

Polyphosphate metabolism and profiling of myristoylation in *Sulfolobus acidocaldarius*

Dissertation

zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften

- Dr. rer. nat. -

vorgelegt von

Svenja Höfmann

Arbeitskreis für Molekulare Enzymtechnologie und Biochemie

Abteilung für Umweltmikrobiologie und Biotechnologie

Fachbereich Chemie

der

Universität Duisburg Essen

Oktober 2022

DuEPublico

Duisburg-Essen Publications online

UNIVERSITÄT
DUISBURG
ESSEN

Offen im Denken

ub | universitäts
bibliothek

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

DOI: 10.17185/duepublico/78420

URN: urn:nbn:de:hbz:465-20230515-112034-5



Dieses Werk kann unter einer Creative Commons Namensnennung - Nicht kommerziell - Keine Bearbeitungen 4.0 Lizenz (CC BY-NC-ND 4.0) genutzt werden.

Die vorliegende Arbeit wurde im Zeitraum von November 2018 bis Oktober 2022 im Labor von Prof. Dr. Bettina Siebers im Arbeitskreis für Molekulare Enzymtechnologie und Biochemie in der Abteilung Umweltmikrobiologie und Biotechnologie an der Fakultät für Chemie der Universität Duisburg-Essen durchgeführt.

Tag der Disputation: 23.02.2023

Gutachter:	Prof. Dr. Bettina Siebers
	Prof. Dr. Markus Kaiser
Vorsitzender:	Prof. Dr. Jochen Gutmann

Table of Content

1.	Introduction.....	1
1.1	Archaea.....	1
1.2	Polyphosphate metabolism	4
1.3	Function of polyphosphates.....	7
1.4	<i>Sulfolobus acidocaldarius</i>	9
1.5	Polyphosphate metabolism in <i>S. acidocaldarius</i>	10
1.6	Post-translational modification of proteins	11
1.7	Lipidation.....	14
1.8	Fatty acids in <i>S. acidocaldarius</i>	17
1.9	<i>In vivo</i> metabolic labeling	18
2	Scope of the thesis	21
3	Chapters.....	22
	Chapter 3.1.....	22
	Discovery of the missing polyphosphate kinase in the thermoacidophilic Crenarchaeon <i>Sulfolobus acidocaldarius</i> by studying polyphosphate metabolism	22
	Chapter 3.2.....	73
	Proteomic analysis of myristoylation in <i>Sulfolobus acidocaldarius</i> by <i>in vivo</i> metabolic labeling.....	73
4	Summary	104
5	Zusammenfassung	106
6	References	109
I.	Acknowledgements.....	128
II.	Curriculum vitae.....	130
III.	Verfassungserklärung.....	131
IV.	List of abbreviations.....	132

1. Introduction

1.1 Archaea

The first discovery of Archaea was reported 140 years ago when Farlow described red halophilic cocci, which were later classified as Archaea [1, 2]. Based on 16S/18S rRNA sequence analysis, Archaea were proposed as the third domain of life only 40 years ago, recognizing the closer phylogenetic relation to the Eukaryotes [3]. Subsequently, the Archaea have been divided physiologically into two kingdoms, the Euryarchaeota, which include methanogens, extremely halophilic, sulfate-reducing, and some (hyper)thermophilic species, and the Crenarchaeota, which encompass mainly (hyper)thermophilic species [4]. Originally, Archaea were thought to occur primarily in extreme environments with respect to pH, temperature, and salinity or possess a unique metabolic trait (methanogenesis), but this later proved not to be true for all archaeal species, so they are ubiquitous. Additionally, they have been shown to be a stable part in the human-associated microbiota and play a role in human disease [5, 6].

With the help of genomic data analysis and culture-independent metagenomics, several deeply branched lineages were discovered within a few years, adding additional phyla to the archaeal domain. This resulted in the establishment of the collective TACK superphylum, which includes several archaeal phyla (Thaum-, Aig-, Cren-, and Korarchaeota) [7]. After this initial implementation, additional phyla were incorporated, namely the Verstraete-, Bathy- and Geoarchaeota. Then following, the superphylum DPANN was launched to refer to Diapherotrites, Parv-, Aenigm-, Nanohalo-, and Nanoarchaeota [8]. The later defined Asgard superphylum (including Loki-, Odin-, Thor- and Heimdallarchaeota) comprises uncultivated Archaea and presumably shares a common archaeal ancestor with the Eukaryotes [9]. Through the investigation of (predominantly) Asgard Archaea, the shared genes with Eukaryotes became increasingly apparent. For instance, the Asgard Archaea encode many key components responsible for the complexity of eukaryotic cells [9]. In particular, a detailed analysis of the genome contents of the Lokiarchaeum revealed eukaryotic signature proteins including cytoskeletal components and moreover, the relation of archaeal cell division proteins to the eukaryotic endosomal sorting complex required for transport (ESCRT-III) as well as various small GTPases [10, 11]. Therefore, it has been proposed that the Eukaryotes evolved within the Asgard superphylum and thus next to the second prokaryotic lineage, the Bacteria and a novel two-domain tree of life was suggested (Fig. 1) [12, 13].

1. Introduction

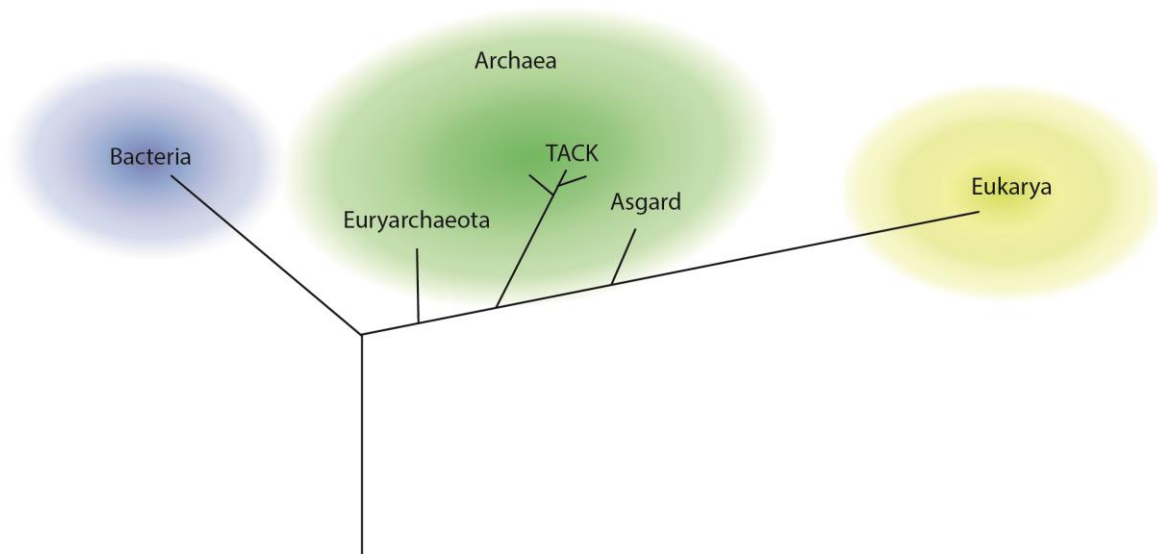


Figure 1: Scheme of the proposed two-domain tree of life.

Phylogenomic approaches provided a two-domain tree of life where Eukaryotes originate from within the Archaea [13].

In agreement to their closer relationship to Eukaryotes, culturable Archaea also share more cellular properties with them, but some are also common with Bacteria. Thus, Archaea contain single circular chromosomes within operon structures and comprise plasmids. Interestingly, cell division is distinct in Cren- and Euryarchaeota, the latter exhibiting a bacterial-like filamenting temperature-sensitive mutant Z (FtsZ)-based cell division and Crenarchaeota possessing a newly discovered machinery, comprising homology to the eukaryotic ESCRT-III (endosomal sorting complex required for transport III) system [14]. In the information-processing mechanism (e.g., replication, transcription, repair, and translation), the proteins involved were found to be more related to eukaryotic ones [15]. Archaea have a simplified version of the eukaryotic transcription machinery, but use bacterial transcriptional regulators to control usage and thus influence elongation and termination rulings [16]. However, Archaea also exhibit a variety of unique properties that they do not share with either Eukaryotes or Bacteria. In particular, the structure of the cell wall of Archaea is universal. First, it does not contain murein and most Archaea feature surface-layer proteins (S-layer) associated with the cytoplasmic membrane [17, 18]. In some cases, pseudomurein was also detected as a component of the cell wall, predominantly occurring in methanogenic species. This oligosaccharide layer is composed of a backbone of *N*-acetylalosaminuronic acid with a β -1,3-linkage to *N*-acetyl-D-glucosamine and thus has a different structure than the murein of Bacteria. Furthermore, extracellular proteins were found to be *N*- or *O*-glycosylated in most archaeal species or lipid modified specifically in Haloarchaea [18]. In Archaea, the motility structures responsible for swimming corresponds functionally to the bacterial flagellum, but the archaellum is structurally resembling to the type IV pilus, unique to Archaea [19, 20].

1. Introduction

The archaeal lipid assembly is very characteristic for this domain of life since their phospholipids exhibit a different composition [21]. Here, the hydrophobic core consists of highly methylated isoprenoid chains attached to a glycerol-1-phosphate backbone via an ether linkage; in contrast, fatty acids (FAs) represent the hydrophobic part in Bacteria and Eukaryotes and are ester-linked to a glycerol-3-phosphate molecule. Due to this special composition of phospholipids, the membrane exhibits significantly higher stability with respect to salinity, pH and temperature, which contributes to the ability of Archaea to occur in extreme environments. In addition, Archaea possess a wide variety of lipids assembling two main lipid cores, C_{20} *sn*-2,3-diacylglycerol diether lipids (DELs or archaeols) and C_{40} glycerol-dialkylglycerol tetraether lipids (TELs) [21]. DELs form a typical bilayer membrane, and two major modifications are found mainly in halophilic and some (hyper)thermophilic Archaea: First, the deviant isoprenoid chain undergoes condensation to a macrocyclic glycerol diether or the second possibility is the appearance of a tetriol diether. The TELs contain two isoprenoid chains, each with 40 carbons attached to two possibly different polar head groups, forming primarily a monolayer. The caldarchaeol (glycerol-dialkyl-glycerol tetraether, GDGT) is the most prominent TEL and can be described as two antiparallel arranged diether lipids connected via their isoprenoid chains. In *Sulfolobales* species, calditoglycerocaldarchaeol (glycerol-dialkyl-nonitol tetraether, GDNT) is mainly found instead of a caldarchaeol, containing mainly a C9 calditol moiety instead of nonitol [22, 23]. The cyclization, thus the formation of up to eight cyclopentane rings is a feature mainly present in (hyper)thermophiles where the degree of cyclization is dependent on species and their cultivation conditions (e.g. growth state and temperature) [24, 25]. Recently, the first isolation of GDNT- β -Glu from *Sulfolobales* species has been published, where GDNT- β -Glu was identified as the major component, consisting mainly of 4 to 6 cyclopentane units, slightly different within these species [23]. The densely packed structure of the hydrophobic core alters the physical properties so that the membrane remains stable even under more extreme conditions such as very high temperatures. This property is an essential aspect that allows hyperthermophilic Archaea to survive under extreme conditions [23]. Since FAs are not part of the archaeal membrane, but are nevertheless present, their function has not yet been fully elucidated. However, it has been reported that some archaeal species have the potential to use FAs as a carbon source, like halophiles [26]. Furthermore, *Sulfolobus acidocaldarius* was demonstrated to grow on hexanoate as the sole carbon source and exhibits a bacterial-like β -oxidation [27, 28]. In addition, FAs can also play a role in protein lipidation, i.e. covalently bound to a protein. Protein modification provides a mean in living cells for rapid response and adaptation to changing environmental conditions or metabolic regulation. A wide variety of post-translational modifications (PTMs) include protein phosphorylation, methylation, glycosylation, ac(et)ylation were already reported for Archaea [29]. In particular, protein phosphorylation in Archaea differs

1. Introduction

among Eury- and Crenarchaeota, the first exhibiting a two-component system (His and Asp phosphorylation) and Hanks-type kinases (Ser, Thr and Tyr phosphorylation), whereas the Crenarchaeota only comprise Hanks-type protein phosphorylation [30]. This work focuses on myristoylation, which describes the covalent attachment of a C14 FA to a glycine, cysteine, or lysine residue of the protein, and is one of the most commonly studied acylations in Eukaryotes and some Bacteria [31]. To date, myristoylation has only been detected in Euryarchaeota, but not in Archaea, which represents a missing link to Eukaryotes [31]. In this study, myristoylome profiling is performed using bioorthogonal labeling of the crenarchaeal model organism *S. acidocaldarius*.

1.2 Polyphosphate metabolism

Another distinctive feature of Archaea is the presence of unique metabolic pathways that differ from those of Bacteria and Eukaryotes [29, 32]. The presence of methanogenesis is ubiquitous in methanogenic Archaea, which can thus reduce CO₂ with H₂ to methane [33]. However, the Archaea resemble Bacteria in respect to complexity [29, 32, 34], but possess modifications of the classical Embden-Meyerhof-Parnas (EMP) and Entner-Doudoroff (ED) pathways and a very altered pentose phosphate and reversed ribulose monophosphate pathway for pentose degradation [32]. Accordingly, some novel enzymes of different enzyme families are integrated into the archaeal variants such as the EMP and ED pathway, which in particular show a divergence with respect to regulatory potential. Thus for example the classical ATP-dependent phosphofructokinase (PFK) in the EMP pathway, a central control point in glycolysis, is substituted by a non-allosteric ATP-, ADP- and PP_i-dependent PFKs [35-37]. Moreover, the formation of 3-phosphoglycerate from glyceraldehyde-3-phosphate (GAP) is catalyzed by the unidirectional glycolytic GAP dehydrogenase (GAPN) or the GAP: ferredoxin oxidoreductase (GAPOR) in the absence of ATP formation, thus substituting the classical enzymes GAP dehydrogenase (GAPDH)/phosphoglycerate kinase (PGK) functioning in gluconeogenesis [38-40]. The classical ED pathway known for Bacteria is absent in Archaea with a modified branched version comprising the nonphosphorylative and semiphosphorylative branch in most Archaea [34].

Like all living cells, Archaea are exposed to a continuously changing environment therefore rely on the storage of certain compounds to provide e.g. carbon for energy or P_i liberation enabling resistance to nutrient limitations [41-44]. To date, three compounds were described to serve as energy storage compound in Archaea, namely inorganic polyphosphate (polyP), polyhydroxyalkanoates (PHAs) and glycogen. The bioinformatic analysis confirmed the presence of PHAs mainly in halophilic Archaea, whereas polyP was so far reported occur in methanogens but was also observed as granules in other archaeal organisms [45, 46].

1. Introduction

Specifically, for extremophilic organisms, polyP, a polymer comprising up to hundreds of orthophosphates linked by a high-energy phosphoanhydride bond, has been described to possess a function, such as persistence under acidic conditions or in copper tolerance. This suggests that polyP plays multiple roles besides adaptation to stressful environmental conditions and thus serves as energy as well as P_i reservoir (see 1.3) [44, 47, 48].

The presence of polyP has been demonstrated for all three kingdoms, but polyP metabolism differs between them. More specifically, Eukaryotes exhibit an altered polyP metabolism compared to Bacteria and Archaea. In Archaea, most of the polyP-metabolizing enzymes described resemble their bacterial counterparts, however only a few enzymes of this domain were characterized so far [43, 46, 49]. For Bacteria, seven enzymes were reported to be directly linked to polyP metabolism (including polyphosphate kinase (PPK1), polyphosphate kinase 2 (PPK2), polyphosphate:AMP phosphotransferase (PAP), 5'/3'-nucleotidase (SurE), exopolyphosphatase (PPX), NAD kinase (NadK) and polyphosphate glucokinase (PPGK) [50]. Interestingly, the PPGK catalyzes the polyP-dependent formation of glucose-6-phosphate, which is the first enzyme reported to be independent from nucleoside phosphate system, but is mainly absent in Archaea [51]. Bioinformatic analysis revealed that only a few archaeal species possess the entire pathway including all seven enzymes [43, 50] and thus, mainly comprise a functional pathway including the PPK1 typically responsible for polyP synthesis as well as one polyP degrading enzyme (Fig. 2) [43]. Bacterial PPK1 responsible for polyP synthesis was initially described as being widely conserved among the Bacteria as well as the Archaea [51], but was demonstrated to be only moderately absent in Archaea [43, 46]. In addition to PPK1, the structurally different PPK2 is also present in many Bacteria and preferentially catalyzes polyP-dependent NDP/NTP formation (Fig. 2) [52]. In addition to the structural difference, the predicted PPK1 mechanism involves a phosphorus-enzyme intermediate that occurs by autophosphorylation of a histidine residue, allowing subsequent phosphate transfer [53, 54]. Thereby, two histidine residues were found to be essential for the enzymatic activity of PPK1 from *Escherichia coli*. The PPK1 enzyme is known to form a tetramer with 80 kDa monomeric units and is membrane associated [54]. For the PPK2 enzyme, belonging to the P-loop kinase superfamily, a completely different mechanism was predicted than for the PPK1 enzymes [55]. Based on structural analysis the coordination of polyP and ADP by positively charged amino acids was assumed [56]. In addition, two magnesium ions are coordinated by two Asp residues and are involved in the catalytic mechanism by coordinating polyP and the nucleotide. The predicted associative PPK2 mechanism describes a nucleophilic attack of the nucleotide on the activated polyP, whereas a second dissociative mechanism was proposed in which the terminal phosphate of the polyP chain first dissociates followed by nucleophilic attack [54]. Interestingly, most eukaryotic

1. Introduction

organisms lack a PPK2s (except *Dictyostelium discoideum*), but encompass a PPK1 homolog with about 30 % similarity to the *E. coli* PPK1 [50, 51, 57-59]. In lower Eukaryotes such as a vacuolar transporter chaperone (VTC) complex was identified to catalyze the polyP synthesis from ATP [60, 61]. For polyP degradation PPX is the most abundant protein in Bacteria and Archaea known to catalyze Mg²⁺-dependent polyP hydrolysis (Fig. 2) [50, 62]. However in Eukaryotes several enzymes are known to catalyze polyP degradation including a PPX belonging to the phosphoesterase superfamily (DHH) [63] as well as endopolyphosphatases, Ppn1, Ppn2 and Ddp1 were identified in *S. cerevisiae*, the latter being a homolog of the Nudix family member diphosphoinositol-polyphosphate phosphohydrolase (DIPP) present in human [64].

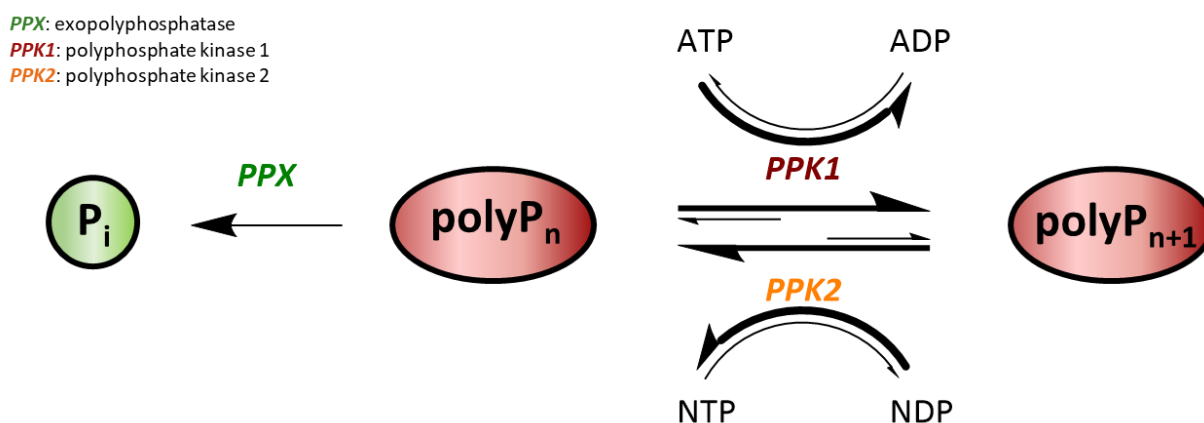


Figure 2: Predominant enzymes catalyzing polyP metabolism in prokaryotes.

PPK1 (red) enzymes catalyze the polyP synthesis by transferring the γ -phosphate from ATP to polyP. The enzymes of the second PPK family (PPK2) are mainly catalyzing the reverse reaction, e.g. NTP synthesis transferring the terminal phosphate from polyP to form NTP. The terminal phosphoanhydride bond of polyP is hydrolyzed by PPX to release free phosphate [52].

The genome organization as polyP operon containing *ppx* and *ppk* genes has been firstly demonstrated in *E. coli* [62]. The presence of two possible promoters (one each of *ppx* and *ppk*) was suggested, and a co-regulation of both genes has been assumed [65]. Moreover, a *ppk-ppx* deletion mutant showed a reduced P_i uptake rate by the low-affinity Pit (phosphate-inorganic transporter) system [51]. Several works also demonstrated some links between polyP synthesis and the phosphate (Pho) regulon, the predominant regulatory mechanisms for inorganic P_i management [51, 66]. In *Streptomyces lividans*, an increased *ppk* expression was detected for P_i-limiting compared to P_i-sufficient conditions and was demonstrated to be regulated by the two-component system PhoR/PhoP [66]. Moreover, an *E. coli* mutant lacking a phosphohistidine phosphatase SixA (signal inhibitory factor X) showed a reduced growth on glycerol minimal medium, independent from the histidine kinase (ArcB) [67]. It was suggested, that SixA modulated phosphorylation of the nitrogen-related phosphotransferase system (PTS^{Ntr}), a conserved bacterial pathway regulating metabolic processes through histidine phosphorylation [67]. The investigation of

1. Introduction

adenylate cyclases class IV (CyaB) revealed a new family also including thiamine triphosphatases. The members of the CyaB, thiamine triphosphatase (CYTH) domain family are known to convert triphosphorylated substrates and exhibit a two-divalent metal cation dependency. It is assumed that these enzymes are related to nucleotide and polyphosphate metabolism. Interestingly, it was reported, that the CYTH domain can occur fused predominantly to the conserved histidine α -helical domain (CHAD), which was demonstrated to be co-localized with polyP granules in *Ralstonia eutropha* [68]. A crystal structure of a CHAD domain with a polyP₉ molecule bound in the center from *Chlorobium tepidum* is available and presents a similar fold as the VTC4-polyP structure from *S. cerevisiae* [69]. In several bacterial and archaeal genomes a conserved gene neighborhood of encoding CYTH, CHAD domains with enzymes involved in nucleotide and polyphosphate metabolism is present [70]. Focusing on *Sulfolobus*, the operon contains the *sixA*, gene followed by the *ppx*, two thymidylate kinase genes and a gene encoding a CHAD domain-containing protein. The gene coding for a CYTH domain containing protein is located at a different position in the genome (*saci_0718*). This enzyme annotated as triphosphate tunnel metalloenzyme (TTM) revealed a substrate preference for triphosphate over NTPs *in vitro* [71]. The function as triphosphatase was already reported for TTM enzymes from *Arabidopsis thaliana* and *Clostridium thermocellum* [72, 73]. The involvement of SaTTM in polyP metabolism and thereby in oxidative stress response has been suggested [71].

1.3 Function of polyphosphates

PolyP was already reported for Eukaryotes, Bacteria and Archaea (Fig. 3). The polymers were firstly isolated from yeast in 1890 and later on mainly studied by the Arthur Kornberg group. Despite its predicted function as energy source of the cell [74] and P_i reservoir [75], the role of polyP as energy source in a pre-ATP world is discussed because polymerization is triggered at high temperatures by dehydration and condensation of orthophosphates [76] and polyP formation was demonstrated under prebiotic conditions, but the resulting yield was very low [77]. The resulting polymer can provide energy by hydrolysis of the phosphoanhydride bond and due to its lower complexity it has been suggested that it may have originated before the ATP molecule. In agreement, some soluble polyphosphates were found in fumaroles at Mount Usu in Japan, including pyro- and triphosphates that can induce divalent ion exchange and in addition, a relation with amino acids is under discussion [78, 79]. Thus, it is suggested that a reaction together with adenosine would have resulted in AMP and subsequently ATP could have derived further acting as metabolic energy source [79]. Another indication for this assumption is the presence of polyphosphate metabolizing enzymes present in most organisms. Despite its involvement in energy charging, several other molecular functions have

1. Introduction

been reported for polyP i.e. to play a role in metal ion chelation, buffering against alkali ions, regulation of development (e.g. fruit body and spore formation), stress and survival, act as channel for DNA entry as well as being a P_i reservoir and is part of cell capsule in *Neissaria* species [75].

