Protein phosphorylation in Archaea: Central carbohydrate metabolism in *Sulfolobus* species

Dissertation

zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

– Dr. rer. nat. –

vorgelegt von

Dominik Esser

geboren in 24.05.1985

Molekulare Enzymtechnologie und Biochemie (Biofilm Centre)

der

Universität Duisburg-Essen

2013

Die vorliegende Arbeit wurde im Zeitraum von Juni 2010 bis Juni 2013 bei Prof. Dr. Bettina Siebers im Arbeitskreis für Molekulare Enzymtechnologie und Biochemie (Biofim Centre) der Universität Duisburg-Essen durchgeführt.

Tag der Disputation: 05.11.2013

Gutachter:	Prof. Dr. Bettina Siebers
	Prof. Dr. Jochen Stefan Gutmann
	Prof. Dr. Phillip Wright
Vorsitzender:	Prof. Dr. M. Walpuski

List of contents

List of contents I

List of figures V

List of tables VII

1. Introduction 1

1.1. Archaea: The third Domain of Life 1

1.2. Protein phosphorylation in Archaea 3

1.2.1. P-proteome studies in Prokaryotes 6

1.2.2. Protein kinases 10

1.2.3. Protein phosphatases 12

1.3. Central carbohydrate metabolism in Sulfolobus species 14

1.3.1. Glycolysis and gluconeogenesis in *Sulfolobus* 14

1.3.2. Pentose degradation in *Sulfolobus* 17

1.3.3. Unraveling pathway complexity and regulation 19

1.3.3.1. Glucose-1-dehydrogenases in Archaea 20

1.3.3.2. Members of aldehyde dehydrogenase superfamily in *S. solfataricus* 21

1.3.3.3. Triose-3-phosphate isomerases 22

1.4. Aim of this work 24

2. Materials and methods 25

2.1. Chemicals and plasmids 25

2.2. Instruments 25

2.3. Strains of Escherichia coli and growth conditions 27

2.4. Strains of Sulfolobus and growth conditions 27

2.4.1. *S. solfataricus* P2 for proteomics 27

2.4.2 *S. acidocaldarius* for proteomics 27

2.4.3 Cultivation of *S. acidocaldarius* 28

2.5. Molecular biological techniques: working with DNA 29

2.5.1. Isolation of *S. acidocaldarius* genomic DNA 29

2.5.2. Polymerase chain reaction (PCR) 29

2.5.3. Agarose gel electrophoresis 30

2.5.4. Plasmid preparation 31

2.5.5. Purification of DNA fragments from agarose gels 31

2.5.6. Restriction of DNA 31

2.5.7. Ligation 31

2.5.8. Transformation 32

2.5.9. Restriction analysis of positive colonies 32

2.5.10. DNA sequencing 32

2.5.2. Disruption of the Saci_0545 and Saci_0884 genes in *S. acidocaldarius* 32

2.6. Molecular biological techniques: working with RNA 34

2.6.1. RNA isolation and sample preparation for RNA-seq analysis 34

2.6.2. Bioinformatic analysis of RNA-seq data 35

2.6.3. Reverse transcription quantitative real time PCR 35

2.7. Biochemical methods 36

2.7.1 Determination of protein concentration 36

2.7.2. Heterologous expression and protein purification 36

2.7.2.1. Aldehyde dehydrogenases from *S. solfataricus* P2 36

2.7.2.2. Triosephosphate isomerase from *S. solfataricus* P2 37

2.7.2.3. Protein phosphatases from *S. acidocaldarius* DSM 639 37

2.7.2.4. Glucose-1-dehydrogenase from *S. acidocaldarius* DSM 639 38

2.7.2.5. Fructose-1,6-bisphosphate aldolase/phosphatase from *S. solfataricus* 39

2.7.3. Determination of the native molecular weight 39

2.7.4 Sodium-dodecylsulfate – polyacrylamide-gel-electrophoresis (SDS-PAGE) 39

2.7.5. Immunoblotting analysis 40

2.8. Enzyme assays 41

2.8.1. Aldehyde dehydrogenase enzyme assay 41

2.8.2 Triosephosphate isomerase enzyme assay 41

2.8.3. Glucose-1-dehydrogenase enzyme assay 42

2.8.4. Protein phosphatase enzyme assays 42

2.9. Kinetic paramteres 43

2.10. Sample preparation for Proteome analysis 43

2.10.1. Protein extraction and quantitation 43

2.10.2. Protein digestion and fractionation 44

2.11. Mass spectrometry analysis and data processing 44

2.11.1 *S. solfataricus* P2 44

2.11.1 *S. acidocaldarius* 45

2.12. Biophysical techniques 47

2.12.1. Flow cytometry analysis 47

2.12.2. Cell size analysis 47

2.12.3. Motility assay on semi-solid gelrite plates 47

2.13. Software, internet databases and tools 48

3. Results and discussion 49

3.1. Change of carbon source causes dramatic effects in the phospho-proteome of the archaeon Sulfolobus solfataricus 49

3.1.1. Introduction 49

3.1.2. Phospho-proteome analysis 50

3.1.3. Classification of p-proteins 53

3.1.4. Comparison with other prokaryotic- and eukaryotic-p-proteomes 56
3.1.5. Glucose vs Tryptone 57
3.1.6. Central carbohydrate metabolism 57
3.1.7. Hexose degradation 59
3.1.8. Pentose degradation (D-arabinose, D-xylose and L-arabinose) 60
3.1.9. Candidates of unknown function 61
3.1.10. Gluconeogenesis, glycogen metabolism and trehalose synthesis 62
3.1.11. RuMP pathway and TCA cyle 63
3.1.12. Conclusion 64

3.2. (Cren)Archaeal Protein Phosphorylation: In Vivo Impact of Protein Phosphatase Deletions on Cell Size, Motility and Energy Metabolism in Sulfolobus acidocaldarius 65
3.2.1. Introduction 65
3.2.2. Expression and enzymatic characterization of Saci-PTP and Saci-PP2A 67
3.2.3. Characterization of *saci_ptp* and *saci_pp2a* deletion mutants 69
3.2.4. Comparative phosphoproteome analysis of MW001, *Δsaci_ptp* and *Δsaci_pp2a* 72
3.2.5. P-proteins of the central carbohydrate metabolism in MW001, *Δsaci_ptp* and *Δsaci_pp2a* 76
3.2.6. Comparative whole-transcriptome profiling of *Δsaci_ptp* and *Δsaci_pp2a* by RNA-seq 79
3.2.7. Proteins of transcriptionally regulated clusters are also phosphorylated 83
3.2.8. Influence of the phosphatases deletions on regulation of cell motility and aerobic respiration 85
3.2.9. Discussion 87

3.3. Atypical protein kinases of the RIO family in Archaea 93
3.3.1. Introduction 93
3.3.1. Distribution of RIO kinases in Archaea 94
3.3.2. Gene neighborhood of RIO kinases 95
3.3.3. RIO gene cluster in Sulfolobaceae 97
3.3.4. First evidence for the function of RIO kinases in *S. solfataricus* P2 98
3.3.5. Conclusion 99

3.4. Unraveling the function of paralogs of the aldehyde dehydrogenase super family from Sulfolobus solfataricus 101
3.4.1. Introduction 101
3.4.2. Heterologous expression of SSO1218, SSO1629 and SSO1842 in *E. coli* and purification 103
3.4.3. Enzyme characterisation of the succinic semialdehyde dehydrogenases 103
3.4.4. Enzyme characterisation of the methylmalonate semialdehyde dehydrogenase 106
3.4.5. Succinic semialdehyde dehydrogenase isoenzymes 107
3.4.6. Methylmalonate semialdehyde dehydrogenase 109
3.4.7. Phylogenetic affiliation of ALDH paralogs in *S. solfataricus* 111

3.5. Characterization of the glucose-1-dehydrogenase (Saci_1079) from Sulfolobus acidocaldarius DSM 639 116
3.5.1. Heterologues expression in *E. coli* BL21 pRIL and purification 116

3.5.2. Substrate specificity of glucose-1-dehydrogenase 117
3.5.3 Temperature optimum of GDH 118
3.5.4. Influence of Me^{2+}-ions on GDH activity 118
3.5.5. Enzyme characterization of the GDH from *S. acidocaldarius* 119
3.5.6. Glucose-1-dehydrogenases in Archaea 122
3.5.7. GDH paralogs in *Sulfolobus acidocaldarius* DSM639 123
3.5.8. Phosphorylation of GDH paralogs in *S. solfataricus* P2 and *S. acidocaldarius* DSM 639 126

3.6. Characterization of the triose-3-phosphate isomerase from Sulfolobus solfataricus P2 in response to temperature change 128
3.6.1. Heterologoues expression in *E. coli* ROSETTA (DE3) and purification 128
3.6.2. Enzymatic characterization of TPI 129
3.6.3. Triosephosphate isomerase from *S. solfataricus* P2 131

4. Summary 134

5. Literatur 137

6. Appendix 153

6.1 Abbreviations 153

6.2. Supporting informations 157
6.2.1. Tabels 157

6.2 Publikationsliste 160

6.3 Lebenslauf 163

6.4 Erklärung 164

6.5 Danksagung 165

List of figures

Figure 1 Maximum likelihood phylogeny of the Archaea based on concatenated ribosomal protein gene sequences (8077 sites) from representative complete genomes according to [7]. 2
Figure 2 The ePK catalytic domain according to [80]. 11
Figure 3 The three families of protein serine/threonine phosphatases according to [93]. 13
Figure 4 The central carbohydrate metabolism in *S. solfataricus* and *S. acidocaldarius*. 15
Figure 5 Established pentose degradation pathways in *Sulfolobus*. 17
Figure 6 Comparison of the total number of detected p-proteins, p-peptides, p-sites and their distribution on the amino acids Ser, Thr, Tyr in *S. solfataricus* P2 (grown on glucose and tryptone, current study [188]), *B. subtilis* [70], *E. coli* [71], *L. lactis* [72], *Zea mays leaf* [189] and *H. salinarium* [46]. 51
Figure 7 Representative MS/MS spectra for phosphorylation on tyrosine. 53
Figure 8 Assignment of p-proteins into the arCOG functional classes of *S. solfataricus* P2. 54
Figure 9 Percentage distribution of p-proteins in all (A) and arCOG functional categories associated with cellular metabolism (B) detected in tryptone (orange) and glucose (black) grown cells. 57
Figure 10 Current understanding of hexose and pentose degradation, glycogen metabolism, trehalose synthesis, TCA cycle and C4/C3 interconversion in *S. solfataricus* P2. 60
Figure 11 Purification of Saci-PP2A and Saci-PTP. 67
Figure 12 Biochemical characterization of the protein phosphatase Saci-PTP. 68
Figure 13 Biochemical characterization of the protein phosphatase Saci-PP2A. 68
Figure 14 Metal-ion dependency of Saci-PP2A in the presence of EGTA. 69
Figure 15 Δ*saci_pp2a* shows a different growth behavior in comparison to MW001 and Δ*saci_ptp*. 69
Figure 16 The phosphatase deletion mutant Δ*saci_pp2a* revealed a phenotype in cell shape and size. 70
Figure 17 Complementation of the Δ*saci_pp2a* phenotype. 71
Figure 18 Flow cytometry of MW001 in comparison to *Δsaci_ptp* and *Δsaci_pp2a*. 71
Figure 19 Comparison of p-proteome studies in Archaea, Bacteria and Eukarya. 73
Figure 20 Distribution of unique p-proteins into their arCOG functional categories from *S. acidocaldarius*. 75
Figure 21 Reconstruction of the CCM including glycolysis via the branched ED pathway, gluconeogenesis, as well as TCA cycle branching points towards amino acids, glycerol, mannose and glycogen metabolism and C_4/C_3 interconversion in *S. acidocaldarius*. 77
Figure 22 Correlation between RNA-seq reads and fold change between sample conditions. 80
Figure 23 Transcriptionally regulated genes of Δ*saci_ptp* and Δ*saci_pp2a* in comparison to the background strain MW001. 81
Figure 24 Bar plot representation of the transcriptomic changes of *Δsaci_ptp* and *Δsaci_pp2a*. 82
Figure 25 Proposed model of *S. acidocaldarius* branched aerobic respiration chain. 83
Figure 26 Differential expression levels of archaella operon genes and respiratory chain genes in Δ*saci_pp2a* and Δ*saci_ptp*. 86
Figure 27 Impact of phosphatase deletion on archaella expression and motility. 87
Figure 28 Model of the involvement of phosphorylation in the regulation of the archaellum in *S. acidocaldarius*. 90
Figure 29 Conserved gene neighborhood of archaeal RIO kinases. 97
Figure 30 RIO kinases in Sulfolobaceae. 98
Figure 31Purification of the recombinant aldehyde dehydrogenases from *S. solfataricus*. 103
Figure 32 Substrate and cosubstrate concentration dependent velocity plots for SSADH-I with DOP (black squares) (A), SSA (black squares) as substrate and NAD^+ (white circles) as cosubstrate (B). 104
Figure 33 Substrate and cosubstrate concentration dependent velocity plots for SSADH-II. 106
Figure 34 Pathway reconstruction of the glutamate (orange framed), polyamine (green framed) and GABA (blue framed) metabolism in *S. solfataricus*. 108
Figure 35 Physiological role of MSDH from *S. solfataricus* in the amino acid degradation of Val, Leu and Ile. 110
Figure 36 Current understanding of the role of the ALDHs in the CCM *S. solfataricus*. 111
Figure 37 Sequence alignment of SSADHs from *S. solfataricus* and GabD from *E. coli*. 112
Figure 38 Sequence alignment of MSDH from *S. solfataricus* and selected MSDHs. 113
Figure 39 Phylogenetic affiliation of selcted ALDH sequences from Archaea (red), Bacteria (blue) and Eukarya (green) constructed using the Neighbor-Joining method [299]. 115
Figure 40 Purification of the recombinant glucose-1-dehydrogenase (Saci_1079) from *S. acidocaldarius* DSM 639. 116
Figure 41 Enzyme activity of Saci-GDH in presence of various sugars and either 5 mM NAD^+ (grey) or 5 mM $NADP^+$ (black) at 70°C (pH 6.5). 117

Figure 42 Temperature dependence of Saci-GDH (Saci_1079) activity. 118
Figure 43 Influence of different Me^{2+}-ions on Saci-GDH activity at 70°C. 119
Figure 44 Cosubstrate concentration dependent velocity plots for Saci-GDH with D-xylose as substrate. 120
Figure 45 Substrate concentration dependent velocity plots for Saci-GDH with NAD^+ as cosubstrate. 120
Figure 46 Substrate concentration dependent velocity plots for Saci-GDH with $NADP^+$ as cosubstrate.. 121
Figure 47 Multiple sequence alignment of GDHs from *S. acidocaldarius* DSM 639 and *S. solfataricus* P2 identified as p-proteins..... 127
Figure 48 Purification of the recombinant triose-3-phosphate isomerase SSO2592 from *S. solfataricus* P2. 128
Figure 49 Effect of temperature on the kinetic properties of SSO-TPI. 129

List of tables

Table 1 Comparison of available prokaryotic and eukaryotic p-proteome studies. 7
Table 2 Composition of Brock's basal salt medium 28
Table 3 Sequences of primers used in PCR mutagenesis. 30
Table 4 PCR protocol including duration, temperature of each phase and cycle numbers. 30
Table 5 Expression of *S. solfataricus* P2 and *S. acidocaldarius* DSM 639 target proteins. 32
Table 6 Primer sets for the construction of deletion mutants in *S. acidocaldarius* 35
Table 7 Comparison of p-proteome studies. 52
Table 8 Comparison of archaeal and bacterial p-proteome analyses. 55
Table 9 Comparison of genomes and p-proteome features of available prokaryotic p-proteome analyses. 56
Table 10 Identification of archaeal target proteins. 58
Table 11 Comparison of different ph-proteome studies with the three *S. acidocaldarius* p-proteome sets. 72
Table 12 CCM enzymes targeted by protein phosphorylation in *S. acidocaldarius*. 78
Table 13 P-proteins related aerobic repiration, motility, transcriptional regulation and phosphorylation identified in the *S. acidocaldarius* p-proteome. 84
Table 14 Kinetic parameters of SSADH-I (SSO1629) for various aldehydes with NAD^+ as cosubstrate. 104
Table 15 Kinetic parameters of SSADH-II (SSO1842) for various aldehydes with $NADP^+$ as cosubstrate. 105
Table 16 Kinetic parameters of SSADH-II (SSO1842) for SSA and DOP with NAD^+ as cosubstrate 105
Table 17 Kinetic parameters of Saci-GDH for the cosubstrates NAD^+ and $NADP^+$ in presence of D-Xylose as substrate. 119
Table 18 Kinetic parameters of Saci-GDH for various substrates with NAD^+ and $NADP^+$ as cosubstrates at 70°C. 121
Table 19 Comparison of kinetic parameters of the characterized members of the MDR branch of the pyridine nucleotide-dependent alcohol/polyol/sugar dehydrogenase superfamily from *S. solfataricus* P2 and GDH (Saci_1079) from *S. acidocaldarius* DSM 639 at 70°C. 122
Table 20 Similarity comparison between the GDH paralogs of *S. solfataricus* P2 and *S. acidocaldarius* DSM639 126
Table 21 Identified *Sulfolobus* GDHs paralogs which are targeted by protein phosphorylation 127
Table 22 Kinetic parameters of the TPI from *S. solfataricus* for its substrates DHAP and D,L-GAP. 131
Table 23 Kinetic parameters for the conversion of DHAP to GAP by TPIs from hyperthermophilic Archaea and Bacteria 132
Table 24 Kinetic parameters for conversion of GAP to DHAP by TPIs from hyperthermophiles. 133
Table SI 1 Identified p-proteins involved in the central carbohydrate metabolism. 157

1. Introduction

1.1. Archaea: The third Domain of Life

The third domain of life, the Archaea, were established beside the Bacteria and Eukaryotes approximately 40 years ago by Woese and co-workers [1]. Archaea are known to dominate extreme environments like thermal hot springs and salt lakes, but occur also in moderate habitats. Since their discovery, major efforts were made to unravel basic and unique features of these organisms. They resemble their prokaryotic brethren in size, shape and DNA organization (e.g. circular chromosome, plasmids, operon structures) whereas, mechanisms involved in information processing (transcription, translation, cell division etc.) appear to be more similar to the respective eukaryotic equivalents; although less complex. Therefore, Archaea are special in regard to their molecular/cellular properties as highlighted for example by the archaeal membrane made of isoprenoid lipids [2]. Furthermore, the Archaea gained a lot of interest due to their unique metabolic properties. In terms of complexity, archaeal metabolism resembles that of Bacteria and primitive Eukaryotes. However, many unusual pathways and enzymes were identified within this domain [3]. Even the generic pathways of sugar degradation, such as the Embden-Meyerhof-Parnas (EMP) and the Entner-Doudoroff (ED) pathway, have modifications, as compared to the classical pathways [3,4]. Many archaeal enzymes share no homology with their bacterial and eukaryotic counterparts and are members of different novel enzyme families.

The improvement in DNA sequencing technologies gave rise to a vast amount of genome sequence information from representatives of all three domains of life as well as environmental communities (metagenomics). Today 196 archaeal genomes have been sequenced (March 2013, integrated microbial genomes database, http://img.jgi.doe.gov/cgi-bin/w/main.cgi). These 196 archaeal species can be divided into six main phyla. The two largest and first established phyla are the Euryarchaeota and the Crenarchaeota [5]. The Euryarchaeota comprise all methanogens and extreme halophiles and some thermoacidophiles and (hyper)thermophiles. Whereas the Crenarchaeota only contain (hyper)thermophilic Archaea [5]. The remaining four phyla are the Nanoarchaeota (under discussion), Thaumarchaeota, Korarchaeota and the recently proposed Aigarchaeota (under discussion, Figure 1). The Nanoarchaeota comprise so far only two species *Nanoarchaeum equitans*, which can only be grown in co-culture with the Crenarchaeum *Ignicoccus hospitalis* [6], and Nst1 were the likely host is a member of the *Sulfolobales* (Acd1) [7]. The Korarchaeota instead contain only uncultivated hyperthermophilic and anaerobic Archaea and were proposed as an separate phylum approximately 15 years ago (based on SSU rRNA sequences) [8]. One special feature derived

from the genome sequence information of Korarchaeotae is that they seem to lack several biosynthetic pathways, which explains to some extend why they could not be cultivated so far [9].

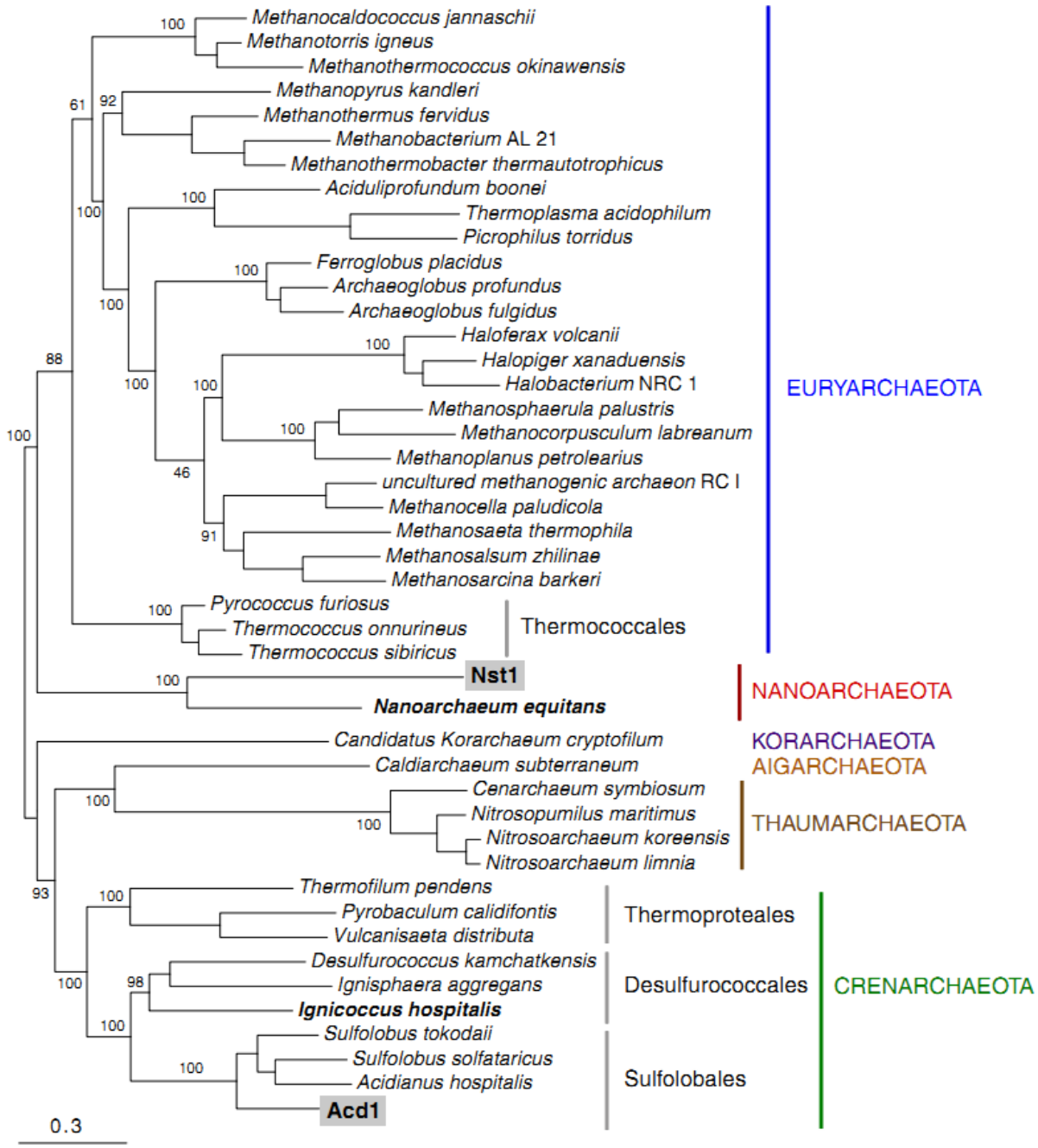


Figure 1 Maximum likelihood phylogeny of the Archaea based on concatenated ribosomal protein gene sequences (8077 sites) from representative complete genomes according to [7].

Members of the Thaumarchaeota are both thermophiles and mesophiles and thrive in very different environmental niches (e.g. freshwater, soil ocean and hot springs) [10]. In addition, some Thaumarchaeota are shown to participate in the aerobic oxidation of ammonia, which is thought to be solely performed by β- and γ-Proteobacteria [10]. The last phylum is the Aigarchaeota. Informations regarding this only recently introduced phylum are scares and only one member

HWCG1 was identified until now [11]. If the Aigarchaeota are indeed a new archaeal phylum or a deeply branching thaumarchaeal lineage is still a matter of debate [12].

Within the domain of Archaea, *Sulfolobus solfataricus* has been established as one of the model organisms [13]. This Crenarchaeon is a thermoacidophilic obligate aerobe with optimal growth at 80°C and pH 2-3. It was first isolated from a solfatara in Pisciarelli, Italy [14] and utilizes a variety of organic compounds as carbon and energy sources such as sugars (e.g. cellulose, cellobiose, D-glucose, D-xylose), alcohols, aldehydes, amino acids or peptides under aerobic conditions [15]. The organism fulfils all requirements for a systems biology organism; the genome sequence has been solved [16], genetic tools have been established [17-21] and *S. solfataricus* is easy to grow (aerobic heterotroph). Therefore, *S. solfataricus* is an interesting candidate for the application in biotechnology and synthetic biology [22,23]. The central carbohydrate metabolism (CCM) has been extensively studied [3,24-28]. *S. solfataricus* uses modifications of the respective hexose and pentose degradation pathways, like the branched ED pathway for glucose and galactose [3,24,28], the Dahms pathway for D-xylose degradation [25] and an oxidative pathway for D-arabinose degradation [27]. Like most other Archaea, *S. solfataricus* uses the modified EMP pathway for gluconeogenesis [3,26]. Although the metabolic pathways have been elucidated, their regulation is still unclear. It is known though that classical bacterial and eukaryotic control points are absent [3] and only minor regulation has been observed at the transcript and proteome level in response to carbon source alteration [26].

In the last few years a close relative of *S. solfataricus* gained a lot of interest, *Sulfolobus acidocaldarius*. It grows optimal under similar conditions (78°C, pH 2-3) as *S. solfataricus* and it also uses the branched ED pathway for glycolysis and the EMP pathway for gluconeogenesis The genetic system is more advanced and versatile compared to the one from *S. solfataricus*, which qualifies the organism for metabolic engineering, biotechnology and synthetic biology applications [29,30].However, *S. acidocaldarius* is less versatile compared to *S. solfataricus* and is only able to use few carbon sources [31].

1.2. Protein phosphorylation in Archaea

In nature, microorganisms have to cope with permanent changes in their environment regarding biotic and abiotic factors such as temperature, pH, carbon and/or nitrogen supply. In order to survive they have to sense and response to these changes. This seems to be especially challenging for organisms thriving in extreme habitats. A large set of post translational modifications (PTMs) are identified in all three domains of life. The most prominent and best investigated PTMs are for example protein phosphorylation, protein glycosylation, protein methylation, protein acetylation and the introduction of disulfide bonds [32]. These modifications can in general alter the biological and/or physico-chemical properties of the protein (e.g. turnover, fold). Originally, it was assumed

that PTMs are restricted to more evolved organisms (i.e. eukaryotes), however later on PTMs were also established in Bacteria as well as in the third domain of life the Archaea.

Among PTMs reversible protein phosphorylation is of special interest due to its reversibility. It plays a major role in signal transduction and allows for a rapid cellular response to diverse external and internal stimuli [33]. In the past, it was assumed that protein phosphorylation on Ser, Thr and Tyr residues is a typical eukaryotic feature, whereas in Prokaryotes protein phosphorylation takes place on His and Asp residues via one/two-component systems. However, today it is well established that one/two-component systems are also present in Eukaryotes and that protein phosphorylation on Ser, Thr and Tyr takes also place in Prokaryotes.

Protein phosphorylation was originally discovered in the 1950s by Krebs and Fischer during their investigation of the rabbit skeletal muscle [34]. They demonstrated that the phosphorylase b is converted to phosphorylase a via autophosphorylation in the presence of [γ-^{32}P-ATP]. However, it took approximately 22 years until Wang and Koshland were able to demonstrate protein kinase (PK) activity in the gram-negative bacterium *Salmonella typhimurium* [35]. One year later it was reported about protein phosphorylation in the gram-negative bacterium *Escherichia coli* by Garnak and Reeves [36]. They identified the first target protein in Prokaryotes, the isocitrate dehydrogenase (IDH), which is even today approximately 35 years later still an exciting phospho-protein (p-protein) due to its regulatory role in the tricarboxylic acid (TCA) cycle. A short time later protein phosphorylation was also reported in the Haloarchaeon *Halobacterium salinarium* by Spudich and Stoeckenius [37] leading to the discovery of the first archaeal two-compounend system CheA and CheY. CheA is a sensor kinase that autophosphorylates and transfers the phosphate to the response regulator CheY which targets the flagellar motor and regulates the direction of the flagellum rotation (clockwise or counterclockwise) [38,39]. In the 1980's Skorko reported the presence of protein phosphorylation in the thermoacidophilic Crenarchaeon *S. acidocaldarius* [40,41]. He measured the PK activity in crude extracts (CE) of cells harvested in the exponential growth phase and stationary growth phase and demonstrated that the CE of the stationary growth phase showed higher PK activity compared to the CE of the stationary growth phase. Furthermore, the influence of divalent Me-ions on the PK activity was demonstrated as well as the protein phosphorylation of the membrane fraction. Therefore, protein phosphorylation is well established in all three domains of life and the previous postulated hypothesis that phosphorylation on Ser, Thr and Tyr residues is a typical feature of Eukaryotes and that in Prokaryotes protein phosphorylation takes solely place on His and Asp residues (one/two-component systems) was disapproved. Comparative genome analysis were not able to identify any His kinases within the Crenarchaeota, whereas several homologs were found in Euryarchaeota [32]. The reason for the missing phosphorylation on His and Asp residues in Crenarchaeota is unclear. Furthermore, it was reported for CheY from *Thermotoga maritima* (optimal growth temperature at

80°C) that the phosphorylation is only stable for less than 30 seconds at 50°C [42], which might represent a great disadvantage and challenge for (hyper)thermophiles.

Only recently it has been proposed that negatively charged amino acids like Asp and Glu are evolutionary linked to phosphorylation sites on Ser and Thr and *vice versa* [43]. Pearlman *et al*., used a comparative genomics approach and proposed that Asp and Glu residues are sometimes (5% of the examined sites) substituted by the neutral amino acids Ser or Thr. Asp and Glu are negatively charges residues similar to p-Ser, p-Thr and p-Tyr residues. However, Asp and Glu are only singly charged, whereas p-Ser, p-Thr and p-Tyr are doubly charged under neutral pH conditions [44]. Furthermore, Asp and Glu are isosteric with p-Ser and p-Thr. Thorsness and Koshland showed something similar already in 1987. They investigated the IDH from *E. coli*. The unphosphorylated IDH is active, whereas the phosphorylated form is inactive [36]. They substituted the usually phosphorylated Ser residue by a negatively charged Asp residue. The IDH mutant showed the same behavior as the phosphorylated IDH and the substitution of the Ser residue with other amino acids resulted in partially active enzymes [45]. The opposite effect can also be imagined, where an Asp or Glu residue is substituted by Ser or Thr, which leads to protein inactivation and the phosphorylation of the Ser or Thr residue would reactivate the protein. A similar effect in case of p-Tyr is unlikely, due to the fact that p-Tyr is not structurally similar to Asp and Glu. However, in cases where the charge of the side chain is more important than the structure such an substitution might be possible [43].

So far, the conducted protein phosphorylation studies in *H. salinarium* indicate that the majority of the phosphorylated archaeal proteins are modified on Thr and Ser, whereas only a small fraction is modified on Tyr [32]. This was confirmed by a recent phosphoproteome (p-proteome) study, which revealed that less than one percent of the detected phospho-peptides (p-peptide) were modified on Tyr [46]. *Sulfolobus* species are the best studied Archaea regarding protein phosphorylation. In *S. solfataricus* two protein phosphatases (PPs) [47,48] and five PKs [49-53] were characterized. Furthermore, in *S. acidocaldarius* and *S. tokodaii* an *in vivo* function of protein phosphorylation in the regulation of the archaellum expression was demonstrated [54,55]. For *S. acidocaldarius* it has been shown that the forkhead-associated (FHA) domain-containing protein ArnA and the von Willebrand (vWA) domain-containing protein ArnB are targeted by protein phosphorylation [56]. Both proteins are capable to repress archaella (former: archaeal flagellum) expression. Similar results regarding the phosphorylation of the FHA domain-containing protein were also shown in *S. tokodaii* [55].

Reports of phosphorylation-mediated post-translational regulation of the CCM are rather rare. So far, only for the gluconate dehydratase (GAD) from *S. solfataricus* inhibition by protein phosphorylation has unambiguously been demonstrated [57], raising the question if PTMs, such as phosphorylation, play a role in CCM regulation in *S. solfataricus* and Archaea in general. In addition, the 2-keto-3-deoxy-gluconate (KDG) kinase from *Thermoplasma acidophilum* and the 2,3-bisphosphoglycerate

independent phosphoglycerate mutase (iPGAM) from *S. solfataricus* were found to be phosphorylated [58,59]. For the KDG kinase it was reported that the removal of the phosphorylation leads to a loss of enzyme activity, whereas the iPGAM underwent autophosphorylation when incubated with [γ-^{32}P-ATP] or [γ-^{32}P-GTP]. However, in both cases, it still needs to be confirmed if the p-proteins represent p-enzyme intermediates formed during enzyme catalysis or if they are real target proteins for regulation via protein phosphorylation.

To unravel signal transduction by reversible protein phosphorylation three major tasks are of great importance: i) investigation of all PKs "the kinome" of the organism of interest, ii) characterization of the PPs and their substrates and iii) mapping the p-proteome, comprising all p-proteins of a cell at a certain time point under defined growth conditions. However, knowledge about protein phosphorylation and signal transduction in Archaea is still rather rare compared to the achievements in Bacteria and Eukaryotes.

1.2.1. P-proteome studies in Prokaryotes

In the past the detection of p-proteins was performed via immunoblotting with antibodies against p-Ser, p-Thr or p-Tyr [60] or by labeling using radioactive [γ-^{32}P-ATP] or [γ-^{33}P-ATP] as substrate for the PK reaction [61]. Both detection methods are reliable but also have major draw backs. The detection via antibodies cannot provide any information of specific phosphorylation sites and the corresponding bands need to be further investigated by MS-based approaches (e.g. MALDI-TOF). The use of labeled [γ-^{32}P-ATP] suffers from the presence of cellular unlabeled ATP pools and artificial phosphorylation events due to the unnatural ratio between the PKs and their substrates [62]. Labeled [γ-^{32}P-ATP] is very often used in combination with gel electrophoresis (i.e. SDS-PAGE, 2-DE gels) to reveal signal transduction pathways by incubation of PKs with putative substrates and subsequent separation via electrophoresis. However, none of these methods provides details about the phosphorylation site. A more advanced detection method is the p-protein staining via Pro Q-Diamond. This fluorescent is able to stain p-Ser, p-Thr and p-Tyr and can be used in combination with gel electrophoresis [63,64]. These are only the most prominent detection methods. Several more methods are available and all of them offer information on protein phosphorylation. The invention of high throughput techniques linked to MS allows for new fast ways to analyze the p-proteome.

Usually p-proteins are enriched prior to the analysis via MS by different enrichment procedures such as metal oxide affinity chromatography (MOAC). MOAC uses metal oxides like TiO_2 or ZrO_2. For TiO_2 a very effective p-peptide enrichment was shown due to high affinity binding of the phosphate moiety in aqueous solutions [65] and today different enrichment techniques via column, titansphere or titania-packed pipette tips have been established. The general principle is that the TiO_2 is

equilibrated under acidic conditions, which leads to a positive charge of the TiO_2 and prevents the binding of non-phosphorylated peptides. By the use of alkaline buffers, the absorbed p-peptides elute from the TiO_2 and can be directly injected in the MS-system [66]. ZrO_2 has similar properties as TiO_2 and has gained a lot of attention during the last years due to its higher thermal stability and physical properties [67]. In general ZrO_2 is positively charged under acidic conditions and its binding affinity towards phosphate is higher than towards carboxylated anions [68].

The first p-proteome mapping was performed for the gram-positive soil bacterium *Corynebacterium glutamaticum* in 2003 by Bendt and co-workers [69]. In this study, two different detection methods were applied: i) *in vivo* labeling by [γ-^{33}P-ATP] and subsequent autoradiography and ii) immune detection via phospho amino acid specific antibodies. Afterwards, the samples were separated via 2-D gel electrophoresis and analyzed via MS.

Table 1 Comparison of available prokaryotic and eukaryotic p-proteome studies.
i = intracellular *Toxoplasma gondii* samples; p = purified *Toxoplasma gondii* samples; PH = anaerobic photoheterotrophic grown *Rhodopseudomonas palustris*; CH = aerobic chemoheterotrophic grown *Rhodopseudomonas palustris*, AG = *Clostridium acetobutylicum* cells in the exponential growth phase when acetic acid and butyric acid are produced, SG = *Clostridium acetobutylicum* cells after the stationary growth phase when acetone, butanol and trace amounts of ethanol are produced. The genome size was calculated by http://img.jgi.doe.gov/cgi-bin/w/main.cgi.

Organism	Domain of Life	Genome size [Mb]	P-proteins [no.]	P-peptides [no.]	P-sites [no.]	p-Ser [%]	p-Thr [%]	p-Tyr [%]
***Halobacterium salinarum* [44]**	Archaea	2.7	69	90	81	86	13	1
***Bacillus subtilis* [67]**	Bacteria	4.2	78	103	78	69.2	20.5	10.3
***Campylobacter jejuni* [73]**	Bacteria	1.6	36	58	35	30.3	72.7	9.1
***Clostridium acetobutyicum* [82]**	Bacteria	4.1	44^{AG}	52^{AG}	70^{AG}	40^{AG}	50^{AG}	10^{AG}
			51^{SG}	70^{SG}	89^{SG}	43.8^{SG}	47.2^{SG}	9.0^{SG}
***Corynebacterium glutamaticum* [66]**	Bacteria	3,174	41	-	-	-	-	-
***Escherichia coli* [68]**	Bacteria	4.6	79	105	81	67.9	23.5	8.6
***Heliobacter pylori* [79]**	Bacteria	1.7	67	80	126	42.8	38.7	18.5
***Klebsiella pneumonia* [77]**	Bacteria	5.5	81	117	93	31.2	15.1	25.8
***Lactococcus lactis* [69]**	Bacteria	2.4	63	102	73	46.5	50.6	2.7
***Listeria monocytogenes* [81]**	Bacteria	2.9	112	155	143	65	30.1	4.9
***Mycobacterium tuberculosis* [80]**	Bacteria	4.4	301	381	516	60	40	0

***Mycoplasma pneumoniae* [78]**	Bacteria	0.8	63	16	16	53.3	46.7	0
***Pseudomonas aeruginosa* [72]**	Bacteria	6.3	23	57	55	52.7	32.7	14.5
***Pseudomonas putida* [72]**	Bacteria	5.9	40	56	53	52.8	39.6	7.5
***Rhodopseudomonas palustris* [86]**	Bacteria	5.5	54^{CH}	100^{CH}	60^{CH}	63.3^{CH}	16.1^{CH}	19.4^{CH}
			42^{PH}	74^{PH}	56^{PH}	58.9^{PH}	23.2^{PH}	17.9^{PH}
***Streptococcus pneumoniae* [74]**	Bacteria	2.1	84	102	163	47.2	43.8	9.0
***Streptomyces coelicolor* [75]**	Bacteria	9.1	40	44	44	34.1	52.3	13.6
***Streptomyces coelicolor* [76]**	Bacteria	9.1	127	260	289	46.8	48	5.2
***Synechococcus sp.* [71]**	Bacteria	3.4	245	2810	410	43.9	42.4	13.6
***Thermus thermophilus* [85]**	Bacteria	2.1	48	52	50	64	26	10
***Arabidopsis thaliana* [89]**	Eukarya	119.7	598	962	609	86	14	0.16
***Plasmodium falciparum* [88]**	Eukarya	23	1,673	7,835	8,463	89.1	10.4	0.5
***Plasmodium falciparum* [83]**	Eukarya	23.3	919	2418	2541	84.4	13.2	2.4
***Trichomonas vaginalis* [84]**	Eukarya	176.4	82	2,418	2,541	84.4	13.2	2.4
***Trypanosoma cruzi* [87]**	Eukarya	67	753	-	2,572	84.1	14.9	1.0
***Toxoplasma gondii* [88]**	Eukarya	80	$2,793^{i}$	$11,822^{i}$	$12,793^{i}$	87.2	12.6	0.25
			$3,506^{p}$	$21,498^{p}$	$24,298^{p}$	89.3	10.5	0.25
***Zea mays* leaf [70]**	Eukarya	2,065	125	149	157	89.8	9.6	0.6

In 2007 the first gel-free p-proteome study was performed for the gram-positive model organism *Bacillus subtilis* via nanoscale liquid chromatography coupled to high resolution hybrid MS with prior p-peptide enrichment by using SCX and TiO_2 chromatographies [70]. Since then, the number of similar p-proteome studies increased and in 2008 the p-proteomes of two other important bacterial model organisms, *E. coli* and *Lactococcus lactis*, were published [71,72]. In 2009 the first archaeal p-proteome from the Haloarchaeon *H. salinarium* was investigated by Aivaliotis and co-workers [46].

So far 24 p-proteomes from seven eukaryotic organisms, 17 Bacteria and only one Archaeon are available (higher Eukaryotes, e.g. human, were neglected, Table 1). For Eukaryotes it was assumed that approximately 30% of the whole proteome are phosphorylated at a certain time and a mean phosphorylation ratio of 90% p-Ser, 10% p-Thr and only 0.05% p-Tyr was estimated [67]. Interestingly, these numbers are not even close to the ones determined in the 24 p-proteome studies. The highest relative amount of p-proteins were detected in the parasites *Toxoplasma gondii* (34.97-43.90%) and *Plasmodium falicarum* (17.35–31.58%), whereas the lowest amount was reported for the parasite *Trichomonas vaginalis* with only 0.08% (Table 1) [73,74]. For the other species a relative amount of p-proteins between 0.12-8.59% was reported. Also the assumed value of 0.05% p-Tyr in Eukaryotes could not be verified by any of the so far conducted studies, instead for the parasite *T. vaginalis* a considerably higher number of 6.8% p-Tyr was reported [74].

For the analyzed Prokaryotes showed (with only four exceptions: *H. salinarium*, *L. lactis*, *Listeria monocytogenes*, *Klebsiella pneumoniae*) a p-Tyr value between 4.9-19.4%, only in case of *Mycobacterium tuberculosis* and *Mycoplasma pneumoniae* no phosphorylation on Tyr was observed at all [61,75]. Evidences for the regulation of the CCM came from several p-proteome studies (Table SI 1). For example in the glycolysis of *B. subtilis* all enzymes with the exception of the hexokinase (HKH) and phosphofructokinase (PFK) were targeted by protein phosphorylation. In *E. coli* at the beginning of the glycolysis the phosphoglucose isomerase (PGI) and all enzymes of the lower shunt of the glycolysis i.e. phosphoglycerate kinase (PGK), PGAM as well as the enolase (ENO) and pyruvate kinase (PyrK) were found to be phosphorylated. In addition, in the thermophilic bacterium *Thermus thermophilus* seven p-proteins, out of 48 identified p-proteins, participating in glycolysis or TCA cycle were detected (three PGM paralogs, FBP aldolase, malate dehydrogenase (MDH), succinyl-CoA synthase α-subunit, IDH) [76]. Furthermore, in the Haloarchaeon *H. salinarium* the PyrK, the pyruvate phosphate dikinase (PPDK), pyruvate:ferredoxin OR, succinate dehydrogenase β-subunit, IDH as well as four phosphomutases and three phosphohexomutases were detected as p-proteins [46]. These are only some examples of CCM enzymes which were identified as p-proteins. In general in almost all available p-proteome studies the CCM is targeted by protein phosphorylation (Table 1). However, detailed comparison is difficult due to the fact that the growth conditions of all investigated species as well as the scientific question and the chosen experimental setup of the different p-proteome studies are quite different.

The identification of many enzymes of the CCM as targets for reversible protein phosphorylation is very interesting but so far the impact in the CCM regulation is still unclear. The best investigated CCM p-protein is the IDH from *E. coli*. The *E. coli* IDH catalyzes the reaction of isocitrate to 2-oxoglutarate at the branching point between the glyoxylate bypass and the TCA cycle. In *E. coli* cells,

grown on acetate or ethanol, approximately 75% of the IDHs are phosphorylated, which inactivates the enzyme and leads to the enhanced flux through the glyoxylate bypass [77].
Unfortunately, such profound knowledge regarding the effect of the phosphorylation is rather scarce for most of the identified CCM p-proteins so far and extensive biochemical studies are required to answer these questions. The probably most challenging task is to identify the corresponding signal transduction pathways involving PKs and PPs, which regulate the (de)phosphorylation.

1.2.2. Protein kinases

Protein phosphorylation is mediated by PKs, which catalyze the phosphorylation of target proteins and protein phosphatases, which specifically remove the respective phosphate group. PKs form a large super family comprising six families [78]. The eukaryotic type like PKs (ePKs) are the most investigated PKs so far. With the emergence of the era of genomics the number of putative ePKs rose with every new genome sequence available. Today there a seven major clusters of ePKs (tyrosine kinase (TK) group, PK A, G and C families (AGC) group, calcium and calmodium regulated PKs (CAMK) group, tyrosine kinase-like (TKL) group, cycline-dependent/mitogen activated/glycogen synthase/ cycline-dependent like PK (CMGC) group, homologs of yeast STE7, STE11 and STE20 PKs (STE) group, cell kinases (CK1) group) known [79]. All ePKs share a conserved catalytic domain containing 12 sub domains (Figure 2). The Mg-ATP molecule binds usually to the amino-terminal lobe (sub domain I-V) and the hinge region, whereas the substrate binds to the carboxy-terminal lobe (sub domain V-XI). The most important residues for catalytic function are the Lys (K) residue in sub domain II, the Asp (D) in sub domain VII, which are involved in the orientation and anchoring of the ATP, and the invariant Asp (D) in the sub domain VIB, which is the likely catalytic base involved in the phospotransfer reaction [80].
A detailed classification of ePKs in Archaea was not performed so far. However, Kennelly analyzed the presence of PKs and PPs in 2003, based on nine available archaeal genomes [33]. His study revealed that *S. solfataricus*, compared to the other eight Archaea, had the highest ePK to PP ratio comprising eight ePKs and two PPs. The other investigated Archaea (e.g. *T. acidophilum*) had between one and four putative ePKs and between one and two PPs. The only exceptions was *Halobacterium* NRC1, where no ePKs and PPs were identified but 14 His kinases, and *M. jannaschii* with four ePKs and three PPs [33]. In general His kinases were only identified in the Euryarchaeota *Archaeoglobus fulgidus* (15), *Methanocaldococcus jannaschii* (16), *Pyrococcus horikoshii* (1) and as already mentioned in *Halobacterium* NRC1 (15). Until now, six ePK from *Sulfolobus* species were investigated. The four ePKs from *S. solfataricus* exhibited only Ser/Thr ePK activity and showed no activity at all towards Tyr substrates [50-53]. The remaining two ePKs were characterized from *S. acidocaldarius* and *S. tokadaii* [54,55].

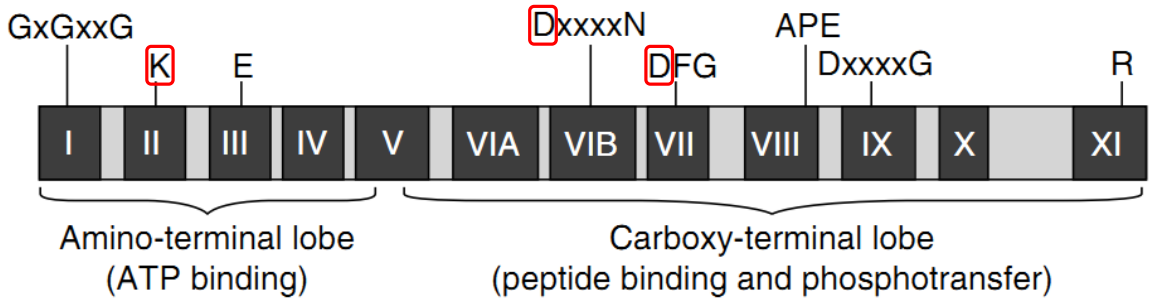


Figure 2 The ePK catalytic domain according to [80].
Above of the sub domains are the motifs and catalytic residues depicted which are conserved in the ePK superfamily, using the single letter amino acid code with x as any amino acid. The catalytic residues are framed in red.

Apart from ePKs [79] the so called atypical PKs (aPKs) of the ABC1, right open reading frame (RIO), piD261, AQ578 and PKN2 family were proposed in 1998 by Leonard and co-workers [78]. Although the presence of aPKs is well established; their physiological function remains largely unknown. Studies of archaeal ABC1, AQ578 and PKN2 were not performed until now. The only investigated aPKs are the piD261 from *S. solfataricus* (SSO0433) and the RIO kinases from *A. fulgidus* and *Haloferax volcanii* [49,81-84]. For the piD261 from *S. solfataricus* it has been shown that ADP-ribose increases the autophosphorylation activity, however no activity was detected in presence of exogenous substrates; therefore its function still remains to be elucidated [49].

Archaeal RIO kinases have been the focus of several studies in the last years. RIO kinases were originally identified in Archaea and Eukaryotes and a common origin with an ancestral RIO gene has been predicted [78]. Until now, four different types of RIO kinases were identified: Rio 1, Rio 2, Rio 3 and Rio B kinase [85]. Phylogenetic analysis confirmed the distribution of RIO kinases in all three domains of life and proposed the presence of a combination of one Rio 1 kinase and one Rio 2 kinase in less complex species (i.e. Prokaryotes and single cellular Eukaryotes) whereas multicellular Eukaryotes including Humans possess in addition one Rio 3 kinases [86].

In general RIO kinases are regarded as a trimmed version of canonical ePKs but lack the subdomains VIII (active loop), X and XI (involved in substrate binding, Figure 2) [83]. RIO kinases possess an insertion of 18-23 amino acids between αC and β3 (flexible loop), which is absent in ePKs. In general Rio 1 and Rio 2 kinases are very similar regarding their overall fold despite the N-terminal domain of the Rio 2 family comprises a wing helix domain, which is absent in all other RIO families. Members of the different RIO families can be distinguished by their specific P-loop (interaction and orientation of the ATP triphosphate moiety) and DFG-loop (metal-ion binding and positioning) sequence [81,85]. For Yeast Rio 1 an essential function in cell cycle progression and chromosome maintenance and furthermore, for both Yeast RIO kinases in 18 S rRNA processing was demonstrated. Deletion of either one of the two RIO kinases in Yeast results in cell death, assuming that both of them possess

different functions [87-92]. A major breakthrough was achieved based on the crystallization of the Rio 1 and Rio 2 kinase from *A. fulgidus* [81,83] and combined bioinformatics analysis [85]. For both archaeal RIO kinases autophosphorylation and activity with common kinase substrates was shown [82,83]. Furthermore, the Rio 1 kinase from *H. volcanii* is capable to phosphorylate the α1 protein of the proteasome 20S core particle [84]. *In vivo* analyses indicate that this event is important for the pigmentation, cell viability and proteasome activity.

1.2.3. Protein phosphatases

The counterparts of PKs are the PPs which are capable to remove the covalently linked phosphate group. Overall, three families of PPs are known (Figure 3): i) phosphatases that remove the phosphorylation from Ser and/or Thr amino acids, ii) the ones which dephosphorylate Tyr (PTP) and iii) aspirate based phosphatases like the TFIIF-associated component of RNA polymerase II carboxy-terminal domain phosphatase (FCP) and small carboxy-terminal domain phosphatases (SCP, not further discussed) [93]. Investigations in Eukaryotes also revealed the presence of atypical PPs, which were so far not detected in Prokaryotes [94].

Ser/Thr PPs can be further classified into two subfamilies the protein Ser/Thr phosphatases (PPPs) and the Mg^{2+}- or Mn^{2+}-dependent protein phosphatases (PPM). In the Archaea both subfamilies are present. PPPs are usually responsible for the dephosphorylation of Ser/Thr in Eukaryotes and Prokaryotes, whereas the PPMs are more common in Bacteria [33,95-97]. A great difference between the PPM and the PPP members is that PPPs usually appear as heterodimer or trimer. The heterodimeric PPPs consist of the catalytic subunit and of a scaffolding forming subunit, whereas the heterotrimeric PPPs have an additional regulatory subunit. The interplay between the different subunits determines substrate specificity (for review see [98]). In contrast, PPMs contain sequence motifs and additional domains, which might mimic the presence of regulatory subunits [98]. Conversely, the PTP family can be subdivided into the PTPs, which are specific for Tyr dephosphorylation, the dual specific PTPs, which can dephosphorylate Ser/Thr as well as Tyr, and the low molecular weight PTPs (<18 kDa). They all share one common amino acid motif, CX_5R [99].

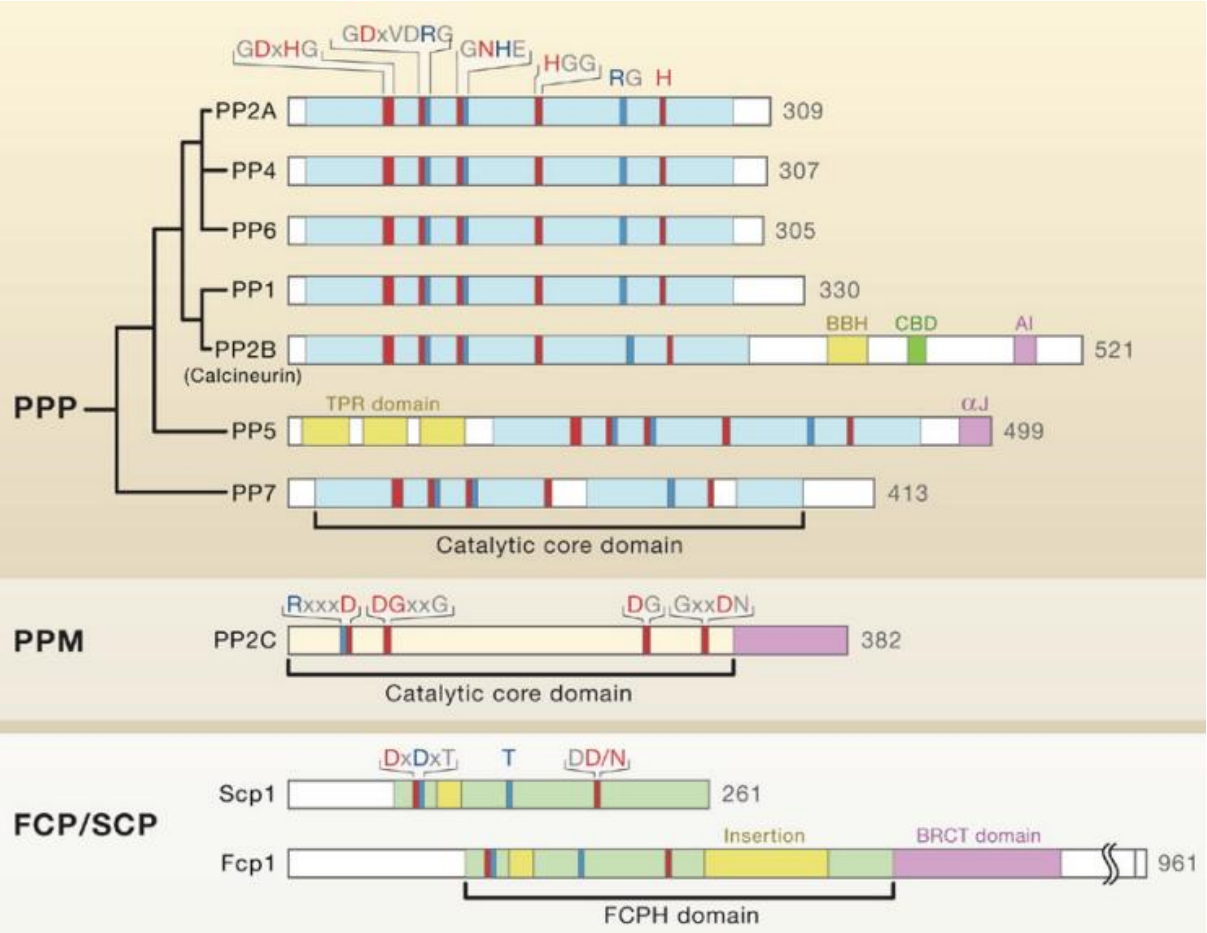


Figure 3 The three families of protein serine/threonine phosphatases according to [93].
The catalytic core domains are indicated below the diagram. Red: residues involved in Me-ion binding, blue: residues involved in phosphate binding. BBH: calcineurin B binding helix, BRCT: breast cancer 1 (BRCA1) C-terminal domain, CTD: carboxy-terminal domain, CBD: Ca^{2+}-calmodulin-binding domain, AI: autoinhibitory sequence, TPR: tetratricopeptide repeat, FCPH: FCP-homology domain. All proteins are from Homo sapiens except for PP7 (for more details see [93]).

So far only the presence of PP1, PP2A, PPM and PTPs was observed in the third domain of life [33]. The PP1 from several Archaea (*S. solfataricus*, *Methanosarcina thermophila* TM-1, *Pyrodictium abyssi* TAG11, *H. volcanii*) were investigated [48,100-104]. In case of *H. volcanii* the PP activity was only verified in cell free crude extract [101]. The PPM from *T. volcanium* is so far the sole archaeal PPM known and harbors dual-specificity, which is uncommon among the other investigated PPMs. Highest activity was reported with pSer/Thr casein followed by pSer/Thr reduced carboxyamidomethylated and maleylated (RCM) lysozyme, pTyr myelin basic protein (MBP), pTyr casein, pTyr RCM lysozyme and pSer/Thr histone as substrates [105]. Furthermore, the crystal structure of PTP from *S. solfataricus* was solved and the specificity towards p-Tyr determined *in vitro* [47]. To our knowledge, only one other archaeal PTP was characterized hitherto, the PTP from *Thermococcus kodakaraensis*. In this analysis, the PTP exhibited p-Tyr (*O*-p-Tyr as well as p-Ser, but no p-Thr activity [106].

1.3. Central carbohydrate metabolism in Sulfolobus species

1.3.1. Glycolysis and gluconeogenesis in *Sulfolobus*

S. solfataricus and *S. acidocaldarius* are both model organisms within the Crenarchaea. *S. solfataricus* is capable to utilize a variety of organic compounds as carbon and energy sources such as sugars (e.g. cellulose, cellobiose, D-glucose, D-xylose), alcohols, aldehydes, amino acids or peptides under aerobic conditions [15]. The glycolysis (ED pathway) and gluconeogenesis (EMP pathway) as well as the degradation pathways for the pentoses D-xylose and D-arabinose were investigated in the past [3,24-28,107]. In comparison, *S. acidocaldarius* can only use few carbon sources [31] and uses also the branched ED pathway for glycolysis and the EMP pathway for gluconeogenesis [108].

For *S. solfataricus* the CCM and especially the ED pathway was extensively studied in the last decade (Figure 4). This pathway is promiscuous and is used to degrade either D-glucose or D-galactose to pyruvate. All enzymes at the beginning of the ED pathway (glucose 1-dehydrogenase (GDH), GAD, 2-keto-3-deoxy(-6-phospho)gluconate (KD(P)G) aldolase, KDG kinase) are promiscuous and can also use D-galactose, galactonate, KDGal and KDPGal, respectively, as substrates [24,57,109-111]. In addition, *S. solfataricus* uses a modified branched ED pathway, consisting of a non- and a semi-phosphorylative branch (spED and npED). The upper part of the EMP pathway (GAP + DHAP → G6P) is used for gluconeogenesis only [24,26,107,112].

The archaeal modified ED pathway is characterized by the missing phosphorylation at the beginning of the pathway, the characteristic intermediate KDG is formed by the activity of GDH and GAD [24,109,113]. In the npED branch, KDG/KDGal is directly cleaved by the bifunctional KD(P)G aldolase [109] forming pyruvate and glyceraldehyde (GA). GA is further converted to 2-phophoglycerate (2-PG) via ferredoxin-dependent GA oxidoreductase (GAOR) [114] and glycerate kinase (GK) [115] and finally via the EMP pathway (i.e. enolase, PyrK) a second molecule of pyruvate is formed. In the spED branch phosphorylation of KDG/KDGal takes place via KDG kinase [111] and the formed KDPG is cleaved by the bifunctional KD(P)G aldolase forming pyruvate and GAP [109]. GAP enters the common lower shunt of the EMP pathway and unique to *S. solfataricus* as well as (hyper)thermophilic Archaea in general, GAP is directly oxidized to 3-PG by an irreversible non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase (GAPN). This enzyme is activated by glucose-1-phosphate (G1P) and represents the major control point in the CCM of *S. solfataricus* and other (hyper)thermopiles like *T. kodakaraensis* and *Thermoproteus tenax* [116-118]. The classical enzyme couple GAPDH and PGK exhibits only gluconeogenetic function (3-PG conversion to 1,3-bisphosphoglycerate (1,3-BPG) and further to GAP) [119]. The next step is the isomerisation of 3-PG

to 2-PG, which can be performed by either iPGAM (SSO0417) or 2,3-bisphosphoglycerate dependent phosphoglycerate mutase (dPGAM, SSO2236). 2-PG is further utilized to PEP by the reversible ENO and PEP is finally transformed to pyruvate via the PyrK and enters the TCA cycle [120].

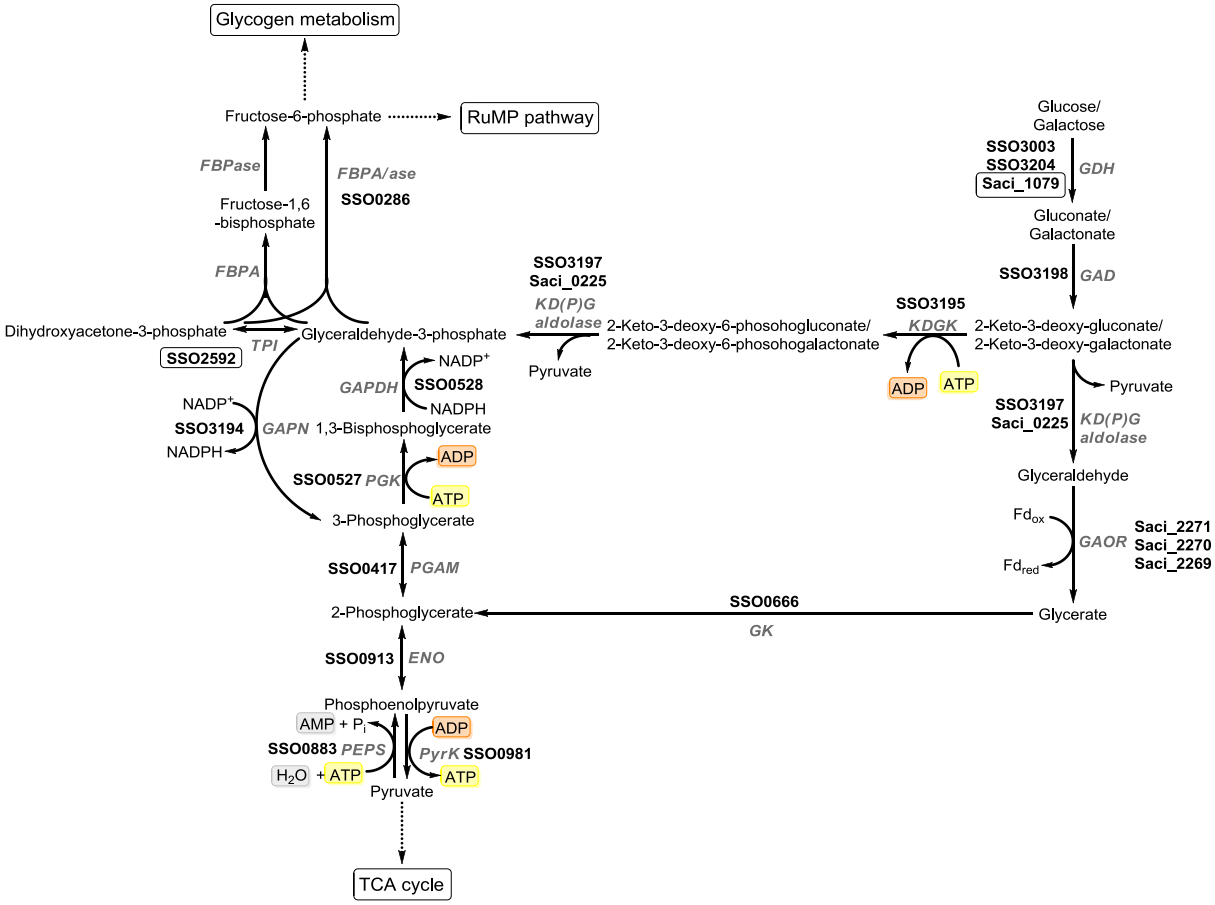


Figure 4 The central carbohydrate metabolism in *S. solfataricus* and *S. acidocaldarius*.
The enzyme names are depicted in italic and gray color. The corresponding gene IDs of characterized enzymes are depicted. Enzymes investigated in this study are marked in a black box. GDH: glucose-1-dehydrogenase, GAD: gluconate dehydratase, KD(P)G aldolase: 2-keto-3-deoxy-(6-phospho)gluconate aldolase, KDGK: 2-keto-3-deoxy-gluconate kinase, GAOR: glyceraldehyde oxidoreductase, GK: glycerate kinase, FBPA/ase: fructose 1,6-bisphosphate aldolase/phosphatase, TPI: triose-3-phosphate isomerase, GAPDH: glyceraldehydes-3-phosphate dehydrogenase, GAPN: non-phosphorylating glyceraldehydes-3-phosphate dehydrogenase, PGK: phosphoglycerate kinase, PGAM: phosphoglycerate mutase, ENO: enolase, PEPS: phosphoenolpyruvate synthetase, PyrK: pyruvate kinase.

Gluconeogenesis is initiated by the action of PEPS which forms PEP from pyruvate and is only active for the anabolic reaction in *S. solfataricus* (Haferkamp *et al.,* 2013 in preperation). PEP is subsequently converted to 2-PG by the reversible ENO and to 3-PG via PGAM. Afterwards 3-PG is transformed to GAP by the gluconeogenic GAPDH/PGK enzyme couple and the TPI catalyzes the reversible isomerisation of GAP to dihydroxyacetone 3-phosphate (DHAP) [119]. The bifunctional FBP aldolase/phosphatase (FBPA/ase) catalyzes the condensation of both triose 3-phosphate intermediates (GAP and DHAP) to fructose 6-phosphate (F6P) [119]. F6P can afterwards enter the ribulose-mono-phosphate (RuMP) pathway, which is used in Archaea to form pentose phosphates

[121]. This pathway has been previously reported in Bacteria for either formaldehyde fixation or detoxification [122]. Apart from the RuMP pathway, F6P can be also converted to glucose 6-phosphate and glucose 1-phosphate to build the carbon storage compound glycogen or the compatible solute trehalose [108].

Biochemical data of the CCM are rather scarce in *S. acidocaldarius*, compared to *S. solfataricus*. Only the KD(P)G aldolase and GAOR from *S. acidocaldarius* were characterized so far. The KD(P)G aldolase revealed activity for the condensation of GA and pyruvate to KDGA, but no activity tests with GAP and pyruvate were conducted [123]. Therefore, if the Saci-KDG aldolase is also promiscuous for KD(P)G and KD(P)Gal still needs to be elucidated. Furthermore, it was shown that the GAOR has highest activity in the presence of GA, followed by formaldehyde, acetaldehyde, isobutyraldehyde, propionaldehyde and GAP. Detailed characterizations were not conducted, however a pH dependency of the substrate specificity and a very high substrate affinity in the µM range (apparent K_m: GA 90 µM, propionaldehyde 30 µM) was reported [114].

By using bioinformatic analysis combined with available biochemical data the ED and EMP pathways of both *Sulfolobus* species can be completely reconstructed.

1.3.2. Pentose degradation in *Sulfolobus*

Apart from the ED and EMP pathway the degradation pathways for the pentoses D-xylose, D-arabinose and L-arabinose were investigated in *Sulfolobus* species (Figure 5) [25,27].

