylating
SOR	sulfur oxygenase reductase
Ta	*Thermoplasma acidophilum*
TAE	Tris-acetate-EDTA-buffer
TATA	TATA box
TEMED	N,N,N',N'-Tetramethylethylenediamine
TIM	Triosephosphate isomerase
TGM	Thermococcales Glycolytic Motif
TM, *Tma*	*Thermotoga maritima*
Tris	Tris-(hydroxymethyl)-aminomethane
TSS	transcription start site
TTX, *Ttx*	*Thermoproteus tenax*
U	international unit of enzyme activity (1 µmol substrate/min)
UV	ultraviolet
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
% (v/v)	percent by volume
% (w/v)	percent by weight

LIST OF PUBLICATIONS

Range S., Hagmeyer D., Rotan O., Sokolova V., **Verheyen J.**, Siebers B., Epple M.

A continuous method to prepare a poorly crystalline silver-doped calcium phosphate ceramics with antibacterial properties. RSC Advances. Submitted 2015 Jan

Wagner M., Wagner A., Ma X., **Kort J.C.**, Ghosh A., Rauch B., Siebers B., Albers S.V.

Investigation of the *malE* promoter and MalR, a positive regulator of the maltose regulon, for an improved expression system in *Sulfolobus acidocaldarius*. Appl Environ Microbiol. 2014 Feb; 80(3): 1072-1081

Kort J.C., Esser D., Pham T.K., Noirel J., Wright P.C., Siebers B.

A cool tool for hot and sour Archaea: proteomics of *Sulfolobus solfataricus*. Proteomics. 2013 Oct; 13(18-19): 2831-2850

Kouril T., Esser D., **Kort J.**, Westerhoff H.V., Siebers B., Snoep J.L.

Intermediate instability at high temperature leads to low pathway efficiency for an in vitro reconstituted system of gluconeogenesis in *Sulfolobus solfataricus*. FEBS J. 2013 Sep; 280(18):4666-4680

Dazert E., Neumann-Haefelin C., Bressanelli S., Fitzmaurice K., **Kort J.**, Timm J., McKiernan S., Kelleher D., Gruener N., Tavis J.E., Rosen H.R., Shaw J., Bowness P., Blum H.E., Klenerman P., Bartenschlager R., Thimme R.

Loss of viral fitness and cross-recognition by CD8+ T cells limit HCV escape from a protective HLA-B27-restricted human immune response. J Clin Invest. 2009 Feb; 119(2):376-386

POSTER PRESENTATIONS

09/2012	Extremophiles (Sevilla, Spain)
	Towards An Improved *S. acidocaldarius* Expression System Useful For Functional Metagenome Analysis
	J. Kort, A. Wagner, S.-V. Albers, B. Siebers
07/2012	MBA3 Marburg
	Hot Expression – Towards An Archaeal Expression Host for Functional Metagenomics
	J. Kort, A. Wagner, S.-V. Albers, B. Siebers
03/2012	VAAM Tübingen
	Hot Metagenomics – Towards an archaeal expression host for metagenome analysis
	J. Kort, A. Wagner, S.-V. Albers, B. Siebers
09/2011	ProkaGenomics Göttingen
	Towards an archaeal expression host for metagenome analysis
	A. Wlodkowski, **J. Kort**, M. Wagner, B. Siebers, S.-V. Albers

09/2011 European Society for Biomaterials 2011 (ESB2011) Dublin
Silver ion-doped calcium phosphate nanoparticles for antibacterial application
A. Peetsch, **J. Kort**, C. Greulich, B. Siebers, M. Köller and M. Epple

04/2011 VAAM Karlsruhe
Towards an archaeal expression host for metagenome analysis
J. Kort, A. Wlodkowski, S.-V. Albers, B. Siebers

03/2011 Innovationsnachmittag „Wirtschaft trifft Wissenschaft“, UDE
Herstellung und antibakterielle Wirkung von Silber-dotiertem Hydroxylapatit
D. Hagmeyer, O. Rotan, **J. Kort**, B. Siebers, M. Epple

ORAL PRESENTATIONS

05/2012 Verbundtreffen (Jülich)
A Novel Archaeal Expression Host: the thermoacidophilic Crenarchaeon *Sulfolobus*

09/2011 Verbundtreffen (ProkaGenomics Göttingen)
A Novel Archaeal Expression Host: the thermoacidophilic Crenarchaeon *Sulfolobus*

LEBENSLAUF

Der Lebenslauf ist in der Online-Version aus Gründen des Datenschutzes nicht enthalten.

ERKLÄRUNG

Hiermit versichere ich, dass ich die vorliegende Arbeit mit dem Titel

***Sulfolobus acidocaldarius* & *Sulfolobus. solfataricus*:**
Exploitation of thermoacidophilic Archaea for biotechnological applications

selbst verfasst und keine außer den angegebenen Hilfsmitteln und Quellen benutzt habe, und dass die Arbeit in dieser oder ähnlicher Form noch bei keiner anderen Universität eingereicht wurde.

Essen, 27.04.2015

Julia Christin Verheyen

ACKNOWLEDGEMENTS

I would like to deeply thank my supervisor Prof. Dr. Bettina Siebers for her support throughout this study, for constructive advice and her trust over the years. Her understanding and patience, especially during my parental leave, is greatly acknowledged.

I greatly thank Prof. Dr. Peter Bayer for supporting me in switching universities and being my co-referee of this thesis.

I wish to warmly thank Dr. Christopher Bräsen for his fast, but patient and detailed proofreading.

I would like to thank the ExpresSys consortium for the numerous scientific discussions. Particularly, Dr. Jennifer Chow (group of Prof. Wolfgang Streit) for providing the metagenomic esterase clones and numerous advice and Dr. Thomas Drepper, Dr. Filip Kovacic and Dr. Anuschka Malach (group of Prof. Karl-Erich Jäger) for their friendly supervision during my lab visits. I also want to give special thanks to Prof. Dr. Sonja-Verena Albers and her group for the very close collaboration throughout the years - particularly to Alexander and Michaela Wagner for opening their home to me on various occasions and for fun times during lab visits, as well as to Dr. Benjamin Meyer for funny congress participations.

Thanks to my bachelor students Yannick Haage and Thomas Meyer for their almost never-ending cloning work. Thanks to Marcel Blum for his RNA preparations and multiple PCR amplifications. Thanks to everyone in the lab for supporting me in experimental work during my pregnancy.

Special thanks to Anna Hagemann and Silke Jachlewski for reviewing my thesis multiple times, for their endless emotional support, even across the country, and their deep friendship which has evolved over the years.

I greatly thank my office mates Silke, Anna, Theresa and Dominik for the wonderful and almost constantly amusing time with lots of coffee, chocolate, waffles and music.

For a wonderful lab time, excellent teamwork, support and friendship I would like to thank Theresa, Verena and “Carlo & the Damen-Labor”.

A big thank you to everyone from the MEB group, former and current members, for an unforgettable time, support and nice working environment: Anna, Silke, Theresa, Jens, Bernadette, Verena, Dominik, Mella, Kohei, Jeanette, Patrick, Britta, Christopher, Sabine, Thomas, Marcel, Frank, Katharina, Rüdiger, Lu, and Christina.

I am deeply grateful for my loving family, especially my mother Katrin and sister Marleen, for supporting me during all times of my life and always believing in me. Last but not least, I express my deep gratitude to my loving husband and daughter.