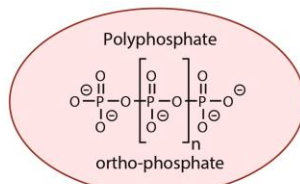
In mammals several additional functions have been reported e.g. regulation in blood coagulation, involvement in cancer progression, function as gliotransmitter and also to act as molecular chaperone [80]. For other eukaryotic organisms it was demonstrated, that polyP takes part in resistance to DNA damage. Moreover, it was described to be associated with membrane channel/transporters as well as its involvement in transcription regulation by decreasing RNA polymerase 1 activity [81]. For *S. cerevisiae* it was demonstrated that a high polyP concentration leads to sol-gel transition [82].

Detailed studies in *E. coli* revealed a physiological function of polyP to prevent protein aggregation [83]. The dependence of chain length was also demonstrated, resulting in long-chain polymers exhibiting higher chaperone efficiency *in vitro* and blocking heat-induced protein aggregation [83]. Another function mainly studied in Bacteria is the involvement in cellular response to various stresses. For example, it was reported that polyP-metal complexes play several roles in resistance to oxidative stress [84].

In addition, a PTM by polyphosphorylation of internal lysine residues was observed in yeast. It was demonstrated to occur non-enzymatically in a polyacidic serine and lysine (PASK)-rich cluster of nuclear signal recognition 1 (Nrs1) and topoisomerase 1 (Top1) proteins. The polyP modification prevents the interaction of both proteins and lowers Top1 activity [85].

In Archaea, the function of polyP has been studied only in some organisms and is not well understood. Nevertheless, some specific functions have been reported and will be described in chapter 3.1 in more detail [47-49, 86-88].

1. Introduction



"Forgotten polymer" [75]

Bacteria	Archaea	Eukaryotes	Mammals
<ul style="list-style-type: none"> Component of DNA-Channel Transcription and translation regulation ATP substitute, energy source P_i homeostasis Stress resistance Stationary phase survival Biofilm formation Chelator of metal ions Persistence Cellular structures Cell motility Virulence Colonization Quorum sensing 	<ul style="list-style-type: none"> Biofilm formation P_i homeostasis Cell motility Cell adhesion Metal tolerance 	<ul style="list-style-type: none"> DNA synthesis and repair Associated with membrane channel/transporters Independent source of ATP (affecting activity of ATP synthase and respiratory chain) Sol-gel transition Buffer against alkali Function of a chaperone Transcription regulation Post-translational modification Arginine sequestration Development and germination 	<ul style="list-style-type: none"> Blood coagulation Cancer progression Glutransmitter Signaling Mitochondrial activity Bone mineralization Inflammation P_i homeostasis Infection Predation Cytokinesis Stress resistance

Figure 3: Functions of polyP among the three kingdoms, Bacteria, Archaea and Eukaryotes. The functions of polyP described for each domain and mammals are shown inside the boxes (blue: Bacteria, green: Archaea, yellow: Eukaryotes, orange: Mammals) [47-49, 75, 80-82, 86-90].

1.4 *Sulfolobus acidocaldarius*

The crenarchaeal organism *S. acidocaldarius* has been firstly isolated from acidic hot springs in the Yellowstone National Park (United States of America) [91]. It belongs to the obligate aerobic and thermoacidophilic Crenarchaeota, as it exhibits optimal growth at around pH 2 - 5 and 70 - 80 °C. This organism is able to grow on various simple organic compounds heterotrophically such as sucrose, D-glucose, D-xylose, L-arabinose [92]. To date, numerous studies on the carbon as well as energy metabolism of this Crenarchaeon are available [32, 93-95]. Important features are the availability of the complete genome including updated gene annotations [96] and the provided versatile genetic toolbox that allows *S. acidocaldarius* to be used as a model organism [97, 98]. Therefore, a *pyrEF* deletion mutant was constructed resulting in the uracil auxotroph strain MW001 for selection. The established genetic system allows the creation of markerless deletion mutants as well as the ectopic integration and expression of foreign DNA [97]. In addition, it is discussed that *Sulfolobus* serves as a source of stable enzymes and biomolecules due to its specific features, and thus possess a great potential for industrial application [92]. In summary, *S. acidocaldarius* is well suited to function as model organism for further biochemical studies.

1.5 Polyphosphate metabolism in *S. acidocaldarius*

One of the first publications dealing with archaeal polyP metabolism described a glycogen-bound polyphosphate kinase from *S. acidocaldarius* [99]. However, a following investigation revealed its enzymatic function as glycogen synthase [100]. Since then, the search for the real PPK remains ongoing and has not been solved for the Crenarchaeota to date. Further publications addressed the characterization of PPX enzymes, like the one from *Saccharolobus solfataricus* [101]. This enzyme showed a high similarity to bacterial PPXs and exhibits the conserved motifs. The biochemical characterization demonstrated that SsoPPX is able to hydrolyze polyP chains with a clear preference for long-chain polyP with up to 800 residues [101]. A gene encoding the homolog was identified in the genome of *S. acidocaldarius* (*saci_2018*) [46]. A *ppx* overexpression strain was constructed, predicted to lack polyP a significant effect on biofilm formation was observed, with thinner biofilm as well as a lower amount of exopolysaccharides (EPS) produced (Fig. 4). Moreover, the *ppx* overexpression strain showed a reduced adhesion to glass surface and the motility assay resulted in only 30 % remaining swarming motility compared to the background strain. Since all phenotypes are related to the surface structures, i.e. the archaellum and adhesive pili, it is assumed, that polyP might affect the function of these structures [88]. However, PPK activity has never been confirmed for *S. acidocaldarius*. Therefore, one aim of this study was the characterization of the enzyme encoded by *saci_2018* gene. Moreover, the elucidation of polyP synthesis in the Crenarchaeon *S. acidocaldarius* is the aim of the first part of this work (Fig. 4).

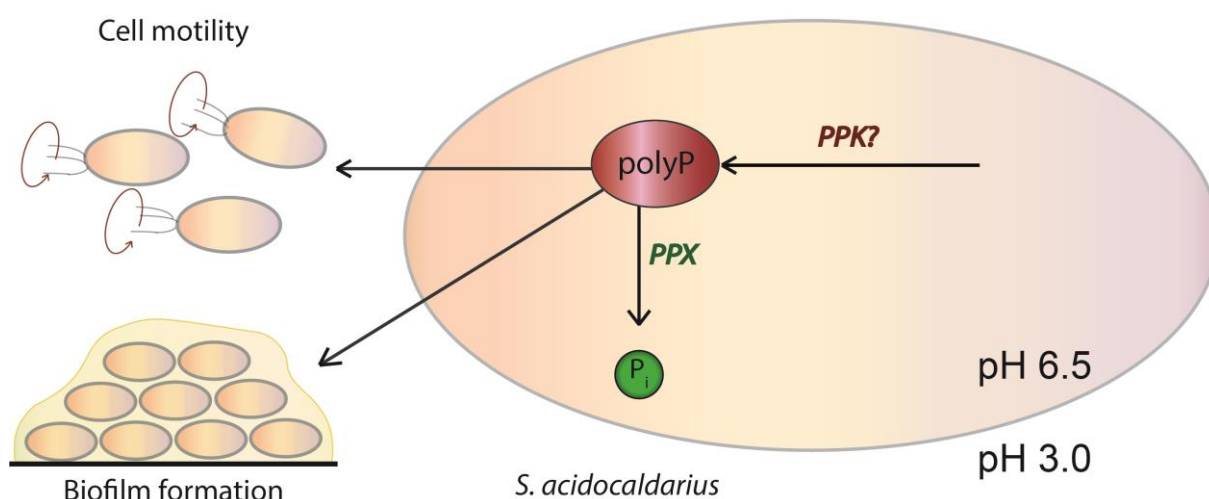


Figure 4: Current knowledge on the function and metabolism of polyP in *S. acidocaldarius*. PolyP in *S. acidocaldarius* cells. The PPX has been identified to degrade polyP and a *ppx* overexpression strain lacks polyP. The lack of polyP resulted in a decreased biofilm formation and less swarming motility was observed [88]. So far, the enzyme(s) involved in polyP synthesis has not been elucidated.

1. Introduction

The study on polyP metabolism in *S. acidocaldarius* is presented in chapter 3.1. First, PPX activity of recombinant Saci_2018 was demonstrated revealing a substrate preference for long chain polyP as well as comprising a very high specific activity (990 U/mg for polyP₇₀₀) compared to characterized bacterial or archaeal homologs. As described above, two thymidylate kinase (dTMPK) genes (*saci_2019*, *saci_2020*) are localized adjacent to the *ppx* gene and were identified to exhibit PPK activity. In addition, a third gene annotated as dTMPK (*saci_0893*) was the only candidate comprising dTMPK activity. Interestingly, the PPK contains one active (Saci_2019) and one inactive subunit (Saci_2020), as the latter itself did not exhibit PPK activity but was able to significantly increase the activity of the homomeric PPK (Saci_2019). In addition, the reversibility of heteromeric PPK (Saci_2019 and Saci_2020) was demonstrated revealing a clear preference for polyP-dependent NTP formation favoring ADP over GDP as substrate. Quantitative ³¹P NMR measurements clearly demonstrated polyP synthesis by PPK, as ADP inhibition could be bypassed by adding an ATP recycling system. Furthermore, an ADP/ATP (25 %/ 75 %) equilibrium was demonstrated to occur regardless of the starting point of the reaction. Based on the features of the heteromeric PPK resembling class I PPK2s and comprising high similarity with dTMPKs, it is thus termed dTMPK-like PPK2 class I.

1.6 Post-translational modification of proteins

Post-translational protein modifications (PTMs) refer to the biochemical process in which single amino acids of a protein are covalently modified. They occur in many possible variants and are known to regulate protein activity and function in response to various changes in extracellular and intracellular conditions. In addition, these modifications result in a greater diversity of side chains, allowing for a broader range of protein properties, including for example protein-protein interaction, protein localization, regulation of enzyme activity, and control of protein lifespan [102].

A wide distributed PTM is phosphorylation, taking place mainly at the hydroxyl group of serine, threonine or tyrosine residues, but is also present on histidine and aspartate, cysteine, arginine, lysine. The reversible protein modification is catalyzed by protein kinases (PKs) and plays a distinct role in signal transduction in Eukaryotes, Bacteria as well as in Archaea. The modification can be reversed by protein phosphatases (PPs). For the attachment of a phosphate group, the PK requires a phosphoryl donor, which is usually a nucleoside triphosphate (regularly ATP), while the removal of the modification by the PP is a hydrolysis reaction [103]. As a distinctive feature in Archaea, Crenarchaeota and Euryarchaeota possess Hanks-type kinases that catalyze Ser, Thr, and Tyr phosphorylation, but only Euryarchaeota

1. Introduction

additionally exhibit a bacterial-like two-component system (TCS) in which His and Asp phosphorylation is present [30].

Another PTM, methylation involves the transfer of a carbon group and can mask the negative charge of carboxylate side chains or increase hydrophobicity when occurring on lysine residues. Moreover, glycosylation can be present as *N*- or *O*-linked glycosylation and can differ in glycan structures [104]. Initial studies reported that *N*-glycosylation was restricted to Eukaryotes, but this was refuted by the detection of attached oligosaccharides on S-layer glycoproteins in *Halobacterium salinarium* [105]. Meanwhile, the presence has also been reported for halophilic as well as methanogenic Archaea and *S. acidocaldarius* [106]. In particular, much attention is paid to the *N*-glycosylation of the S-layer glycoproteins of *S. acidocaldarius*, since the detected *N*-glycans resemble those of Eukaryotes due to the structure starting from two *N*-acetylglucosamines and branching characteristics [29]. Within the *Sulfolobus* species, the composition of the *N*-glycan differs in the quantity of D-mannose and D-glucose units, but interestingly includes 6-sulfoquinovose, which was previously detected only in photosynthetic Bacteria [29, 107]. Furthermore, *N*-glycosylation appears to be essential, as attempts to generate deletion mutants of genes encoding the enzymes involved failed or resulted in restricted growth on media with high salt concentrations [29].

Protein acetylation defines the covalent attachment of short-, medium and long-chain fatty acids, with the shortest version of this modification being called acetylation. This modification occurs in two major types regarding the acetylation site. The acetylation of the N-terminus of the protein (*N* α -acetylation) can arise co-translational, for example by ribosome-associated enzymes as well as post-translational and as either complete or partial [108, 109]. N-terminal acetyltransferases (NATs) belong to the Gcn5-related NATs (GNAT) protein superfamily [110] and catalyze the transfer of an acetyl group onto the α -amino group of the polypeptide by acetyl-CoA in Eukaryotes [111]. This acetylation group covers the negative charge of the amino group and on the one hand blocks the N-terminus from further modifications but on the other hand is a subject to propionylation or arginylation modifications [109, 112, 113]. In Eukaryotes various NATs are known and exhibit different substrate specificities with respect to the first amino acids present in the polypeptide chain [114]. In comparison, N-terminal acetylation is rarely observed in Bacteria [115]. In *E. coli*, it has been observed on ribosomal proteins and occurs post-translationally [116]. Some additional studies also showed that acetylation can affect mRNA transfer efficiency by various mechanisms [117]. The first studies on N-terminal acetylation in Archaea were carried out in halophiles and were followed by proteomic analysis in *H. salinarum* and *Natronomonas pharaonis*. These studies revealed that about 15 % of proteins were N-terminal acetylated and about 60 % of the proteins undergo methionine cleavage [118, 119]. For the Crenarchaeon *S. solfataricus*, this modification was reported for thermosome, proteasome as well as ribosome subunits. Additionally, a conserved NAT was

1. Introduction

identified to acetylate N-terminal sequences *in vitro* as well as the structure was solved resembling eukaryotic NATs [120, 121].

Alternatively, an internal lysine residue can also be acetylated, resulting in N ϵ -acetylation. The acetyl group masks the charged lysine residue and thus reduces the DNA binding affinity, e.g. for histones. Thus, gene expression can also be decisively affected by N ϵ -lysine acetylation. So far, several enzymes were identified to catalyze lysine acetylation and deacetylation and are termed histone acetyl transferases (HAT) and histone deacetylases (HDAC) although they do not modify histones alone [115, 122, 123]. The HAT enzyme families were grouped regarding their catalytic domains, where the most prominent GNAT superfamily includes Gcn5, PCAF, Eip3 Hat1, Hpa2 and Nut1 enzymes [123, 124]. In addition, the MYST family was initially named according to its four founding members in yeast and mammals: MOZ, YBF2/SAS3, SAS2, Iip60 [125]. The MYTH family enzymes contain the conserved MYTH domain, which comprises an acetyl-CoA binding motif and a zinc finger as well as they and have been reported to regulate long-range/chromosome-wide gene expression [124, 126, 127]. In Bacteria, both N α - and N ϵ -acetylation is catalyzed by homologs of the yeast GNAT by using acetyl-CoA, which can be generated e.g. from acetate by acetyl-CoA synthetase [117]. Acetylation can also occur non-enzymatically on N ϵ -lysine by high-energy acetyl-phosphate [117]. Bacterial genomes have been reported to contain a large number of genes encoding GNATs that may act on different targets [117]. Structural analysis of GNATs demonstrated a conserved core domain composed of β -sheets, whereas various reported domain organizations led to the definition of four types (I – V) of GNAT [117].

On the other hand, the histone deacetylases are more diverse and comprise classes of the classical HDAC family as well as the NAD⁺-dependent Sir2 family. In Eukaryotes, the classical HDAC family enzymes contain a zinc-dependent active site and are divided into three classes. Class I HDAC (human HDAC 1, 2, 3, 8) is closely related to the transcriptional regulator Rpd3 from *S. cerevisiae*, and several members act as transcriptional corepressors. Class II (human HDAC 4, 5, 6, 7, 9, and 10) comprises domains resembling HDA1, another deacetylase from *S. cerevisiae*, and exhibits tissue-specific expression in mammals. In 2004, class IV HDACs were defined, related to human HDAC11, which differs from the classical class I and II HDACs, whereas sirtuins (Sir2, silent information regulator 2) display the class III HDACs [128]. Sirtuins catalyze the NAD⁺-dependent deacetylation of an internal lysine residue by releasing 2'-O-acetyl-ADP-ribose as well as nicotinamide and were identified in Eukaryotes, Bacteria and Archaea revealing a high conservation of this class of deacetylases. In mammals, seven members of sirtuins were identified differing in subcellular localization and enzymatic activity [129]. Sirtuins are essential for a molecular linkage between regulation of cellular functions and metabolism. For instance, they were reported to play a role in stress resistance, vascular biology and tumorigenesis in mammals [130]. In Bacteria as well as in Eukaryotes,

1. Introduction

the regulation of the acetyl-CoA synthetase (ACS) by acetylation was described [117, 131]. For example in Bacteria, where a protein acetyltransferase and a CobB (sirtuin) enzyme were demonstrated to be involved in the acetylation and deacetylation process [131-133]. The genome of *Bacillus subtilis* encompassed only one gene encoding sirtuin, where the enzyme was shown to be active on various acyl-lysine substrates with decreasing enzyme activity for increasing acyl-chain length [134, 135].

In Archaea, one of the first studies in *Haloferax volcanii* revealed three genes encoding acetyltransferases belonging to the Gcn5 family (Pat1, Pat2, Elp3), as well as two genes coding for deacetylases, a Sir2 and a Hda1 homolog. During construction of deletion mutants, the single acetyltransferase mutants ($\Delta pat1$, $\Delta pat2$, $\Delta elp3$) were possible, but the double mutant (*pat2*, *elp3*) was not viable. In addition, the *sir2* deletion mutant could grow normally, but the deletion of the *hda1* gene was lethal [136]. A Sir2 deacetylase has also been studied in the Crenarchaeon *S. solfataricus* in the context of deacetylation of the chromatin protein Alba (acetylation lowers binding affinity). Alba was obtained to harbor two acetylation modifications, an N-terminal one as well as a N ϵ -acetylation at K16. The K16 acetylation could be removed by the NAD⁺-dependent Sir2 deacetylase, whose antagonist the protein acetyltransferase (SsoPat) was later identified to acetylate K16 *in vitro* [137, 138]. Acetylation of Alba by SsoPat decreased the DNA binding affinity of Alba, and in addition, a lower affinity of acetylated Alba for double-stranded RNA was also noted [138]. In contrast, another study described that Alba is not acetylated but methylated [139]. Detailed proteomic analysis in *Sulfolobus islandicus* and *Thermococcus gammatolerans* revealed the presence of protein acetylation in Archaea and are described in chapter 3.2 in more detail [140, 141].

1.7 Lipidation

Protein lipidation is mainly investigated in Eukaryotes and is known to influence membrane interactions of proteins as well as protein-protein interactions, protein stability, enzymatic activities, membrane association and subcellular localization [111, 142]. The attached lipids may include fatty acyl chains, but also isoprenoid chains, which are mainly attached to nucleophilic side chains like cysteine, serine or lysine residues, but can also occur N-terminally. While cysteine prenylation, palmitoylation, and N-terminal glycine myristoylation are the most studied lipid modifications, there are other reported modifications such as serine acylation, N-terminal cysteine palmitoylation, lysine acylation, and C-terminal cholesterol modification [142]. With particular attention to lysine acylation with various long-chain acyl-CoAs acting as donors has been reported, whereas short-chain acylation, i.e., acetylation,

1. Introduction

has already been described above. Acylation includes propionyl-, butyryl-, succinyl-, crotonyl- and long-chain fatty acyl-CoA like myristoyl-CoA [111, 142].

In this work, focus will be first on myristoylation, which defines the covalent attachment of a C14 FA, the presence in Bacteria, Eukaryotes and also Archaea has been reported [111, 143]. The modification is known to occur on the N-terminus of the protein, where it is linked via an amide bond to a cysteine or glycine residue. Moreover, it can also be attached to an internal lysine or cysteine residue by formation of an amide or thioester bond (Fig. 5) [31].

First addressing N-terminal myristoylation, which is mainly investigated in Eukaryotes and occurs on the N-terminus of the protein (termed G-myristoylation for glycine present at the N-terminus) and can take place co- or post-translationally [31]. A conserved consensus sequence (G-XXX-S/T/C) was identified for the co-translational modification after cleavage of the initial methionine by methionine aminopeptidase [144]. This myristoylation is catalyzed by a protein N-myristoyltransferase (NMT) belonging to the GNAT superfamily [145]. This enzyme was investigated in many eukaryotic organisms, like plants, yeast, fungi and human. The NMT from yeast reveals a Bi-Bi mechanism, where the activated myristoyl-CoA binds to the apoprotein and leads to a conformational change of NMT to enable binding of the target protein. Through a nucleophilic attack of the N-terminal glycine on the thioester bond of myristoyl-CoA, the acyl residue is transferred and subsequently, free CoA and the myristoylated peptide are released [146]. By resolving several structures, a deeper insight into the mechanism could be gained, which, for example, explains the substrate specificity by the binding of the C5 and/or C6 unit of myristoyl-CoA resulting in the orientation of the remaining acyl-chain [147-149]. However, much lower enzymatic activity has been reported when palmitoyl-CoA acts as an acyl donor [148]. The importance of this enzyme was demonstrated, for instance, by the fact that it has been proven to be essential in a number of mammals [142]. Interestingly, two NMTs have been identified in humans (NMT1 and NMT2) that are involved in cell survival via the regulation of apoptotic mechanisms [150]. In mammals N-glycine myristoylation has been identified on proteins including mainly involved in information processing, apoptotic mechanisms and in addition, structural viral proteins were observed [142]. This modification influences the membrane association of the target proteins as well as cellular localization in Eukaryotes. An involvement in protein-protein interaction was also described, for instance, the N-terminal myristoylation of calmodulin is obligatory for the binding to the protein kinase C substrate CAP-23/NAP-22 [151]. The regulation of enzymatic activity was reported for various enzymes. One example given is the c-Abl tyrosine kinase, which exhibits a decreased activity when myristoylated by inducing conformational changes resulting in autoinhibition [152].

1. Introduction

In addition, myristoylation is also known to occur on cysteine residues, termed S-myristoylation, which is probably due to S-palmitoylation using a C14 substrate [153]. S-palmitoylation was reported to usually occur in the neighborhood of transmembrane domains or PTMs influencing membrane association like N-glycine myristoylation or C-terminal prenylation. It was suggested that these additional features direct the proteins to membranes, where palmitoyltransferase containing an aspartic acid-histidine-histidine-cysteine domain (DHHC) is localized. One function reported is the effect on protein stability, as S-palmitoylation prevents ubiquitination, which leads to protein degradation [31, 142, 154].

In the case of N ϵ -lysine acylation, where a long-chain acyl-CoA is attached to an internal lysine residue, the most prominent example is the myristoylation of the tumor necrosis factor- α (TNF α). This protein exhibits two internal lysine residues (K19, K20) where myristoylation was obtained. The demyristoylation is catalyzed by a member of sirtuins (NAD⁺-dependent lysine deacetylase) the SIRT6 enzyme which exhibits long-chain deacylase activity. Removal of the long-chain acyl modification leads to secretion of TNF α , whereas myristoylated TNF α causes lysosomal degradation [155, 156].