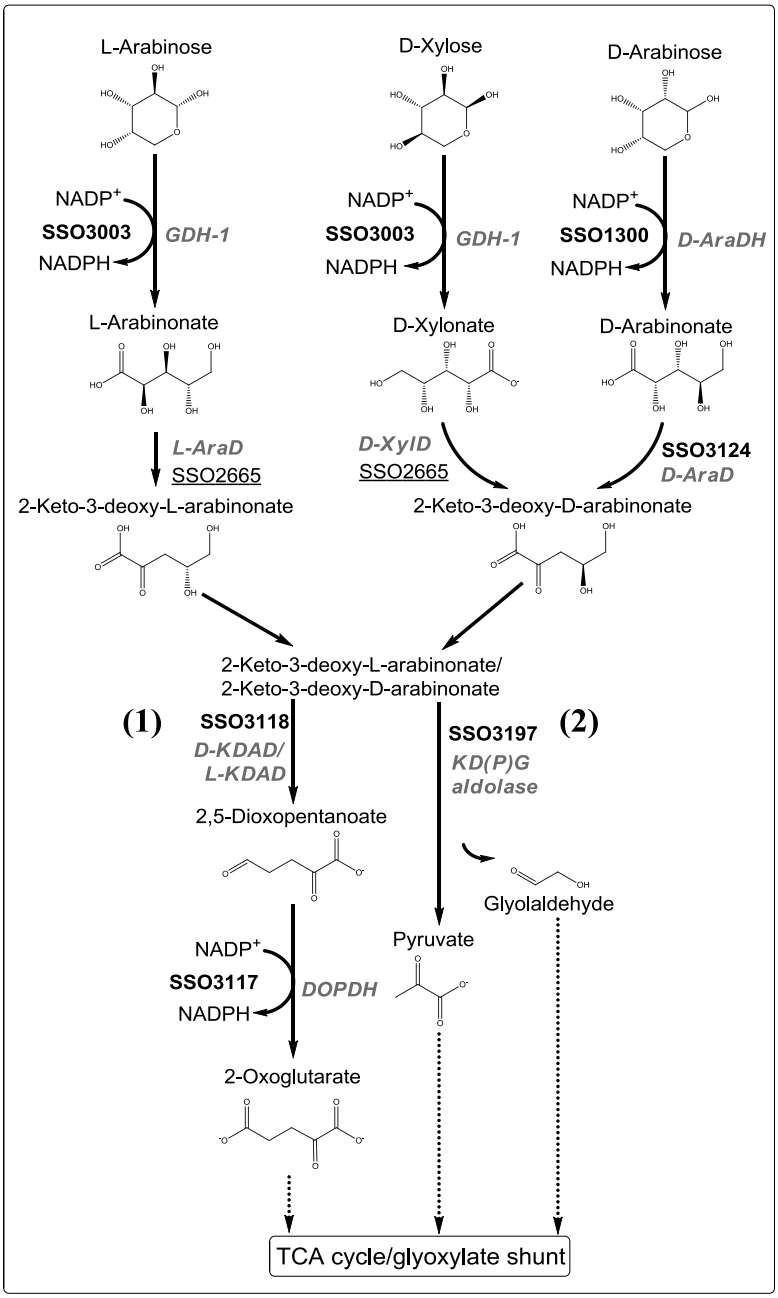


Figure 5 Established pentose degradation pathways in *Sulfolobus*.
The degradation pathways for L-arabinose, D-arabinose and D-xylose via aldolase-independent (1) and aldolase-dependent (2) pathways are depicted. The identified and characterized enzymes participating in the pentose degradation are depicted in bold and enzymes for which the function is only predicted are underlined. GDH-1: glucose 1-dehydrogenase, D-AraDH: D-arabinose dehydrogenase, AraD: arabinoate dehydratase, XylD: xylonate dehydratase, KDAD: 2-keto-3-deoxy-arabinoate dehydratase, KD(P)G aldolase: 2-keto-3-deoxy-(6-phospho)gluconate aldolase, DOPDH: 2,5-dioxopentanoate dehydrogenase.

Brouns and coworkers investigated the degradation pathway of D-arabinose in *S. solfataricus* with great efforts by using a combined transcriptome, proteome approach as well as crude extract measurements and the biochemical investigation of the recombinant enzymes (Figure 5) [27]. Besides the exploration of this pathway, they solved the crystal structure of two novel dehydratases involved in pentose degradation the D-arabinoate dehydratase [124] and D-2-keto-3-deoxy-arabinoate dehydratase [125]. The crude extract measurements revealed the existence of an D-arabinose induced D-arabinose dehydrogenase and D-arabinoate dehydratase as well as the presence of an $NADP^+$-dependent 2,5-dioxopentanoate dehydrogenase (DOPDH). Furthermore, they verified the absence of enzymes, which are required for the conversion of D-arabinose to xylulose-5-phosphate as known from Bacteria or Eukaryotes. The combination of the transcriptome and proteome data enabled the identification of putative candidates of this pathway. 16 genes were highly expressed (log 2 ratio between 1.8 and 4.0) in the D-arabinose grown cells and from these 16 genes were five of the corresponding proteins (SSO3066, SSO1300, SSO3124, SSO3117, SSO3118) detected in the proteome study with an 20-fold difference in the expression level (compared to D-glucose grown cells). SSO3066 was proposed to act as the arabinose-binding protein (AraS) of the arabinose ABC transporter, which was not further investigated. The remaining four proteins were recombinantly expressed and characterized, revealing that SSO1300 is active as D-AraDH and uses L-fucose (K_{cat} 26.8 s-1) or D-arabinose (K_{cat} 23.8 s-1) as substrate and $NADP^+$ as well as NAD^+ as cosubstrate (k_m = 0.04 mM, k_m = 1.25 mM). Furthermore, they proved that SSO3124 acts as D-AraD (K_{cat} 1.8 s-1) and showed that the enzyme is inhibited at higher substrate concentrations. The activity of SSO3118 as D-KDAD was only demonstrated in a coupled enzyme assay with D-AraD using D-arabinoate as substrate. Finally, SSO3117 revealed activity with DOP (K_{cat} 8.6 s-1) as substrate and $NADP^+$ was the preferred cosubstrate compared to NAD^+ [27].

D-xylose degradation in *S. solfataricus* and *S. acidocaldarius* was shown to either proceed in an aldolase-independent pathway via DOPDH, as shown for *H. volcanii* [126], where the end product 2-oxoglutarate enters directly the TCA cycle or by using an aldolase-dependent route via 2-keto-3-deoxy-D-arabinoate aldolase in which pyruvate and glyoxylate are formed. The latter enters the TCA cycle via the glyoxalate shunt and malate is formed [25]. Both pathways were verified by CE measurements, [^{13}C]-labeling studies and enzyme characterizations, similar to the experimental setup used in *H. volcanii* [126]. CE measurements of *S. solfataricus* cells grown on D-xylose revealed D-xylose dehydrogenase, D-xylonate dehydratase, 2-keto-3-deoxy-D-arabinoate (KDA) aldolase, glycolaldehyde-OR, glycolate dehydrogenase, malate synthase as well as DOPDH activities, confirming the presence of both degradation pathways. The labeling experiments performed with D-[1-^{13}C]- and D-[2-^{13}C]-xylose in *S. acidocaldarius* confirmed that both proposed routes are active simultaneously and that the classical pentose phosphate pathway is absent [25].

As discussed above *S. solfataricus* enzymes involved in the CCM were shown to be promiscuous (e.g. SSO3003 GDH-1, SSO3197 GAD, SSO3197 KD(P)G aldolase [24,109,110]). The native promiscuous GDH-1 was purified from *S. solfataricus* cells grown on D-glucose, and activity was reported with D-xylose as substrate [109]. The kinetic properties of GDH-1 reveal that this enzyme is also active in the degradation of the D-xylose and L-arabinose. As candidate for the D-xylonate dehydratase active in D-xylose/L-arabinose degradation SSO2665 was proposed. However, the activity of the recombinant enzyme could not be confirmed and this prediction is only based on sequence similarity (61%) with the characterized SSO-GAD [24]. For L-KDA/D-KDA catalysis the well characterized KD(P)G aldolase SSO3197 was proposed. The enzyme was shown to catalyze the condensation of pyruvate and GA/GAP to KDG/KDPG, respectively, and also the cleavage of KDG/KDPG [127]. Recently, a detailed enzyme characterization with D-KDA and L-KDA, revealed a similar catalytic efficiency of the enzyme for the cleavage of D-KDA and L-KDA as for KD(P)G [128]. Furthermore, SSO3197 is also capable to catalyze the condensation of glycolaldehyde and pyruvate, which is essential for aldolase-dependent pathway (Figure 5) [25]. Next, GAOR was proposed as candidate for glycolaldehyde oxidation to glycolate. The GAOR of *S. acidocaldarius* was shown to possess DCPIP-dependent activity with GA [114] and the respective homolog was identified in *S. solfataricus* (SSO2636, SSO2637 and SSO2639). CE measurements revealed that the artificial DCPIP-dependent GAOR activity with GA as substrate is 7-10 fold higher as the determined $NAD(P)^+$-dependent GADH activity [25]. Finally, it was shown that the recombinant enzyme SSO3187 harbors glycolate dehydrogenase activity and that SSO1334 and SSO1333 possesses malat synthase and isocitrate lyase activity as already demonstrated in the past for the respective *S. acidocaldarius* homologs [129]. Therefore, the previously determined activities of SSO3003 and SSO3197 in combination with the confirmed activities of SSO3187 and SSO1334 and the predicted activities for SSO2665, SSO2636, SSO2637 and SSO2639 (DCPIP-dependent GA-OR) completed the proposed aldolase-dependent and aldolase-independent D-xylose and L-arabinose degradation pathway by Nunn and coworkers. Further on, the previously reported activities of SSO3118 (L-KDAD/D-KDAD) and SSO3117 (DOPDH) were confirmed by CE measurements [27].

1.3.3. Unraveling pathway complexity and regulation

The above mentioned degradation pathways are well investigated and for most of the proposed candidates the suggested functions were verified by biochemical investigations. However, for many of the CCM enzymes several paralogs were identified in the genome (e.g. three KDPG aldolases) and their function still remains to be elucidated for CCM reconstruction. Especially members of the medium chain alcohol/polyol dehydrogenase/reductase super family, ALDH super family and members of the enolase super family are of great interest for the CCM reconstruction. The medium

chain alcohol/polyol dehydrogenase/reductase super family contains sugar dehydrogenases as well as alcohol dehydrogenases, which usually exhibit broad substrate specificity and are the first enzymes of alcohol or sugar degradation pathways. The enolase super family contains mandelate racemase /muconate lactonizing enzymes, which are putative sugar acid dehydratases and are present in all so far identified hexose or pentose degradation pathways. Finally, members of the ALDH super family play a major role in aldehyde conversion and also detoxification as shown for the GAPN and DOPDH in *S. solfataricus*. So far, not all members of the different enzyme super families have been analyzed and therefore their role in the CCM of *Sulfolobus* species is still unclear.

In addition, the glycolysis in *Sulfolobus* species was investigated by several scientists. However, the EMP pathway used for gluconeogenesis is poorly understood. The probable most important gluconeogenic enzyme, which still needs to be investigated in *Sulfolobus* species, might be the TPI. The TPI channels the triose 3-phosphate intermediates DHAP and GAP between the sp-ED branch and the upper part of the EMP pathway. Furthermore, DHAP can not only be used for the gluconeogenesis via FBPA, it is also the first intermediate of the glycerol metabolism. It still needs to be elucidated if the *S. solfataricus* TPI has a role in anabolism or catabolism.

1.3.3.1. Glucose-1-dehydrogenases in Archaea

GDH (EC 1.1.1.47) is the first enzyme of the modified ED-pathway in Archaea. The first report of GDH in the third domain of life was in 1974 by Tomlinson and co-workers. They were able to measure GDH activity in homogenates of *H. saccharovorum* cells [130]. Until now, GDHs from several (hyper)thermophilic Archaea (*S. solfataricus* MT-4 and P2, *Thermoplasma acidophilum*, *T. tenax*, *Picrophilus torridus*) and from the Haloarchaeon *H. mediterranei* were investigated [109,113,131-139]. In addition, the D-AraDH from *S. solfataricus* P2 was characterized and the crystal structure was solved and the D-XylDHs from *Haloarcula marismortui* and *H. volcanii* were investigated [126,140].

In general two different types of GDHs have been described so far: i) pyrroloquinoline-quinone (PQQ)-dependent GDHs (EC 1.1.99.17) and ii) pyridine-nucleotide-dependent GDHs (EC 1.1.1.47.). So far no PQQ-dependent GDHs were identified in any sequenced archaeal genome and therefore these enzymes seem to be restricted to gram-negative bacteria (i.e. *Acinetobacter*) [141,142]. Instead pyridine-nucleotide-dependent GDHs were found in members of all three domains of life. Structural comparison between pyridine-nucleotide-dependent GDHs and PQQ-dependent GDHs revealed that they are not related to each other [141].

All (hyper)thermophilic GDHs identified so far represent homotetramers with exception of the *T. tenax* GDH and the GDH from *H. mediterranei* [139,143], which were described as homodimers [109,131,134-136]. GDHs are members of the medium chain alcohol/polyol dehydrogenase/ reductase branch of the superfamily of pyridine-nucleotide-dependent alcohol/polyol/sugar

dehydrogenases [144]. In contrast, bacterial GDHs (EC 1.1.1.47.) belong to a different enzyme superfamily the short-chain dehydrogenase/reductase superfamily [145].

Two GDH isoenzymes which differ in substrate specificity as well as kinetic properties have been characterized from *S. solfataricus* [109,113,132]. The initial enzyme for glucose conversion via the branched ED pathway in *S. acidocaldarius* is still unknown. In the genome of *S. acidocaldarius* 12 members of the medium chain alcohol/polyol dehydrogenase/reductase branch of the superfamily of pyridine-nucleotide-dependent alcohol/polyol/sugar dehydrogenases were identified and Saci_1079 shows highest similarity (57% amino acid identity) to GDH-1 (SSO3003) of *S. solfataricus*, which exhibits broad substrate specificity.

In order to address pathway promiscuity as well as metabolic complexity in *S. acidocaldarius* aim of this study was to characterize the GDH (Saci_1079) from *S. acidocaldarius* in respect to (co)substrate specificity as well as kinetic parameters.

1.3.3.2. Members of aldehyde dehydrogenase superfamily in *S. solfataricus*

ALDHs play an essential role in intermediary metabolism as well as cellular detoxification by converting aldehydes to the corresponding activated or non-activated acids in all three domains of life. In addition, an important cellular role in osmotic protection and cellular differentiation is discussed [146]. Two distinct structural families have been characterized, which catalyze the oxidation of aldehydes to the respective acids via a two-step mechanism. Both structural families share the initial acylation step including the nucleophilic attack of the catalytic cysteine residue on the aldehyde group, hydride transfer and the formation of the thioacylenzyme intermediate as well as NAD(P)H. However, they differ in the deacylation step. In phosphorylating ALDHs like GAPDH (type 2, TIGR01546) inorganic phosphate serves as acyl acceptor, whereas in non-phosphorylating ALDHs the attack of the thioacylenzyme intermediate can either be performed by a water or CoA molecule leading to the formation of the respective non-activated or CoA-activated acids [147].

The non-phosphorylating ALDH superfamily comprises a group of distantly related enzymes which metabolize a wide variety of endogenous and exogenous aldehydes [148,149]. The superfamily consists of 20 different subfamilies [150] catalyzing the oxidation of aliphatic and/or aromatic aldehydes to the respective acids and therefore have an important role in cellular detoxification and defense systems [151,152]. They are present in all sequenced genomes known today and previous bioinformatic studies revealed that the number of genes encoding ALDHs in Archaea is significantly reduced (1-5 gens) compared to Bacteria (1-26 gens) and Eukaryotes (8-17 gens) [150].

In the genome of *S. solfataricus* five genes encoding paralogs of the ALDH superfamily were identified (SSO1218, *gapN*-1 (SSO1629), *gapN*-2 (SSO1842), *gapN*-3 (SSO3194), *aldhT* (SSO3117))

[150] and so far only two paralogs SSO3194 and SSO3117 have been characterized [27,116]. For SSO3117, the enzymatic characterization revealed broad substrate specificity with significant activities towards DOP, glycolaldehyde and GA [27]. A combined transcriptomic-proteomic approach demonstrated a significant up-regulation in *S. solfataricus* cells grown on D-arabinose and therefore a role as DOPDH in D-arabinose degradation has been proposed [27]. For SSO3194 GAPN activity was demonstrated [116]. GAPN catalyzes the unidirectional oxidation of GAP forming 3-PG and is, due to its allosteric properties (i.e. activation by glucose 1-phosphate), a key enzyme of the spED branch in *S. solfataricus* (lower common shunt of the EMP pathway) and (hyper)thermophilic Archaea in general [116].

The three remaining ALDHs in *S. solfataricus* are annotated as i) methylmalonate semialdehyde dehydrogenase (MSDH, SSO1218) and ii) GAPN isoenzymes (SSO1629, SSO1842). SSO1218 shows high similarity with the investigated MSDH from *B. subtilis* [146,147,153]. The physiological role of MSDHs is the degradation of branched amino acids like valine, leucine and isoleucine as shown in various bacteria like Pseudomonas spp [154]. Instead SSO1629 and SSO1842 are assumed to be involved in the branched ED-pathway as GADH or GAPN isoenzyme.

1.3.3.3. Triose-3-phosphate isomerases

TPI (EC 5.3.1.1) is one of the most studied enzymes. The $(\beta\alpha)_8$-barrel fold, first found in TPI and therefore also known as the TPI barrel fold, represents one of nature's most common protein architecture [155]. TPI has been the subject of several analyses for example concerning the linkage between structure and thermostability [156-158] and detailed enzyme characterization from enzymes of all three domains of life [159-161]. Its glycolytic/gluconeogenetic role is the reason that it is ubiquitous and possesses a conserved function across all three domains of life. Contrary to eukaryal and bacterial TPIs, knowledge about archaeal TPIs is rather scarce. Sequence analyses revealed that archaeal TPIs exhibit only low similarity to their bacterial and eukaryal counterparts (20-25 % identity), whereas eukaryal and bacterial TPIs are much more similar to each other (approximately 40% identity and higher) [162].

TPI catalyzes the interconversion of the thermolabile intermediates DHAP (5.3 min at 80°C) and GAP (1.6 min at 80°C) in the EMP pathway [163]. In mesophilic organisms TPI is generally found to exist as a homo-dimer whereas in hyperthermophiles homo-tetramers were observed. Furthermore, most mesophilic dimeric TPIs comprise approximately 250 amino acid residues per monomer and a significantly reduced size (approximately 25 amino acids) is observed in hyperthermophiles [162,164,165].

Higher oligomerization states of proteins are generally considered as one mean to achieve thermostability [156-158], however, this property is not obligatory for adaptation to high temperature it can also represent a regulatory mean to control catalytic activity as reported for the TPI from the mesophilic Eukaryote *Giardia lamblia* [159]. Specific properties were found for the TPIs from the hyperthermophilic Bacteria *T. maritima* and *T. neapolitana,* their TPIs form bifunctional fusion proteins with the N-terminus covalently linked to the C-terminus of the PGKs [160,166]. For the TPI from *T. tenax* was shown that an equilibrium between inactive dimers and active tetramers exists. Furthermore, interaction with glycerol-1-phosphate dehydrogenase from *T. tenax* resulted in a shift of the equilibrium towards the active tetrameric form [165]. Therefore, TPIs seem to be highly regulated in thermophilic Prokaryotes. Interestingly, for eukaryotic TPIs inhibition by a large set of intermediates (e.g. PEP, 2-PG, 3-PG, 2,3-BPG, 6-phospho-gluconate, G1P, G6P, F6P, FBP, citrate, succinate, ATP) was reported [130,167], which has not investigated so far for (hyper)thermophilic TPIs.

As mentioned before the EMP pathway of *Sulfolobus* species is only poorly investigated and TPI is the key enzyme which regulates the flux of the triose 3-phosphate intermediates (GAP and DHAP) between glycolysis and gluconeogenesis. Due to this key role in *Sulfolobus* species the detailed characterization of TPI, catalyzing the reversible interconversion of GAP to DHAP, is of great interest.

1.4. Aim of this work

The aim of this work is to gain a deeper insight into the complexity as well as regulation of the CCM in *S. solfataricus* and *S. acidocaldarius*. The first aim is to address the role of reversible protein phosphorylation via PKs and PPs in both *Sulfolobus* species. The p-proteomes of *S. solfataricus* and *S. acidocaldarius* will be analyzed by using precursor acquisition independent from ion count (PAcIFIC), which is a gel and enrichment free MS method. *S. solfataricus* cells were grown on either D-glucose (glycolytic growth) or tryptone (peptidolytic growth), harvested in the exponential growth phase and analyzed for the impact of the offered carbon source on protein phosphorylation. The main focus of the thesis will be the reconstruction of the CCM and the possible regulation via reversible protein phosphorylation.

A similar approach was used to investigate the p-proteome of three *S. acidocaldarius* strains. Apart from the parent strain MW001 also strains lacking either one of the two predicted PPs (Saci_0884, Saci_0545) will be analyzed (grown on sucrose and NZ-amine, harvested in exponential growth phase) to investigate the impact of both PPs on protein phosphorylation and to further increase the number of p-proteins. To gain a deeper understanding of their enzymatic properties, both putative PP genes were cloned, heterologously expressed in *E. coli*, purified and biochemical characterized.

The second aim is to identify and characterize players of the CCM in both *Sulfolobus* species *S. solfataricus* and *S. acidocaldarius*. The putative GDH from *S. acidocaldarius* Saci_1079 will be characterized to investigate the initial reaction of glycolysis in *S. acidocaldarius*. The corresponding gene will be cloned, heterologously expressed in *E. coli* and purified. The enzyme will be biochemical characterized in presence of various sugars to determine the physiological substrate and role in the CCM. Furthermore, the ALDHs SSO1218, SSO1629 and SSO1842 from *S. solfataricus* will be characterized with aldehydes as substrates to clarify the physiological function of them. Further on, bioinformatic tools will be used to unravel the role of all members of the ALDH superfamily in *S. solfataricus*. Finally, to further investigate how *S. solfataricus* channels both triose 3-phosphate intermediates GAP and DHAP from glycolysis to gluconeogenesis and *vice versa*, the putative TPI SSO2592 will be investigated. The corresponding gene will be cloned, heterologosly expressed in *E. coli* and purified. Furthermore, the enzyme will be biochemical characterized in the catabolic direction as well as anabolic direction and inhibitor studies with various CCM intermediates will be performed to investigate if crenarchaeal TPIs are similar regulated as eukaryal TPIs.

2. Materials and methods

2.1. Chemicals and plasmids

All chemicals were purchased from Sigma-Aldrich, VWR (St. Louis, MO, U.S.A) or Roche Diagnostics in analytical grade. For heterologous expression, the pET vector systems (pET-11c (Novagen), pET-324 [168]) was used. The DOP was synthesized from 1,4-dibromobutane as previously described [169-171]. The synthesis was performed by Dr. Jolanta Polkowska (AG Schrader, University of Duisburg-Essen, Essen, Germany). Methylmalonate semialdehyde (MMSA) was synthesized as described by Kupiecki and Coon [172] and malonate semialdehyde (MSA) was prepared from ethyl ester diethyl acetal of MSA (Acros Organics, Noisy-le-Sec, France). Both substrates were enzymatically titrated with *B. subtilis* MSDH. The synthesis of MMSA and MSA was performed by Dr. Francoise Talfournier (CNRS-Université de Lorraine, Vandoeuvre-lès-Nancy, France).

2.2. Instruments

Device	Manufacturer
Balance	**KERN EW 4200-2NM (Kern & Sohn GmbH); TE124S (Sartorius AG, Goettingen); TE601 (Sartorius AG, Goettingen)**
Centrifuge	**Table centrifuge Mini Spin plus (Eppendorf AG, Hamburg); Centrifuge SORVALL® RC6 (Kendro Laboratory Products GmbH, Langensbold); Hettich® Universal 320 R (Hettich AG, Bäch)**
Clean bench	**HERAsafe®KSP Class II Bio-safety Cabine (Kendro Laboratory Products GmbH, Langensbold)**
Fermenter	**Infors HT Minifors (Infors AG, Bottmingen)**
Gel-documentation	**ChemDoc Gel-documentationsystem (BioRad Laboratories GmbH, München); Video copy processor: Mitsubishi P91W; Software: Quantity One 4.6.5 (BioRad); Molecular Imager® Versa Doc™ MP 4000 System (Bio Rad Laboratories GmbH, München)**
Heater/Stirring device	**POWERTHERM VARIOMAG® (H+P Labortechnik AG); Stirring device: MR 3001 (Heidolph Instruments GmbH & Co. G,**

	Schwabach)
Membrane-vacuum pump	**LABOPORT Typ:N816.3KN.18 (KNF Neuberger GmbH, Freiburg)**
Microwave	**Type HF1612 (Siemens AG, München)**
pH-Meter	**WTW Series inoLab pH 720 (WTW GmbH, Weilheim); pH-Electrode: SenTIX 81 pH0-14/0-100°C/3mol/KCl (WTW GmbH, Weilheim)**
Photometer	**BioPhotometer Plus (Eppendorf AG, Hamburg); Specord®200 (Analytik Jena, Jena); Compact regulator: JUMO dTRON 308 (Analytik Jena, Jena)**
Shaker	**MULTITRON (INFORS GmbH, Eisbach)**
SDS Polyacrylamid gelelectrophorese system	**Hoefer Mighty small II Model SE250-10A-75 (Hoefer Pharacia Biotech Inc., San Franscisco); Energy supply: Consort E835 (MS Laborgeräte)**
Thermostate	**Thermostat 5320 (Eppendorf AG, Hamburg)**
Ultrasonic processor	**UP 200s (Hielscher Ultrasonics GmbH, Stuttgart, Deutschland)**
Shacker	**MS 2 Minishaker (IKA® Werke GmbH & Co. KG)**
Waterbath	**THERMOMIX® UB (B.Braun Melsungen AG, Melsungen); Thermostation: THERMO-BOY (LAUDA GmbH & Co.KG); Container: Medico 6 (LAUDA GmbH & Co.KG)**
Agarose gel electrophoresis system	**B1A EasyCast™, Owl Seperation Systems (Portsmouth (USA)); Power supply: Consort E835 (MS Laborgeräte)**
Anion exchange column	**UNO Q-12, BioRad laboratories GMBH, Munich, Germany**
Size exclusion column	**HiLoad 26/60 Superdex 200 prep grad, Amersham Biosciences, GE Healthcare, Glattbrugg, Switzerland**
Centrifugal concentrators	**Vivaspin6, 10 kDa MW cut off, Satorius Stedim Biotech, Göttingen, Germany**
Gel documentation system	**ChemiDoc System, BioRad laboratories GMBH, Munich, Germany**
Fast protein liquid chromatography system	**Biologic DuoFlow Pathfinder 20 system (BioRad laboratories GMBH, Munich, Germany). System: F10 work station, MX-1 mixer, 3-Tray rack, AVR7-3 sample inject valve, QuadTec UV/Vis detector with 3 mm PEEK flow cell, system cable 25 (RS-232), BioFrac fraction collector**

2.3. Strains of Escherichia coli and growth conditions

E. coli K12 DH5α strain (Life Technologies, Breda, The Netherlands) was used for the molecular cloning and the strains BL21(DE3) (Merck (Novagen) Darmstadt, Germany), BL21 (DE3)-RIL and ROSETTA(DE3) (Agilent technologies (Stratagene) Waldbronn, Germany) for heterologous over expression of the investigated recombinant proteins. All strains were grown under aerobic conditions in 1–5 mL batch cultures in reaction tubes (up to 5 mL) or Erlenmeyer flasks (0.4-2.0 L) at 37°C with aeration by gyratory shaking (180 rpm). The cultures were grown in Luria Bertani broth (LB) medium (20 g/1 L, Sigma-Aldrich, Taufkirchen, Germany). Antibiotics were added to the cultures according to the respective plasmid: ampicillin (Amp) 100 mg/ml, kanamycin (Kan) 50 mg/ml and chloramphenicol 34 mg/mL (RIL plasmid). Growth of the cultures was monitored at 600 nm (Biophotometer, Eppendorf, Hamburg, Germany).

2.4. Strains of Sulfolobus and growth conditions

2.4.1. *S. solfataricus* P2 for proteomics

Growth of *S. solfataricus* P2 for the p-proteome analysis was performed by Julia Reimann (AG Albers, Max Plank Institute for Terrestrial Microbiology, Marburg, Germany). *S. solfataricus* P2 DSM1617, obtained from DSMZ (Braunschweig, Germany), was grown aerobically as a shaking culture at 76°C in Brocks medium pH 3 [173] supplemented with 0.2% tryptone or 0.2% glucose. Cells for proteomic analysis were grown in two biological replica in a volume of three times 400 ml for each carbon source. At an optical density (OD600) of 0.65 - 0.7, cells were pelleted (6700 x g, 20 min), resuspended in 2 ml 20 mM HEPES pH 8 and snap frozen in liquid nitrogen.

2.4.2 *S. acidocaldarius* for proteomics

Growth of the *S. acidocaldarius* strains MW001, *Δsaci_ptp* and *Δsaci_pp2a* for the p-proteome analysis was performed by Julia Reimann (AG Albers, Max Plank Institute for Terrestrial Microbiology, Marburg, Germany). The markerless in-frame deletion mutants and the uracil auxotrophic background strain *S. acidocaldarius* MW001 [51] were grown at 76 °C in Brock`s basal medium at pH 3.5 [52]. The medium was supplemented with 0.1 % (w/v) NZ-amine, 0.2 % sucrose and 10 µg/ml uracil. The trans-complementation strains of the deletion mutants were grown without uracil since the uracil auxotrophic marker was encoded by the complementation plasmid.

2.4.3 Cultivation of *S. acidocaldarius*

Brock's basal salt medium with final concentrations given in Table 2 and a pH adjusted to 3.5 using concentrated sulfuric acid was chosen as basic growth medium. For the growth of the wild-type strain DSM 639 the medium was supplemented with 0.1% tryptone and 0.2% dextrin. To avoid evaporation at a temperature of 80°C the cultures were filled into long necked Erlenmeyer flasks. Continuous aeration was realized by placement in a shaking incubator at 180 rpm at a temperature of 80°C. The growth was monitored by measurement of the optical density at a wavelength of 600 nm (OD600).

Table 2 Composition of Brock's basal salt medium.

Compound	Concentration [µM]
$(NH_4)_2SO_4$	9840
KH_2PO_4	2060
$MgSO_4 \cdot 7\ H_2O$	284.00
$CaCl_2 \cdot 2\ H_2O$	476.12
$FeCl_3 \cdot 6\ H_2O$	73.99
$Na_2B_4O_7 \cdot 10\ H_2O$	11.80
$MnCl_2 \cdot 4\ H_2O$	9.10
$ZnSO_4 \cdot 7\ H_2O$	0.77
$CuCl_2 \cdot 2\ H_2O$	0.29
$NaMoO_4 \cdot 2\ H_2O$	0.12
$VOSO_4 \cdot 2\ H_2O$	0.15
$CoSO_4 \cdot 7\ H_2O$	0.04

2.5. Molecular biological techniques: working with DNA

2.5.1. Isolation of *S. acidocaldarius* genomic DNA

S. acidocaldarius DSM 639 cells were grown as described in section 2.4.3 Cultivation of S. acidocaldarius. Cells were collected from approximately 2 ml culture at an optical density of about 0.8 by centrifugation at 21,130 rcf (Table centrifuge Mini Spin plus). For the isolation of genomic DNA the DNeasy Blood and Tissue Kit (Qiagen) was used following the instructions of the manufacturer. The cell pellet was resuspended in 180 µl ATL buffer. Afterwards the mixture was incubated at 56°C for 1 h. During this time it was blended several times. To remove any residual RNA a subsequent RNase digestion for 2 min at room temperature was carried out using RNase A (Fermentas). Before loading the sample to the DNeasy spin column it was first mixed with 200 µl of buffer AL and afterwards with 200 µl of 96-100% ethanol. The column was then centrifuged for 1 min at 21,130 rcf. After replacement of the collection tube the column was washed with 500 µl of buffer AW1 applying 12,074 rcf for 1 min and in a following step with 500 µl of buffer AW2. The column was dried for 3 min at 21,130 rcf. Elution of the DNA was carried out by applying of 100 µl of nuclease free water onto the column and centrifugation into a clean collection tube for 1 min at 12,074 rcf. The DNA quantity was measured by ultraviolet absorbance spectrophotometry (BioPhotometer). At 260 nm, an absorbance (A_{260}) of 1.0 corresponds to 50 µg of double stranded DNA per mL. The ratio $OD_{260/280}$ (BioPhotometer) was used for qualitative analysis. For a pure DNA sample the ratio of absorbance (A_{260}/A_{280}) is 1.8-2.0. Other ratios indicate contaminations by either proteins or phenol of the DNA sample.

2.5.2. Polymerase chain reaction (PCR)

Desired DNA fragments for cloning were amplified using PCR mutagenesis.

Therefore a reaction mix (50 µl total volume) was prepared containing 2.5 µl KOD buffer #1, 3 µl $MgCl_2$ (1.5 mM), 5 µl dNTPs (0.2 mM each) and 0.02 U of KOD was set to 0.02 U (0.4 µl KOD Polymerase suspension) (all Novagen® KOD Hot Start Polymerase Kit, Merck4Biosciences, Germany). Approximately 150 ng of *S. acidocaldarius* DSM 639 genomic DNA (gDNA) served as DNA template. The primers used for PCR are indicated in Table 3 below. The underlined sequences indicate restriction sites that were introduced and the respective restriction enzymes are given.

The protocol of PCR amplification used for all genes in this study is shown in Table 4 below. The PCR outcome was inspected by agarose gel electrophoresis (see section 2.5.3). Depending on the experiment either the whole PCR product was loaded to the gel or a small aliquot.

Table 3 Sequences of primers used in PCR mutagenesis.
The sections underlined indicate the introduced restriction sites.

Primer	Endonucleases	Sequence (5'-3')	Plasmid
FP SSO1218[(1)]	*Nco*I	GGGGGAATTCCATGGGTAAGGTATTGG	pET324
RP SSO1218[(1)]	*Bam*HI	GGCGGGGATCCTCACCACCTGGTTATTAC TACC	
FP SSO1629[(1)]	*Nde*I	GGGGGAATTCATATGATGAGTCAAGAGG	pET11c
RP SSO1629[(1)]	*Bam*HI	GGCGGGGATCCTCAAAGAAGCGTTATTGC	
FP SSO1842[(1)]	*Nde*I	GGGGGAATTCATATGATGAGCTTAGAGG	pET11c
RP SSO1842[(1)]	*Bam*HI	GGCGGGGATCCTCAAAGTAATGTGATAGC	
FP SSO2592	*Nde*I	GGGGGAATTCATATGAACCTCCTATTATTAT AATAAAC	pET11c
RP SSO2592	*Bam*HI	GGCGGGATCCTCACGAAGAGATCGCCCTCAA TGC	
FP Saci_0545	*Nde*I	AAACATATGATGTATTGGGTAAAAAAGCATG TC	pET15b
RP Saci_0545	*Bam*HI	AAAGGATCCTCATAAAATCTTCCATTTATCTTT CAT	
FP Saci_0884[(2)]	*Nco*I	GGGCCATGGCTAACATTGAAGAAACGTATGA G	pET DueT-1
RP Saci_0884[(2)]	*Bam*HI	GGGGGATCCTTAGTGGTGATGATGGTGATGT ACTATCTCTTCTATTAGTTGATCGTTCAC	
FP Saci_1079	*Nco*I	GCGCCATGGATGAGAGCTATAATAATAAGAC CTCCTAATG	pET15b
RP Saci_1079	*Xho*I	CGCGCGCTCGAGTTAGTCCCAAATTATTCTAA CTTTTATTTC	

[(1)] The genes SSO1218, SSO1629 and SSO1842 were cloned by Theresa Kouril (AG Siebers, University Duisburg-Essen, Germany).
[(2)] The gene Saci_0884 was cloned by Julia Reimann (AG Albers, Max Plank Institute for Terrestrial Microbiology, Marburg, Germany).

Table 4 PCR protocol including duration, temperature of each phase and cycle numbers.

Phase	Temperature [° C]	Time period [s]	Cycle numbers
Start	95	180	1
Denaturation	95	30	30
Annealing	55	30	
Extension	72	60	
Termination	72	50	1
Storage	8	infinite	1

2.5.3. Agarose gel electrophoresis

Agarose gel electrophoresis was used for diverse applications, such as the estimation of DNA yield and purity, determination of the size of PCR reaction products and DNA gel extraction. 1% Agarose gels were prepared in TBE-buffer (90 mM Tris-Acetat, 90 mM boric acid, 2 mM EDTA) and were employed for all applications. The DNA samples were mixed with loading buffer (6 x: 0.21% bromophenolblue, 0.21% Orange, 50% glycerol and 0.2 M EDTA (pH 8)) and applied to the agarose

gel. Depending on the designated separation quality of DNA bands and the chosen size of the gel chamber (distance of the electrodes) voltage was adjusted at 60-100 V. The gel run was performed at room temperature. DNA samples were run in parallel with a size marker (GeneRulerTM 1 kb DNA Ladder, MBI Fermentas GmbH). For fluorescent visualisation of DNA under UV light, the gels were stained with ethidium bromide (200 µL/L, Sigma-Aldrich). The results were documented under UV light using the ChemiDoc-Gel Documentation System (BioRad).

2.5.4. Plasmid preparation

Extraction of the plasmid-DNA from an *E. coli* DH5α over night culture was achieved by using the GeneJET™ Plasmid Miniprep Kit (MBI Fermentas GmbH), following the manufacturer's instruction.

2.5.5. Purification of DNA fragments from agarose gels

Extraction and purification of DNA fragments from agarose gels and of PCR fragments were achieved by using the Wizard® SV Gel and PCR clean up system (Promega GmbH), following the manufacturer's instruction.

2.5.6. Restriction of DNA

Restriction of PCR products was performed with 0.2 µg of dsDNA or 1 µg plasmid DNA. DNA was incubated with 1 U of the respective endonuclease shown in Table 3 in a total volume of 30 µl. The reaction was performed at 37°C for 30min. For the isolation of restriction products, samples were separated via agarose gel electrophoresis and DNA was purified using the Promega PCR clean up system according to the manufacturer's instruction (Promega).

2.5.7. Ligation

For cloning of the restricted and purified PCR products comprising the desired genes were ligated into the restricted vector (Table 3). The ligation was performed in a molar ratio of gene to plasmid of 1:3. The amount of used plasmid DNA was between 50 ng to 100 ng. The reaction mixture containing the PCR product and plasmid DNA was incubated at 45°C for 5 min and afterwards 1 µL of the T4-ligase buffer and 2.5 U T4-ligase (Fermentas) were added. The ligation reaction was carried out overnight at 16°C. Afterwards, T4-ligase was inactivated by incubation at 70°C for 5min and the reaction mixture was stored at -20°C or immediately used for transformation into *E. coli* DH5α.

2.5.8. Transformation

The *E. coli* strain K12 DH5α was used for storage and amplification of plasmid DNA, and for overexpression different *E. coli* expression strain were used (Table 5).

Table 5 Expression of *S. solfataricus* P2 and *S. acidocaldarius* DSM 639 target proteins.

ORF ID	Enzyme	Plasmid	Host strain
SSO1218	MSDH	pET324^{Amp}	BL21-CodonPlus (DE3) RIL^{Cam}
SSO1629	SSADH-I	pET11c^{Amp}	BL21-CodonPlus (DE3) RIL^{Cam}
SSO1842	SSADH-II	pET11c^{Amp}	BL21-CodonPlus (DE3) RIL^{Cam}
SSO2592	TPI	pET11c^{Amp}	ROSETTA (DE3)^{Cam}
Saci_0545	PTP	pET15b^{Amp}	ROSETTA (DE3)^{Cam}
Saci_0884	PP2Ac	pET DueT-1^{Amp}	ROSETTA (DE3)^{Cam}
Saci_1079	GDH	pET15b^{Kan}	BL21-CodonPlus (DE3) RIL^{Cam}

For transformation 70 µL of chemical competent cells *E. coli* DH5α or of the respective expression strains were gently mixed with 2.5 µL plasmid DNA or 10 µL of the ligation reaction mixture, and incubated 30 min on ice. Afterwards, cells were incubated for 90 seconds at 42°C and afterwards 400 µL of sterile LB-medium was added and the cells were incubated in a rotary shaker at 37°C and 180 rpm for 55 min. The cell suspension was then completely transferred to 5 mL LB-medium contained 100 µg/mL Amp in the case of *E. coli* DH5α and 50 µg/mL Cam in case of *E. coli* ROSETTA or BL21-CodonPlus (DE3) RIL or streaked onto LB agar plates containing 100 µg/ml Amp or Kan. Finally, the cells were incubated at 37°C and 180 rpm in the shaker overnight.

2.5.9. Restriction analysis of positive colonies

In order to identify positive clones plasmid DNA was prepared from 5 ml overnight cultures derived from culture plates after transformation. A total number of about 15 colonies was usually chosen for this purpose. The plasmid vector DNA was restricted as described in section 2.5.6, using the respective restriction enzymes (Table 3).

2.5.10. DNA sequencing

DNA sequencing was performed to verify the sequence information of the cloned gene. Automated DNA sequencing was performed by AGOWA (Berlin).

2.5.2. Disruption of the Saci_0545 and Saci_0884 genes in *S. acidocaldarius*

Construction of the *Δsaci_ptp* and *Δsaci_pp2a* gene disruption mutants was performed by Julia Reimann (AG Albers, Max Plank Institute for Terrestrial Microbiology, Marburg, Germany). Plasmids for deletion mutant construction were cloned using the PCR product of the upstream- and downstream-region of the gene of interest (Table 6), overlap PCR was performed to fuse both fragments and the overlap product was ligated into the gene targeting plasmid pSVA406. The constructed plasmids were methylated using *E. coli* ER1821 and transformed into the background strain *S. acidocaldarius* MW001 (uracil auxotrophic strain). The transformants were first selected on gelrite plates lacking uracil and then counterselected on plates with uracil and 5-FOA. The markerless in-frame deletion mutants of both phosphatase genes (saci0545 and saci0884), designated *Δsaci_ptp* and *Δsaci_pp2a*, respectively, were identified by colony PCR and confirmed by sequencing (MWG Operon, Ebersberg, Germany).

Table 6 Primer sets for the construction of deletion mutants in *S. acidocaldarius*.
FP: forwards primer, RP: revers primer, ol: overlap.

ORF ID	Enzyme	Sequence (5´ - 3´)	Purpose
Saci_0545	**PTP**	**Primers for pSVA1016**	
		CCCCGGATCCGTTTTCCGATTAGAACTATT	Upstr FP *Bam*HI
		CATAAAATCTTCCATGTCTTCATCACTCTGAAGAT	Upstr RP ol
		CAGAGTGATGAAGACATGGAAGATTTTATGATAGA	Downstr FP ol
		CCAAACTGCAGAGCCTTATGAATTAAGCTC	Downstr RP *Pst*I
		AACTCATAGCGTGAGATCC	Control primer FP
		ATCCAGCTAATGCATGTTCC	Control primer RP
Saci_0884	**PP2Ac**	**Primers for pSVA1017**	
		CCCCGGATCCGTATTTTCTTAAACCTTC	Upstr FP *Bam*HI
		TGTCTGCTATACTATAATGTTCACAATATTGTGGT	Upstr RP ol
		AATATTGTGAACATTATAGTATAGCAGACAAAAAA	Downstr FP ol
		CCCCCTGCAGCCATAACTTATCCTTAAT	Downstr RP *Pst*I
		TTCCTGCCCACTGATATTCC	Control primer FP
		CGGTTGGTTAAATCAATTAG	Control primer RP

2.6. Molecular biological techniques: working with RNA

2.6.1. RNA isolation and sample preparation for RNA-seq analysis

RNA isolation and sample preparation for deep sequencing analysis were performed by Fabian Amman (Institute for Theoretical Chemistry, University of Vienna, Vienna, Austria). *S. acidocaldarius* MW001, *Δsaci_ptp* and *Δsaci_pp2a* were inoculated in triplicate in Brock medium supplemented with 0.1 % NZ-amine, 0.2 % sucrose and 10 µg/ml uracil. Cultures were grown to reach an OD600 of 0.6 (exponential growth phase) and subsequently samples of the three independent cultures of each strain were pooled in equal amounts to generate one mixed sample per strain. Cell pellets of 10 ml culture volume were taken for RNA isolation. TRIzol reagent (Invitrogen) was used for total RNA isolation following the instructions of the manufacturer. Residual chromosomal DNA present in RNA samples was removed by RNase-free DNase I (Roche) treatment for 2 h at 37°C. DNA-free RNA samples were confirmed by PCR amplification using saci0574 (secY) primer pairs. Prior cDNA synthesis DNA-free RNA samples were fragmented to achieve a molecule size range of 50 to 500 nucleotides. Thus, 6 µg of DNA-free RNA samples were mixed with 4 µL of 5X fragmentation buffer (200 mM Tris-acetate, pH 8.2; 500 mM potassium acetate; 150 mM magnesium acetate). The reaction mix was filled up with DEPC H_2O up to 20 µL, incubated at 95°C for 2.5 min and immediately transferred on ice. The reaction was then cleaned up by using sephadex columns (illustra MicroSpin™ G-25, GE Healthcare). RNA fragmentation range was checked by polyacrylamide gel electrophoresis. Six hundred ng of fragmented RNA was used for cDNA synthesis via the SuperScript Double-stranded cDNA Synthesis kit (Invitrogen) following the manufacturer's instruction. The reaction was cleaned by using phenol:chloroform:isoamyl alcohol and cDNA samples were then precipitated by using 3M NaOAc/ 100% ethanol and incubated over night at -20°C. Finally, cDNA samples were washed with 500 µL 70% Ethanol. Five ng of cDNA were used as starting material for the generation of single-end sequencing libraries with the NEBnext DNA sample preparation kit, following the manufacturer's instruction. The DNA was ligated to adaptors, which carried a unique four letter barcode sequence. DNA fragments of 150-500 bp were selected for sequencing. To retain strand specificity the DNA was treated with Uracil DNA-Glycosilase (UDGase) prior to the PCR amplification. The sequencing was done by an Illumina Genome Analyzer IIx, multiplexed together with a total of eight samples. The reads were partitioned according to their barcode, stripped from adapter sequences as well as barcodes and mapped, using default settings, to the *S. acidocaldarius* DSM 639 (NC_007181) reference genome, with segemehl [174]

2.6.2. Bioinformatic analysis of RNA-seq data

Analysis of the RNA deep sequencing data was performed by Alvaro Orell (AG Albers, Max Plank Institut for Terrestrial Microbiology, Marburg, Germany) and Fabian Amman (Institute for Theoretical Chemistry, University of Vienna, Vienna, Austria). For each protein coding and ncRNA gene, according to the NCBI database, the associated reads were counted. The overall Illumina sequencing represented 26.6 %, 30.2 %, 26.5 % of genome coverage by the mapped reads for MW001, *Δsaci_ptp* and *Δsaci_pp2a*, respectively. To identify differentially expressed genes between the sample conditions, we used DESeq (version 1.5) [175],which allows to estimate the dispersion within one condition, also without having replicas at ones disposal. This estimation is based on the assumption that only a few genes are truly differential expressed; hence variance across different conditions can be seen as a too conservative estimation of the variance within one condition. From the estimated expected variance and the observed fold change in read counts, the statistical significance of altered RNA abundance can be calculated for every gene. Genes with a p-value below 0.054 were considered to be significantly altered in their expression.

2.6.3. Reverse transcription quantitative real time PCR

Reverse transcription quantitative real time PCR was performed by Julia Reimann (AG Albers, Max Plank Institut for Terrestrial Microbiology, Marburg, Germany). First Strand cDNA Synthesis Kit (Fermentas) was used for cDNA synthesis according to the manufacturer's instructions. The synthesis was thus performed using random hexamer primers and 1 µg of DNA-free RNA as template. Quantitative PCR (qPCR) analysis was carried out using Maxima SYBR Green/ROX qPCR Master Mix (Fermentas) based on the Sybr green detection system (Real-Time 7300 PCR machine; Applied Biosystems). The efficiency of each primer pair was calculated from the average slope of a linear regression curve, which resulted from qPCRs using a 10-fold dilution series (10 pg - 10 ng) of *S. acidocaldarius* chromosomal DNA as template. Cq values (quantification cycle) were automatically determined by Real-Time 7300 PCR software (Applied Biosystems) after 40 cycles. Cq values of each transcript of interest were standardized to the Cq value of the housekeeping gene saci0574 (secY) [55]. qPCR reactions with DNA-free RNA samples as template were performed as control. At least 3 biological replicates of each assessed condition and 2 technical replicates per qPCR reaction were performed.

2.7. Biochemical methods

2.7.1 Determination of protein concentration

Protein concentrations were determined (modified according to Bradford, 1976) using the BioRad Protein Assay. For this purpose, a calibration curve was used with bovine serum albumin as standard (BSA; standard II, BioRad, concentrations: 0 µg/mL, 1 µg/mL, 2.5 µg/mL, 5 µg/mL, 10 µg/mL, 15 µg/mL and 25 µg/mL).
5, 10 or 20 µL sample, respectively, were applied and a. bidest was added to a final volume of 600 µL. Subsequently, 400 µL Bradford-reagent was added to standards as well as samples and incubated for 15 min at room temperature. Absorbance was determined photometrically at 600 nm. All measurements were performed as triplicates.

2.7.2. Heterologous expression and protein purification

2.7.2.1. Aldehyde dehydrogenases from *S. solfataricus* P2

Heterologous expression of SSO1218, SSO1629 and SSO1842 were performed using *E. coli* BL21-CodonPlus (DE3)-RIL as expression strain. The protein overexpression was induced after reaching an optical density of 0.6 -0.8 by the addition of isopropyl-β-thiogalactopyranoside (IPTG, final concentration 1 mM). The cells were harvested by centrifugation at 4°C, 17000 rpm for 20 min. The cells were resuspended into 0.1 M HEPES/KOH buffer, pH 6.5 at 70°C (1 g in 3 mL buffer). Cell disruption was performed via French press (three times, 1200 psi) and afterwards cell debris was removed by centrifugation at 4°C, 21130 rcf for 45 min to remove the insoluble proteins. For enrichment of the recombinant enzymes, the resulting *E. coli* crude extract was diluted 1:1 with 0.1 M HEPES/KOH buffer, pH 6.5 at 70°C and subjected to a heat precipitation for 20 min at 80°C. After heat precipitation, the samples were cleared by centrifugation to remove the denaturated heat instable *E. coli* proteins (21130 rcf for 30 min at 4° C). The supernatant was dialyzed overnight against 20 mM HEPES/KOH (pH 6.5, 70°C), subjected to anion exchange chromatography (UNO Q-12, Bio-Rad Laboratories, pre-equilibrated in 20 mM HEPES/KOH (pH 6.5, 70°C)), and eluted with a linear salt gradient from 0 to 1 M NaCl in buffer. Fractions containing the recombinant enzymes (checked by SDS-PAGE and activity tests) were pooled and concentrated via centrifugal concentrators (Vivaspin6, 10 kDa MW cut off, Sartorius Stedim Biotech). Afterwards, the sample was dialyzed overnight against

50 mM HEPES/KOH, 300 mM KCl (pH 6.5, 70°C) and applied to size exclusion chromatography (HiLoad 26/60 Superdex 200 prep grade, Amersham Biosciences, pre-equilibrated in the same buffer (50 mM HEPES/KOH, 300 mM KCl (pH 6.5, 70°C)). The fractions containing the recombinant enzymes were analyzed via SDS-PAGE and enzyme activity tests, pooled and stored at 4°C.

2.7.2.2. Triosephosphate isomerase from *S. solfataricus* P2

Heterologous expression of SSO2592 was performed by using *E. coli* ROSETTA (DE3) as expression strain. The protein overexpression was induced after reaching an optical density of 0.6 -0.8 by the addition of IPTG (final concentration 1 mM). The cells were harvested by centrifugation at 4°C, 17000 rpm for 20 min and stored at -80°C. The cell pellets were resuspended in 0.1 M Tris/HCl, 5 mM dithiothreitol (DTT) buffer, pH 6.5 at 70°C (1 g in 3 mL buffer). Cell disruption was performed via French press (three times, 1200 psi) and afterwards cell debris was removed by centrifugation at 4°C, 21130 rcf for 45 min. For enrichment of the recombinant enzymes, the resulting *E. coli* crude extract was diluted 1:1 with 0.1 M Tris/HCl, 5 mM DTT buffer, pH 6.5 at 70°C and subjected to a heat precipitation for 20 min at 90°C. After heat precipitation, the samples were cleared by centrifugation to remove the denaturated heat instable *E. coli* proteins (21130 rcf for 30 min at 4° C). The supernatant was dialyzed overnight against 20 mM Tris/HCl, 5 mM DTT (pH 6.5, 70°C), subjected to anion exchange chromatography (UNO Q-12, Bio-Rad Laboratories, pre-equilibrated in 20 mM Tris/HCl, 5 mM DTT (pH 6.5, 70°C)), and eluted with a linear salt gradient from 0 to 1 M NaCl in buffer. Fractions containing the recombinant enzymes (checked by SDS-PAGE and activity tests) were pooled and concentrated via centrifugal concentrators (Vivaspin6, 10 kDa MW cut off, Sartorius Stedim Biotech). Afterwards, the sample was dialyzed overnight against 50 mM Tris/HCl, 5 mM DTT, 300 mM KCl (pH 6.5, 70°C) and applied to size exclusion chromatography (HiLoad 26/60 Superdex 200 prep grade, Amersham Biosciences, pre-equilibrated in the same buffer (50 mM Tris/HCl, 5 mM DTT, 300 mM KCl (pH 6.5, 70°C)). The fractions containing the recombinant enzymes were analyzed via SDS-PAGE and enzyme activity tests, pooled and stored at 4°C.

2.7.2.3. Protein phosphatases from *S. acidocaldarius* DSM 639

Heterologous expression of Saci_0545 and Saci_0884 were performed by using *E. coli* ROSETTA (DE3) as expression strain. Both proteins had a hexa-His-tag at the N-terminus. The protein overexpression was induced after reaching an optical density of 0.6 -0.8 by the addition of IPTG (final concentration 1 mM). The cells were harvested by centrifugation at 4°C, 21130 rcf for 20 min and stored at -80°C.

The cell pellets were resuspended in 50 mM NaH_2PO_4, 300 mM NaCl buffer, pH 6.5 at room temperature (1 g in 3 mL buffer). Cell disruption was performed via French press (three times, 1200 psi) and afterwards cell debris was removed by centrifugation at 4°C, 21130 rcf for 45 min. For enrichment of the recombinant enzymes, the resulting *E. coli* crude extract was diluted 1:1 with 50 mM NaH_2PO_4, 300 mM NaCl buffer, pH 6.5 at room temperature and subjected to a heat precipitation for 20 min at 80°C. After heat precipitation, the sample was cleared by centrifugation to remove the denaturated heat instable *E. coli* proteins (21130 rcf for 30 min at 4° C) and subjected to immobilized metal affinity chromatography (Ni-TED 2000, Machery-Nagel, pre-equilibrated in 50 mM NaH_2PO_4, 300 mM NaCl (pH 8.0, room temperature)), and eluted in 50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imidiazole (pH 8.0, room temperature). Fractions containing the recombinant enzymes (analyzed by SDS-PAGE, enzyme activity) were pooled and concentrated via centrifugal concentrators (Vivaspin6, 10 kDa MW cut off, Sartorius Stedim Biotech).

Saci_0884 was further purified by ion exchange chromatography and size exclusion chromatography. Briefly, Saci_0884 was dialyzed overnight against 20 mM Tris/HCl (pH 6.5, 70°C), subjected to ion exchange chromatography (UNO Q-17, Bio-Rad Laboratories, pre-equilibrated in 20 mM Tris/HCl (pH 6.5, 70°C)), and eluted with a linear salt gradient from 0 to 1 M NaCl. Afterwards, the sample was dialyzed overnight against 50 mM Tris/HCl, 300 mM KCl (pH 6.5, 70°C) and applied to size exclusion chromatography (HiLoad 26/60 Superdex 200 prep grade, Amersham Biosciences, pre-equilibrated in 50 mM Tris/HCl, 300 mM KCl (pH 6.5, 70°C)). The fractions containing the recombinant enzymes were analyzed via SDS-PAGE and enzyme activity tests, pooled and stored at 4°C.

2.7.2.4. Glucose-1-dehydrogenase from *S. acidocaldarius* DSM 639

Heterologous expression of Saci_1079 was performed by using *E. coli* BL21-CodonPlus (DE3)-RIL as expression strain. The protein overexpression was induced after reaching an optical density of 0.6 - 0.8 by the addition of IPTG (final concentration 1 mM). The cells were harvested by centrifugation at 4°C, 21130 rcf for 20 min and stored at -80°C. The cell pellets were resuspended in 0.1 M Tris/HCl, 5 mM $MgCl_2$ buffer, pH 6.5 at 70°C (1 g in 3 mL buffer). Cell disruption was performed via French press (three times, 1200 psi) and afterwards cell debris was removed by centrifugation at 4°C, 21130 rcf for 45 min. For enrichment of the recombinant enzymes, the resulting *E. coli* crude extract was diluted 1:1 with 0.1 M Tris/HCl, 5 mM $MgCl_2$ buffer, pH 6.5 at 70°C and subjected to a heat precipitation for 20 min at 70°C. After heat precipitation, the samples were cleared by centrifugation to remove the denaturated heat instable *E. coli* proteins (21130 rcf for 30 min at 4° C). The supernatant was dialyzed overnight against 20 mM Tris/HCl, 5 mM $MgCl_2$ (pH 6.5, 70°C), subjected to anion exchange chromatography (UNO Q-12, Bio-Rad Laboratories, pre-equilibrated in 20 mM Tris/HCl, 5 mM $MgCl_2$

(pH 6.5, 70°C)), and eluted with a linear salt gradient from 0 to 1 M NaCl in buffer. Fractions containing the recombinant enzymes (checked by SDS-PAGE and activity tests) were pooled and concentrated via centrifugal concentrators (Vivaspin6, 10 kDa MW cut off, Sartorius Stedim Biotech). Afterwards, the sample was dialyzed overnight against 50 mM Tris/HCl, 5 mM $MgCl_2$, 300 mM KCl (pH 6.5, 70°C) and applied to size exclusion chromatography (HiLoad 26/60 Superdex 200 prep grade, Amersham Biosciences, pre-equilibrated in the same buffer (50 mM Tris/HCl, 5 mM $MgCl_2$, 300 mM KCl (pH 6.5, 70°C)). The fractions containing the recombinant enzymes were analyzed via SDS-PAGE and enzyme activity tests, pooled and stored at 4°C.

2.7.2.5. Fructose-1,6-bisphosphate aldolase/phosphatase from *S. solfataricus*

The FPBA/ase from *S. solfataricus* (SSO0286) was required as auxiliary enzyme for the anabolic TPI assay. The *E. coli* Rosetta (DE3) cells containing the FBPA/ase plasmid were grown until the optical density of 0.6 was reached and over-expression was induced by adding 1mM IPTG to the medium. After 3 h of incubation, the cells were harvested by centrifugation (6000 x g, 20 min, 4°C) and resuspended (1:3) in 100 mM Tris/KOH buffer (pH 7.2, room temperature) and cell disruption was carried out by sonication (5 min, 50 % amplitude). After centrifugation (21130 rcf, 45 min, 4°C), the soluble cell extract was heat treated for 20 min at 80°C and the denaturated host proteins were removed by centrifugation (21130 rcf, 30 min, 4°C).

2.7.3. Determination of the native molecular weight

For the determination of the native molecular weight of the recombinant enzymes, 500 µL protein solution with a concentration of 2 mg/mL was subjected to size exclusion chromatography (HiLoad 26/60 Superdex™ 200 prep grade, Amersham Biosciences, pre-equilibrated in 50 mM HEPES/KOH, 300 mM KCl (pH 6.5 , 70 °C). The calibration curve was generated using the gel filtration calibration kit HMW (43-669 kDa, GE Healthcare).

2.7.4 Sodium-dodecylsulfate – polyacrylamide-gel-electrophoresis (SDS-PAGE)

Sodium-dodecyl-sulfate-polyacrylamide-gel-electrophoresis (SDS-PAGE) allows to seperate proteins according to their molecular weight. SDS-PAGE is applied to control the purity of protein fractions during purification and to determine the molecular weight of protein subunits.

The detergent SDS denaturates proteins and provides them with a uniform negative charge, therefore the proteins can be separated by a polyacrylamidegel-matrix exclusively according to their size and independent of their native charge and form.

For analyzing the recombinant enzymes a 12% SDS-polyacrylamide-gel was used. The gel is made up of a lower separating gel, which is covered with a stacking gel. The separating gel is composed of 12% (m/v) acrylamide-bisacrylamide (30%), 375 mM Tris (pH 8.8 (RT)), 0.1% (m/v) SDS, 0.67% (m/v) ammoniumpersulfate (APS) (100 mg/mL) and 0.067% (v/v) N,N,N,N-Tetramethylethylenediamine (TEMED), whereas the stacking gel contains 4% (m/v) acrylamide-bisacrylamide (30%), 125 mM Tris (pH 6.5 (room temperature)), 0.1 % (m/v) SDS, 0.45% (m/v) APS (100 mg/mL) and 0.15% (v/v) TEMED. First, the separation gel was poured. After a polymerisation time of about 20-30 min, the stacking gel was poured at top of the separating gel (polymerisation time ca. 15 min) and a comb was inserted to form the gel-pockets. Aliquots of 10 µL from each sample were mixed with 4 µL sample buffer (62.5 mM Tris-HCl, pH 6.8, 10% glycerol, 2% SDS, 5% DTT and 0.005% bromine phenol blue). Afterwards the samples were incubated for denaturation at 94°C for 4 min and 10 µL of the samples and 5 µL protein marker (PageRuler™ (10-200 kDa) Fermentas) were applied to the gel. SDS-PAGE was performed at 20 mA (per gel) in electrophoresis buffer (25 mM Tris-HCl, 192 mM glycine and 3.5 M SDS (pH 8.3)). To visualize the protein-bands the gel was stained for 30 min with coomassie-solution (0.05% Coomassie Brilliant Blue, 40% MeOH and 10% HAc) at room temperature. Subsequently, the coomassie-solution was removed and the gel was covered with water and destained using a microwave (2-3 min). Gel documentation was then performed with the ChemiDoc system (BioRad) and the Quantity One software packet (BioRad).

2.7.5. Immunoblotting analysis

In order to demonstrate an *in vivo* effect of the FlaB and FlaX expression in the *S. acidocaldarius* strains MW001, *Δsaci_ptp* and *Δsaci pp2a* immunoblotting analysis were performed by Julia Reimann (AG Albers, Max Plank Institut for Terrestrial Microbiology, Marburg, Germany). The expression levels of the archaella components FlaB and FlaX were detected in an immunoblotting experiment. Therefore the background strain *S. acidocaldarius* MW001 and both phosphatase deletion mutants (*saci_ptp* and *Δsaci_pp2a*) were transferred to nutrient-limited medium. The same amount of cells referring to OD600 with and without starvation stress treatment were lysed and the proteins separated by SDS-PAGE. The expression levels of FlaB and FlaX were detected by immunoblotting using specific antibodies.

2.8. Enzyme assays

2.8.1. Aldehyde dehydrogenase enzyme assay

Aldehyde dehydrogenase activity was determined in a continuous enzyme assay at 70°C. The standard assay was performed in 0.1 M HEPES/KOH (pH 6.5 at 70°C). Enzymatic activity was determined by monitoring the change in absorbance at 340 nm due to the formation of NADH or NADPH (εNADPH (70°C) = 5.71 $mM^{-1}cm^{-1}$; εNADH (70°C) = 5.8 $mM^{-1}cm^{-1}$). For each assay three independent measurements were performed. SSADH-I activity was determined with NAD^+ (0.05 – 25 mM) as cosubstrate and D,L-GA (2.5 – 40 mM), D,L-GAP (1 – 15 mM), D,L-glycolaldehyde (2.5 -12.5 mM), SSA (0.01 – 12.5 mM) and DOP (0.1 – 10 mM). SSADH-II assays were performed with NAD^+ (0.1 – 25 mM) or $NADP^+$ (0.005 – 20 mM) as cosubstrates and D,L-GA (0.5 – 16 mM), D,L-GAP (0.5 – 28 mM), D,L-glycolaldehyde (0.5 – 45 mM), SSA (0.005 – 20 mM) and DOP (0.1 – 10 mM).
MSDH: Prior to kinetic measurments, MSDH was preincubated at 30°C for 30 min in the presence of 2 mM NAD^+ to eliminate the lag-phase exhibited by progress curves for enzymatic turn-over, as already described for the *B. subtilis* MSDH [147]. Initial rate measurements were carried out at 30°C on a SAFAS UV mc2 spectrometer in 50 mM potassium phosphate (pH 8.2) in the presence of NAD^+ (2 mM), D,L-methylmalonate semialdehyde (MMSA) (0.5 – 1 mM), malonate semialdehyde (MSA) (0.5 – 1 mM) or propion aldehyde (PA) (50 – 100 mM) and CoA (0.5 – 1 mM).

2.8.2 Triosephosphate isomerase enzyme assay

Assays were performed at 60°C, 65°C, 70°C and 80°C in the catabolic direction (DHAP to GAP) using GAPN of *S. solfataricus* (SSO3194) as auxiliary enzyme. Enzyme activity was followed photometrically at 340 nm, monitoring the change in absorbance due to the formation of NADPH (εNADPH (70 C) = 5.71 $mM^{-1}cm^{-1}$). Enzyme assays were performed in the presence of 100 mM Tris/HCl (pH 6.5 (70°C)), 170 µg GAPN (after heat precipitation) and 5 mM $NADP^+$ in a final volume of 500 µL. TPI, GAPN and $NADP^+$ were incubated for 2 min at the respective temperature before DHAP was added to start the reaction. For each assay three independent measurements and a control without TPI were performed.
In the anabolic direction (GAP to DHAP) assays were performed at 70°C using the bifunctional FBPase from *S. solfataricus* P2 (SSO0286), the PGI from *T. tenax* and G6PDH from *T. maritima* as auxiliary enzymes. The PGI and G6PDH were provided from Theresa Kouril (AG Siebers, Universitiy of Duisburg-Essen, Essen, Germany) as described elsewhere [119] and the FBPA/ase was prepared as

described in section 2.7.2.5. The reaction progress was followed photometrically at 340 nm, following the change in absorbance due to the formation of NADPH (εNADPH (70 C) = 5.71 $mM^{-1}cm^{-1}$). Assays were performed in the presence of 100 mM Tris/HCl, 20 mM $MgCl_2$ (pH 6.5 (70°C)), 50 µL FBPase, 10 µL PGI, 10 µL G6PDH (after heat precipitation) and 1 mM $NADP^+$ in a final volume of 500 µL. TPI, FBPase, PGI, G6PDH and $NADP^+$ were incubated for 2 min at 70°C. The reaction was started by adding GAP as substrate. For each assay three independent measurements and a control without TPI were performed

2.8.3. Glucose-1-dehydrogenase enzyme assay

Glucose-1-dehydrogenase (Saci_1079) activity was determined in a continuous enzyme assay at 70°C. The standard assay was performed in 0.1 M Tris/HCl, 5 mM $MgCl_2$ (pH 6.5 at 70°C). Enzymatic activity was determined by monitoring the change in absorbance at 340 nm due to the formation of NADH or NADPH (εNADPH (70 °C) = 5.71 $mM^{-1}cm^{-1}$; εNADH (70 °C) = 5.8 $mM^{-1}cm^{-1}$). The enzyme and $NAD(P)^+$ were incubated for 2 min at 70°C and the reaction was started by the addition of the substrate. For each assay three independent measurements were performed.

2.8.4. Protein phosphatase enzyme assays

Phosphatase activity was determined in a continuous enzyme assay at 70°C with para-nitrophenylphosphate as substrate. For Saci-PTP the standard assay was performed in 0.1 M HEPES/KOH (pH 6.5) with 7 µg protein, whereas for Saci_0884 the assay was performed in 20 mM Tris/HCl in the presence of 300 mM KCl (pH 6.5), 5 mM Mn^{2+}, 1 mM EGTA containing 0.25 µg protein. Enzymatic activity was determined by monitoring the change in absorbance at 403 nm due to the formation of para-nitrophenyl (εp-NPP = 18 $mM^{-1}cm^{-1}$). For each assay three independent measurements were performed.