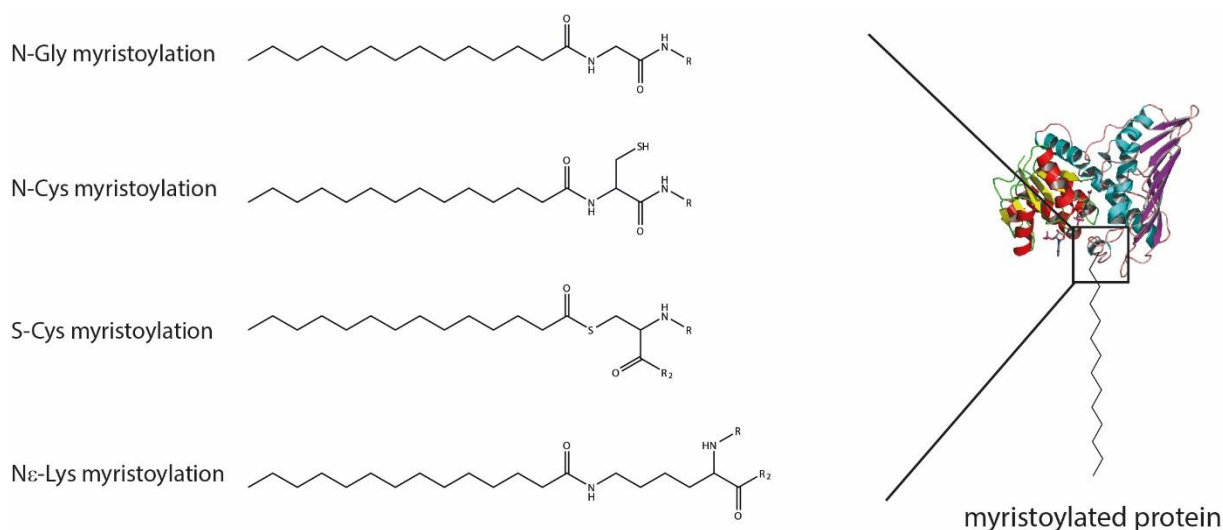


Figure 5: Different types of protein myristoylation.

N-terminal glycine and cysteine myristoylation, internal S-myristoylation and N ϵ -lysine myristoylation, where R represents the polypeptide chain.

Besides numerous studies in Eukaryotes, protein myristoylation is not well understood in Bacteria as well as in Archaea. In Bacteria, N-terminal cysteine myristoylation can lead to a triacylation of the cysteine, allowing the lipoprotein to migrate through the periplasm to the outer membrane for integration using a specialized machinery [157]. Examples for N ϵ -myristoylation are the glycine-rich virulence factors, the Repeat in Toxin class (RTX), are also known occur with two modified lysine residues essential for their cytotoxic activity [158, 159]. The enzyme catalyzing RTX myristoylation (e.g. HlyC) belongs to the GNAT superfamily

1. Introduction

as other acetyltransferases present in Bacteria. They can be classified by the used acyl-donor being either an acyl-CoA or an acyl carrier protein (ACP) carrying the acyl moiety [31].

In 1994, myristoylation was detected for the first time in Euryarchaeota, but since then no further report for Archaea has been published [143]. However, due to the presence of sirtuin homologs in Crenarchaeota as well as FAs [26, 160], protein myristoylation is also assumed to exist in Archaea. The examination of this modification would provide very important insights into the change in membrane composition from Archaea to Eukaryotes, which describes the phenomenon of the lipid divide. Therefore, one part of this work will relate to myristoylation in the Crenarchaeon *S. acidocaldarius* with a focus on the identified target proteins.

1.8 Fatty acids in *S. acidocaldarius*

Although the membrane composition of Archaea does not include FAs, the presence of FAs has been demonstrated for some archaeal organisms. Moreover, the utilization of FAs as carbon and energy source has been reported e.g. for halophilic Archaea indicating the presence of a degradation pathway for FAs [161, 162] and bioinformatic analysis suggested that homologs of all β -oxidation enzymes are present in several Archaea [26]. Focusing on *S. acidocaldarius*, it was also demonstrated to grow on hexanoate and butyrate as sole carbon source [27]. This organism exhibits a 30-kb gene cluster, which comprises genes encoding enzymes involved in lipid and FA metabolism. A TetR-family transcriptional regulator (FadR_{Sa}) was described to repress the transcription of FA gene cluster. In addition FadR_{Sa} was reported to be inhibited by acyl-CoAs which then causes a higher expression of β -oxidation enzymes and thus leads to FA degradation [27]. The FA metabolism comprises a degradation pathway mainly similar to the bacterial-like β -oxidation pathway. The first step includes the activation of FAs to a CoA ester by an AMP-forming acyl-CoA synthetase (Saci_1122). This enzyme exhibited a substrate specificity towards medium chain FAs (C5–C8). The following oxidation step resulting in the enoyl-CoA ester is catalyzed by an acyl-CoA dehydrogenase (Saci_1123) which needs FADs. The resulting reduced FAD cofactor was demonstrated to be re-oxidized by the electron transferring flavoprotein (ETF, Saci_0315). A fusion protein, the 3-hydroxyacyl-CoA dehydrogenase/enoyl-CoA hydratase (Saci_1109), is catalyzing the hydration of the enoyl-CoA as well as the oxidation of the resulting 3(S)-hydroxyacyl-CoA to 3-ketoacyl-CoA. The final step of β -oxidation is the thiolitic cleavage, which releases an acetyl-CoA and a saturated acyl-CoA ester with a reduced chain length of two carbons. The enzyme responsible is a β -ketothiolase (Saci_1114) also exhibiting a medium chain specificity [28]. Interestingly, the last step represents the main driving force for FA degradation and also drives the acyl-CoA dehydrogenase (ACAD) catalyzed reaction by consuming its product,

1. Introduction

since the chemical equilibrium is thus influenced in favor of FA degradation. However, since this pathway was demonstrated to be not fully reversible, a different pathway independent from the β -oxidation must be present for the FA synthesis. It is suggested to exhibit bacterial as well as eukaryotic features and was reconstructed *in vitro* for C8 FA synthesis. The enzyme cascade included a bacterial-like short chain dehydrogenase/reductase (SDR) superfamily 3-ketoacyl-ACP reductase (fabG) homolog, a medium-chain dehydrogenase/reductase (MDR) superfamily enoyl-CoA reductase and a MaoC like dehydratase with an *R*-specificity [28].

In summary, it can be discussed that FAs act as storage compounds in *S. acidocaldarius*. They can penetrate into the cell through their hydrophobic character and can there be activated by conversion into a CoA ester. Subsequently, degradation is possible by bacterial-like β -oxidation, or the CoA ester can be used for PTM of proteins, namely acylation. This is only one proposal how FAs can influence metabolic processes. To identify enzymes modified by FA acylation, we used the method of *in vivo* metabolic labeling in *S. acidocaldarius* cells.

1.9 *In vivo* metabolic labeling

To identify acylated proteins a bioorthogonal chemical proteomic approach was established. The metabolic labeling of cells can take place by adding a natural product inspired probe (NP) like a myristoyl-alkyne (\equiv MyA) to the culture *in vivo*. During incubation under optimal growth conditions, the covalent binding of the probe to the target proteins can occur. Subsequently, the cells are harvested and the produced protein lysate is used for the following copper-catalyzed azide-alkyne cycloaddition reaction (CuAAC). Here, the alkyne moiety enables the attachment of a fluorophore like rhodopsin via click chemistry [163, 164]. Thereby an in-gel fluorescence analysis can visualize the outcome of the experiment. For identification of target proteins, the probe is clicked onto a biotin tag to allow an avidin affinity enrichment with ongoing tryptic on-bead digestion [165]. The direct digestion minimizes environmental contamination of the samples compared to the initial elution of the enriched proteins. Subsequent LC-MS/MS analysis using dedicated software packages allows identification of the proteins (Fig. 6). In direct comparison with control samples treated with DMSO or ethanol, the proteins binding nonspecifically were determined for the same experiment. Thereby the abundant proteins in the samples containing the natural product-inspired probe derive from the labeled proteins. However, probe degradation via cellular processes (e.g. β -oxidation) was discussed to occur during the metabolic labeling procedure resulting in a reduced FA [166] probe and thereby false positive candidates [31]. Some mechanisms are known, for example *N*-propargylamines can inhibit FAD-dependent enzymes [167], as well as alkynes are known to inhibit cytochrome P450 enzymes [165, 168, 169].

1. Introduction

This work describes in chapter 3.2 the generation of a myristoylome profile by *in vivo* metabolic labeling (\equiv MyA) of the Crenarchaeon *S. acidocaldarius* MW001. The results of four independent labeling experiments showed an overlap of identified candidates, but also comprised a high biological variance. The identified proteins (enriched proteins comparing probe and non-probe samples) are predicted to be involved in metabolism or genetic information processing. When considering the candidates predicted to be involved in metabolism, it is apparent from the archaeal Clusters of Orthologous Gene (arCOG) classification [170] that most proteins have a function in energy production and conversion (arCOG class: C) as well as lipid transport and metabolism (arCOG class: I). We identified three proteins in all experiments, annotated as chromatin protein Alba (Saci_1322), β -ketothiolase (Saci_1114), and alkyldihydroxyacetonephosphate synthase (alkyl-DHAP synthase, Saci_1188). For the Alba homolog of *S. solfataricus* the N ϵ -lysine acetylation by SsoPat was already demonstrated [137, 138]. Moreover, Saci_1114 was demonstrated to catalyze one reaction of the β -oxidation pathway responsible for FA degradation in *S. acidocaldarius* [28]. The last candidate, Saci_1188, is predicted to catalyze the cleavage of an ester-linked fatty acid and thereby replacing it by an ether linkage and a long-chain alcohol. However, this class of enzymes has mainly been described for Eukaryotes, which leaves the question of the function of this protein open. In addition some more candidates were identified, including an ACS, which was already reported to be regulated by acetylation in several organisms [117]. Comparing the identified proteins with reported myristoylome and acetylome data, we obtained a certain overlap with archaeal lysine-acetylome data from *S. islandicus* and *T. gammatolerans* [140, 141], suggesting that matching candidates can also comprise acylation modifications. Overall, our results provide initial insights into the distribution of myristoyl-modified proteins, which could suggest a role of FAs in Archaea to address the governing lipid divide.

1. Introduction

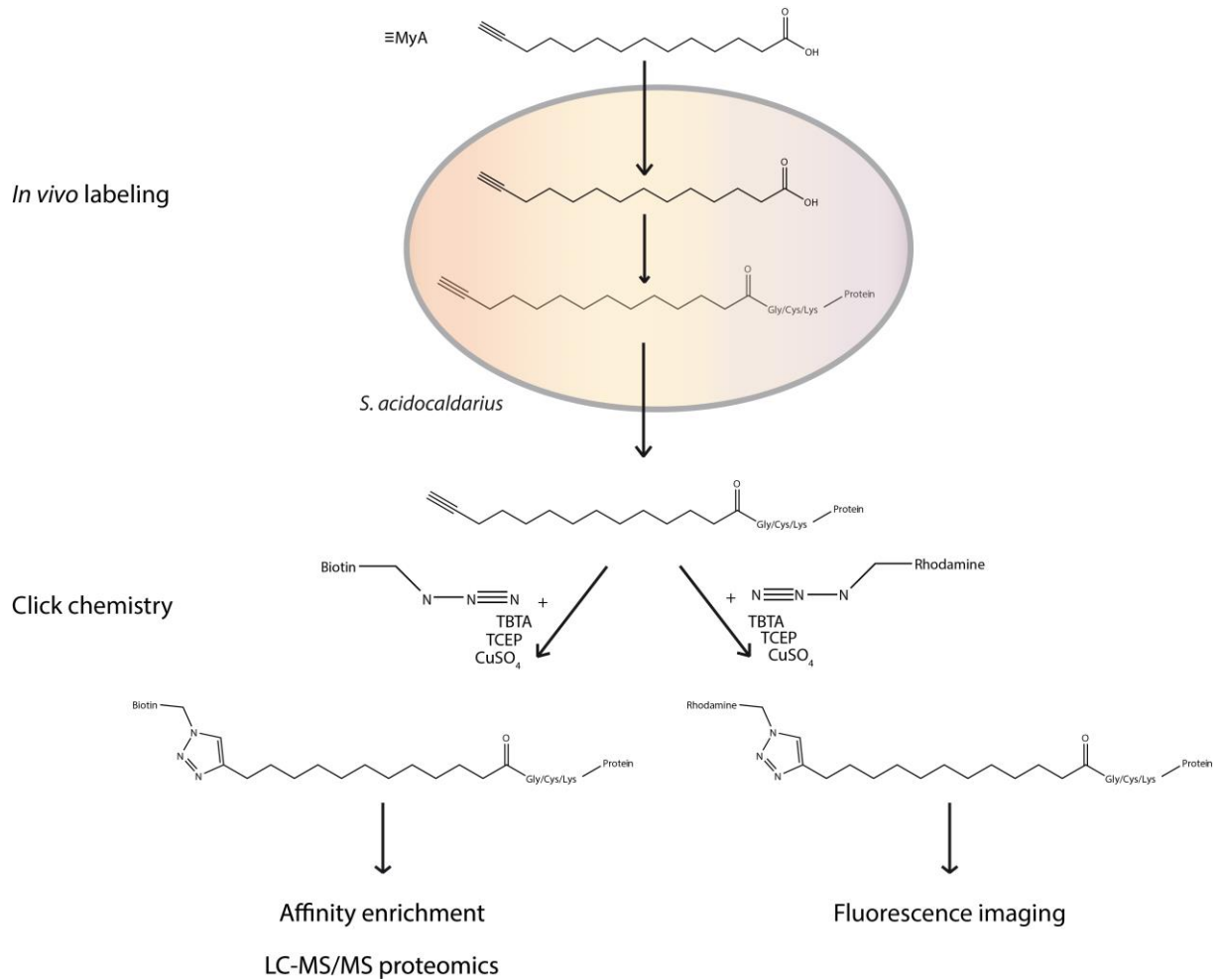


Figure 6: Workflow of bioorthogonal labeling in *S. acidocaldarius* MW001.

The myristic acid alkyne probe was added to *S. acidocaldarius* MW001 cells during late exponential with a final concentration of 20 μM . After 6 hours of incubation the cell lysates were used for copper catalyzed azide-alkyne cycloaddition (click chemistry), enabling fluorescence detection or avidin pulldown with following LC-MS/MS analysis [165].

2 Scope of the thesis

In the present dissertation, the respective research topics are organized into the following chapters, and the contribution I have made to these manuscripts is elucidated below.

In chapter 3.1, the manuscript titled “**Discovery of the missing polyphosphate kinase in the thermoacidophilic Crenarchaeon *Sulfolobus acidocaldarius* by studying polyphosphate metabolism**” addresses the characterization of enzymes involved in polyphosphate metabolism and in particular the identification of the missing polyphosphate kinase. Therefore, I performed heterologous expression of genes encoding the exopolyphosphatase and thymidylate kinases, the protein purification and characterization of relevant enzymes. On the one hand, biochemical characterization was carried out using spectrophotometric measurements to determine the enzymatic properties, revealing a dTMPK-like heteromeric PPK. Furthermore, I used ^{31}P NMR analysis for qualitative investigation, which provided an important contribution to the functions, regulation and kinetic properties of the respective enzymes. With the help of an established ATP recycling system, the polyP synthesis of the heteromeric PPK could be visualized.

In the chapter 3.2, the manuscript with the title “**Proteomic analysis of myristoylation in *Sulfolobus acidocaldarius* using *in vivo* metabolic labeling**”, we performed a bioorthogonal chemical proteomics approach to study protein myristoylation. For this, I prepared the *S. acidocaldarius* MW001 cultures and performed the *in vivo* metabolic labeling using the myristoyl-alkyne probe. For in-gel visualization, we performed copper-catalyzed azide-alkyne cycloaddition reaction (CuAAC) with a Cy3 fluorophore, whereas a biotin tag was used for the pulldown approach to identify the target proteins. The avidin affinity enrichment was performed as initial step during sample preparation for protein identification using LC-MS/MS analysis. The reproducibility of the data is demonstrated by the overlap of the results from four independent experiments, which provided some proteins as promising candidates for protein myristoylation in *S. acidocaldarius*.

3 Chapters

Chapter 3.1

**Discovery of the missing polyphosphate
kinase in the thermoacidophilic**

Crenarchaeon

***Sulfolobus acidocaldarius* by studying
polyphosphate metabolism**

Discovery of the missing polyphosphate kinase in the thermoacidophilic crenarchaeon *Sulfolobus acidocaldarius* by studying polyphosphate metabolism

Authors

Svenja Höfmann¹, Felix Niemeyer², Torsten Schaller², Christopher Bräsen¹, Bettina Siebers^{1*}

¹Molecular Enzyme Technology and Biochemistry (MEB), Environmental Microbiology and Biotechnology (EMB), Centre for Water and Environmental Research (CWE), Department of Chemistry, University of Duisburg-Essen, Universitätsstrasse 5, 45141 Essen, Germany

²Faculty of Chemistry (Organic Chemistry) and Center for Nanointegration Duisburg-Essen (CENIDE), University of Duisburg-Essen, Universitätsstrasse 7, 45141 Essen, Germany

*Corresponding authors: Prof. Dr. Bettina Siebers (Biochemistry): bettina.siebers@uni-due.de,
Tel.: +49 (0)201 183-7061, Fax: +49 (0)201 183-7062

Keywords

Archaea – *Sulfolobus acidocaldarius* – Inorganic polyphosphates – Polyphosphate metabolism – Exopolyphosphatase – Polyphosphate kinase – Heteromeric PPK - Thymidylate kinase

3.1. Polyphosphate metabolism

Abbreviations

(polyP) polyphosphate, (VTC complex) vacuolar transporter chaperone complex, (PPK) polyphosphate kinase, (PPN) endopolyphosphatase, (PPX) exopolyphosphatase, (dTMPK) thymidylate kinase, (dTMP) deoxythymidine monophosphate, (dTDP) deoxythymidine diphosphate

3.1. Polyphosphate metabolism

Abstract

Inorganic polyphosphate, a linear polymer of orthophosphate residues linked by phosphoanhydride bonds, occurs in all three domains of life and plays a diverse and prominent role in metabolism and cellular regulation. While the polyphosphate metabolism and its physiological significance have been well studied in bacteria and eukaryotes including human, there are only few studies in archaea available so far. In Crenarchaeota including members of *Sulfolobaceae*, the presence of polyphosphate and degradation via exopolyphosphatase has been reported and there is some evidence for a functional role in metal ion chelation, biofilm formation and motility, however, the nature of the crenarchaeal polyphosphate kinase is still unknown. Here we used the crenarchaeal model organism *Sulfolobus acidocaldarius* to study the enzymes involved in polyphosphate degradation and synthesis. The exopolyphosphatase (Saci_2018) releases P_i from polyP and exhibited high specific activity with small, medium and long chain polyphosphates (333 (polyP₃), 808 (polyP₄₅) and 994 (polyP₇₀₀) U/mg protein, respectively). The two genes annotated as thymidylate kinase (*saci_2019* and *saci_2020*), localized downstream of the exopolyphosphatase, were identified as the missing polyphosphate kinase (PPK) in *S. acidocaldarius*; only for Saci_0893 thymidylate kinase activity was confirmed. Notably Saci_2020 showed no polyphosphate kinase activity itself but was able to enhance polyphosphate kinase activity of Saci_2019. Heteromeric polyphosphate kinase activity is reversible and shows a clear preference for polyP-dependent nucleotide kinase activity, i.e. polyP-dependent formation of ATP from ADP (13.4 U/mg protein) and to a lower extent of GDP to GTP (0.45 U/mg protein) whereas AMP does not serve as substrate. PPK activity in the direction of ATP-dependent polyP synthesis is rather low (214 mU/mg); GTP was not used as phosphoryl donor. Quantitative ³¹P NMR analysis allowed to follow the enzyme reaction for both ATP and polyP synthesis, for the latter using an ATP recycling system, resulting in a ADP/ATP (25 %/ 75 %) equilibrium independent from the starting point of the reaction. Regarding its enzymatic properties the novel polyphosphate kinase is a PPK2 class I enzyme with unusual heteromeric structure resembling some PPK2 class II enzymes. Notably the enzyme shows no sequence similarity to members of the PPK1, but does with the PPK2 family. Phylogenetic analysis group the enzyme within the thymidylate kinase family of the P-loop kinase superfamily. Our studies show that polyP, in addition to its function as phosphate storage, has a special importance for the energy homeostasis of *S. acidocaldarius* in order to respond to changes in cellular demand and to external fluctuations.

Introduction

Polyphosphate – Distribution and physiological function

Inorganic polyphosphate (polyP) is a linear polymer of three up to hundreds of orthophosphates linked by high-energy phosphoanhydride bonds [171]. Polyphosphate in particular serves as an energy source, reservoir of P_i essential for metabolism and cell growth (P_i homeostasis) as well as ATP substitute in the cell and a function as primordial energy currency in a pre-ATP world is under discussion [75].

PolyP was first isolated from yeast [172]; it has been identified as dense, metachromatic granules, so-called volutins, in many microorganisms in the early nineteen-hundreds [75] and today it is known to occur in all three domains of life [51]. However, the research on “the forgotten polymer” [75] was delayed until the discovery of polyP kinase (PPK), the enzyme involved in polyP synthesis, in *Escherichia coli* by Arthur Kornberg [74, 173]. Since then the metabolism and physiological role of polyP has been studied in many bacteria, using novel polyP quantification methods, enzyme characterization and genetic manipulation, (for review see [44, 51, 52] and this work has been extended to eukaryotes including yeast and humans [174-178] and only more recently to archaea [46, 49, 88, 179]).

Since the in-depth study of polyP metabolism, a multitude of physiological functions of polyP have been reported. PolyP serves as metal ion chelator, buffer against alkali ions and channel for DNA entering (genetic competence). Additionally, polyP is an important regulator for stress, survival and development influencing a multitude of central cellular processes including stress response (e.g. towards heat, UV, antibiotics, metal ions, and oxidative stress, maybe by preventing protein aggregation), bacterial homeostasis, motility and biofilm formation, quorum sensing, genetic competence, virulence and invasion, predator-prey relations, sporulation and germination, DNA replication and phage production, cytokinesis and cell division as well as capsule formation in *Neisseria* species (for reviews see [51, 52, 75, 89]. In addition, polyP is known to influence gene expression (e.g. in *E. coli* and *Streptomyces lividans*) and as inorganic polyanion functions as a primordial chaperone as well as pro-aggregation molecule [89]. In eukaryotes, polyP shows different cellular localizations and influences various metabolic processes such as DNA and RNA synthesis [180, 181]. In mammals, a regulation of bone healing, nerve transduction, blood clotting and cancer metastasis dependent on chain length and concentration of polyP was reported [89].

3.1. Polyphosphate metabolism

Enzymes involved in polyP metabolism.

For polyP synthesis so far two different families of PPKs (EC 2.7.4.1; IPR003414), PPK1 and PPK2, have been described in Bacteria. PPK1 catalyzes the reversible transfer of the terminal γ -phosphate from ATP to polyP with clear preference for polyP synthesis (Fig. 1) [173]. PPK1 homologs are found in many bacteria and few eukaryotes such as the social slime mold *Dictyostelium discoideum* [182].

PPK2s (IPR022486) are also capable of synthesizing polyP, but preferentially catalyze the reverse reaction, the phosphorylation of nucleoside mono- or diphosphates, and thus function as polyP-dependent nucleotide kinases (Fig. 1) [183] (for review see [52]). PPK2s do not share any sequence similarity to PPK1 and phylogenetic, structural as well as sequence analysis suggests that they evolved from P-loop kinases and possess a common evolutionary origin with bacterial dTMPKs [55, 184]. Several bacterial PPK2s have been characterized and large differences were found within the classes in terms of reversibility, substrate specificity and metal cofactor specificity/dependence; several crystal structures were solved (for review and original literature see [52]). In general, PPK2s are divided into three subfamilies/classes based on phylogenetic studies and nucleoside phosphate preference [183] (Fig.1): The class I PPK2s with only one PPK domain utilize polyP to phosphorylate nucleoside diphosphates mainly ADP or GDP [184]. The class II PPK2s with either one or two PPK domains catalyze the phosphorylation of nucleoside monophosphates e.g. AMP to ADP by polyP (poly-AMP phosphotransferase) [185] so that a combination of both classes can convert AMP to ATP [184]. Finally, the class III PPK2s with one PPK domain either phosphorylate nucleoside mono- or diphosphates and thus are able to synthesize the direct formation of ATP from AMP and polyP [183]. The distribution of PPK1s and PPK2s in the bacteria is very diverse. While some organisms have only one PPK1 or PPK2, some possess up to five PPK2 homologues (i.e. *Ralstonia eutropha*) and others have both PPK1 and PPK2 [51, 52].