Protein tyrosine phosphatase activity was determined by using the tyrosine phosphatase assay system (Promega) following the manufacturer's instruction. The artificial p-peptide RRA(pT)VA was used as substrate. The enzyme activity was determined in a discontinuous enzyme assay at 70°C in 0.1 M HEPES/KOH (pH 6.5) containing 2 µg Saci-PTP. The phosphate release was measured and visualized by using malachite green. For each assay three independent measurements were performed.

Protein serine/threonine phosphatase activity was determined by using the serine/threonine phosphatase assay system (Promega) following the manufacturer's instruction. The artificial p-peptide TEVGKRI(pY)RLVGDKN was used as substrate. This p-peptide corresponds to the putative ALDH Saci_1938,which was detected as p-protein during the first p-proteome analysis from

S. acidocaldarius (not mentioned in this study). The p-peoptide was synthesized by Protein Mods, LLC (Madison, Unites states of America). The enzyme activity was determined in a discontinuous enzyme assay at 70°C in 20 mM Tris/HCl, in the presence of 300 mM KCl (pH 6.5), 5 mM Mn^{2+}, 1 mM EGTA and 0.1 µg Saci-PP2A. The phosphate release was measured by using malachite green (Promega). For each assay, three independent measurements were performed.

2.9. Kinetic paramteres

For enzyme reactions following Michaelis Menten kinetics the kinetic parameters (V_{max} and K_m) were calculated by iterative curve-fitting (Hanes) using the program Origin (Microcal Software, Northampton, MA, USA). For enzyme reactions, which do not follow Michaelis Menten kinetics the following equations were used to calculate the kinetic parameters:

$v = \frac{(V_{max}*[S])}{K_m + [S] - k_{irr}*[S]}$ for irreversible substrate inhibition or $v = \frac{V_{max}*[S]}{K_m + [S] + \frac{[S]^2}{k_i}}$ for reversible substrate inhibition, as indicated.

For the calculation of the K_i-values of 3-PG and PEP for the *S. solfataricus* TPI the equation [119]:

$$v_{TPI} = \frac{V_{Mf}*\frac{GAP}{K_{m,GAP}} - V_{Mr}*\frac{DHAP}{K_{m,DHAP}}}{(1+\frac{GAP}{K_{m,GAP}}+\frac{DHAP}{K_{m,DHAP}}+\frac{3PG}{K_{i,3PG}}+\frac{PEP}{K_{i,PEP}})}$$

2.10. Sample preparation for Proteome analysis

2.10.1. Protein extraction and quantitation

The protein extraction and quantitation was performed by Dominik Esser (AG Siebers, Universitiy of Duisburg-Essen, Essen, Germany) and Trong Khoa Pham (ChELSI Institue, Department of Chemical and Biological Engineering, The University of Sheffield, United Kingdom). .*Sulfolobus* cells, as prepared in section 2.4.1 and 2.4.2, were washed with cold water and then resuspended in lysis phosphate buffer containing 50 mM Tris-HCl (pH 7.5), 5 mg/ml lysozyme and phosphatase inhibitors consisting of (5 mM each) sodium fluoride, 2-glycerol phosphate, sodium vanadate, and sodium pyrophosphate [70]. Cell lysis was performed at 37°C for 15 min, and then protein extraction was carried out using sonication as detailed elsewhere [176,177]. To improve protein extraction as well as membrane protein solubilization, DNAse I (100 µg/ml) was added and incubated at 37°C for 15 min, and N-Octyl-glucoside was added to a final concentration of 1% [70]. Samples were then centrifuged at 21000 x g for 20 min at 4°C, the supernatant was collected and protein concentration was

quantified using the RC-DC Protein Quantification Assay (Bio-Rad, UK). All chemicals were purchased from Sigma.

2.10.2. Protein digestion and fractionation

The sample preparations were performed by Dominik Esser (AG Siebers, Universitiy of Duisburg-Essen, Essen, Germany) and Trong Khoa Pham (ChELSI Institue, Department of Chemical and Biological Engineering, The University of Sheffield, United Kingdom). Two mg of protein (per condition) was dissolved and denatured in 8 M urea, reduced with 10 mM DTT in 50 mM ammonium bicarbonate at 56°C for 1 h, then alkylated with 55 mM idoacetamide in 50 mM ammonium bicarbonate at 37°C for 30 min in the dark. Samples were then diluted with 40 mM ammonium bicarbonate in 9% acetonitrile (ACN) to reach a final <1 M urea concentration before protein digestion [178]. Proteins were digested by trypsin with a ratio (trypsin/protein) of 1/25 overnight at 37°C. Samples were then dried by vacuum concentration before resuspension in buffer A containing 10 mM KH2PO4, 30% acetonitrile and 0.1% trifluoroacetic acid (TFA) pH 3.0 for fractionation as well as sample cleaning. Strong cation exchange chromatography was performed as described elsewhere. Briefly, buffer B used for peptide elution comprised 10 mM KH_2PO_4, 30% ACN, 0.1% TFA, and 500 mM KCl, pH 3.0 [176]. Fractioned peptides were collected every 2 min, samples were then dried by vacuum concentration.

2.11. Mass spectrometry analysis and data processing

2.11.1 *S. solfataricus* P2

The analysis and data processing were performed by Dominik Esser (AG Siebers, Universitiy of Duisburg-Essen, Essen, Germany) and Trong Khoa Pham (ChELSI Institue, Department of Chemical and Biological Engineering, The University of Sheffield, United Kingdom). All dried peptides were analysed on an ion trap using the PAcIFIC technique as described in detail by Panchaud et al. [179] with some modifications. All fractionated peptides were dissolved (and then combined) in buffer A containing 3% ACN and 0.1% formic acid (FA) before submitting multiple times on a HCT-Ultra ion trap (QIT) (Bruker Daltonics, UK) coupled with an Ultimate 3000 nano-flow HPLC (Dionex, UK) comprising a trap and an analytical column (3 mm C18, Dionex-LC Packings) operating at a flow rate of 300 nL/min. Buffer B contained 97% ACN and 0.1% FA. The gradient was 90 min, ramping from 5% B to 35% B in the first 70 min, then ramping to 90% B in 5 min, then switching back to 3% ACN for 15 min. Peptides eluting from the HPLC column were directly submitted to a Bruker low-flow

electrospray needle operating at 3600 V (Bruker Daltonics, UK). For PAcIFIC, in the first injection, the QIT was used to perform a PACiFIC-type analysis with collision induced dissociation (CID) at each of 10 continuous 1.0 m/z intervals across a range of 10 m/z per LC-MS analysis using a 2.0 m/z isolation width with a dynamic ion charge threshold (ICC) of 180000, fragment ions were scanned from 100 to 1600 m/z with averages of 3. A subsequent analysis started with a new (+10 m/z forward shifted) 10 m/z range in the same format as the first run and this continued until the considered precursors covered the range from 400 to 1200 m/z.

All mass spectrometry data were then extracted to mgf format using Bruker Data Analysis V4.0 with a MRM script. Subsequently, all MS data were merged and searched within the appropriate *Sulfolobus solfataricus* P2 database (downloaded in June 2009) using the Phenyx V 2.6 search engine (Genebio, Geneva). The parameters for searching were set up as follows: fixed modification of cys_CAM (carbamidomethylation) variations of Oxidation_M (methione) and PHOS (serine, tyrosine and threonine), trypsin was used with miss cleavages of 2. Other parameters were: parent tolerance of 1.5 and MS/MS tolerance of 0.8 Da, a minimum peptide length of 5 and z-score of 5.5, p-value 10-5, AC score of 5.5. All p-peptides were then inspected manually to ensure that a neutral loss of phosphoric acid was found either from the precursors or (and) fragment ions as well as a good coverage of b, b++ and (or) y, y++ ion series was required (at least 5 either b or y series). A biological duplicate was applied for each condition.

2.11.1 *S. acidocaldarius*

The analysis and data processing were performed by Dominik Esser (AG Siebers, Universitiy of Duisburg-Essen, Essen, Germany) and Trong Khoa Pham (ChELSI Institue, Department of Chemical and Biological Engineering, The University of Sheffield, United Kingdom). All dried peptides were cleaned using a C18 Discovery DSC-18 SPE column (Supelco, Sigma) as described elsewhere [180]; all fractions were subsequently combined, before analysis on a HCT-Ultra ion trap (Bruker Daltonics, UK) coupled with an Ultimate 3000 nano-flow HPLC (Dionex, UK) consisting of an analytical column (3 mm C18, Dionex-LC Packings) operating at a flow rate of 300 nL/min. Buffer A contained 3% acetonitrile and 0.1% formic acid, whilst buffer B contained 97% acetonitrile and 0.1% formic acid. The 90 min gradient comprised ramping from 3% B to 35% B in the first 70 min, then ramping to 90% B in 5 min, then switching back to 3% of buffer A for 15 min. Peptides eluting from HPLC column were directly submitted to the ion trap.

The PAcIFIC technique was applied for analysing (combined) cleaned peptides. Details of this technique can be found elsewhere [10,59]. Briefly, resuspended peptides were injected multiple times on the ion trap with collision induced association at each of 10 continuous 1.0 m/z intervals

across a range of 10 m/z for each LC-MS analysis using a 2.0 m/z width and fragmented ions scanned from 100 to 1600 m/z with a dynamic ion charge threshold of 160000. Subsequently, a next (forward) new ten ions was started in the same format as the first run, then this process continued until the considered precursors covered the range from 450 to 1100 m/z.

Data from mass spectrometry were then extracted to mgf format using Bruker Data Analysis V4.0 with a MRM script, these were then searched against the *S. acidocaldarius* database downloaded from NCBI in March 2010 using Phenyx V 2.6 (Genebio, Geneva). The searches were performed using parameters as follows: carbamidomethylation of cysteine (fixed modification), oxidation of methionine (variation), and phosphorylation of serine, tyrosine, threonine (variation), trypsin with 2 missed cleavages. Furthermore, other parameters such as parent, MS/MS, tolerances were set at 2.0 and 0.8 Da, respectively, whilst minimum peptide length, z-score, p-value and AC score were set at 5, 5.5, 10-5, and 5.5 respectively.

The identified p-peptides were then manually inspected to ensure that a neutral loss of phosphoric acid was observed either from the precursors or (and) fragment ions with an essential presence of at least 5 (either) b or y (ions) series. Moreover, an independent biological duplicate sample was generated for each strain and an identical analysis carried out.

2.12. Biophysical techniques

2.12.1. Flow cytometry analysis

Flow cytometry analysis of the *S. acidocaldarius* strains MW001, *Δsaci_ptp* and *Δsaci_pp2a* were performed by Ann-Christin Lindås and Rolf Bernander (Department of Molecular Evolution, Biomedical Center, Uppsala University). Aliquots from growing cultures of strains MW001, *Δsaci_ptp* and *Δsaci_pp2a* were transferred to ice-cold ethanol (final concentration 70%). For DNA staining, approximately 0.5 ml of cell suspension in ethanol was centrifuged at 4°C for 10 min at 16 000 xg. The pellet was resuspended in 1 ml of 10 mM Tris-buffer (pH 7.4) containing 10 mM $MgCl_2$, centrifuged again, and then resuspended in 70 µl of the same buffer. Equal volumes (65 µl) of cell suspension and DNA-specific stain (0.2 mg/ml mithramycin A and 0.04 mg/ml ethidium bromide, also in Tris-MgCl2 buffer) were mixed. The samples were incubated on ice for 30 min before analysis in an Apogee A40 Analyzer flow cytometer equipped with a 405 nm solid-state laser.

2.12.2. Cell size analysis

Cell size analysis of the *S. acidocaldarius* strains MW001, *Δsaci_ptp* and *Δsaci_pp2a* were performed by Julia Reimann (Max Plank Institut for Terrestrial Microbiology, Marburg, Germany), Ann-Christin Lindås and Rolf Bernander (Department of Molecular Evolution, Biomedical Center, Uppsala University). Microscopy pictures were taken using an Axio Imager.M1 microscope (Zeiss) equipped with a Zeiss Plan Apochromat x 100/1.40 Oil DIC objective and a Cascade:1K CCD camera (Photometrics). Cell size was analyzed automatically with ImageJ and ObjectJ using a modified version of the filaments-91i.ojj project. The results were evaluated with Origin 6.1 (OriginLab Corporation, Northampton, MA, USA).

2.12.3. Motility assay on semi-solid gelrite plates

Motility assays of the *S. acidocaldarius* strains MW001, *Δsaci_ptp* and *Δsaci_pp2a* were performed by Julia Reimann (Max Plank Institut for Terrestrial Microbiology, Marburg, Germany). To compare the motility of the background strain *S. acidocaldarius* MW001 to the phosphatase deletion strains *(Δsaci_ptp* and *Δsaci_pp2a*) the two control strains *ΔaapF* (hypermotile strain) and *ΔaapFΔflaH* (non-motile strain) were used. The strains were spotted on semi-solid gelrite plates (0.15% (w/v) gelrite)

containing only 0.005% (w/v) tryptone, 0.2% (w/v) dextrin and 10 µg/ml uracil as described before [35,37]. After five days at 76°C pictures were taken of the plates.

2.13. Software, internet databases and tools

BioEdit: Open source software used for the preparation and modification of sequence alignments.

BLASTP: For sequence comparison the Blast tool at the website of The National Centre for Biotechnology Information (NCBI; http://www.ncbi.nlm.nih.gov/BLAST) was used.

BRENDA: For enzyme functional data and information about the respective literature the BRENDA database was used (http://www.brenda-enzymes.info/index.php4).

Clone Manager 7.0: Software used to design oligonucleotides for molecular cloning and to investigate the restriction sites of the investigated genes and vectors.

Clustal W2: Internet tool used to generate multiple sequence alignments or phylogenetic trees (http://www.ebi.ac.uk/Tools/phylogeny/clustalw2_phylogeny/).

ChemDraw Pro 12.0: Software used to draw chemical structures and metabolic pathways.

HHpred: Internet tool used to perform profile-sequence comparison (http://toolkit.tuebingen.mpg.de/hhpred#).

IMG: Integrated microbial genomes database, used for sequence comparison and analysis of gene neighborhoods (http://img.jgi.doe.gov/cgi-bin/w/main.cgi).

JalView: Open source program used for interactive editing, analysis and visualization of multiple sequence alignments, sequence annotation, secondary structure information, phylogenetic trees and 3D molecular structures.

KEGG: Kyoto Encyclopaedia of Genes and Genomes database (http://www.genome.jp/kegg/pathway.html).

Microcal Origin 8.1: Software used to plot curves and calculate kinetic parameters.

Quantity One (BioRad): Software used for gel documentation.

SMART: Protein database which allows the identification and annotation of genetically mobile domains and the analysis of domain architectures (http://smart.embl.de/).

Splits Tree4: Application used for construction of unrooted phylogenetic networks ,phylogenetic trees or networks.

STRING: Database used to predict protein-protein interactions, which are either direct interactions (physical) or indirect interactions (functional, http://string.embl.de) [181].

UniProt: Protein database containing information's of protein sequences and functional information's (www.uniprot.org).

WIN ASPECT: Software used to measure and monitor the performed enzyme activity assays.

3. Results and discussion

3.1. Change of carbon source causes dramatic effects in the phospho-proteome of the archaeon Sulfolobus solfataricus

This study was performed in collaboration with Phillip Craig Wright and Trong Khoa Pham (ChELSI Institute, Department of Chemical and Biological Engineering, The University of Sheffield) and Sonja-Verena Albers and Julia Reimann (AG Albers, Max Planck Institut for Terrestrial Microbiology, Marburg, Germany). Dominik Esser performed the MS/MS analysis under the guidance of Trong Khoa Pham at the University of Sheffield and later on processed, analyzed and interpreted. these data. Julia Reimann performed the *S. solfataricus* growth experiments on D-glucose and tryptone and send the cells for MS/MS analyses to the University of Sheffield. The results of this study were published in Journal of Proteome research (PMID: 22639831).

3.1.1. Introduction

As well as constituting the third domain of life, Archaea offer exciting opportunities for biotechnological exploitation. The key model organism for the Crenarchaeota is *S. solfataricus*, a thermoacidophile that grows optimally at 80°C and pH 3.2 [182]. It grows heterotrophically on complex peptidolytic constituents and a wide variety of sugars [31]. The CCM has been extensively studied [3,24-28] and due to new, unusual enzymes, pathways and regulatory potential it shows great potential for White Biotechnology and Synthetic Biology [22,23].

S. solfataricus uses non-phosphorylative modifications of the respective hexose and pentose degradation pathways, like the branched ED pathway for glucose and galactose [3,24,28], the Dahms pathway for D-xylose degradation [25] and an oxidative pathway for D-arabinose degradation [27]. Like most other Archaea, *S. solfataricus* uses the modified EMP pathway for gluconeogenesis [3,26]. Despite these metabolic pathway investigations, their regulation is still unclear. It is known though that classical bacterial and eukaryotic control points are absent [3] and only minor regulation has been observed at the transcript and proteome level in response to carbon source alteration [26].

Reversible protein phosphorylation/dephosphorylation plays a major role in signal transduction and allows for a rapid appropriate cellular response to diverse external and internal signals/stimuli [33]. This post-translational modification takes usually place at His and Asp residues (one/two-component systems) or at Ser, Thr or Tyr residues (ePKs).

Reversible protein phosphorylation was established in the Archaea with a report on the halophilic Euryarchaeon *H. salinarium* [38], where CheA and CheY represent the best-studied archaeal two component system. Archaeal p-proteins with Ser, Thr and Tyr modifications have since been identified in the thermoacidophilic Crenarchaeon *S. acidocaldarius* [40]. It is also clear that ePK and the respective PP are present in Archaea [183]. Further, bioinformatic analyses revealed the presence of ePKs and PPs in all completed archaeal genomes, whereas His kinases –two component systems- are restricted to Euryarchaeota [33,184,185]. It is also proposed that novel PK families exist among the Archaea [78]. In terms of phosphorylation-mediated post-translational regulation of the CCM, only for GAD has this unambiguously been demonstrated [57], raising the question as to whether post-translational modifications such as phosphorylation play a major role in CCM regulation in *S. solfataricus* and the Archaea in general.

The largest archaeal p-proteome study to date, the p-proteome from *H. salinarum* strain R1 was analyzed using TiO_2 enrichment of p-peptides and mass spectrometry revealed 69 broadly functionally distributed p-proteins [46]. Despite this background, knowledge on post-translational modifications in Archaea is rather scarce. Only five PKs and six PPs have been characterized in detail and only for one of these regulatory function has been demonstrated [33,47,51,52,55,105,186,187].

We demonstrate that the Precursor Acquisition Independent From Ion Count (PAcIFIC) approach [179] can be extended to deeply mine p-proteomes. We exemplify this by studying a member of the third domain of life. Importantly, using PAcIFIC, compared to previous studies depending on p-peptide pre-enrichment approaches we detected a 3-7-fold higher number of p-proteins with a so far unique/unexpected preference for Tyr-P (1318 phosphorylation sites located on 690 p-peptides from 540 p-proteins with a Ser/Thr/Tyr-ratio of 25.8/20.6/53.6%). Using the pre-enrichment free PACiFIC approach we dramatically extend knowledge of archaeal p-proteins from 80 to 620. Taking advantage of our workflow, we demonstrate the substantial impact of offered carbon source on phosphorylation events occurring in *S. solfataricus*, in particular those in the CCM.

3.1.2. Phospho-proteome analysis

Two biological replicates, with both carbon sources, were performed by using PAcIFIC [179]. The comparison of the unique p-proteins which were detected in the glucose grown cells of the first and second biological replicate revealed that 88.7% (276 p-proteins) of the 311 unique p-proteins were found in both biological replicates.

Furthermore, 6.4% (20 p-proteins) were only detected in the first biological replicate, and the remaining 4.8% (15 p-proteins) in the second biological replicate. In case of the tryptone grown cells

63.3% (197 p-proteins) were detected in both biological replicates and 1.9% (6 p-proteins) only in the first and; 34.7% (108 p-proteins) in the second biological replicate, respectively.

Overall, 1200 p-peptides were identified in *S. solfataricus* P2 corresponding to 690 unique p-peptides and furthermore to 622 p-proteins (list of p-peptides see Table S1 of [188]). 82 of these p-proteins were detected in both, glucose and tryptone grown cells, resulting in 540 unique p-proteins (Figure 6).

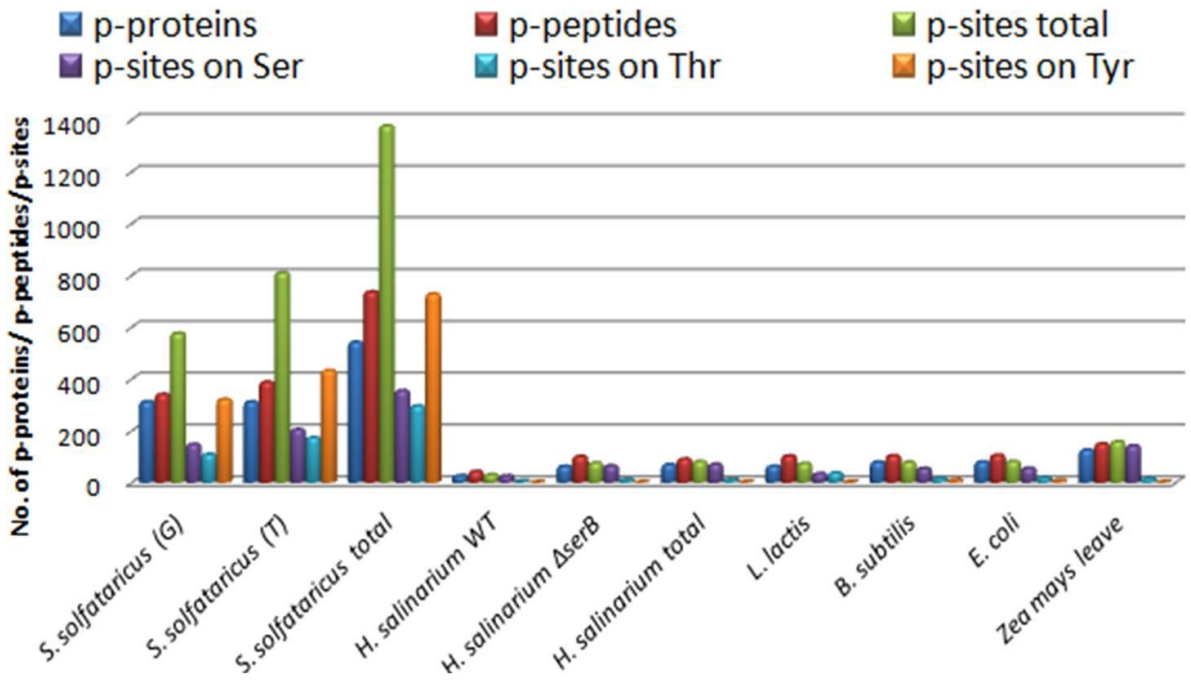


Figure 6 Comparison of the total number of detected p-proteins, p-peptides, p-sites and their distribution on the amino acids Ser, Thr, Tyr in *S. solfataricus* P2 (grown on glucose and tryptone, current study [188]), *B. subtilis* [70], *E. coli* [71], *L. lactis* [72], *Zea mays leaf* [189] and *H. salinarium* [46].

More than half of the identified p-peptides (402; 58.3%) had multiple (i.e. 231 doubly, 125 tri-, 39 tetra-, 5 quintuple- and 2 hexa-phosphorylated) p-sites and 288 p-peptides with a single p-site were identified. Overall, 1318 p-sites were detected with 340 p-sites on Ser, 272 on Thr and 706 on Tyr, revealing a total Ser/Thr/Tyr ratio of 25.8/20.6/53.6% (Table 7). The high proportion of phosphorylation of either Ser or Thr might be a result of their high abundance in the genome. However, using the PACiFIC workflow here, we were able to detect low abundance phosphorylation of Tyr. It has been known that Ser or Thr phosphorylated peptides can be distinguished from those Tyr phosphorylated peptides, since Ser or Thr phosphorylated peptides lose predominantly (98 Da) H_3PO_4 while Tyr phosphorylated peptides are more likely to lose (80 Da) HPO_3 [190]. Therefore, we manually carried out a check of the loss of HPO_3 from intact peptide precursors to ensure that the identification of peptides containing Tyr was actually phosphorylated.

As mentioned in the methods section (see 2.11.1) for the identification/detection of phosphoryated peptides, it was required to see a neutral loss of 80 Da or 98 Da, as well as a good coverage of the fragmentation b and y ion series. Examples of phosphorylated Tyr are shown in Figure 7 where

single, double and triple phosphorylation sites are presented. Furthermore, the presence of p-Tyr was already verified in the past in the crude extract of *S. solfataricus*, *M. thermophila* and *H. volcanii* [191] and only the dominance of Tyr phosphorylation in *S. solfataricus* was rather unexpected. However, it is not possible to preclude that a fraction of the detected p-Tyr was a result of the acid treatment during the sample preparation. During this step it is possible that nucleotidylated Tyr was transformed into p-Tyr [192]. In order to prevent false positive p-Tyr p-peptides, all MS/MS spectra were controlled manually and no evidence for nucleotidylated Tyr peptides was found.

Table 7 Comparison of p-proteome studies.

Organism	P-proteins	P-peptides	P-sites total	P-sites S	P-sites T	P-sites Y	S/T/Y %-ratio
	[No.]	[No.]	[No.]	[No.]	[No.]	[No.]	[%]
***S. solfataricus* (G)**	311	343	580	146	107	327	25.2/18.5/55.4
***S. solfataricus* (T)**	311	384	810	203	174	433	25.1/21.5/53.5
***S. solfataricus* total**	540	690	1318	340	272	706	25.8/20.6/53.6
***H. salinarium* WT**	26	42	30	25	5	0	84/16/0
***H. salinarium* ΔserB**	62	100	75	64	10	1	86/13/1
***H. salinarium* total**	69	90	81	70	10	1	87/12/1
L. lactis	63	102	73	34	37	2	46.6/50.7/2.7
B. subtilis	78	103	78	54	16	8	69.2/20.5/10.3
E. coli	79	105	81	55	19	7	67.9/23.5/8.6
***Zea mays* leaf**	125	149	157	141	15	1	89.8/9.6/0.6

In the *S. solfataricus* genome, eight ORFs encoding potential ePKs (SSO0197, SSO0433, SSO2291, SSO2387, SSO2605, SSO3182, SSO3207, SSO0469) have been identified [51,52]. Three out of these eight ePKs (SSO0433 (PK4), SSO0469 (SSOPK3), SSO2387 (SSOPK2)) and a membrane associated PK (SSOPK1), where the encoding gene is unknown, have been investigated. For these three, Ser/Thr kinase activity has be demonstrated [49,51,52]. No phosphorylation on Tyr rich compounds (i.e. poly(Glu4:Tyr), poly(Glu:Tyr)) was detected. Interestingly, the three predicted ePKs (SSO2291, SSO3182, and SSO3207) show the characteristic signature of protein Tyr kinases (a characteristic Arg residue in subdomain VIb, i.e. Asp-Xaa-Arg-Xaa-Xaa-Asn or Asp-Xaa-Xaa-Arg-Asn [193]), whereas all residual archael ePK ORFs lack this feature [33]. However, it cannot be excluded that potentially novel archael-type PKs are responsible for Tyr phosphorylation in *S. solfataricus*. In accordance with our finding of Tyr phosphorylation, a Tyr specific phosphatase was identified in *S. solfataricus* [47]. This is in addition to the homolog of the Ser/Thr specific phosphatase from *S. solfataricus* P1, which shares 91% amino acid sequence identity to the P2 enzyme (PPP, PP-1arch [33]).

3.1.3. Classification of p-proteins

The *S. solfataricus* P2 genome was recently revised and the genome was updated by using the arCOG functional code and available biochemical information (BRENDA, http://www.brenda-enzymes.info/) [194]. The identified p-proteins were categorized according to the arCOG functional code [195]. P-proteins were identified in 21 out of 26 arCOG functional categories in *S. solfataricus*, and results are summarized in Figure 8, where strikingly most of the p-proteins are either hypothetical proteins, proteins of unknown function or proteins with no annotation.

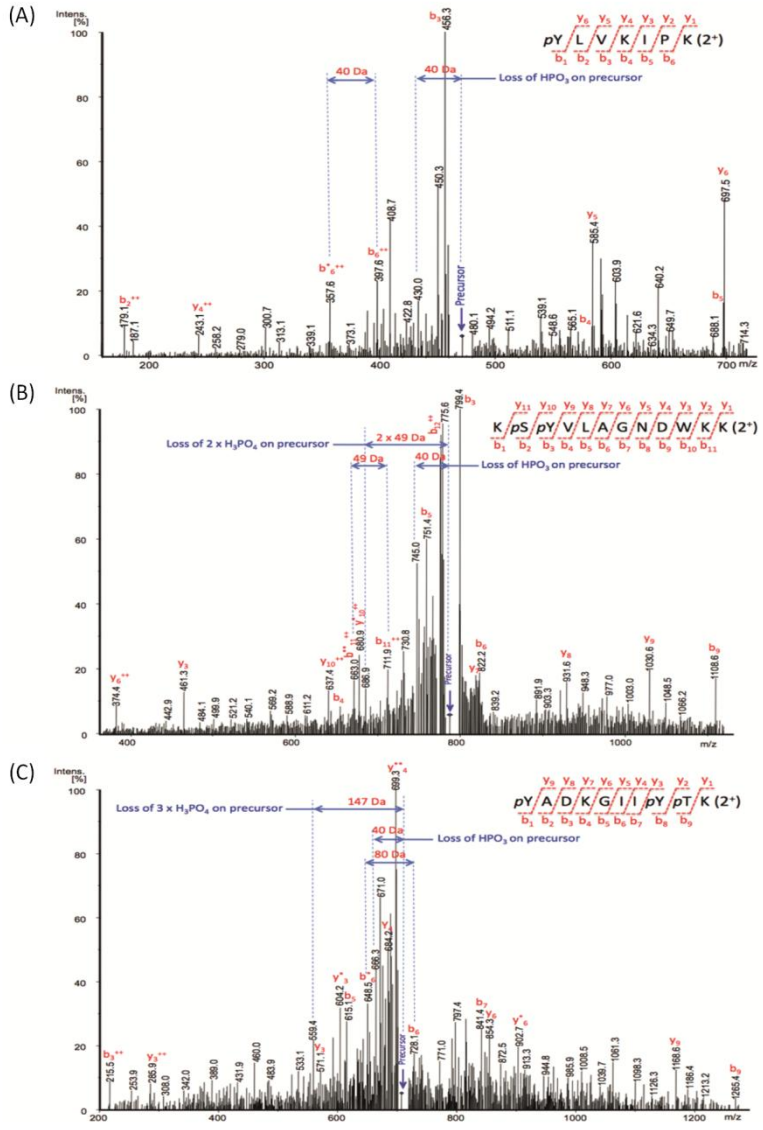


Figure 7 Representative MS/MS spectra for phosphorylation on tyrosine.
(A) for phosphorylated peptide pYLVKIPK (2^+) from glucose 1-dehydrogenase SSO3003 carrying single phosphorylated tyrosine, (B) for phosphorylated peptide KpSpYVLAGNDWKK SSO0304 (2^+) from 3-deoxy-D-arabino-helptulosonate 7-phosphate synthase carrying double phosphorylated sites including serine and tyrosine, and (C) for phosphorylated peptide pYADKGIIpYpTK SSO1440 (2^+) from ATP-dependent RNA-helicase carrying triple phosphorylated sites including a single threonine and double tyrosine. Only neutral loses of HPO_3 or/and H_3PO_4 and ions b and y series are annotated.

The broad distribution of p-proteins suggests that regulatory protein phosphorylation plays a major role in almost every cellular process. Interestingly, the comparison of p-proteins identified from prokaryotic p-proteome studies (*E. coli*, *B. subtilis*, *L. lactis*, *H. salinarium*) revealed several significant matches with an expected dominance for archaeal orthologs (Table 8).

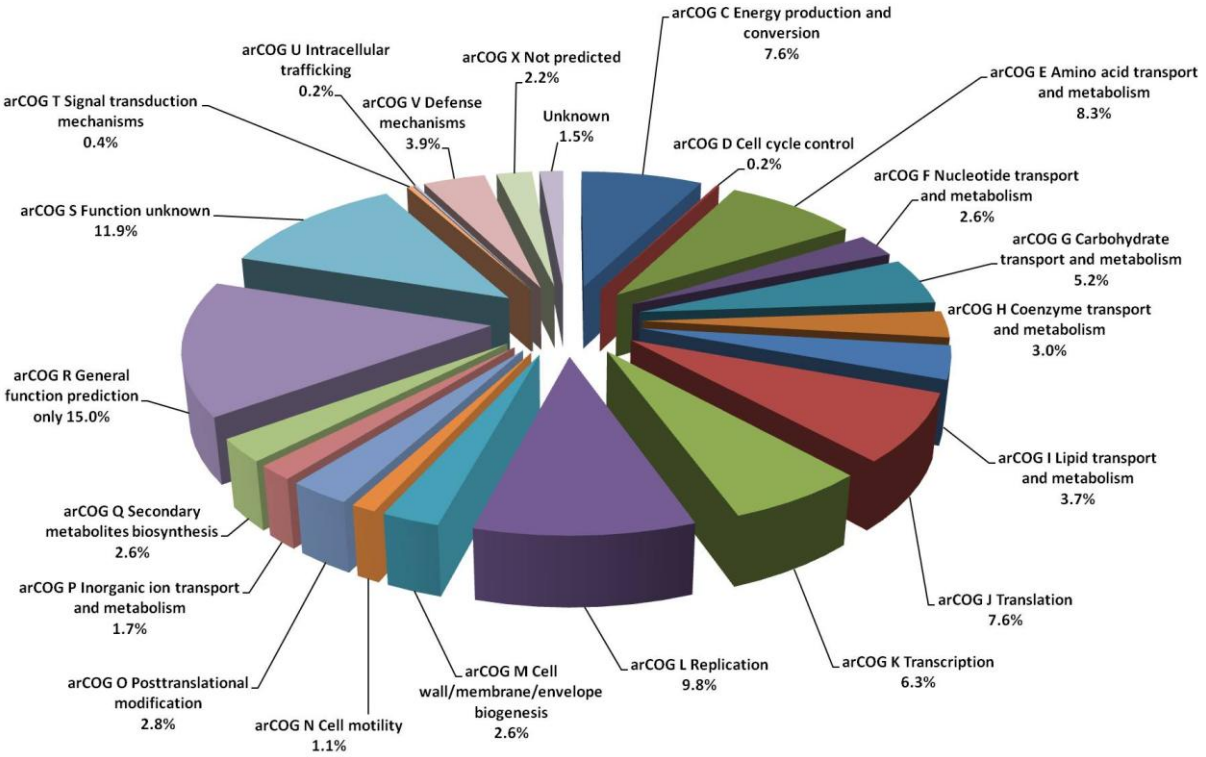


Figure 8 Assignment of p-proteins into the arCOG functional classes of *S. solfataricus* P2.
The distribution of all detected p-proteins (glucose and tryptone) in the respective arCOG functional categories [194,195] is shown and the amount of p-proteins as well as the percental distribution is given.

The comparison of p-proteins identified from prokaryotic p-proteome studies (*E. coli*, *B. subtilis*, *L. lactis*, *H. salinarium*) revealed 29 significant matches (e-value ≤ e^{-14}). According with their phylogenetic distance Bacteria comprise only 3-5 phosphorylated orthologs, whereas the Archaeon *H. salinarium* possesses 15 phosphorylated orthologs. The majority (six) of the orthologs found in *H. salinarium* belongs to the arCOG J (translation) category; the remaining orthologs are distributed in different arCOG categories. Disregarding different experimental set ups and growth conditions IDH was phosphorylated in three species (*S. solfataricus*, *H. salinarium* and *E. coli*) highlighting its important function in the regulation of carbon flux into the glyoxylate and TCA cycle (see also section 3.1.11.).

Table 8 Comparison of archaeal and bacterial p-proteome analyses.
The comparisons were performed with BLAST-P analyses (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Only hits with E-values lower than E^{-14} are shown.

Organism	Gen ID	Predicted function	Orthologe	arCOG annotation	E-value
H. salinarium strain R1 & *ΔserB*	OE4169F	Conserved hypothetical protein	SSO0258	Predicted DNA-binding protein	$2E^{-40}$
	OE4671R	Conserved hypothetical protein	SSO0572	ATPase (PilT family)	$3E^{-95}$
	OE3943R	Conserved hypothetical protein	SSO0408	Ribosomal protein S15P	$1E^{-22}$
	OE2629F	Ribosomal protein S11	SSO0069	Ribosomal protein L13	$3E^{-19}$
	OE3388F	Ribosomal protein L3	SSO0719	Ribosomal protein L3	$7E^{-54}$
	OE3412F	Ribosomal protein L32.eR	SSO0701	Ribosomal protein L32E	$8E^{-14}$
	OE4101R	Tryptophanyl-tRNA synthetase	SSO0452	Tryptophanyl-tRNA synthetase	$2E^{-17}$
	OE4736R	Ribosomal protein S12	SSO0219	Ribosomal protein S12	$3E^{-46}$
	OE2866R	Succinate dehydrogenase subunit B	SSO2357	Succinate dehydrogenase /fumarate reductase	$2E^{-32}$
	OE3634F	Isocitrate dehydrogenase	SSO2182	Isocitrate dehydrogenase	$1E^{-92}$
	OE2458R	IMP dehydrogenase	SSO0613	Aspartate carbamoyl-transferase	$8E^{-37}$
	OE4301R	ABC-type transport system ATP-binding protein	SSO2615	ABC-type oligopeptide transport system	$3E^{-70}$
	OE4230F	Probable acylaminoacyl-peptidase	SSO2693	Dipeptidyl aminopeptidase /acylaminoacyl-peptidase	$2E^{-47}$
	OE1385F	Thiamine biosynthesis protein thiN	SSO0468	Predicted transcriptional regulator fused phosphomethylpyrimidine kinase	$2E^{-15}$
	OE5243F	Transducer protein car	SSO0669	Mn-dependent transcriptional regulator (DtxR family)	$2E^{-18}$
E. coli K12 isolate: MG1665	icd	Isocitrate dehydrogenase	SSO2182	Isocitrate dehydrogenase	$1E^{-108}$
	lysS	Lysine tRNA synthetase, constitutive; suppressor of ColE1 mutation in primer RNA	SSO0090	Lysyl-tRNA synthetase (class II)	$8E^{-103}$
	lysU	Lysyl-tRNA synthetase, heat inducible	SSO0090	Lysyl-tRNA synthetase (class II)	$4E^{-107}$
	tpx	Heat shock protein, integral membrane protein	SSO2613	Peroxiredoxin	$7E^{-14}$
L. lactis IL1403	dltA	D-alanine activating enzyme	SSO2523	Acyl-CoA synthetase (AMP-forming)/ AMP-acid ligase II	$3E^{-15}$
	hisS	Histidyl-tRNA synthetase	SSO0279	Histidyl-tRNA synthetase	$3E^{-48}$
	mleS	Malolactic enzyme	SSO2869	Malic enzyme	$4E^{-13}$
	optC	Oligopeptide ABC trasporter permease protein	SSO1283	ABC-type dipeptide/oligopeptide/ nickel transport system, permease component	$3E^{-16}$
	serS	Seryl-tRNA synthetase	SSO0602	Seryl-tRNA synthetase	$2E^{-67}$
B. subtilis strain 168	BSU1242	Unknown	SSO0176	ATPase of the AAA+ class	$4E^{-23}$
	BSU00940	Cysteinyl-tRNA synthetase	SSO2280	Cysteinyl-tRNA synthetase	$1E^{-98}$
	BSU1322	Related to the activation of sigma-H	SSO1893	ABC-type cobalt transport system, ATPase component	$2E^{-74}$
	BSU1676	Aspartate-semialdehyde dehydrogenase	SSO0876	Aspartokinase	$3E^{-24}$

3.1.4. Comparison with other prokaryotic- and eukaryotic-p-proteomes

In currently available prokaryotic p-proteome analyses of *B. subtilis* strain 168 [70], *E. coli* K12 isolate MG1655 [71], *L. lactis* IL1403 [72] and from the halophile *H. salinarum* R1 [46], 69-79 p-proteins were observed. In the eukaryotic p-proteome study of *Zea mays* L. ecotype B73 leaves, 125 unique p-proteins were identified [189]. In all these studies, the p-proteins were enriched via TiO_2 beads or TiO_2 MOAC prior to p-peptide identification via MS. With the PAcIFIC approach [179] applied here a 4.3 up to 7.9 fold higher number of p-proteins was detected, suggesting that either PAcIFIC is superior to the TiO_2-enrichment or that the number of p-proteins in *S. solfataricus* is exceptionally high. To get a first estimate for prokaryotes, the total gene count was compared with the total number of identified p-proteins (Table 9 ,Figure 6). Of course the number of genes does not represent the actual number of proteins available at a certain time under specific growth conditions, but compared to 1-3% p-proteins identified in previous studies, the 17.3% phosphorylation rate is tremendously high in *S. solfataricus*.

Table 9 Comparison of genomes and p-proteome features of available prokaryotic p-proteome analyses.

Organism	D.o.l. (*)	Genome size [bp]	Gene count	P-proteins [Nr.]	Phosphorylation [%]
E. coli	Bacteria	4,639,675	4497	69	4.5
B. subtilis	Bacteria	4,215,606	4354	78	1.8
L. lactis	Bacteria	2,365,589	2434	63	2.6
H. salinarium	Archaea	2,668,776	2801	79	2.8
S. solfataricus	Archaea	2,992,245	3141	542	17.3

Furthermore, the calculated Ser/Thr/Tyr ratios from previous studies revealed a clear preference for Ser and Thr phosphorylation, with minimal P-Tyr (<10%). In contrast, as detailed earlier, Tyr is by far the most abundant detected p-site (53.6 %) in *S. solfataricus*, followed by Ser (25.8 %) and Thr (20.6 %, Table 8). The question arises if this is an evolutionary adaptive trait of the thermoacidophilie *S. solfataricus*, the first (hyper)thermophile studied so far, or if this bias is based on the enrichment method used in previous studies. The method might discriminate against p-Tyr peptides resulting in the enrichment for/of p-Ser and p-Thr. Unfortunately, no other information about p-proteomes of (hyper)thermophiles and Archaea, except *H. salinarum,* which comprises a different life style, is available. Therefore, future studies have to be awaited in order to decide whether this is a method- or species- or a (hyper)thermophile-specific feature.

3.1.5. Glucose vs Tryptone

As described above 540 unique p-proteins (622 including 82 common p-proteins) were identified. In response to carbon source, an equal number of unique p-proteins (311 p-proteins) were detected in glucose versus tryptone grown cells resulting in no significant change in the Ser/Thr/Tyr ratio (Table 7). Comparison of the arCOG functional code profiles in the p-proteome under glucose and tryptone conditions revealed only slight variations (Figure 9 A). By focusing on the arCOGs, associated with cellular metabolism (arCOG C, E-I, P-Q), analysis revealed that numbers of phosphorylated members of the arCOGs amino acid (arCOG E), nucleotide (arCOG F), lipid and inorganic ion (arCOG P) transport/metabolism are higher under glucose conditions. In contrast, phosphorylation of members of the arCOGs carbohydrate (arCOG G) and coenzyme (arCOG H) transport/metabolism were enhanced with tryptone as carbon source. In addition, the members of the arCOGs secondary metabolite (arCOG Q) and energy production/conversion (arCOG C) seem to be unaffected regardless of carbon source (Figure 9 B).

Together with the previously reported finding that the *S. solfataricus* CCM is not significantly controlled at the transcriptional or translational level in response to glucose and tryptone [26], our p-proteome results let suggest that *S. solfataricus* tunes metabolism and many other cellular processes by changing the phosphorylation profiles of the proteins involved.

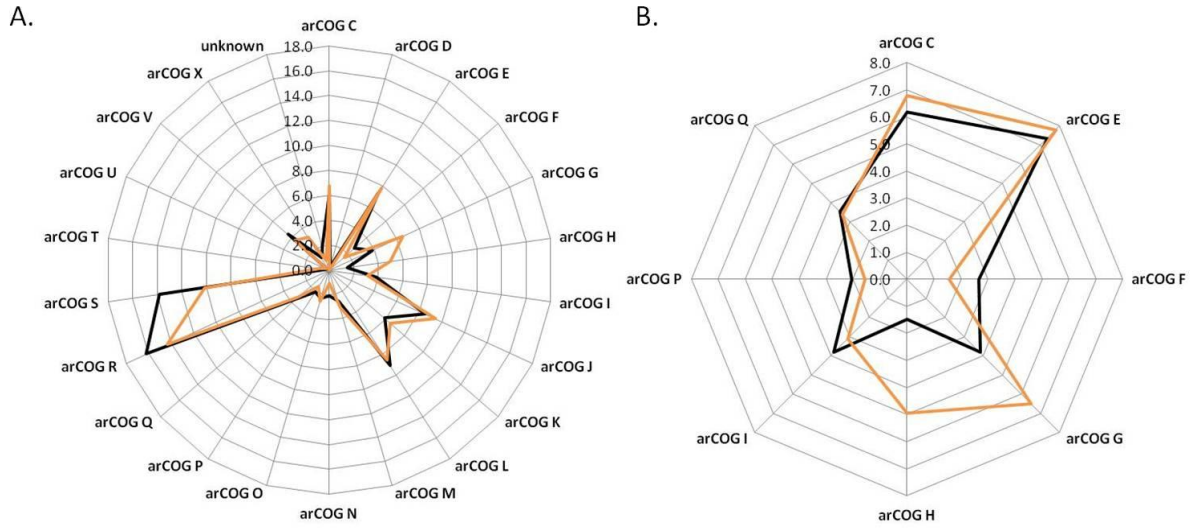


Figure 9 Percentage distribution of p-proteins in all (A) and arCOG functional categories associated with cellular metabolism (B) detected in tryptone (orange) and glucose (black) grown cells.
For functional classes see Figure 8 or for more details [195].

3.1.6. Central carbohydrate metabolism

Although the CCM in Archaea has been studied in some detail, its regulation is still ambiguous. Bioinformatics and biochemistry has revealed that Archaea harbor many new unusual enzymes and

pathways (e.g. the reversible ribulose monophosphate pathway for pentose generation) and well established classical control points from Bacteria and Eukaryotes are missing [3].

Table 10 Identification of archaeal target proteins.

Accession Nr.	Target protein	Predicted function	Organism	Information	Reference
Q973R9	ST0829	FHA-domain containing protein	*Sulfolobus tokodaii*	Phosphorylated by ST1565	[55]
-	-	Methyltransferase activating protein	*Methanosarcina barkerii*	Autophosphorylation with ^{32}P	[196]
Q76KA7	TK0925	Phenylalanyl-tRNA synthase subunit β-chain	*Thermococcus kodakaraensis*	Antiphospho-Tyr antibody	[106]
Q5JH20	TK1404	Phosphomannomutase	*Thermococcus kodakaraensis*	Antiphospho-Tyr antibody	[106]
Q980S1	SSO0207	(*) Hexosephosphate mutase	*Sulfolobus solfataricus*	Phosphorylated on a Ser-residue	[197]
D4GYZ1	HVO_1562	β-subunit of 20S proteasome	*Haloferax volcanii*	Phosphorylated on Ser129	[198]
B0R4J9	CheA	Taxis sensor His kinase cheA	*Halobacterium salinarium*	Phosphorylated with [γ-^{32}P]ATP & Mg^{2+}	[38]
B0R4K1	CheY	Response regulator cheY	*1. Halobacterium salinarium*	2. Phosphorylated on Asp-residue	[38]
D9PYR8	MTH1412	Cell division control protein 6 homolog 1	*Methanobacterium thermoautothrophicum*	Autophosphorylation with ^{32}P on Ser-residue	[199]
D9PU97	MTH1599	Cell division control protein 6 homolog 2	*Methanobacterium thermoautothrophicum*	Autophosphorylation with ^{32}P on Ser-residue	[199]
Q980N4	SSO0257	Cell division control protein 6 homolog 1	*Sulfolobus solfataricus*	Autophosphorylation with ^{32}P on Ser-residue	[200]
Q8ZYK1	PAE0737	Cell division control protein 6	*Pyrobaculum aerophilum*	Phosphorylated on Ser-residue	[199]
O58655	PH0961	Translation initiation factor 2 subunit α	*Pyrococcus horikoshii*	Phosphorylated by hPKR from human	[201]
P95928	SSO2154	Zn-dependent aminopeptidase	*Sulfolobus solfataricus*	Phosphorylated on Ser- or Thr-residue	[202]
-	-	(*) α-subunit of succinyl-CoA synthetase	*Sulfolobus solfataricus*	^{32}P incorporation	[203]
-	-	Glycogen synthase	*Sulfolobus acidocaldarius*	^{32}P incorporation	[204]
Q97U27	SSO3198	Gluconate dehydratase	*Sulfolobus solfataricus*	^{32}P-labelled, no enzyme activity after incubation with phosphatase	[57]
Q980A0	SSO0417	(*) Phosphoglycerate mutase	*Sulfolobus solfataricus*	Phosphorylation at Ser59	[58]
Q9HLV3	Ta0122	2-keto-3deoxy-gluconate kinase	*Thermoplasma acidophilum*	No enzyme activity after incubation with phosphatase	[59]

Previously, reversible protein phosphorylation has been demonstrated in *S. solfataricus*, however, until now the respective signal transduction pathways are unknown and only few p-proteins involved

in CCM have been identified and some were shown to represent p-enzyme intermediates (e.g. hexosephosphate mutase SSO0207, α subunit of succinyl-CoA synthetase SSO2482 , iPGAM SSO0417 [33,197]) rather than targets of regulatory protein modification (e.g. GAD SSO3198 [57]) (Table 10).

3.1.7. Hexose degradation

S. solfataricus uses an unusual branched, promiscuous ED pathway for glucose and galactose degradation [3,24,28,109] and the EMP pathway for gluconeogenesis [26] (Figure 10). In contrast to the classical ED pathway in Bacteria, in the archaeal modification, substrate phosphorylation takes place either at the level of KDG/KDGal (spED branch) or glycerate forming 2-PG (npED branch).

Our PACIFIC p-proteome analysis reveals that well characterized ED enzymes are targeted by protein phosphorylation (Figure 10). The initial enzyme of the pathway the multifunctional GDH-1 (SSO3003), which is active in addition to a glucose specific GDH-2 (SSO3204), was phosphorylated on Tyr134 (p-peptide –YLVKIPK-) under both growth conditions [109,113]. GDH-2, which is specific for glucose [113], is phosphorylated on Tyr324, Ser326, Thr333 (-YPSVLNRMITR-) under tryptone conditions (only one biological replicate). This seems to support previous suggestions, that GDH-1 is active in different sugar degradation pathways (e.g. D- galactose, D-xylose and L-arabinose [113]), whereas GDH-1 is involved in glucose degradation. Therefore GDH-1 might be down regulated on both glucose and tryptone and GDH-2 only on tryptone.

At the branch point of the branched ED pathway KDG/KDGal is either phosphorylated by the KDG kinase forming KDPG/KDPGal in the spED branch or cleaved directly by the multifunctional KD(P)G aldolase to GAP and pyruvate (npED branch). In our study KDG kinase (SSO3195) was phophorylated in the tryptone grown cells (Thr313; p-peptide -FLNEFKT-) again suggesting an inhibitory effect. The KDG kinase of *T. acidophilum*, which also utilizes the branched ED pathway, has already been shown to be phosphorylated [59]. However, the *T. acidophilum* KDG kinase is a member of a different enzyme family (Pfam 00294).

For the remaining reactions of the branched ED pathway, no functional characterized p-proteins were detected under the chosen growth conditions. The previously reported phosphorylation of iPGAM (SSO0417) [58] on the catalytically important Ser59 (p-enzyme intermediate) and GAD (SSO3198; regulatory protein phosphorylation, p-protein active [57]) was not observed under both growth conditions.

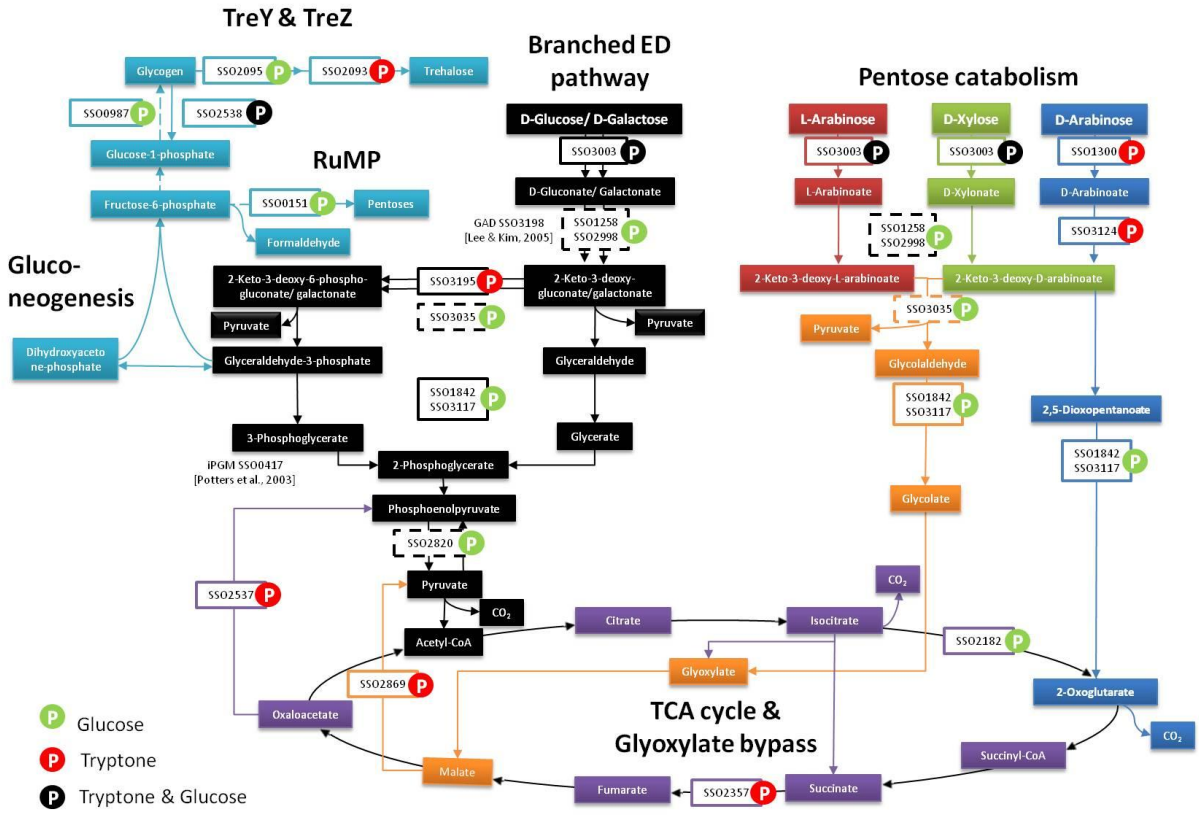


Figure 10 Current understanding of hexose and pentose degradation, glycogen metabolism, trehalose synthesis, TCA cycle and C4/C3 interconversion in *S. solfataricus* P2.
Shown is the branched ED pathway for glucose and galactose degradation [24,28,57,109-111,132] as well as the pathways for D-arabinose [27] and D-xylose and L-arabinose [25] catabolism. Phosphorylated enzymes are indicated by; color indicates the growth conditions under which the p-protein was observed: green, glucose; red, tryptone; violet, both growth conditions. Boxes with dashed lines indicate candidates of unknown function. Only p-proteins identified in both replicates are depicted.

In the p-proteome of *H. salinarum* [46], which is also assumed to degrade glucose via a spED variant, the only glycolytic enzyme shown to be phosphorylated is the PyrK common to both the ED and EMP pathway. P-proteome studies in Bacteria (i.e. *B. subtilis*, *E. coli* and *L. lactis*) and Eukaryotes (*Zea mays*) relying on the EMP pathway for glucose degradation revealed that a majority of their glycolytic enzymes are targets of phosphorylation [70-72,189]. For example, in the case of *B. subtilis* PGI, FBPA, TPI, GAPDH, PGK, iPGAM, ENO and PyrK were shown to be phosphorylated [70]. However, in contrast to the ED enzymes of *S. solfataricus*, most of these enzymes except for PyrK are also functional in the gluconeogenetic direction (see below).

3.1.8. Pentose degradation (D-arabinose, D-xylose and L-arabinose)

Protein phosphorylation also seems to play a major role in pentose degradation in *S. solfataricus*. This is particularly evident for the pathway for D-arbinose degradation to 2-oxoglutarate in the TCA cycle [27]. Notably, three out of four enzymes were found to be phophorylated D-arabinose

dehydrogenase (SSO1300, Tyr320 (p-peptide -IKPYIIK-) in glucose grown cells), D-arabinonate dehydratase (SSO3124, multiple-phosphorylated on Tyr12 (p-peptide -LCYEGINDER-), Tyr88 and Thr91 (p-peptide –MYKATLR-) in tryptone grown cells), and DOPDH (SSO3117, Tyr336 and Tyr339 (p-peptide -KDLEYIEYGK-) in glucose grown cells; Thr244 (-NRMTR-) in tryptone grown cells (detected in only one biological replicate)). Since cells were grown in the presence of glucose and tryptone, it is tempting to speculate that the observed phosphorylation pattern shut down the D-arabinose pathway under the assayed growth conditions.

The pathways for degradation of D-xylose (Dahms pathway) and L-arabinose were proposed recently, which are characterized by the action of the multifunctional ED GDH-1 and aldol cleavage by the multifunctional ED KD(P)G aldolase. In addition, D-xylose seems to be degraded in parallel via the Weimberg pathway, which omits aldol cleavage [25]. The pathway was recently demonstrated in the halophilic Archaeon *H. volcanii* [126]. Of the proposed candidates, only enzymes involved in the branched ED pathway and the D-arabinose degradation pathway, i.e. GDH-1 SSO3003 and 2,5-dioxopentanoate dehydrogenase SSO3117, were observed to be phosphorylated. However, as detailed below there are several paralogs in the *S. solfataricus* genome of unknown function and the final elucidation of these candidates has to be awaited.

3.1.9. Candidates of unknown function

In the course of the reconstruction of the CCM pathway in *S. solfataricus* (SulfoSYS, http://www.sulfosys.com/), several additional paralogs of glycolytic enzymes were identified for which a functional prediction is not possible from sequence alone [108]. Among these candidates, three sugar acid dehydratases (SSO1258 (Tyr355, Ser356 and Tyr357; p-peptide -YSYKNRLR-), SSO2998 (Tyr6, p-peptide -MELKITDFK-), SSO1303 (Tyr20, Thr23, Thr24; p-peptide -IDKIYVVTTER-; Tyr109, p-peptide - INMVWDLLYR-; only in one replicate detected) of unknown substrate specificity, which might catalyze the second step in C_6 or C_5 degradation (Figure 10) were found to be phosphorylated. SSO1258 and SSO2998 were detected in glucose grown cells and SSO1303 phosphorylation was detected in tryptone grown cells.

Notably, also, one homolog (SSO3035) of the multifunctional KD(P)G aldolase (SSO3197), reveals two p-sites (Tyr323 and Tyr328; p-peptide -LYDLLLEYR-) detected only in glucose grown cells. Blast-P searches reveal high similarity to the recently characterized KDG aldolase of *P. torridus* (Pto1279, 5e-25; 26% amino acid identity), which is specific for dephosphorylated substrates (i.e. GA and pyruvate) [205]. This suggests the presence of a second aldolase, which is specific for the np-ED branch to be active in this organism in addition to the multifunctional KD(P)G aldolase. The latter has been proposed to be active in both ED branches [28,110] as well as D-xylose and L-arabinose degradation

[25]. It is tempting to speculate that regulation of KDG kinase (see above) and the predicted KDG aldolase by phosphorylation might allow for the regulation of carbon flow between both ED branches. The characterization of the respective candidates is still in progress.
In addition, another member of the aldehyde dehydrogenase superfamily (SSO1842, gapN-2) [206], a paralog of the characterized GAPN [116], and DOPDH of *S. solfataricus* [27] was detected to be phosphorylated in glucose grown cells (Ser436 (p-peptide -WDSLPFGGFKK-)).
Nonetheless, the functional analysis for the homologs (SSO1258, SSO1303, SSO2998, SSO3035) has to be awaited (Figure 10) in order to clarify their possible role in hexose or pentose degradation.

3.1.10. Gluconeogenesis, glycogen metabolism and trehalose synthesis

In the gluconeogenic EMP pathway, no p-proteins were detected in *S. solfataricus* [188]. This finding is consistent with the p-proteome analysis of *H. salinarum* [46]. In contrast, observations of bacterial p-proteomes, showed that most EMP enzymes, especially those working in both the gluconeogenetic and glycolytic directions, are phosphorylated, which let suggest that phosphorylation might trigger the direction of the pathway flux [70-72]. Th regulationmight not be necessary in the case in *S. solfataricus* and *H. salinarum* though, since they use different pathways for glycolysis and gluconeogenesis. However, many of the enzyme homologs predicted to be involved in the first steps of gluconeogenesis, i.e. PEP synthesis from pyruvate and oxaloacetate as well as pyruvate synthesis from malate, were found to be phosphorylated in *S. solfataricus*. The GTP-PEP carboxykinase homolog (SSO2537, Tyr536 and Tyr541, p-peptide –EYGKEEYEK-) and the malic enzyme homolog (SSO2869, (Tyr3 and Ser13, p-peptide –IYFFMVDAVELSR-; Tyr16 and Tyr23, p-peptide –KYEGKIEIYPK-; Tyr242 and Thr248, p-peptide -WKYELAIKTNK-) were phosphorylated only in the presence of tryptone, pointing to an activation under gluconeogenetic conditions. Furthermore, SSO2820, which is annotated as PEPS/pyruvate phosphate dikinase (PPDK), was four times phosphorylated (Ser22, Tyr24 and Tyr29,p-peptide -KSAYLGELYK-; Tyr298, p-peptide -VYLLEVR-) under glucose grown condition and three times (Ser22, Tyr24 and Tyr29,p-peptide -KSAYLGELYK-; only detected in one replicate) under tryptone conditions. In accordance with these structural observations, biochemical studies revealed no PEPS or PPDK enzyme activity [Haferkamp *et al.,* manuscript in preparation]. Thus, the function of SSO2820 remains unclear. Furthermore, the functions of the other phosphorylated homologs of gluconeogenetic enzymes need to be proven and the effect of phosphorylation/dephosphorylation on protein functions remainto be elucidated.
In the metabolism of the long term carbon and energy storage compound glycogen [108] the glycogen synthase (SSO0987, Thr532, p-peptide VRENAITRVK) was found to be phosphorylated on glucose. One sugar phosphate nucleotydyl transferase (SSO0989), the encoding gene is organized in a

conserved gene cluster with the gene encoding glycogen synthase, as well as the degrading glycogen phosphorylase (SSO2538) possesses a different phosphorylation pattern under both growth conditions indicating a subtle regulation (SSO0989: Tyr235, Tyr238 and Tyr240, p-peptide –GYPVYGYPMK-, tryptone (only one replicate), Tyr215, p-peptide -EMYKMGK-; glucose); SSO2538: Tyr438, Tyr439, Tyr443 and Tyr446, p-peptide –YYEVGYNAYK-, tryptone (only one replicate), Tyr462, Tyr464 and Tyr465, p-peptide -LMKEYGYY-, glucose and tryptone).

The compatible solute trehalose is synthesized from glycogen via the TreY-TreZ pathway [108]. Maltooligosyltrehalose synthase, which catalyzes the the conversion of the terminal α-1,4-linkage of the glucose polymer into an α-1,1-linkage (TreY, SSO2095, Tyr190 and Tyr191, -LLLMQYYR-,) was phosphorylated under glycolytic growth conditions. On tryptone maltooligosyltrehalose hydrolase, which releases the terminal α-1,1-linked disaccharide, i.e. trehalose (TreZ, SSO2093, Tyr550 and Tyr559, -YELDKGFALYK- (both replicates); Ser341, Tyr348 –SYKDVFIYDGK- (one replicate) and Thr2 and Tyr5 - MTFGYK- (one replicate)) was found to be phophorylated. Therefore, it is suggested that glycogen metabolism as well as trehalose synthesis are profoundly regulated by phosphorylation.

3.1.11. RuMP pathway and TCA cyle

Pentoses are generated in *S. solfataricus* P2 *via* the reversed RuMP pathway *via* 3-hexulose-6-phosphate isomerase and 3-hexulose-6-phosphate synthase, which catalyze the formation of ribulose 5-phosphate and formaldehyde from F6P [108]. The first enzyme of this pathway, 3-hexulose-6-phosphate isomerase was phosphorylated in glucose grown cells (SSO0151, Tyr45, -MVNELENFYK-), which might channel the flux into the RuMP pathway.

In accordance with its aerobic life style, *S. solfataricus* possesses a fully operative TCA cycle [108] and two of its enzymes, IDH (Ser84, -FPKESEELIEK-) and succinate dehydrogenase (SDH, Thr263 and Tyr272, -TGKLAEAPVYLK-) are phosphorylated. IDH is also phosphorylated in *H. salinarum* [46] (Table 8). In *E. coli*, IDH is regulated via the two-component-system AceK, which possesses PyrK and PP activity (p-site at Ser113) [207,208]. The dephosphorylated IDH is active in the TCA cycle; upon phosphorylation the enzyme is inactivated and isocitrate is converted to glyoxylate by the isocitrate lyase [209]. But so far, no AceK homologs were found in the third domain of life [33]. The p-sites are different, suggesting different underlying regulatory principles. SDH phosphorylation was observed in *S. solfataricus* grown on tryptone as well as in *H. salinarum*. This might be a special archaeal feature of TCA cycle regulation since SDH is not phosphorylated in any of the bacterial p-proteomes [46,70,71]. Conversely, in Bacteria, the malate dehydrogenase in particular seems to be under the control of phosphorylation [70,71].

3.1.12. Conclusion

This comparative enrichment free p-proteome study from *S. solfataricus* grown on glucose and tryptone identified 690 p-peptides from 540 unique p-proteins. It is the first PACiFIC-mediated p-proteome study. 82 overlapping p-proteins were identified in glucose and tryptone grown cells. 1318 p-sites were detected overall with a Ser/Thr/Tyr ratio of 25.8/20.6/53.6%. Categorization of p-proteins using the new *S. solfataricus* genome annotation (according to arCOGs) reveals that proteins of 21 out of 26 functional categories are protein phosphorylation targets. This highlights the likelihood that many processes in *S. solfataricus* are regulated by reversible protein phosphorylation. Interestingly, compared to previous p-proteome studies, we observed strong Tyr phosphorylation. The reasons for this are not yet clear.

In light of missing information on archaeal CCM regulation, our focus was on the phosphorylation of CCM enzymes in response to offered carbon source. Our results demonstrate thus sugar metabolism in *S. solfataricus* seems to be highly controlled by protein phosphorylation and dephosphorylation. This particularly applies to enzymes at the beginning of the different sugar degradation pathways (dehydrogenases, dehydratases), as well as of gluconeogenesis (PEP carboxykinase (GTP), malic enzyme) and to glycogen metabolism and trehalose synthesis via the TreY-TreZ pathway. Further, the enzymes at the branch point between the npED and the spED branch (i.e. KDG kinase and predicted KDG aldolase) as well as between the TCA cycle and the glyoxylate bypass (i.e. IDH) and the beginning of the RuMP pathway (i.e. 3-hexulose-6-phosphate isomerase) were the subject of (de)phosphorylation, suggesting a tight regulation of pathways in respect to different carbon sources and utilization of different branches. The observed significant changes in the phosphorylation pattern in response to the carbon source suggest that sugar degradation pathways are blocked in the presence of tryptone, whereas gluconeogenetic enzymes are activated. However, so far, the effect of phosphorylation is only reported for the *S. solfataricus* GAD, which is only active as a p-enzyme [57]. Thus, future studies need to elucidate the role of protein phosphorylation on target proteins and to identify the respective protein kinases and phosphatases in order to unravel the signal transduction pathways.