However, the processes and enzymes involved in polyP synthesis are not universally conserved and in many eukaryotes the mechanism for polyP synthesis is still unknown. In yeast polyP is synthesized by a vacuolar transporter chaperone 4 (VTC4) complex, which has also been identified in other eukaryotes such as *Trypanosoma brucei* [89]. In *D. dictyostelium*, in addition to the bacterial type PPK1 an actin-related protein (Arps) complex was identified in acidocalcisomes that synthesizes an actin-like filament in conjunction with polyP [59].

The most common enzyme for the degradation of polyP found in all three domains of life is the exopolyphosphatase (PPX), which hydrolyzes the terminal phosphoanhydride bond in the linear polymer and releases P_i . Two types of PPXs have been reported that are active on long chain polyP but differ in their structure and evolutionary/family affiliation. The *S. cerevisiae* PPX1 is a member of the DHH phosphoesterase superfamily (evolutionary

3.1. Polyphosphate metabolism

closely related to family II pyrophosphatases) and homologous are found in fungi and protozoa [186]. The two *E. coli* enzymes, PPX1 and PPX2, are members of the actin-like ATPase domain superfamily (IPR043129) comprising sugar kinases, actin, heat shock protein hsp70 and prokaryotic cell cycle proteins. An ATPase domain common to prokaryotic cell cycle proteins, sugar kinases, actin, and hsp70 heat shock proteins. While PPX1 catalyzes the degradation of polyP, PPX2 also possesses guanosine pentaphosphate phosphohydrolase (GPPA) activity liberating P_i from polyP or guanosine-5'triphosphate-3'-diphosphate (pppGpp) and thus plays an important role in the bacterial stringent response [62, 187]. In addition, an endopolyphosphatase (PPN) that cleaves inside the polymer forming polyPs of different chain length has been reported from yeast or animal cells (Fig. 1) [51, 188, 189]. Last but not least, several enzymes that employ both, either polyP or ATP, as phosphate donor such as NAD kinase and glucokinase have been described [51] also from archaea [190].

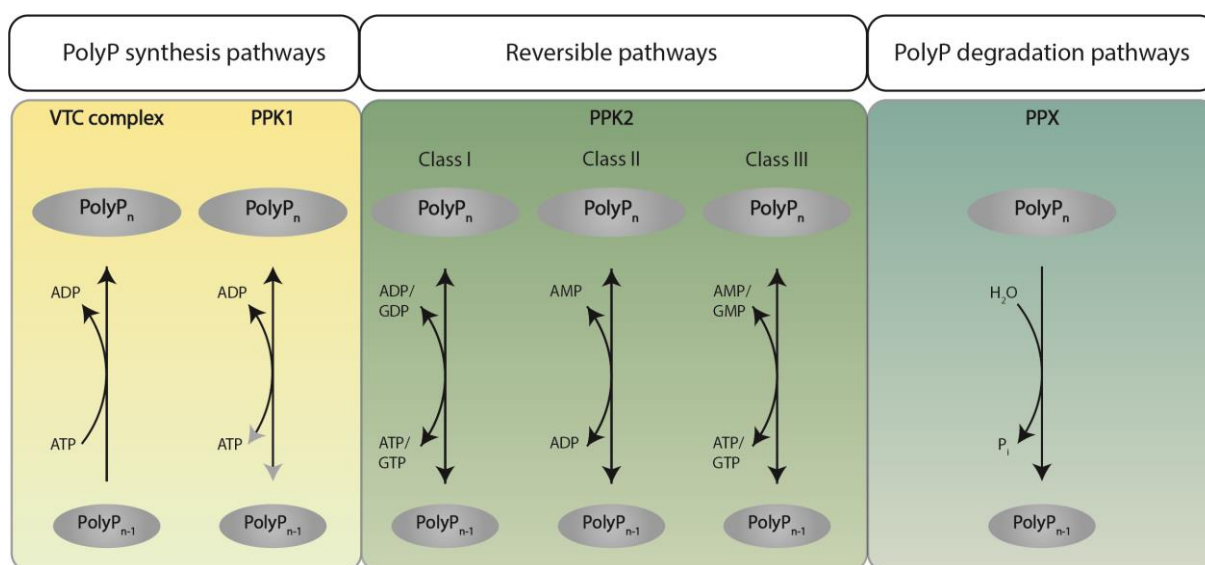


Figure 1: Enzymes involved in polyphosphate synthesis and degradation.

The enzymes involved in polyP conversion with substrates and co-substrates are depicted.

Abbreviations; VTC complex, vacuolar transporter chaperone complex; PPK, polyphosphate kinase; PPN, endopolyphosphatase; PPX, exopolyphosphatase; polyP, polyphosphate; P_i , orthophosphate.

Compared to bacteria and eukaryotes, relatively little is known about polyP function and metabolism in archaea. The presence of polyP in this kingdom was already described in 1983 [191] and currently polyP granules are known to be present in several archaea e.g. *Sulfuracidifex metallicus* and *Metallosphaera sedula* [46]. For the methanogenic archaeon *Methanosarcina mazei* P_i starvation experiments with subsequent addition of P_i resulted in a significant increase in *ppk1* gene transcripts and induced polyP accumulation [49]. In the Crenarchaeota, *S. metallicus* and *Saccharolobus solfataricus* (formerly *Sulfolobus solfataricus*) the influence of polyP accumulation in response to metal toxicity (e.g. copper) was reported [47, 48, 179]. *PPX* overexpression in *S. solfataricus* and *Sulfolobus acidocaldarius* resulted in a lack of polyP accompanied by decreased motility,

3.1. Polyphosphate metabolism

adhesion and biofilm formation [88]. However, the respective PPX has not been biochemically characterized so far.

Notably, bioinformatic analyses in extremophiles revealed the presence of key enzymes for polyP synthesis and degradation in all extremophilic bacteria whereas in many extremophilic archaea the set of encoded polyP metabolizing proteins turned out to be incomplete. While predominantly methanogenic archaea contain a complete polyP pathway, many halophilic Euryarchaeota, i.e. Haloferacales, Natribiales and Halobacteriales, lack genes encoding catabolic PPX whereas all Crenarchaeota lack genes encoding homologs for polyP synthesis (e.g. PPK1s, PPK2, Arps) [43, 46, 49]. Thus, although good evidence has been provided for polyP degradation via PPX and polyP formation could be detected using electron microscopy [46, 47], no enzymes for polyphosphate synthesis could be so far identified in *Sulfolobales* and Crenarchaeota in general [101].

Herein, the polyP metabolism in the crenarchaeal, thermoacidophilic model organism *S. acidocaldarius* with optimal growth at 76 °C and pH 3 [91] was studied. Besides the characterization of the recombinant PPX (Saci_2018) from *S. acidocaldarius*, we present the identification and characterization of the missing PPK, a heteromeric, dTMPK-like PPK (Saci_2019 and Saci_2020), which represents a novel class within the P-loop kinase family. The heteromeric PPK catalyzes the reversible conversion of ATP and polyP with clear preference for the catabolic direction, i.e. polyP-dependent nucleotide kinase activity. The lacking dTMPK activity, which is essential for DNA synthesis, could be shown for Saci_0893 of *S. acidocaldarius*.

Results

Polyphosphate degradation in *S. acidocaldarius*

Cloning, expression, purification and characterization of exopolyphosphatase (PPX, Saci_2018)

Even though the physiological role of polyP was investigated by *ppx* overexpression in *S. acidocaldarius* [88], a detailed characterization of the PPX is still missing. Therefore, we cloned the gene *saci_2018* encoding the PPX (EC 3.6.1.11) into the vector pET15b for heterologous expression in *E. coli*. The N-terminal His-tagged protein was purified via immobilized metal affinity chromatography (IMAC) (SI Fig. 1) and stored in 15 % (v/v) glycerol at -80 °C. In total 12 mg of purified PPX were gained out of 4.4 g (wet weight) *E. coli* cells.

The PPX was characterized using the malachite green assay following P_i liberation. In order to address substrate specificity polyPs of different chain lengths, i.e. polyP₃, polyP₄₅ and polyP₇₀₀, and pppGpp were tested as substrates in the presence of Mg²⁺ as divalent cation (Fig. 2B). The highest specific activity was observed for the long-chain polyphosphates polyP₇₀₀ with 994.4 U/mg protein (1 μM substrate) followed by medium-chain polyP₄₅ (807.6 U/mg protein, 20 μM substrate) and triphosphate (polyP₃) (333.4 U/mg protein, 500 μM substrate). No activity was observed using pppGpp as substrate (100 μM).

Due to solubility issues with polyP₇₀₀, the kinetic parameters of the PPX were determined using polyP₄₅ as substrate. First, we investigated the PPX activity at different temperatures from 50 °C to 90 °C (SI Fig. 2A). The specific activity increased with temperature and highest activity (1234 U/mg) was observed at 90 °C with 44 % (539 U/mg protein) and 16 % (197 U/mg protein) residual activity at 70 and 50 °C, respectively. To match the physiological growth conditions of *S. acidocaldarius*, all further enzyme assays were carried out at 70 °C.

The PPX activity was dependent on divalent metal ions with only slight preference for Mg²⁺ over Mn²⁺ (each 2.5 mM, residual activity of 100 % and 98.6 % respectively (SI Fig. 2B)), but no activation of PPX activity was observed when adding NH₄⁺ or K⁺ (SI Fig. 2C). Moreover, no inhibitory effect of NaF (1mM) on PPX enzymatic activity could be observed. The enzyme kinetics showed a sigmoidal behavior and at higher substrate concentration (above 50 μM) a slight inhibition was observed (Fig. 2C). The kinetic characterization of the PPX revealed a $v_{\max,app}$ of 890 U/mg protein and a $K_{M,app}$ value of 5.24 μM, corresponding to a $k_{cat,app}$ of 655.2 s⁻¹ and catalytic efficiency of 1.16 x 10⁸ M⁻¹ s⁻¹ using polyP₄₅ as substrate.

3.1. Polyphosphate metabolism

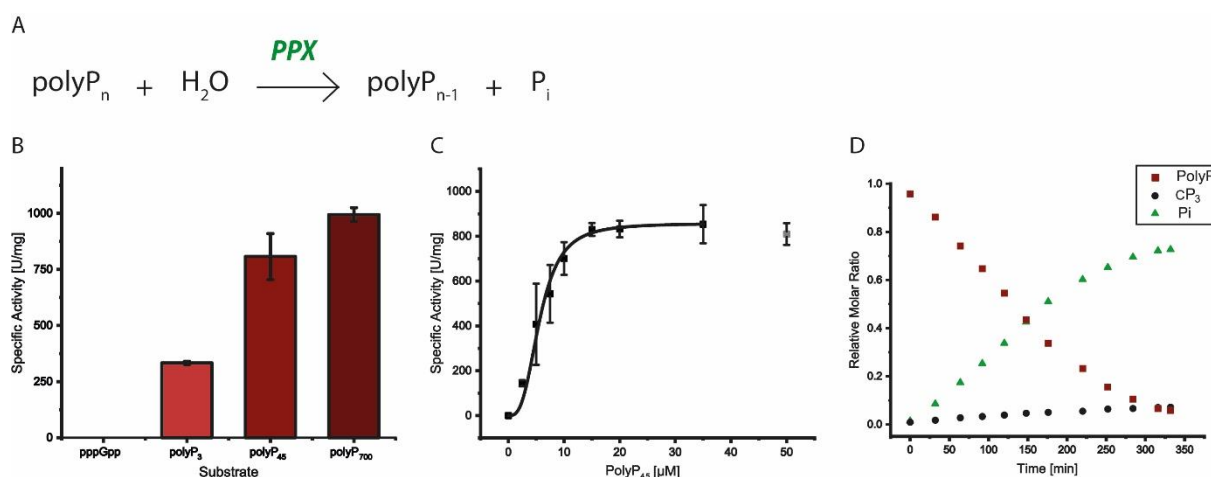


Figure 2: Characterization of the recombinant exopolyphosphatase (PPX, Saci_2018) from *S. acidocaldarius*.

(A) Enzyme reaction catalyzed by the PPX. (B) Substrate specificity of PPX. PolyPs of different chain lengths (polyP₃ (500 μM), polyP₄₅ (20 μM), polyP₇₀₀ (1 μM)) as well as pppGpp (100 μM) were tested as substrate. The enzymatic activity was determined at 70 °C with 1.5 μg of purified PPX and the P_i release was followed discontinuously via the malachite green assay. (C) Kinetic properties of the PPX. The enzyme characterization was performed at 70 °C in the presence of 0 to 50 μM polyP₄₅, by following P_i liberation using the discontinuous malachite green assay. (D) ³¹P NMR analysis of the PPX reaction. The conversion of polyP₄₅ (500 μM) was followed over time at 70 °C in the presence of 14 μg PPX. Both, substrate depletion (polyP, red squares) and product formation (P_i, green triangles) as well as the non-enzymatic production of cyclotriphosphate (cP₃, black circles) were observed (for ³¹P-NMR spectra see SI Fig. 3). Three independent measurements (n=3) were performed (except for the ³¹P NMR measurement) and error bars indicate the standard error of the mean (SEM).

Due to the high standard deviations obtained in the malachite green assay, we established ³¹P NMR as alternative qualitative experimental approach. The reaction was carried out using the same assay conditions at 70 °C, however, in 10 % D₂O and the conversion of polyP₄₅ (500 μM) and the product, P_i, formation was followed over time in the presence of PPX (14 μg protein) (Fig. 2D). PolyP₄₅ was completely converted to free orthophosphate by PPX. Unexpectedly, an additional signal belonging to an unknown phosphate compound was detected (singlet at -20.5 ppm, SI Fig. 3) that could be identified as cyclotriphosphate (cP₃). The non-enzymatic formation of cP₃ from polyP at high temperature in the presence of metal ions has been reported previously [192]. Further *in vitro* studies in absence of PPX using ³¹P NMR spectroscopy confirmed the increase of non-enzymatic cP₃ and P_i formation from polyP in presence of divalent metal ions (MgSO₄, 2.5 mM, 6-fold) and monovalent ions (KH₂PO₄, 500 μM, 2-fold) at 70 °C (SI Fig. 4). cP₃ was stable under the chosen assay conditions and time-dependent hydrolysis to orthophosphate was not observed. Notably, cP₃ itself did not serve as substrate for PPX (³¹P NMR, SI Fig. 5).

In conclusion, the PPX shows a very high specific activity, preferably towards long-chain polyP chains. In this process, polyP is completely consumed as substrate, producing orthophosphate and the heat-stable by-product cP₃, which does not serve itself as substrate for PPX.

Polyphosphate synthesis in *S. acidocaldarius*

In Crenarchaeota, the enzyme(s) involved in polyP synthesis are still unknown. No homologues of bacterial/eukaryotic PPKs have been identified so far [43, 46, 49]. Closer inspection of the *ppx* gene (*saci_2018*) neighborhood revealed two adjacent genes annotated as dTMPKs (*saci_2019* and *saci_2020*) (Fig. 3A). In addition, a separate, third gene annotated as dTMPK (*saci_0893*) was identified in the genome. Notably, the solved crystal structures of the single and two domain fused PPK2s from *Pseudomonas aeruginosa* and *Sinorhizobium meliloti*, respectively, identified dTMPKs as the closest structural homologs [56, 184]. Therefore, all three proteins were recombinantly expressed and examined for their respective enzymatic activities.

Cloning, expression, purification and characterization of the three annotated thymidylate kinases (Saci_0893, Saci_2019, Saci_2020)

The genes encoding the three annotated thymidylate kinases were cloned into the expression vector pET15b (*saci_0893*, *saci_2019*) and pET28b (*saci_2020*), respectively, for heterologous expression in *E. coli* Rosetta (DE3). The N-terminal His-tagged proteins were purified via IMAC (SI Fig. 1). In total 200 µg Saci_0893, 250 µg Saci_2019 and 120 µg Saci_2020 were obtained from 1.0 g, 6.0 g and 9.0 g (wet weight) *E. coli* cells, respectively.

Thymidylate kinases (EC: 2.7.4.9, dTMPK) catalyze the Mg²⁺-dependent phosphorylation of deoxythymidine monophosphate (dTMP) to deoxythymidine diphosphate (dTDP) by using ATP as phosphoryl donor (Fig. 3B). First, a qualitative analysis via ³¹P NMR spectroscopy was used to investigate dTMPK activity of the three enzymes (Fig. 3C). Thymidylate kinase activity was tested with dTMP (1 mM) as substrate and ATP (2 mM) as phosphoryl donor in 50 mM TRIS/HCl pH 7.4 containing 50 mM KCl, 4 mM MgCl₂, 10 % D₂O and 10 µg enzyme at 70 °C. After eight hours of incubation, the ³¹P NMR analysis revealed for Saci_0893 and Saci_2019, but not for Saci_2020, ADP formation from ATP. However, only for Saci_0893 the ATP-dependent formation of dTDP from dTMP was detected (Fig. 3D). Moreover, only for Saci_0893 a clear dependence of the enzyme activity on the substrate (dTMP) (SI Fig. 6A) was detected. Since only Saci_0893 possess dTMPK activity, the specific activity was determined using a continuous coupled PK-LDH assay at 55 °C by following ADP formation. The determination of the enzymatic activity of Saci_0893 revealed a dTMP-dependent specific activity of 2.52 ± 0.03 U/mg (Tab. 1).

3.1. Polyphosphate metabolism

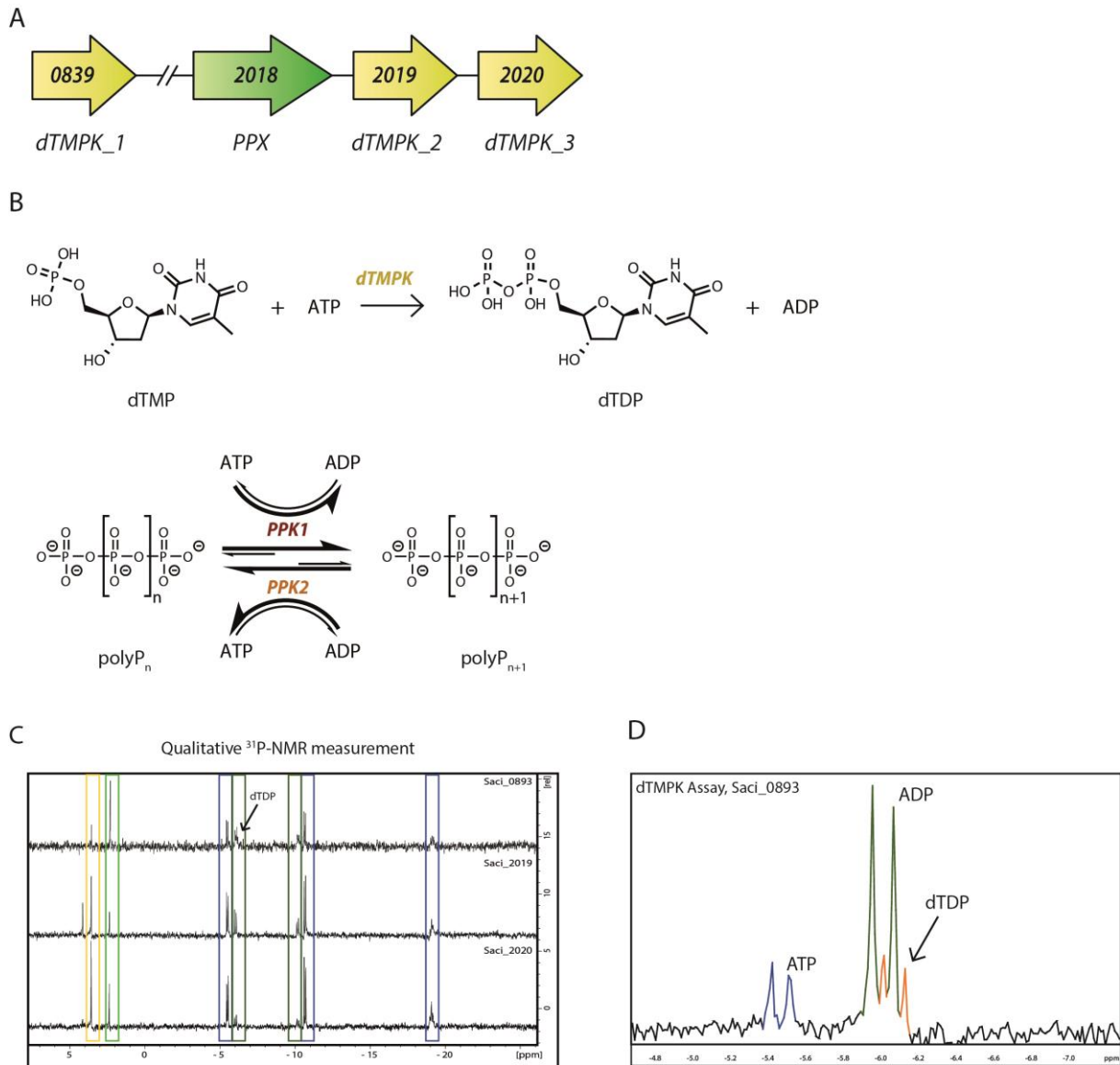


Figure 3: Investigation of polyphosphate kinase and thymidylate kinase candidates in *S. acidocaldarius*.

(A) Localization of annotated dTMPKs in the *S. acidocaldarius* genome. Two thymidylate kinase genes (yellow) *saci_2019* (dTMPK_2) and *saci_2020* (dTMPK_3) are located directly adjacent to the *ppx* gene (*saci_2018*, green) while the third candidate is separate in the genome (*saci_0893*, dTMPK_1). (B) Enzyme reactions catalyzed by dTMPKs and PPKs. (C) ³¹P NMR spectra of dTMPK assay of the recombinant Saci_0893, Saci_2019 and Saci_2020, respectively. The assay was performed at 70 °C for 8 hours containing 2 mM ATP, 1 mM dTMP, 4 mM MgCl₂, 50 mM KCl with 10 µg of the respective enzyme in 50 mM TRIS/HCl pH 7.4 with 10 % D₂O. The respective signals for ATP (blue box), ADP (dark green box), dTMP (yellow box) and P_i (light green) are depicted. (D) Magnification showing the formation of dTDP in the presence of Saci_0893. The dTDP signal (orange) overlaps with one ADP signal (dark green).

Next, we were curious about a possible PPK function of the annotated dTMPKs and followed the ATP-dependent formation of polyP using the MicroMolar Polyphosphate Assay Kit (ProFoldin) as well as the formation of ADP (PK-LDH assay) and P_i (malachite green assay) over time (SI Fig. 6). In accordance with our previous observations, ADP but also polyP formation from ATP was observed for Saci_2019 but not for Saci_0893 and Saci_2020.

3.1. Polyphosphate metabolism

Thus, of the three annotated dTMPKs only Saci_0893 possesses dTMPK activity, whereas Saci_2019 shows PPK activity directly starting from ATP without the need of a precursor and for Saci_2020 neither dTMPK nor PPK activity could be confirmed.

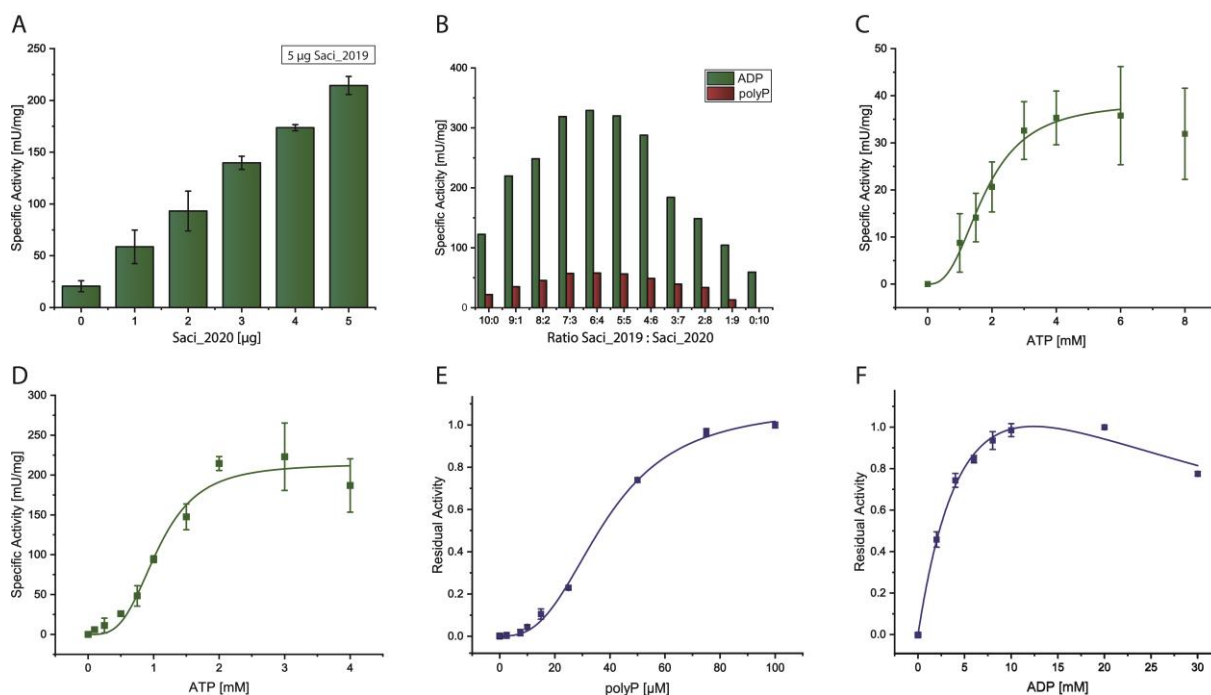


Figure 4: Characterization of recombinant, homomeric PPK (Saci_2019) and heteromeric PPK (Saci_2019 and Saci_2020) from *S. acidocaldarius*.

(A) The specific PPK activity (following ADP formation) of homomeric PPK (Saci_2019, 5 µg protein) upon addition of 0 - 5 µg Saci_2020, and (B) by using different molar ratios of Saci_2019 and Saci_2020 (from 10:0 to 0:10, as indicated; 10 µg enzyme in total) are depicted. The formation of ADP (green) and polyP (red) (B) are depicted. Kinetic properties of the recombinant, homomeric PPK (Saci_2019, C) and heteromeric PPK (Saci_2019 and Saci_2020, molar ratio 1:1, D) from *S. acidocaldarius* with respect to ATP-dependent polyP synthesis following ADP formation. PolyP-dependent nucleotide kinase activity, i.e. ATP formation, of the heteromeric PPK with different polyP (E) or ADP concentrations (F). For all experiments (except B) three independent measurements ($n=3$) were performed and error bars indicate the standard error of the mean (SEM).

Since a two-domain structure, with an inactive N-terminal domain, was reported for the class II PPK2 from *P. aeruginosa* [184] we further investigated the influence of the Saci_2020 on Saci_2019 PPK activity. First a fixed amount of Saci_2019 (5 µg) was used and 0 - 5 µg Saci_2020 were added and the specific PPK activity was determined based on ADP formation (70 °C) using a discontinuous PK-LDH assay at 37 °C (Fig. 4A). A steady increase in PPK activity was observed with increasing amounts of Saci_2020. To determine the optimal molar ratio of the two enzymes *in vitro* an additional titration experiment was performed. For this, a molar ratio of 10: 0 to 0:10 (Saci_2019 to Saci_2020) was used and the specific activity was determined based on polyP and ADP formation (Fig. 4B). The highest PPK activity was observed at a molar ratio of 7:3 to 5:5 with a specific activity of 320 mU/mg protein. Therefore, for all further PPK experiments with the heteromeric enzyme a molar ratio of 1:1 for Saci_2019 and Saci_2020 was used.

Next, we performed a more detailed characterization of the homomeric PPK (Saci_2019) as well as the heteromeric PPK (Saci_2019 and Saci_2020 (1:1)) at a reaction temperature of

3.1. Polyphosphate metabolism

70 °C using the discontinuous PK-LDH assay at 37 °C. The homomeric enzyme showed a pH- and temperature optimum at pH 8.0 and 80 °C, respectively (SI Fig. 7). For the heteromeric PPK a shift in the pH optimum to 7.0 and of the temperature optimum to 70 °C, with only 12 % residual activity at 80 °C, was observed (SI Fig. 8).

The kinetic characterization with ATP as substrate (following ADP formation) revealed an allosteric behavior with a $v_{\max,app}$ of 37 mU/mg and an apparent $K_{M,app(ATP)}$ of 1.73 mM (70 °C), resulting in a k_{cat} of 0.015 s⁻¹ and a catalytic efficiency of 8.77 M⁻¹s⁻¹ for the homomeric PPK (Saci_2019) (Fig. 4C). Substrate inhibition was observed at ATP concentrations above 8 mM ATP. For the heteromeric PPK with ATP as substrate (following ADP formation) a specific activity ($v_{\max,app}$) of 214 mU/mg and a $K_{M,app(ATP)}$ of 1.07 mM was determined, corresponding to a k_{cat} of 0.173 s⁻¹ and a catalytic efficiency of 161.6 M⁻¹s⁻¹. It should be noted that due to the low protein yield, the measured values showed a certain deviation between the respective experiments, which resulted in partly deviating activities. Nonetheless, an about 18-fold higher catalytic efficiency is observed for the heteromeric PPK compared to the homomeric PPK. No enzyme activity could be detected using GTP as substrate. Utilization of the polyP assay kit for characterization of the heteromeric PPK revealed similar kinetic properties, however, an about four-fold decreased specific activity ($v_{\max,app}$ of 63.3 mU/mg) was observed whereas the $K_{M,app(ATP)}$ (0.95 mM) was in a similar range (k_{cat} of 0.05 s⁻¹, catalytic efficiency of 53.9 M⁻¹s⁻¹) indicating the limitations of the assay (SI Fig. 8C). To address if the formed ADP might influence the enzyme activity of heteromeric PPK, polyP formation was followed over time with 1.5 mM ATP in absence and presence of 2 mM ADP. The addition of ADP resulted in complete inhibition of polyP synthesis (SI Fig. 9) suggesting either inhibition or reversibility of the PPK reaction.

Reversibility of the PPK reaction (polyP-dependent nucleotide kinase activity)

Hence, to address the reversibility of the PPK reaction ATP or GTP formation from ADP and GDP, respectively, with polyP (50 µM polyP₄₅) as phosphate donor was examined. For initial enzyme assays 5 µg of heteromeric PPK (SI Fig. 10A (ADP), B (GDP)) as well as homomeric PPK (Saci_2019 or Saci_2020) (SI Fig. 10C, D (ADP)) were used. The utilization of polyP as well as ATP or GTP formation at 70 °C was followed discontinuously via the MicroMolar Polyphosphate Assay Kit (ProFoldin) and the hexokinase/glucose-6-phosphate dehydrogenase assay at 37 °C, respectively. For the heteromeric PPK a specific activity with ADP of 56.3 mU/mg (polyP detection) and 1.40 U/mg (ATP detection) was determined. The enzyme also shows activity with GDP, albeit much lower (6.13 mU/mg (polyP detection) and 0.455 U/mg (GTP detection)).

3.1. Polyphosphate metabolism

For the homomeric enzymes using ADP as substrate, only Saci_2019 catalyzes ATP synthesis with a specific activity of 148 mU/mg, which is 10 times lower than for the heteromeric PPK (SI Fig. 10C and D). Notably no decrease in polyP concentration could be observed, in these assays. For Saci_2020, as for the polyP synthesis, no ATP formation from polyP was observed.

The detailed kinetic characterization of the heteromeric PPK with polyP₄₅ and ADP as (co)substrate revealed a v_{max} of 13.4 U/mg and K_M -values of 38.29 μ M for polyP₄₅ (k_{cat} : 10.8 s⁻¹, k_{cat}/K_M : 282.7 mM⁻¹s⁻¹) and of 7.76 mM for ADP (k_{cat} : 5.05 s⁻¹, k_{cat}/K_M : 650.6 M⁻¹s⁻¹) (Fig. 4 E and F). Hence, the heteromeric PPK shows a clear preference (4-fold higher catalytic efficiency, comparing ADP and ATP dependence) for polyP-dependent nucleotide formation.

Previous studies have reported that the direction of the PPK reaction can be controlled by either different divalent metal ions or their concentration (for review see [52]). Therefore, we tested the effect of different metal ions on heteromeric PPK activity (SI Fig. 11). In absence of metal ions as well as in presence of Ca²⁺, Co²⁺ and Ni²⁺ (10 mM) no PPK activity (i.e. polyP or nucleotide formation) was observed. In direction of polyP formation, the enzyme precipitated by addition of Mn²⁺ to the reaction mixture. However, for nucleotide formation the enzyme showed a similar specific activity with Mg²⁺ or Mn²⁺. Additionally, the optimal Mg²⁺ concentration was determined for both reactions, resulting in the highest specific activity with 10 mM MgCl₂ (4.5 U/mg for NTP-formation, 147 mU/mg for polyP-formation) while using 20 mM ADP and 100 μ M polyP₄₅ (NTP-formation) or 3 mM ATP (polyP formation) as substrates (SI Fig. 11B and D).

Finally, in order to analyze the reversibility of the PPK reaction we performed quantitative ³¹P NMR (Fig. 5, SI Fig. 12) to follow either nucleotide formation from polyP (Fig. 5A) as well as polyP formation from ATP (Fig. 5B) over time. First, we followed ATP formation from polyP, i.e. polyP-dependent nucleotide kinase activity. The assay mixture (ATP formation) contained 10 mM ADP, 500 μ M polyP₄₅, 10 mM Mg²⁺, and 20 μ g heteromeric PPK and 10 % D₂O was analyzed at 70 °C, where one merged spectrum was recorded every 15 min for about 750 min. The obtained ratios of the phosphorous signals are displayed as a function of the reaction time (Fig. 5A). In agreement with our enzymatic studies, a continuous consumption of polyP and ADP as well as the formation of ATP could be observed over time. After about 75 min, the enzymatic reaction reached an equilibrium resulting in a ratio of ATP to ADP of 75 % to 25 %. The remaining polyP was hydrolyzed non-enzymatically resulting in formation of cP₃ and P_i (SI Fig. 12A).

Next, we studied polyP synthesis from ATP (Fig. 5B, SI Fig. 12B). Therefore, 2 mM ATP and 4 mM Mg²⁺ were incubated in presence of 10 % D₂O and 20 μ g heteromeric PPK at 70 °C, where formation of ADP and consumption of ATP were followed up to 360 min. One merged

3.1. Polyphosphate metabolism

spectrum was recorded every 30 min. The obtained ratios of the signals of ADP and ATP are displayed as a function of the reaction time (Fig. 5B). The experiment shows that ATP is consumed and converted to ADP, however, after approx. 200 min the same equilibrium is reached as for ATP synthesis (Fig. 5A) and no more conversion of ATP is observed. No formation of polyP is observed under these conditions supporting the clear enzymatic preference for polyP-dependent nucleotide synthesis.

In order to shift the equilibrium of the reaction by circumventing formation of ADP and thus enable polyP synthesis, an enzymatic ATP recycling system was added based on the recombinant pyruvate kinase (PK) from *S. solfataricus*. This enzyme showed a specific activity of 78.5 U/mg at 55 °C (continuous LDH assay) and catalyzes the recycling of ADP to ATP by conversion of phosphoenolpyruvate (PEP) to pyruvate. In preliminary assays, different PEP and MgCl₂ concentrations and SsoPK amounts were tested following ADP and polyP formation (SI Tab. 1). The highest polyP concentration formed was detected for 20 mM MgCl₂, 10 mM PEP and 8 µg/ml SsoPK resulting in synthesis of 920 µM polyP after 90 min. Therefore, the ³¹P NMR measurement (following polyP synthesis from ATP) was repeated in the presence of the optimized recycling system and 20 µg heteromeric PPK (Fig. 5C, SI Fig. 12C) for 750 min at 70 °C. The experiment nicely shows that until the depletion of PEP at 450 min the ATP concentration is constant due to PK recycling system and under these conditions polyP formation is observed. After PEP depletion, only a slight decrease in ATP and increase in ADP is observed indicating that the equilibrium is reached whereas polyP shows a dramatic decrease to almost zero. Notably, as in previous experiments, a non-enzymatic formation of cP₃ from polyP is observed which continues to increase even after the decrease of polyP. The obtained free phosphate results primarily from the hydrolysis of PEP and polyP (SI Fig. 13), as demonstrated in previous experiments (SI Fig. 4). The same reaction was carried out with 20 µg of homomeric PPK (Saci_2019), but although the ATP concentration was stable, no production of polyP could be detected after 750 min (Fig. 5D, SI Fig. 12D). This again illustrates the significantly higher PPK activity of the heteromeric enzyme.

3.1. Polyphosphate metabolism

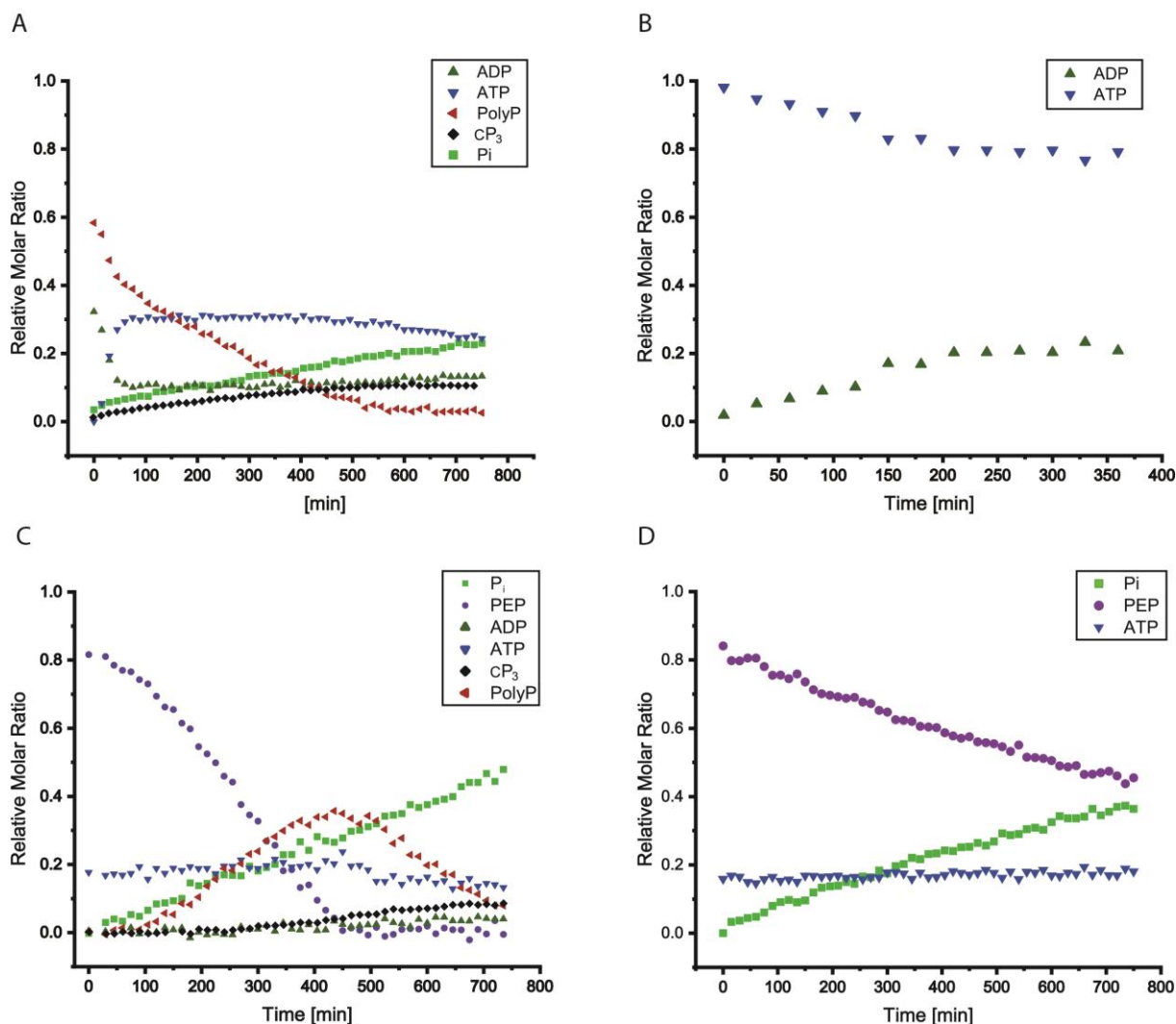


Figure 5: Analysis of the reversible heteromeric PPK from *S. acidocaldarius* using ^{31}P NMR spectroscopy. The polyP-dependent ATP formation (A) and ATP-dependent polyP formation without (B) and with ATP recycling system for heteromeric (C) and homomeric PPK (Saci_2019) was analyzed.

All experiments were performed at 70 °C with 20 μg of heteromeric (A-C) or homomeric (D) PPK. ATP formation (A) was assayed with 10 mM ADP, 10 mM MgCl_2 and 500 μM polyP₄₅. PolyP synthesis (B, C and D) was carried out with 2 mM ATP, 4 mM MgCl_2 and for ATP recycling with heteromeric (C) or homomeric PPK (D), 10 mM PEP, 8 μg SsoPK and 20 mM MgCl_2 (final concentration) were added. The obtained signals in the ^{31}P NMR spectra are shown: ATP (blue triangles), ADP (green triangles), polyP (red triangles), PEP (purple dots), cP₃ (black squares) and P_i (green squares) (for NMR spectra see SI Fig. 12).

Discussion

PolyP has been found in all three domains of life, however, although its function as phosphate and energy storage as well as regulator of a multitude of cellular processes is well established in Bacteria and Eukaryotes the knowledge of polyP metabolism in Archaea is only limited. The presence of polyP has been demonstrated in archaea and the effect of low polyP levels (*ppx* overexpression) on motility, biofilm formation and heavy metal resistance has been proven in *Sulfolobus spp.*. Contrary to this functional analysis, however, the enzymes involved in polyP metabolism have hardly been studied so far. Bioinformatic studies provide an incomplete picture of enzymes involved in polyP metabolism in many archaea, suggesting the presence of novel enzymes [43, 46]. In all members of the Crenarchaeota homologues for known PPK of the PPK1, PPK2 and VTC4 family are missing.

Polyphosphate degradation by PPX

Starting with the knowledge that polyP is present in *S. acidocaldarius* and exhibits a PPX homolog in the genome a first insight into polyP metabolism was gained by characterization of *SaciPPX*. The *SaciPPX* comprised highest activity at 90 °C and as reported for archaeal PPXs and for the enzyme from *E. coli* [62] as well as the PPX2 from *M. tuberculosis* [193], a preference for long-chain polyP (here polyP₇₀₀) with only 81.2 % and 33.5 % residual activity for polyP₄₅ and polyP₃, respectively, was detected. No activity was observed using pppGpp as substrate. Notably, ppGpp/pppGpp has been shown to be generally absent in archaea [194, 195], although stringent response [195] and an *in vitro* effect of ppGpp on elongation factor 1 α , i.e. protein synthesis, have been reported in *S. solfataricus* [196]. Within the archaea, only the activity of recombinant PPXs from *S. solfataricus* [101] and *M. sedula* [48] have been described, but a detailed enzyme characterization has been missing. The PPX from *S. solfataricus* revealed highest activity at 50 - 60 °C with long-chain polyP (P₇₀₀₋₈₀₀) (SI Tab. 2) in the presence of 10 mM Mn²⁺ resulting in a specific activity of 0.6 U/mg protein using an end point determination with [³³P]polyP₇₅₀ (acid precipitation/filtration assay). For the recombinant *M. sedula* PPX the P_i release was determined after 30 min of incubation with highest activity at 65 °C and also a clear preference for long chain polyP (P₇₀₀) as well as a 3.6-fold increased activity in presence of 20 mM Mg²⁺ and inhibition by P_i (1 mM) were reported.

The kinetic characterization of the *SaciPPX* with polyP₄₅ showed a specific activity of 890 U/mg, which is higher than reported for the most active bacterial PPX1s, e.g. from *Chlorobium tepidum* TLS (590 U/mg polyP₃, K_M: 97.7 μ M) [197] (SI Tab. 2). In addition, high substrate affinity (5.24 μ M) was observed for polyP₄₅, which is similar to the *paPPX* from

3.1. Polyphosphate metabolism

P. aeruginosa with 7.14 μM [198]. Moreover, the *SaciPPX* requires divalent cations (Mg^{2+} or Mn^{2+}) for enzymatic activity, which is consistent with previous studies that PPX enzymes are metal ion dependent mainly accepting Mg^{2+} or Mn^{2+} for polyP hydrolysis [62, 197, 199]. In contrast to other PPXs from bacteria such as *E. coli* and *P. aeruginosa*, where activation of PPX activity was observed by addition of NH_4^+ or K^+ , no increased PPX activity was detected for *SaciPPX*. A slight inhibition on *SaciPPX* activity was detected for polyP₄₅ concentrations above 50 μM , but not by adding NaF, which matches the properties reported for *E. coli* PPX [62]. Moreover, during the ^{31}P NMR analysis no product inhibition was detected, contrary to the recognized P_i inhibition for the *E. coli* PPX, which showed a complete inhibition at 20 μM P_i [62].

The *SaciPPX* is closely related to the characterized PPX from *S. solfataricus* (45 % sequence identity to SSO1193, BlastX, $1e^{-121}$), which was reported to exhibit highest sequence similarity to bacterial-like PPX1s (25-45 %) [101]. The *E. coli* PPX1 presents a dimer, with each monomer (58.1 kDa) comprising four domains [200]. In particular, domain I and II are structurally related and share structural similarity with the ribonuclease H-like fold (ASHKA domain I and II), whereas domain III resembles the HD domain of SpoT and domain IV cold shock-associate RNA binding proteins. Notably, a similar four-domain structure was reported for the PPX2/GPPA of *E. coli*, and *Helicobacter pylori* [201], whereas a reduced size containing only domain I and II was reported for the PPX/GPPA from *Aquifex aeolicus* [202, 203]. Moreover, structural comparison of characterized enzymes revealed that the enzymes also differ in their oligomerization state [201]. While the PPX2/GPPA from *A. aeolicus* occurs as monomer the PPX1 from *E. coli* and PPX2/GPPA from *H. pylori* form dimers and the PPX2/GPPA of *E. coli* comprises a tetramer. In addition, significant structural differences in domain localization are found and thus the formation of different subfamilies within the bacterial PPX/GPPA enzyme family is suggested [201]. A previously constructed structural homology based model for SsoPPX (417 aa) based on EcoPPX1 (513 aa) showed that domain IV is lacking, whereas the suggested polyP binding sites, domain II and III, are well conserved [46].

The *SaciPPX* exhibits a dimeric structure and thus resembles the *E. coli* PPX1. Moreover, sequence alignments (SI Fig. 15) indicate that in the *SaciPPX* the P1, P2 and P3 loops comprising the catalytic residues responsible for hydrolysis of polyP are conserved [201]. *SaciPPX* exhibits a shortened C-terminal domain as reported for SsoPPX, where domain IV is lacking [46]. The enzymatic properties revealed a substrate specificity for long-chain polyPs in contrast to other PPX enzymes, thus a correlation between the structural difference and the enzymatic characteristics for *Saci* and SSoPPX can be discussed.

In addition to PPX, a highly active triphosphatase (triphosphate tunnel metalloenzyme (TTM), *Saci_0718*) of the CYTH (CyaB, thiamine triphosphatase) domain-like superfamily was

3.1. Polyphosphate metabolism

recently reported from *S. acidocaldarius* [71]. The enzyme is capable to catalyze the dephosphorylation of ATP and GTP to ADP and GDP, however, polyP₃ was identified to be the preferred substrate. Besides polyP₃ also polyP₄ was converted to a lower extent with the final formation of PP_i and P_i. The reaction product PP_i revealed a strong inhibitory effect on the TTM, where an equimolar substrate and inhibitor concentration resulted in less than 40 % residual activity. PP_i is generally regarded as waste product of the cell that needs to be hydrolyzed by pyrophosphatases in order to drive anabolic reactions such as DNA synthesis. A cytoplasmic inorganic pyrophosphatase (family I pyrophosphatase) was already described from *S. acidocaldarius* [204, 205] and *S. islandicus* [206]. Both enzymes showed an Mg²⁺-dependent activity with a pH optimum between 6 and 7 and were sensitive to NaF. The highest activity was observed for the *S. islandicus* (*Sisl*) enzyme at 50 °C, whereas the *Saci* enzyme was reported to be highly thermostable with a temperature optimum of 75 °C. The specific activity of *Sisl* and *Saci* enzymes was 850 U/mg and 860 U/mg, respectively. The K_M-values revealed a higher substrate affinity of *Saci* enzyme to PP_i compared to *Sisl* enzyme with K_M-values of 5.4 μM and 1.02 mM, respectively.