3.2. (Cren)Archaeal Protein Phosphorylation: In Vivo Impact of Protein Phosphatase Deletions on Cell Size, Motility and Energy Metabolism in Sulfolobus acidocaldarius

The cloning of the putative PTP gen *saci_0545* as well as the heterologous expression in *E. coli* and the purification of both protein phosphatases (Saci_0545, Saci_0884) was performed by Dominik Esser. Also the detailed biochemical characterization of both protein phosphatases with para-nitrophenylphosphate and the different p-peptides was performed by Dominik Esser. Furthermore, the arCOG evaluation of the RNA-Seq data and analysis of the p-proteome data was performed and discussed by Dominik Esser under supervision of Bettina Siebers. Deletion mutant construction, microscopic size measurement, determination of growth rates, complementation studies, growth of cells for RNA-Seq, fermentation of all three strains for p-proteome analysis, nutrient limitation assays and motility assays were done by Julia Reimann (AG Albers, Max Planck Institut for Terrestria Microbiology, Marburg, Germany), supervised by Sonja-Verena Albers. Trong Khoa Pham (ChELSI Institute, Department of Chemical and Biological Engineering, The University of Sheffield) performed the p-proteome analysis via PAcIFIC. Alvaro Orell (AG Albers, Max Planck Institut for Terrestria Microbiology, Marburg, Germany) prepared RNA for the RNA-Seq approach, which was performed together with Fabian Amman (Institute for Theoretical Chemistry, University of Vienna), who prepared the RNA-Seq dataset. q-PCR studies were performed by Alvaro Orell. Influence of the deletion mutants on cell cycle and cell size via flow cytometry was investigated by Ann-Christin Lindås and Rolf Bernander (Department of Molecular Evolution, Biomedical Center, Uppsala University). The results of this study were submitted in Molecular and Cellular Proteomics and are currently under revision by the authors.

3.2.1. Introduction

Phosphorylation is one of the most important PTMs of proteins together with methylation, acetylation, glycosylation,ubiquitination and sampylation. In particular, the reversible character of protein phosphorylation allows for subtle and immediate regulation by modulating protein activity in various pathways and processes of the cell [210-212]. The addition of a phosphorylgroup to proteins is catalyzed by protein kinases and dephosphorylation by their cognate phosphatases. Phosphorylation has been reported in all three domains of life, the Eukarya, the Bacteria and the Archaea [46,70,212]. Interestingly, in the Crenarchaeota, one of the phyla of the Archaea, only Ser/Thr/Tyr phosphorylation is predicted by bioinformatics analyses and two-component systems (involving His/Asp phosphorylation) identified in Euryarchaeota are absent [33,184,185,213].

A p-proteome study of the thermoacidophilic crenarchaeon *S. solfataricus* P2 revealed a vast amount of Ser/Thr/Tyr-phosphorylated proteins (540 detected in total) in almost all arCOGs categories with an unexpected high number of Tyr phosphorylation. A differential phosphorylation pattern was observed in cells grown on glucose versus tryptone. Focused on the CCM, an important role of protein phosphorylation in regulating the glycolytic flux in *Sulfolobus* was proposed [188]. Besides the p-proteome of the mesophilic euryarchaeon *H. salinarum* (69 p-proteins in total) [46], this was the first high throughput study highlighting the importance of regulatory protein phosphorylation in Archaea. Although a few archaeal eukaryotic-like protein kinases and phosphatases were investigated in more detail [33,47,48,50-53,55,101,103,106], there is still a knowledge gap regarding the impact of Ser/Thr/Tyr protein phosphorylation on *in vivo* cellular processes and the signal transduction pathways involved in Archaea. Comparative database searches revealed only two protein phosphatases encoded in the genomes of all *Sulfolobales*, in comparison to 8 predicted eukaryotic protein kinases for example in *S. solfataricus* [33]. Similarly, in Eukaryotes, fewer protein phosphatases are encoded in the genome compared to protein kinases, though the phosphatases of the PPP family often act as multimeric proteins with different catalytic, regulatory and core subunits [99,214]. Protein phosphatases can be categorized into different families. On one hand, Ser/Thr specific protein phosphatases (PPPs) contain the 220 amino acid long catalytic domain including motif I (GDXHG), motif II (GDXXDRG), and motif III (GNHE) [215], and the Mg^{2+} or Mn^{2+} dependent Ser/Thr phosphatases (PPMs), which contain 11 specific motifs [95,216]. Conversely, the phosphotyrosine phosphatase family (PTP) can be subdivided into the PTPs, which are specific for Tyr dephosphorylation, the dual specific PTPs, which can dephosphorylate Ser/Thr, and Tyr, and the low molecular weight PTPs. They all share one common amino acid motif, CX_5R [99]. Only a few archaeal Ser/Thr or Tyr specific protein phosphatases are characterized so far. One example is the Ser/Thr phosphatase PP1-arch1 from *S. solfataricus*, which revealed Mn^{2+} dependent protein phosphatase activity *in vitro* [48]. Furthermore, the crystal structure of SsoPTP from *S. solfataricus* was solved and the specificity towards p-Tyr determined *in vitro* [47]. To our knowledge, only one other archaeal PTP was characterized hitherto, the *Tk*-PTP from the euryarchaeon *T. kodakaraensis* KOD. In this analysis, *Tk*-PTP exhibited phoshotyrosine- as well as phosophoserine-, but no p-Thr activity [106]. Finally, the only determined archaeal member of the PPM family is evident in *Thermoplasma volcanium* and revealed an *in vitro* divalent metal-ion dependent dual-specificity towards phosphorylated Ser/Thr as well as Tyr residues [105].

Like in all Crenarchaea, the model organism *S. acidocaldarius*, which grows optimally at pH 3 and 76°C, comprises only two protein phosphatases, one predicted Tyr phosphatase Saci-PTP (encoded by the gene *saci_0545,* called *saci_ptp*) and one Ser/Thr phosphatase, Saci-PP2A (encoded by the

gene *saci_0884*, called *saci_pp2a*), which is a homolog to the PP2A catalytic subunit of eukaryotic phosphatases.

In this study, the two *S. acidocaldarius* protein phosphatases, Saci-PTP and Saci-PP2A, were biochemically characterized with respect to their substrate specificity and kinetic properties *in vitro*. Furthermore *in vivo* analyses were performed with both phosphatase deletion mutants (Δ*saci_ptp* and Δ*saci_pp2a)* to determine the corresponding phenotypes. For additional characterization, p-proteome analysis and RNA-sequencing were used to elucidate the effect of single phosphatase deletions compared to the background strain MW001, respectively. A major effect on cell motility and energy metabolism was observed which could be confirmed by in vitro/ and in vivo analysis of cell motility indicating the importance of protein phosphorylation based signal transduction cascades in crenarchaea.

3.2.2. Expression and enzymatic characterization of Saci-PTP and Saci-PP2A

The two protein phosphatases, Saci-PTP and Saci-PP2A, were heterologously expressed (Figure 11) in *E. coli*, purified (see section 2.7.2.3.) and characterized.

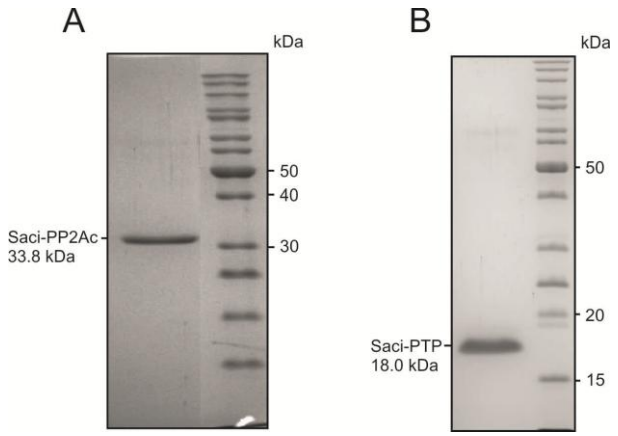


Figure 11 Purification of Saci-PP2A and Saci-PTP.
2µg of Saci-PP2A (A) after size exclusion chromatography and Saci-PTP (B) after immobilized metal affinity chromatography were applied to SDS-PAGE. The SDS-PAGE was stained with Coomassie brilliant blue and PageRuler^{TM} Unstained Protein Ladder (Fermentas) was used as size ladder.

General phosphatase activity of Saci-PTP was observed in the presence of para-nitrophenylphosphate (pNPP) as substrate. At 70°C a V_{max}-value of 3.7 U/mg and a K_m-value of 0.53 mM was determined. Furthermore, substrate inhibition was observed at higher pNPP concentrations (> 2 mM) with a k_{irr}-value of 0.17 (U/mg)mM^{-1} (Figure 12). Saci-PTP activity was verified by using the p-Tyr peptide TEVGKRI(pY)RLVGDKN as substrate, revealing a V_{max}-value of 9.2 U/mg, a K_m-value of 0.53 mM and a Hill coefficient of 2.0 (Figure 12). In addition, activity with the p-

Thr peptide RRA(pT)VA was observed (0.2 U/mg, 250 nM), indicating dual protein phosphatase activity (data not shown).

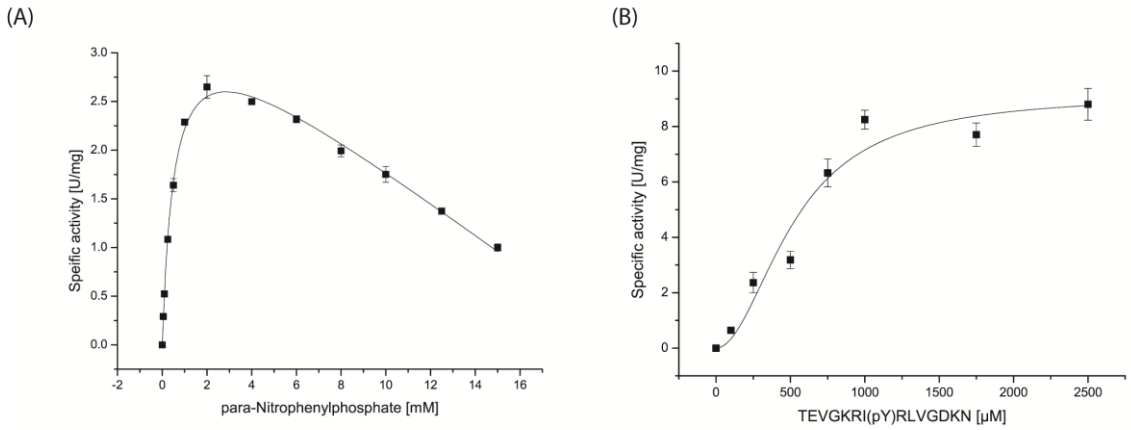


Figure 12 Biochemical characterization of the protein phosphatase Saci-PTP.
The activity of Saci-PTP was verified in the presence of pNPP (A) and of the p-Tyr peptide TEVGKRI($_p$Y)RLV (B). All experiments were performed at 70 °C.

The putative protein serine/threonine phosphatase Saci-PP2A was analyzed for activity with pNPP and the p-Thr peptide RRA(pT)VA (Figure 13). For pNPP, a V_{max}-value of 4.3 U/mg and a K_m-value of 2.5 mM were determined. Furthermore, a lag phase at low substrate concentrations (> 2 mM) was observed with a Hill coefficient of 4.9. For the p-Thr peptide RRA(pT)VA, a V_{max}-value of 192.6 U/mg and a K_m-value of 53.6 µM were detected. No activity was detected with the p-Tyr peptide TEVGKRI(pY)RLVGDKN as substrate.

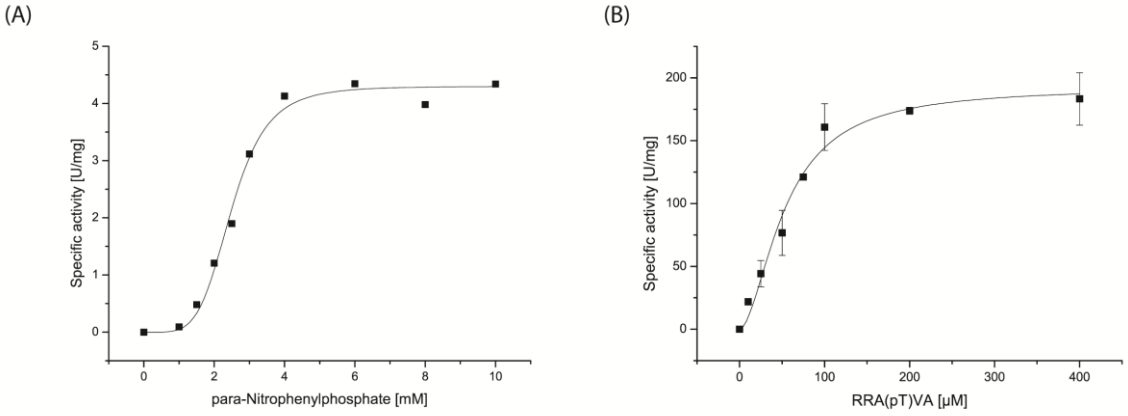


Figure 13 Biochemical characterization of the protein phosphatase Saci-PP2A.
The activity of Saci-PP2A was verified in the presence of pNPP (A) and of the p-Thr peptide RRA(pT)VA (B). All experiments were performed at 70 °C.

In addition, the analysis revealed that Saci-PP2A requires Me^{2+}-ions for activity. The highest activity was observed with Cu^{2+} followed by Mn^{2+}, Ni^{2+}, Mg^{2+}, Cd^{2+} and Co^{2+} in the presence of EGTA

(Figure 14). When EGTA was replaced by EDTA, no PP2Ac activity could be observed. In the absence of Me^{2+}-ions Saci-PP2A was inactive. Activity tests in the presence of okadaic acid, a typical inhibitor of PPPs, revealed a strong inhibitory effect (56 % activity in presence of 10 nM okadaic acid).

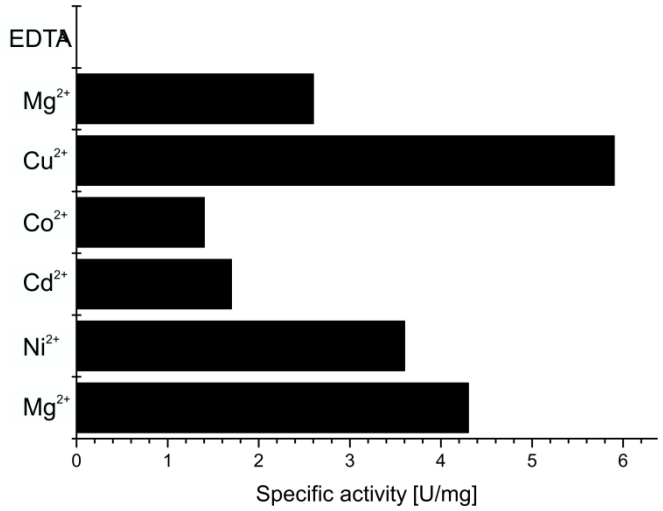


Figure 14 Metal-ion dependency of Saci-PP2A in the presence of EGTA.
Enzyme activity was measured in the presence of 10 mM p-NPP, 5 mM Me^{2+} and 1 mM EGTA. EDTA: EGTA was substituted by EDTA. The highest activity was observed with Cu^{2+} and Mn^{2+}.

3.2.3. Characterization of *saci_ptp* and *saci_pp2a* deletion mutants

Using the uracil auxotrophic background strain of *S. acidocaldarius* MW001, markerless deletion mutants of both phosphatase encoding genes were generated. *Δsaci_pp2a* showed slower growth (doubling time 9.5 h) and reached the stationary phase after 130 h in contrast to MW001 and *Δsaci_ptp*.

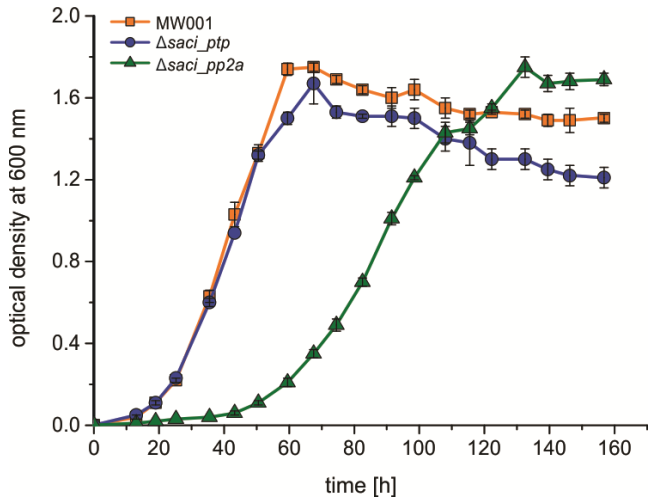


Figure 15 Δ*saci_pp2a* shows a different growth behavior in comparison to MW001 and Δ*saci_ptp*.
All three strains were grown for 160 hours on Brock medium supplemented with 0.1% (w/v) NZ-amine, 0.2% (w/v) sucrose and 10 µg/ml uracil at 76°C. Δ*saci_ptp* was comparable in growth behavior to the background strain MW001, whereas the *saci_pp2a* deletion mutant displayed an extended lag-phase and also a slower growth rate.

These revealed doubling times of 5.7 h and 6.3 h, respectively, and reached the stationary phase after 60 h (Figure 15). Notably, the *Δsaci_ptp* cells appeared normal in cell size, whereas the Ser/Thr phosphatase deletion mutant showed an aberrant cell size distribution in comparison to MW001 (Figure 16). Although the average cell length of MW001 and *Δsaci_ptp* compared to *Δsaci_pp2a* only differed by about 0.2 µm (length of 1.27 µm, 1.30 µm and 1.46 µm for MW001, *Δsaci_ptp* and *Δsaci_pp2a*, respectively), the maximal and minimal cell diameter revealed considerable differences (Figure 16).

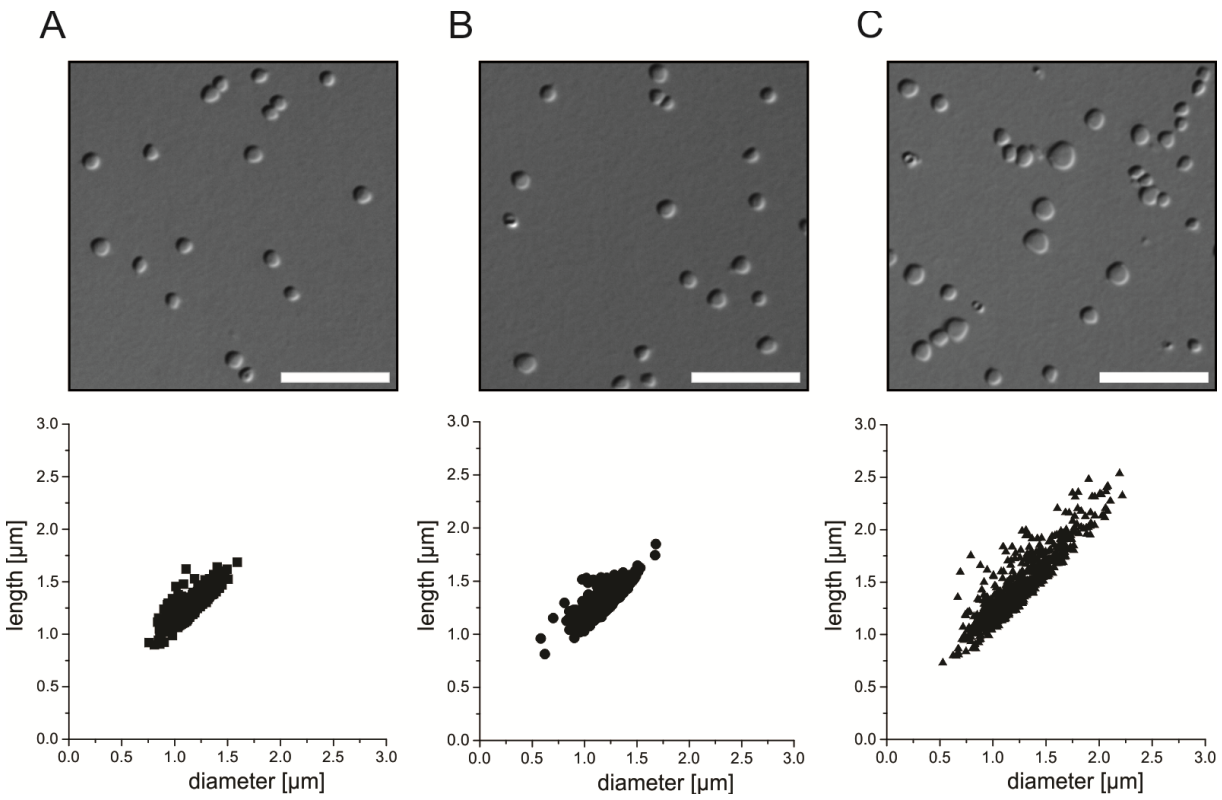


Figure 16 The phosphatase deletion mutant Δ*saci_pp2a* revealed a phenotype in cell shape and size.
MW001 and Δ*saci_ptp* show the regular round-like cell shape known for Sulfolobales under normal growth conditions (A and B, upper panel), whereas the *saci_pp2a* deletion mutant appeared to be more diverse in size with a significantly increased amount of larger cells (C, upper panel). Evaluation of the actual cell size using ObjectJ revealed a spread distribution in Δ*saci_pp2a* especially towards bigger cell sizes (C, lower panel), if compared to the more defined cell size range of the wild type strain MW001 (A, lower panel) and the *saci_ptp* deletion strain (B, lower panel). The calibration bar indicates 10 µm.

The highest divergence was observed in the maximal cell length, which was 1.69 µm for MW001 and 1.85 µm for *Δsaci_ptp*, in contrast to 2.54 µm of the *Δsaci_pp2a* cells. Additionally, in a trans-complementation experiment, the phenotype of *Δsaci_pp2a* could be partially restored with a growth curve and cell shape distribution comparable to the background strain (Figure 17).

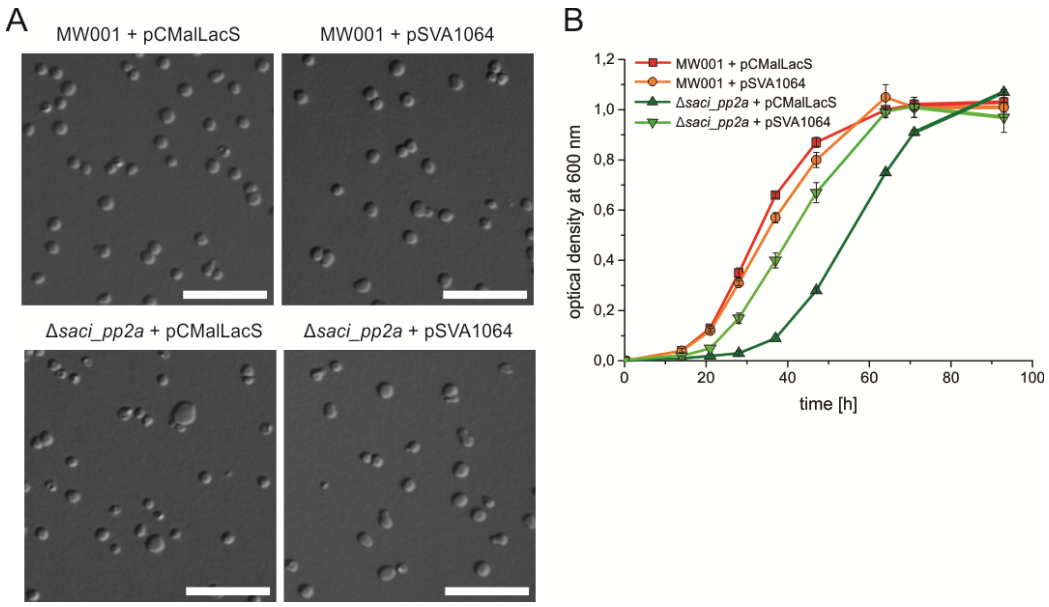


Figure 17 Complementation of the Δ*saci_pp2a* phenotype.
Δ*saci_pp2a* and MW001 were transformed with the control plasmid pCMalLacS and the complementation plasmid pSVA1064. (A) DIC light microscopy images of Δ*saci_pp2a* in comparison to MW001 after transformation with pCMalLacS or pSVA1064. The phenotype in shape was reduced in the complemented strain Δ*saci_pp2a* + pSVA1064, cells are smaller and more homogeneous. (B) A growth curve was performed at 76°C on Brock medium supplemented with 0.1% (w/v) NZ-amine and 0.2% (w/v) sucrose. The complemented strain Δ*saci_pp2a* + pSVA1064 showed a reduced phenotype in growth and grew almost like the wild type strain. The scalebar indicates 10 µm in all pictures.

In order to address if cell cycle control might be affected, flow cytometry was performed. However, no changes in DNA content distributions, reflecting the relative lengths of the cell cycle periods [217], could be detected by flow cytometry (fluorescence; DNA content) in all three strains (background strain and phosphatase mutants, Figure 18).

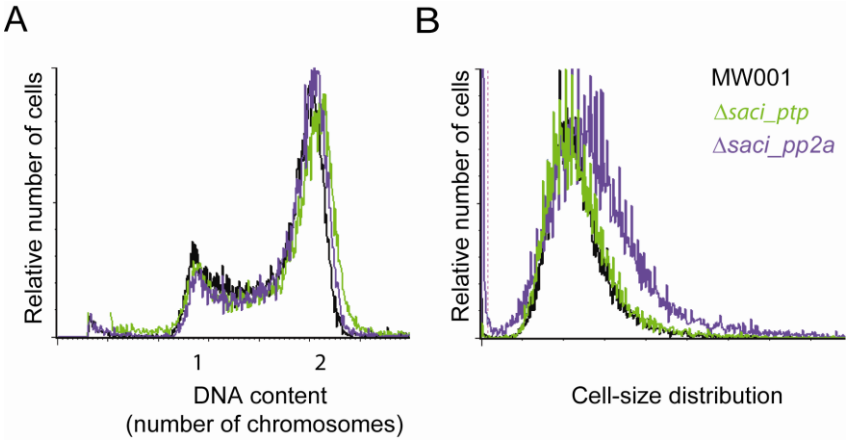


Figure 18 Flow cytometry of MW001 in comparison to *Δsaci_ptp* and *Δsaci_pp2a*.
(A) All three strains show the typical DNA content distribution of exponentially growing *Sulfolobus* cultures [30]. (B) In cell size, however, the Ser/Thr phosphatase deletion mutant (Δ*saci_pp2a*) shows significant larger and an increased proportion of smaller cells than the wild type strain or the Tyr phosphatase deletion mutant *(Δsaci_ptp*).

The cell-size distribution, on the other hand, confirmed a significant difference in the wt and *Δsaci_ptp* compared to the *saci_pp2a* deletion mutant (Figure 18).

3.2.4. Comparative phosphoproteome analysis of MW001, *Δsaci_ptp* and *Δsaci_pp2a*

It was expected that deletion of the phosphatases would lead to the accumulation of p-peptides within the cell, allowing the identification of associated signaling pathways. To examine this possibility, the p-proteomes (including cytoplasmic and membrane proteins) of the three *S. acidocaldarius* strains (MW001, *Δsaci_pp2a* and *Δsaci_ptp*) were analyzed from late exponential phase (optical density of 0.8) using PAcIFIC as described previously for *S. solfataricus* P2 [188]. Two biological replicates were analyzed of the three investigated strains.

Table 11 Comparison of different ph-proteome studies with the three *S. acidocaldarius* p-proteome sets.

Organism	P-proteins	P-peptides	P-sites	p-S	p-T	p-Y	Ratio p-S/p-T/p-Y
	No.	No.	No.	No.	No.	No.	[%]
***S. acidocaldarius* Total**	538	1084	1552	551	418	583	35.5/26.9/37.6
***S. acidocaldarius* MW001**	69	117	164	54	35	75	32.9/21.3/45.7
S. acidocaldarius* Δ*saci_pp2a	143	253	382	117	89	176	30.6/23.3/46.1
S. acidocaldarius* Δ*saci_ptp	407	718	1006	380	294	332	37.8/29.2/33.0
***S. solfataricus* Total [188]**	540	690	1318	340	272	706	25.8/20.6/53.6
***S. solfataricus* glucose**	311	343	580	146	107	327	25.17/18.45/55.38
***S. solfataricus* tryptone**	311	384	810	203	174	433	25.06/21.48/53.46
***Halobacterium salinarum* [46]**	69	90	81	70	10	1	87/12/1
***Lactococcus lactis* [72]**	63	102	73	34	37	2	46.5/50.6/2.7
***Bacillus subtilis* [70]**	78	103	78	54	16	8	69.2/20.5/10.3
***Escherichia coli* K12 [71]**	79	105	81	55	19	7	67.9/23.5/8.6
***Zea mays* leaf [189]**	125	149	157	141	15	1	89.8/9.6/0.6

Overall 1084 p-peptides from 538 unique p-proteins were identified in the three investigated strains with an overall pS/pT/pY %-ratio of 35.5/26.9/37.6. In MW001 (background strain) 117 p-peptides

from 69 p-proteins (pS/pT/pY %-ratio of 32.9/21.3/45.7), in the *Δsaci_pp2a* strain 253 p-peptides from 143 p-proteins (pS/pT/pY %-ratio of 30.6/23.3/46.1) and in the *Δsaci_ptp* strain 718 p-peptides from 407 p-proteins (pS/pT/pY %-ratio of 37.8/29.2/33.0) were identified (Table 11 ,Figure 19, list of p-peptides see Table S1 of [218]).

The number of p-proteins detected in the MW001 and *Δsaci_pp2a* strains are in the same range as the numbers detected in other p-proteome studies (69-143 vs. 69-125; Table 11) [46,70-72,188,189]. However, the number of p-proteins detected in the *Δsaci_ptp* strain is considerably higher than in the other strains (69 vs. 143 vs. 407). The total number of unique p-proteins detected in this study is close to the number of unique p-proteins detected in *S. solfataricus* P2 (538 vs. 540) [188].

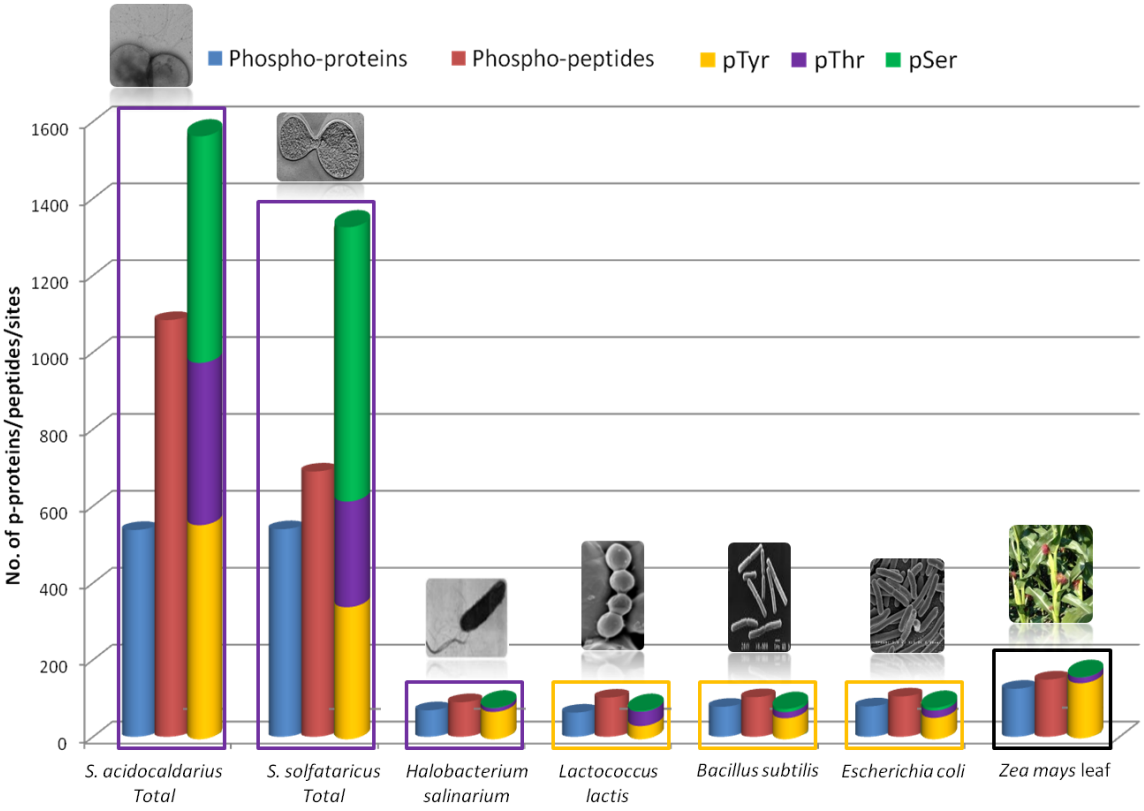
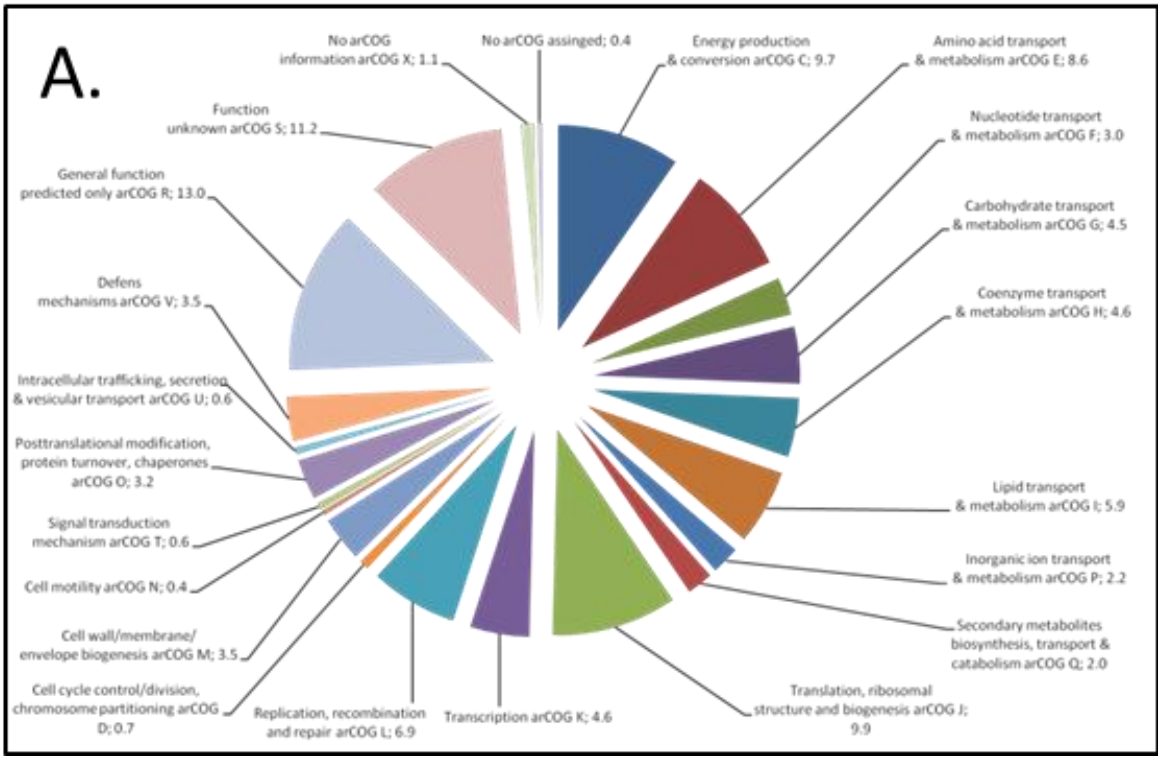
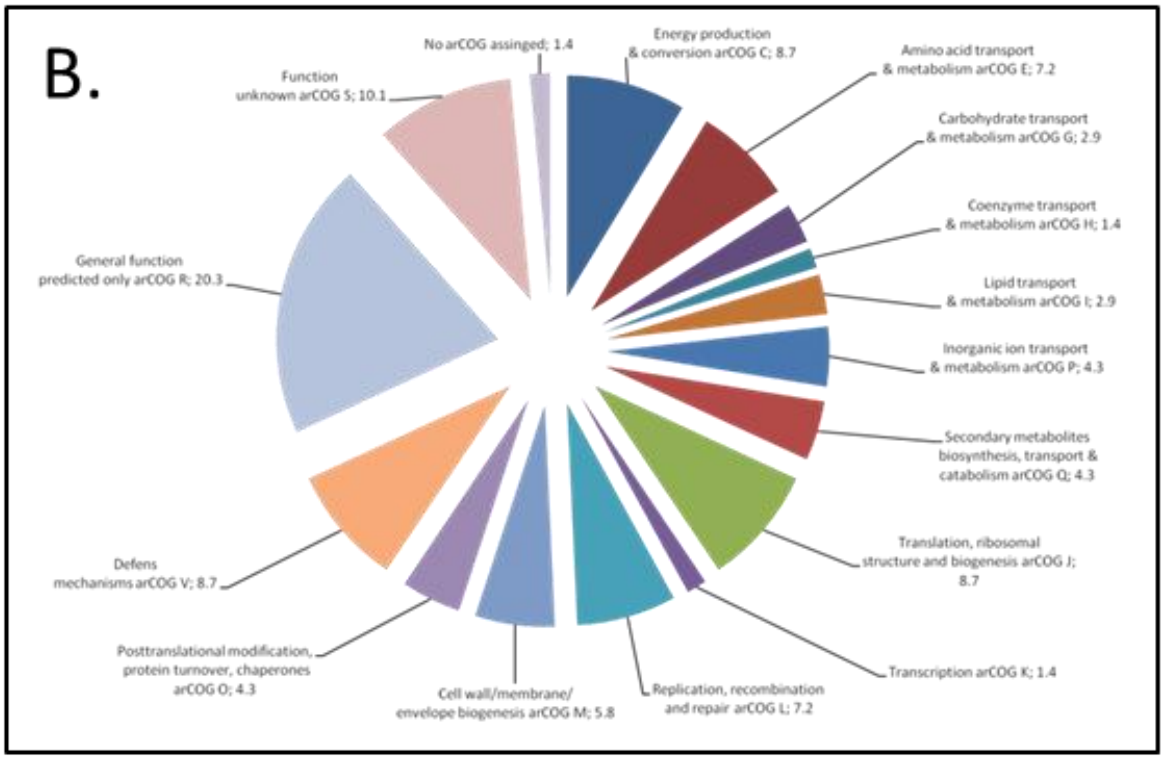


Figure 19 Comparison of p-proteome studies in Archaea, Bacteria and Eukarya.
S. acidocaldarius Total (MW001, *Δsaci_ptp* and *Δsaci_pp2a*) *S. solfataricus* P2 Total (glucose and tryptone) [188], *B. subtilis* [70], *E. coli* [71], *L. lactis* [72] and *H. salinarium* [46].

Interestingly, in all three *S. acidocaldarius* strains an unusual pS/pT/pY %-ratio was observed. The amount of pY is extremely high (33.0-46.1 %) in comparison with other p-proteome studies (0.6-10.3 %), with the exception of the p-proteome study from *S. solfataricus* (53.6 %). The pS/T/Y %-ratios of the strains MW001 and *Δsaci_pp2a* are very much alike. It seems that the deletion of the *saci_pp2a* had only an effect on the amount of p-proteins (2.1 times; 143 vs. 69) but not on the ratio of pS/pT/pY. Instead, the deletion of the *saci_ptp* increased the amount of p-proteins 5.9-times (69 vs. 407) and had also a tremendous influence on the pS/pT/pY %-ratio.


A.
No arCOG information arCOG X; 1.1
No arCOG assinged; 0.4
Energy production & conversion arCOG C; 9.7
Amino acid transport & metabolism arCOG E; 8.6
Function unknown arCOG S; 11.2
Nucleotide transport & metabolism arCOG F; 3.0
General function predicted only arCOG R; 13.0
Carbohydrate transport & metabolism arCOG G; 4.5
Defens mechanisms arCOG V; 3.5
Coenzyme transport & metabolism arCOG H; 4.6
Intracellular trafficking, secretion & vesicular transport arCOG U; 0.6
Posttranslational modification, protein turnover, chaperones arCOG O; 3.2
Lipid transport & metabolism arCOG I; 5.9
Signal transduction mechanism arCOG T; 0.6
Inorganic ion transport & metabolism arCOG P; 2.2
Cell motility arCOG N; 0.4
Secondary metabolites biosynthesis, transport & catabolism arCOG Q; 2.0
Cell wall/membrane/ envelope biogenesis arCOG M; 3.5
Cell cycle control/division, chromosome partitioning arCOG D; 0.7
Translation, ribosomal structure and biogenesis arCOG J; 9.9
Replication, recombination and repair arCOG L; 6.9
Transcription arCOG K; 4.6


B.
No arCOG assinged; 1.4
Energy production & conversion arCOG C; 8.7
Amino acid transport & metabolism arCOG E; 7.2
Function unknown arCOG S; 10.1
Carbohydrate transport & metabolism arCOG G; 2.9
Coenzyme transport & metabolism arCOG H; 1.4
General function predicted only arCOG R; 20.3
Lipid transport & metabolism arCOG I; 2.9
Inorganic ion transport & metabolism arCOG P; 4.3
Secondary metabolites biosynthesis, transport & catabolism arCOG Q; 4.3
Translation, ribosomal structure and biogenesis arCOG J; 8.7
Defens mechanisms arCOG V; 8.7
Transcription arCOG K; 1.4
Posttranslational modification, protein turnover, chaperones arCOG O; 4.3
Cell wall/membrane/ envelope biogenesis arCOG M; 5.8
Replication, recombination and repair arCOG L; 7.2

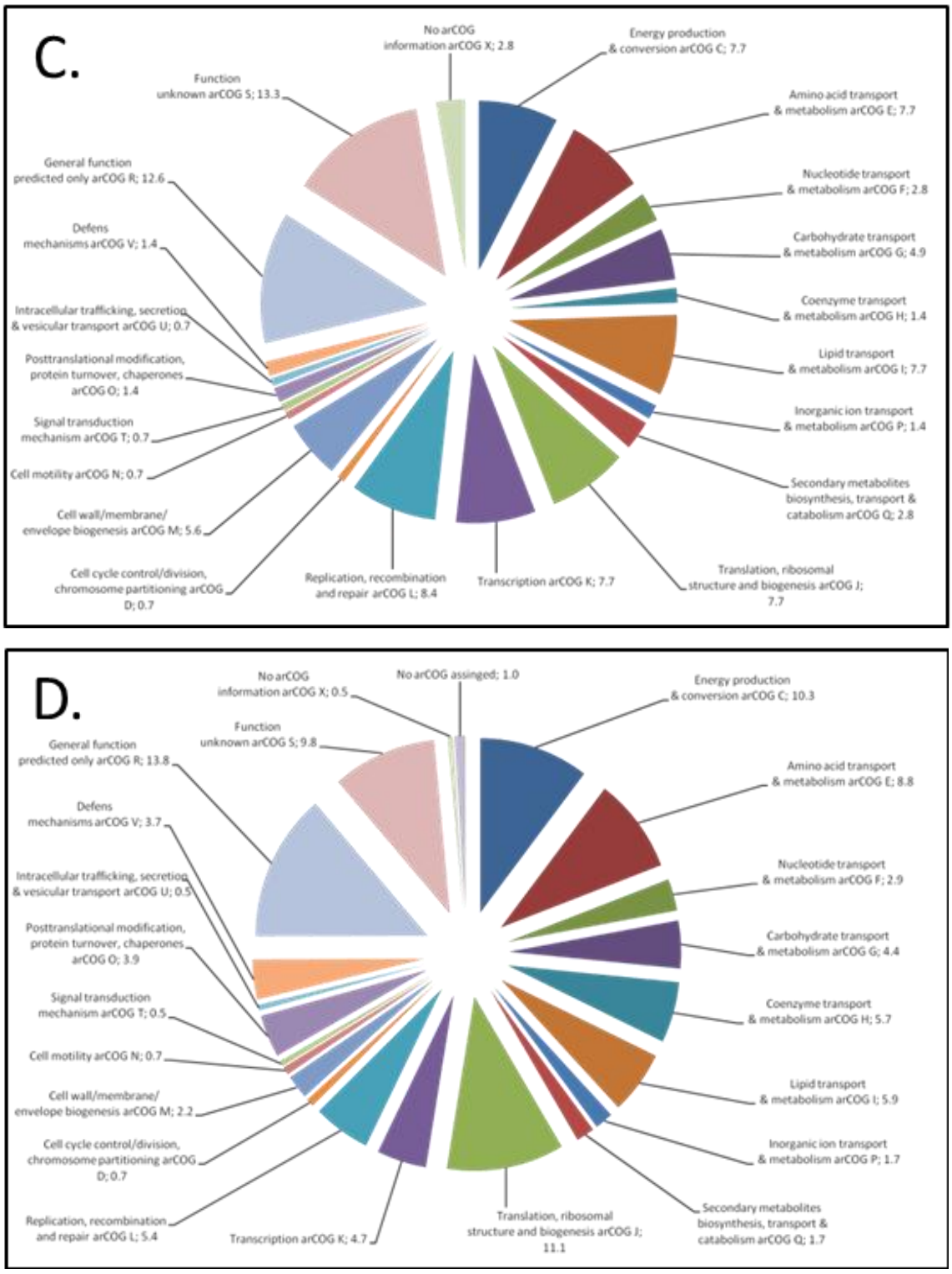


Figure 20 Distribution of unique p-proteins into their arCOG functional categories from *S. acidocaldarius*.
(A) Distribution of the unique p-proteins detected in all three strains; (B) distribution of the p-protein detected in MW001; (C) distribution of the p-protein detected in the Δ*saci_pp2a* strain; (D) distribution of the p-protein detected in the Δ*saci_ptp* strain.

The high amount of pY was reduced from 45.7-46.1 % (Ø 45.9 %) to 33.0 % (12.9 % reduced) and furthermore, the amount of pS and pT (was equally increased (6.1 % and 6.9 %, respectively)). The impact on the overall ratio, not only on pY, verifies the *in vitro* experiments of the Saci-PTP, which

revealed dual phosphatase activity. The identified p-proteins were categorized according to the arCOG functional code [194,219]. In general, all arCOG categories can be summarized into four major categories. The largest major category is comprised of cellular metabolism (arCOG C, E, F, G, H, I, P and Q), followed by cellular processes/signaling (arCOG D, M, N, T, O, U and V), information storage/processing (J, K and L) and the poorly characterized proteins (R, S and X). P-proteins were identified in 21 out of 26 arCOG functional categories in *S. acidocaldarius* (unique p-proteins) (Figure 20, Table S1 of [218]), indicating an important role of phosphorylation in almost all processes in the cell.

3.2.5. P-proteins of the central carbohydrate metabolism in MW001, *Δsaci_ptp* and *Δsaci_pp2a*

The investigated *S. acidocaldarius* strains were grown on 0.1% (v/w) NZ-amine and 0.2% sucrose. NZ-amine delivers various amino acids and is used for peptidolytic growth, whereas sucrose is a disaccharide which is composed of the monosacharides glucose and fructose and serves as substrate for glycolysis. Grogan reported in the past that *S. acidocaldarius* is capable to use sucrose and also the monosacharides D-glucose and D-fructose as sole carbon source [31].

However, growth was only observed after approximately one week adaptation and it was proposed that this long lag phase was caused by regulatory or mutational adaptation. Hence, *S. acidocaldarius* prefers L-amino acids as sole carbon source and is also capable to grow on mixtures of amino acid sources, like NZ-amine; and hexoses or disaccharides [31].

The reconstruction of the CCM in *S. acidocaldarius* revealed that several enzymes of the TCA cycle and the EMP pathway are targeted by reversible protein phosphorylation (Figure 21, Table 12). Most of them were identified in the *Δsaci_pp2a* strain (29), whereas only two were identified in the *Δsaci_ptp* strain and another two in MW001, respectively.

From the glycolytic branched ED pathway only the PyrK (Saci_1648) was identified as p-protein, whereas several gluconeogenic enzymes (EMP pathway, Saci_0671 FBPA/ase, Saci_1355 PGK, Saci_1356 GAPDH) were targeted by reversible protein phosphorylation. PGK and GAPDH are known to be reversible, however studies in *S. solfataricus* showed that both enzymes are used for gluconeogenesis only (EMP pathway) and are substituted in the glycolytic direction (ED pathway) by GAPN [119].

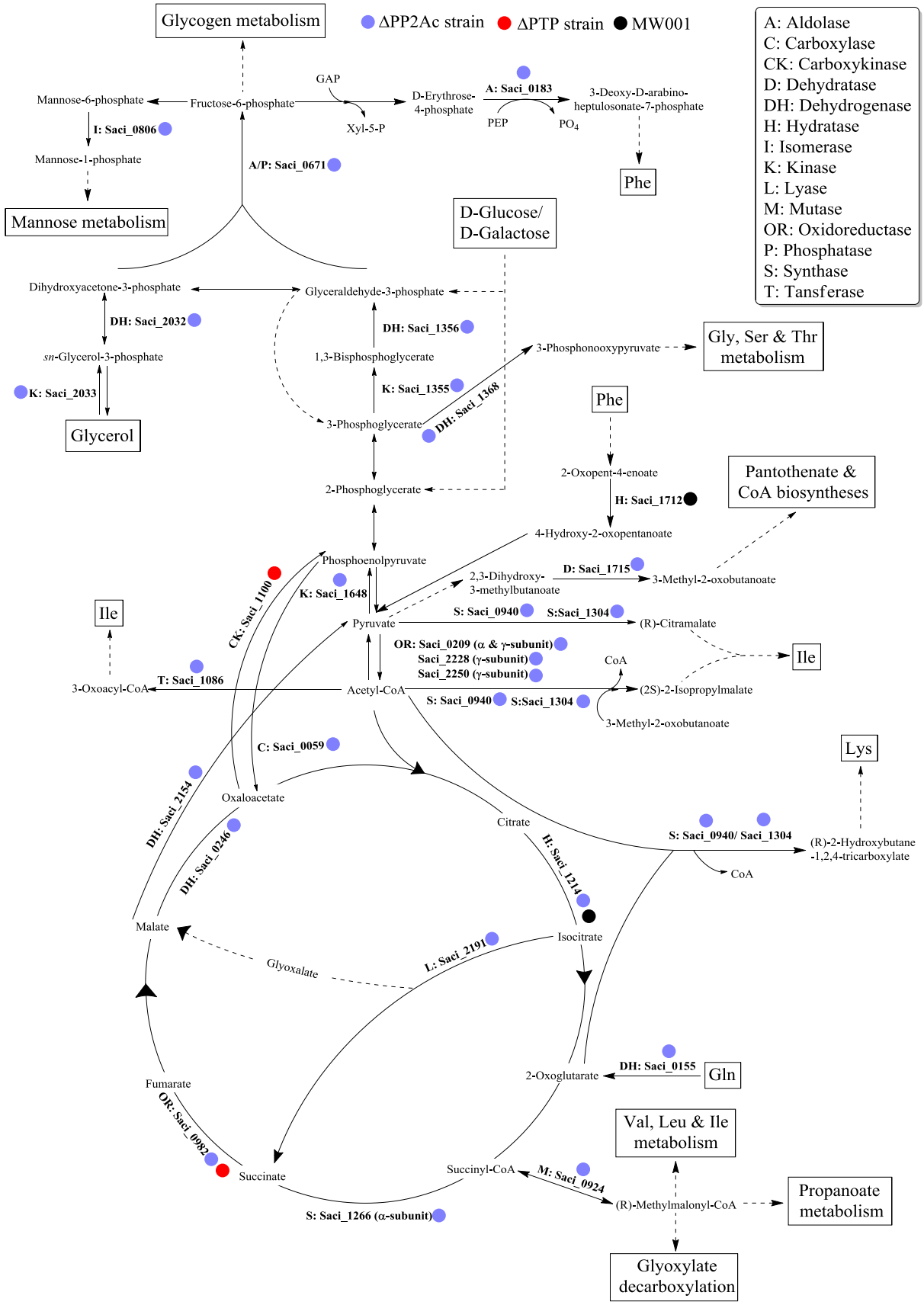


Figure 21 Reconstruction of the CCM including glycolysis via the branched ED pathway, gluconeogenesis, as well as TCA cycle branching points towards amino acids, glycerol, mannose and glycogen metabolism and C_4/C_3 interconversion in *S. acidocaldarius*.

Enzymes targeted by protein phosphorylation as well as the respective catalyzed reactions are depicted. Dashed lines indicate reactions, where the corresponding enzyme was not identified as p-proteins.

Table 12 CCM enzymes targeted by protein phosphorylation in *S. acidocaldarius*.

Gen ID	Strain	Propossed function	Pathway	P-peptide
Saci_0059	*Δsaci_pp2a*	PEP carboxylase	C_4/C_3-interconversion	-pYpTQGYQNK-
Saci_0155	*Δsaci_pp2a*	Glutamate dehydrogenase	Amino acid metabolism	-LpTQDELEQLpSRR-
Saci_0183	*Δsaci_pp2a*	3-Deoxy-D-arabino-heptulosonate -7-phosphate synthase	RuMP pathway	-KpYADMLQIGAR-
Saci_0209	*Δsaci_pp2a*	Pyuvate/2-oxoacid OR SU	Amino acid metabolism	-MNpTGIEApTNYILK-
Saci_0246	*Δsaci_pp2a*	Malate dehydrogenase	TCA cycle	-EVpYNVpYK-
Saci_0671	*Δsaci_pp2a*	FBPA/ase	EMP pathway	-GELDpTK-
Saci_0806	*Δsaci_pp2a*	Phosphomannomutase	Mannose transition point	-MGKLFGpTDGVR-
Saci_0924	*Δsaci_pp2a*	Methylmalonyl-CoA mutase	Amino acid & propanoate metabolism, glyoxylate decarboxylation	-MpTMDpTDDK-
Saci_0940	*Δsaci_pp2a*	Isopropylmalate/homocitrate/ citramalate synthase	Amino acid metabolism	-pYKLKMpTR-
Saci_0982	*Δsaci_ptp*	α-SU succinate dehydrogenase	TCA cycle	-DEpSRVR-
	Δsaci_pp2a			–AVVIAAGGMGMLpYR-
Saci_1100	*Δsaci_ptp*	Phosphoenolpyruvate carboxy kinase	C_4/C_3-interconversion	-VYVpSFGEEKDR-
Saci_1214	MW001	Aconitate hydratase	TCA cycle	-EISEpYIK-
	Δsaci_pp2a			-QpYGSGSpSR-
Saci_1266	*Δsaci_pp2a*	α -SU Succinyl-CoA synthase	TCA cycle	-KMIApYIAGMTAPREK-
Saci_1304	*Δsaci_pp2a*	Isopropylmalate/homocitrate/ citramalate synthase	Amino acid metabolism	-GVKSTLTpSLILKK-
Saci_1355	*Δsaci_pp2a*	PGK	Lower EMP shunt	-GLLpSLVPR-
Saci_1356	*Δsaci_pp2a*	GAPDH	Lower EMP shunt	-ApSLNIMSRDEpSIR-
Saci_1368	*Δsaci_pp2a*	3PGDH	Amino acid metabolism	-MKNNVIIVNpTSR-
Saci_1648	*Δsaci_pp2a*	PyrK	Lower EMP shunt	-AVATVIEGGpTITpSR-
Saci_1712	MW001	2-Keto-4-pentenoate hydratase	Amino acid metabolism	-KTIpSVKFV-
Saci_1715	*Δsaci_pp2a*	Dihydroxyacid dehydratase	Pantothenate & CoA metabolism	-EpTGKAVpSR-
Saci_2032	*Δsaci_pp2a*	Glycerol-3-phosphate dehydrogenase	Glycerol metabolism	-RDIGSGTpSGR-
Saci_2033	*Δsaci_pp2a*	Glycerol kinase	Glycerol metabolism	-GLIIGIpTR-
Saci_2154	*Δsaci_pp2a*	Malic enzyme	C_4/C_3-interconversion	-DPEKSFELTpSR-
Saci_2191	*Δsaci_pp2a*	Isocitrate lyase	TCA cycle	-DpTAFLpTGK-
Saci_2228	*Δsaci_pp2a*	Pyuvate/2-oxoacid OR SU	TCA cycle	-AFQpSAK-
Saci_2250	*Δsaci_pp2a*	Pyuvate/2-oxoacid OR SU	TCA cycle	-KIITVNApTK-

Furthermore, five enzymes of the TCA cycle (Saci_0246 MDH, Saci_0982 SDH α-SU, Saci_1214 aconitase hydratase, Saci_1266 succinyl-CoA synthase α-SU, Saci_2191 isocitrate lyase) and three enzymes involved in C_4/C_3-interconversion (Saci_0059 PEP carboxylase, Saci_1100 PEPCK, Saci_2154 malic enzyme) were found to be phosphorylated. In addition, several enzymes at branching points were found to be phosphorylated, eight enzymes of amino acid metabolism (Saci_0155 glutamate dehydrogenase, Saci_0209 pyruvate/2-oxoacid OR SU, Saci_0924 methylmalonyl-CoA mutase, Saci_0940 isopropylmalate/homocitrate/citramalate synthase, Saci_1304 isopropylmalate/

homocitrate/citramalate synthase, Saci_1712 2-keto-4-pentenoate hydratase, Saci_2228 pyruvate/2-oxoacid OR SU, Saci_2250 pyruvate/2-oxoacid OR SU), two enzymes of glycerol metabolism (Saci_2032 glycerol-3-phosphate dehydrogenase, Saci_2033 glycerol kinase), the phosphomannomutase (Saci_0806) from mannose metabolism and the dihydroxyacid dehydratase (Saci_1715) from pantothenate and CoA metabolism as well as the 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase (Saci_0183) from the RuMP pathway are targeted by protein phosphorylation.

3.2.6. Comparative whole-transcriptome profiling of *Δsaci_ptp* and *Δsaci_pp2a* by RNA-seq

To elucidate the influence of the protein phosphatase deletion the transcriptomes of the background strain MW001 and the strains *Δsaci_ptp* and *Δsaci_pp2a* were investigated via RNA deep sequencing. Thus, the transcriptome of the *Δsaci_ptp* and *Δsaci_pp2a* strains in comparison to the background strain MW001 were determined using an RNA-seq approach. Three cDNA libraries corresponding to each strain were constructed to perform Illumina deep sequencing. We obtained 1,835,618; 1,472,609; 1,560,420 qualified Illumina read pairs for MW001, *Δsaci_ptp* and *Δsaci_pp2a*, respectively. An average of 82.6 % of these reads were mapped to rRNA and tRNA sequences, whereas 14.5% were mapped to protein-coding sequences (CDS) and 2.8% to intergenic regions in the *S. acidocaldarius* genome.

A differential gene expression analysis was then performed for each mutant strain in comparison to the wild-type strain MW001. The transcript levels of 84 and 71 genes were significantly (considering a threshold of 2-fold changes) altered in *Δsaci_ptp* and *Δsaci_pp2a*, respectively (Figure 23 A and Figure 24,list of RNA sequences are given in Table S4 [218]).

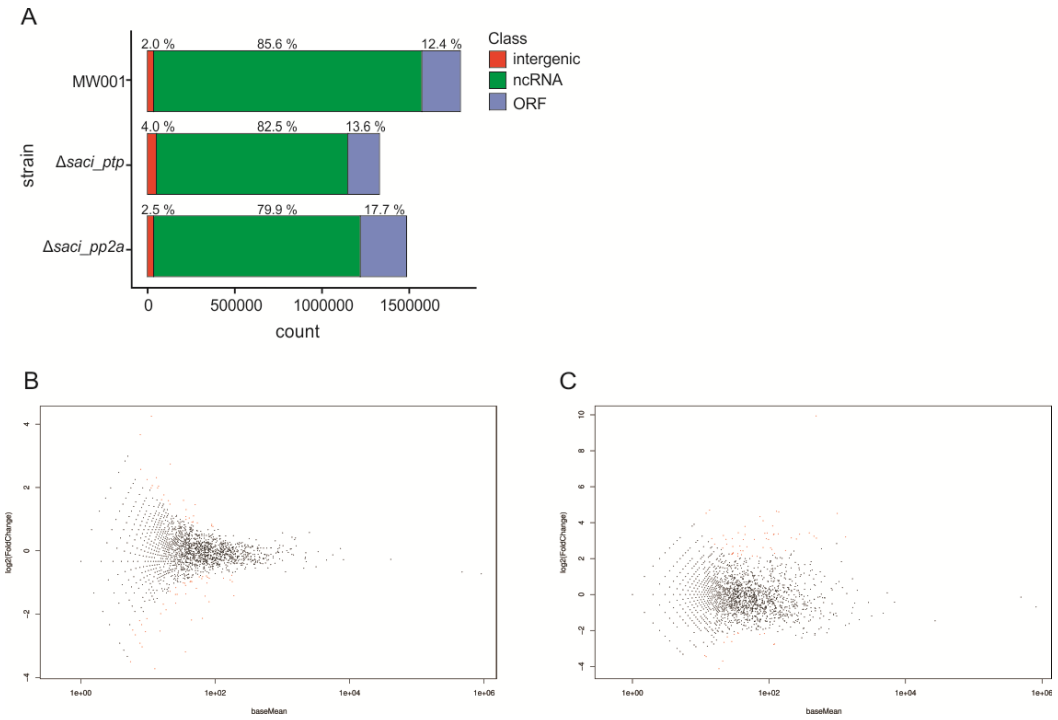


Figure 22 Correlation between RNA-seq reads and fold change between sample conditions.
(A) Summary of obtained RNA-seq reads, giving the number of reads and the genomic regions to which they were mapped, for samples from the wild type strain MW001, Δ*saci_ptp* and Δ*saci_pp2a. (B+C)* Plot of normalized mean expression intensity versus $\log_2$ fold change for wild type against Δ*saci_pp2a (B)* and wild type against Δ*saci_ptp (C).* Significantly altered genes (p-value < 0.055) are marked red.

While the majority of these changes corresponded to down-regulated genes in *Δsaci_ptp* (52 of 84 genes), the opposite was observed in *Δsaci_pp2a* (56 of 71 genes up-regulated) (Figure 23 A and Figure 24). In order to functionally categorize the regulated genes, we subjected them to the arCOG category analysis [194,219].

Additionally, we further determined that only eight ORFs were found to be commonly regulated in both deletion strains (Figure 23 B). Whereas five of the commonly regulated genes were found to be altered in the same manner (either up- or down-regulated), the other three ORFs were not (Figure 23 B and). These three ORFs, *saci_1858*, *saci_1859* and *saci_1860*, corresponded to genes encoding SoxNL complex components (i.e. complex III in the respiratory chain). They were either down-regulated in Δsaci_ptp or up-regulated in Δsaci_pp2a. Whereas *saci_1858* and *saci_1859* encode for mono-heme cytochrome b558-566 (CbsA and CbsB), *saci_1860* encodes for a Riske Fe-S protein (SoxL) [220]. Also the two down-stream genes of the soxNL complex operon, *saci_186*1 (soxN) and *saci_1862* (odsN), were found to be up-regulated in *Δsaci_pp2a* (Figure 24). Moreover, *Δsaci_pp2a* displayed augmented transcript levels of one component of the terminal oxidase (i.e. complex IV), the DoxBCE complex (*saci_0098* and *saci_0099*) (Figure 24). Four additional genes belonging to the energy metabolism and conversion arCOG category were determined as up-regulated in *Δsaci_pp2a*.

These regulated genes encode for a cytochrome/quinol oxidase (Saci_0097), a polyferredoxin (Saci_1802), a thioredoxin (Saci_1823) and the sulfocyanine SoxE-like protein (Saci_2096) (Figure 24).

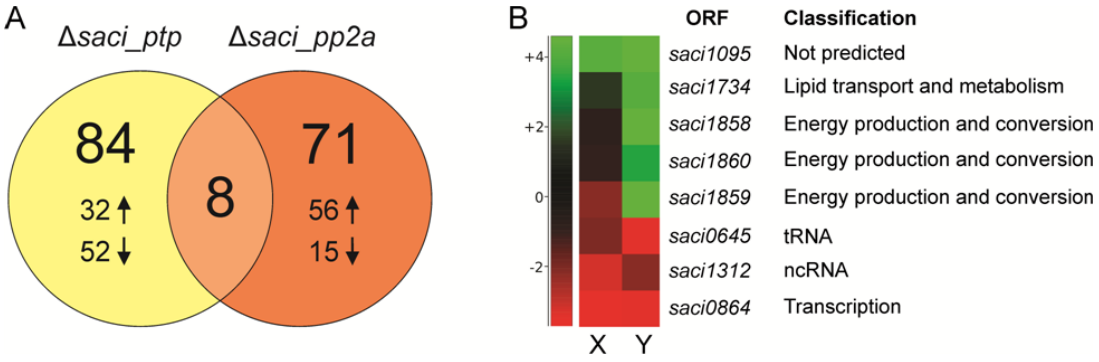


Figure 23 Transcriptionally regulated genes of Δ*saci_ptp* and Δ*saci_pp2a* in comparison to the background strain MW001.
(A) The number of regulated gene products in *Δsaci_ptp* and *Δsaci_pp2a* are depicted in a Venn diagram. Up and down arrows represent the number of genes which were found to be either up- or down-regulated for each deletion strain. In *Δsaci_ptp*, most genes were down-regulated, whereas in *Δsaci_pp2a* most were up regulated. Eight genes were regulated in both phosphatase deletion strains in common. (B) Heat-map representation of all eight commonly regulated genes between *Δsaci_ptp* (X) and *Δsaci_pp2a* (Y). This graph is based on the differential gene expression analysis by RNA-seq experiments, showing all genes which are significantly altered in their expression in both deletion strains. Right side panel shows the functional classification for each gene product.

The majority of the regulated transcripts encode for proteins belonging to the arCOG categories of unknown function (arCOG S), energy production and conversion (arCOG C) and transcription (arCOG K). In contrast to *Δsaci_pp2a*, *Δsaci_ptp* showed a general down-regulation pattern of gene expression, including the genes encoding for components of the alternative terminal SoxABCD oxidase complex [33]. Thus, it was further observed that *Δsaci_ptp* displayed diminished transcript levels of *saci_2088* (soxB) and *saci_2086* (soxD). Additionally, transcript levels of ORFs encoding for a thioredoxin-like (Saci_0359), a Rieske-like protein (Saci_1963) and an isocitrate-lyase (Saci_2191) were found to be down-regulated. In order to illustrate the transcriptomic changes on terminal oxidases, a proposed model of *S. acidocaldarius* branched respiratory electron transport chain and the transcriptional response of each phosphatase mutant on their respective components are depicted in Figure 25. Among the up-regulated genes in *Δsaci_ptp*, three predicted transcriptional regulators were the highest expressed genes (*saci_0065*, *saci_0455* and *saci_1787*). However, the regulatory function of any of these putative transcriptional regulators is still unknown as target genes have not yet been reported.

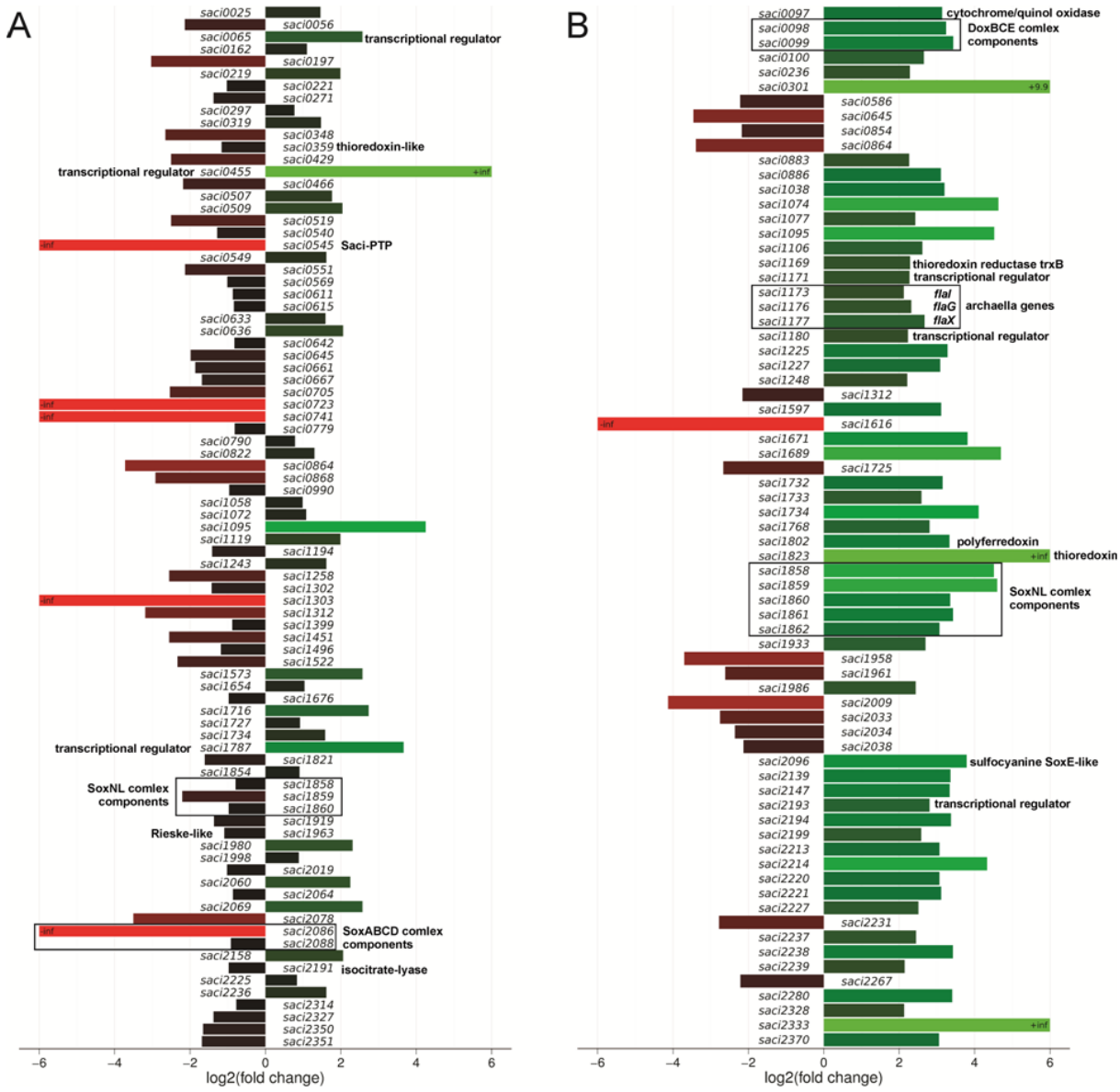


Figure 24 Bar plot representation of the transcriptomic changes of *Δsaci_ptp* and *Δsaci_pp2a*.
A differential gene expression analysis was performed for each deletion strain, Δ*saci_ptp* (A) and Δ*saci_pp2a* (B), in comparison to the wild-type strain MW001. All genes of whose expression was significantly altered ≥ 2-fold are shown. Red and green bars represent down- and up-regulated genes, respectively. Genes with a p-value below 0.054 were considered to be significantly regulated. The annotation of interesting genes is indicated beside.