Our ³¹P NMR studies using polyP clearly demonstrated the non-enzymatic formation of cyclicP₃ (cP₃) in presence of metal ions at high temperature [192]. CP₃ was stable at high temperature and did not serve as substrate for PPX or triphosphatase (data not shown). Whether cP₃ is also formed *in vivo* and might be a substrate for other phosphatases remains to be elucidated.

Hence, *S. acidocaldarius* seems to possess at least three enzymes involved in polyP and PP_i degradation and thus phosphate liberation to enable phosphate homeostasis. They differ in preference for long chain polyP (PPX), short chain polyP (P₃- P₄, TTM) and PP_i (Fig. 6). The *Saci*PPX exhibits a high specific activity with up to 1200 U/mg for the preferred long-chain polyP₇₀₀, forms a dimer and lacks domain IV as reported for *Sso*PPX [46].

P_i is an essential nutrient and due to fluctuations in the environment, changes in the internal demands as well as its central role in cellular processes (e.g. protein phosphorylation, metabolism, as biological buffer) and as constituent of biological (macro)molecules (e.g. DNA, RNA, phospholipids, nucleotides) organisms have developed sophisticated mechanisms for P_i uptake and accumulation to ensure P_i homeostasis. The so-called pho regulon in bacteria [207-209] and the PHO pathway in yeast [210, 211] have been well studied. In the Euryarchaeon, *M. mazei* two high affinity ABC-type P_i transport complexes (*pstSCAB-phoU*) related to the bacterial pho regulon have been identified and their expression in response to P_i was analyzed [49]. In contrast, members of the Crenarchaeota were previously reported to lack the *pstSCAB* high affinity transport system (Pit low affinity system absent in *M. sedula*) and interestingly, homologs of the *S. cerevisiae* Pho84 major facilitator superfamily of

3.1. Polyphosphate metabolism

substrate transporters were identified (structure [212]) in *S. solfataricus*, *M. sedula* as well as *S. acidocaldarius* [46, 48, 213, 214]. However, detailed studies on the regulatory network in archaea are still missing.

The conserved gene organization of genes involved in polyphosphate metabolism has been observed in several bacterial and a few archaeal (methanogenic) species [62, 70]. By studying the gene neighborhood of the *ppx* gene in *S. acidocaldarius* we noticed the presence of two genes annotated as thymidylate kinases (*saci_2019* and *saci_2020*) in addition to a third candidate (*saci_0893*). Since dTMPKs have been characterized as the closest structural homologs of PPK2 [183, 184] we expressed and analyzed the three recombinant *Saci*dTMPKs for their enzymatic activity. Of the three dTMPKs only *Saci_0893* possessed thymidylate kinase activity, whereas *Saci_2019* revealed low PPK activity starting from ATP only. Although *Saci_2020* showed neither dTMPK nor PPK activity itself, however, PPK activity of *Saci_2019* (polyP synthesis) was significantly enhanced (6-fold) by the addition of *Saci_2020* (Tab. 1). Titration experiments revealed an optimal ratio of 1:1 for both proteins, referred to as heteromeric PPK, to achieve maximal specific activity (Tab. 1).

The strong inhibition of polyP synthesis by ADP (2.5 mM) pointed to reversibility of the reaction and indeed the homomeric enzyme *Saci_2019* as well as the heteromeric PPK showed a clear preference for ATP formation from ADP using polyP as phosphate donor, i.e. polyP-dependent nucleotide kinase activity. As reported for polyP synthesis the heteromeric enzyme comprised a significant increase of specific activity (90-fold) compared to *Saci_2019* alone.

Table 1: Summary of kinetic constants determined for recombinant enzymes involved in polyP metabolism (PPX and PPKs) and for the dTMPKs from *S. acidocaldarius*.

Enzyme	Oligomerization	Substrate/Cofactor	Specific activity [U/mg]	K _M [mM]	V _{max} [U/mg]	k _{cat} [1/s]	k _{cat} /K _M [M ⁻¹ s ⁻¹]
PPX (<i>Saci_2018</i>)	Dimer	polyP ₄₅ polyP ₇₀₀ P ₃	580.2 994.4 334.4	5.24 x 10 ⁻³	890	655.2	1.16 x 10 ⁸
<i>Saci_0893</i>	N.D.	dTMP/ATP	2.52				
<i>Saci_2019</i>	Tetramer	ATP		1.73	37 x 10 ⁻³	0.015	8.76
Heteromeric PPK (<i>Saci_2019</i> and <i>Saci_2020</i>)	Tetramer	ATP ADP polyP ₄₅		1.07 7.76 38.29	0.214 13.4	0.173 5.05	161.6 282.7 x 10 ³ 650.6

3.1. Polyphosphate metabolism

In order to analyze the reversibility of the PPK reaction we performed quantitative ^{31}P NMR (Fig. 5). Our *in vitro* studies allowed us to follow the enzymatic reaction resulting in the same ADP/ATP (25 %/75 %) equilibrium. Whereas the polyP-dependent ATP formation could easily be detected, the synthesis of polyP from ATP was observed only in the presence of an ATP recycling system. Notably, as soon as PEP for ATP recycling by PK was depleted the polyP synthesis stagnated and polyP hydrolysis took over by releasing cP_3 and P_i . Moreover, the same ADP/ATP ratio appeared like in the PPK assay (polyP formation) in the absence of the ATP recycling system (Fig. 5B) as well as in the polyP-dependent ATP formation reaction (Fig. 5A). Although, the enzyme had a higher specific activity for ATP formation, ADP was not completely converted even though polyP was still present. Therefore, it can be stated that regardless of the starting point and the reaction, the enzyme always remains in the same final ADP/ATP (25 %/75 %) ratio. This observed ADP/ATP equilibrium was also detected for other PPK enzymes (*E. coli* PPK1 and *S. meliloti* PPK2) via HPLC analysis, resulting in 70 % ATP and 30 % ADP independent from starting with ATP or ADP [54]. Thus, suggesting that the PPK enzyme is highly dependent on the cellular ADP/ATP ratio and is capable of keeping it constant through reversibility. This allows the cell to respond very quickly to a changing energy demand of the cell and to switch between energy storage and energy supply.

The need for an efficient ATP recycling system for polyP synthesis has been previously reported for the second PPK2 class I from *P. aeruginosa* (PA2428, named PPK2B) [215]. The observed ADP inhibition resulted in a decreased polyP synthesis and PPK specific activity, whereas the ATP recycling system led to a 10-fold increase in specific activity.

Thus, according to its enzymatic characteristics, the heteromeric *SacI*PPK is a typical class I PPK2 that catalyzes nucleoside diphosphate phosphorylation with strong preference for ADP. As recently reported [52, 183], class I enzymes differ significantly in their (co)substrate specificity and metal ion requirement. The first enzyme isolated from *P. aeruginosa* (PPK2A, class I) showed a 75-fold preference for nucleoside diphosphate kinase (NDK) activity (compared to NMP) although with GDP favored over ADP [185]. Notably Mg^{2+} was the preferred cofactor for GDP/ADP phosphorylation whereas Mn^{2+} significantly enhanced polyP synthesis. Moreover, PPK2 from *A. johnsonii* was reported to switch between NDP and polyP synthesis in response to low or high concentrations of Mg^{2+} (100 mM MgCl_2 for AMP phosphorylation) [216]. By contrast, PPK2B from *Corynebacterium glutamicum* exhibited preference for Mn^{2+} -dependent polyP synthesis from ATP/GTP (e.g. [217]) and the reversible PPK2c from *R. eutropha* possessed broad substrate specificity and accepted different purine and pyrimidine nucleotides [218]. However, we could not observe this effect for the *SacI*PPK and identified Mg^{2+} as the best cofactor for both reactions, polyP and ATP synthesis.

3.1. Polyphosphate metabolism

In terms of its heteromeric structure ($\alpha_2\beta_2$) and activation of Saci_2019 by Saci_2020, the SaciPPK resembles two domain enzymes found so far only in class II PPK2s. PPK2 class II enzymes comprise both one PPK2 domain and two PPK2 domain fusion proteins (e.g. PPK2C from *P. aeruginosa* and PPK2 from *A. johnsonii* [219]). Interestingly the *P. aeruginosa* class II PPK2C has been demonstrated to contain a catalytically inactive N-terminal and active C-terminal domain [184], similar to the subunit composition of the heteromeric SaciPPK with an active (Saci_2019) and inactive subunit (Saci_2020).

Moreover, class III PPK2 enzymes were discussed to share a common ancestor with dTMPKs due to their high structural similarity [183]. Sequence comparison of both SaciPPK subunits with sequences of PPK2 and dTMPK enzymes with solved structures (SI Fig. 16) revealed the difference between the two enzymes with respect to the presence of the catalytically important residues. First, it should be mentioned that PPK2 and dTMPK enzymes share some catalytic residues and motifs, such as the P-loop, DRX motif and the LID region. Several structural analyses revealed that residues inside the P-loop (e.g. Lys, Val, Asp, Pro) are involved in nucleotide binding in dTMPKs and PPK2s [220, 221]. In addition, the DRX motif contains an Asp residue that is discussed to play a central role in phosphate transfer between phosphorus donor and acceptor in both enzyme families [220, 222]. Comparing the sequences of the two subunits of heteromeric SaciPPK both with each other and with the PPK2 and dTMPK sequences, it becomes evident that the two proteins differ significantly in terms of the presence of catalytically important residues. The Asp residue within the P-loop (position 3) is present in all PPK2 and dTMPK sequences considered (except *S. aureus* dTMPK, comprising a Glu) and in Saci_2019, except in Saci_2020 where it is replaced by Met. Moreover, the Arg residue in the DRX motif is conserved among the two enzyme families, but only Saci_2019 possess this catalytic residue in contrast to Saci_2020. The LID region of PPK2 enzymes is elongated compared to dTMPK enzymes, with a higher number of conserved residues (Arg, Lys Asp) for PPK2 enzymes involved in Mg²⁺ coordination and polyP binding. For example, *P. aeruginosa* PPK2 contains two Arg and one Lys residue involved in polyP binding. In contrast, dTMPKs typically exhibit only one Arg residue responsible for dTMP binding present at a non-conserved position (*P. falciparum* at position 3) [223-225]. Interestingly, the LID region of *S. acidocaldarius* and *Sulfurisphaera tokodaii* enzymes is longer than the dTMPK and shorter than PPK2 LIDs, where Saci_2019 and Saci_2020 contain more than one Arg residue, which might interact with polyP. Despite the high sequence similarity of Saci_2019 and Saci_2020, the findings from the sequence comparisons and the biochemical characterization of the recombinant enzymes are consistent, since Saci_2020 is lacking most of the conserved residues. Regarding classification of the heteromeric PPK, the sequence analysis revealed a closer relation to PPK2 class III enzymes (e.g. *M. ruber*) for both subunits (Saci_2019 and Saci_2020), although the characteristic Asp residue, involved in ATP binding, is not present

3.1. Polyphosphate metabolism

at +4 position adjacent to the DRX motif. Moreover, the *Saci*PPK shows a closer sequence similarity to dTMPKs than other known PPK2 class III enzymes.

In summary, the heteromeric PPK from *S. acidocaldarius* shares some individual features with the known PPK2 class proteins. Due to the enzymatic characteristics, it resembles a class I enzyme by only accepting NDP for polyP-dependent NTP synthesis. In addition, the heteromeric enzyme includes a two-domain structure, also reported for class II PPK2 enzymes. However, sequence analysis clearly showed a closer relationship with class III PPK2s and high similarity with dTMPKs. A more detailed phylogenetic analysis will provide further information on the correct classification, but here, according to the biochemical property, the enzyme is termed a dTMPK-like PPK2 class I.

Conclusion

In the current study, we address the physiological function of polyP in the crenarchaeal model organism *S. acidocaldarius* by identification and characterization of enzymes involved in polyP metabolism (Fig. 6). Our studies show that polyP possesses a fundamental significance as phosphate and energy storage in *S. acidocaldarius*. Phosphate is an essential nutrient and therefore sophisticated mechanisms are required to ensure P_i homeostasis in the cell. So far, P_i uptake has not been studied in Crenarchaeota but there is some evidence that they utilize homologs of the *S. cerevisiae* Pho84 transporter [213, 214]. Depending on the cellular demand P_i will be either directly used in metabolic and cellular processes or if the energy charge of the cell is high P_i surplus can be stored as polyP via the ATP-dependent *SacI*PPK2 activity. When the P_i demand of the cell increases long chain polyP can be degraded by the PPX and/or short chain polyP by the triphosphatase (TTM, [71]) under renunciation of energy gain. However, depending on the energy requirements of the cell, polyP can also be used as a source of energy. At low energy charge with ATP demand, the polyP-dependent nucleotide kinase activity of the *SacI*PPK2 enables ATP formation. Thereby the energy stored in the phosphoanhydride bonds of polyP can be directly used to phosphorylate ADP without any energy loss. PPKs thus represent an interesting energy recycling system that is already used in various biotechnological applications such as enzyme cascades [226]. The dual function of polyP as P_i storage or energy storage requires a sophisticated synthesis/degradation regulation in order to adapt to fluctuations in the environment and internal demands, which still has to be elucidated in archaea. Moreover, the PPK2 reaction yielded the same ratio of ADP/ATP (25%/75%) regardless of the starting point. The observed phenotypes in response to *ppx* overexpression with polyP depletion resulting in decreased energy-dependent motility and thus adhesion and biofilm formation [88] support this scenario. Finally, the identification of the novel thymidylate kinase-like PPK2 class I in members of the *Sulfolobales* highlights once again the endowment of archaea with novel enzymes from different enzyme families that perform the same function as their bacterial or eukaryotic counterparts.

3.1. Polyphosphate metabolism

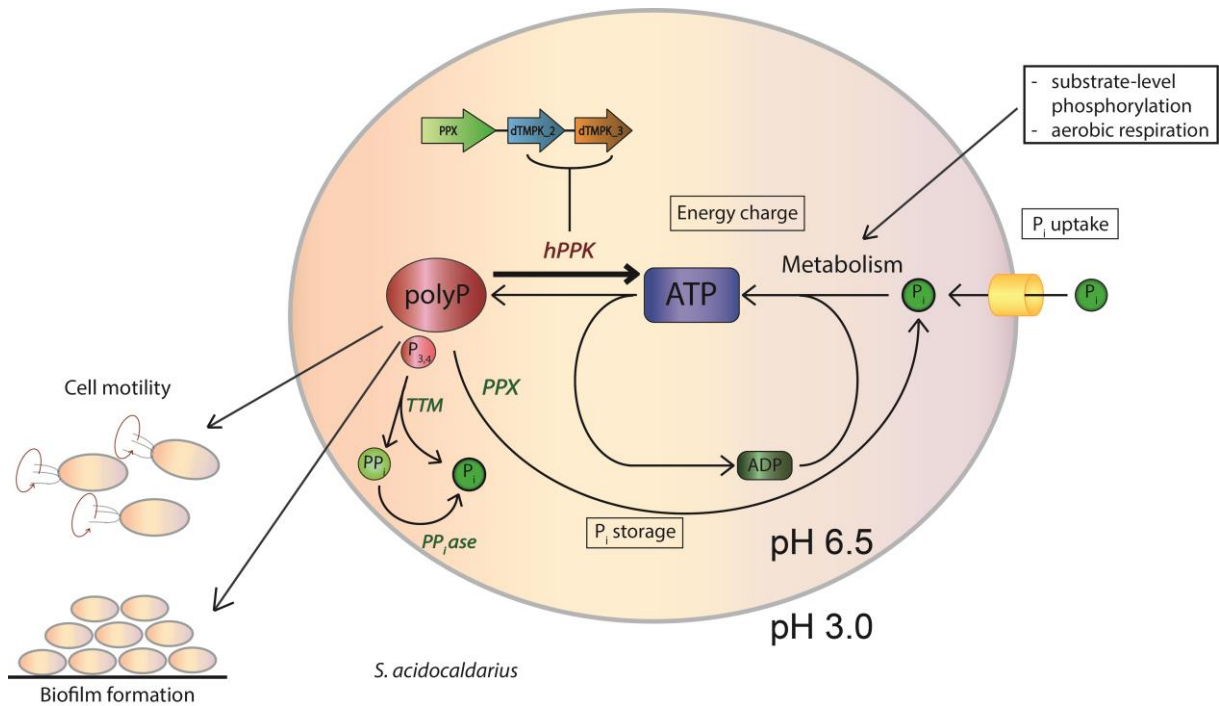


Figure 6: PolyP metabolism in *S. acidocaldarius*.

PolyP is synthesized by heteromeric PPK by using ATP as phosphoryl donor as well as utilized by the same enzyme to produce ATP from ADP and polyP. Preferably, long-chain polyP degradation is catalyzed by PPX where short-chain polyP₃ and polyP₄ is hydrolyzed by TTM releasing PP_i and Pi. A cytoplasmic PP_iase was identified to catalyze PP_i degradation. Thus, polyP is involved in Pi storage and energy charge in *S. acidocaldarius* and a lack of polyP was reported to reduce energy-dependent cell motility and enhance adhesion and biofilm formation.

Materials and methods

Strains and growth conditions

The *Escherichia coli* strains DH5 α for cloning were grown in Luria-Bertani (LB) medium, whereas Rosetta (DE3) strains were grown in Terrific broth (TB) medium (22 g/L yeast extract, 12 g/L tryptone/peptone from casein, 0.4 % (v/v) glycerol), containing the appropriate antibiotics (100 μ g/mL ampicillin (pET15b), 50 μ g/mL kanamycin (pET28b), 25 μ g/mL chloramphenicol (Rosetta (DE3))) at 37 °C.

Cloning, heterologous expression and purification of the recombinant proteins

The *ppx* (*saci_2018*, GenBank: AAY81314.1), *dtmpk* (*saci_0893*, GenBank: AAY80256, *saci_2019*, GenBank: AAY81315.1, and *saci_2020*, GenBank: AAY81316.1) genes were amplified via PCR using genomic DNA of *S. acidocaldarius* DSM639 as template (for primers see SI Tab. 3). PCR fragments were cloned into the vector pET15b (*ppx*, *dtmpk:saci_0893*, *saci_2019*) and pET28b (*dtmpk:saci_2020*) and successful cloning was confirmed by sequencing. The genes were cloned with the sequence of an N-terminal 6x His-Tag which originates from the pET vector system. The expression vectors were used for heterologous gene expression in *E. coli* Rosetta (DE3) cells grown on TB-medium and the induction was performed at OD₆₀₀ 0.5-0.6 with 1 mM isopropyl- β -d-thiogalactopyranoside (IPTG, Carl Roth GmbH + Co. KG, Karlsruhe, Germany). After incubation overnight at 18 °C with shaking (180 rpm) and cell harvest (6.000 x g, 15 min, 4 °C), the cell pellet (1 g cells (wet weight)) was resuspended in 3 ml 50 mM TRIS/HCl, 300 mM NaCl, pH 8.0 supplemented with protease inhibitor (Merck KGaA, Darmstadt, Germany) and disrupted by french press (20.000 psi). In addition, thermally unstable *E. coli* proteins were removed by heat precipitation (20 min at 70 °C with gentle agitation) followed by centrifugation (25.000 x g, 45 min, 4 °C). The cleared lysate with the His-tagged proteins were purified via Protino[®] Ni-TED column (Machery & Nagel, Düren, Germany) following the instructions of the manufacturer. 10 mM TRIS/HCl, 300 mM NaCl, 250 mM imidazole, pH 8.0 was used as elution buffer. Purified proteins (PPX, dTMPKs) were concentrated via ultrafiltration (Vivaspin concentrator MW cut off 10 kDa, Sartorius AG, Göttingen, Germany) and then applied to size exclusion chromatography (Superose[®] 6 Increase 10/300 GL, Cytiva Lifescience™, Freiburg, Germany) using 50 mM TRIS/HCl with 300 mM NaCl pH 8.0 as buffer for the separation of *Saci_2018*, homomeric PPK (*Saci_2019* or *Saci_2020*) as well as the heteromeric PPK (*Saci_2019* and *Saci_2020* (1:1 ratio)). A calibration curve was generated by running ribonuclease A (13.7 kDa) ovalbumin (44 kDa), conalbumin (75 kDa), aldolase (158 kDa) and ferritin (440 kDa). The oligomeric state of native *Saci_2018*, *Saci_2019*, *Saci_2020* was calculated with the observed elution volumes, corresponding to a molecular weight of 81 kDa (*Saci_2018*),

3.1. Polyphosphate metabolism

88 kDa (Saci_2019) and 107 kDa (Saci_2020). The pure proteins were stored at -80 °C with 15 % glycerol and used for further enzymatic characterization. Purification and molecular mass of subunits were monitored by SDS-polyacrylamide gel electrophoresis.

Enzyme assays

Exopolyphosphatase (PPX) - PPX activity was determined by detecting P_i liberation with a discontinuous malachite green assay as described previously [177]. Briefly, the reaction was performed in 50 mM TRIS/HCl, pH 7.2 with 200 mM NH_4Cl and 2.5 mM $MgSO_4$ in the presence of 1.45 μ g PPX in a total volume of 200 μ L. After preincubation for 90 seconds at 70 °C the reaction was started by addition of the substrate poly P_{45} (0 - 50 μ M). During the reaction 20 μ L aliquots were removed from the samples every 5 min and the reaction was stopped by adding 300 μ L of the malachite green reagent. The reagent was prepared using 1.5 mL of 1 % (w/v) malachite green oxalate, 30 mL of 1 % (w/v) ammonium molybdate both solved in distilled H_2O , 4.64 mL of 70 % (v/v) $HClO_4$, 13.86 mL of H_2O . After stirring 30 min the solution was supplemented with 0.45 mL of 5 % (v/v) Triton X-450 solved in H_2O and filtered [227]. A standard curve was prepared using 20 μ L of 0-500 μ M KH_2PO_4 (following [177]) and 300 μ L of the malachite green reagent. After 20 min of incubation at room temperature (RT) the absorption was determined at 650 nm [227]. One unit was defined as one μ mol P_i released per min. To determine substrate specificity pppGpp (0.1 mM), poly P_3 (0.5 mM), poly P_{45} (20 μ M) and poly P_{700} (1 μ M) were tested. The temperature dependence was assayed with 20 μ M poly P_{45} at 50 – 90 °C, the pH was adjusted at the respective temperature. The dependence of divalent cations was determined with 25 μ M poly P_{45} and $MgCl_2$ and $MnCl_2$ (2.5 mM) or without divalent cations. To investigate activating or inhibiting effects NaF (1 mM), KCl (80 mM) or NH_4Cl (25 mM) were added to the PPX reaction mixture in 50 mM TRIS/HCl, pH 7.2 The kinetic properties were assayed with 0 – 50 μ M poly P_{45} with 2.5 mM $MgCl_2$ at 70 °C.

For the PPX ^{31}P NMR measurements the same reaction mixture was used containing 500 μ M poly P_{45} as substrate, 14 μ g PPX and 10 % D_2O in a total volume of 1 ml. The reaction was followed over 330 min at 70 °C, where one spectrum was recorded every 30 min. The signals obtained were assigned (see SI) and integrated over a list of the spectra with normalizing the sum of integrals (set to 1).

3.1. Polyphosphate metabolism

Thymidylate kinase (dTMPK) – For qualitative analysis via ^{31}P NMR the reaction mixtures contained 2 mM ATP, 1 mM dTMP, 4 mM MgCl_2 , 50 mM KCl, 10 μg of the respective enzyme (Saci_0893, Saci_2019, Saci_2020, or 5 μg of Saci_2019 and Saci_2020, respectively) in 50 mM TRIS/HCl pH, 7.4 with 10 % D_2O . The samples were incubated at 70 °C for 8 hours and analyzed via ^{31}P NMR spectroscopy. Moreover, dTMPK activity was assayed with 1 mM dTMP, 2 mM ATP, 50 mM KCl, 4 mM MgCl_2 and 10 μg Saci_0893 in 50 mM TRIS/HCl, pH 7.4 at 70 °C [228] and the ADP, P_i and polyP formation were determined discontinuously with the PK-LDH assay at 37 °C, malachite green assay and polyP assay, respectively.