On the other hand, *Δsaci_pp2a* showed that the transcript levels of three different transcriptional regulators (Saci_1171, Saci_1180 and Saci_2193) were increased when compared with the wild-type strain MW001. Interestingly, saci_1180 and saci_1171 loci are found directly flanking up- and down-stream the archaellum operon (formerly archaeal flagellum [221]). It has recently been demonstrated that Saci_1180 and Saci_1171 act positively on regulating archaella gene expression (Lassak and Albers, submitted). Intriguingly, three additional genes of the archaellum operon coding for FlaI, FlaG and FlaX components (*saci_1173*, *saci_1176* and *saci_1177*) were also found to be up-

regulated in *Δsaci_pp2a* cells. This finding implies a regulatory role of protein phosphorylation in archaella expression, as has previously been suggested by Reimann et al. [54].

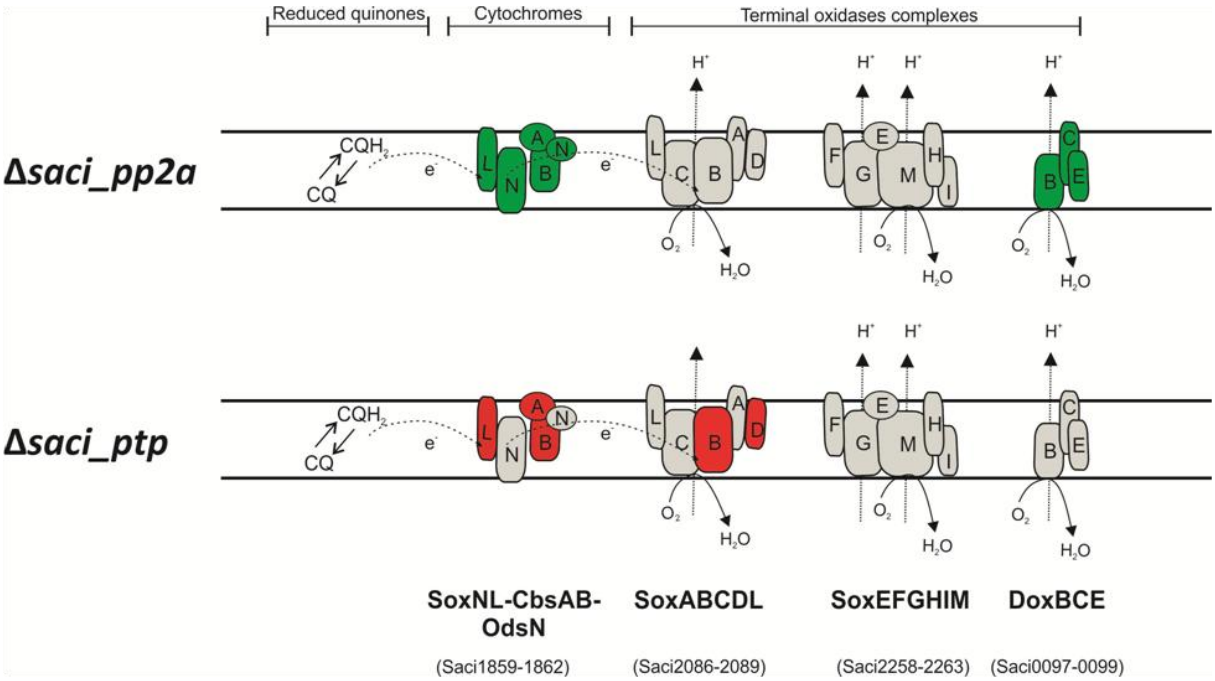


Figure 25 Proposed model of *S. acidocaldarius* branched aerobic respiration chain.
Respiratory electron transport chains model of *S. acidocaldarius* is based primarily on separate biochemical characterization studies. Electrons from NADH or succinate oxidation presumably enter ETC via NADH:quinone oxidoreductases or succinate:quinone oxidoreductases (not shown). Reduced quinones (CQH_2) then deliver electrons to cytochrome oxidase bc1 complex SoxLN-CbsAB-OdsN [220] and consecutively to one of the three terminal oxidase complexes. These terminal complexes are SoxABCD-SoxL [222-224], the bb3 terminal oxidase complex SoxEFGHIM [225-227], and the DoxBCE [228]. The components whose transcript levels were altered in each deletion strain are depicted either in red or green for down- and up-regulated, respectively. Gen IDs for each component are shown below.

3.2.7. Proteins of transcriptionally regulated clusters are also phosphorylated

The transcriptomic analysis revealed a regulatory effect on transcripts encoding for proteins of the aerobic respiration chain and the archaellum in the phosphatase deletion strains. Therefore and because motility has been studied in detail previously the p-proteome data were specifically screened for proteins involved in these cellular processes (Table 13). Intriguingly, six proteins of the arCOG C family of energy production and conversion are phosphorylated, including SoxC (Saci_2087, SoxABCDL complex) and SoxI (Saci_2258, SoxEFGHIM complex), which belong to terminal oxidase complexes (Figure 25). Furthermore, all of the three von Willebrand domain-containing proteins encoded in *S. acidocaldarius* (Saci_0977, Saci_1209 and Saci_1211) are phosphorylated at different residues. One of these proteins, ArnB (Saci_1211), was previously shown to be involved in negative archaellum regulation *in vivo* [54]. Also the membrane associated archaellum protein FlaJ (Saci_1172), involved in archaella assembly, revealed four different phosphorylation sites and was

identified in both phosphatase deletion mutants. Interestingly, the predicted positive transcriptional regulator ArnR1 (Saci_1171) in direct vicinity to the archaellum operon (Lassak and Albers, submitted) was also phosphorylated. Notably, seventeen additional predicted transcriptional regulators revealed phosphorylation sites, mostly in the phosphatase deletion mutants, indicating the importance of this posttranslational modification in transcriptional regulation in *S. acidocaldarius* and may be in Archaea in general. As all of the detected proteins must have been modified by protein kinases, it was interesting to see also five predicted serine/threonine kinases to be phosphorylated in the phosphatase deletion strains.

Table 13 P-proteins related aerobic repiration, motility, transcriptional regulation and phosphorylation identified in the *S. acidocaldarius* p-proteome.

Protein	arCOG category	Functional arCOG annotation	Strain
Saci_2087	C	Cytochrom b subunit of the bc complex (SoxC)	*Δsaci_ptp*
Saci_2250	C	Pyruvate:ferredoxin oxidoreductase or related 2-oxoacid:ferredoxin oxidoreductase, gamma subunit	*Δsaci_ptp*
Saci_2258	C	Predicted subunit of heme/copper-type cytochrome/quinol oxidase (SoxI)	*Δsaci_ptp*
Saci_2263	C	Heme/copper-type cytochrome/quinol oxidase (SoxI)	*Δsaci_pp2a*
Saci_2273	C	Fe-S oxidoreductase	*Δsaci_pp2a*
Saci_2316	C	Fe-S oxidoreductase	*Δsaci_pp2a* *Δsaci_ptp*
Saci_0977	R	Uncharacterized protein containing a vWA factor type A domain	MW001 *Δsaci_ptp*
Saci_1209	R	Uncharacterized protein containing a vWA factor type A domain	*Δsaci_ptp*
Saci_1211	R	Protein containing a vWa factor type A domain ArnB	*Δsaci_ptp*
Saci_1172	NU	Archaeal flagella assembly protein J	*Δsaci_pp2a* *Δsaci_ptp*
Saci_1171	K	Predicted transcriptional regulator	*Δsaci_ptp*
Saci_0446	K	Transcriptional regulator, contains HTH domain	*Δsaci_ptp*
Saci_0501	K	Predicted transcriptional regulator, PadR family	*Δsaci_pp2a*
Saci_0882	K	Predicted transcriptional regulator	*Δsaci_pp2a*
Saci_1068	K	Transcriptional regulator, MarR family	*Δsaci_ptp*
Saci_1161	K	Sugar-specific transcriptional regulator TrmB	*Δsaci_ptp*
Saci_1242	K	Transcriptional regulator, contains HTH domain	*Δsaci_pp2a*
Saci_1344	K	Predicted transcriptional regulator	MW001 *Δsaci_ptp*
Saci_1502	K	Predicted transcriptional regulator, contains C-terminal CBS domains	*Δsaci_ptp*
Saci_1523	K	Predicted transcriptional regulator, contains ATP-binding and HTH domains	*Δsaci_pp2a* *Δsaci_ptp*
Saci_1528	K	Predicted transcriptional regulator, contains ATP-binding and HTH domains	*Δsaci_pp2a*
Saci_1588	K	Transcriptional regulator (IcIR family)	*Δsaci_ptp*
Saci_1658	K	Transcriptional regulator (Lrp/AsnC family)	*Δsaci_ptp*
Saci_2116	K	Predicted transcriptional regulator, C-terminal HTH-like domain	*Δsaci_pp2a*

Saci_2352	K	Predicted transcriptional regulator, PadR family	*Δsaci_ptp*
Saci_0635	K	Transcriptional regulator, xre family	*Δsaci_ptp*
Saci_1505	K	Transcriptional regulator, xre family	*Δsaci_ptp*
Saci_2005	K	Predicted transcriptional regulator	*Δsaci_pp2a*
Saci_0965	TD	Serine/threonine protein kinase involved in cell cycle control	*Δsaci_pp2a*
Saci_1193	RTKL	Membrane associated serine/threonine protein kinase	*Δsaci_pp2a* *Δsaci_ptp*
Saci_1694	RTKL	Membrane associated serine/threonine protein kinase	*Δsaci_pp2a* *Δsaci_ptp*
Saci_1869	RTKL	Membrane associated serine/threonine protein kinase	*Δsaci_ptp*
Saci_1875	RTKL	Membrane associated serine/threonine protein kinase	*Δsaci_pp2a*

3.2.8. Influence of the phosphatases deletions on regulation of cell motility and aerobic respiration

Previous studies in *S. acidocaldarius* showed an important impact of reversible protein phosphorylation on cell motility [54]. The main motility device of *S. acidocaldarius* is the archaellum. The conserved operon comprises seven genes, encoding the archaellin subunit FlaB, the membrane protein FlaJ, the ATPase FlaI, and the accessory proteins FlaX, FlaG, FlaF and FlaH [221]. To confirm the regulatory impact of the phosphatase deletions on cell motility as well as aerobic respiration, which was indicated in the RNA-seq and p-proteome data, the expression of all archaellum operon genes and selected respiratory chain genes was first determined in both *Δsaci_pp2a* and *Δsaci_ptp* compared to MW001 by quantitative RT-PCR. The cells were treated the same way as for the RNA-seq experiments and harvested at OD_{600} 0.6. Whereas no significant changes in expression levels could be observed in the *Δsaci_ptp* deletion strain, all seven genes of the archaellum operon were found to be highly up-regulated in the *Δsaci_pp2a* cells (Figure 26), which confirmed the result of the transcriptomic analysis (Figure 24).

Also the analysis of genes encoding for SoxNL complex components and the SoxABCDL terminal oxidase confirmed induced transcription levels in *Δsaci_pp2a* (Figure 26). Next, the up-regulation of the archaella genes in *Δsaci_pp2a*, which was observed on transcriptional level (Figure 24 and Figure 26), was further investigated at protein level. To that end, the background strain and both phosphatase deletion strains were first cultivated in Brock medium rich in nutrients and then transferred to basic Brock medium without supplemented nutrients for five hours. This treatment has been shown before to highly induce archaella biosynthesis. Samples of all cultures before and after stress induction were analyzed by immunoblotting with specific antibodies against the archaellin subunit FlaB and a second archaellum protein FlaX. Both proteins were induced by nutrient limitation stress in the wild type strain MW001 and to the same extent in *Δsaci_ptp*. Intriguingly,

Δsaci_pp2a showed already a signal for FlaB before induction and highly increased levels of FlaB and FlaX after nutrient limitation if compared to MW001 and *Δsaci_ptp* (Figure 27 A).

Finally, the swimming behavior of the wild type strain and both phosphatase deletion mutants were investigated using a motility assay on semi-solid gelrite plates. Therefore, MW001, *Δsaci_ptp*, *Δsaci_pp2a*, a hypermotile positive control (*ΔaapF*) and a non-motile negative control (*ΔaapFΔflaH*) were spotted on semi-solid gelrite-plates containing low amounts of nutrients [229,230]. After five days of incubation at 76 °C, MW001 and *Δsaci_ptp* showed only slight motility. In contrast, the motility of *Δsaci_pp2a* was comparable to the hypermotile positive control *ΔaapF*, which is lacking the aap pili (Figure 27 B).

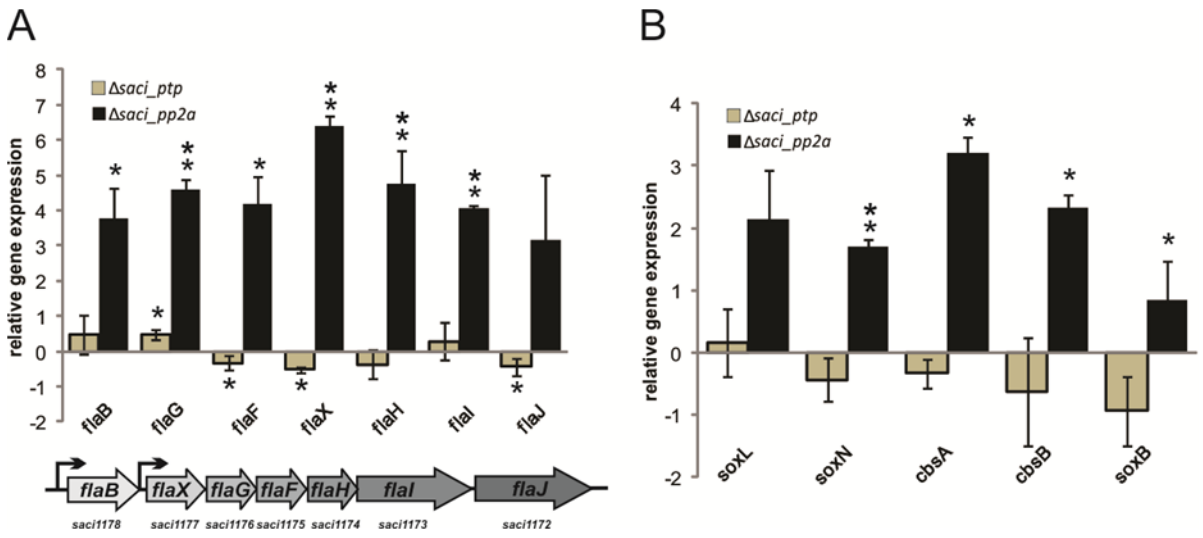


Figure 26 Differential expression levels of archaella operon genes and respiratory chain genes in *Δsaci_pp2a* and *Δsaci_ptp*.

Total RNA isolated from *S. acidocaldarius* MW001, *Δsaci_pp2a* and *Δsaci_ptp* cultures were used for cDNAs synthesis. qRT-PCR analysis was performed using specific primers for each archaella component (A) and components of the terminal oxidase complexes (B). Relative transcript expression levels of each gene were normalized to an internal control gene secY. The values reflect the fold change in expression compared to cDNA prepared from MW001 cells, which is designated as baseline. The means and standard deviations of 3 biological replicates are shown. Up-regulation of all archaella and the terminal oxidase genes in the *saci_pp2a* deletion strain reflect the RNA-seq results. * significant (p-value ≤ 0.05), ** highly significant (p-value ≤ 0.01).

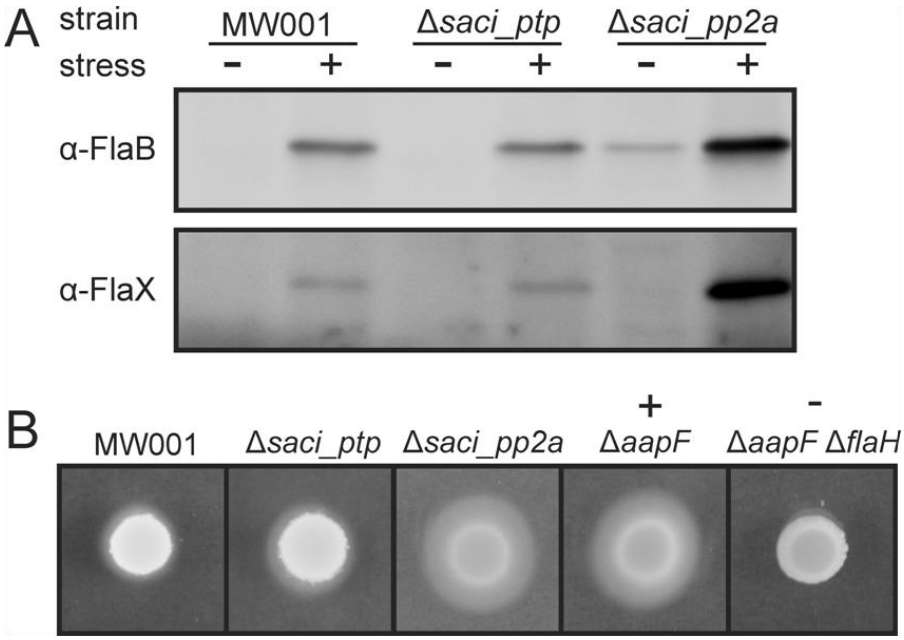


Figure 27 Impact of phosphatase deletion on archaella expression and motility.
(A) For the nutrient limitation assay MW001, Δ*saci_ptp* and Δ*saci_pp2a* were grown till OD_{600} of about 0.4 in Brock medium supplemented with 0.1% (w/v) NZ-amine, 0.2% (w/v) sucrose and 10 µg/ml Uracil and then transferred to medium lacking the nutrient source. After 6 hours the stress samples were taken and immunoblotting was performed with specific antibodies against the archaellum proteins FlaB and FlaX. The amount of protein of Δ*saci_ptp* is comparable to the background strain MW001, whereas for Δ*saci_pp2a* highly elevated expression of FlaB and FlaX was observed after starvation stress. Interestingly, already under normal growth conditions the FlaB level was increased in the Saci-PP2A phosphatase deletion mutant. (B) The wild type strain MW001, the two phosphatase deletion strains Δ*saci_ptp* and Δ*saci_pp2a*, the hypermotile positive control strain Δ*aapF* and the negative non-motile strain Δ*aapF*Δ*flaH* were spotted on semi-solid gelrite plates and incubated for five days at 76°C. A small halo is visible for MW001 and Δ*saci_ptp*, whereas Δ*saci_pp2a* was as hyper motile as the positive control strain.

3.2.9. Discussion

In this study, we performed a detailed enzyme characterization of the two identified protein phosphatases in *S. acidocaldarius* Saci-PTP and Saci-PP2A. Further on we demonstrate that Saci-PTP and Saci-PP2A are acting as phosphatases with activity to specific p-proteins in *S. acidocaldarius*. Therefore both phosphatases seem to act on different signal transduction pathways in the cell. Whereas the deletion of Saci_PTP did not result in a significantly changed growth phenotype, the deletion of Saci-PP2A led to a response on mRNA and proteome level that was similar to conditions when *S. acidocaldarius* experiences severe nutrient limitations. High up-regulation of all archaella components and the respiratory chain enable the organisam to evade these unfavorable conditions by assembling archaella and providing the energy for archaella driven motility.

The two identified genes of *S. acidocaldarius* encoding Saci-PTP and Saci-PP2A were heterologously expressed in *E. coli*, purified and enzyme activity was tested with pNPP as well as phopho-peptides. In contrast to the SsoPTP of *S. solfataricus* [231], Saci-PTP displayed a dual-specific phosphatase

activity using p-Tyr peptides as well as p-Thr peptides as substrate. However, the used substrates to determine the enzyme activity of SsoPTP and Saci_PTP were very different, which might explain the observed difference of both PTPs regarding their substrate preference to some extent. In case of SsoPTP either p-Ser/p-Thr were used as substrates or the small p-peptides ApYR/NKpYGN [47]. Unfortunately, detailed information regarding the assay was not reported, only the absence of phosphate release in presence of pSer/pThr was mentioned. Saci-PP2A, resemble the typical catalytic subunit of eukaryotic PP2A type phosphatases, was specific for Ser/Thr dephosphorylation. The enzyme required divalent metal ions for activity and could be strongly inactivated by ocadaic acid. The dual activity of the Saci-PTP and thereby possible overlap in substrates with the Saci-PP2A might explain that the ratios of phosphorylated Ser/Thr/Tyr peptides did not differ in the p-proteome of the *Δsaci_pp2a* deletion mutant in comparison to the background strain cells. On the other hand it was unexpected that the ratio of p-Tyr peptides dropped so dramatically in the *Δsaci_ptp* deletion mutant. A possible explanation is that due to the deletion of the Saci-PTP a regulatory circuit was intercepted that is important for the expression of the Tyr specific kinases. However, as so far none of the *S. acidocaldarius* kinases has been shown to phosphorylate Tyr future studies have to be awaited.

In general the p-proteome data showed that phosphorylated proteins were found from pathways involved in all cellular processes without a detectable preference, highlighting the essential role of regulatory protein phosphorylation in *S. acidocaldarius*.

The CCM from *S. acidocaldarius* was one of the main processes in the cell which was found to be highly targeted by protein phosphorylation (Figure 21). Comparison pf the phosphorylation pattern of the CCM detected in *S. solfataricus* P2 (for details see section 3.1. and Figure 10) and *S. acidocaldarius* revealed that the TCA cycle and the C_4/C_3-interconversion were targeted in both Sulfolobales species. However, an accurate comparison between the phosphorylation impact on the CCM in both Archaea is rather difficult. The p-proteome of *S. solfataricus* was analyzed in respect to different C-sources either grown on tryptone or on D-glucose [188] whereas the three *S. acidocaldarius* strains were grown on a combination of 0.1% (v/w) NZ-amine and 0.2% (v/w) sucrose. Furthermore, deletion of PPs will require specific cell response. This is also visible when comparing the CCM of both Sulfolobales (Figure 10 and Figure 21). For *S. solfataricus* a high regulation of the ED pathway and enzymes involved of the ED like pathways is assumed. Especially the sugar dehydrogenases, sugar acid dehydratases and the ALDH participating in those pathways were targeted. Instead, no enzyme of the ED pathway in *S. acidocaldarius* was detected as p-protein whereas the EMP pathway was highly targeted in *S. acidocaldarius* and several branching points into other metabolic pathways as well. The TCA cycle alone is only slight targeted in *S. solfataricus* (two reactions) compared to *S acidocaldarius* were four reactions of the TCA cycle and also the isocitrate

lyase are targeted. In both cases it was only one subunit of the succinate dehydrogenase enzyme complex (Saci_0982, SSO2357) targeted. However, they are not homolog to each other and represent different subunits of the succinate dehydrogenase enzyme complex. The only part of the CCM which is similar targeted in both species is the C_4/C_3-interconversion. In both Archaea the malic enzyme (Saci_2154, SSO2869, e-value 0.0, 87% identity) and the PEP carboxy kinase (Saci_1100, SSO2537, e-value 0.0, 75% identity) are phosphorylated. The phosphorylation sites however are not conserved in the enzymes of both Sulfolobales.

In the transcriptomics data from the *Δsaci_pp2a* strain most genes were down regulated, whereas three regulators (*saci_0065*, *saci_0455* and *saci_1787*) were highly induced, but since for none of these the regulatory targets are known, it is difficult to predict pathways that might be affected.

However, the Hiseq data for *Δsaci_pp2a* strain clearly indicated that in response to the deletion of Saci-PP2A, genes of the archaellum assembly operon and respiration chain components were highly upregulated. The upregulation of the genes involved in archaellum assembly found in the HiSeq data were confirmed by quantitative RT-PCR. Additional experiments at protein level confirmed the elevated expression levels of the archaellum proteins FlaB and FlaX and finally *in vivo* motility assays demonstrated hyper motility of *Δsaci_pp2a* compared to MW001 and *Δsaci_ptp*. Interestingly, the two up- and downstream located genes (HTH-containing transcriptional regulators Saci_1171 and Saci_1180, ArnR1 and ArnR, respectively) of the archaellum operon) were also up-regulated in *Δsaci_pp2a* and in addition ArnR1 (Saci_1171) was found to be phosphorylated in the *Δsaci_ptp* p-proteome analysis. In Archaea, current knowledge concerning signaling pathways involved in archaella regulation is limited. The activation of flaB expression by these two regulators (ArnB, ArnA) was already experimentally confirmed (Lassak and Albers, submitted). Furthermore, we recently reported that two interacting proteins, ArnA (Saci_1210) and ArnB (Saci_1211), negatively regulate archaella expression in *S. acidocaldarius* and both proteins were phosphorylated by the eukaryotic-like kinase Saci_1193 (ePK) and dephosphorylated by the phosphatase Saci-PP2A *in vitro* (for a model see Figure 28) [54].

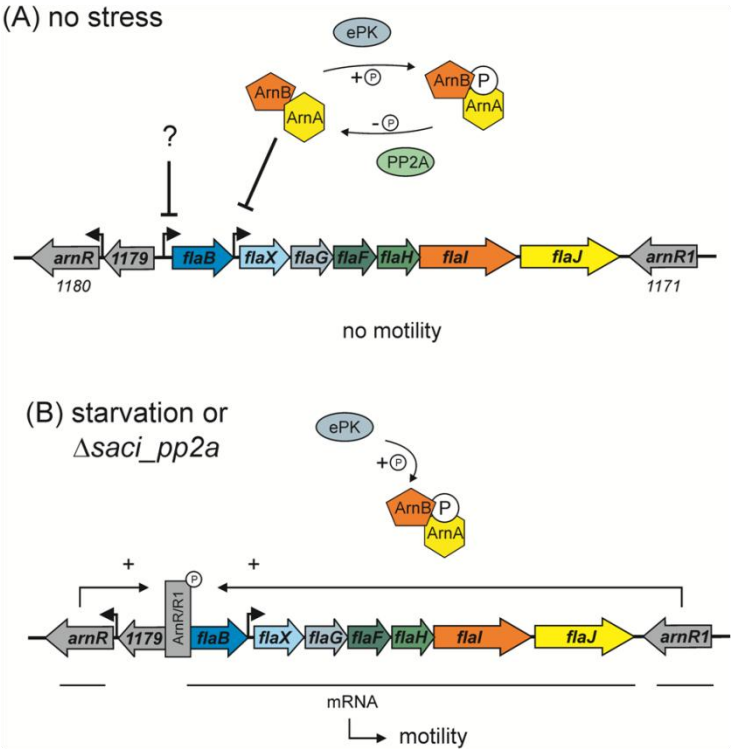


Figure 28 Model of the involvement of phosphorylation in the regulation of the archaellum in *S. acidocaldarius*.
During exponential growth (A) archaellum synthesis from the archaellum operon *flaB-flaJ* (*saci_1178-1172*) is repressed. Part of the repression network are ArnA (Saci_1210) and ArnB (Saci_1211) that are phosphorylated and dephospohrylated by the ePK (Saci_1193) and PP2A, respectively. ArnA and B closely interact and repress the *flaX* promoter [35]. Probably another, yet unknown, factor is repressing the *flaB* promoter. When S. *acidocaldarius* cells enter the stationary growth phase and undergo starvation or as seen in the *Δsaci_pp2a* deletion mutant (B), the ArnA/B complex exist only in its phosphorylated form and repression is released. Moreover, the expression of ArnR (Saci_1180) and ArnR1 (Saci_1171) is induced and these factors then drive high expression of FlaB which leads to the assembly of archaella on the cell surface and therefore to increased motility of the cells (Lassak et al, submitted).

Intriguingly, this kinase and ArnB revealed *in vivo* phosphorylation in the kinase domain and the C-terminal region, respectively, in the current p-proteome data set (see Table 13). In addition, the ePK gene (*saci_1193*) displays strong cell-cycle-specific induction of transcription [232], with >16 fold upregulation around the cell division stage, just prior to a 2 - 4 fold upregulation of the archaellum encoding genes (*saci_1174 – saci_1178*), further linking these functions together. Whereas in *S. tokodaii* ArnA was binding to the FlaX promoter [233], in *S. acidocaldarius* no clear impact on transcriptional level of the archaella components FlaB and FlaX was observed in the arnA deletion strain, hinting that regulation occurs via protein-protein interactions as a result of phosphorylation [54]. Therefore the here observed C-terminal phosphorylation sites of ArnB could be the docking sites for ArnA leading to the strong interaction of both proteins [54]. Taken together, the data suggest that in *saci_pp2a* a situation of nutrient limited growth conditions is mimicked, at least regarding to archaella expression, where the negative regulators (ArnA/B) are inactive due to permanent phosphorylation and the positive regulators (ArnR/R1) can enhance archaella expression. In line with this assumption is the fact that a thioredoxin (Saci_1823), a thioredoxin reductase

(Saci_1169) and a peroxiredoxin (Saci_2227) encoding genes were found to be up-regulated in the *Δsaci_pp2a* RNA-seq data. These proteins are well known in all three domains of life to act during oxidative stress conditions and to be involved in disulfide bond shuffling of proteins. Although the thioredoxin enzyme system is not yet fully understood in Sulfolobales, some of these proteins were investigated in more detail in respect to their function and structure [234-239]. Interestingly, the thioredoxin reductase (Saci_1169) is genomically in direct vicinity to the archaella genes and one of the positive regulators (ArnR1), and was also phosphorylated in the p-proteome analysis. Hence, nutrient limitation could not only lead to archaella expression, but also to oxidative stress based on excess of oxygen, which in turn would require proteins like peroxiredoxin and the thioredoxin system to avoid cellular damage. Additionally, since the exact mechanisms of the regulators located up- and down-stream of the archaellum operon are not yet known, regulatory mechanisms via thiol-disulfide bonds are another possibility. In summary, the ArnA/ArnB and ArnR/R1 regulatory system of archaella expression is not yet fully understood, but it constitutes the first complex regulatory network involving protein phosphorylation which is investigated in the Crenarchaeota.

Surface structure driven motility is a very energy-consuming process and therefore assembly of these structures has to be tightly controlled which has been studied for the bacterial flagellum in great detail. Interestingly, many genes involved in energy conversion like the genes encoding proteins of the SoxNL, SoxABCDL and DoxBCE respiratory chain complexes were differently regulated in the phosphatase deletion strains compared to the background strain as depicted in Figure 25. Expression and assembly of surface appendages like the bacterial flagella or the archaella are energy-consuming processes. Recently, an association of the bacterial flagella switch-motor complex to the NADH-ubiquinone oxidoreductase and the F0F1 ATP synthase was demonstrated in *E. coli*, and predicted to function as a local energy supply for flagella movement [240]. Although regulation of the archaellum is based on ATP hydrolysis and not on proton motive force (as for the bacterial flagellum) [221], energy must be provided for archaella assembly and function. Therefore the up-regulation of the archaella genes might go along with the up-regulation of genes involved in the respiratory chain in *S. acidocaldarius*. A possible interaction of these energy generating complexes and the archaella basal body has not been investigated so far, but cannot be ruled out. Furthermore, it is important to note that phosphorylation of respiratory chain components was reported in different p-proteomic studies including Pseudomonas spp. [241], *Saccharomyces cerevisiae* mitochondria [242], human muscle mitochondria [243] and *S. solfataricus* [188]. In line with these reports, two subunits of the terminal oxidase complexes (SoxC and SoxI) and other proteins involved in energy conservation were also found to be phosphorylated. However, a possible connection to the up-regulation of respiratory chain components in the phosphatase lacking strain *Δsaci_pp2a* still needs to be investigated.

In conclusion, this study provides first insights into the *in vitro* and *in vivo* function of the two *S. acidocaldarius* protein phosphatases, Saci-PTP and Saci-PP2A. The RNA-sequencing approach and the p-proteome study of both phosphatase deletion mutants have shed light into the multitude of cellular processes reversible protein phosphorylation is involved in. In particular, it could be demonstrated *in vivo* that Saci_PP2A is involved in the very complex signal transduction pathway that regulates the expression of the *S. acidocaldarius* archaellum.

3.3. Atypical protein kinases of the RIO family in Archaea

The results of this study were published in Biochemical Society Transactions (PMID: 23356318).

3.3.1. Introduction

The final aim in systems biology is to build a silicon cell, a precise replicate of the living cell that is able to predict *in vivo* function. Besides regulation at the transcription and translation level PTMs play a major role in cellular regulation, since they allow for a fast response to environmental changes. The probably most important PTM is reversible protein phosphorylation, which has a crucial role in signal transduction. This modification takes usually place at His or Asp (two component systems, most common in Bacteria) or Ser, Thr, Tyr residues (most common in Eukaryotes). Protein phosphorylation is mediated by PKs, which catalyze the phosphorylation of respective target proteins and protein phosphatases, which specifically remove the respective phosphate group. PKs form a large super family comprising six families [78] and apart from the well investigated ePKs [79] so called atypical PKs (aPKs) of the ABC1, RIO, piD261, AQ578 and PKN2 family were proposed. Although the presence of aPKs is well established their physiological function remains largely unknown.

RIO (right open reading frame) kinases have gained a lot of attention during the last decade. The RIO families were originally identified in Archaea and Eukaryotes and a common ancestor with an ancestral RIO gene has been predicted [78]. Until now, four different types of RIO kinases were identified: Rio 1, Rio 2, Rio 3 and Rio B kinase [85]. Phylogenetic analysis confirmed the distribution of RIO kinases in all three domains of Life and proposed the presence of a combination of one Rio 1 kinase and one Rio 2 kinase in less complex species (i.e. prokaryotes and single cellular eukaryotes) whereas multicellular Eukaryotes including Humans possess in addition one Rio 3 kinases [86].

For yeast-Rio 1 an essential function in cell cycle progression and chromosome maintenance and furthermore, for both yeast RIO kinases in 18 S rRNA processing was demonstrated. Deletion of either of the two RIO kinases results in cells death, assuming that both of them possesses different functions [87-92]. However, whereas the function in yeast is well established the role of their prokaryotic counterparts remains obscure. Although PK activity could be verified, their physiological target proteins are still unknown. A major breakthrough was achieved based on the crystallization of the Rio 1 and Rio 2 kinase from *A. fulgidus* [81,83] and combined bioinformatics analysis [85]. For both archaeal RIO kinases autophosphorylation and activity with common kinase substrates (Rio 1 kinase: histone H1 and myeline basic protein; Rio 2 kinase: histone H1, myeline basic protein and α-casein) was shown [82,83].

In general RIO kinases are regarded as a trimmed version of canonical ePKs but lack the subdomains VIII (active loop), X and XI (involved in substrate binding) [83]. In addition RIO domains possesses an insertion of 18-23 amino acids between αC and β3 (flexible loop), which is absent in ePKs. In general Rio 1 and Rio 2 kinases are very similar regarding their overall fold despite that the N-terminal domain of the Rio 2 family comprises a wing helix domain, which is absent in all other RIO families. As outlined previously members of the different RIO families can be distinguished by their specific P-loop (interaction and orientation of the ATP triphosphate moiety) and DFG-loop (metal-ion binding and positioning) sequence [81,85].

In this article we focus on the phylogenetic distribution of RIO kinase subfamilies within the third domain of Life -the Archaea- and by using comparative genomics (i.e. BLAST analysis, gene neighborhood), interaction predictions (i.e STRING,[181]) as well as the available p-proteome of *Sulfolobus solfataricus* P2 [188] we aim to gain first insights into their possible physiological function in Archaea.

3.3.1. Distribution of RIO kinases in Archaea

The distribution of RIO kinases within the Archaea was analyzed using AF1804 and AF2426 from *A. fulgidus* DSM 4304 as template [81-83]. All sequences with an e-value smaller than $1e^{-10}$ were considered as putative RIO kinases and for classification within the RIO subfamily the Rio 1 (STGKEA), Rio 2 (GxGKES), Rio 3 (STGKES) or Rio B (SGKEA) P-loop sequences were used [85].

In the available 121 archaeal genomes (http://img.jgi.doe.gov/cgi-bin/w/main.cgi version 3.5, August 2012) 224 RIO kinases were identified confirming the assumption that RIO kinases are conserved in all three domains of Life [89]. According to their P-loop sequences 72 of the 224 RIO kinases were classified as Rio 1 kinases, 122 as Rio 2 kinases and 30 as Rio B kinases (see supplement for analysis Table SI 1 [244]). 55 of the 72 Rio 1 kinases harbor the canonical GxxSTGKEANVY/F P-loop sequence or the short variant STGKEA, whereas the remaining 17 Rio 1 kinases showed variations in their P-loop sequence (for details see Table SI 1 [244]). Among 122 Rio 2 kinases 81 possess the Rio 2 kinase P-loop sequence GxGKES [81,83,85] and in 41 Rio 2 kinases an alternative Rio 2 P-loop was identified (i.e. GVGKEG, for details see Table SI 1 [244]). All 30 members of the Rio B kinase subfamily contain the conserved Rio B kinase P-loop sequence SGKEA [85] and no alterations of this sequence were observed.

In most Archaea (100 of 121) two RIO kinases were present. Only in 19 Archaea one RIO kinase was identified and in *Aciduliprofundum boonei* T469 three Rio 1 kinases and in *Methanopyrus kandleri* AV19 four RIO kinases (two Rio 1 and two Rio 2) were identified. Combinations of Rio 1 and Rio 2 kinases were identified in 58 Archaea (most members of the Desulfurococcales, Archaeoglobales, Halobacteriales, Methanomicrobia and Thermococcales) and of Rio B and Rio 2 in 30 Archaea (all

members of the Sulfolobales except *S. acidocaldarius* (Rio 1 and Rio 2) and most members of the Methanococcales). Two copies of Rio 2 were found in 12 Archaea (most members of the Thermoproteales). Only one copy of Rio 1 was detected in 10 Archaea (all members of the Methanobacteriaceae) and one copy of Rio 2 in 9 Archaea (all members of Thermoplasmatales, Nanoarchaeum, Cenarchaeum, Nitrosopumilaceae) (Table SI 1 [244]). Interestingly no Archaeon with a combination of Rio 1 and Rio B was identified. Furthermore, no Rio 3 kinases could be identified in Archaea, which is in accordance with their unique presence in multicellular Eukaryotes [86].

3.3.2. Gene neighborhood of RIO kinases

To gain a better understanding of the physiological function of the RIO kinase family in Archaea we compared the gene neighborhood of all identified archaeal RIO kinases (up to 10 genes up/down stream). Remarkably, for most RIO kinases a conserved gene neighborhood was observed (164 of 224 RIO kinases (73.2%)) mainly with genes encoding proteins involved in biological processes related to transcription and translation (Figure 29, Table SI 1 [244], see below for detailed discussion).

The majority of the Rio 1 kinases genes (43 of 68, 59.7%) as well as Rio B kinases genes (29 of 30) are found in a conserved neighborhood with genes encoding the human heterogeneous nuclear RNP K homology (KH) domain protein and the archaeal translation initiation factor 1A (aeIF-1A). 14 (19.4%) Rio 1 kinase genes cluster with the gene encoding the KH domain protein alone and four (5.6%) with genes encoding the KH domain protein and the DEAD/DEAH box protein (5, 6.9%) or with the domain of unknown function protein (DUF) 460 gene alone (2, 2.8%). The remaining Rio B kinase gene of *Methanococcus voltae* (Mvol_0151) is located in the gene neighborhood of the aeIF-1A and the DEAD/DEAH-box protein gene. The majority of the Rio 2 kinase genes present as additional copies (46 of 122, 37.7%) are located in a conserved gene neighborhood with a gene encoding for the DUF 460 protein (PFAM) (31, 24.5%) or a combination of genes encoding DUF460 and small nuclear ribonucleoproteins (snRNP) (13, 10.7%, restricted to *Sulfolobus* species) or cobyrinic acid a,c-diamide synthase CbiA (2, 1.6%, restricted to Metalosphaera species). The remaining 18 Rio 2 kinase genes were either located in the gene neighborhood of the gene encoding the KH domain protein alone (9) or in combination with the archaeal translation initiation factor aeIF-1A encoding gene (8) and the last Rio 2 kinase of *Vulcanisaeta distributa* (Vdis_2273) was located upstream from aeIF-1A.

Remarkably, the Archaea that show a similar gene organization of Rio 2 as found for Rio 1 and Rio B are with few exceptions the ones that harbor two Rio 2 genes/paralogs or only one copy of Rio 2. Therefore, suggesting that the combination of RIO kinases with aeIF-1A with the KH domain protein encodes an essential function that is present in all Archaea. In none Archaeon with two Rio kinases,

except *A. fulgidus*, a similar gene neighborhood was observed and it is tempting to speculate that both atypical protein RIO kinases might have acquired different functions.

As outlined above the majority of proteins encoded in the RIO gene cluster has an predicted function in processes related to transcription and translation. KH domain proteins are present in all three domains of Life. In general, the function of KH domain proteins is to recognize RNA or ssDNA and KH motifs are found in various proteins, which are involved in a vast number of biological processes (e.g. transcriptional regulation, translation control) [245-247]. Although the knowledge about archaeal translation is rather scarce [248] for the archaeal translation initiation factor aeIF-1a from *Sulfolobus solfataricus* P2 binding to the small ribosomal subunit and furthermore stimulation of the binding of Met-tRNAi (initiator tRNA) and mRNA to the ribosome was demonstrated [249] [248]. DEAD/DEAH box helicases are involved for example in ribosome biogenesis, translation initiation and mRNA decay [250,251]. Their general function is to separate short RNA duplexes or RNA-DNA heteroduplexes and they are further more important for RNA structure modification [252]. For the cold shock induced DEAD-box helicase CsdA from *E. coli* an involvement in translation initiation, biogenesis of the 50S ribosomal subunit or/and degradasome formation with RNase E was demonstrated [253-255]. Intriguingly, studies of archaeal DEAD-box helicases revealed also a significant induction by cold shock and cell growth at low temperature [256,257]. DUF460 proteins are only present in Archaea and no homologs were found in Bacteria or Eukaryotes (PFAM 26.0), however, their function remains to be elucidated [258]. snRNPs are distributed in all three domains of Life [259]. Although archaeal snRNPs are structurally more closely related to their eukaryotic as to their bacterial homologs (Hfq proteins) in terms of function they resemble Hfq proteins. Whereas the eukaryal snRNPs are involved in spliceosome formation, Hfq proteins are involved in RNA processing and are of interest for RNA-RNA interaction. An interaction of DEAD box proteins with snRNPs has been reported in *M. jannaschii* [252,260]. Only the cbiA gene in Metallosphaera species, which encodes for the cobyrinic acid a,c-diamide synthases involved in cobalamin biosynthesis [261] shows no obvious link to transcription and translation.

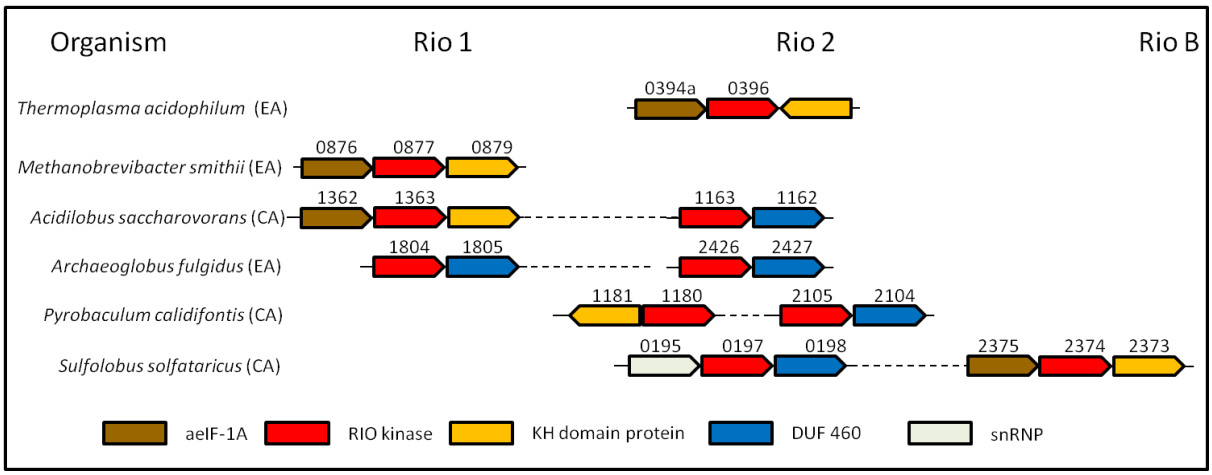


Figure 29 Conserved gene neighborhood of archaeal RIO kinases.
The characteristic organization of members of the Rio 1, Rio 2 and Rio B subfamily in most Archaea is depicted. The two studied enzymes of *A. fulgidus* are an exception regarding their gene organization. Genes and their orientation are indicated as arrows, the color code is given below.

3.3.3. RIO gene cluster in Sulfolobaceae

In all seven sequenced Sulfolobaceae except *Acidianus hospitalis* and Metallosphaera cuprina a larger conserved RIO kinase gene region was observed (Figure 30). The cluster comprises nine genes and for all encoded proteins (except acetyl-CoA c-acetyltransferase, acaB-4) a possible function in RNA or DNA processing, modification and degradation can be predicted. Ribonuclease HII (SSO2384) exhibits ribonucleotide specific endonuclease activity and degrades the RNA strand of DNA/RNA heteroduplexes. A function in DNA replication, transcription, repair and development in all three domains of Life is predicted [262], and some archaeal enzymes have been investigated (e.g. *T. kodokaraensis*) [263].

For archaeal nonsense-mediated decay 3 (NMD3) domain proteins (SSO2383) a function in biogenesis of the large ribosomal subunit has been predicted [264]. For DUF 424 domain proteins (SSO2382) no functional information is available. The role of universal translation initiation factor 2 (subunit B; aeIF2: alpha, beta (aeIF-2B, SSO2381), and gamma subunit) is well established in *Sulfolobus* and Archaea in general and a crucial role in binding of charged initiator tRNA to the small ribosomal subunit was demonstrated [265]. SSO2380 encodes a UPF004, eukaryotic elongation subunit 3 (Elp3) and TRM2 and MiaB (TRAM) domain protein. TRAM domains are common in Archaea and a function in thiolation of tRNA and ribosomal proteins has been predicted [266]. UPF004 and Elp3 domains are commonly present in methyltransferases [267,268] and an N-terminal UPF004 domain and C-terminal TRAM domain was identified in methylthiotransferases (MtaB) [267,268]. Besides the RIO kinase gene (SSO2375) as discussed above also the genes encoding aeIF-1A and KH domain protein are present in the gene cluster. Therefore except the acaB-4 genes all genes encode

proteins, which seem to be tightly connected to biological processes related to transcription and translation regulation, which is also assumed for RIO kinases.

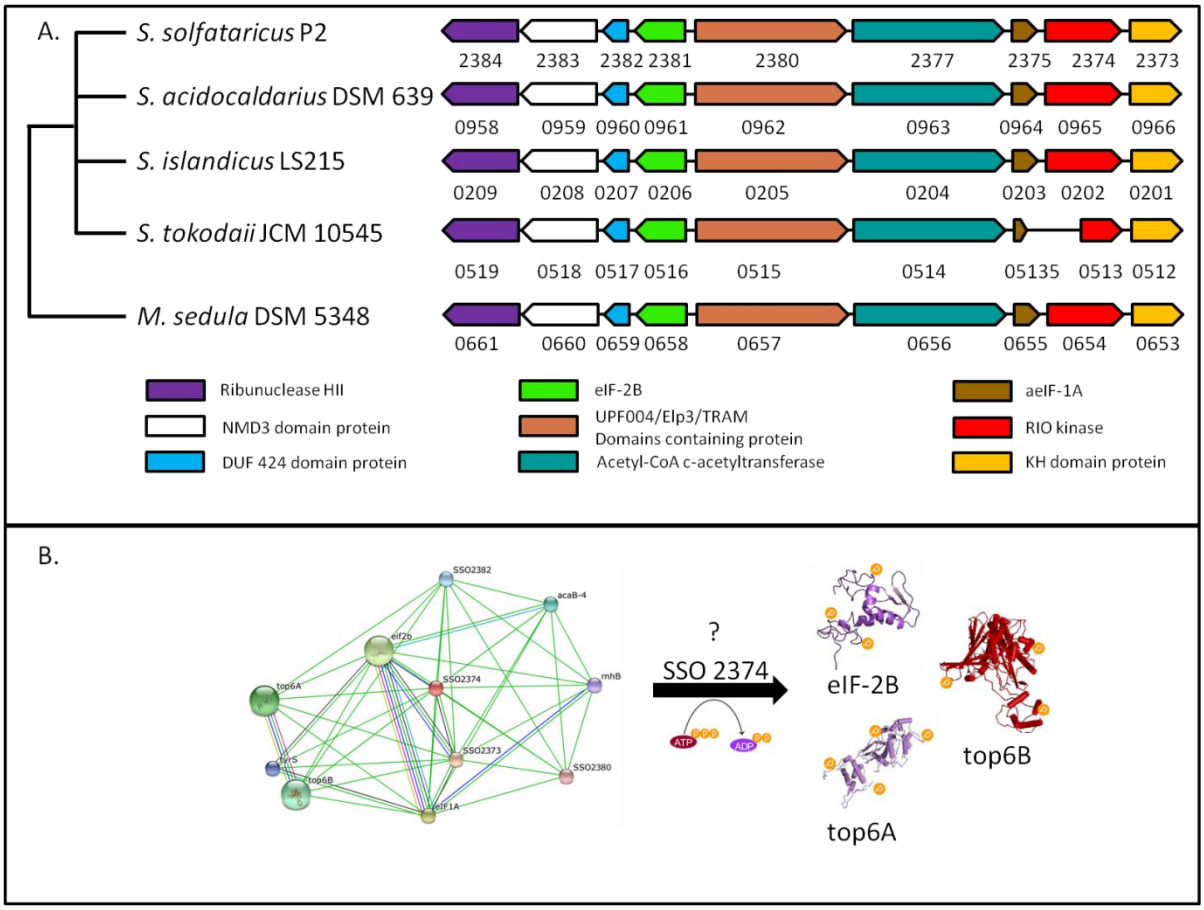


Figure 30 RIO kinases in Sulfolobaceae.
The conserved RIO kinase gene cluster in all *Sulfolobus* species are depicted (A). Genes and their orientation are indicated as arrows, the colour code is given below. Predicted protein-protein interaction map (STRING, [181]) for RIO B kinase of *S. solfataricus* P2 (SSO2374) and p-proteins of interacting candidates aeIF-2B (SSO2381), Top6A (SSO0907) and Top6B (SSO0968) in the p-proteome analysis of *S. solfataricus* P2 [188].

3.3.4. First evidence for the function of RIO kinases in *S. solfataricus* P2

S. solfataricus P2 is probably the best studied Archaeon in terms of protein phosphorylation. Two protein phosphatases as well as several protein kinases were characterized and the p-proteome in response to changes of carbon source (glucose/tryptone) was determined [47,50-53,100,188]. Besides the conserved gene organization of Rio B kinase (SSO0197) and Rio 2 kinase (SSO2374) we used “predicted protein-protein interactions” (STRING, [181]) as well as available p-proteome data in order to identify possible target proteins for RIO kinase phosphorylation.

For SSO Rio B kinase STRING analysis predicted protein-protein interaction with the translation initiation factor aeIF-1A (SSO2375) [269], the KH domain containing protein (SSO2373), the translation initiation factor (aeIF-2B) (SSO2381) and the ribunuclease HII (SSO2384), which reside in

the same gene cluster as well as for several other proteins involved in translation or transcription (e.g. DNA topoisomerase VI type IIB 6a and 6b, tyrosyl-tRNA synthetase and DNA topoisomerase III). Intriguingly, in accordance with this observation aeIF-2B was identified to be phosphorylated on tyrosine, serine and also threonine (-ApYVECpSpTCK-) in glucose grown cells [188]. In addition both DNA topoisomerase 6 subunits (Top6A (SSO0907), Top6B (SSO0968)) were detected to be multiple phosphorylated (Top6A: -DIVpTpYMLpSpSE-, glucose and -MpTADMEpSK-, tryptone; Top6B: -KpYEDFR- glucose and -RLpYpTFK-, tryptone).

For Rio 2 kinase (SSO0195) STRING analysis (STRING 9.0) [181] predicts protein-protein interaction with the DUF460 domain containing protein (SSO0198), which is located in the conserved gene cluster. SSO0198 was also detected as p-protein in the p-proteome study of *S. solfataricus* P2 in glucose grown cells [188]. Also the predicted interacting transcription regulator SSO0942 was detected to be phosphorylated in tryptone grown cells (-DpYNKpSMK-). Furthermore, besides several hypothetical proteins and NAD synthetase (nadE) and the putative TATA binding (TBP) interacting protein (TIP49-like) (SSO2450) an interaction between both SSO RIO kinases is predicted as well as between the SSO Rio 2 kinase and two ePKs (SSO3182 and SSO3207) suggesting a more complex cellular regulatory network.

3.3.5. Conclusion

Atypical protein kinases of the RIO kinase family are found in all three domains of Life underlining their essential function. In all archaeal sequenced genomes at least one (Rio1 or Rio 2 kinase) but typically two members of the RIO kinase family (Rio 1 and Rio 2 kinase, Rio B and Rio 2 kinase, in few cases two times Rio 2 kinase) were identified. In all Archaea a conserved clustering of one RIO kinase gene (usually of the Rio 1 or Rio B kinase gene or of one of the two Rio 2 kinase genes) with genes encoding the KH domain protein and often the translation initiation factor aeIF-1A is observed. The second Rio 2 kinase gene typically clusters with a gene encoding the DUF460 protein. In some Archaea an additional association of DEAD/DEAH helicases, snRNP or CbiA is observed. An even larger conserved gene cluster was identified in most members of the Sulfolobaceae comprising for example in addition ribonuclease H II, aeIF-2B and a predicted methythiotransferase for thiolation of tRNA or ribosomal proteins. Although the function of RIO kinases in Archaea remains to be resolved, the presented bioinformatics analysis including comparison of the phylogenetic distribution, gene neighborhood analysis as well as protein-protein interaction prediction provides some first hints for a possible function. Further on the combination with the available p-proteome study of *S. solfataricus* [188] allows for the identification of possible target proteins for RIO kinase phosphorylation in

S. solfataricus. Our studies suggest a regulatory function of archaeal RIO kinases in biological processes related to transcription and translation resembling the assumed physiological role in yeast. The first *in vivo* evidence for the important role of RIO kinases comes from a study of the haloarchaeon *H. volcanii* [84]. The α1 protein of the proteasome 20S core particle was identified as a target for *in vitro* phosphorylation of Rio1 kinase (Rio 1p, HVO 0135). Site-directed mutagenesis of the p-sites of the α1 protein and expression in *H. volcanii* displayed a link between *in vivo* phosphorylation of the proteasome and cell viability, as well as pigmentation [84].

3.4. Unraveling the function of paralogs of the aldehyde dehydrogenase super family from Sulfolobus solfataricus

The initial kinetic parameters of SSADH-I and SSADH-II with D,L-GAP, D,L-GA, glycolaldehyde and SSA were determined during the Master Thesis of Dominik Esser. In the PhD thesis all three aldehyde dehydrogenases (SSADH-I, SSADH-II and MSDH) were purified and characterized in detail. Dr. Francois Talfournier (CNRS-Université de Lorraine, Vandoeuvre-lès-Nancy, France) kindly synthesized the thermo-instable substrates methylmalonate semialdehyde and malonate semialdehyde and furthermore performed the initial enzyme activity measurements of the *S. solfataricus* MSDH (SSO1218). Dr. Jolanta Polkowska (AG Schrader, University of Duisburg-Essen, Essen, Germany) synthesized the aldehyde 2,5-dioxopentanoate, which was used for the remaining experiments to determine the kinetic parameters of SSADH-I and SSADH-II with this aldehyde. The physiological interpretation and pathway reconstruction of the gained data was performed by Dominik Esser under the guidance of Prof. Dr. Bettina Siebers. Dr. Christopher Bräsen (AG Siebers, University of Duisburg-Essen, Essen Germany) performed the construction of phylogenetic trees to investigate the evolutionary relationship of members of the aldehyde dehydrogenase superfamily. The results of this study were published in Extremophiles (PMID: 23296511).

3.4.1. Introduction

Aldehyde dehydrogenases (ALDHs) play an essential role in intermediary metabolism as well as cellular detoxification by converting aldehydes to the corresponding activated or non-activated acids in all three domains of life. In addition an important cellular role in osmotic protection and cellular differentiation is discussed [146]. Two distinct structural families have been characterized, which catalyze the oxidation of aldehydes to the respective acids via a two-step mechanism. Both structural families share the initial acylation step including the nucleophilic attack of the catalytic cysteine residue on the aldehyde group, hydride transfer and the formation of the thioacylenzyme intermediate as well as NAD(P)H). However, they differ in the deacylation step. In phosphorylating ALDHs like glyceraldehyde-3-phosphate dehydrogenase (GAPDH; type 2, TIGR01546) inorganic phosphate serves as acyl acceptor, whereas in non-phosphorylating ALDHs the attack of the thioacylenzyme intermediate can either be performed by a water or CoA molecule leading to the formation of the respective non-activated or CoA-activated acids [147].

The non-phosphorylating ALDH superfamily comprises a group of distantly related enzymes which metabolize a wide variety of endogenous and exogenous aldehydes [148,149]. The superfamily

consists of 20 different subfamilies [150] catalyzing the oxidation of aliphatic and/or aromatic aldehydes to the respective acids and therefore have an important role in cellular detoxification and defense systems [151,152]. They are present in all sequenced genomes known today and previous bioinformatic studies revealed that the number of genes encoding ALDHs in Archaea is significantly reduced (1-5 gens) compared to Bacteria (1-26 gens) and Eukaryotes (8-17 gens) [150].

S. solfataricus P2 is a thermoacidophilic Crenarchaeon which belongs to the order of Sulfolobales. *S. solfataricus* thrives at temperatures from 70°C to 90°C (optimal 80°C) at pH 2-4 (optimal 3.2) [182]. The genome sequence is available [270] and most importantly the organism is genetically tractable [29]. *S. solfataricus* grows on a large variety of different sugars (e.g. cellulose, starch, glucose, arabinose and fructose as well as complex peptidolytic compounds [31]), and can easily be grown in the laboratory under defined conditions.

S. solfataricus uses an unusual branched ED pathway for glycolysis, comprising a semi-phosphorylative (sp) and a non-phosphorylative (np) branch [28,57,109-111] and the EMP pathway for gluconeogenesis [26,271,272]. The ED pathway was shown to be promiscuous for glucose and galactose degradation [109,110]. Also the pathways for D-arabinose [27], D-xylose and L-arabinose degradation in *S. solfataricus* were reported recently [25].

In the genome of *S. solfataricus* five genes encoding paralogs of the ALDH superfamily were identified (SSO1218; *gapN*-1 (SSO1629), *gapN*-2 (SSO1842), *gapN*-3 (SSO3194), *aldhT* (SSO3117)) [150] and so far only two paralogs SSO3194 and SSO3117 have been characterized [27,116]. For SSO3117 enzymatic characterization revealed broad substrate specificity and significant activity with DOP, glycolaldehyde and GA has been reported. A combined transcriptomic-proteomic approach demonstrated a significant up-regulation in *S. solfataricus* cells grown on D-arabinose and therefore a role as DOPDH in D-arabinose degradation has been proposed [27]. For SSO3194 non-phosphorylating GAPN activity was demonstrated [116]. GAPN catalyzes the unidirectional oxidation of GAP forming 3-PG and is, due to its allosteric properties (i.e. activation by G1P), a key enzyme of the spED branch in *S. solfataricus* (lower common shunt of the EMP pathway) and (hyper)thermophilic Archaea in general [116]. In order to unravel the function of the unknown ALDH paralogs in *S. solfataricus* the three encoding genes (SSO1629, SSO1842 annotated as GAPN isoenzymes, SSO1218 annotated as MSDH), were cloned, the enzymes were recombinantly overexpressed in *E. coli* and purified.

The recombinant enzymes were characterized as MSDH (SSO1218) as well as two succinic semialdehyde dehydrogenase isoenzymes (SSADH, SSO1629, SSO1842), which differ in their enzymatic properties. In addition the physiological role of the ALDHs in *S. solfataricus* is addressed by pathway reconstruction.

3.4.2. Heterologous expression of SSO1218, SSO1629 and SSO1842 in *E. coli* and purification

The ALDHs SSO1218 (MSDH), SSO1629 (SSADH-I) and SSO1842 (SSADH-II) were expressed in *E. coli* BL21 (DE3) RIL via IPTG induction and purified to apparent homogeneity using heat precipitation (80°C, 20 min), ion exchange chromatography and gel filtration (Figure 31). From 8 g of recombinant cells (wet weight) 10.4 mg SSADH-I, 4.2 mg SSADH-II and 5.9 mg MSDH were gained, respectively. The subunit size of all three ALDHs was around 50 kDa as determined via SDS-PAGE, which agrees well with their calculated molecular subunit weights (i.e. 54.6 kDa (SSO1218), 51.5 kDa (SSO1629), 51.6 kDa (SSO1842)). Furthermore by using gel filtration a native molecular weight of 130 kDa was determined for MSDH, 157 kDa for SSADH-I and 114 kDa for SSADH-II, suggesting a homodimeric, homotrimeric and homodimeric structure, respectively.

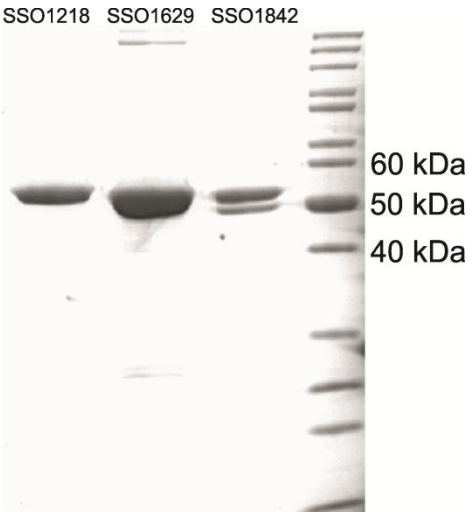


Figure 31Purification of the recombinant aldehyde dehydrogenases from *S. solfataricus*.
The SDS-PAGE (12.5%) was loaded with 2 µg of SSO1218 (MSDH), SSO1629 (SSADH-I) and SSO1842 (SSADH-II) after purification via heat precipitation, ion exchange chromatography and size exclusion chromatography and stained with Coomassie Blue.

3.4.3. Enzyme characterisation of the succinic semialdehyde dehydrogenases

SSO1629 (SSADH-I) showed hydrolytic ALDH activity with clear preference for succinic semialdehyde (SSA, k_{cat}/K_m 842 $mM^{-1}s^{-1}$), reduced efficiency with DOP (k_{cat}/K_m 21.3 $mM^{-1}s^{-1}$) and only minor activities with D,L-GAP (k_{cat}/K_m 0.5 $mM^{-1}s^{-1}$), D,L-GA (k_{cat}/K_m 0.3 $mM^{-1}s^{-1}$) and glycolaldehyde (k_{cat}/K_m 0.1 $mM^{-1}s^{-1}$). Formaldehyde, betaine aldehyde and benzaldehyde could not serve as substrates. SSADH-I is only capable to use NAD^+ (k_{cat}/K_m 7.6 $mM^{-1}s^{-1}$) as cosubstrate, in presence of $NADP^+$ no enzyme activity was detected (Table 14).

Table 14 Kinetic parameters of SSADH-I (SSO1629) for various aldehydes with NAD^+ as cosubstrate.
Enzyme activity in presence of NAD^+ was determined in presence of SSA (0.2 mM) at 70°C.

Substrate	$V_{max(app)}$	$K_{m(app)}$	$k_{cat(app)}$	$k_{cat}/K_{m(app)}$
	[U/mg protein]	[mM]	[s^{-1}]	[$mM^{-1}s^{-1}$]
Succinic semialdehyde (C_4) (1)	26.3	0.03	22.6	842.0
2,5-Dioxopentanoate (C_5) (2)	8.3	0.3	7.1	21.4
D,L-Glyceraldehyde 3-phosphate (C_3)	0.5	0.9	0.4	0.5
D,L-Glyceraldehyde (C_3)	0.8	2.6	0.7	0.3
Glycolaldehyde (C_2)	0.4	2.8	0.4	0.1
NAD^+ (1)	80.4	9.1	69.0	7.6

(1) The following equation was used: $v = (V_{max}*[S])/(k_m+[S]-k_{irr}*[S])$.
(2) The following equation was used: $v = (V_{max}*[S])/(k_m+[S]+([S]^2/k_i)$.
The characterizations for the different substrates were performed in presence of 10 mM NAD^+.

For SSA, DOP and NAD^+ substrate inhibition was observed at concentrations higher than 2 mM (SSA, DOP) and 10 mM (NAD^+), respectively (Figure 32). From the curve fittings using the equations given in Table 14 and materials and methods the kinetic parameters were estimated. For D,L-GA(P) and glycolaldehyde classical Michealis-Menten kinetics were observed. The determined catalytic efficiencies indicate that SSA is the physiological substrate of SSADH-I, although significant activity with α-KGSA was observed (Table 14).

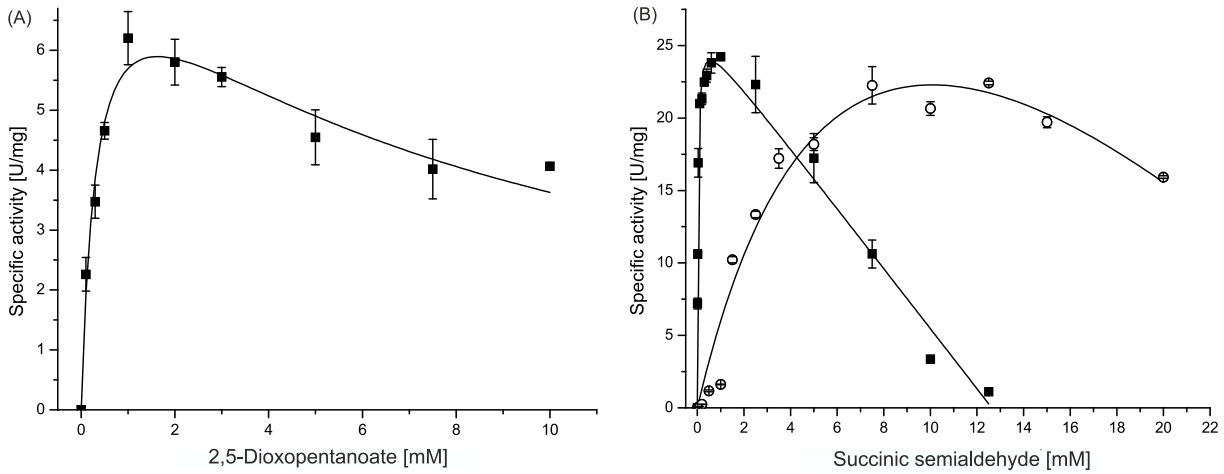


Figure 32 Substrate and cosubstrate concentration dependent velocity plots for SSADH-I with DOP (black squares) (A), SSA (black squares) as substrate and NAD^+ (white circles) as cosubstrate (B).
SSADH-I activity was determined in a continuous assay at 70°C by following the NADH formation at 340 nm.