Additionally, a dTMPK assay was performed with 1 mM dTMP, 2 mM ATP, 4 mM MgCl_2 and 10 μg Saci_0893 in 50 mM TRIS/HCl, pH 7.4 coupled with PK-LDH for investigation of Saci_0893 specific activity. The detection of ADP by the continuous PK-LDH assay is described below, which is described below. One unit was defined as one μmol NADH oxidized per min.

Polyphosphate kinase (PPK) - PPK activity was determined in a discontinuous assay as described previously [229]. Briefly, the reaction was performed in 50 mM TRIS/HCl, pH 7.4 with 50 mM KCl and 4 mM MgCl_2 in the presence of 10 μg enzyme (Saci_2019, Saci_2020) in a total volume of 1 ml. After preincubation for 90 seconds at 70 °C, the reaction was started by addition of the substrate ATP (2 mM). During the reaction, samples were taken from each reaction every 30 min and the produced polyP, ADP and P_i were determined discontinuously using the polyP assay, a discontinuous PK-LDH assay and the malachite green assay, respectively. One Unit was defined as one μmol polyP produced or ADP released per min.

To determine the optimal molar ratio of both enzymes (Saci_2019 and Saci_2020) the PPK activity was determined in relation to ADP formation adding 0 – 5 μg Saci_2020 to 5 μg Saci_2019 in a total reaction volume of 0.5 mL. Additionally, a titration experiment was performed using both enzymes in different ratios from 10:0 to 0:10 (10 μg enzyme in total) for the PPK assay, and polyP (polyP assay) and ADP formation (PK-LDH assay) were detected as described.

The temperature dependence of homomeric and heteromeric PPK (10 μg Saci_2019 and 5 μg Saci_2019 and Saci_2020, respectively) was determined in the range of 50 to 90 °C at pH 7.4 adjusted at the respective temperatures. The pH dependence was analyzed for homomeric and heteromeric PPK in a mixed buffer (50 mM TRIS/HCl, Citrate/HCl, BICINE/NaOH and MES/NaOH with the respective pH (4 – 9) at 70 °C. The kinetic properties of PPK were assayed with 0 - 8 mM ATP (homomeric PPK) and 0 - 4 mM ATP (heteromeric PPK) with 4 mM MgCl_2 , 50 mM KCl and 10 μg of enzyme (Saci_2019, heteromeric PPK

3.1. Polyphosphate metabolism

(5µg Saci_2019 and Saci_2020, respectively)) in 50 mM Tris/HCl, pH 7.4 at 70 °C. ADP formation was determined discontinuously by the PK-LDH assay (see below). Additionally, enzymatic activity was tested using 2 mM GTP as substrate for heteromeric PPK. To investigate ADP inhibition 2.5 mM ADP were added to the reaction mixture of heteromeric PPK containing 1.5 mM ATP as substrate and polyP formation was determined discontinuously.

Reverse PPK activity was assayed using a discontinuous assay detecting ATP formation and polyP degradation. The reaction mixture contained 2 mM ATP/GTP, 50 µM polyP₄₅, 10 mM MgCl₂, 50 mM KCl and 5 µg of homomeric or heteromeric PPK in 50 mM TRIS/HCl, pH 7.4. The polyP and ATP concentration was determined via polyP assay and HK-G6PDH assay (see below), respectively. One Unit was defined as one µmol polyP utilized or one µmol NADP⁺ reduced per min.

The kinetic properties were determined with 0-40 mM ADP with 100 µM polyP₄₅, whereas polyP₄₅ dependence (0-100 µM) was assayed with 20 mM ADP (10 mM MgCl₂, 50 mM KCl and 5 µg of homomeric or heteromeric PPK in 50 mM TRIS/HCl, pH 7.4). Heteromeric PPK activity was tested with different divalent cations, namely Mg²⁺, Mn²⁺, Ca²⁺, Co²⁺ and Ni²⁺ with a final concentration of 10 mM, in regard to ADP (3 mM ATP, 10 µg heteromeric PPK) and ATP formation (20 mM ADP, 100 µM polyP₄₅, 5 µg heteromeric PPK). The Mg²⁺-dependence of the PPK activity was analyzed using 0 - 12.5 mM MgCl₂ (ATP formation) and 0 – 100 mM (polyP formation) in 50 mM TRIS/HCl, pH 7.4, 50 mM KCl at 70 °C. The specific activity was determined via the discontinuous HK-G6PDH (ATP formation) or PK-LDH (ADP formation) assay at 37 °C.

For qualitative analysis, the PPK enzyme reactions and control samples were followed via ³¹P NMR measurements at 70 °C for 12 hours. The reaction mixture following ATP formation contained 10 mM ADP, 500 µM polyP₄₅, 10 mM MgCl₂, 50 mM KCl, 20 µg heteromeric PPK in 50 mM TRIS/HCl, pH 7.4 with 10 % D₂O. The reaction mixture following polyP synthesis contained 2 mM ATP, 4 mM MgCl₂, 50 mM KCl, 20 µg heteromeric PPK in 50 mM TRIS/HCl, pH 7.4 with 10 % D₂O. The PPK assay for polyP formation with ATP recycling system included 2 mM ATP, 10 mM PEP, 20 mM MgCl₂, 50 mM KCl, 8 µg SsoPK [230], 20 µg heteromeric PPK or 20 µg homomeric PPK in 50 mM TRIS/HCl, pH 7.4 with 10 % D₂O. To follow PEP hydrolysis a sample, without polyP or ATP, containing 10 mM PEP, 20 mM MgCl₂, 50 mM KCl, 20 µg heteromeric PPK in 50 mM TRIS/HCl, pH 7.4 with 10 % D₂O was analyzed. To investigate polyP and ADP hydrolysis a sample with 10 mM ADP, 500 µM polyP₄₅, 10 mM MgCl₂, 50 mM KCl, in 50 mM TRIS/HCl, pH 7.4 with 10 % D₂O was analyzed. The reactions were followed over 700 min (for PPK reaction (polyP formation) without ATP recycling only 300 min) at 70 °C, where one merged spectrum was recorded every 15 or

3.1. Polyphosphate metabolism

30 min. The signals obtained were assigned (see SI) and integrated over a list of the spectra, where the sum of integrals was normalized.

PK-LDH assay following ADP formation (polyP formation)

The enzymatic activity homomeric and heteromeric PPK for polyP formation was determined by coupling ADP formation to NADH oxidation via pyruvate kinase (PK, rabbit muscle, (Merck, Darmstadt, Germany)) and L-lactate dehydrogenase (LDH, rabbit muscle, (Merck, Darmstadt, Germany)). The assay contained 2 mM phosphoenolpyruvate (PEP), 4 mM MgCl₂, 50 mM KCl, 0.5 mM NADH, 8 U pyruvate kinase (PK, rabbit muscle, (Merck, Darmstadt, Germany)), 4 U L-lactate dehydrogenase (LDH, rabbit muscle, (Merck, Darmstadt, Germany)) in 50 mM TRIS/HCl pH 7.4. The continuous assay for dTMPK assay (Saci_0893) was carried out at 55 °C, whereas the discontinuous assay (PPK assay in polyP formation direction) was performed at 37 °C for 15 min. The NADH consumption was detected spectrophotometrically at 340 nm (5.8 mM⁻¹ cm⁻¹). One unit was defined as one μmol NADH oxidized per min.

HK-G6PDH assay following NTP formation

The enzymatic activity of homomeric and heteromeric PPK for NTP formation was determined by coupling the ATP/GTP formation to the NADP⁺ reduction via the hexokinase (HK) and glucose-6-phosphate dehydrogenase (G6P-DH) (Roche, Basel, Switzerland). The reaction mixture contained 2 mM glucose, 10 mM MgCl₂, 50 mM KCl, 500 μM NADP⁺, 1 U hexokinase (HK, yeast) (Roche, Basel, Switzerland) and 1 U glucose-6-phosphate dehydrogenase (G6P-DH, yeast) (Roche, Basel, Switzerland) in 50 mM TRIS/HCl, pH 7.4. The reaction was performed in a total volume of 200 μL at 37 °C for 15 min and NADPH formation was detected spectrophotometrically at 340 nm to determine the final ATP or GTP concentration.

PolyP assay

The MicroMolar Polyphosphate Assay Kit (ProFoldin, Massachusetts, United States) was used for determination of polyP concentration according to the manufacturer's instructions. Briefly, 30 μL of sample were mixed with 30 μL of the fluorescent dye (PPD dye) and fluorescence intensity was determined (emission 550 nm, excitation 415 nm), where 0 – 50 μM polyP₄₅ were used for the standard curve.

3.1. Polyphosphate metabolism

³¹P NMR analysis

For detailed analysis of the phosphate compounds, ³¹P NMR analysis was used. Therefore, ¹H- and ³¹P NMR spectra were recorded on an DRX 300, DRX 500 or AVANCE NEO 400 MHz nuclear magnetic resonance spectrometer (Bruker, Billerica, USA) at ambient temperature in 90 % H₂O and 10 % D₂O (for NMR lock), where ¹H-NMR spectra were recorded only for NMR lock. The ³¹P NMR chemical shifts are given in ppm (for AVANCE NEO 400 MHz); splitting patterns are given as singlet (s), doublet (d), triplet (t), plus coupling constants (J) are reported in Hz. The spectra were analyzed with TopSpin Software (Bruker, Billerica, USA). The NMR spectra are shown in the supplemental material (SI). The spectra were integrated using the multiple method for a list of spectra with normalization setting the sum of integrals to 1 for each spectrum. Since the relaxation time was not determined, the concentrations were not recalculated. Starting from the integrals obtained, the relative molar ratio was calculated by dividing the integrals by the number of phosphorus atoms in the molecules giving the respective signal (e.g. three P atoms in cP₃). For each compound standard ³¹P NMR was performed, revealing the signals summarized in the SI part.

Supplementary Materials

Discovery of the missing polyphosphate kinase in the thermoacidophilic Crenarchaeon *Sulfolobus acidocaldarius* by studying polyphosphate metabolism

Svenja Höfmann¹, Felix Niemeyer², Torsten Schaller², Christopher Bräsen¹, Bettina Siebers^{1*}

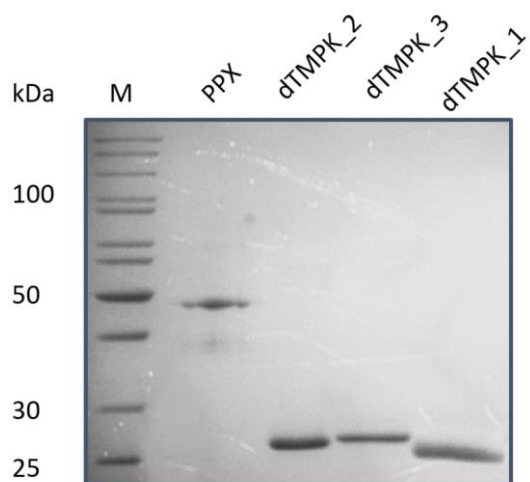
¹Molecular Enzyme Technology and Biochemistry (MEB), Environmental Microbiology and Biotechnology (EMB), Centre for Water and Environmental Research (CWE), Department of Chemistry, University of Duisburg-Essen, Universitätsstrasse 5, 45141 Essen, Germany

²Faculty of Chemistry (Organic Chemistry) and Center for Nanointegration Duisburg-Essen (CENIDE), University of Duisburg-Essen, Universitätsstrasse 7, 45141 Essen, Germany

*Corresponding authors: Prof. Dr. Bettina Siebers (Biochemistry): bettina.siebers@uni-due.de,
Tel.: +49 (0)201 183-7061, Fax: +49 (0)201 183-7062

3.1. Polyphosphate metabolism

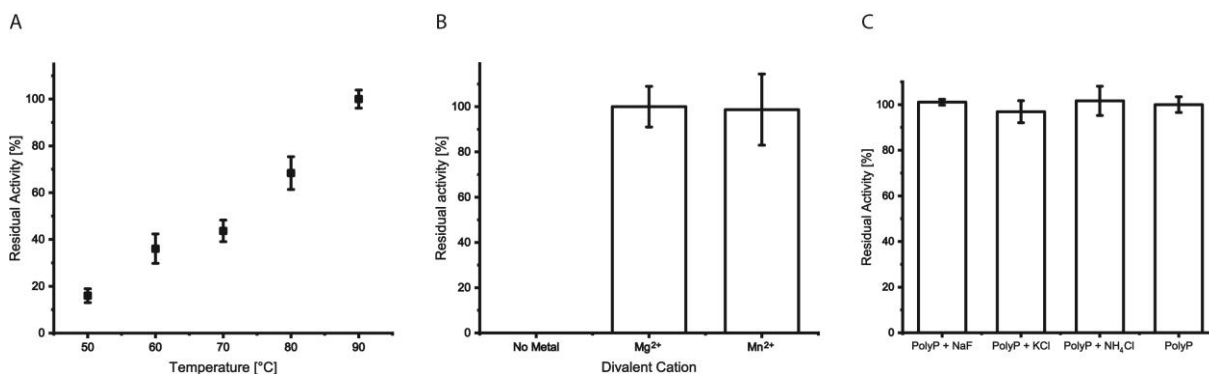
Protein purification



SI Figure 1: The PPX and three annotated dTMPKs (Saci_2019, Saci_2020 and Saci_0893) from *S. acidocaldarius* after recombinant expression in *E. coli* Rosetta (DE3) and purification via affinity chromatography (Ni-TED column) (SDS-PAGE and Coomassie staining).

PPX (*saci_2018*), the two subunits of PPK dTMPK_2,3, (*saci_2019*, *saci_2020*) and thymidylate kinase dTMPK_1 (*saci_0893*). For all enzymes 2 μ g of protein were applied. M, protein marker (PageRuler Unstained Protein Ladder 10-180 kDa, Thermo Fisher Scientific, Schwerte, Germany).

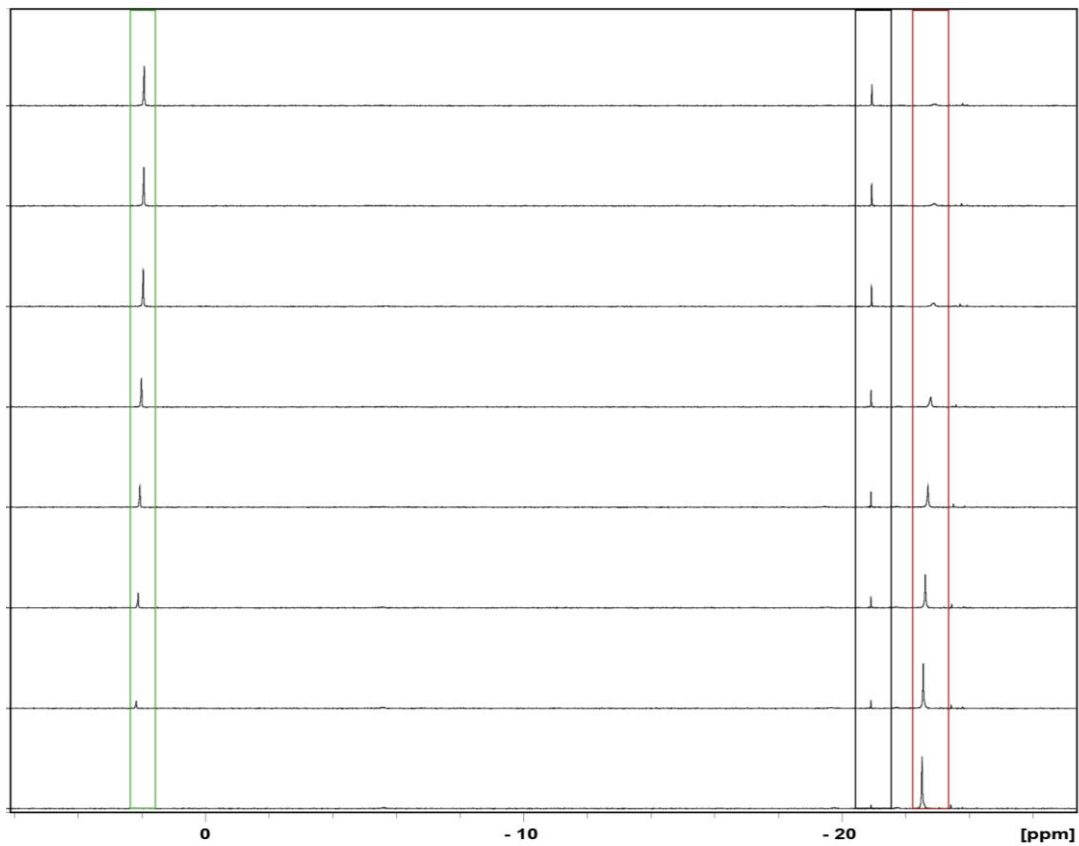
Enzyme characterization



SI Figure 2: Temperature and metal ion dependence as well as effect of inhibitors/activators on enzyme activity of the recombinant PPX from *S. acidocaldarius*.

The enzymatic activity was determined at 70 °C with 1.5 μ g of purified enzyme, where P_i liberation was determined discontinuously via the Malachite Green Assay at RT. (A) The temperature optimum of the PPX was investigated in a temperature range from 50 to 90 °C using 20 μ M polyP₄₅ as substrate. The residual activity is shown with highest specific activity (1234 U/mg protein) at 90 °C. (B) Dependence on divalent cations was tested in absence as well as presence of Mg²⁺ and Mn²⁺ (2.5 mM). Highest activity is observed in presence of Mg²⁺ (2.26 U/mg protein). (C) Effect of addition of 1 mM NaF, 80 mM KCl or 25 mM NH₄Cl in 50 mM TRIS/HCl, pH 7.2 on PPX activity was determined (100 % residual activity was 381 U/mg). Three independent measurements (n=3) were performed and error bars indicate the standard error of the mean (SEM).

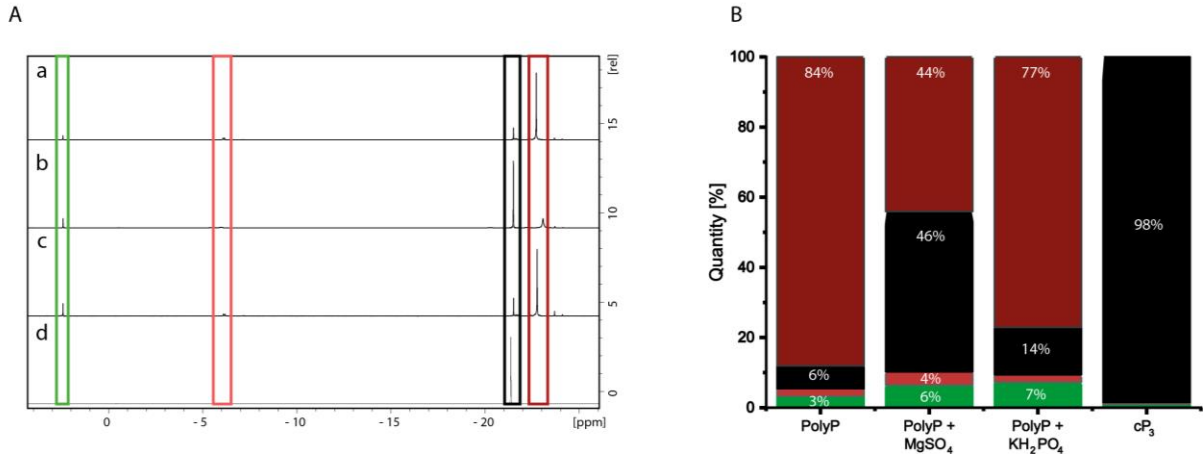
3.1. Polyphosphate metabolism



SI Figure 3: Stacked ^{31}P NMR spectra of PPX kinetic measurement.

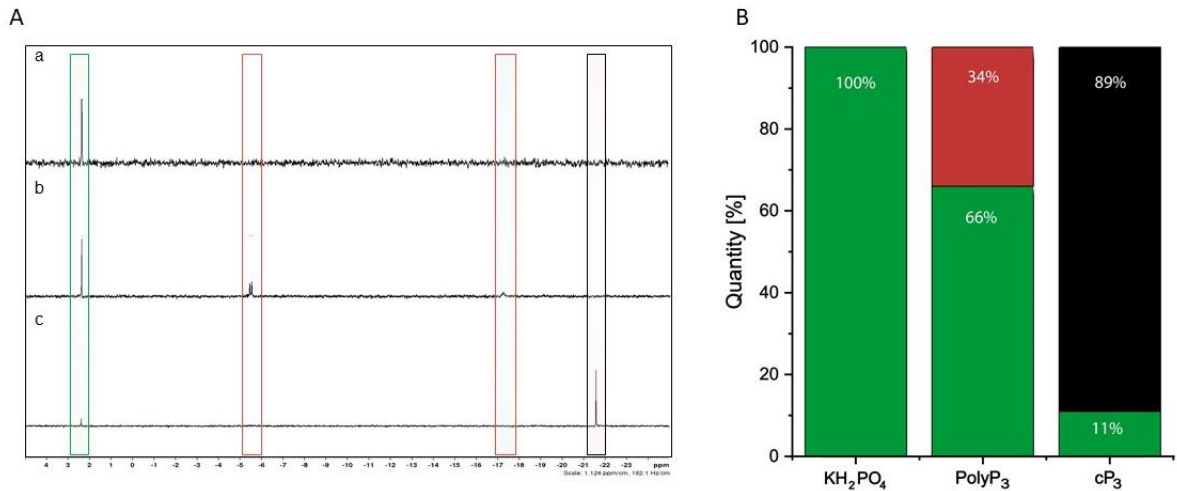
The reaction was performed in 50 mM TRIS/HCl, pH 7.2 with 200 mM NH_4Cl , 2.5 mM MgSO_4 , 500 μM poly P_{45} in the presence of 14 μg PPX in a total volume 1 ml with 10 % D_2O . The P_i (green box) and cP_3 (black box) formation as well as the polyP consumption (red box) could be detected via ^{31}P NMR. The reaction was followed from 0 min (bottom spectrum) to 330 min (top spectrum) at 70 °C.

3.1. Polyphosphate metabolism



SI Figure 4: Stability of different phosphate compounds in presence and absence of divalent (MgSO₄) and monovalent ions (KH₂PO₄) detected via ³¹P NMR (A) and quantification of obtained signals (B).

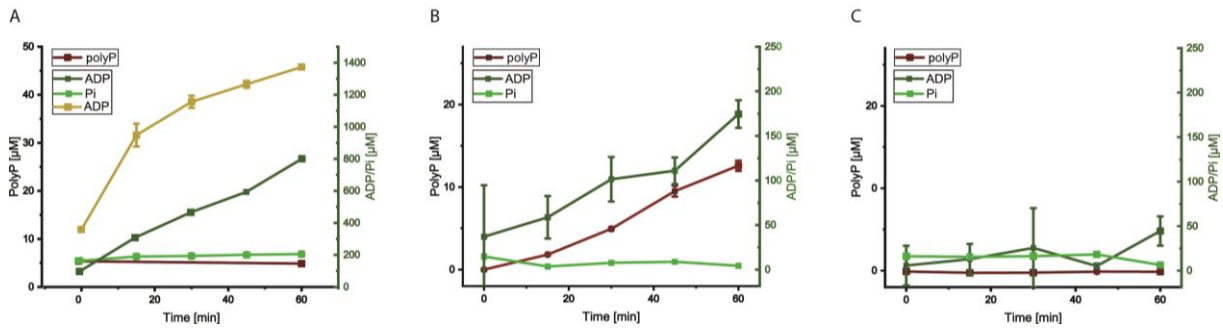
The samples were incubated at 70 °C overnight in 50 mM TRIS/HCl pH 7.2, 200 mM NH₄Cl with 500 μM polyP₄₅ (a), 500 μM polyP₄₅ with 2.5 mM MgSO₄ (b), 500 μM polyP₄₅ with 500 μM KH₂PO₄ (b) and 100 mM cP₃ (d) as reaction compounds. (A) The NMR data of the different samples are shown and (B) the quantification of phosphates was determined by evaluation of ³¹P NMR spectra, where P_i (green), polyP₃ (light red), cP₃ (black) and polyP (red) were detected.