In contrast to SSADH-I, SSO1842 (SSADH-II) uses both pyridine nucleotides $NADP^+$ (k_{cat}/K_m 3066.4 $mM^{-1}s^{-1}$) and NAD^+ (k_{cat}/K_m 12.3 $mM^{-1}s^{-1}$) as cosubstrate with a clear preference for $NADP^+$ (determined in the presence of 0.2 mM SSA) (Table 15 and Table 16, Figure 33).

Table 15 Kinetic parameters of SSADH-II (SSO1842) for various aldehydes with $NADP^+$ as cosubstrate.
Enzyme activity in presence of $NADP^+$ was determined in presence of 0.2 mM SSA at 70°C.

Substrate	$V_{max(app)}$	$K_{m(app)}$	$k_{cat(app)}$	$k_{cat}/K_{m(app)}$
	[U/mg protein]	[mM]	[s^{-1}]	[$mM^{-1}s^{-1}$]
Succinic semialdehyde (C_4) (1)	21.4	0.004	18.4	3454.2
2,5-Dioxopentanoate (C_5) (2)	24.7	0.2	21.2	134.1
D,L-Glyceraldehyde (C_3)	3.5	114	3.0	0.03
D,L-Glyceraldehyde 3-phosphate (C_3)	3.2	23.3	2.8	0.1
Glycolaldehyde (C_2)	2	21.2	1.8	0.1
$NADP^+$ (1)	19.2	0.01	16.5	3066.4

(1) The following equation was used: $v = (V_{max}*[S])/(k_m+[S]-k_{irr}*[S])$.
(2) The following equation was used: $v = (V_{max}*[S])/(k_m+[S]+([S]^2/k_i)$.
The characterizations for the different substrates were performed in presence of 1 mM $NADP^+$.

The enzyme most effectively converted SSA (k_{cat}/K_m 3454 $mM^{-1}s^{-1}$), followed by DOP (k_{cat}/K_m 134 $mM^{-1}s^{-1}$), D,L-GAP (k_{cat}/K_m 0.1 $mM^{-1}s^{-1}$), glycolaldehyde (k_{cat}/K_m 0.1 $mM^{-1}s^{-1}$) and D,L-GA (k_{cat}/K_m 0.03 $mM^{-1}s^{-1}$) at 70°C in the presence of both cosubstrates (Table 15 and Table 16). No enzyme activity was observed with formaldehyde, betaine aldehyde and benzaldehyde.

Table 16 Kinetic parameters of SSADH-II (SSO1842) for SSA and DOP with NAD^+ as cosubstrate.
Kinetic parameters for NAD^+ were determined in presence of 0.2 mM SSA at 70°C.

Substrate	$V_{max(app)}$	$K_{m(app)}$	$k_{cat(app)}$	$k_{cat}/K_{m(app)}$
	[U/mg protein]	[mM]	[s^{-1}]	[$mM^{-1}s^{-1}$]
Succinic semialdehyde (C_4) (1)	19.6	0.001	16.8	1789.1
2,5-Dioxopentanoate (C_5) (2)	31.8	0.1	27.3	215.3
NAD^+ (1)	158.6	11.1	136.3	12.3

(1) The following equation was used: $v = (V_{max}*[S])/(k_m+[S]-k_{irr}*[S])$.
(2): The following equation was used $v = (V_{max}*[S])/(k_m+[S]+([S]2/k_i)$.
The characterizations for the different substrates were performed in presence of 5 mM NAD^+.

Also, SSADH II showed inhibition kinetics with SSA and DOP at concentrations higher than 1 mM and also for both cosubstrates at concentrations above 1 mM ($NADP^+$) and 10 mM (NAD^+), respectively (Figure 33). With glycolaldehyde, D,L-GA and D,L-GAP the enzyme followed classical Michealis-Menten kinetics. The determined catalytic efficiencies suggest a physiological function in SSA and DOP conversion with a clear preference for SSA (16-fold higher K_{cat}/k_m, Table 15 and Table 16).

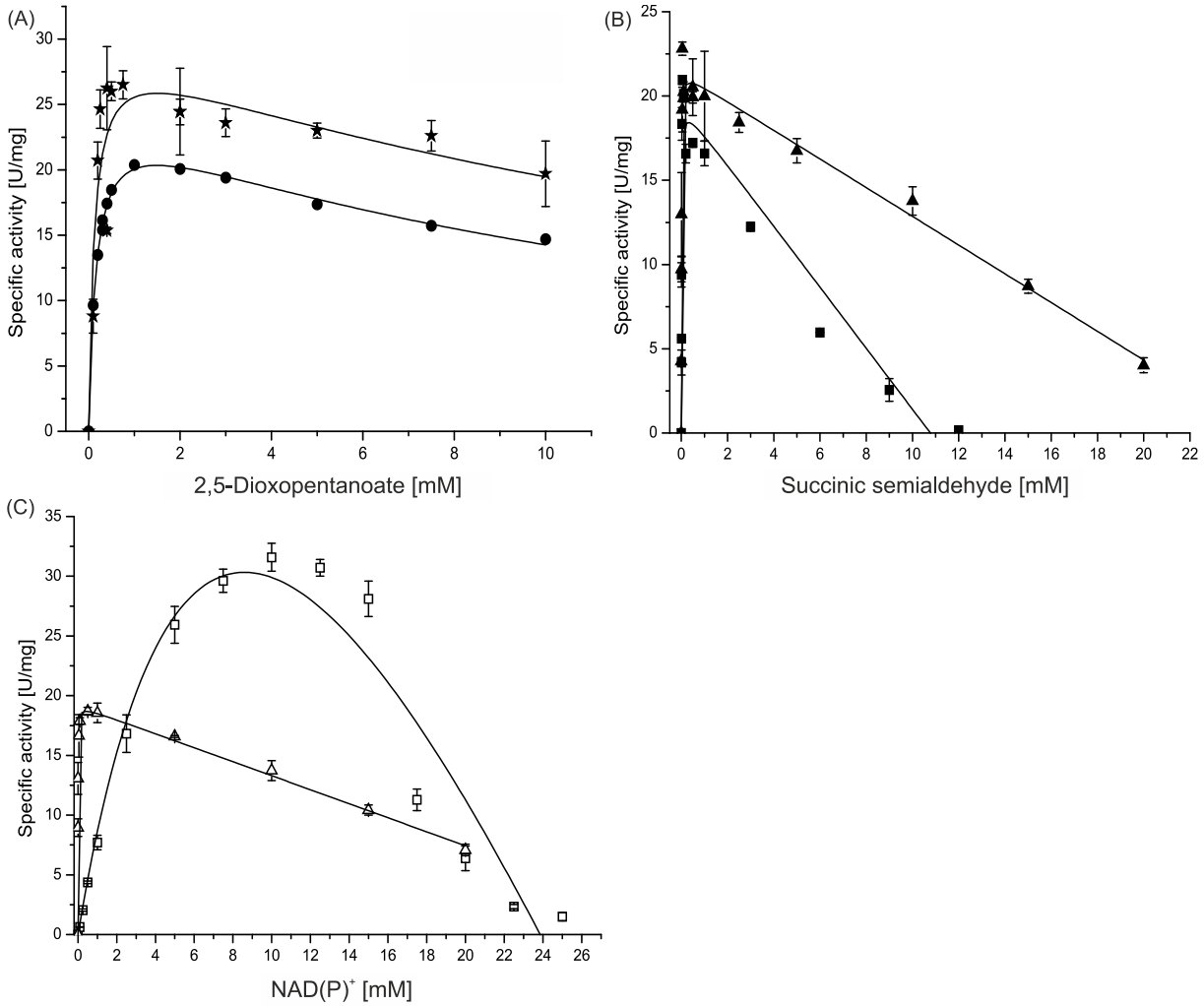


Figure 33 Substrate and cosubstrate concentration dependent velocity plots for SSADH-II.
For the substrate DOP the kinetic parameters were determined in the presence of NAD^+ (black stars) or $NADP^+$ (black circles) (A). For the substrate SSA the kinetic parameters were determined in the presence of NAD^+ (black squares) or $NADP^+$ (black triangles) (B). For the cosubstrates NAD^+ (white square) and $NADP^+$ (white triangle) kinetic parameters were determined in presence of SSA. SSADH-II activity was determined in a continuous assay at 70°C by following the NAD(P)H formation at 340 nm.

3.4.4. Enzyme characterisation of the methylmalonate semialdehyde dehydrogenase

SSO1218 (MSDH) showed strictly CoA-dependent ALDH activity with methylmalonate semialdehyde (MMSA), malonate semialdehyde (MSA) and propionaldehyde (PA) as substrates in the presence of the cosubstrate NAD^+ at 60°C. No activity could be detected in presence of $NADP^+$ as cosubstrate. The highest activity was determined in presence of PA (k_{cat} 0.03 s^{-1}) followed by MSA (k_{cat} 0.02 s^{-1}) and D,L-MMSA (k_{cat} 0.003 s^{-1}). No activity was observed for the substrates D,L-GA, D,L-GAP and D,L-glycolaldehyde neither with NAD^+ nor with $NADP^+$ as cosubstrate. However, PA is no physiological substrate for MSDH. It only forms the same thiopropionyl enzyme intermediate as MMSA and MSA, which makes it useful for the investigation of MSDHs. Unfortunately a detailed characterization was

impossible due to the thermolability of MSA and D,L-MMSA [172] and the low boiling/evaporating point of PA (46-50°C).

3.4.5. Succinic semialdehyde dehydrogenase isoenzymes

Our detailed enzymatic analysis revealed that, due to the clear preference for SSA conversion, SSADH-I can be classified as a succinic semialdehyde: NAD^+ oxidoreductase (EC 1.2.1.24) [273] and SSADH-II (SSO1842), based on its dual cosubstrate specificity for both pyridine nucleotides, as succinic semialdehyde: $NAD(P)^+$ oxidoreductase (EC 1.2.1.16). Thus, to our knowledge the two isoenzymes represent the first archaeal and (hyper)thermophilic SSADHs reported so far. For SSADH-I a native molecular mass of 157 kDa was determined suggesting a homotrimeric structure composed of 51.5 kDa subunits. Trimers are unusual among ALDH superfamily members, which mostly represent homodimers [274] or homotetramers [275]. However, a similar trimeric oligomerisation state has also been reported for the SSADH from Pseudomonas spp. [276]. In contrast, the homodimeric structure determined for SSADH-II (native molecular mass 114 kDa) is in agreement with the oligomeric structures of several characterized SSADHs e.g. from *Rattus norvegicus* [277].

Both *S. solfataricus* SSADHs are inhibited by elevated SSA concentrations, which is well established for various eukaryal enzymes (e.g. *Rhipicephalus microphilus* [278]). Furthermore, both SSADHs were inhibited by their cosubstrates, which was also shown for bacterial (e.g. *P. putida* inhibited by $NADP^+$ [279]) and eukaryal homologs (e.g. Human inhibited by NAD^+ [280]). So far only few SSADHs with dual cosubstrate specificity were reported (e.g. *E. coli* YneI, human ALDH8A1) [151]. Most eukaryotic SSADHs are specific for NAD^+ as cosubstrate and for the bacterial isoenzymes either NAD^+ or $NADP^+$ specificity was demonstrated.

Whereas the presence of SSADH isoenzymes in Eukaryotes is rather scarce (e.g. *Euglena gracilis*, [281]) they have been demonstrated in several bacterial species and different physiological functions have been demonstrated in metabolism of glutamate and arginine as well as of polyamines and γ-aminobutyric acid (GABA) metabolism (e.g. *Klebsiella pneumonia* [282], *E. coli* [151,283], *P. putida* [279], *Ralstonia eutropha* [284]). In *E. coli* one of two SSADH isoenzymes, the $NADP^+$-dependent SSADH (GabD) is primarily induced upon growth on GABA as sole nitrogen source and under nitrogen limitation [151]. In addition to growth on GABA, the second SSADH isoenzyme, the $NAD(P)^+$-dependent SSADH (Ynel) was significantly induced upon addition of exogeneous SSA (2 mM) as well as growth on arginine and putrescine (polyamine) as sole nitrogen sources, which results in the accumulation of endogenous SSA [151]. Bioinformatic analyses and reconstruction of the GABA, polyamine and glutamate/aginine metabolic pathways in *S. solfataricus* using the *E. coli* homologs revealed a functional association of both SSADH isoenzymes with two γ–aminobutyrate transaminase

(gabT) homologs (SSO2727, SSO3211) catalyzing the conversion of GABA to SSA (score value ~ 0.9, STRING 8.3)[181], which is further metabolized via the SSADH isoenzymes to succinate.

Furthermore, all homologs of the glutamate to arginine pathway as depicted in Figure 34 were identified in *S. solfataricus* and N-acetylglutamate synthase (ArgA) as well as argininosuccinate lyase (ArgH) activities have been demonstrated in crude extracts [285]. Additionally, the key enzyme of polyamine synthesis the polyamine synthase propylamine transferase (SSO0445) has been purified and characterized from *S. solfataricus* cells and [286] candidates encoding two further enzymes in polyamine metabolic pathway (i.e. speB-2 and speE) are also present in the *S. solfataricus* genome (Figure 34). Thus, those pathways in amino acid/polyamine metabolism finally resulting in the production of SSA via the GABA shunt are likely operative in *S. solfataricus* suggesting that both SSADHs characterized in this study function in these pathways catalyzing the final SSA oxidation to succinate which then enters the TCA cycle.

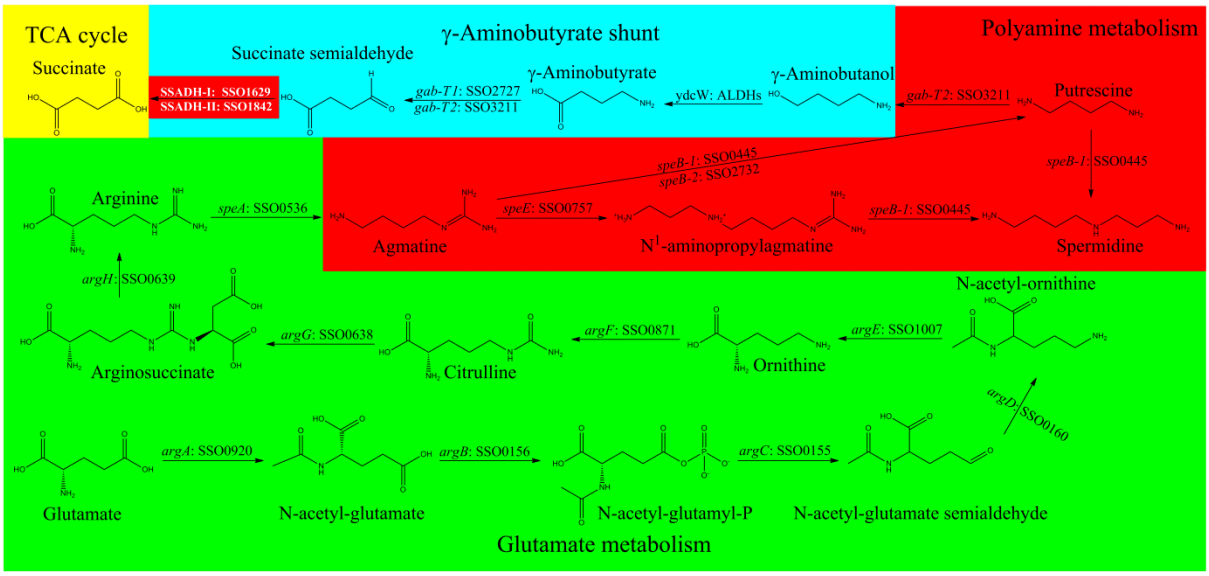


Figure 34 Pathway reconstruction of the glutamate (orange framed), polyamine (green framed) and GABA (blue framed) metabolism in *S. solfataricus*.

argA: amino-acid acetyltransferase; araB: acetylglutamate kinase; argC: N-acetyl-γ-glutamyl-phosphate reductase; argD: Acetylornithine/ succinyldiaminopimelate aminotransferase; argE: acetylornithine deacetylase; argF: ornithine carbamoyltransferase; argG: argininosuccinic synthase; argH: argininosuccinic lyase: speA: arginine decarboxylase; speB: agmatine ureohydrolase/ polyamine synthase; speE: aminopropyl transferase; ydcW: γ-aminobutanol dehydrogenase; gabT: γ-aminobutyrate transaminase/ putrescine aminotransferase; puuA: γ-glutamylputrescine synthase; puuB: γ-glutamylputrescine oxidoreductase; puuC: γ-glutamyl-γ-aminobutyraldehyde dehydrogenase; puuD: γ-glutamyl-γ-aminobutyrate hydrolase. The investigated enzymes are depicted in the red frame.

However, the enzyme catalyzing the γ-aminobutanol oxidation step in the GABA shunt in *S. solfataricus* remains to be identified and homologs of the second pathway for putrescine degradation in *E. coli*, the so called Puu bypass involving the SSADH YneI, appear to be absent in *S. solfataricus*. In addition to the presence of the respective pathways the presence of two SSADH

isoenzymes, which differ in cosubstrate specificity, resembles the situation found in *E. coli* and other Bacteria [151] and it is tempting to speculate that also in *S. solfataricus* the two SSADHs are regulated in response to nitrogen source and availability. However, since both enzymes might also have an important function in pentose degradation as discussed below respective growth, expression and genetic studies still have to be awaited in order to elucidate the physiological function of both isoenzymes.

Both characterized SSADHs show significant activity with DOP as substrate. The observed catalytic efficiency with DOP is compareable to those determined for mesophilic bacterial homologs (e.g. 190 $mM^{-1}s^{-1}$ *P. putida* [287]). Noticeable, the k_{cat}-value of the DOPDH (SSO3117), which was shown to function in D-arabinose degradation in *S. solfataricus* [27], is 8.6 s^{-1} and comparable to SSADH-I (7.13 s^{-1}) and SSADH-II (27.3 s^{-1}, NAD^+; 21.2 s^{-1}, $NADP^+$). Additionally, a function of DOPDH as glycolaldehyde dehydrogenase (k_{cat} 5.3 s^{-1}) in the D-xylose and L-arabinose degradation pathways was proposed [25]. The k_{cat} value of SSO3117 for glycolaldehyde is three times and 14 times higher compared to SSADH-II (k_{cat} 1.8 s^{-1}) and SSADH-I (k_{cat} 0.4 s^{-1}), respectively. These data might indicate that also SSADH-I and SSADH-II could play an additional role in pentose degradation, i.e. in DOP oxidation, in *S. solfataricus*.

However, both SSADHs (and also of the MSDH, see below) clearly showed no or only minor activities towards GAP or GA indicating that all three paralogs do not function as GAPN or GADH isoenzymes in the branched ED pathway in *S. solfataricus*. Instead, from the data now available for all five ALDH paralogs in *S. solfataricus* it can be concluded that the previously characterized GAPN (SSO3194) exclusively catalyzes the GAP oxidation to 3-PG in the course of the spED branch [116] and that the GA oxidation to glycerate in the npED branch is indeed carried out by the ferredoxin-dependent GA oxidoreductase (hetero trimer) encoded by SSO2636, SSO2637 and SSO2639, which has been characterized from the close relative *S. acidocaldarius* [114].

3.4.6. Methylmalonate semialdehyde dehydrogenase

SSO1218 was characterized as NAD^+ and CoA-dependent MSDH preferentially utilizing PA (k_{cat} 0.03 s^{-1}), followed by MSA (k_{cat} 0.02 s^{-1}) and MMSA (k_{cat} 0.003 s^{-1}). No ALDH activity was observed with D,L-GA, D,L-GAP and D,L-glycolaldehyde in presence of $NAD(P)^+$.The observed preference for PA (k_{cat} 12.6 s^{-1}) over MSA (k_{cat} 7.4 s^{-1}) and MMSA (k_{cat} 2.2 s^{-1}) was also shown in the case of the MSDH from *B. subtilis* (Bsu-MSDH) [146,147]. Nonetheless, MMSA and MSA are the only physiological substrates for MSDH. The unusual low specific activities of the SSO-MSDH for PA, MSA and MMSA are presumably due to the physico-chemical properties of these aldehydes, which make an appropriate biochemical investigation rather difficult. To our knowledge this is the first report of an archaeal and

(hyper)thermophilic MSDH, so far these enzymes have only been predicted by bioinformatic analysis [116,150]. MSDH from *S. solfataricus* represents a 130 kDa homodimer (54.6 kDa subunits), which was also reported for the MSDH from *Pseudomonas aeruginosa* [288].

The number of members of the ALDH superfamily which are CoA-dependent is quiet small compared to the number of the CoA-independent members. The sole CoA dependent members are the CoA-dependent SSADH (EC 1.2.1.76) investigated from *Clostridium kluyveri* and *Metallosphaera sedula* [289,290], acetaldehyde dehydrogenase (EC 1.2.1.10) investigated in several prokaryotes (e.g. *E. coli, B. subtilis*) [291,292] and eukaryotes (e.g. *Giardia intestinalis*, Human) [293,294]. These ALDHs are quiet unique among the ALDH superfamily with respect to their ping-pong type reaction mechanism. For MSDH it was shown that in case of MMSA and MSA a β-decarboxylation (HCO^{3-} release) prior to the transthioesterification step is the rate limiting step [146]. In the presence of PA, a decarboxylated analog of MMSA, due to the absence of the carboxyl group β-decarboxylation does not take place and the transthioesterification of the thiopropionylenzyme intermediate is the rate-limiting step.

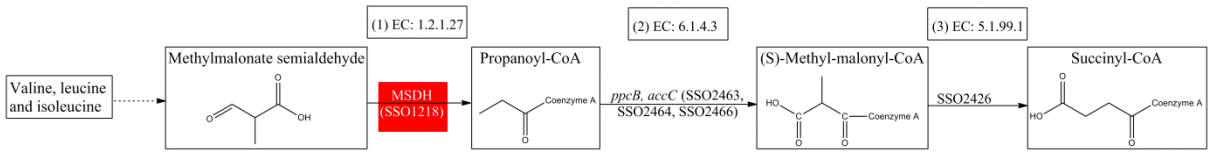


Figure 35 Physiological role of MSDH from *S. solfataricus* in the amino acid degradation of Val, Leu and Ile.
(1) Methylmalonate semialdehyde is converted by MSDH to propanoyl-CoA. (2) Propanoyl-CoA is further converted to (S)-metyhl-malonyl-CoA via acetyl-CoA/propanoyl-CoA carboxylase and (3) methyl-malonyl-CoA epimerase finaly generates succinyl-CoA which enters the TCA cycle. The investigated enzyme is depicted in the red frame.

MSDHs are involved in the degradation of the branched amino acids valine, leucine and isoleucine as known from bacteria (e.g Pseudomonas spp. [154]) oxidatively decarboxylating the key intermediate of the pathway MMSA to propionyl-CoA, which is further converted to succinyl-CoA and finally enters the TCA cycle (Figure 35). Bioinformatic analyses using the characterized *E. coli* homologs revealed that all candidates of the common part of the degradation pathway for branched amino acids are also present in *S. solfataricus* (Figure 35) suggesting a similar function of MSDHs in thermoacidophilic Archaea (Figure 36).

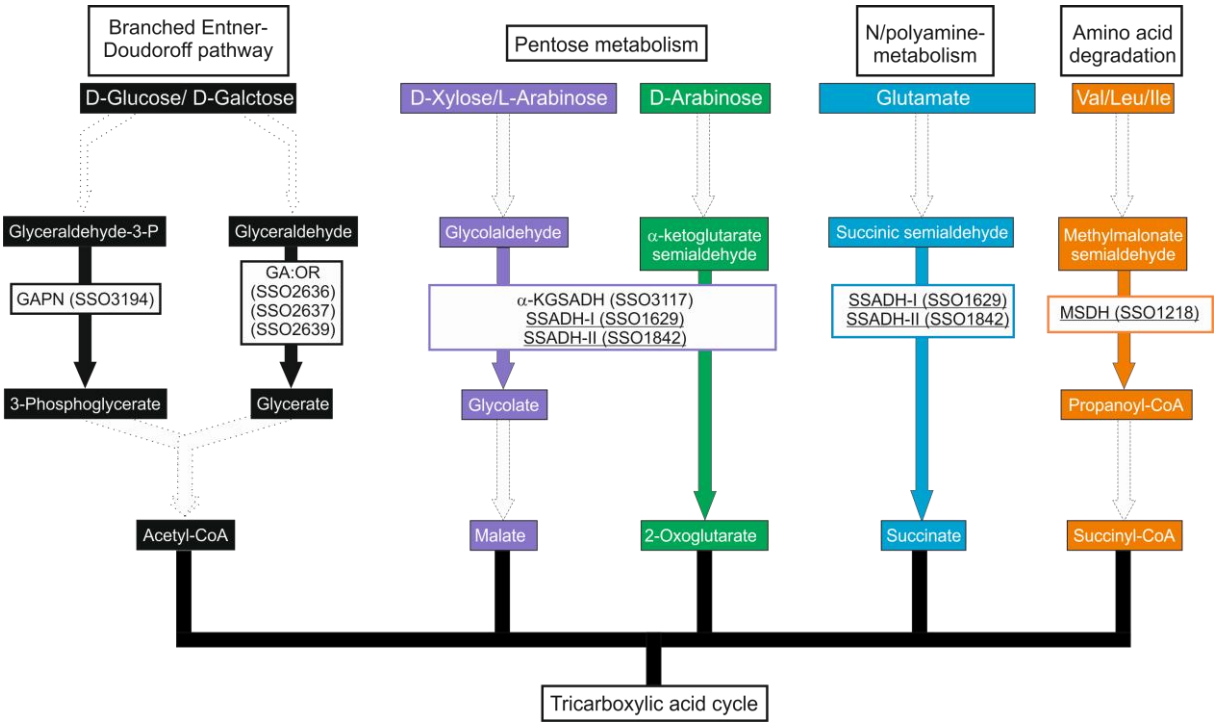


Figure 36 Current understanding of the role of the ALDHs in the CCM *S. solfataricus*.
The pathways for the degradation of D-glucose/D-galactose [28,57,109,110,132], D-xylose and L-arabinose [25], D-arabinose [27], N/polyamine-metabolism and valine, leucine and isoleucine are depicted. The investigated enzymes are underlined.

3.4.7. Phylogenetic affiliation of ALDH paralogs in *S. solfataricus*

All five SSO ALDH paralogs (SSADH-I (SSO1629), SSADH-II (SS1842) and MSDH (SSO1218) reported herein as well as the previously characterized GAPN and DOPDH) appear to belong to two of three main ALDH classes present in Archaea. As shown in the neighbor-joining phylogenetic tree in Figure 39 (constructed using the MEGA 5.1 software package based on a Clustal W alignment) the two SSADH paralogs from *S. solfataricus* are most closely related to each other (94% sequence identity) and belong to a crenarchaeal subcluster of ALDHs within the main class of GAPN related proteins. A second subcluster within this class contains the previously crystallized lactaldehyde dehydrogenase from *M. jannaschii* [295] (as well as some bacterial ALDH homologs) and a third subcluster comprises the GAPN homologs from all three domains of life (including the *S. solfataricus* GAPN SSO3194 as well as the GAPN from *T. tenax*) [116,118]. Conversely, the apparently functionally equivalent *E. coli* SSADHs (GabD and YneI) [151] belong to another main class of ALDHs which also contains archaeal homologs mainly from methanogens and halophiles as well as the GADH enzymes characterized from thermoacidophiles, e.g. *P. torridus* and *T. acidophilum* [296] previously reported to form a novel ALDH enzyme family. This indicates a convergent evolution of the *E. coli* SSADHs in the GADH/YenI/GabD/betaineALDH class and the SSO-SSADHs which thus might have adopt their

function through diversification within the ALDH class of GAPN related enzymes (the differences between *E. coli* SSADHs and SSO-SSADHs is also depicted in the sequence alignment in Figure 37).

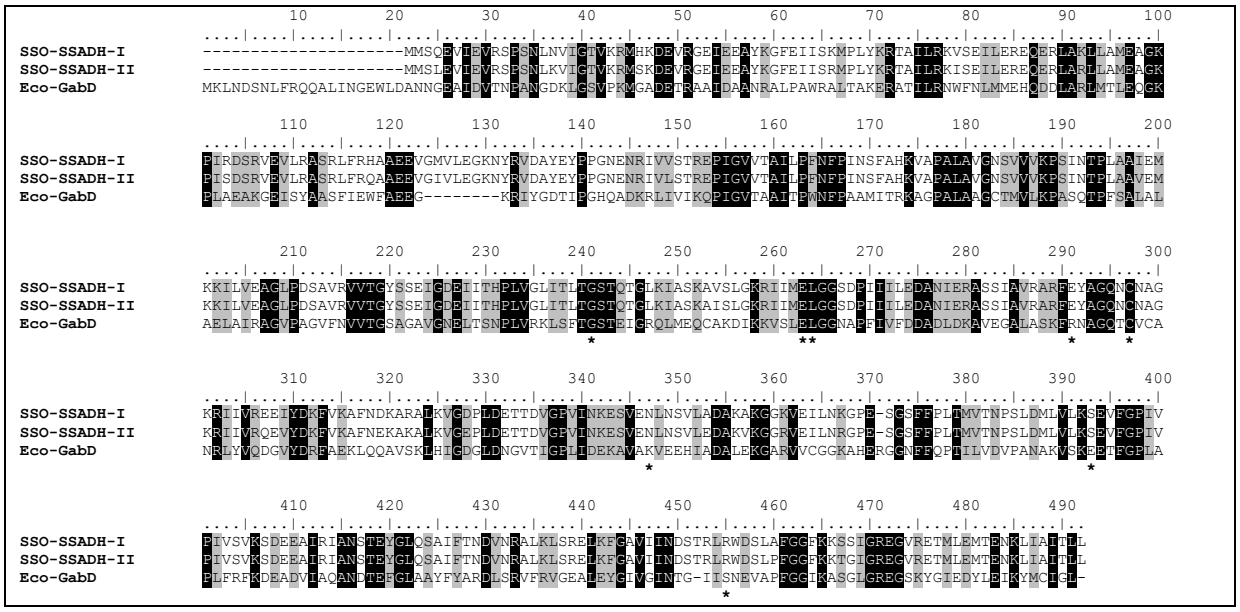


Figure 37 Sequence alignment of SSADHs from *S. solfataricus* and GabD from *E. coli*.
$NADP^+$-dependent SSADH from *E. coli* GabD (AAC75708.1), NAD^+-dependent SSADH-I (NP_343053.1) and $NADP^+$-dependent SSADH-II (NP_343247.1) from *S. solfataricus*. The sequence alignment was performed using ClustalW2 [297]. The crystal structure of gabD from *E. coli* was used to identify the SSADH conserved residues [298], which are marked with an asterix. Structurally important regions have been marked in either gray (similar residues) or black (conserved residues).

The third main class of ALDHs comprises the MSDH subgroup as well as a subgroup containing the DOPDH from *S. solfataricus*. The MSDH cluster also contains the characterized MSDH from *B. subtilis* [146]. This homology is also reflected by the high amino acid conservation particularly of functionally important residues in both enzymes shown in the sequence alignment in Figure 38. Figure 38 Sequence alignment of MSDH from *S. solfataricus* and selected MSDHs.

This again supports the assumption that SSO-MSDH fulfill a similar function in the metabolism of branched chain amino acids. Interestingly, also the *S. solfataricus* DOPDH seems not or only distantly related to the DOPDH-I characterized from *Azospirillum brasilense* which appear to cluster with the GADH/YenI/GabD/betaineALDH related enzymes again indicating convergent evolution.

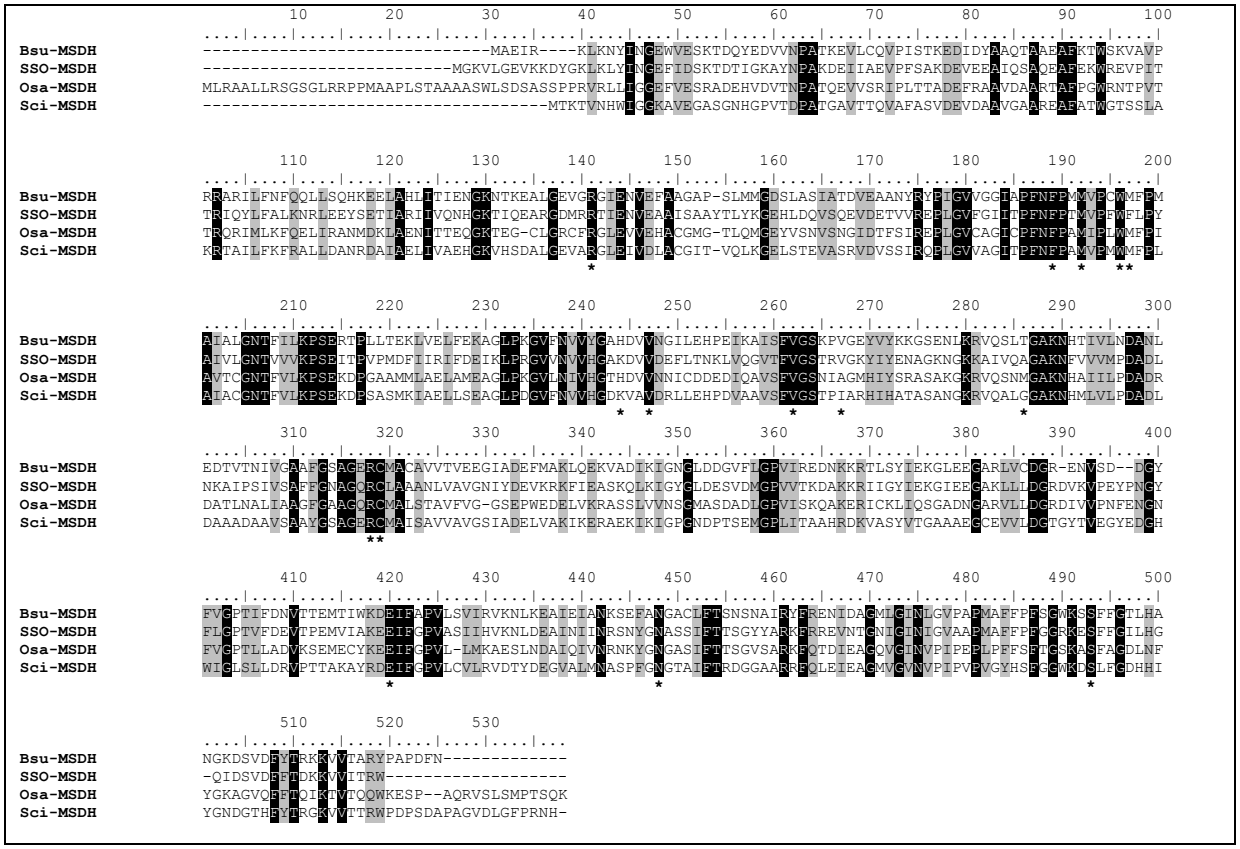


Figure 38 Sequence alignment of MSDH from *S. solfataricus* and selected MSDHs.
MSDH from *B. subtilis* (EHA31950.1), *S. solfataricus* (AAK41465.1), *Oryza sativa* (AAC03055.1) and *Streptomyces cinnamonensis* (DQ005575.1).The sequence alignment was performed using ClustalW2 [297]. The crystal structure of MSDH from *B. subtilis* was used identify the MSDH conserved residues [292], which are marked with an asterix. Structurally important regions have been marked in either gray (similar residues) or black (conserved residues).

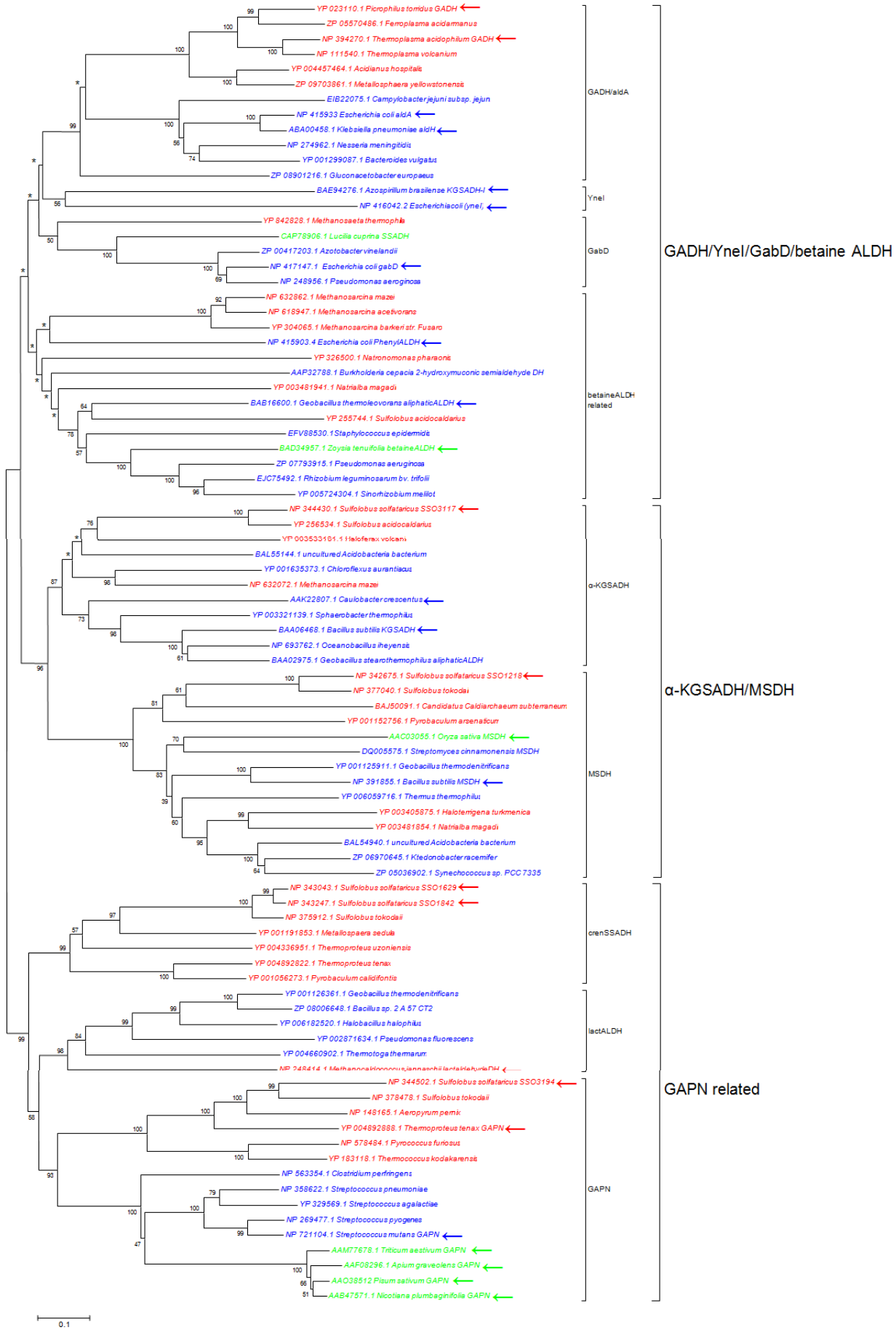

YP 023110.1 Picrophilus torridus GADH
ZP 05570486.1 Ferroplasma acidarmanus
NP 394270.1 Thermoplasma acidophilum GADH
NP 111540.1 Thermoplasma volcanium
YP 004457464.1 Acidianus hospitalis
ZP 09703861.1 Metallosphaera yellowstonensis
EIB22075.1 Campylobacter jejuni subsp. jejuni
NP 415933 Escherichia coli aldA
ABA00458.1 Klebsiella pneumoniae aldH
NP 274962.1 Nesseria meningitidis
YP 001299087.1 Bacteroides vulgatus
ZP 08901216.1 Gluconacetobacter europaeus
GADH/aldA
BAE94276.1 Azospirillum brasilense KGSADH-I
NP 416042.2 Escherichia coli (ynel)
Ynel
YP 842828.1 Methanosaeta thermophila
CAP78906.1 Lucilia cuprina SSADH
ZP 00417203.1 Azotobacter vinelandii
NP 417147.1 Escherichia coli gabD
NP 248956.1 Pseudomonas aeroginosa
GabD
NP 632862.1 Methanosarcina mazei
NP 618947.1 Methanosarcina acetivorans
YP 304065.1 Methanosarcina barkeri str. Fusaro
NP 415903.4 Escherichia coli PhenylALDH
YP 326500.1 Natronomonas pharaonis
AAP32788.1 Burkholderia cepacia 2-hydroxymuconic semialdehyde DH
YP 003481941.1 Natrialba magadii
BAB16600.1 Geobacillus thermoleovorans aliphaticALDH
YP 255744.1 Sulfolobus acidocaldarius
EFV88530.1 Staphylococcus epidermidis
BAD34957.1 Zoysia tenuifolia betaineALDH
ZP 07793915.1 Pseudomonas aeruginosa
EJC75492.1 Rhizobium leguminosarum bv. trifolii
YP 005724304.1 Sinorhizobium meliloti
betaineALDH related
GADH/Ynel/GabD/betaine ALDH
NP 344430.1 Sulfolobus solfataricus SSO3117
YP 256534.1 Sulfolobus acidocaldarius
BAL55144.1 uncultured Acidobacteria bacterium
YP 001635373.1 Chloroflexus aurantiacus
NP 632072.1 Methanosarcina mazei
AAK22807.1 Caulobacter crescentus
YP 003321139.1 Sphaerobacter thermophilus
BAA06468.1 Bacillus subtilis KGSADH
NP 693762.1 Oceanobacillus iheyensis
BAA02975.1 Geobacillus stearothermophilus aliphaticALDH
α-KGSADH
NP 342675.1 Sulfolobus solfataricus SSO1218
NP 377040.1 Sulfolobus tokodaii
BAJ50091.1 Candidatus Caldiarchaeum subterraneum
YP 001152756.1 Pyrobaculum arsenaticum
AAC03055.1 Oryza sativa MSDH
DQ005575.1 Streptomyces cinnamonensis MSDH
YP 001125911.1 Geobacillus thermodenitrificans
NP 391855.1 Bacillus subtilis MSDH
YP 006059716.1 Thermus thermophilus
YP 003405875.1 Haloterrigena turkmenica
YP 003481854.1 Natrialba magadii
BAL54940.1 uncultured Acidobacteria bacterium
ZP 06970645.1 Ktedonobacter racemifer
ZP 05036902.1 Synechococcus sp. PCC 7335
MSDH
α-KGSADH/MSDH
NP 343043.1 Sulfolobus solfataricus SSO1629
NP 343247.1 Sulfolobus solfataricus SSO1842
NP 375912.1 Sulfolobus tokodaii
YP 001191853.1 Metallosphaera sedula
YP 004336951.1 Thermoproteus uzoniensis
YP 004892822.1 Thermoproteus tenax
YP 001056273.1 Pyrobaculum calidifontis
crenSSADH
YP 001126361.1 Geobacillus thermodenitrificans
ZP 08006648.1 Bacillus sp. 2 A 57 CT2
YP 006182520.1 Halobacillus halophilus
YP 002871634.1 Pseudomonas fluorescens
YP 004660902.1 Thermotoga thermarum
lactALDH
NP 344502.1 Sulfolobus solfataricus SSO3194
NP 378478.1 Sulfolobus tokodaii
NP 148165.1 Aeropyrum pernix
YP 004892888.1 Thermoproteus tenax GAPN
NP 578484.1 Pyrococcus furiosus
YP 183118.1 Thermococcus kodakarensis
NP 563354.1 Clostridium perfringens
NP 358622.1 Streptococcus pneumoniae
YP 329569.1 Streptococcus agalactiae
NP 269477.1 Streptococcus pyogenes
NP 721104.1 Streptococcus mutans GAPN
AAM77678.1 Triticum aestivum GAPN
AAF08296.1 Apium graveolens GAPN
AAO38512 Pisum sativum GAPN
AAB47571.1 Nicotiana plumbaginifolia GAPN
GAPN
GAPN related
0.1

Figure 39 Phylogenetic affiliation of selcted ALDH sequences from Archaea (red), Bacteria (blue) and Eukarya (green) constructed using the Neighbor-Joining method [299].

NCBI accession or genebank numbers of the respective proteins as well as the organism names are given. Characterized sequences are indicated by arrows. The optimal tree with the sum of branch length = 24.36501941 is shown. Bootstrap values (1000 replicates) are shown at the nodes [300], values below 50 are replaced by asterisks. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method [301] and are in the units of the number of amino acid substitutions per site. The analysis involved 86 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 389 positions in the final dataset. Evolutionary analyses were conducted in MEGA5 [302].

3.5. Characterization of the glucose-1-dehydrogenase (Saci_1079) from Sulfolobus acidocaldarius DSM 639

In order to investigate the CCM of *S. acidocaldarius* and to address pathway promiscuity as well as metabolic complexity in *S. acidocaldarius* the first enzyme of the ED-pathway (glycolysis) the GDH (Saci_1079) was characterized in respect to (co)substrate specificity as well as kinetic parameters.

3.5.1. Heterologues expression in *E. coli* BL21 pRIL and purification

The Saci-GDH (Saci_1079) was expressed in *E. coli* BL21-CodonPlus (DE3)-RIL via induction by IPTG (see section 2.7.2.4.). The enzyme was found in the soluble fraction and was purified to apparent homogeneity via heat precipitation (70°C), anion exchange chromatography and size exclusion chromatography. After purification a homogeneous band at approximately 40 kDa was visible (Figure 40), which agrees with the theoretical molecular weight of 40.4 kDa Saci_1079. From 5.5 g recombinant cells (wet weight) 0.6 mg Saci-GDH were obtained.

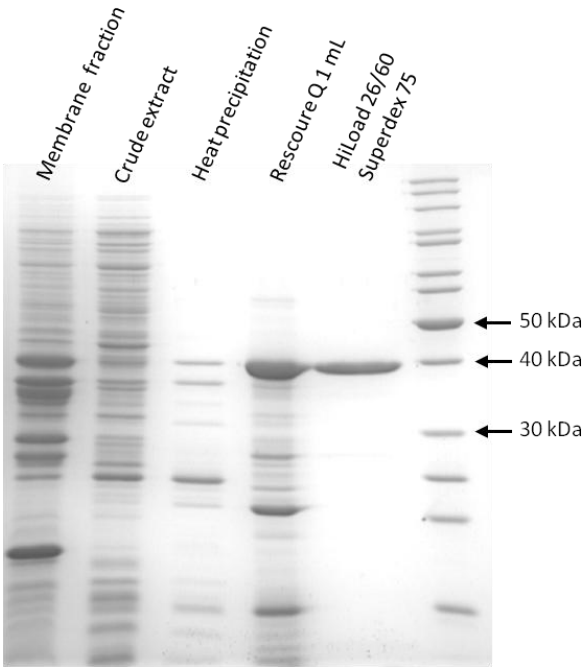


Figure 40 Purification of the recombinant glucose-1-dehydrogenase (Saci_1079) from *S. acidocaldarius* DSM 639.
10 µg membrane fraction, crude extract, heat precipitation and protein fraction after Resource Q (6 mL) and 2 µg protein fractions after HiLoad 26/60 Superdex 75, were applied on SDS-PAGE (12.5%). The gel was stained with Coomassie brilliant blue.

3.5.2. Substrate specificity of glucose-1-dehydrogenase

GDH activity of Saci_1079 was determined for sixteen different substrates in the presence of either 5 mM NAD^+ or $NADP^+$ in Tris/HCl, pH 6.5 at 70°C (Figure 41). Highest specific activity (> 100 U/mg) was observed with D-xylose, L-arabinose, 6-deoxy-D-glucose, D-fucose, D-galactose and D-glucose as substrate with both NAD^+ and $NADP^+$ as cosubstrate. Minor activity was observed with 2-deoxy-D-glucose, maltose, D-glucosamine, D-ribose, D-lyxose, D-lactose, D-arabinose, D-mannose, L-xylose and sucrose with both cosubstrates (Figure 41) .

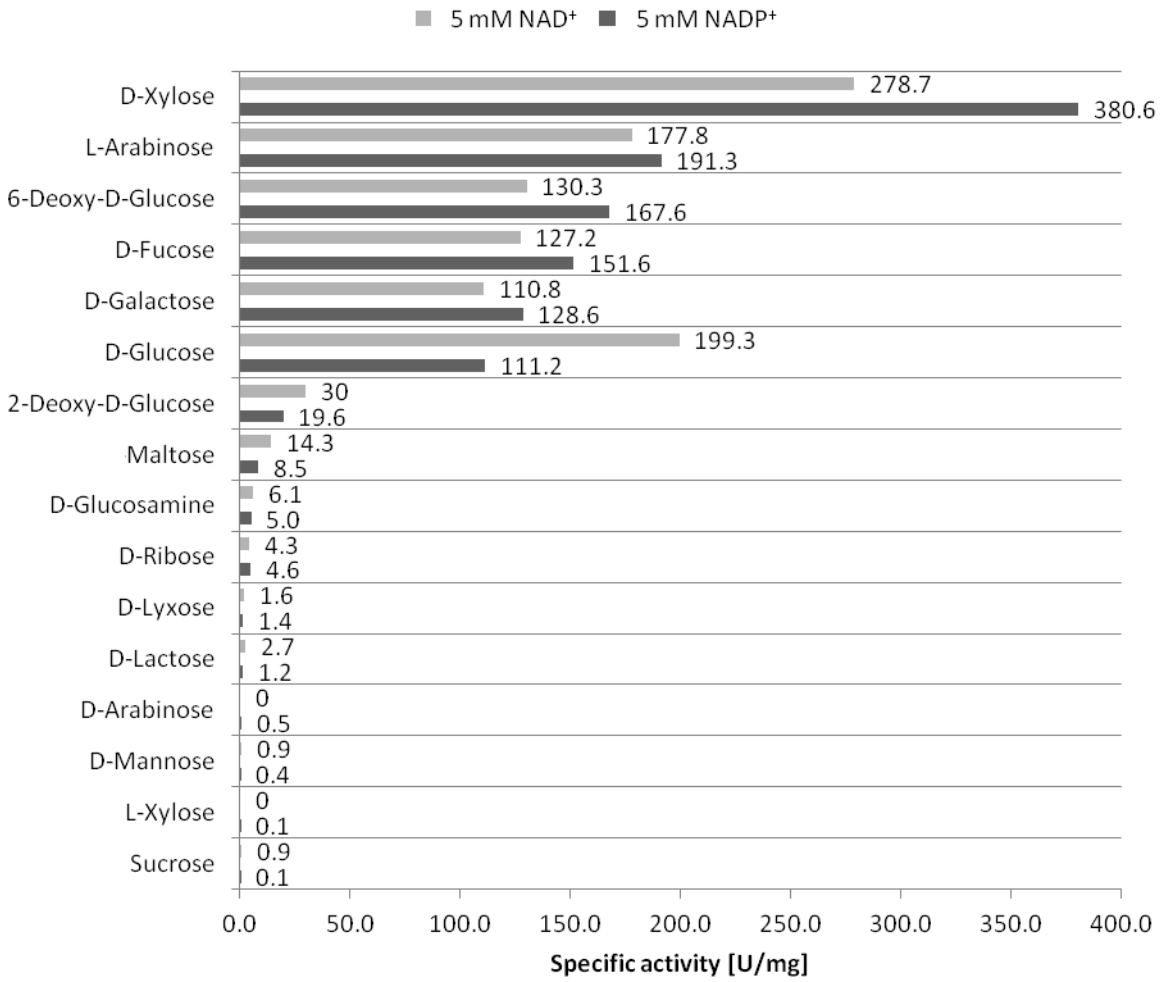


Figure 41 Enzyme activity of Saci-GDH in presence of various sugars and either 5 mM NAD^+ (grey) or 5 mM $NADP^+$ (black) at 70°C (pH 6.5).

Some changes in substrate specificity were observed depending on the cosubstrate. In the presence of $NADP^+$ the highest specific activity was detected for D-xylose (380.6 U/mg) followed by L-arabinose (191.3 U/mg), 6-deoxy-D-glucose (167.6 U/mg), D-fucose (151.6 U/mg), D-galactose (128.6 U/mg) and D-glucose (11.2 U/mg), whereas in the presence of NAD^+ the highest specific activity was determined with D-xylose (278.7 U/mg) followed by D-glucose (199.3 U/mg), L-arabinose (177.8 U/mg), 6-deoxy-D-glucose (130.3 U/mg), D-fucose (127.2 U/mg) and D-galactose (110.8 U/mg).

3.5.3 Temperature optimum of GDH

Analysis of the GDH activity in a temperature range from 20°C up to 100°C revealed that Saci_1079 possesses a very broad temperature range. At low temperature only minor enzyme activity was detectable (20-40°C, 1.6-9.1% relative activity). The temperature optimum of Saci_1079 was determined at 90°C (126%, 272 U/mg), which is approximately 12°C above the optimal growth temperature of *S. acidocaldarius* DSM 639 [31]. However, high enzyme activity was only observed in a narrow temperature range (82-92°C) and at 100°C the enzyme lost more than 50% of its maximal activity.

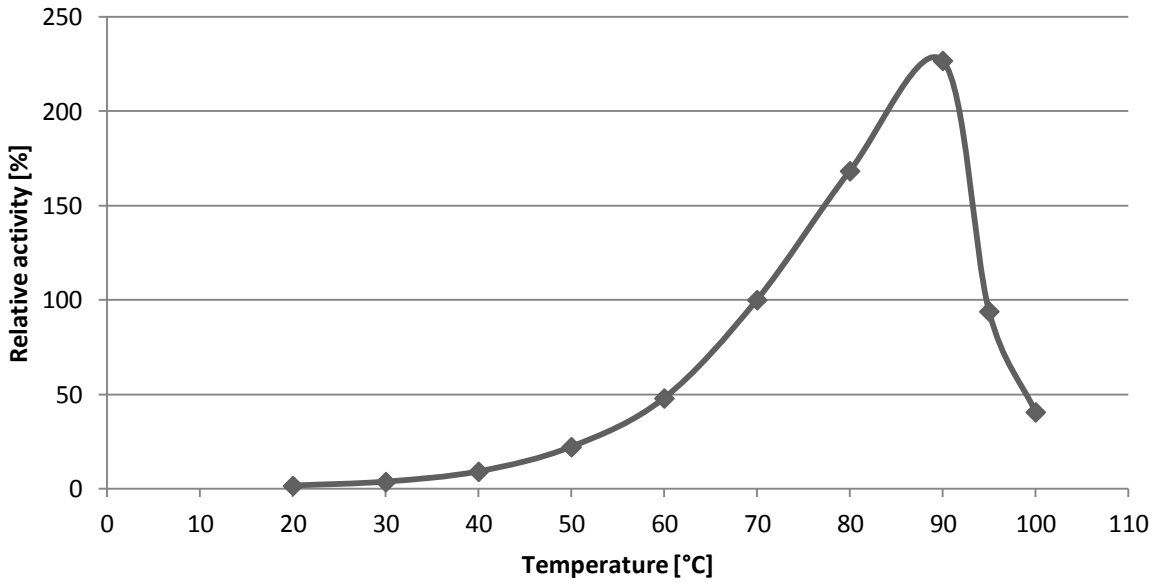


Figure 42 Temperature dependence of Saci-GDH (Saci_1079) activity.
The enzyme activity was determined in 100 mM Tris/HCl, (pH 6.5 at the respective temperature, 5 mM $MgCl_2$, 2.5 mM $NADP^+$ and 10 mM D-xylose). Enzyme activity was determined in a continuous assay monitoring the NAD(P)H formation at 340 nm.

3.5.4. Influence of Me^{2+}-ions on GDH activity

Most of the investigated GDHs require Me^{2+}-ions for activity, which are bound into their active center. Saci-GDH showed highest activity in the presence of Mn^{2+}- followed by Mg^{2+}-,Co^{2+}-, Ni^{2+}- and Cd^{2+}-ions. No enzyme activity was observed in the presence of Cu^{2+}-ions. Whereas the highest activity was detected in presence of 2.5 mM $MnCl_2$. However, only a slight reduced activity was observed with $MgCl_2$, which was used in the assays (Figure 43).

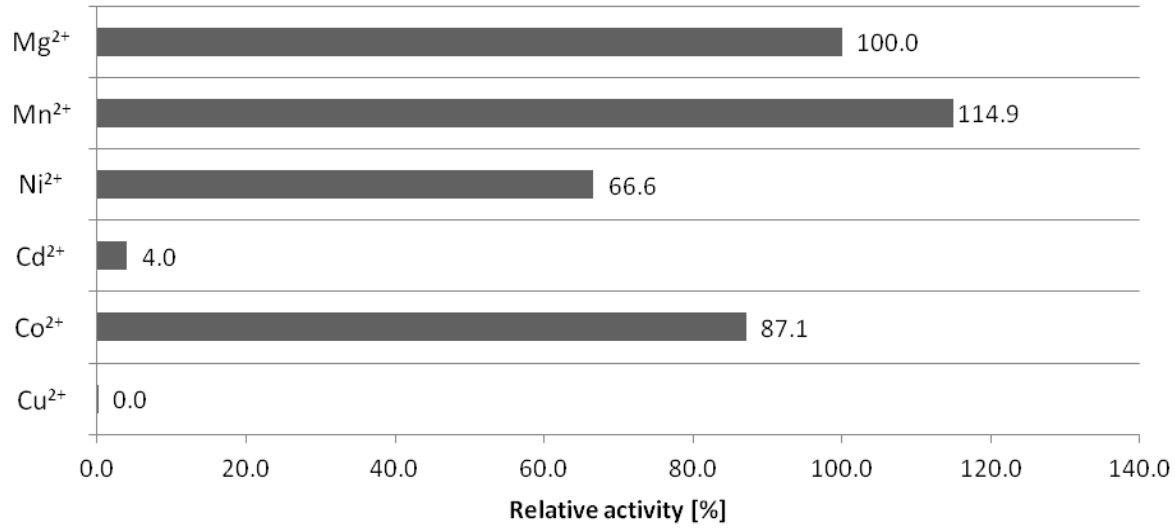


Figure 43 Influence of different Me^{2+}-ions on Saci-GDH activity at 70°C.
The enzyme activity was measured in the presence of 2.5 mM of the corresponding $MeCl_2$ compound in 100 mM Tris/HCl, pH 6.5 at 70°C. Enzyme activity was determined in a continuous assay at 70°C by monitoring the NAD(P)H formation at 340 nm.

3.5.5. Enzyme characterization of the GDH from *S. acidocaldarius*

Saci-GDH revealed activity with both cosubstratres NAD^+ and $NADP^+$, but $NADP^+$ was the preferred cosubstrate (k_{cat}/K_m: $NADP^+$ 1987 $mM^{-1}s^{-1}$ compared NAD^+ 367 $mM^{-1}s^{-1}$; Figure 44, Table 17). Inhibition was observed for both cosubstrates at elevated concentrations (> 8 mM NAD^+, > 1 mM $NADP^+$, Figure 44). However, the determined K_m-values (K_m 1.9 mM NAD^+, K_m 0.08 mM $NADP^+$) were far below the concentrations where the cosubstrate inhibition occured.

Table 17 Kinetic parameters of Saci-GDH for the cosubstrates NAD^+ and $NADP^+$ in presence of D-Xylose as substrate.

Cosubstrate	V_{max}	K_m	k_{cat}	k_{cat}/K_m
	[U/mg]	[mM]	[s^{-1}]	[$mM^{-1}s^{-1}$]
$NAD^{+(1)}$	1040.1	1.9	699.9	368.34
$NADP^{+(1)}$	236.2	0.08	158.9	1986.6

(1) The following equation was used: $v = (V_{max}*[S])/(k_m+[S]-k_{irr}*[S])$.

Saci-GDH was characterized in presence of $NAD(P)^+$ for all six substrates, which resulted in specific activities above 100 U/mg. The highest catalytic efficiency in presence of 7.5 mM NAD^+ was observed for 6-deoxy-D-glucose (557.3 $mM^{-1}s^{-1}$) followed by D-xylose (143.3 $mM^{-1}s^{-1}$), D-fucose (23.2 $mM^{-1}s^{-1}$), D-glucose (22.0 $mM^{-1}s^{-1}$), D-galactose (16.1 $mM^{-1}s^{-1}$) and L-arabinose (6.5 $mM^{-1}s^{-1}$, Figure 45, Table 18).

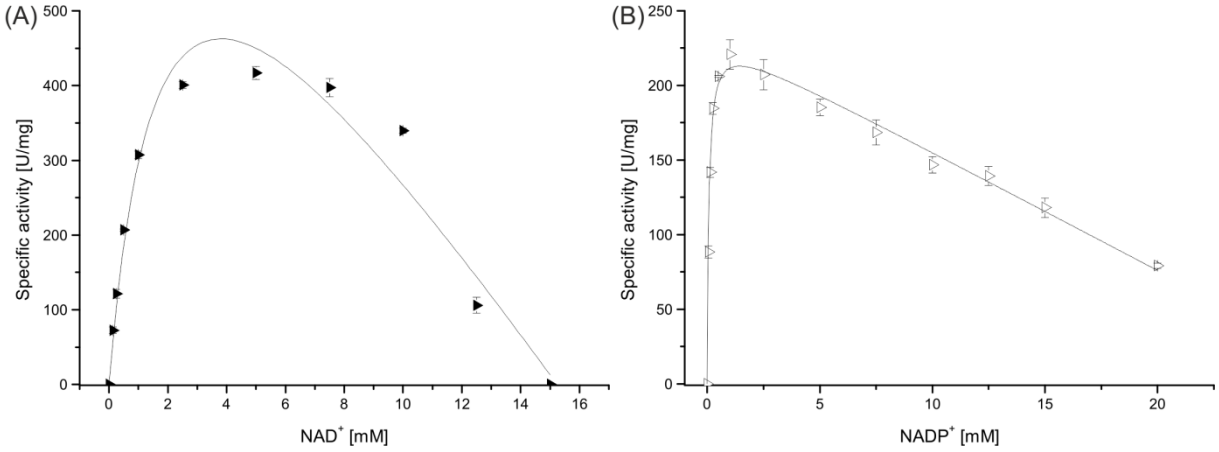


Figure 44 Cosubstrate concentration dependent velocity plots for Saci-GDH with D-xylose as substrate.
For the cosubstrates NAD^+ (A) and $NADP^+$ (B), kinetic parameters were determined in the presence of D-xylose (10 mM D-xylose in case of $NADP^+$, 30 mM D-xylose in case of NAD^+). Enzyme activity was determined in a continuous assay at 70°C by monitoring the NAD(P)H formation at 340 nm. Three independent measurements were performed. Mean-values are depicted and the standard deviation is given.

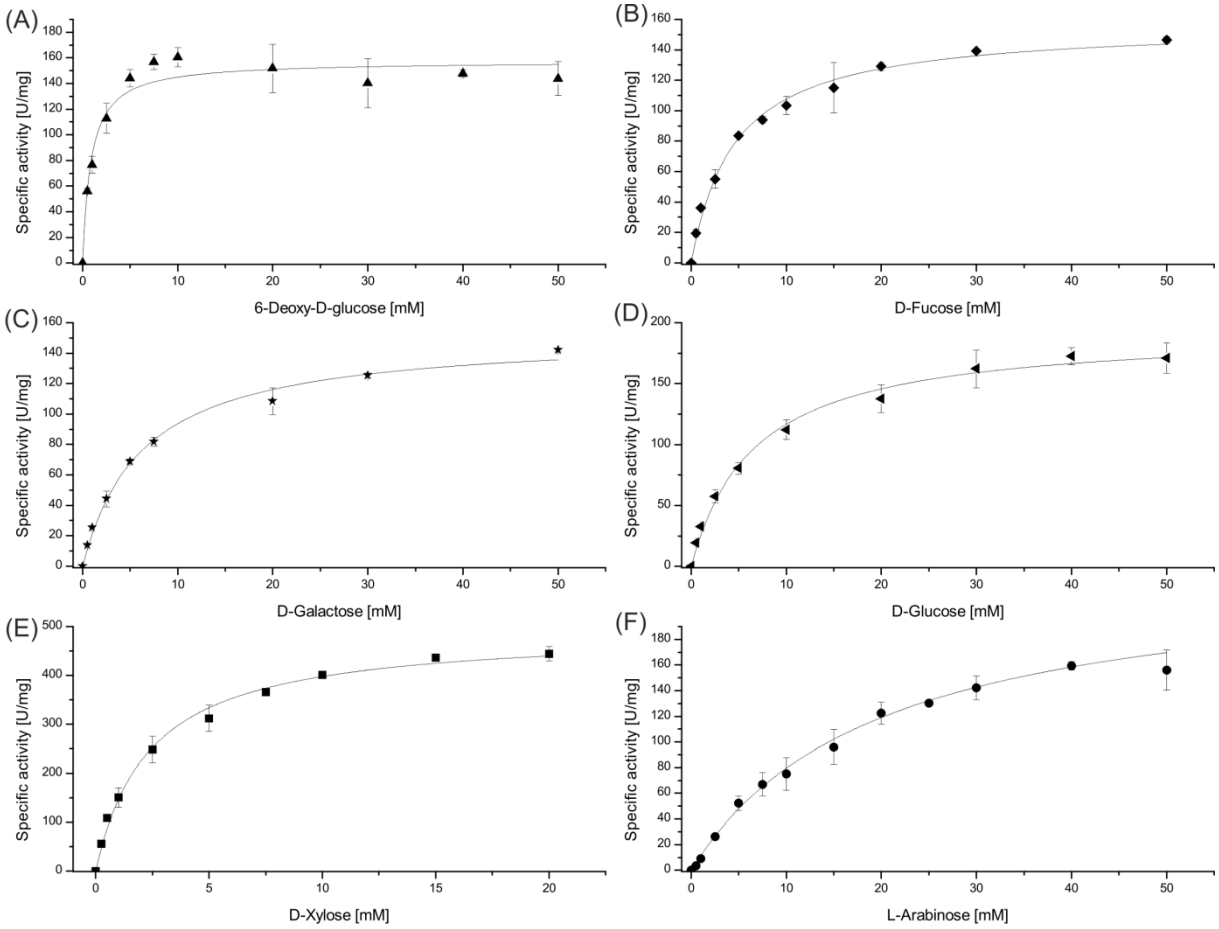


Figure 45 Substrate concentration dependent velocity plots for Saci-GDH with NAD^+ as cosubstrate.
For the substrates 6-deoxy-D-glucose (A), D-fucose (B), D-galactose (C), D-glucose (D), D-xylose (E) and L-arabinose (F) kinetic parameters were determined in the presence of 7.5 mM NAD^+. GDH activity was determined in a continuous assay at 70°C by following the NADH formation at 340 nm. Three independent measurements were performed. Mean-values are depicted and the standard deviation is given.

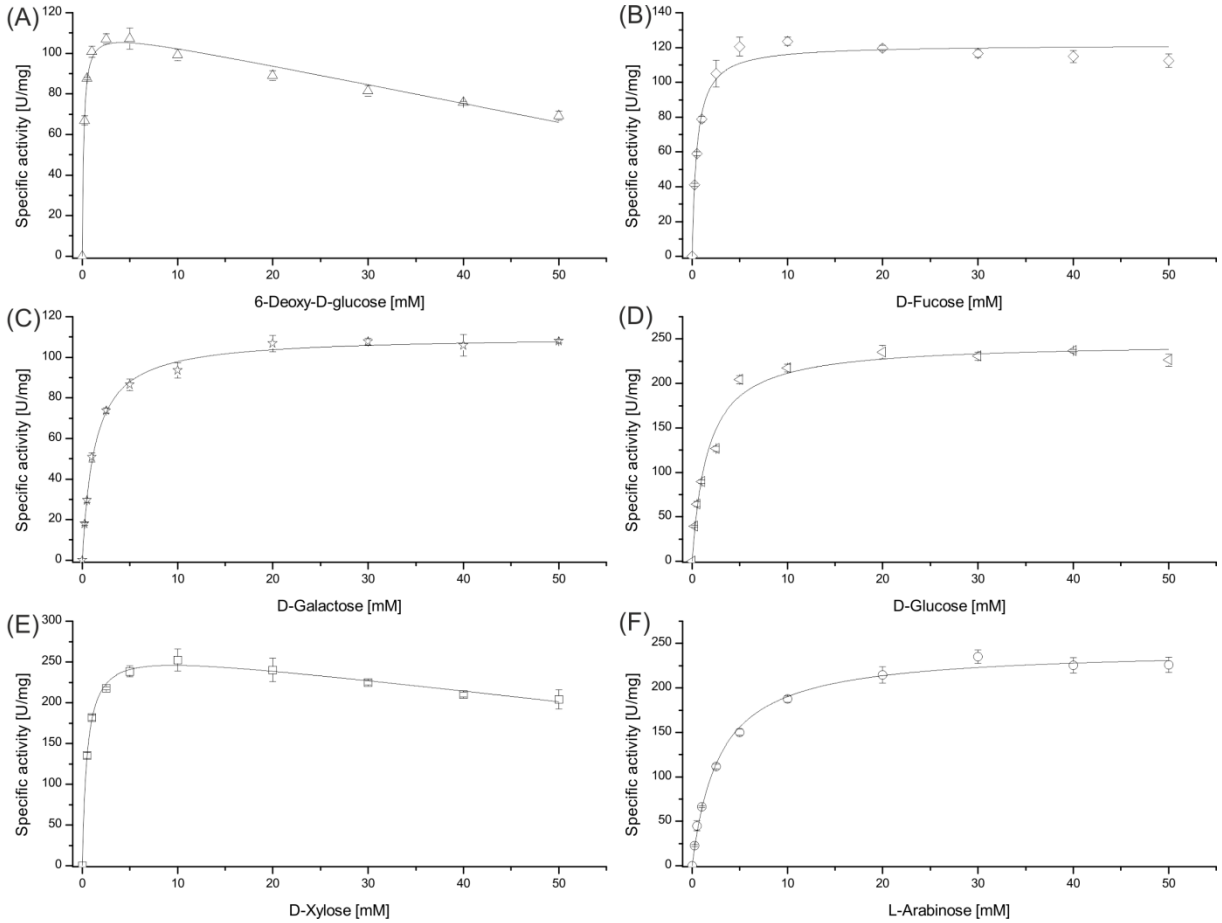


Figure 46 Substrate concentration dependent velocity plots for Saci-GDH with $NADP^+$ as cosubstrate..
For the substrates 6-deoxy-D-glucose (A), D-fucose (B), D-galactose (C), D-glucose (D), D-xylose (E) and L-arabinose (F) kinetic parameters were determined in presence of 2.5 mM NAD^+. Saci_1079 activity was determined in a continuous assay at 70°C by following the NADH formation at 340 nm. Three independent measurements were performed. Mean-values are depicted and the standard deviation is given.

Table 18 Kinetic parameters of Saci-GDH for various substrates with NAD^+ and $NADP^+$ as cosubstrates at 70°C.

Substrate	V_{max}	K_m	k_{cat}	k_{cat}/K_m	Cosubstrate
	[U/mg]	[mM]	[s^{-1}]	[$mM^{-1}s^{-1}$]	
D-Xylose	491.8	2.3	330.8	143.3	NAD^+ (2.5 mM)
L-Arabinose	221.9	23.1	149.3	6.5	
6-Deoxy-D-glucose	157.4	0.2	105.8	557.3	
D-Fucose	156.9	4.6	105.6	23.2	
D-Galactose	152.7	6.3	102.8	16.1	
D-Glucose	194.2	5.8	130.7	22.0	
D-Xylose[(1)]	272.6	0.4	183.3	359.7	$NADP^+$ (2.5 mM)
L-Arabinose	244.0	3.8	164.2	43.3	
6-Deoxy-D-glucose	113.1	0.2	76.0	507.2	
D-Fucose	121.6	0.5	81.7	132.0	
D-Galactose	110.4	2.7	74.3	27.6	
D-Glucose	246.0	2.6	165.4	64.2	

[(1)] The following equation was used: $v = (V_{max}*[S])/(k_m+[S]-k_{irr}*[S])$.

The characterization of Saci-GDH in the presence of $NADP^+$ revealed a similar preference for the tested substrates (Table18, Figure 46) only the preference for L-arabinose was higher than for D-galactose compared to $NADP^+$. The highest catalytic efficiency was observed for 6-deoxy-D-glucose (507.2 $mM^{-1}s^{-1}$) followed by D-xylose (359.7 $mM^{-1}s^{-1}$), D-fucose (132.0 $mM^{-1}s^{-1}$), D-glucose (64.2 $mM^{-1}s^{-1}$), L-arabinose (43.3 $mM^{-1}s^{-1}$) and D-galactose (27.6 $mM^{-1}s^{-1}$).

3.5.6. Glucose-1-dehydrogenases in Archaea

Sulfolobus species harbor several GDH paralogs. Most GDHs are known to use a variety of different sugars as substrate and possesses dual cosubstrate specificity for both pyridine nucleotides [27,109,113,124,303]. BLAST P analysis in *S. solfataricus* P2 revealed the presence of 14 putative GDH paralogs of the medium-chain alcohol/polyol dehydrogenase/reductase (MDR) branch of the pyridine nucleotide-dependent alcohol/polyol/sugar dehydrogenase superfamily [124].

Table 19 Comparison of kinetic parameters of the characterized members of the MDR branch of the pyridine nucleotide-dependent alcohol/polyol/sugar dehydrogenase superfamily from *S. solfataricus* P2 and GDH (Saci_1079) from *S. acidocaldarius* DSM 639 at 70°C.

Substrate	SSO-GDH-1 (SSO3003)[109,132]		SSO-GDH-2 (SSO3204)[113]		SSO-D-AraDH (SSO1300)[27,124]		Saci-GDH (Saci_1079)	
	NAD^+	$NADP^+$	NAD^+	$NADP^+$	NAD^+	$NADP^+$	NAD^+	$NADP^+$
D-Glucose k_{cat}/K_m ($mM^{-1}s^{-1}$)	49.9	36.7	14.9	115.2	-	-	22.0	64.2
D-Galactose k_{cat}/K_m ($mM^{-1}s^{-1}$)	108	85	-	-	-	-	22.0	27.7
D-Xylose k_{cat}/K_m ($mM^{-1}s^{-1}$)	245	261	-	-	-	-	143.3	359.7
L-Arabinose k_{cat}/K_m ($mM^{-1}s^{-1}$)	-	82	-	-	-	-	6.5	43.4
D-Fucose k_{cat}/K_m ($mM^{-1}s^{-1}$)	-	-	-	-	-	-	23.2	131.9
L-Fucose k_{cat}/K_m ($mM^{-1}s^{-1}$)	-	-	-	-	-	67.0	-	-
D-Arabinose k_{cat}/K_m ($mM^{-1}s^{-1}$)	-	-	-	-	-	10.6	-	-

All enzyme assays were performed at 70°C.

In the genome of *S. acidocaldarius* DSM639 Saci_1079 is the only annotated GDH [304]. However, the functional prediction based on BLAST analysis alone is very difficult. As shown in case of the SSO-D-AraDH (SSO1300), which was originally annotated as alcohol dehydrogenase (ADH) IV [27].

Three GDHs from *S. solfataricus* have been characterized in the past (GDH-1 SSO3003, GDH-2 SSO3204, D-AraDH SSO1300) [27,109,113,124]. SSO-GDH-1 was purified from *S. solfataricus* P2 cells grown on D-glucose as C-source. The enzyme shows broad substrate specificity with NAD^+ and $NADP^+$ as cosubstrate, whereas the highest catalytic efficiency was detected with D-xylose in presence of $NADP^+$ (261 $mM^{-1}s^{-1}$) and catalytic efficiencies < 100 $mM^{-1}s^{-1}$ were determined for D-glucose, D-galactose and L-arabinose [25,109,132]. In contrast, SSO-GDH-2 was specific for D-glucose as substrate and a higher catalytic efficiency compared to SSO-GDH-1 was observed in presence of $NADP^+$ (115.2 $mM^{-1}s^{-1}$ compared 50.0 $mM^{-1}s^{-1}$) [113]. For the SSO-D-AraDH an increased expression level in D-arabinose grown cells was shown [27]. The enzyme exhibits, like SSO-GDH-1, broad substrate specificity and prefers $NADP^+$ over NAD^+ as cosubstrate. Enzyme activity was detected in the presence of several pentose and hexose sugar, however physiological relevant activity was only determined with L-galactose, L-fucose, D-ribose and D-arabinose [27,124]. Enzyme characterization was only performed with D-arabinose and L-fucose as substrate, which revealed a 6.7 fold higher catalytic efficiency for L-fucose as substrate (Table 19).

The Saci-GDH is compared to the investigated *S. solfataricus* GDHs the most efficient *Sulfolobus* enzyme in terms of D-xylose utilization (Table 19). In respect to its kinetic properties, Saci-GDH shows the highest similarity to SSO-GDH-1. Both enzymes are capable to convert D-glucose, D-galactose, D-xylose and L-arabinose but exhibit highest catalytic efficiency with D-xylose. As reported for SSO-GDH-1 and SSO-GDH-2 the enzyme of *S. acidocaldarius* possesses dual cosubstrate specificity.

3.5.7. GDH paralogs in *Sulfolobus acidocaldarius* DSM639

In *S. solfataricus* P2 14 members of the MDR branch of the pyridine nucleotide-dependent alcohol/polyol/sugar dehydrogenase superfamily were predicted [124]. Those 14 sequences were used as template to evaluate the number of putative GDHs in *S. acidocaldarius* DSM 639 by BLAST searches (Table 20). BLAST analysis with the reported *S. solfataricus* P2 GDH paralogs revealed the presence of 12 putative members of the MDR branch of the pyridine nucleotide-dependent alcohol/polyol/sugar dehydrogenase superfamily in *S. acidocaldarius*. The highest similarity of Saci-GDH was observed with SSO-GDH-1 (57%), followed by SSO-GDH-2 (38%) and SSO-D-AraDH (29%). The comparison of the amino acid sequences of Saci-GDH with the *S. solfataricus* P2 paralogs SSO-GDH-1 [132] and SSO-D-AraDH [124], where the crystal structure was solved in the past, revealed several differences (Figure 47).

The comparison revealed that the identified residues involved in the binding of the catalytic Zn-ion and of the structural Zn-ion are conserved in SSO-GDH-1 and Saci-GDH, whereas changes occur in SSO-GDH-2 and SSO-D-AraDH (Figure 47) [132]. Furthermore, the residues identified to be involved in the cosubstrate (NAD^+ and $NADP^+$) binding are conserved in SSO-GDH-1 and Saci-GDH and changes occur in SSO-GDH-2 and SSO-D-AraDH. The residue Arg213, which binds to the 3'-phosphate of the nucleotide, is conserved in SSO-GDH-2 and Saci-GDH. In SSO-D-AraDH, which uses only NAD^+ as cosubstrate, is this residue substituted by a Ser residue. It is assumed that the residue Arg213 determines if an enzyme is able to use both pyridine nucleotides or only NAD^+ [132]. Hence, the identified residue in SSO-GDH-1, which interact with the substrates D-glucose and D-xylose, are also conserved in Saci-GDH, whereas changes occur in SSO-GDH-2 and SSO-D-AraDH (Figure 47).

In summary, the important residues identified in SSO-GDH-1 proposed for the Zn-ion, cosubstrate and substrate interaction are completely conserved in Saci-GDH and further supports the observed dual cosubstrate and broad substrate specificity. In comparison to SSO-GDH-2 several differences are evident, which might reflect the narrow substrate specificity of SSO-GDH-2 with strong preference for D-glucose as substrate (for more details see [113]). Furthermore, SSO-D-AraDH shares only some structural similarities with SSO-GDH-1, SSO-GDH-2 and Saci-GDH. These differences could be the reason why SSO-D-AraDH is only capable to use L-fucose and D-arabinose as substrates and NAD^+ as cosubstrate (Table 19).

Table 20 Similarity comparison between the GDH paralogs of *S. solfataricus* P2 and *S. acidocaldarius* DSM 639.
Paralogs which have a identity [%] to each other ≥ 50% are shown in bold.

Identity [%]	Saci_0410	Saci_0557	Saci_0911	Saci_1079	Saci_1115	Saci_1393	Saci_1690	Saci_2057	Saci_2126	Saci_2145	Saci_2205	Saci_2224
GDH-1[107,130] (SSO3003)	-	-	-	**57**	-	-	-	-	-	-	-	-
GDH-2[111] (SSO3204)	-	-	26	38	-	-	-	-	-	-	-	-
D-AraDH[27,122] (SSO1300)	-	33	27	-	28	34	27	39	33	30	29	30
SSO2441	25	**56**	28	29	26	29	28	36	32	33	33	33
SSO1646	28	30	29	-	26	26	28	36	**73**	29	28	32
SSO1220	28	30	24	-	29	31	27	33	32	25	26	30
SSO0472	29	28	30	-	29	31	31	31	-	**63**	-	37
SSO2494	27	-	**62**	-	**70**	30	31	32	30	28	31	278
SSO0764	29	27	**73**	-	**58**	31	30	29	28	27	31	26
SSO2501	28	32	27	-	28	31	30	35	36	38	31	**61**
SSO2878	-	30	33	-	30	29	26	32	28	-	**74**	31
SSO2334	25	27	28	-	27	**65**	**59**	31	32	27	31	32
SSO2800	27	26	28	-	27	**59**	**79**	35	32	29	-	-
SSO2717	25	-	-	-	-	-	-	-	-	-	30	-
SSO3237	**71**	25	25	-	-	-	28	29	29	29	-	26

3.5.8. Phosphorylation of GDH paralogs in *S. solfataricus* P2 and *S. acidocaldarius* DSM 639

As discussed in section 3.1 and 3.2 the p-proteomes of both *Sulfolobus* species were determined and several enzymes of the CCM have been identified to be targeted by reversible protein phosphorylation in *S. solfataricus* and *S. acidocaldarius* [188,218], as already predicted in the past by Snijders *et al.,* for *S. solfataricus* [26]. In *S. solfataricus* five GDH paralogs were identified as p-proteins and of them four (SSO0472, SSO2441, SSO-D-AraDH, SSO-GDH-1) were phosphorylated on one residue, whereas GDH-2 was found to be phosphorylated on three residues (Table 21).

Table 21 Identified *Sulfolobus* GDHs paralogs which are targeted by protein phosphorylation.
* = only identified in one biological replicate, # = only one proteome study was performed.

ORF	P-peptide	P-site	Growth condition & strain
SSO0472	-IIAVGpTRK-	Thr^{224}	Tryptone, P2
D-AraDH (SSO1300)	-IKPpYIIK-	Tyr^{320}	D-Glucose, P2
SSO2441	-pYEIKVELPK-	Tyr^{48}	D-Glucose, P2
GDH-1 (SSO3003)	-pYLVKIPL-	Tyr^{134}	D-Glucose & Tryptone; P2
GDH-2 (SSO3204)	-pYPpSVLNRMIpTR-	Tyr^{324}, Ser^{326}, Thr^{333}	Tryptone*, P2
Saci_1079	-pYLVKIPK-	Tyr^{132}	OD 1.2#, MW001
Saci_1115	-LGpYLVLK-	Tyr^{271}	OD 1.2#, ΔPTP
Saci_1393	-VpYNLEDGK-	Tyr^{302}	OD 1.2#, ΔPTP
Saci_2205	-FDAVFELVGpSK-	Ser^{213}	OD 0.8, MW001

In *S. acidocaldarius* DSM 639 four GDH paralogs were identified as p-proteins in two different studies (Saci_1079 (MW001), Saci_1115 (ΔPTP) and Saci_1393 (ΔPTP)) in *S. acidocaldarius* cells harvested at OD 1.2 (stationary growth phase, preliminary results which are not part of this study) and all three proteins were modified on Tyr. Saci_2205 was phosphorylated on Ser in MW001 harvested at OD 0.8 (exponential growth phase, chapter 3.2.). The detected p-peptides and phosphorylation sites of SSO-GDH-1 and Saci-GDH were identical (Figure 47), which suggest a similar, although unknown, effect on both GDHs,.The phosphorylation sites in the other GDHs were not conserved. In case of SSO0472 a Thr residue was phosphorylated, which is proposed to interact with the adenyl phosphate group from the cosubstrate. For all the other phosphorylation sites no functional prediction was possible.

In summary, Saci-GDH possesses the highest similarity to SSO-GDH-1, which is also reflected by its broad substrate specificity and kinetic properties. Strikingly, even the detected phosphorylation site seems to be conserved in both GDHs, which might point towards an important role of protein phosphorylation in the regulation of GDHs.

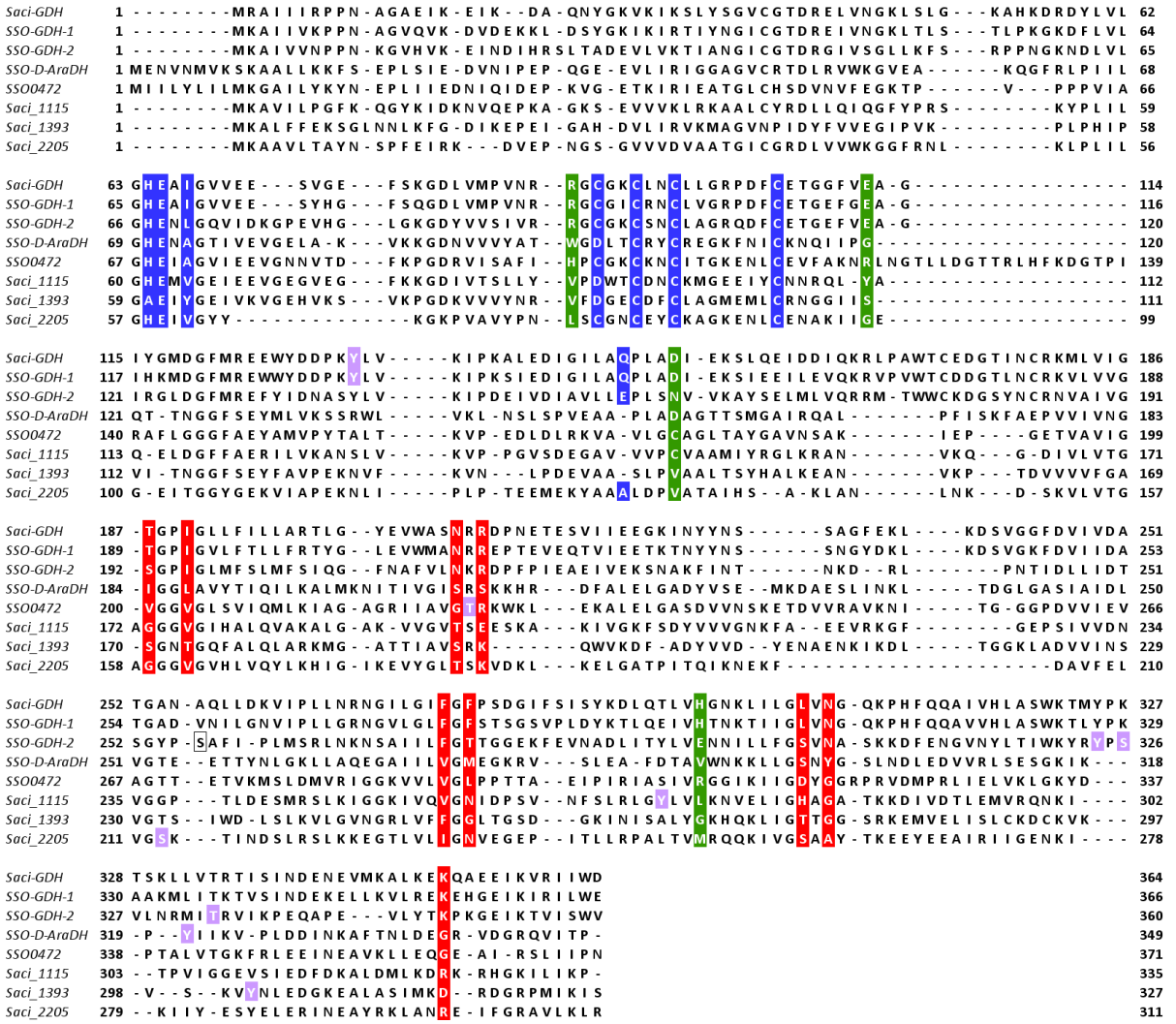


Figure 47 Multiple sequence alignment of GDHs from *S. acidocaldarius* DSM 639 and *S. solfataricus* P2 identified as p-proteins

Amino acid functions derived from the GDH-1 (SSO3003) crystal structure are indicated by: blue = catalytic Zn binding, red = cosubstrate binding pocket, orange = adenyl phosphate pocket, green = substrate binding site. P-sites on Ser, Thr or Tyr are indicated in purple [124,132].

3.6. Characterization of the triose-3-phosphate isomerase from Sulfolobus solfataricus P2 in response to temperature change

The biochemical characterization of the *S. solfataricus* triose-3-phosphate isomerase at 70°C for the catabolic and anabolic reaction are part of the publication "Intermediate instability at high temperature leads to low pathway efficiency for *in vitro* reconstituted gluconeogenesis of *Sulfolobus solfataricus*" submitted in FEBS Journal and currently under revision [119].

3.6.1. Heterologoues expression in *E. coli* ROSETTA (DE3) and purification

The TPI was expressed in *E. coli* ROSETTA (DE3) via induction by IPTG. After cell disruption the enzyme was found in the soluble fraction and was purified to apparent homogeneity using heat precipitation (90°C), anion exchange chromatography and size exclusion chromatography. After purification and separation via SDS-PAGE one homogeneous band at approximately 25 kDa was visible (Figure 48), which agrees well with the calculated theoretical molecular weight of 24.7 kDa. From 7.2 g recombinant cells (wet weight) 9.5 mg of purified TPI were gained. After size exclusion chromatography a native molecular weight of 72.3 kDa was determined suggesting an unusual homotrimeric structure.

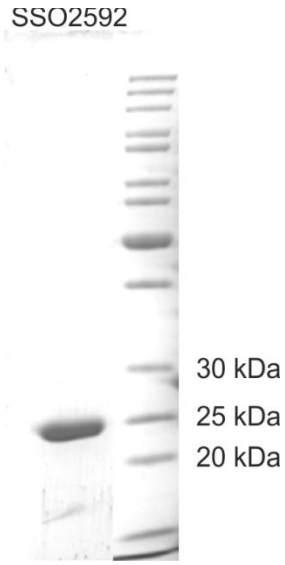


Figure 48 Purification of the recombinant triose-3-phosphate isomerase SSO2592 from *S. solfataricus* P2.
2 µg SSO2592 after purification via heat precipitation, ion exchange chromatography and size exclusion chromatography were applied on SDS-PAGE (12.5%). The gel was stained with Coomassie Blue.

3.6.2. Enzymatic characterization of TPI

TPI catalyzes the interconversion of DHAP, and GAP. TPI activity with DHAP was assayed at 60°C, 65°C, 70°C and 80°C in presence of GAPN (SSO3194) from *S. solfataricus* P2 as auxiliary enzyme using 5 mM $NADP^+$ as cosubstrate. All reactions were performed with purified recombinant SSO-TPI in 100 mM Tris/HCl buffer (pH 6.5 at 70°C). Enzyme activity tests in 100 mM HEPES/KOH buffer (pH 6.5 at 70°C) resulted in 50% decreased TPI activity.

TPI followed classical Michaelis-Menten kinetics at 60°C, 65°C and 70°C (Figure 49 B-D, Table 22), whereas cooperativity was observed at 80°C (sigmoidal curve, Figure 49 E). The Hanes Woolf plot at 80°C revealed no linear correlation at low DHAP concentrations (Figure 49 F).

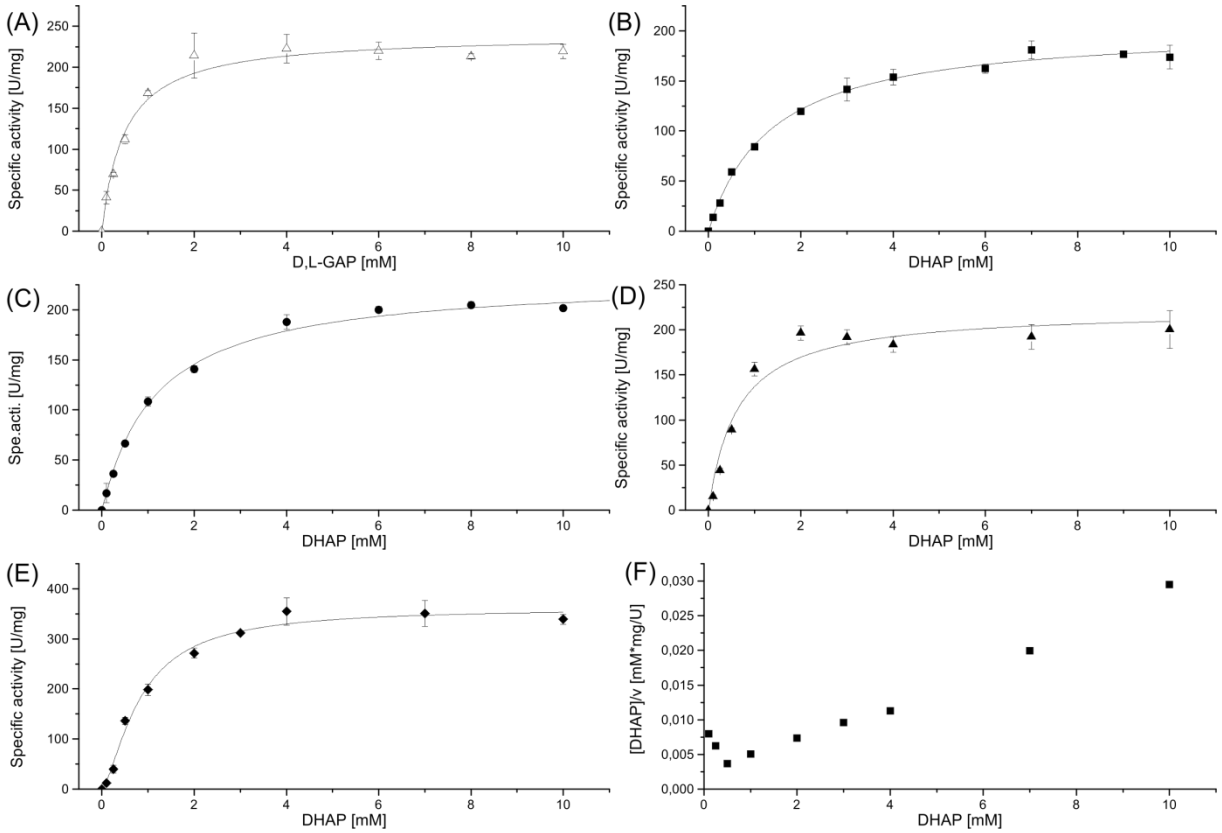


Figure 49 Effect of temperature on the kinetic properties of SSO-TPI.
Activity was determined in a continuous assay at 70°C by varying the D,L-GAP concentration (A) or at 60°C (B), 65°C (C), 70°C (D) and 80°C (E) by varying the DHAP concentration. The Hanes Woolf plot at 80°C by varying the DHAP concentration is depicted in (F). NADPH formation was monitored at 340 nm. Three independent measurements were performed. Mean-values are depicted and the standard deviation is given.

Therefore the apparent k_m-values at 80°C was calculated by iterative curve-fitting using the Hill equation. The Hill equation considers allosteric regulation and the Hill-coefficient gives an approximation for the amount of active sites present in the enzyme. At 80°C a Hill-coefficient of 2.3 +/- 0.5 was determined indicating a positive cooperativity.

The catalytic efficiency (k_{cat}/K_m) increases with temperature and is 3-fold higher at 80°C compared to 60°C. There is no significant difference in k_{cat} between 60° and 70°C, but the affinity for DHAP is 2.3-fold higher at 70°C compared to 60°C and thus the catalytic efficiency at 70°C increases. The affinity for DHAP at 80°C is slightly reduced compared to 70°C but this is compensated by the enhanced reaction velocity V_{max} at 80°C.

Table 22 Kinetic parameters of the TPI from *S. solfataricus* for its substrates DHAP and D,L-GAP.

Substrate	Temperature	V_{max}	K_m	k_{cat}	k_{cat}/K_m
	[°C]	[U/mg]	[mM]	[s^{-1}]	[$mM^{-1}s^{-1}$]
DHAP	60	200.7	1.3	83.1	62.7
DHAP	65	235.8	1.2	97.2	78.4
DHAP	70	195.8	0.6	80.6	141.6
DHAP	80[(1)]	361.6	0.8	149.2	188.8
D,L-GAP	70	227.8	0.4	93.9	240.9

[(1)] The Hill equation was used for calculation: $v=V_{max}*x^n/(k^n+x^n)$.

The kinetic parameters for the interconversion of D,L-GAP to DHAP were assayed in a continuous assay using FBPA/ase (SSO0286) from *S. solfataricus* P2, glucose-6-phosphate isomerase (PGI) from *T. tenax* and glucose-6-phosphate dehydrogenase (G6PDH) from *T. maritima* as auxiliary enzymes in the presence of 1 mM $NADP^+$ as cosubstrate for G6PDH at 70°C (pH 6.5, Figure 49 A, Table 22). TPI follows classical Michaelis-Menten kinetics at 70°C for D,L-GAP. The determined V_{max}-value was similar to the one determined for DHAP at 70°C, whereas the substrate affinity was slightly reduced from K_m = 0.39 mM (D,L-GAP) to K_m = 0.6 mM (DHAP). This resulted in an increased catalytic efficiency for the conversion of D,L-GAP to DHAP (240.9 $mM^{-1}s^{-1}$) compared to DHAP to D,L-GAP (141.6 $mM^{-1}s^{-1}$, Table 22).

For the eukaryotic TPI from pea seed an efficient inhibition by the CCM intermediates PEP (K_i = 0.1 µM), 2-PG (K_i = 6 µM), 3-PG (K_i = 50 µM) and also 2,3-BPG (K_i = 10 µM) was reported. Also for metabolites a significant inhibition at higher concentrations (5mM) was observed (6-phospho-gluconate 9%, G1P 10%, G6P 25%, F6P 44%, FBP 69%, citrate 25%, succinate 9%, ATP 39%). Furthermore, several anions resulted in a decreased activity (e.g. phosphate, Cl^-) [130]. Similar inhibition was also reported for the TPI of the parasitic protist *Trypanosoma brucei*. The *T. bruceri* TPI was inhibited by the metabolites D-GAP, DHAP, 2-phosphoglycolate, D,L-glycerol-1-phosphate, PEP, D,L-3-PG, L-3-PG and D,L-2-PG and also by the anions sulphate, phosphate and arsenate [167].

Notably, so far for none of the investigated archaeal TPIs inhibition by CCM intermediates was analyzed and reported. The catabolic activity of the SSO-TPI was determined in the presence of fructose-1-phosphate (F1P), G1P, ATP, FBP, 3-PG and PEP. Only in presence of 3-PG and PEP a decrease of the enzyme activity was observed. The other tested metabolites had not effect on TPI activity. For 3-PG a K_i-value of 0.4 mM was calculated and for PEP a K_i-value of 0.66 mM (calculated

by Prof. Dr. J. L. Snoep) [119]. Inhibtion studies for the anabolic direction (GAP to DHAP) were not possible due to the precipitation of the FBPA/ase serving as auxiliary enzyme.

3.6.3. Triosephosphate isomerase from *S. solfataricus* P2

TPI catalyzes the interconversion of DHAP and GAP in the EMP pathway [163] and is of special interest as both substrates are instable at high temperature. In *S. solfataricus*, due to the absence of PFK, the EMP-pathway is solely used for gluconeogenesis thus TPI serves as gluconeogenic enzyme and is not involved in glycolysis, which proceeds via the branched ED-pathway (Figure 4). The reaction catalyzed by TPI represents an important step in gluconeogenesis, which might be important for regulation of fluxes. The genome of *S. solfataricus* P2 harbors only one gene (SSO2592) encoding TPI.

Kinetic properties of the SSO-TPI were determined for the conversion of DHAP to D,L-GAP in a continuous assay at 60°C, 65°C, 70°C and 80°C and for the conversion of D,L-GAP to DHAP at 70°C.

As discussed above, the catalytic efficiency increased with temperature and is 3-fold higher at 80°C compared to 60°C (Table 22). Since 80°C is the optimal growth temperature for *S. solfataricus* P2, the enzyme seems to be perfectly adapted to the optimal growth temperature of the organism. However, the catalytic efficiency of SSO-TPI is relatively low compared to TPIs of other (hyper)thermophilic archaea (Table 23). Kinetic parameters of the catabolic direction at 70°C were reported for the TPI from *T. tenax* (Ttx, optimal growth temperature 86°C) [164], *Pyrococcus woesei* (Pw, optimal growth temperature 100°C), *Methanothermus fervidus* (Mf, optimal growth temperature 83°C) [161] and for the hyperthermophilic bacterium *T. maritima* (Tm, optimal growth temperature 80°C).

Table 23 Kinetic parameters for the conversion of DHAP to GAP by TPIs from hyperthermophilic Archaea and Bacteria.

Reported data are shown in bold, the remaining data were calculated on the basis of reported data for the respective temperature. For *T. maritime* values for the isolated TPI (underlined) and TPI-PGK fusion protein are given. References: *T. tenax*, *P. woesei* and *M. fervidus*, *T. maritima*.

Parameter	T [°C]	*S. solfataricus*	*T. tenax* [164]	*P. woesei* [161]	*M. fervidus* [161]	*T. maritime* [305]
Optimal growth temperature		80°C	86°C	100°C	83°C	80°C
K_m [mM]	70	0.6	**0.9**	1.5	2.6	n.d.
	80	0.8	n.d.	n.d.	n.d.	**2.6/ 2.7**
k_{cat} [s^{-1}]	70	80.7	**776.8**	**1,680**	**1,019.3**	n.d.
	80	149.2	n.d.	n.d.	n.d.	**583.3/ 3360.7**
k_{cat}/K_m [$mM^{-1}s^{-1}$]	70	141.6	863.0	**1,120.0**	**400.1**	n.d.
	80	188.8	n.d.	n.d.	n.d.	224.4/ 1244.7

The Pw-TPI has the highest catalytic efficiency for the conversion of DHAP to GAP (catabolic) of all so far characterized (hyper)thermophlic TPIs (Table 23 and Table 24) [161]. The kinetic parameters clearly indicate that in *P. woesei* the TPI has a role in gluconeogenesis (250,000 $mM^{-1}s^{-1}$, GAP to DHAP) rather than in glycolysis (1,120 $mM^{-1}s^{-1}$, DHAP to GAP). The Ttx-TPI shows a similar behavior and for the GAP to DHAP conversion a 21-times higher catalytic efficiency was observed, indicating a gluconeogenic function as well (863.0 $mM^{-1}s^{-1}$ compared to 12,038.3/18,445.8 $mM^{-1}s^{-1}$) [164]. For the SSO-TPI only a slight preference for the anabolic reaction was observed. For the catabolic reaction (DHAP to GAP) a catalytic efficiency of 141.6 $mM^{-1}s^{-1}$ (at 70°C) and for the anabolic reaction (GAP to DHAP) of 240.9 $mM^{-1}s^{-1}$ were determined. Therefore, the SSO-TPI possesses slight preference for the anabolic direction. However, there are some evidence from TPIs from *T. tenax* and *T. maritime* that protein-protein interaction strongly influences the enzyme activity.

Table 24 Kinetic parameters for conversion of GAP to DHAP by TPIs from hyperthermophiles.
Reported data are shown in bold, the remaining data were calculated on the basis of reported data. For *T. tenax* values for TPI assayed with *T. tenax* GLPDH (underlined) with *M. thermoautothrophicus* GLPDH are given. References: *T. tenax* [164], *P. woesei* [161].

Parameter	*S. solfataricus*	*P. woesei*	*T. tenax*
Temperatur [°C]	70	60	70
K_m [mM]	0.39	**0.02**	**0.2/ 0.2**
k_{cat} [s^{-1}]	93.9	3,440.0	2,407.7/ <u>3,689.2</u>
k_{cat}/K_m [$mM^{-1}s^{-1}$]	240.9	250,000.0	12,038.3/ <u>18,445.8</u>

For the Ttx-TPI an equilibrium between inactive dimers and active tetramers was observed in solution, which is shifted towards the active state by interaction with glycerol-1-phosphate dehydrogenase (GLPDH) of *T. tenax.* This is discussed as a regulatory mean in metabolic thermoadaptation for the reduction of intracellular pools of labile intermediates of the EMP pathway under anabolic conditions as GLPDH reduces DHAP to the thermostable glycerol 1-phosphate. The regulation of TPI via protein-protein interaction is also shown for the Tm-TPI. The Tm-TPI represents a bifunctional fusion protein with PGK and kinetic data for the isolated TPI as well as for the fusion protein were reported [305].

For the SSO-TPI a strong inhibition was observed in the presence of HEPES/KOH compared to Tris/HCl buffer. For the native SSO-TPI a molecular weight of 72.33 kDa and a subunit size of 25 kDa were determined, suggesting an unusual trimeric structure in the native state. No other additional peak was observed after size exclusion chromatography, therefore an equilibrium between different oligomeric states as reported for *T. tenax* can be excluded. However, we cannot rule out that higher protein concentrations or other *S. solfataricus* enzymes like SSO-GLPDH might influence the

oligomeric state and thus enzyme activity of the SSO-TPI. In order to address this issue, the enzyme crystallization is currently on the way.

For the TPI from pea seed a strong inhibition by numerous intermediates of the CCM (e.g. PEP, 3-PG) and anions (e.g. Cl^-) was reported. Tomlinson and Turner concluded based on intracellular intermediate concentrations that the TPI in the living Pea seed cells is in general strongly inhibited and thus the high reaction velocity of 12,000 U/mg will not be achieved under *in vivo* conditions [130]. The inhibition of the SSO-TPI in presence of PEP and 3-PG suggests a similar reduced activity under *in vivo* conditions, which might allow the enzyme to adapt a similar reaction velocity as reported for other CCM enzymes (e.g. GAPN V_{max} ~ 5 U/mg) [116].

In summary, the TPI from *S. solfataricus* P2 shows unusual properties compared to other hyperthermophilic TPIs. It is much less efficient compared to all investigated archaeal TPIs and shows a strong inhibition in presence of PEP and 3-PG. The determined trimeric oligomerization state is also unique compared to other bacterial and archaeal homologs and requires further experimental validation. In comparison to other hyperthermophilic TPIs the preference for the gluconeogenetic direction is significantly reduced.

4. Summary

The Crenarchaea *Sulfolobus solfataricus* and *S. acidocaldarius* are among the best investigated Archaea and are well established as thermoacidophilic model organisms. The central carbohydrate metabolism (CCM) of *Sulfolobus* species has been well studied using classical biochemical and modern high throughput approaches such as genomics, proteomics and transcriptomics. However, the regulation of the CCM is still far from being understood. Like other archaea, *Sulfolobus* species lack the "classical" control points known from their eukaryotic or bacterial brethren and the only well established control point is the non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase, which is activated at low levels of the metabolite glucose 1-phosphate. Furthermore, the complexity of the CCM in *Sulfolobus* species as well as other Archaea is still unsolved and there are often several paralogs in the archaeal genomes with unknown function.

In the first part of this thesis the impact of regulation by post translational modifications via reversible protein phosphorylation was analyzed in *S. solfataricus* and *S. acidocaldarius*. Crenarchaea like *Sulfolobus* species lack bacterial type one/two-component systems (phosphorylation on His and Asp) but rely on eukaryal like protein phosphorylation on Ser, Thr and Tyr. In *S. solfataricus* we investigated how the offered carbon source glucose (glycolytic growth) or tryptone (gluconeogenic growth) alter the global phosphorylation pattern. Analysis by precursor aquisition independent from ion count (PAcIFIC) revealed a high number of p-proteins (540) with a so far unique enrichment for p-Tyr (>50%) compared to previous studies in all three domains of life (Table 1). Furthermore, p-proteins were detected in all functional categories indicating a major impact on all cellular processes. In response to the carbon source dramatic changes in the phosphorylation pattern were observed. In the CCM especially proteins at the beginning of various sugar degradation pathways and gluconeogenic EMP and TCA and at branching points to other metabolic routs were targeted by protein phosphorylation.

In *S. acidocaldarius* we used a similar experimental approach which combined transcriptomics, biochemistry and growth studies to investigate the effect of deletion of the two protein phosphatases. The respective genes of the putative protein phosphatases were cloned and the recombinant enzymes in regard to their substrate specificity investigated. The catalytic subunit of the protein Ser/Thr phosphatase (PP2Ac) resembles their eukaryotic counterparts and was inhibited by okadaic acid and requires Mn^{2+}-ions for catalytic activity and showed strict activity towards p-Ser/p-Thr peptides. Whereas the protein tyrosine phosphatase (PTP) revealed dual substrate specificity with p-Tyr and p-Thr peptides. The phenotypic characterization of the three strains (MW001, *Δsaci_pp2a*, *Δsaci_ptp*) revealed that the parant strain MW001 and *Δsaci_ptp* resembled each other, where as the *Δsaci_pp2a* strain showed a differences in growth and cell size variation. The number of

p-proteins in the background strain MW001 was lower (69) compared to *S. solfataricus*, but similar to that reported for other prokaryotic and eukaryotic organisms (Table 1). According to the expectation the number of p-proteins was increased in the two phosphatase deletion strains (*Δsaci_pp2a* (143), *Δsaci_ptp* (407)) and comparable to the number determined in *S. solfataricus*. However, the deletion of the two protein phosphatases not only resulted in an increase of p-proteins, the whole phosphorylation pattern was significant changed. The combination of the three data sets revealed that similar to *S. solfataricus* all cellular processes are targeted by protein phosphorylation. Furthermore, the investigation of the transcriptome, via RNA deep sequencing analysis, revealed a major impact on the transcription of genes of the respiratory chain as well as genes involved in cell motility (e.g. archaellum). These results were further supported by the identification of respective p-proteins and by the finding that the *Δsaci_pp2a* strain showed a different motility behavior compared to the strains MW001 and *Δsaci_ptp*. The CCM showed also a complete different phosphorylation pattern in MW001 compared to the two deletion strains. Especially enzymes of the tricarboxylic acid cycle as well as enzymes located at branching points (e.g. amino acid metabolism) seem to be targeted by protein phosphorylation in the *Δsaci_pp2a* strain.

The high amount of p-Tyr detected in both *Sulfolobus* species let suggest that phosphorylation on Tyr is a crenarchaeal feature. The high growth temperature seems not to be the main reason of the high p-Tyr amount since only a low amount of p-Tyr was identified in the thermophilic gram-negative bacterium *Thermus thermophilus* (grown at 70°C, 10% p-Tyr). Also in the Euryarchaeota *Halobacterium salinarium* only one phosphorylation site on Tyr was identified, which excluds that this is a typical archaeal feature. However, compared to *Sulfolobus* species, *T. thermophilus* and *H. salinarium* also possess His kinases, which might provide alternative possibilities to regulate cellular functions by reversible protein phosphorylation. Therefore, suggesting that the high amount of p-Tyr is a typical feature of the crenarchaeal *Sulfolobus* species.

The second part of this thesis focused on unraveling the complexity and regulation of the CCM in *Sulfolobus* species. From *S. solfataricus* were the triose-3-phosphate isomerase (TPI) and members of the aldehyde dehydrogenase super family characterized and from *S. acidocaldarius* the putative glucose-1-dehydrogenase (GDH). To unravel the function of the *S. solfataricus* TPI in gluconeogenesis kinetic parameters for both triose-3-phosphates glyceraldehydes-3-phosphate and dihydroxyacetone-3-phosphate were determined. The determined catalytic efficiencies were rather small compared to previously characterized archaeal TPIs. The reason is still unclear and we expect further insight from the available crystal structure. The *S. solfataricus* TPI showed inhibition by phosphoenolpyruvate and 3-phosphoglycerate which was not reported for any archaeal TPI so far but is well established for eukaryotic TPIs.

For the two so far uncharacterized members of the aldehyde dehydrogenase super family succinic semialdehyde dehydrogenase activity was demonstrated (SSO1629, SSO1842) and a major role in polyamine and γ-aminobutyric acid metabolism was proposed based on bioinformatic analysis. For SSO1218 methylmalonate semialdehyde dehydrogenase activity demonstrated and a role in the degradation of branched amino acids was proposed.

The characterization of the GDH from *S. acidocaldarius* (Saci_1079), one of the 12 of the medium-chain alcohol/polyol dehydrogenase/reductase branch of the pyridine nucleotide-dependent alcohol/polyol/sugar dehydrogenase superfamily in *S. acidocaldarius*, revealed that this enzyme exhibits broad substrate specificity a broad range of different sugars and might play an important role not only in D-glucose, but also in D-xylose, D-galactose and D-fucose degradation. The enzyme shows highest similarity to GDH-1 (SSO3003) of *S. solfataricus* and resembles the enzyme in respect to its broad substrate specificity, kinetic parameters as well as structural features.

Notably, some of the enzymes (e.g. SSO1842, Saci_1079) have been identified as p-proteins and future studies aim to unravel the impact of phosphorylation on enzyme function in order to resolve the respective signal transduction pathways.

5. Literatur

1. Woese CR, Fox GE (1977) Phylogenetic structure of the prokaryotic domain: The primary kingdoms. Proceedings of the National Academy of Sciences 74: 5088-5090.
2. Albers SV, Meyer BH (2011) The archaeal cell envelope. Nat Rev Micro 9: 414-426.
3. Siebers B, Schönheit P (2005) Unusual pathways and enzymes of central carbohydrate metabolism in Archaea. Current Opinion in Microbiology 8: 695-705.
4. Sato T, Atomi H (2011) Novel metabolic pathways in Archaea. Current Opinion in Microbiology 14: 307-314.
5. Woese CR, Kandler O, Wheelis ML (1990) Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria and Eucarya. Proceedings of the National Academy of Sciences USA 87: 4576-4579.
6. Morris B, Henneberger R, Huber H, Moissl-Eichinger C (2013) Microbial syntrophy: interaction for the common good. FEMS Microbiology Reviews.
7. Podar M, Makarova KS, Graham DE, Wolf YI, Koonin EV, et al. (2013) Insights into archaeal evolution and symbiosis from the genomes of a nanoarchaeon and its inferred crenarchaeal host from Obsidian Pool, Yellowstone National Park. Biology direct 8: 9.
8. Barns SM, Delwiche CF, Palmer JD, Pace NR (1996) Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences. Proceedings of the National Academy of Sciences 93: 9188-9193.
9. Elkins JG, Podar M, Graham DE, Makarova KS, Wolf Y, et al. (2008) A korarchaeal genome reveals insights into the evolution of the Archaea. Proceedings of the National Academy of Sciences 105: 8102-8107.
10. Schleper C, Nicol GW (2010) Ammonia-oxidising archaea—physiology, ecology and evolution. Advances in microbial physiology 57: 3.
11. Nunoura T, Takaki Y, Kakuta J, Nishi S, Sugahara J, et al. (2011) Insights into the evolution of Archaea and eukaryotic protein modifier systems revealed by the genome of a novel archaeal group. Nucleic Acids Research 39: 3204-3223.
12. Brochier-Armanet C, Forterre P, Gribaldo S (2011) Phylogeny and evolution of the Archaea: one hundred genomes later. Current Opinion in Microbiology 14: 274-281.
13. Albers SV, Birkeland NK, Driessen AJM, Gertig S, Haferkamp P, et al. (2009) SulfoSYS (Sulfolobus Systems Biology): Towards a silicon cell model for the central carbohydrate metabolism of the archaeon *Sulfolobus solfataricus* under temperature variation. Biochemical Society Transactions 37: 58-64.
14. Zillig W, Stetter KO, Wunderl S, Schulz W, Priess H, et al. (1980) The Sulfolobus-"Caldariella" group: Taxonomy on the basis of the structure of DNA-dependent RNA polymerases. Arch Microbiol 125: 259-269.
15. Grogan DW (1989) Phenotypic characterization of the archaebacterial genus Sulfolobus: comparison of five wild-type strains. J Bacteriol 171: 6710-6719.
16. She Q, Singh RK, Confalonieri F, Zivanovic Y, Allard G, et al. (2001) The complete genome of the crenarchaeon Sulfolobus solfataricus P2. Proc Natl Acad Sci U S A 98: 7835-7840.
17. Albers SV, Driessen AJM (2008) Conditions for gene disruption by homologous recombination of exogenous DNA into the Sulfolobus solfataricus genome. Archaea 2: 145-149.
18. Albers SV, Jonuscheit M, Dinkelaker S, Urich T, Kletzin A, et al. (2006) Production of recombinant and tagged proteins in the hyperthermophilic archaeon Sulfolobus solfataricus. Applied and Environmental Microbiology 72: 102-111.
19. Jonuscheit M, Martusewitsch E, Stedman K, 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.

20. Wagner M, Berkner S, Ajon M, Driessen AJM, Lipps G, et al. (2009) Expanding and understanding the genetic toolbox of the hyperthermophilic genus Sulfolobus. Biochemical Society Transactions 37: 97-101.
21. 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.
22. Egorova K, Antranikian G (2005) Industrial relevance of thermophilic Archaea. Current Opinion in Microbiology 8: 649-655.
23. Unsworth LD, Van Der Oost J, Koutsopoulos S (2007) Hyperthermophilic enzymes – stability, activity and implementation strategies for high temperature applications. FEBS Journal 274: 4044-4056.
24. Lamble HJ, Milburn CC, Taylor GL, Hough DW, Danson MJ (2004) Gluconate dehydratase from the promiscuous Entner-Doudoroff pathway in *Sulfolobus solfataricus*. FEBS Letters 576: 133-136.
25. Nunn CEM, Johnsen U, Schönheit P, Fuhrer T, Sauer U, et al. (2010) Metabolism of pentose sugars in the hyperthermophilic archaea *Sulfolobus solfataricus* and *Sulfolobus acidocaldarius*. Journal of Biological Chemistry 285: 33701-33709.
26. Snijders APL, Walther J, Peter S, Kinnman I, De Vos MGJ, et al. (2006) Reconstruction of central carbon metabolism in *Sulfolobus solfataricus* using a two-dimensional gel electrophoresis map, stable isotope labelling and DNA microarray analysis. PROTEOMICS 6: 1518-1529.
27. Brouns SJJ, Walther J, Snijders APL, Van De Werken HJG, Willemen HLDM, et al. (2006) Identification of the missing links in prokaryotic pentose oxidation pathways: Evidence for enzyme recruitment. Journal of Biological Chemistry 281: 27378-27388.
28. Ahmed H, Ettema TJG, Tjaden B, Geerling ACM, Van Der Oost J, et al. (2005) The semi-phosphorylative Entner-Doudoroff pathway in hyperthermophilic archaea: A re-evaluation. Biochemical Journal 390: 529-540.
29. Wagner M, Berkner S, Ajon M, Driessen AJM, Lipps G, et al. (2009) Expanding and understanding the genetic toolbox of the hyperthermophilic genus Sulfolobus. Biochemical Society Transactions 37: 97-101.
30. Wagner M, van Wolferen M, Wagner A, Lassak K, Meyer BH, et al. (2012) Versatile genetic tool box for the crenarchaeote *Sulfolobus acidocaldarius*. Frontiers in Microbiology 3.
31. Grogan DW (1989) Phenotypic characterization of the archaebacterial genus Sulfolobus: Comparison of five wild-type strains. Journal of Bacteriology 171: 6710-6719.
32. Eichler J, Adams MWW (2005) Posttranslational protein modification in Archaea. Microbiology and Molecular Biology Reviews 69: 696.
33. Kennelly PJ (2003) Archaeal protein kinases and protein phosphatases: Insights from genomics and biochemistry. Biochemical Journal 370: 373-389.
34. Krebs EG, Fischer EH (1956) The phosphorylase b to a converting enzyme of rabbit skeletal muscle. Biochimica et biophysica acta 20: 150-157.
35. Wang JY, Koshland Jr DE (1978) Evidence for protein kinase activities in the prokaryote *Salmonella typhimurium*. Journal of Biological Chemistry 253: 7605-7608.
36. Garnak M, Reeves HC (1979) Phosphorylation of Isocitrate dehydrogenase of *Escherichia coli*. Science (New York, NY) 203: 1111.
37. Spudich JL, Stoeckenius W (1980) Light-regulated retinal-dependent reversible phosphorylation of Halobacterium proteins. Journal of Biological Chemistry 255: 5501-5503.
38. Rudolph J, Oesterhelt D (1995) Chemotaxis and phototaxis require a CheA histidine kinase in the archaeon *Halobacterium salinarium*. EMBO Journal 14: 667-673.
39. Falke JJ, Bass RB, Butler SL, Chervitz SA, Danielson MA (1997) The two-component signaling pathway of bacterial chemotaxis: A molecular view of signal transduction by receptors, kinases, and adaptation enzymes. Annual Review of Cell and Developmental Biology. pp. 457-512.
40. Skorko R (1984) Protein phosphorylation in the archaebacterium Sulfolobus acidocaldarius. European Journal of Biochemistry 145: 617-622.

41. Skorko R (1989) Polyphosphate as a source of phosphoryl group in protein modification in the archaebacterium *Sulfolobus acidocaldarius*. Biochimie 71: 1089-1093.
42. Swanson RV, Sanna MG, Simon MI (1996) Thermostable chemotaxis proteins from the hyperthermophilic bacterium *Thermotoga maritima*. Journal of Bacteriology 178: 484-489.
43. Pearlman SM, Serber Z, Ferrell JE (2011) A mechanism for the evolution of phosphorylation sites. Cell 147: 934-946.
44. Cooper JA, Sefton BM, Hunter T (1983) Detection and quantification of phosphotyrosine in proteins. Methods in enzymology 99: 387.
45. Thorsness PE, Koshland DE (1987) Inactivation of isocitrate dehydrogenase by phosphorylation is mediated by the negative charge of the phosphate. Journal of Biological Chemistry 262: 10422-10425.
46. Aivaliotis M, Macek B, Gnad F, Reichelt P, Mann M, et al. (2009) Ser/Thr/Tyr protein phosphorylation in the archaeon *Halobacterium salinarum* - A representative of the third domain of life. PLoS ONE 4.
47. Chu HM, Wang AHJ (2007) Enzyme-substrate interactions revealed by the crystal structures of the archaeal Sulfolobus PTP-fold phosphatase and its phosphopeptide complexes. Proteins: Structure, Function and Genetics 66: 996-1003.
48. Leng J, Cameron AJM, Buckel S, Kennelly PJ (1995) Isolation and cloning of a protein-serine/threonine phosphatase from an Archaeon. Journal of Bacteriology 177: 6510-6517.
49. Haile JD, Kennelly PJ (2011) The activity of an ancient atypical protein kinase is stimulated by ADP-ribose in vitro. Archives of Biochemistry and Biophysics 511: 56-63.
50. Lower BH, Bischoff KM, Kennelly PJ (2000) The archaeon *Sulfolobus solfataricus* contains a membrane-associated protein kinase activity that preferentially phosphorylates threonine residues *in vitro*. Journal of Bacteriology 182: 3452-3459.
51. Lower BH, Kennelly PJ (2003) Open reading frame sso2387 from the archaeon *Sulfolobus solfataricus* encodes a polypeptide with protein-serine kinase activity. Journal of Bacteriology 185: 3436-3445.
52. Lower BH, Potters MB, Kennelly PJ (2004) A Phosphoprotein from the Archaeon *Sulfolobus solfataricus* with Protein-Serine/Threonine Kinase Activity. Journal of Bacteriology 186: 463-472.
53. Lower BH, Kennelly PJ (2002) The membrane-associated protein-serine/threonine kinase from *Sulfolobus solfataricus* is a glycoprotein. Journal of Bacteriology 184: 2614-2619.
54. Reimann J, Lassak K, Khadouma S, Ettema TJG, Yang N, et al. (2012) Regulation of archaella expression by the FHA and von Willebrand domain-containing proteins ArnA and ArnB in *Sulfolobus acidocaldarius*. Molecular Microbiology.
55. Wang B, Yang S, Zhang L, He ZG (2010) Archaeal eukaryote-like serine/threonine protein kinase interacts with and phosphorylates a forkhead-associated-domain-containing protein. Journal of Bacteriology 192: 1956-1964.
56. (!!! INVALID CITATION !!!).
57. Kim S, Lee SB (2005) Identification and characterization of *Sulfolobus solfataricus* D-gluconate dehydratase: A key enzyme in the non-phosphorylated Entner-Doudoroff pathway. Biochemical Journal 387: 271-280.
58. Potters MB, Solow BT, Bischoff KM, Graham DE, Lower BH, et al. (2003) Phosphoprotein with phosphoglycerate mutase activity from the archaeon *Sulfolobus solfataricus*. Journal of Bacteriology 185: 2112-2121.
59. Jung JH, Lee SB (2005) Identification and characterization of *Thermoplasma acidophilum* 2-keto-3-deoxy-D-gluconate kinase: A new class of sugar kinases. Biotechnology and Bioprocess Engineering 10: 535-539.
60. Kaufmann H, Bailey JE, Fussenegger M (2001) Use of antibodies for detection of phosphorylated proteins separated by two-dimensional gel electrophoresis. Proteomics 1: 194-199.
61. Su HC, Hutchison CA, Giddings MC (2007) Mapping phosphoproteins in *Mycoplasma genitalium* and *Mycoplasma pneumoniae*. BMC microbiology 7: 63.
62. Reinders J, Sickmann A (2005) State of the art in phosphoproteomics. Proteomics 5: 4052-4061.

63. Clarke SJ, Khaliulin I, Das M, Parker JE, Heesom KJ, et al. (2008) Inhibition of mitochondrial permeability transition pore opening by ischemic preconditioning is probably mediated by reduction of oxidative stress rather than mitochondrial protein phosphorylation. Circulation research 102: 1082-1090.
64. Eymann C, Becher D, Bernhardt J, Gronau K, Klutzny A, et al. (2007) Dynamics of protein phosphorylation on Ser/Thr/Tyr in *Bacillus subtilis*. Proteomics 7: 3509-3526.
65. Ikeguchi Y, Nakamura H (1997) Determination of organic phosphates by column-switching high performance anion-exchange chromatography using on-line preconcentration on titania. Anal Sci 13.
66. Koizumi Y, Taya M (2002) Kinetic evaluation of biocidal activity of titanium dioxide against phage MS2 considering interaction between the phage and photocatalyst particles. Biochemical Engineering Journal 12: 107-116.
67. Tichy A, Salovska B, Rehulka P, Klimentova J, Vavrova J, et al. (2011) Phosphoproteomics: searching for a needle in a haystack. Journal of proteomics 74: 2786-2797.
68. Blackwell JA, Carr PW (1991) Fluoride-modified zirconium oxide as a biocompatible stationary phase for high-performance liquid chromatography. Journal of Chromatography A 549: 59-75.
69. Bendt AK, Burkovski A, Schaffer S, Bott M, Farwick M, et al. (2003) Towards a phosphoproteome map of *Corynebacterium glutamicum*. Proteomics 3: 1637-1646.
70. Macek B, Mijakovic I, Olsen JV, Gnad F, Kumar C, et al. (2007) The serine/threonine/tyrosine phosphoproteome of the model bacterium *Bacillus subtilis*. Molecular and Cellular Proteomics 6: 697-707.
71. Macek B, Gnad F, Soufi B, Kumar C, Olsen JV, et al. (2008) Phosphoproteome analysis of *E. coli* reveals evolutionary conservation of bacterial Ser/Thr/Tyr phosphorylation. Molecular and Cellular Proteomics 7: 299-307.
72. Soufi B, Gnad F, Jensen PR, Petranovic D, Mann M, et al. (2008) The Ser/Thr/Tyr phosphoproteome of *Lactococcus lactis* IL1403 reveals multiply phosphorylated proteins. Proteomics 8: 3486-3493.
73. Treeck M, Sanders JL, Elias JE, Boothroyd JC (2011) The Phosphoproteomes of *Plasmodium falciparum* and *Toxoplasma gondii* Reveal Unusual Adaptations Within and Beyond the Parasites' Boundaries. Cell host & microbe 10: 410-419.
74. Yeh YM, Huang KY, Richie Gan RC, Huang HD, Wang TCV, et al. (2012) Phosphoproteome profiling of the sexually transmitted pathogen *Trichomonas vaginalis*. Journal of Microbiology, Immunology and Infection.
75. Schmidl SR, Gronau K, Pietack N, Hecker M, Becher D, et al. (2010) The Phosphoproteome of the Minimal Bacterium *Mycoplasma pneumoniae* Analysis of the Complete known Ser/Thr Kinome suggests the Existence of Novel Kinases Molecular & Cellular Proteomics 9: 1228-1242.
76. Takahata Y, Inoue M, Kim K, Iio Y, Miyamoto M, et al. (2012) Close proximity of phosphorylation sites to ligand in the phosphoproteome of the extreme thermophile *Thermus thermophilus* HB8. Proteomics 12: 1414-1430.
77. Laporte DC (1993) The isocitrate dehydrogenase phosphorylation cycle: regulation and enzymology. Journal of cellular biochemistry 51: 14-18.
78. Leonard CJ, Aravind L, Koonin EV (1998) Novel families of putative protein kinases in bacteria and archaea: Evolution of the 'eukaryotic' protein kinase superfamily. Genome Research 8: 1038-1047.
79. Taylor SS, Kornev AP (2011) Protein kinases: Evolution of dynamic regulatory proteins. Trends in Biochemical Sciences 36: 65-77.
80. Hanks SK (2003) Genomic analysis of the eukaryotic protein kinase superfamily: a perspective. Genome Biol 4: 111.
81. LaRonde-LeBlanc N, Guszczynski T, Copeland T, Wlodawer A (2005) Autophosphorylation of *Archaeoglobus fulgidus* Rio2 and crystal structures of its nucleotide-metal ion complexes. FEBS Journal 272: 2800-2810.

82. LaRonde-LeBlanc N, Guszczynski T, Copeland T, Wlodawer A (2005) Structure and activity of the atypical serine kinase Rio1. FEBS Journal 272: 3698-3713.
83. Laronde-Leblanc N, Wlodawer A (2004) Crystal structure of *A. fulgidus* Rio2 defines a new family of serine protein kinases. Structure 12: 1585-1594.
84. Humbard MA, Reuter CJ, Zuobi-Hasona K, Zhou G, Maupin-Furlow JA (2010) Phosphorylation and Methylation of Proteasomal Proteins of the Haloarcheon *Haloferax volcanii*. Archaea 2010.
85. LaRonde-LeBlanc N, Wlodawer A (2005) A family portrait of the RIO kinases. Journal of Biological Chemistry 280: 37297-37300.
86. Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S (2002) The protein kinase complement of the human genome. Science 298: 1912-1934.
87. Angermayr M, Roidl A, Bandlow W (2002) Yeast Rio1p is the founding member of a novel subfamily of protein serine kinases involved in the control of cell cycle progression. Molecular Microbiology 44: 309-324.
88. Geerlings TH, Faber AW, Bister MD, Vos JC, Raué HA (2003) Rio2p, an evolutionarily conserved, low abundant protein kinase essential for processing of 20 S pre-rRna in *Saccharomyces cerevisiae*. Journal of Biological Chemistry 278: 22537-22545.
89. Vanrobays E, Gleizes PE, Bousquet-Antonelli C, Noaillac-Depeyre J, Caizergues-Ferrer M, et al. (2001) Processing of 20S pre-rRNA to 18S ribosomal RNA in yeast requires Rp10p, an essential non-ribosomal cytoplasmic protein. EMBO Journal 20: 4204-4213.
90. Vanrobays E, Gelugne JP, Gleizes PE, Caizergues-Ferrer M (2003) Late cytoplasmic maturation of the small ribosomal subunit requires RIO proteins in *Saccharomyces cerevisiae*. Molecular and Cellular Biology 23: 2083-2095.
91. Giaever G, Chu AM, Ni L, Connelly C, Riles L, et al. (2002) Functional profiling of the *Saccharomyces cerevisiae* genome. Nature 418: 387-391.
92. Schäfer T, Strauß D, Petfalski E, Tollervey D, Hurt E (2003) The path from nucleolar 90S to cytoplasmic 40S pre-ribosomes. EMBO Journal 22: 1370-1380.
93. Shi Y (2009) Serine/threonine phosphatases: mechanism through structure. Cell 139: 468.
94. Sadatomi D, Tanimura S, Ozaki K, Takeda K (2013) Atypical Protein Phosphatases: Emerging Players in Cellular Signaling. International Journal of Molecular Sciences 14: 4596-4612.
95. Barford D (1996) Molecular mechanisms of the protein serine/threonine phosphatases. Trends in Biochemical Sciences 21: 407-412.
96. Cohen PTW (1997) Novel protein serine/threonine phosphatases: Variety is the spice of life. Trends in Biochemical Sciences 22: 245-251.
97. Kennelly PJ (2002) Protein kinases and protein phosphatases in prokaryotes: A genomic perspective. FEMS Microbiology Letters 206: 1-8.
98. Virshup DM, Shenolikar S (2009) From promiscuity to precision: protein phosphatases get a makeover. Molecular cell 33: 537.
99. Shi L, Potts M, Kennelly PJ (1998) The serine, threonine, and/or tyrosine-specific protein kinases and protein phosphatases of prokaryotic organisms: a family portrait. FEMS Microbiology Reviews 22: 229-253.
100. Kennelly PJ, Oxenrider KA, Leng J, Cantwell JS, Zhao N (1993) Identification of a serine/threonine-specific protein phosphatase from the archaebacterium *Sulfolobus solfataricus*. Journal of Biological Chemistry 268: 6505-6510.
101. Oxenrider KA, Kennelly PJ (1993) A protein-serine phosphatase from the halophilic archaeon *Haloferax volcanii*. Biochemical and Biophysical Research Communications 194: 1330-1335.
102. Oxenrider KA, Rasche ME, Thorsteinsson MV, Kennelly PJ (1993) Inhibition of an archaeal protein phosphatase activity by okadaic acid, microcystin-LR, or calyculin A. FEBS Letters 331: 291-295.
103. Solow B, Young JC, Kennelly PJ (1997) Gene cloning and expression and characterization of a toxin-sensitive protein phosphatase from the Methanogenic archaeon *Methanosarcina thermophila* TM-1. Journal of Bacteriology 179: 5072-5075.

104. Mai B, Frey G, Swanson RV, Mathur EJ, Stetter KO (1998) Molecular cloning and functional expression of a protein- serine/threonine phosphatase from the hyperthermophilic archaeon *Pyrodictium abyssi* TAG11. Journal of Bacteriology 180: 4030-4035.
105. Dahche H, Abdullah AS, Ben Potters M, Kennelly PJ (2009) A PPM-family protein phosphatase from the thermoacidophile *Thermoplasma volcanium* hydrolyzes protein-bound phosphotyrosine. Extremophiles 13: 371-377.
106. Jeon SJ, Fujiwara S, Takagi M, Tanaka T, Imanaka T (2002) Tk-PTP, protein tyrosine/serine phosphatase from hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1: Enzymatic characteristics and identification of its substrate proteins. Biochemical and Biophysical Research Communications 295: 508-514.
107. Siebers B, Schonheit P (2005) Unusual pathways and enzymes of central carbohydrate metabolism in Archaea. Curr Opin Microbiol 8: 695-705.
108. Zaparty M, Siebers B (2011) Reconstruction of the central carbon metabolic network of thermoacidophilic Archaea In: Horikoshi KA, G.; Bull, A.T.; Robb, F.T.; Stetter, K. O., editor. Extremophiles Handbook. 1st ed: Spriner
109. Lamble HJ, Heyer NI, Bull SD, Hough DW, Danson MJ (2003) Metabolic Pathway Promiscuity in the Archaeon *Sulfolobus solfataricus* Revealed by Studies on Glucose Dehydrogenase and 2-Keto-3-deoxygluconate Aldolase. Journal of Biological Chemistry 278: 34066-34072.
110. Lamble HJ, Theodossis A, Milburn CC, Taylor GL, Bull SD, et al. (2005) Promiscuity in the part-phosphorylative Entner-Doudoroff pathway of the archaeon Sulfolobus solfataricus. FEBS Letters 579: 6865-6869.
111. Kim S, Lee SB (2006) Characterization of *Sulfolobus solfataricus* 2-keto-3-deoxy-D-gluconate kinase in the modified Entner-Doudoroff pathway. Bioscience, Biotechnology and Biochemistry 70: 1308-1316.
112. Ahmed H, Ettema TJ, Tjaden B, Geerling AC, van der Oost J, et al. (2005) The semi-phosphorylative Entner-Doudoroff pathway in hyperthermophilic archaea: a re-evaluation. Biochem J 390: 529-540.
113. Haferkamp P, Kutschki S, Treichel J, Hemeda H, Sewczyk K, et al. (2011) An additional glucose dehydrogenase from *Sulfolobus solfataricus*: Fine-tuning of sugar degradation? Biochemical Society Transactions 39: 77-81.
114. Kardinahl S, Schmidt CL, Hansen T, Anemüller S, Petersen A, et al. (1999) The strict molybdate-dependence of glucose-degradation by the thermoacidophile *Sulfolobus acidocaldarius* reveals the first crenarchaeotic molybdenum containing enzyme - An aldehyde oxidoreductase. European Journal of Biochemistry 260: 540-548.
115. Kouril T, Wieloch P, Reimann J, Wagner M, Zaparty M, et al. (2013) Unraveling the function of the two Entner–Doudoroff branches in the thermoacidophilic Crenarchaeon *Sulfolobus solfataricus* P2. FEBS Journal 280: 1126-1138.
116. Ettema TJG, Ahmed H, Geerling ACM, 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.
117. Matsubara K, Yokooji Y, Atomi H, Imanaka T (2011) Biochemical and genetic characterization of the three metabolic routes in *Thermococcus kodakarensis* linking glyceraldehyde 3-phosphate and 3-phosphoglycerate. Molecular Microbiology 81: 1300-1312.
118. Brunner NA, Siebers B, Hensel R (2001) Role of two different glyceraldehyde-3-phosphate dehydrogenases in controlling the reversible Embden–Meyerhof–Parnas pathway in *Thermoproteus tenax*: regulation on protein and transcript level. Extremophiles 5: 101-109.
119. Kouril T, Esser D, Kort J, Westerhoff HV, Siebers B, et al. (2013) Intermediate instability at high temperature leads to low pathway efficiency for in vitro reconstituted gluconeogenesis of *Sulfolobus solfataricus* FEBS Journal (Submitted).
120. Haferkamp P (2011) Biochemical studies of enzymes involved in glycolysis of the thermoacidophilic crenarchaeon Sulfolobus solfataricus: Universität Duisburg-Essen, Fakultät für Chemie.

121. Kato N, Yurimoto H, Thauer RK (2006) The Physiological Role of the Ribulose Monophosphate Pathway in Bacteria and Archaea. Bioscience, Biotechnology, and Biochemistry 70: 10-21.
122. Quayle JR, Ferenci T (1978) Evolutionary aspects of autotrophy. Microbiological Reviews 42: 251-273.
123. Wolterink-van Loo S, Siemerink MAJ, Perrakis G, Kaper T, Kengen SW, et al. (2009) Improving low-temperature activity of *Sulfolobus acidocaldarius* 2-keto-3-deoxygluconate aldolase. Archaea 2: 233-239.
124. Brouns SJJ, Turnbull AP, Willemen HLDM, Akerboom J, van der Oost J (2007) Crystal Structure and Biochemical Properties of the D-Arabinose Dehydrogenase from *Sulfolobus solfataricus*. Journal of Molecular Biology 371: 1249-1260.
125. Brouns SJJ, Barends TRM, Worm P, Akerboom J, Turnbull AP, et al. (2008) Structural Insight into Substrate Binding and Catalysis of a Novel 2-Keto-3-deoxy-d-arabinonate Dehydratase Illustrates Common Mechanistic Features of the FAH Superfamily. Journal of Molecular Biology 379: 357-371.
126. Johnsen U, Dambeck M, Zaiss H, Fuhrer T, Soppa J, et al. (2009) D-xylose degradation pathway in the halophilic archaeon *Haloferax volcanii*. Journal of Biological Chemistry 284: 27290-27303.
127. Buchanan CL, Connaris H, Danson MJ, Reeve CD, Hough DW (1999) An extremely thermostable aldolase from *Sulfolobus solfataricus* with specificity for non-phosphorylated substrates. Biochemical Journal 343: 563-570.
128. Archer RM, Royer SF, Mahy W, Winn CL, Danson MJ, et al. (2013) Syntheses of 2-Keto-3-deoxy-D-xylonate and 2-Keto-3-deoxy-L-arabinonate as Stereochemical Probes for Demonstrating the Metabolic Promiscuity of *Sulfolobus solfataricus* Towards D-Xylose and L-Arabinose. Chemistry-A European Journal 19: 2895-2902.
129. Uhrigshardt H, Walden M, John H, Petersen A, Anemüller S (2002) Evidence for an operative glyoxylate cycle in the thermoacidophilic crenarchaeon *Sulfolobus acidocaldarius*. FEBS Letters 513: 223-229.
130. Tomlinson JD, Turner JF (1979) Pea seed triose phosphate isomerase. Phytochemistry 18: 1959-1962.
131. Giardina P, De Biasi MG, De Rosa M, Gambacorta A, Buonocore V (1986) Glucose dehydrogenase from the thermoacidophilic archaebacterium *Sulfolobus solfataricus*. Biochemical Journal 239: 517.
132. Milburn CC, Lamble HJ, Theodossis A, Bull SD, Hough DW, et al. (2006) The structural basis of substrate promiscuity in glucose dehydrogenase from the hyperthermophilic archaeon *Sulfolobus solfataricus*. Journal of Biological Chemistry 281: 14796-14804.
133. Bright JR, Byrom D, Danson MJ, Hough DW, Towner P (2005) Cloning, sequencing and expression of the gene encoding glucose dehydrogenase from the thermophilic archaeon *Thermoplasma acidophilum*. European Journal of Biochemistry 211: 549-554.
134. Smith LD, Budgen N, Bungard SJ, Danson MJ, Hough DW (1989) Purification and characterization of glucose dehydrogenase from the thermoacidophilic archaebacterium *Thermoplasma acidophilum*. Biochemical Journal 261: 973.
135. Siebers B, Wendisch VF, 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. Archives of Microbiology 168: 120-127.
136. Angelov A, Fütterer O, Valerius O, Braus GH, Liebl W (2005) Properties of the recombinant glucose/galactose dehydrogenase from the extreme thermoacidophile, *Picrophilus torridus*. FEBS Journal 272: 1054-1062.
137. Bonete MJ, Pire C, Llorca FI, Camacho ML (1996) Glucose dehydrogenase from the halophilic Archaeon *Haloferax mediterranei*: Enzyme purification, characterisation and N-terminal sequence. FEBS Letters 383: 227-229.
138. Pire C, Camacho ML, Ferrer J, Hough DW, Bonete MJ (2000) $NAD(P)^+$-glucose dehydrogenase from *Haloferax mediterranei*: kinetic mechanism and metal content. Journal of Molecular Catalysis B: Enzymatic 10: 409-417.

139. Pire C, Esclapez J, Ferrer J, Bonete MJ (2006) Heterologous overexpression of glucose dehydrogenase from the halophilic archaeon *Haloferax mediterranei*, an enzyme of the medium chain dehydrogenase/reductase family. FEMS Microbiology Letters 200: 221-227.
140. Johnsen U, Schönheit P (2004) Novel Xylose Dehydrogenase in the Halophilic Archaeon *Haloarcula marismortui*. Journal of Bacteriology 186: 6198-6207.
141. Cleton-Jansen AM, Goosen N, Odle G, Van de Putte P (1988) Nucleotide sequence of the gene coding for quinoprotein glucose dehydrogenase from *Acinetobacter calcoaceticus*. Nucleic Acids Research 16: 6228.
142. Cleton-Jansen AM, Goosen N, Vink K, van de Putte P (1989) Cloning, characterization and DNA sequencing of the gene encoding the M r 50000 quinoprotein glucose dehydrogenase from *Acinetobacter calcoaceticus*. Molecular and General Genetics MGG 217: 430-436.
143. Serrano JA, Camacho ML, Bonete MJ (1998) Operation of glyoxylate cycle in halophilic archaea: presence of malate synthase and isocitrate lyase in *Haloferax volcanii*. FEBS Letters 434: 13-16.
144. Edwards KJ, Barton JD, Rossjohn J, Thorn JM, Taylor GL, et al. (1996) Structural and sequence comparisons of quinone oxidoreductase, zeta-crystallin, and glucose and alcohol dehydrogenases. Archives of Biochemistry and Biophysics 328: 173.
145. Asada Y, Endo S, Inoue Y, Mamiya H, Hara A, et al. (2009) Biochemical and structural characterization of a short-chain dehydrogenase/reductase of *Thermus thermophilus* HB8: A hyperthermostable aldose-1-dehydrogenase with broad substrate specificity. Chemico-biological interactions 178: 117-126.
146. Talfournier F, Stines-Chaumeil C, Branlant G (2011) Methylmalonate semialdehyde dehydrogenase from *Bacillus subtilis*: substrate specificity and coenzyme A binding. Journal of Biological Chemistry.
147. Stines-Chaumeil C, Talfournier F, Branlant G (2006) Mechanistic characterization of the MSDH (methylmalonate semialdehyde dehydrogenase) from *Bacillus subtilis*. Biochemical Journal 395: 107-115.
148. Lindahl R (1992) Aldehyde dehydrogenases and their role in carcinogenesis. Critical Reviews in Biochemistry and Molecular Biology 27: 283-335.
149. Vasiliou V, Pappa A, Petersen DR (2000) Role of aldehyde dehydrogenases in endogenous and xenobiotic metabolism. Chemico-Biological Interactions 129: 1-19.
150. Sophos NA, Vasiliou V (2003) Aldehyde dehydrogenase gene superfamily: The 2002 update. Chemico-Biological Interactions 143-144: 5-22.
151. Fuhrer T, Chen L, Sauer U, Vitkup D (2007) Computational prediction and experimental verification of the gene encoding the NAD^+/$NADP^+$-dependent succinate semialdehyde dehydrogenase in *Escherichia coli*. Journal of Bacteriology 189: 8073-8078.
152. Ludewig F, Hüser A, Fromm H, Beauclair L, Bouche N (2008) Mutants of GABA transaminase (POP2) suppress the severe phenotype of succinic semialdehyde dehydrogenase (ssadh) mutants in arabidopsis. PLoS ONE 3.
153. Dubourg H, Stines-Chaumeil C, Didierjean C, Talfournier F, Rahuel-Clermont S, et al. (2004) Expression, purification, crystallization and preliminary X-ray diffraction data of methylmalonate-semialdehyde dehydrogenase from *Bacillus subtilis*. Acta Crystallographica Section D: Biological Crystallography 60: 1435-1437.
154. Massey LK, Sokatch JR, Conrad RS (1976) Branched chain amino acid catabolism in bacteria. Bacteriological Reviews 40: 42-54.
155. Sterner R, Höcker B (2005) Catalytic versatility, stability, and evolution of the (β/!) 8-barrel enzyme fold. Chem Rev 105: 4038-4055.
156. Jaenicke R, Böhm G (1998) The stability of proteins in extreme environments. Current Opinion in Structural Biology 8: 738-748.
157. Sterner R, Liebl W (2001) Thermophilic adaptation of proteins. Critical reviews in biochemistry and molecular biology 36: 39-106.
158. Vieille C, Zeikus GJ (2001) Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. Microbiology and Molecular Biology Reviews 65: 1-43.

159. López-Velázquez G, Molina-Ortiz D, Cabrera N, Hernández-Alcántara G, Peon-Peralta J, et al. (2004) An unusual triosephosphate isomerase from the early divergent eukaryote *Giardia lamblia*. Proteins: Structure, Function, and Bioinformatics 55: 824-834.
160. Schurig H, Beaucamp N, Ostendorp R, Jaenicke R, Adler E, et al. (1995) Phosphoglycerate kinase and triosephosphate isomerase from the hyperthermophilic bacterium *Thermotoga maritima* form a covalent bifunctional enzyme complex. The EMBO journal 14: 442.
161. Schramm A, Kohlhoff M, Hensel R (2001) Triose-phosphate isomerase from *Pyrococcus woesei* and *Methanothermus fervidus*. In: Michael W. W. Adams RMK, editor. Methods in Enzymology: Academic Press. pp. 62-77.
162. Kohlhoff M, Dahm A, Hensel R (1996) Tetrameric triosephosphate isomerase from hyperthermophilic Archaea. FEBS Letters 383: 245-250.
163. Lodi PJ, Knowles JR (1991) Neutral imidazole is the electrophile in the reaction catalyzed by triosephosphate isomerase: structural origins and catalytic implications. Biochemistry 30: 6948-6956.
164. Walden H, Taylor GL, Lorentzen E, Pohl E, Lilie H, et al. (2004) Structure and function of a regulated archaeal triosephosphate isomerase adapted to high temperature. Journal of Molecular Biology 342: 861-875.
165. Walden H, Bell GS, Russell RJ, Siebers B, Hensel R, et al. (2001) Tiny TIM: a small, tetrameric, hyperthermostable triosephosphate isomerase. Journal of Molecular Biology 306: 745.
166. Yu JS, Noll KM (2006) The hyperthermophilic bacterium *Thermotoga neapolitana* possesses two isozymes of the 3-phosphoglycerate kinase/triosephosphate isomerase fusion protein. FEMS Microbiology Letters 131: 307-312.
167. Lambeir AM, Opperdoes FR, Wierenga RK (1987) Kinetic properties of triose-phosphate isomerase from *Trypanosoma brucei*. European Journal of Biochemistry 168: 69-74.
168. Van Der Does C, Den Blaauwen T, De Wit JG, Manting EH, Groot NA, et al. (1996) SecA is an intrinsic subunit of the *Escherichia coli* preprotein translocase and exposes its carboxyl terminus to the periplasm. Molecular Microbiology 22: 619-629.
169. Kraus GA, Landgrebe K (1984) A Direct Synthesis of ω-Bromo-1-alkenes. Synthesis 10: 885.
170. Macritchie JA, Silcock A, Willis CL (1997) Enantioselective synthesis of unsaturated α-hydroxy acids. Tetrahedron Asymmetry 8: 3895-3902.
171. Crestia D, Guerard C, Bolte J, Demuynck C (2001) Rabbit muscle aldolase (RAMA) as a catalyst in a new approach for the synthesis of 3-deoxy-D-manno-2-octulosonic acid and analogues. Journal of Molecular Catalysis - B Enzymatic 11: 207-212.
172. Kupiecki FP, Coon MJ (1960) Methylmalonic semialdehyde. Biochem Prep 7: 69-71.
173. Brock TD, Brock KM, Belly RT, Weiss RL (1972) Sulfolobus: A new genus of sulfur-oxidizing bacteria living at low pH and high temperature. Archives of Microbiology 84: 54-68.
174. Hoffmann S, Otto C, Kurtz S, Sharma CM, Khaitovich P, et al. (2009) Fast mapping of short sequences with mismatches, insertions and deletions using index structures. PLoS computational biology 5: e1000502.
175. Anders S, Huber W (2010) Differential expression analysis for sequence count data. Genome Biol 11: R106.
176. Zaparty M, Esser D, Gertig S, Haferkamp P, Kouril T, et al. (2009) "Hot standards" for the thermoacidophilic archaeon *Sulfolobus solfataricus*. Extremophiles 14: 119-142.
177. Pham TK, Sierocinski P, van der Oost J, Wright PC (2009) Quantitative Proteomic Analysis of *Sulfolobus solfataricus* Membrane Proteins. Journal of Proteome Research 9: 1165-1172.
178. Gan CS, Reardon KF, Wright PC (2005) Comparison of protein and peptide prefractionation methods for the shotgun proteomic analysis of *Synechocystis* sp. PCC 6803. Proteomics 5: 2468-2478.
179. Panchaud A, Scherl A, Shaffer SA, Von Haller PD, Kulasekara HD, et al. (2009) Precursor acquisition independent from ion count: How to dive deeper into the proteomics ocean. Analytical Chemistry 81: 6481-6488.
180. Chong PK, Wright PC (2005) Identification and Characterization of the Sulfolobus solfataricus P2 Proteome. Journal of Proteome Research 4: 1789-1798.

181. Szklarczyk D, Franceschini A, Kuhn M, Simonovic M, Roth A, et al. (2011) The STRING database in 2011: Functional interaction networks of proteins, globally integrated and scored. Nucleic Acids Research 39: D561-D568.
182. Zillig W, Stetter KO, Wunderl S (1980) The Sulfolobus-'Caldariella' group: Taxonomy on the basis of the structure of DNA-dependent RNA polymerases. Archives of Microbiology 125: 259-269.
183. Smith CM, Shindyalov IN, Veretnik S, Gribskov M, Taylor SS, et al. (1997) The protein kinase resource. Trends in Biochemical Sciences 22: 444-446.
184. Kim D, Forst S (2001) Genomic analysis of the histidine kinase family in bacteria and archaea. Microbiology 147: 1197-1212.
185. Koretke KK, Lupas AN, Warren PV, Rosenberg M, Brown JR (2000) Evolution of two-component signal transduction. Molecular Biology and Evolution 17: 1956-1970.
186. Bhaduri A, Sowdhamini R (2005) Genome-wide survey of prokaryotic O-protein phosphatases. Journal of Molecular Biology 352: 736-752.
187. Duan X, He ZG (2011) Characterization of the specific interaction between archaeal FHA domain-containing protein and the promoter of a flagellar-like gene-cluster and its regulation by phosphorylation. Biochemical and Biophysical Research Communications.
188. Esser D, Pham TK, Reimann J, Albers SV, Siebers B, et al. (2012) Change of carbon source causes dramatic effects in the phospho-proteome of the Archaeon *Sulfolobus solfataricus*. Journal of Proteome Research.
189. Bi YD, Wang HX, Lu TC, Li Xh, Shen Z, et al. (2010) Large-scale analysis of phosphorylated proteins in maize leaf. Planta: 1-10.
190. Metzger S, Hoffmann R (2000) Studies on the dephosphorylation of phosphotyrosine-containing peptides during post-source decay in matrix-assisted laser desorption/ionization. Journal of Mass Spectrometry 35: 1165-1177.
191. Smith SC, Kennelly PJ, Potts M (1997) Protein-tyrosine phosphorylation in the Archaea. Journal of Bacteriology 179: 2418-2420.
192. Foster R, Thorner J, Martin GS (1989) Nucleotidylation, not phosphorylation, is the major source of the phosphotyrosine detected in enteric bacteria. Journal of Bacteriology 171: 272-279.
193. Hanks SK, Lindberg RA (1991) Use of degenerate oligonucleotide probes to identify clones that encode protein kinases. Methods in Enzymology 200: 525-532.
194. Esser D, Kouril T, Zaparty M, Sierocinski P, Chan P, et al. (2011) Functional curation of the *Sulfolobus solfataricus* P2 and *S. acidocaldarius* 98-3 complete genome sequences. Extremophiles: 1-2.
195. Makarova KS, Sorokin AV, Novichkov PS, Wolf YI, Koonin EV (2007) Clusters of orthologous genes for 41 archaeal genomes and implications for evolutionary genomics of archaea. Biology Direct 2.
196. Daas PJH, Wassenaar RW, Willemsen P, Theunissen RJ, Keltjens JT, et al. (1996) Purification and properties of an enzyme involved in the ATP-dependent activation of the methanol:2-mercaptoethanesulfonic acid methyltransferase reaction in *Methanosarcina barkeri*. Journal of Biological Chemistry 271: 22339-22345.
197. Ray WK, Keith SM, DeSantis AM, Hunt JP, Larson TJ, et al. (2005) A phosphohexomutase from the archaeon *Sulfolobus solfataricus* is covalently modified by phosphorylation on serine. Journal of Bacteriology 187: 4270-4275.
198. Humbard MA, Stevens Jr SM, Maupin-Furlow JA (2006) Posttranslational modication of the 20S proteasomal proteins of the archaeon *Haloferax volcanii*. Journal of Bacteriology 188: 7521-7530.
199. Grabowski B, Kelman Z (2001) Autophosphorylation of archaeal Cdc6 homologues is regulated by DNA. Journal of Bacteriology 183: 5459-5464.
200. De Felice M, Esposito L, Rossi M, Pisani FM (2006) Biochemical characterization of two Cdc6/ORC1-like proteins from the crenarchaeon *Sulfolobus solfataricus*. Extremophiles 10: 61-70.

201. Tahara M, Ohsawa A, Saito S, Kimura M (2004) *In Vitro* Phosphorylation of Initiation Factor 2α (aIF2α) from Hyperthermophilic Archaeon *Pyrococcus horikoshii* OT3. Journal of Biochemistry 135: 479-485.
202. Condò I, Ruggero D, Reinhardt R, Londei P (1998) A novel aminopeptidase associated with the 60 kDa chaperonin in the thermophilic archaeon *Sulfolobus solfataricus*. Molecular Microbiology 29: 775-785.
203. Solow B, Bischoff KM, Zylka MJ, Kennelly PJ (1998) Archaeal phosphoproteins. Identification of a hexosephosphate mutase and the α-subunit of succinyl-CoA synthetase in the extreme acidothermophile *Sulfolobus solfataricus*. Protein Science 7: 105-111.
204. Cardona S, Remonsellez F, Guiliani N, Jerez CA (2001) The Glycogen-Bound Polyphosphate Kinase from *Sulfolobus acidocaldarius* Is Actually a Glycogen Synthase. Applied and Environmental Microbiology 67: 4773-4780.
205. Reher M, Fuhrer T, Bott M, Schönheit P (2010) The nonphosphorylative entner-doudoroff pathway in the thermoacidophilic euryarchaeon *Picrophilus torridus* involves a novel 2-Keto-3-deoxygluconate- specific aldolase. Journal of Bacteriology 192: 964-974.
206. Esser D, Kouril T, Talfournier F, Polkowska J, Schrader T, et al. (2013) Unraveling the function of paralogs of the aldehyde dehydrogenase super family from *Sulfolobus solfataricus*. Extremophiles: 1-12.
207. Cortay JC, Bleicher F, Rieul C, Reeves HC, Cozzone AJ (1988) Nucleotide sequence and expression of the aceK gene coding for isocitrate dehydrogenase kinase/phosphatase in *Escherichia coli*. Journal of Bacteriology 170: 89-97.
208. Klumpp DJ, Plank DW, Bowdin LJ, Stueland CS, Chung T, et al. (1988) Nucleotide sequence of aceK, the gene encoding isocitrate dehydrogenase kinase/phosphatase. Journal of Bacteriology 170: 2763-2769.
209. Cozzone AJ (1988) Protein phosphorylation in prokaryotes. Annual Review of Microbiology 42: 97-125.
210. Johnson LN, Lewis RJ (2001) Structural basis for control by phosphorylation. Chemical Reviews-Columbus 101: 2209-2242.
211. Westheimer FH (1987) Why nature chose phosphates. Science 235: 1173-1178.
212. Pawson T, Scott JD (2005) Protein phosphorylation in signaling--50 years and counting. Trends in Biochemical Sciences 30: 286.
213. Galperin MY (2005) A census of membrane-bound and intracellular signal transduction proteins in bacteria: Bacterial IQ, extroverts and introverts. BMC Microbiology 5.
214. Hubbard MJ, Cohen P (1993) On target with a new mechanism for the regulation of protein phosphorylation. Trends in Biochemical Sciences 18: 172-177.
215. Barton GJ, Cohen PTW, Barford D (1994) Conservation analysis and structure prediction of the protein serine/threonine phosphatases. European Journal of Biochemistry 220: 225-237.
216. Bork P, Brown NP, Hegyi H, Schultz J (1996) The protein phosphatase 2C (PP2C) superfamily: detection of bacterial homologues. Protein Science 5: 1421-1425.
217. Bernander R, Poplawski A (1997) Cell cycle characteristics of thermophilic archaea. Journal of Bacteriology 179: 4963-4969.
218. Reimann J, Esser D, Pham KT, Orell A, Amman F, et al. (2013) (Cren)Archaeal Signal Transduction: In Vivo Impact of Protein Phosphatase Deletions on Cell Size, Motility and Energy Metabolism in *Sulfolobus acidocaldarius*. Molecular and cellular proteomics.
219. Makarova KS, Omelchenko MV, Gaidamakova EK, Matrosova VY, Vasilenko A, et al. (2007) *Deinococcus geothermalis*: the pool of extreme radiation resistance genes shrinks. PLoS ONE 2: e955.
220. Hiller A, Henninger T, Schäfer G, Schmidt CL (2003) New Genes Encoding Subunits of a Cytochrome bc 1–Analogous Complex in the Respiratory Chain of the HyperThermoacidophilic Crenarchaeon *Sulfolobus acidocaldarius*. Journal of bioenergetics and biomembranes 35: 121-131.
221. Jarrell KF, Albers SV (2012) The archaellum: an old motility structure with a new name. Trends in Microbiology.

222. Lübben M, Warne A, Albracht SPJ, Saraste M (2006) The purified SoxABCD quinol oxidase complex of *Sulfolobus acidocaldarius* contains a novel haem. Molecular Microbiology 13: 327-335.
223. Gleißner M, Kaiser U, Antonopoulos E, Schäfer G (1997) The archaeal SoxABCD complex is a proton pump in *Sulfolobus acidocaldarius*. Journal of Biological Chemistry 272: 8417-8426.
224. Lübben M, Kolmerer B, Saraste M (1992) An archaebacterial terminal oxidase combines core structures of two mitochondrial respiratory complexes. The EMBO journal 11: 805.
225. Castresana J, Lübben M, Saraste M (1995) New archaebacterial genes coding for redox proteins: implications for the evolution of aerobic metabolism. Journal of Molecular Biology 250: 202-210.
226. Komorowski L, Verheyen W, Schäfer G (2002) The archaeal respiratory supercomplex SoxM from *S. acidocaldarius* combines features of quinole and cytochrome c oxidases. Biological chemistry 383: 1791-1799.
227. Lübben M, Arnaud S, Castresana J, Warne A, Albracht SPJ, et al. (1994) A second terminal oxidase in *Sulfolobus acidocaldarius*. European Journal of Biochemistry 224: 151-159.
228. Purschke WG, Schmidt CL, Petersen A, Schäfer G (1997) The terminal quinol oxidase of the hyperthermophilic archaeon *Acidianus ambivalens* exhibits a novel subunit structure and gene organization. Journal of Bacteriology 179: 1344-1353.
229. Henche AL, Koerdt A, Ghosh A, Albers SV (2012) Influence of cell surface structures on crenarchaeal biofilm formation using a thermostable green fluorescent protein. Environmental Microbiology 14: 779-793.
230. Lassak K, Neiner T, Ghosh A, Klingl A, Wirth R, et al. (2012) Molecular analysis of the crenarchaeal flagellum. Molecular Microbiology 83: 110-124.
231. Chu HM, Wang AHJ (2006) Enzyme–substrate interactions revealed by the crystal structures of the archaeal Sulfolobus PTP-fold phosphatase and its phosphopeptide complexes. Proteins: Structure, Function, and Bioinformatics 66: 996-1003.
232. Lundgren M, Bernander R (2007) Genome-wide transcription map of an archaeal cell cycle. Proceedings of the National Academy of Sciences 104: 2939-2944.
233. Duan X, He Z (2011) Characterization of the specific interaction between archaeal FHA domain-containing protein and the promoter of a flagellar-like gene-cluster and its regulation by phosphorylation. Biochemical and Biophysical Research Communications 407: 242-247.
234. Esposito L, Ruggiero A, Masullo M, Ruocco MR, Lamberti A, et al. (2011) Crystallographic and spectroscopic characterizations of *Sulfolobus solfataricus* TrxA1 provide insights into the determinants of thioredoxin fold stability. Journal of Structural Biology.
235. Ming H, Kato Y, Miyazono K, Ito K, Kamo M, et al. (2007) Crystal structure of thioredoxin domain of ST2123 from thermophilic archaea *Sulfolobus tokodaii* strain7. Proteins: Structure, Function, and Bioinformatics 69: 204-208.
236. Grimaldi P, Ruocco MR, Lanzotti MA, Ruggiero A, Ruggiero I, et al. (2008) Characterisation of the components of the thioredoxin system in the archaeon *Sulfolobus solfataricus*. Extremophiles 12: 553-562.
237. Ruocco MR, Ruggiero A, Masullo L, Arcari P, Masullo M (2004) A 35 kDa NAD(P)H oxidase previously isolated from the archaeon *Sulfolobus solfataricus* is instead a thioredoxin reductase. Biochimie 86: 883-892.
238. Ruggiero A, Lanzotti MA, Ruocco MR, Grimaldi P, Marasco D, et al. (2009) Crystallization and preliminary X-ray crystallographic analysis of two dimeric hyperthermostable thioredoxins isolated from *Sulfolobus solfataricus*. Acta Crystallographica Section F: Structural Biology and Crystallization Communications 65: 604-607.
239. Maaty WS, Wiedenheft B, Tarlykov P, Schaff N, Heinemann J, et al. (2009) Something old, something new, something borrowed; how the thermoacidophilic archaeon *Sulfolobus solfataricus* responds to oxidative stress. PLoS ONE 4.
240. Zarbiv G, Li H, Wolf A, Cecchini G, Caplan SR, et al. (2011) Energy Complexes Are Apparently Associated with the Switch–Motor Complex of Bacterial Flagella. Journal of Molecular Biology.

241. Ravichandran A, Sugiyama N, Tomita M, Swarup S, Ishihama Y (2009) Ser/Thr/Tyr phosphoproteome analysis of pathogenic and non-pathogenic Pseudomonas species. Proteomics 9: 2764-2775.
242. Ohlmeier S, Hiltunen JK, Bergmann U (2010) Protein phosphorylation in mitochondria–A study on fermentative and respiratory growth of *Saccharomyces cerevisiae*. ELECTROPHORESIS 31: 2869-2881.
243. Zhao X, León IR, Bak S, Mogensen M, Wrzesinski K, et al. (2011) Phosphoproteome analysis of functional mitochondria isolated from resting human muscle reveals extensive phosphorylation of inner membrane protein complexes and enzymes. Molecular & Cellular Proteomics 10.
244. Esser D, Siebers B (2013) Atypical protein kinases of the RIO family in archaea. Biochemical Society transactions 41: 399-404.
245. Valverde R, Edwards L, Regan L (2008) Structure and function of KH domains. FEBS Journal 275: 2712-2726.
246. Benelli D, Maone E, Londei P (2003) Two different mechanisms for ribosome/mRNA interaction in archaeal translation initiation. Molecular Microbiology 50: 635-643.
247. Tolstrup N, Sensen CW, Garrett RA, Clausen IG (2000) Two different and highly organized mechanisms of translation initiation in the archaeon *Sulfolobus solfataricus*. Extremophiles 4: 175-179.
248. Londei P (2005) Evolution of translational initiation: New insights from the archaea. FEMS Microbiology Reviews 29: 185-200.
249. Hasenöhrl D, Benelli D, Barbazza A, Londei P, Bläsi U (2006) *Sulfolobus solfataricus* translation initiation factor 1 stimulates translation initiation complex formation. RNA 12: 674-682.
250. Linder P (2006) Dead-box proteins: A family affair - Active and passive players in RNP-remodeling. Nucleic Acids Research 34: 4168-4180.
251. Tarn WY, Chang TH (2009) The current understanding of Ded1p/DDX3 homologs from yeast to human. RNA Biology 6: 17-20.
252. Jarmoskaite I, Russell R (2011) DEAD-box proteins as RNA helicases and chaperones. Wiley interdisciplinary reviews RNA 2: 135-152.
253. Prud'homme-Généreux A, Beran RK, Iost I, Ramey CS, Mackie GA, et al. (2004) Physical and functional interactions among RNase E, polynucleotide phosphorylase and the cold-shock protein, CsdA: Evidence for a 'cold shock degradosome'. Molecular Microbiology 54: 1409-1421.
254. Charollais J, Dreyfus M, Iost I (2004) CsdA, a cold-shock RNA helicase from *Escherichia coli*, is involved in the biogenesis of 50S ribosomal subunit. Nucleic Acids Research 32: 2751-2759.
255. Owttrim GW (2006) RNA helicases and abiotic stress. Nucleic Acids Research 34: 3220-3230.
256. Boonyaratanakornkit BB, Simpson AJ, Whitehead TA, Fraser CM, El-Sayed NMA, et al. (2005) Transcriptional profiling of the hyperthermophilic methanarchaeon *Methanococcus jannaschii* in response to lethal heat and non-lethal cold shock. Environmental Microbiology 7: 789-797.
257. Lim J, Thomas T, Cavicchioli R (2000) Low temperature regulated DEAD-box RNA helicase from the Antarctic archaeon, *Methanococcoides burtonii*. Journal of Molecular Biology 297: 553-567.
258. Bateman A, Coggill P, Finn RD (2010) DUFs: Families in search of function. Acta Crystallographica Section F: Structural Biology and Crystallization Communications 66: 1148-1152.
259. Murina VN, Nikulin AD (2011) RNA-binding Sm-like proteins of bacteria and archaea. Similarity and difference in structure and function. Biochemistry (Moscow) 76: 1434-1449.
260. Nielsen JS, Bøggild A, Andersen CBF, Nielsen G, Boysen A, et al. (2007) An Hfq-like protein in archaea: Crystal structure and functional characterization of the Sm protein from *Methanococcus jannaschii*. RNA 13: 2213-2223.
261. Bertrand EM, Saito MA, Jeon YJ, Neilan BA (2011) Vitamin B12 biosynthesis gene diversity in the Ross Sea: the identification of a new group of putative polar B12 biosynthesizers. Environmental Microbiology 13: 1285-1298.

262. Hou J, Liu Y, Lu Z, Liu X, Liu J (2012) Biochemical characterization of RNase HII from *Aeropyrum pernix*. Acta Biochimica et Biophysica Sinica 44: 339-346.
263. Muroya A, Tsuchiya D, Ishikawa M, Haruki M, Morikawa M, et al. (2001) Catalytic center of an archaeal type 2 ribonuclease H as revealed by X-ray crystallographic and mutational analyses. Protein Science 10: 707-714.
264. Greber BJ, Boehringer D, Godinic-Mikulcic V, Crnkovic A, Ibba M, et al. (2012) Cryo-EM Structure of the Archaeal 50S Ribosomal Subunit in Complex with Initiation Factor 6 and Implications for Ribosome Evolution. Journal of Molecular Biology 418: 145-160.
265. Vasile F, Pechkova E, Nicolini C (2008) Solution structure of the β-subunit of the translation initiation factor aIF2 from archaebacteria *Sulfolobus solfataricus*. Proteins: Structure, Function, and Bioinformatics 70: 1112-1115.
266. Makarova KS, Koonin EV (2010) Archaeal Ubiquitin-Like Proteins: Functional Versatility and Putative Ancestral Involvement in tRNA Modification Revealed by Comparative Genomic Analysis. Archaea 2010.
267. Sofia HJ, Chen G, Hetzler BG, Reyes-Spindola JF, Miller NE (2001) Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: functional characterization using new analysis and information visualization methods. Nucleic Acids Research 29: 1097-1106.
268. Arragain S, Handelman SlK, Forouhar F, Wei F, Tomizawa K, et al. (2010) Identification of Eukaryotic and Prokaryotic Methylthiotransferase for Biosynthesis of 2-Methylthio-N6-threonylcarbamoyladenosine in tRNA. Journal of Biological Chemistry 285: 28425-28433.
269. Stolboushkina E, Nikonov S, Nikulin AD, Bläsi U, Manstein DJ, et al. (2008) Crystal Structure of the Intact Archaeal Translation Initiation Factor 2 Demonstrates Very High Conformational Flexibility in the α- and β-Subunits. Journal of Molecular Biology 382: 680-691.
270. She Q, Singh RK, Confalonieri F, Zivanovic Y, Allard G, et al. (2001) The complete genome of the crenarchaeon Sulfolobus solfataricus P2. Proceedings of the National Academy of Sciences of the United States of America 98: 7835-7840.
271. Oost JVD, Siebers B (2007) The Glycolytic Pathways of Archaea: Evolution By Tinkering: Blackwell Publishing Ltd. 247-259 p.
272. Zaparty M, Zaigler A, Stamme C, Soppa J, Hensel R, et al. (2008) DNA microarray analysis of central carbohydrate metabolism: Glycolytic/gluconeogenic carbon switch in the hyperthermophilic crenarchaeum Thermoproteus tenax. Journal of Bacteriology 190: 2231-2238.
273. Cash CD, Maitre M, Ossola L, Mandel P (1978) Purification and properties of two succinate semialdehyde dehydrogenases from human brain. Biochimica et biophysica acta 524: 26-36.
274. Sreerama L, Sladek NE (1993) Identification and characterization of a novel class 3 aldehyde dehydrogenase overexpressed in a human breast adenocarcinoma cell line exhibiting oxazaphosphorine-specific acquired resistance. Biochemical pharmacology 45: 2487-2505.
275. Tamaki N, Nakamura M, Kimura K, Hama T (1977) Purification and Properties of Aldehyde Dehydrogenase from *Saccharomyces cerevisiae*. Journal of Biochemistry 82: 73-79.
276. Rosemblatt MS, Callewaert DM, Tchen TT (1973) Succinic semialdehyde dehydrogenase from a *Pseudomonas species*. II. Physical and immunochemical properties of the enzyme. Journal of Biological Chemistry 248: 6014-6018.
277. Cash C, Ciesielski L, Maitre M, Mandel P (1977) Purification and properties of rat brain succinic semialdehyde dehydrogenase. Biochimie 59: 257-268.
278. Rothacker B, Werr M, Ilg T (2008) Succinic semialdehyde dehydrogenase from the parasitic cattle tick *Rhipicephalus microplus*: Gene identification, biochemical characterization and comparison with the mouse ortholog. Molecular and Biochemical Parasitology 161: 32-43.
279. Sànchez M, Alvarez MA, Balaña R, Garrido-Pertierra A (1988) Properties and functions of two succinic-semialdehyde dehydrogenases from *Pseudomonas putida*. Biochimica et Biophysica Acta (BBA)/Protein Structure and Molecular 953: 249-257.

280. Kang JH, Park YB, Huh TL, Lee WH, Choi MS, et al. (2005) High-level expression and characterization of the recombinant enzyme, and tissue distribution of human succinic semialdehyde dehydrogenase. Protein Expression and Purification 44: 16-22.
281. Tokunaga M, Nakano Y, Kitaoka S (1976) Separation and properties of the NAD linked and NADP linked isozymes of succinic semialdehyde dehydrogenase in *Euglena gracilis* z. Biochimica et Biophysica Acta 429: 55-62.
282. Sanchez M, Fernandez J, Martin M, Gibello A, Garrido-Pertierra A (1989) Purification and properties of two succinic semialdehyde dehydrogenases from *Klebsiella pneumoniae*. Biochimica et Biophysica Acta - General Subjects 990: 225-231.
283. Donnelly MI, Cooper RA (1981) Two succinic semialdehyde dehydrogenases are induced when *Escherichia coli* K-12 is grown on γ-aminobutyrate. Journal of Bacteriology 145: 1425-1427.
284. Lütke-Eversloh T, Steinbüchel A (1999) Biochemical and molecular characterization of a succinate semialdehyde dehydrogenase involved in the catabolism of 4-hydroxybutyric acid in *Ralstonia eutropha*. FEMS Microbiology Letters 181: 63-71.
285. Van de Casteele M, Demarez M, Legrain C, Glansdorff N, Pierard A (1990) Pathways of arginine biosynthesis in extreme thermophilic archaeo- and eubacteria. Journal of General Microbiology 136: 1177-1183.
286. Cacciapuoti G, Porcelli M, Carteni-Farina M (1986) Purification and characterization of propylamine transferase from *Sulfolobus solfataricus*, an extreme thermophilic archaebacterium. European Journal of Biochemistry 161: 263-271.
287. Koo PH, Adams E (1974) α-Ketoglutaric semialdehyde dehydrogenase of Pseudomonas. Properties of the separately induced isoenzymes. Journal of Biological Chemistry 249: 1704-1716.
288. Bannerjee D, Sanders LE, Sokatch JR (1970) Properties of purified methylmalonate semialdehyde dehydrogenase of *Pseudomonas aeruginosa*. Journal of Biological Chemistry 245: 1828-1835.
289. Sohling B, Gottschalk G (1993) Purification and characterization of a coenzyme-A-dependent succinate-semialdehyde dehydrogenase from *Clostridium kluyveri*. European Journal of Biochemistry 212: 121-127.
290. Kockelkorn D, Fuchs G (2009) Malonic semialdehyde reductase, succinic semialdehyde reductase, and succinyl-coenzyme A reductase from *Metallosphaera sedula*: Enzymes of the autotrophic 3-hydroxypropionate/4-hydroxybutyrate cycle in Sulfolobales. Journal of Bacteriology 191: 6352-6362.
291. Lei Y, Pawelek PD, Powlowski J (2008) A shared binding site for NAD^+ and coenzyme A in an acetaldehyde dehydrogenase involved in bacterial degradation of aromatic compounds. Biochemistry 47: 6870-6882.
292. Bchini R, Dubourg-Gerecke H, Rahuel-Clermont S, Aubry A, Branlant G, et al. (2012) Adenine binding mode is a key factor in triggering the early release of NADH in coenzyme A-dependent methylmalonate semialdehyde dehydrogenase. Journal of Biological Chemistry 287: 31095-31103.
293. Sánchez LB (1998) Aldehyde dehydrogenase (CoA-acetylating) and the mechanism of ethanol formation in the amitochondriate protist, *Giardia lamblia*. Archives of Biochemistry and Biophysics 354: 57-64.
294. Kedishvili NY, Goodwin GW, Popov KM, Harris RA (2000) Mammalian methylmalonate-semialdehyde dehydrogenase. pp. 207-218.
295. Grochowski LL, Xu H, White RH (2006) Identification of Lactaldehyde Dehydrogenase in *Methanocaldococcus jannaschii* and Its Involvement in Production of Lactate for F420 Biosynthesis. Journal of Bacteriology 188: 2836-2844.
296. Reher M, Schönheit P (2006) Glyceraldehyde dehydrogenases from the thermoacidophilic euryarchaeota *Picrophilus torridus* and *Thermoplasma acidophilum*, key enzymes of the non-phosphorylative Entner-Doudoroff pathway, constitute a novel enzyme family within the aldehyde dehydrogenase superfamily. FEBS Letters 580: 1198-1204.
297. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, et al. (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23: 2947-2948.

298. Langendorf CG, Key TLG, Fenalti G, Kan WT, Buckle AM, et al. The X-ray crystal structure of *Escherichia coli* succinic semialdehyde dehydrogenase; structural insights into $NADP^+$/enzyme interactions. PLoS ONE 5.
299. Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular Biology and Evolution 4: 406-425.
300. Sanderson MJ (1989) Confidence limits of Phylogenies: the Bootstrap revisitied. Cladistics 5: 113-129.
301. Zuckerkandl E (1965) Evolutionary divergence and convergence in proteins. Evolving genes and proteins.
302. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, et al. (2011) MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Molecular Biology and Evolution 28: 2731-2739.
303. Chien L, Duan Y, Yu N (2012) Purification and characterization of recombinant glucose dehydrogenase isolated from a hyperthermophilic Sulfolobus-like bacterium. African Journal of Biotechnology 11: 14227-14235.
304. Chen L, Brägger K, Skovgaard M, Redder P, She Q, et al. (2005) The genome of *Sulfolobus acidocaldarius*, a model organism of the Crenarchaeota. Journal of Bacteriology 187: 4992-4999.
305. Beaucamp N, Hofmann A, Kellerer B, Jaenicke R (1997) Dissection of the gene of the bifunctional PGK-TIM fusion protein from the hyperthermophilic bacterium *Thermotoga maritima*: Design and characterization of the separate triosephosphate isomerase. Protein Science 6: 2159-2165.
306. Ravichandran A, Sugiyama N, Tomita M, Swarup S, Ishihama Y (2009) Ser/Thr/Tyr phosphoproteome analysis of pathogenic and non-pathogenic Pseudomonas species. Proteomics 9: 2764-2775.
307. Yang M, Qiao Z, Zhang W, Xiong Q, Zhang J, et al. (2013) Global Phosphoproteomic Analysis Reveals Diverse Functions of Serine/Threonine/Tyrosine Phosphorylation in the Model Cyanobacterium *Synechococcus* sp. Strain PCC 7002. Journal of Proteome Research.
308. Sun X, Ge F, Xiao C, Yin X, Ge R, et al. (2009) Phosphoproteomic analysis reveals the multiple roles of phosphorylation in pathogenic bacterium *Streptococcus pneumoniae*. Journal of Proteome Research 9: 275-282.
309. Hu C, Lin M, Huang H, Ku W, Yi T, et al. (2012) Phosphoproteomic Analysis of *Rhodopseudomonas palustris* Reveals the Role of Pyruvate Phosphate Dikinase Phosphorylation in Lipid Production. Journal of Proteome Research 11: 5362-5375.
310. Voisin S, Watson DC, Tessier L, Ding W, Foote S, et al. (2007) The cytoplasmic phosphoproteome of the Gram-negative bacterium *Campylobacter jejuni*: Evidence for modification by unidentified protein kinases. Proteomics 7: 4338-4348.
311. Parker JL, Jones AME, Serazetdinova L, Saalbach G, Bibb MJ, et al. (2010) Analysis of the phosphoproteome of the multicellular bacterium *Streptomyces coelicolor* A3 (2) by protein/peptide fractionation, phosphopeptide enrichment and high-accuracy mass spectrometry. Proteomics 10: 2486-2497.
312. Misra SK, Milohanic E, Aké F, Mijakovic I, Deutscher J, et al. (2011) Analysis of the serine/threonine/tyrosine phosphoproteome of the pathogenic bacterium *Listeria monocytogenes* reveals phosphorylated proteins related to virulence. Proteomics 11: 4155-4165.
313. Bai X, Ji Z (2012) Phosphoproteomic investigation of a solvent producing bacterium *Clostridium acetobutylicum*. Applied Microbiology and Biotechnology: 1-11.
314. Prisic S, Dankwa S, Schwartz D, Chou MF, Locasale JW, et al. (2010) Extensive phosphorylation with overlapping specificity by *Mycobacterium tuberculosis* serine/threonine protein kinases. Proceedings of the National Academy of Sciences 107: 7521-7526.

6. Appendix

6.1 Abbreviations

Abbreviation	Full name
2-PG	2-Phosphoglycerate
3-PG	3-Phosphoglycerate
ACN	Acetonitrile
ADH	Alcohol dehydrogenase
ADP	Adenosine-5'-diphosphate
aeIF	Archaeal translation initiation factor
AGC	A, G and C families
AH	Aconitase hydratase
ALDH	Aldehyde dehydrogenase
Amp	Ampicillin
APS	Ammoniumpersulfate
arCOG	Archaeal Cluster of Orthologous Genes
ATP	Adenosine 5'-triphosphate
BPG	1,3-Bisphosphoglycerate
BRCA1	Breast cancer 1
BRCT	BRCA1 C-terminal domain
Cam	Chloramphenicol
CAMK	Calcium and calmodium regulated PKs
CCM	Central carbohydrate metabolism
CDS	Protein-coding sequences
CE	Crude extract
CID	Collision induced dissociation
CK1	Cell kinases
CMGC	Cycline-dependent/mitogen activated/glycogen synthase/ cycline-dependent like kinases
CoA	Coenzyme A
CTD	Carboxy-terminal domain
DNA	Deoxyribonucleic acid
DH	Dehydrogenase
DHAP	Dihydroxyacetone-3-phosphate
D-KDA	2-Dehydro-3-deoxy-D-arabinoate
dNTP	Deoxy-nucleoside triphosphate
DOP	2,5-Dioxopentanoate
dPGAM	2,3-bisphosphoglycerate dependent phosphoglycerate mutase
DTT	Dithiothreitol
DUF	Domain of unknown function
D-XylD	D-Xylonate dehydratase

e.g.	example given
EC	Enzyme Commission number
ED	Entner-Doudoroff
Elp3	Eukaryotic elongation sub unit 3
EMP	Embden-Meyerhoff-Parnas
ENO	Enolase
etc.	*et cetera*
F1P	Fructose-1-phosphate
F6P	Fructose-6-phosphate
FA	Formic acid
FBP	Fructose 1,6-bisphosphate
FBPA	Fructose-1,6-bisphosphate aldolase
FBPase	Fructose-1,6-bisphosphate phosphatase
FHA	Forkhead-associated
G6PDH	Glucose 6-phosphate dehydrogenase
G6PI	Glucose 6-phosphate isomerase
GA	Glyceraldehyde
GABA	γ-Aminobutyric acid
GAD	Gluconate dehydratase
GAP	Glyceraldehyde 3-phosphate
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GAPN	Non-phosphorylation glyceraldehyde-3-phosphate dehydrogenase
GDH	Glucose-1-dehydrogenase
GK	Glycerate kinase
GLPDH	Glycerol-1-phosphate dehydrogenase
GluK	Glucokinase
GlyDH	Glycerol dehydrogenase
GlyP	Glycogen phosphorylase
GTP	Guanosine-5'-triphosphate
Hac	Acetic acid
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HK	Hexokinase
IDH	Isocitrate dehydrogenase
iPGAM	2,3-bisphosphoglycerate independent phosphoglycerate mutase
IPTG	Isopropyl-β-thiogalactopyranoside
KAN	Kanamycin
KDG	2-Keto-3-deoxy-gluconate
KDGal	2-Keto-3-deoxy-galactonate
KDPG	2-Keto-3-deoxy-6-phosphogluconate
KDPGal	2-Keto-3-deoxy-6-phosphogalactonate
KH	Human heterogeneous nuclear RNP K homology
LB	Luria-Betani
LC	Liquid chromatography
L-KDA	2-Dehydro-3-deoxy-L-arabinoate

m/z	Mass to charge ratio
MALDI	Matrix assided laser desorption ionization
MBP	Myelin basic protein
MDH	Malate dehydrogenase
MDR	Medium-chain alcohol/ polyol dehydrogenase/ reductase
MeOH	Methanol
MMSA	Methylmalonate semialdehyde
MOAC	Metal oxide affinity chromatography
MS	Mass spectrometry
MS	Maolante semialdehyde
MSDH	Methylmalonate semialdehyde dehydrogenase
NMD	Nonsense-mediated decay
np	Non-phosphorylative
OR	Oxidoreductase
ORF	Open reading frame
PA	Propion aldehyde
PCR	Polymerase chain reaction
PEP	Phosphoenolpyruvate
PEPCK	Phosphoenolpyruvate carboxykinase
PEPS	Phosphoenolpyruvate synthase
PFK	Phosphofructokinase
PGAM	Phosphoglycerate mutase
PGI	Phosphoglucose isomerase
PGK	Phosphoglycerate kinase
PGM	Phosphomannomutase
PK	Protein kinase
PyrK	Pyruvate kinase
pNPP	Para-Nitrophenylphosphate
PP	Protein phosphatase
PPDK	Pyruvate water dikinase
P-peptide	Phosphopeptide
PPM	Metal dependent Proteine-serine/threoning-phosphatase
PPP	Proteine-serine/threoning-phosphatase
P-protein	Phosphoprotein
PQQ	Pyrroloquinoline-quinone
P-site	Phosphorylation site
PTM	Post translational modification
PTP	Phospho-tyrosine-phosphatase
qPCR	Quantitative real time polymerase chain reaction
rcf	Relative centrifugal force
RCM	Reduced carboxyamidomethylated and maleylated
RIO	right open reading frame
RNA	Ribonucleic acid
RuMP	Ribulose mono phosphate

Saci	*Sulfolobus* acidocaldarius
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
snRNP	Small nuclear ribonucleoproteins
sp	Semiphosphorylative
SSA	Succinic semialdehyde
SSADH	Succinic semialdehyde dehydrogenase
SSO	*Sulfolobus* solfataricus
STE	Homologs of yeast STE7, STE11 and STE20 PKs
SU	Subunit
TCA	Tricarboxylic acid
TEMED	N,N,N,N-Tetramethylethylenediamine
TFA	Trifluoroacetic acid
TK	Tyrosine kinase
TKL	Tyrosine kinase-like
TOF	Time of flight
TPI	Triose-3-phosphate isomerase
T_{opt}	Temperature optimum
Tris	Tris-(hyroxymethyl)-aminomethane
Ttx	Thermoproteus tenax
v	Volume
vWA	von Willebrand
w	Weight

6.2. Supporting informations

6.2.1. Tabels

Table SI 1 Identified phosphoproteins involved in the central carbohydrate metabolism.
ADH = alcohol dehydrogenase, ALDH = aldehyde dehydrogenase, AH = aconitase hydratase, CCM = central carbohydrate metabolism, ENO = enolase, FBPA = fructose 1,6-bisphosphate aldolase, FBPase = fructose 1,6-bisphosphate phosphatase, G6PDH = glucose 6-phosphate dehydrogenase, G6PI = glucose 6-phosphate isomerase, GAPDH = glyceraldehyde 3-phosphate dehydrogenase, GluK = glucokinase, GlyDH = glycerol dehydrogenase, GlyP = glycogen phosphorylase, IDH = isocitrate dehydrogenase, mDH = Malate dehydrogenase, OR = oxidoreductase, PEPS = phosphoenolpyruvate synthetase, PEPCK = phosphoenolpyruvate carboxykinase, PFK = phosphofructokinase, PGAM = phosphoglyceratemutase, PGI = phosphoglucose isomerase, PGK = phosphoglyceratekinase, PGM = phosphomannomutase, PyrK = pyruvate kinase, PPDK = pyruvate phosphate dikinase, TCA = tricarboxylic acid, TPI = triose 3-phosphate isomerase.

Organism	Class	Enzyme	Pathway
***P. putida* [306]**	γ-Proteobacterium	PEPS	Glycolysis
***T. vaginalis* [74]**	Parabasalia	GluK	Glycolysis
		GlyP	Gluconeogenesis
***Synechococcus* sp. [307]**	Synechococcus	PGM	Glycolysis
		PEPS	Glycolysis
		AH	TCA cycle
***S. pneumoniae* [308]**	Bacilli	PGK	Glycolysis
		PyrK	Glycolysis
		PEPCK	TCA cycle
***E. coli* [71]**	γ-Proteobacterium	PGI	Glycolysis
		PGK	Glycolysis
		PGAM	Glycolysis
		ENO	Glycolysis
		PyrK	Glycolysis
		PEPS	Glycolysis
		G6PI	Glycolysis
		PGM	Glycolysis
		IDH	TCA cycle
		PEPCK	TCA cycle
		AH	TCA cycle
		MDH	TCA cycle
		3-PGDH	Amino acid metabolism
***L. lactis* [72]**	Bacilli	PGI	Glycolysis
		FBPA	
		GAPDH	
		PGK	
		PGAM	
***T. thermophilus* [76]**	Deinococci	Three PGMs	Glycolysis

		FBPA	TCA cycle
		MDH	TCA cycle
		Succinyl-CoA synthase α-subunit	TCA cycle
		IDH	TCA cycle
***H. salinarium* [46]**	Halobacteria	PyrK	Glycolysis
		PPDK	Glycolysis
		Pyruvate:ferredoxin OR	TCA cycle
		Succinate dehydrogenase β-subunit	TCA cycle
		IDH	TCA cycle
		Three PGMs	Glycolysis
		Four PGAMs	Glycolysis
***R. palustris* [309]**	α-Proteobacterium	FBPA	Glycolysis
		PGAM	Glycolysis
		PGM	Glycolysis
		PEPCK	TCA cycle
		Succinyl-CoA synthase α-subunit	TCA cycle
		ALDHs	CCM
		ADHs	CCM
***C. jejuni* [310]**	ε-Proteobacterium	PyrK	Glycolysis
		FBPase	
		FBPA	
		TPI	
		PGM	
		PGI	
		ENO	
***S. coelicolor* [311]**	Actinobacteria	PFK	Glycolysis
		PGAM	Glycolysis
		Pyruvate DH	TCA cycle
		G6PDH	Glycolysis
		PGM	Glycolysis
		GlyDH	Glycerol metabolism
		Succinyl-CoA synthase β-subunit	TCA cycle
***L. monocytogenes* [312]**	Bacilli	FBPA	Glycolysis
		GAPDH	
		PGAM	
		PGI	
		TPI	
		ENO	
		PyrK	
***C. acetobutylicum* [313]**	Clostridia	PGK	Glycolysis
		PyrK	Glycolysis

		GAPDH	Glycolysis
		G6PI	Glycolysis
		Malic enzyme	TCA cycle
		Pyruvate:ferrodoxin OR	TCA cycle
		ADH/ALDH	CCM
***B. subtilis* [70]**	Bacilli	PGI	Glycolysis
		FBPA	
		TPI	
		GAPDH	
		PGK	
		PGAM	
		ENO	
		PyrK	
***M. pneumoniae* [75]**	Mollicutes	FBPA	Glycolysis
		PGM	Glycolysis
		PFK	Glycolysis
		TPI	Glycolysis
		GAPDH	Glycolysis
		PGK	Glycolysis
		PyrK	Glycolysis
		Glycerol kinase	Glycerol metabolism
		Pyruvate DH E1-E3	TCA cycle
***M. tuberculosis* [314]**	Actinobacteria	FBPA	Glycolysis
		PGM	Glycolysis
		IDH	TCA cycle
		Citrate synthase	TCA cycle
		MDH	TCA cycle
		2-oxoglutarate DH	TCA cycle
		Malate OR	TCA cycle
		ALDHs	CCM
		Glycogen phosphorylase	Gluconeogenesis
		3-PG DH	Amino acid metabolism
		Glutamate DH	Amino acid metabolism

6.2 Publikationsliste

Referierte Publikationen

Kort J. C., Esser D., Pham T. K, Noirel J., Wright P. C., Siebers B. (2013). A Cool Tool for Hot and Sour Archaea: Proteomics of *Sulfolobus solfataricus*. Accepted in Journal of Proteomics.

Reimann J., Esser D., Orell A., Amman F., Pham T. K., Lindås A. C., Bernander R., Wright P. C., Siebers B., Albers S. V. (2013). (Cren)Archaeal Signal Transduction: *In Vivo* Impact of Protein Phosphatase Deletions on Cell Size, Motility and Energy Metabolism in *Sulfolobus acidocaldarius.* Submitted in Molecular and Cellular Proteomics (under revison).

Kouril T., Esser D., Kort J., Westerhoff H., Siebers B., Snoep J. (2013). Intermediate instability at high temperature leads to low pathway efficiency for an in vitro reconstituted system of gluconeogenesis in *Sulfolobus solfataricus.* Revised manuscript submitted to FEBS Journal (30.05.2013)
Esser D., Siebers B. (2013). Atypical protein kinases of the RIO family in archaea. Biochem Soc Trans. 2013 Feb 1. PMID:23356318

Esser D., Kouril T., Talfournier F., Polkowska J., Schrader T., Bräsen C., Siebers B. (2013). Unraveling the function of paralogs of the aldehyde dehydrogenase super family from *Sulfolobus solfataricus*. Extremophiles. 2013 Jan 8. PMID:23296511

Esser D., Pham T K., Reimann J., Albers S. V., Siebers B., Wright P. C. (2012). Change of carbon source causes dramatic effects in the phospho-proteome of the archaeon *Sulfolobus solfataricus*. J Proteome Res. 2012 Oct 5;11(10):4823-33. doi: 10.1021/pr300190k. Epub 2012 Sep 14. PMID:22639831

Esser D, Kouril T, Zaparty M, Sierocinski P, Chan PP, Lowe T, Van der Oost J, Albers SV, Schomburg D, Makarova KS, Siebers B (2011). Functional curation of the *Sulfolobus solfataricus* P2 and *S. acidocaldarius* 98-3 complete genome sequences. Extremophiles. 2011 Nov;15(6):711-2. doi: 10.1007/s00792-011-0392-1. Epub 2011 Sep 13. PMID: 21912952

Zaparty M., Esser D., Gertig S., Haferkamp P., Kouril T., Manica A., Pham T. K., Reimann J., Schreiber K., Sierocinski P., Teichmann D., van Wolferen M., von Jan M., Wieloch P., Albers S. V., Driessen A. J., Klenk H.-P. P., Schleper C., Schomburg D., van der Oost J., Wright P. C., and Siebers B. (2009)."Hot

standards" for the thermoacidophilic archaeon *Sulfolobus solfataricus*. Extremophiles. 2010 Jan;14(1):119-42. Epub 2009 Oct 4.

Poster

Phosphoproteome analysis of the Archaeon Sulfolobus acidocaldarius with a special focus on central carbohydrate metabolism, Annual Conference of the Association for General and Applied Microbiology (VAAM), **D. Esser**, J. Reimann, T. K. Pham, P. C. Wright, S. V. Albers, B. Siebers, March 2013, Bremen, Germany

Change of carbon source causes dramatic effects in the phospho-proteome of the Archaeon *Sulfolobus solfataricus*, International Study Group on Systems Biology (ISGSB), **D. Esser**, T. K. Pham, J. Reimann, S. V. Albers, P. C. Wright, B. Siebers, September 2012, Ameland, The Netherlands

Signal transduction in the thermoacidophilic creanarchaeum *Sulfolobus acidocaldarius*: Investigation of the protein phosphatases PP2A and PTP, Molecular Biology of Archaea III, **D. Esser**, S. V. Albers, B. Siebers, July 2012, Marburg, Germany

Hot protein phosphorylation: carbon source dependent phosphor-proteome mapping in *Sulfolobus solfataricus* P2, poster presentation, Annual Conference of the Association for General and Applied Microbiology (VAAM), **D. Esser**, T. K. Pham, J. Reimann, S. V. Albers, P. C. Wright, B. Siebers, April 2011, Karlsruhe, Germany

Life in the hot lane: How (hyper-)thermophiles cope with the instability of metabolites, poster presentation, WE-Heraeus-Seminar 472 on 'Biothermodynamics of Metabolic and Ecological Networks', T. Kouril, A. Kolodkin, **D. Esser**, L. Abel, M. Zaparty, H. V. Westerhoff, B. Siebers, January 2011, Bad Honnef, Germany

Network reconstruction: From Genome to Function, poster presentation, WE-Heraeus-Seminar 472 on 'Biothermodynamics of Metabolic and Ecological Networks', **D. Esser**, T. Kouril, M. Zaparty, B. Siebers, January 2011, Bad Honnef, Germany

Life in the hot lane: How (hyper-)thermophiles cope with the instability of metabolites, poster presentation, Annual Conference of the Association for General and Applied Microbiology (VAAM), T. Kouril, **D. Esser**, M. Zaparty, B. Siebers, SulfoSYS consortium, March 2010, Hannover, Germany

Biochemical analyses of the central carbohydrate metabolism of *Sulfolobus solfataricus* in response to temperature change, poster presentation, Gordon Research Conferences on Archaea, T. Kouril, **D. Esser**, N. Eling, B. Siebers, SulfoSYS consortium, July 2009, Waterville Valley Resort, USA

SulfoSYS- Identification of players in the CCM network of *S. solfataricus*, poster presentation, 2nd Evaluation Conference SysMO, T. Kouril, **D. Esser**, M. Zaparty, P. Haferkamp, J. Reimann, S.-V. Albers, B. Siebers, May 2009, Vienna, Austria

Role of members of the aldehyde dehydrogenase superfamily in *Sulfolobus solfataricus*, poster presentation, Annual Conference of the Association for General and Applied Microbiology (VAAM), **D. Esser**, T. Kouril, M. Zaparty, B. Siebers, March 2008, Bochum, Germany

Vorträge

New phosphoproteomics: PACiFIC reveales dramatic phosphoproteome changes in *Sulfolobus solfataricus* P2 in response to carbon shift, Gordon Research Conferences: Archaea Ecology, Metabolism and Molecular Biology, July 2011, Waterville Valley, United States of America

Signal transduction in the thermoacidophilic crenarchaeum Sulfolobus acidocaldarius: Investigation of the protein phosphatase PP2A, oral presentation, Annual Conference of the Association for General and Applied Microbiology (VAAM), March 2012, Tübingen, Germany

6.3 Lebenslauf

Name:	Dominik Esser
Geburtsdatum:	24.05.1985
Geburtsort:	Oberhausen
Staatsangehörigkeit:	Deutsch
Familienstand:	Ledig

Studium

04/2010 – 07/2013	Promotion im Fachbereich Chemie, Universität Duisburg-Essen bei Prof. Dr. Bettina Siebers
10/2008 – 03/2010	Master Studiengang Water Science mit Schwerpunkt Umweltchemie, Analytische Chemie und Mikrobiologie, Fachbereich Chemie, Universität Duisburg-Essen
10/2006 – 09/2008	Bachelor Studiengang Water Science mit Schwerpunkt Analytische Chemie, Mikrobiologie und Biochemie, Fachbereich Chemie, Universität Duisburg-Essen
10/2005 – 09/2006	Bachelor Studiengang Chemie, Fachbereich Chemie, Universität Duisburg-Essen

Ausbildung

09/2004– 06/2005	Berufskolleg West, Essen Abschluss Fachabitur in Chemietechnik
09/2001 – 07/2004	Bayer AG/Healthcare, Wuppertal Elberfeld 3 1/2 jährige Ausbildung zum Chemikanten im dualen System Verkürzung der Ausbildung um ½ Jahr
08/1996 – 06/2001	Erich-Kästner Gesamtschule, Bochum Realschulabschluss
09/1995 - 06/1996	Haupschule Altstaden, Oberhausen
09/1991 – 06/1995	Josefschule, Oberhausen

6.4 Erklärung

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

„Protein phosphorylation in Archaea: Central carbohydrate metabolism in *Sulfolobus* species ”

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, im Juni 2013

6.5 Danksagung

Besonderer Dank gilt meiner Doktormutter Prof. Dr. Bettina Siebers für die Bereitstellung des Arbeitsplatzes, ihre Unterstützung und den nötigen Freiraum um mich als Wissenschaftler zu definieren und weiter zu entwickeln. Ich danke ihr auch für ihre Diskussionsbereitschaft und Anregungen ohne die das Verfassen dieser Arbeit nicht möglich gewesen wäre.

Herrn Prof. Dr. Phillip Craig Wright (University of Sheffield) danke ich sehr für die freundliche Aufnahme in seiner Arbeitsgruppe, die Bereitstellung des Arbeitsplatzes und die sehr gute Zusammenarbeit. Mein besonderer Dank gilt an dieser Stelle auch Dr. Trong Khoa Pham, Dr. Saw Yen Ow und Dr. Malinda Salim für die Unterstützung bei der Durchführung der Massenspektrometrie Experimente, sowie für hilfreichen Ratschläge und die Erweiterung meines Horizontes auf diesem Themengebiet.

Ganz besonders danke ich auch Dr. Sonja-Verena Albers (Max Planck Institut, Marburg) für anregende Diskussionen, die Zusammenarbeit und zur Verfügungstellung der *Sulfolobus acidocaldarius* Stämme *Δsaci_pp2a* und *Δsaci_ptp*. Außerdem möchte ich mich noch bei Julia Reimann (Max Planck Institut, Marburg) für die gute Zusammenarbeit bedanken.

Desweiteren bedanke ich mich bei Dr. Francois Talfournier (Nancy University) für die Aktivitäts-Messungen der MSDH und Dr. Jolanta Polkowska (Universität Duisburg-Essen) für die Synthese des Substrates 2,5-Dioxopentanoate.

Ich möchte mich auch besonders bei meinen ehemaligen Bachelor/Master Studenten Deniz Dalcig, Lu Shen, Marina Horstkott und Wolfgang Pfeifer für ihre hervorragende Arbeit im Labor bedanken.

Ein besonderes Dankeschön gilt vor allem den ehemaligen und momentanen Mitgliedern der Arbeitsgruppe Siebers: Anna Hagemann, Julia Kort, Dr. Theresa Kouril, Dr. Patrick Haferkamp, Khoei Matsubara, Silke Jachlewski, Jens Benninghof, Dr. Christopher Bräsen, Sabine Dietel, Thomas Knura, Bernardette Rauch, Dr. Verena Kallnik, Marcel Blum, Katharina Kruse, Frank Schult, Claus-Rüdiger Wallis, Dr. Melanie Zaparty und Dr. Jeannette Marrero Coto, für die Hilfestellung und angenehme Arbeitsatmosphäre. Ganz besonders möchte ich mich an dieser Stelle bei Dr. Theresa Kouril und Dr. Patrick Haferkamp bedanken für die Betreuung während meiner Anfangszeit in der Arbeitsgruppe, für die notwendigen Kaffeepausen um den Kopf frei zu bekommen und für ihre Freundschaft.

Der größte Dank gilt meiner Familie und meinen Freunden. Danke Sebastian und Viktor für eure Freundschaft, nicht nur während der Promotion sondern auch während des Studiums. Ohne euch hätte ich ganz bestimmt häufiger einen klaren Kopf gehabt ☺, aber gerade dafür dass ihr das vermeiden konntet danke ich euch und für vieles vieles mehr. Danke Niko, im Dojo habe ich mit Sicherheit mehr als nur einmal meinen Arbeitsalltag und meine überschüssige Energie an dir

ausgelassen. Deine Funktion als Energie Ventil hat mich davor bewahrt einen zu hohen Blutdruck zu bekommen, oder gar zu explodieren ☺. Ganz besonderer Dank gilt auch meinen Eltern Regina und Olaf, meiner Schwester Nadine und meinen beiden süßen Nichten Mina und Sophie für ihre Unterstützung und ihr Verständnis während der letzten drei Jahre. Vor allem Mina und Sophie danke ich, denn deren gebrabbel, gequassel und rum gekrabbel haben sie es immer wieder geschafft, dass ich die Arbeit für einen kurzen Moment vergessen konnte. Mein größter Dank gilt meiner Verlobten Patrizia. Es gibt so vieles wofür ich ihr an dieser Stelle danken will, aber besonders bedanke ich mich dafür das sie mich stehts unterstützt hat während dieser Zeit und dafür, dass sie immer wieder Verständnis hatte wenn ich am Wochenende nicht ruhig sitzen konnte und so manches Wochenende dann doch wieder im Labor verbracht habe.