SI Figure 5: Activity of recombinant PPX from *S. acidocaldarius* with P_i as control (a), polyP₃ (b) and cP₃ (c) detected via ³¹P NMR (A) and quantification of obtained signals (B).

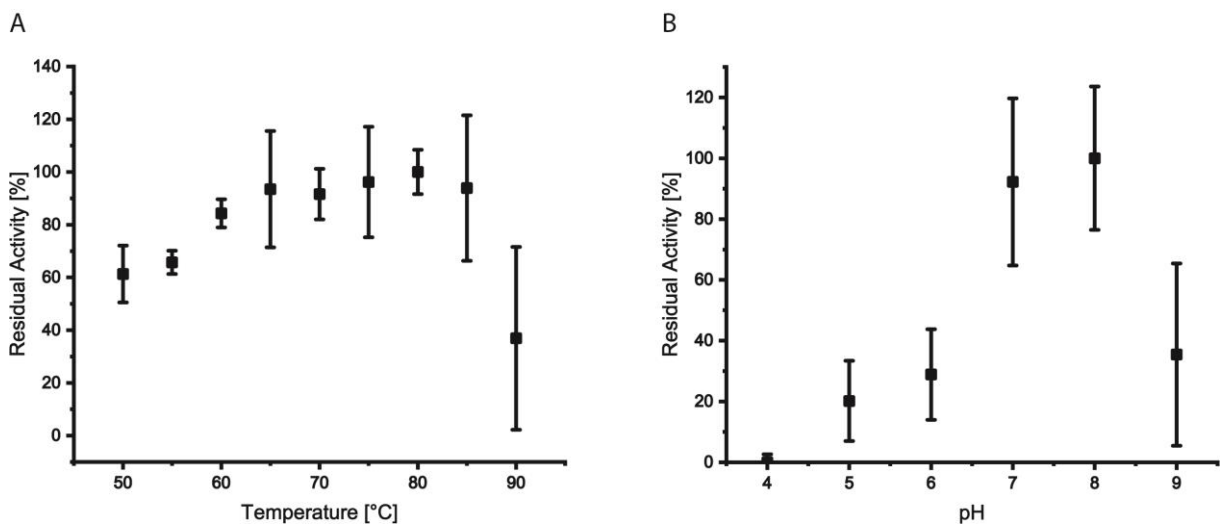
PPX activity assay was performed at 70 °C overnight in 50 mM TRIS/HCl pH 7.2, 200 mM NH₄Cl, 2.5 mM MgSO₄ with 2.5 μg PPX and as substrate 500 μM KH₂PO₄ (a), polyP₃ (b) and cP₃ (c). (A) ³¹P NMR spectra were recorded and signals from P_i (green box), terminal and middle phosphates from polyP₃ (light red), cP₃ (black box) could be detected. (B) Quantification of signals for KH₂PO₄ as substrate (control), polyP₃ and cP₃. Notably, using cP₃ as substrate, P_i formation did not increase after longer incubation, indicating polyP₃/PP_i impurities.

3.1. Polyphosphate metabolism



SI Figure 6: Enzymatic activity of the three thymidylate kinase homologs Saci_0893 (dtMPK_1), Saci_2019 (dtMPK_2) and Saci_2020 (dtMPK3).

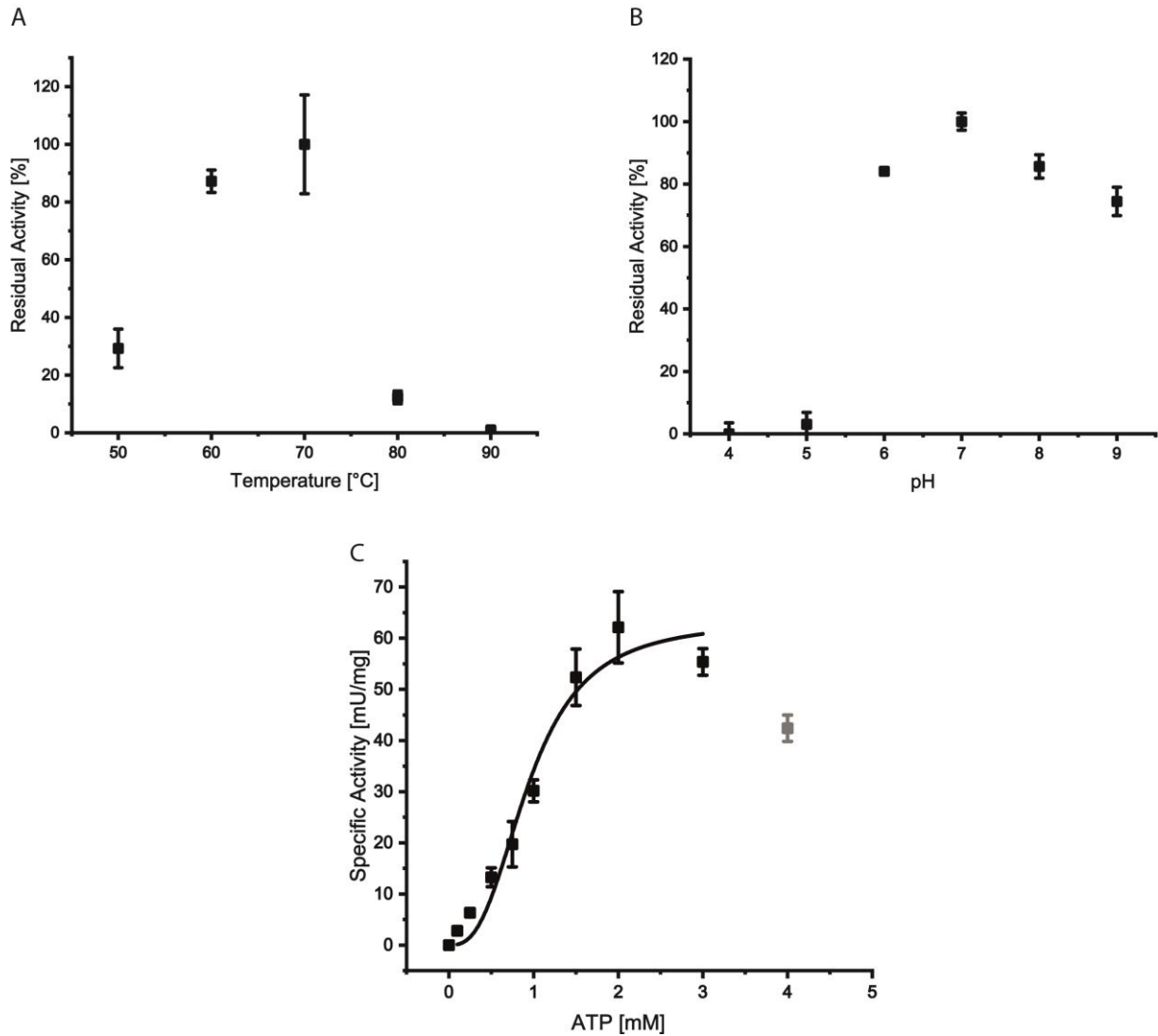
The dTMPK assay was performed with 1 mM dTMP (only for Saci_0893), 2 mM ATP, 50 mM KCl, 4 mM MgCl₂ and 10 μg enzyme in 50 mM TRIS/HCl, pH 7.4, where PPK activity was assayed for Saci_2019 and Saci_2020 in the same reaction mixture not containing dTMP. During the enzymatic assay (dTMPK assay with Saci_0893 and PPK assay with Saci_2019 and Saci_2020), ADP (dark green, yellow with dTMP), P_i (light green) and polyP (red) formation were detected discontinuously. (A) Saci_0893 enzymatic activity was dependent on dTMPK, by revealing an increased ADP-formation (yellow; without dTMP: dark green). (B) For Saci_2019 ADP- as well as polyP formation were detected, whereas (C) Saci_2020 showed almost no enzymatic activity. For all experiments three independent measurements (n=3) were performed and error bars indicate the standard error of the mean (SEM).



SI Figure 7: Characterization of the homomeric PPK (Saci_2019) from *S. acidocaldarius* with respect to polyP formation.

The temperature (A) and pH (B) dependence are shown. The enzymatic activity was determined discontinuously using 10 μg of purified homomeric PPK by detecting ADP formation using PK-LDH assay. The temperature dependence (50 – 90 °C) was analyzed in 50 mM TRIS/HCl, pH 7.4 adjusted at the respective temperature. The pH dependence of the PPK was assayed in mixed buffer (50 mM TRIS/HCl, Citrate/HCl, BICINE/NaOH and MES/NaOH, respectively) adjusted to the respective pH (4-9) at 70 °C. For all experiments three independent measurements (n=3) were performed and error bars indicate the standard error of the mean (SEM).

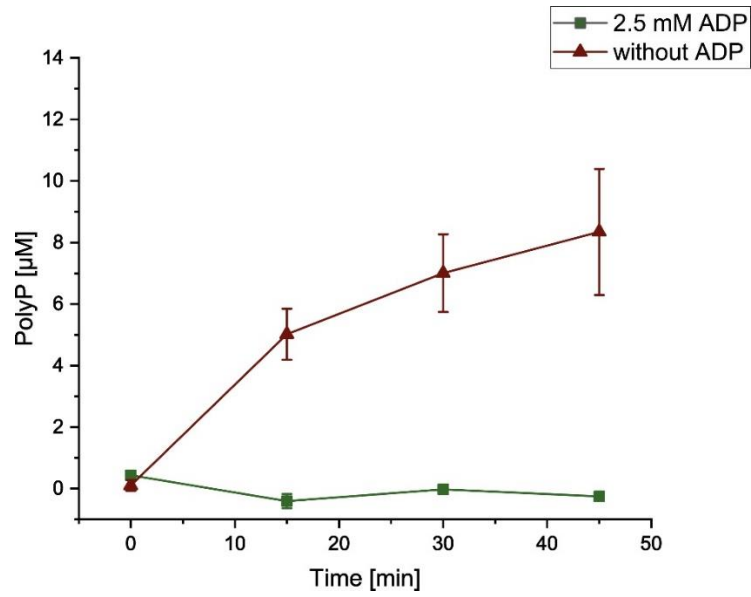
3.1. Polyphosphate metabolism



SI Figure 8: Characterization of the recombinant heteromeric PPK (Saci_2019 and Saci_2020, 1:1) from *S. acidocaldarius* with respect to polyP formation.

The temperature (A) as well as the pH (B) dependence and the kinetic properties following polyP formation (C) are shown. For (A) and (B) the same conditions were tested as for the homomeric PPK (Saci_2019, SI Fig. 7). (C) The kinetic properties were determined following polyP formation in the presence of 0 – 4 mM ATP, 4 mM MgCl₂, 50 mM KCl and 5 µg of heteromeric PPK in 50 mM TRIS/HCl, pH 7.4 at 70 °C, where polyP concentration was determined discontinuously using polyP Assay. For all experiments three independent measurements (n=3) were performed and error bars indicate the standard error of the mean (SEM).

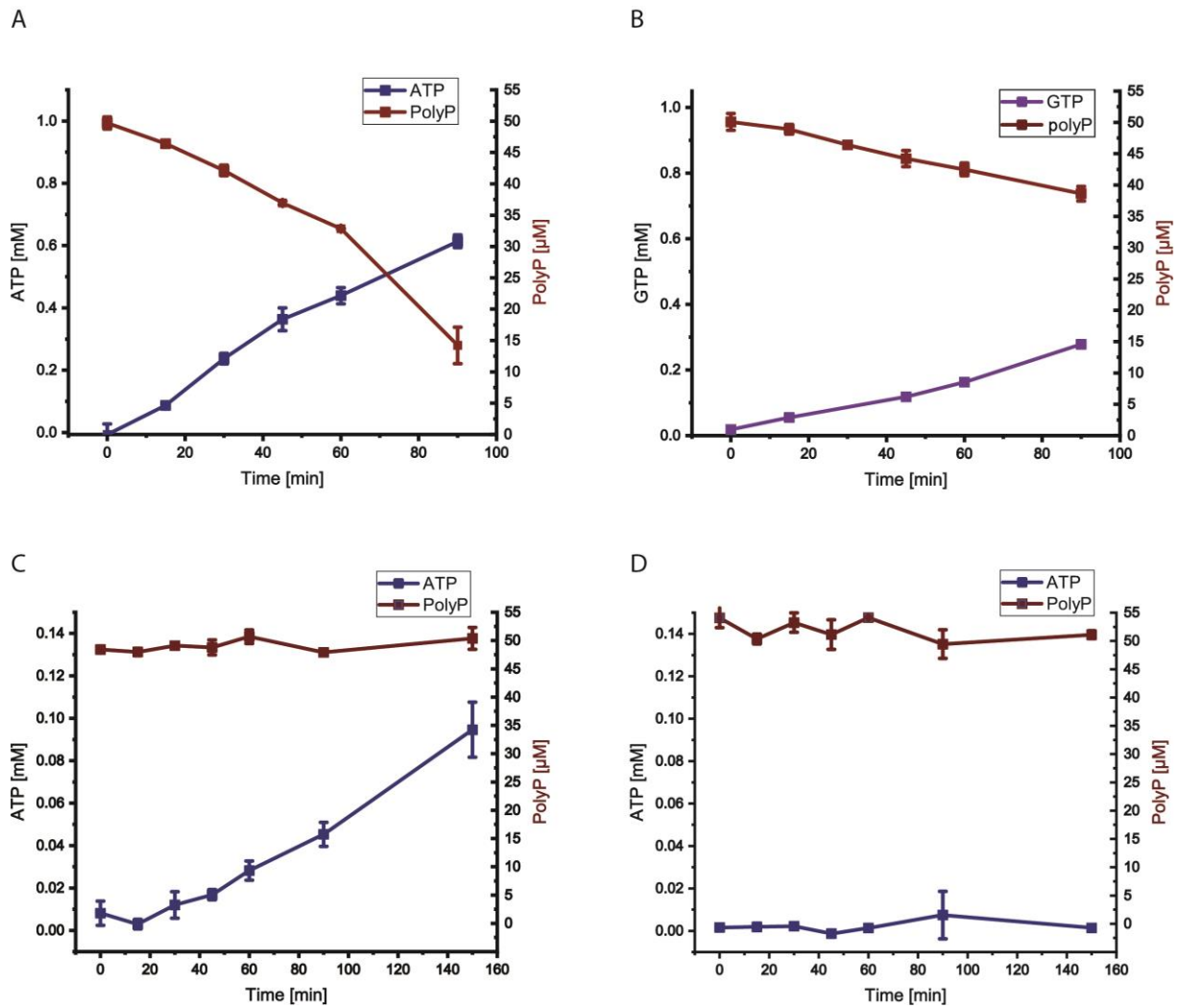
3.1. Polyphosphate metabolism



SI Figure 9: Inhibitory effect of ADP on polyP formation by the heteromeric PPK.

The polyP formation was followed during PPK assay (as described) in absence as well as presence of 2.5 mM ADP to the reaction mixture containing 1.5 mM ATP as substrate. The polyP formation was followed using polyP assay. For all experiments three independent measurements ($n=3$) were performed and error bars indicate the standard error of the mean (SEM).

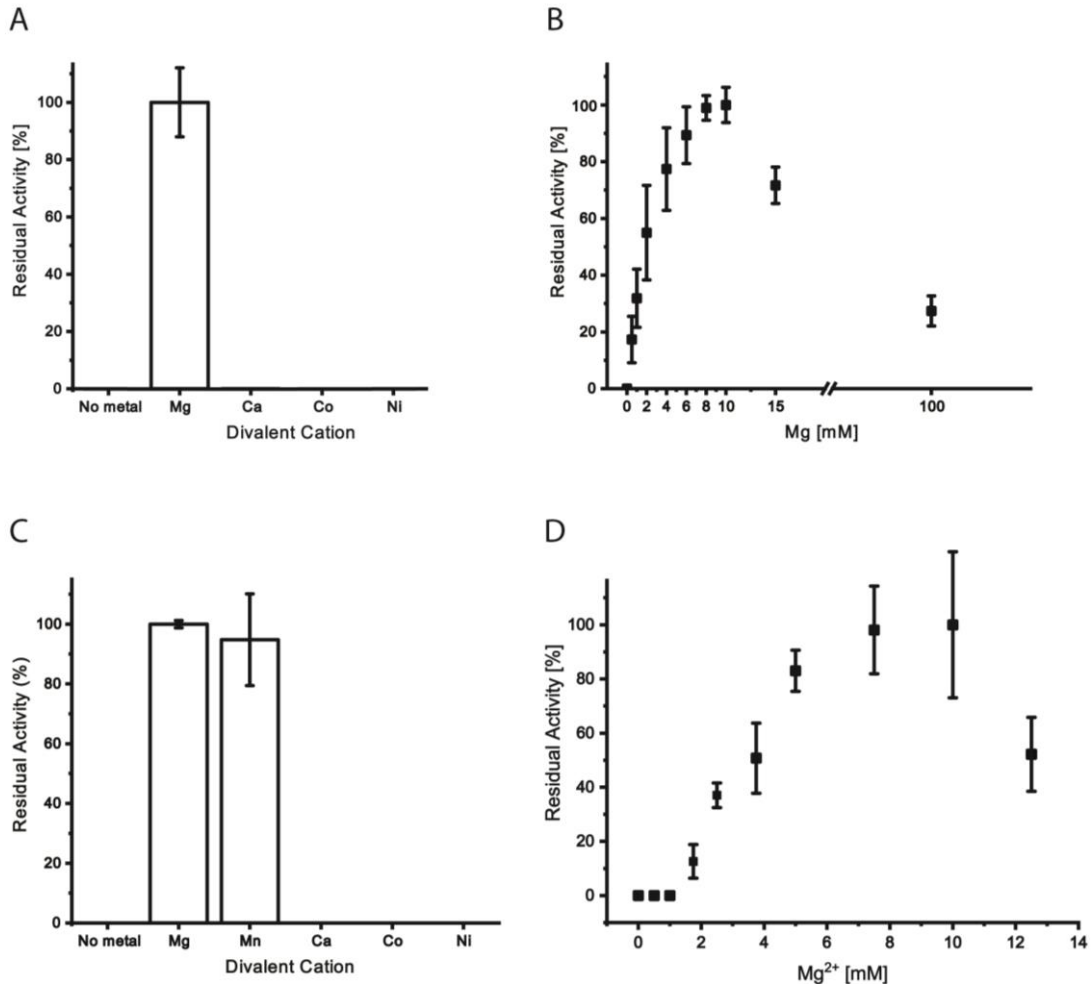
3.1. Polyphosphate metabolism



SI Figure 10: PolyP-dependent nucleotide kinase activity of the recombinant heteromeric PPK and the homomeric PPKs (Saci_2019 and Saci_2020) from *S. acidocaldarius* with respect to ATP/GTP formation.

The polyP -dependent formation of ATP from ADP (A) and GTP from GDP (B) with 5 μg of heteromeric PPK as well as the ATP formation with 5 μg of the homomeric enzymes (Saci_2019 (C), Saci_2020 (D)) are shown. The enzymatic activity (i.e. ATP/GTP formation and polyP utilization) was determined with 50 μM polyP₄₅ as phosphoryl donor and 2 mM ADP (A, C, D) or GDP (B) as phosphoryl acceptor at 70 °C in 10 mM MgCl₂, 50 mM KCl, 50 μM polyP₄₅ in 50 mM TRIS/HCl, pH 7.4. The ATP (blue squares) or GTP formation (purple squares) was determined by the discontinuous HK-G6PDH assay and polyP degradation was quantified via polyP assay. For all experiments three independent measurements (n=3) were performed and error bars indicate the standard error of the mean (SEM).

3.1. Polyphosphate metabolism

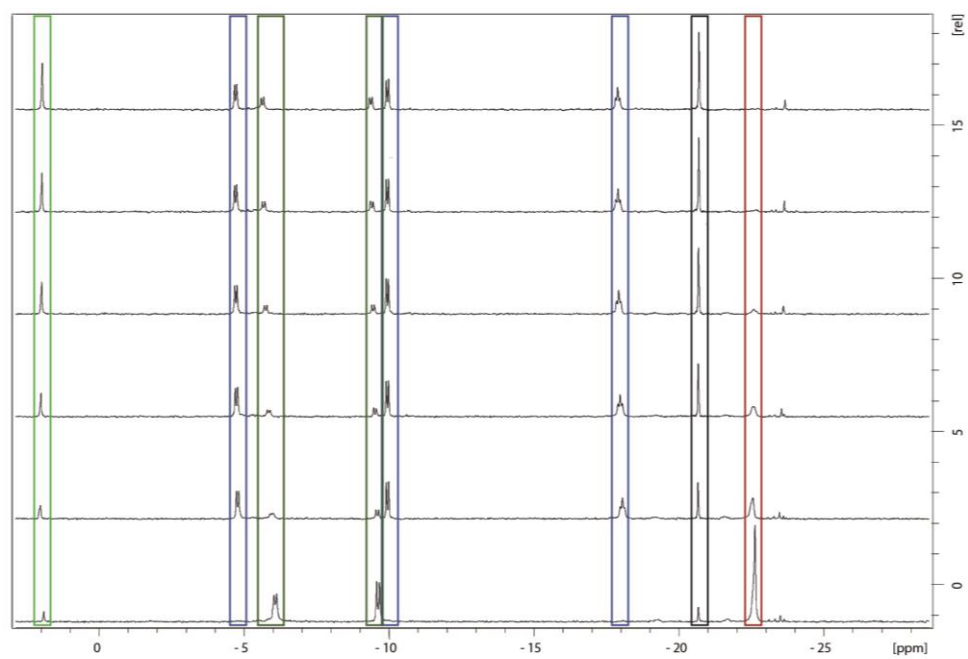


SI Figure 11: Metal ion dependence of the recombinant heteromeric PPK from *S. acidocaldarius* with respect to ATP and polyP formation.

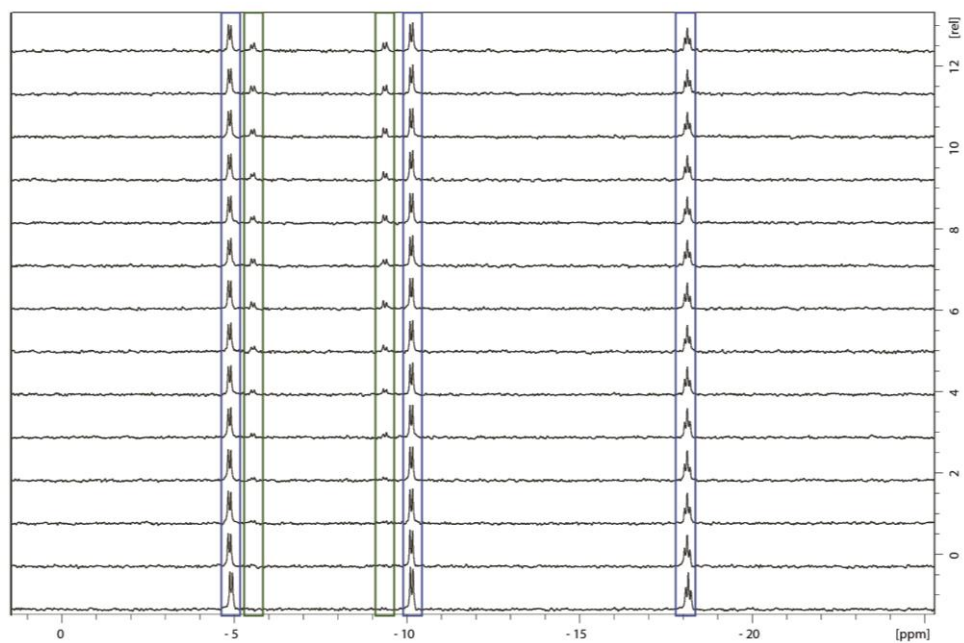
The enzymatic activity of PPK was assayed with 3 mM ATP (polyP formation) or 20 mM ADP, 500 μ M polyP₄₅ (ATP formation) with 50 mM KCl in 50 mM TRIS/HCl, pH 7.4 at 70 °C. The specific activity was determined using the discontinuous PK-LDH (ADP formation) as well as HK-G6PDH (ATP formation) assays with 10 μ g (ADP formation, A and B) and 5 μ g (ATP formation, C and D) of heteromeric PPK. Both reactions, i.e. for polyP (A) and ATP (C) formation, were performed in absence of metal ions or in presence of 10 mM of MgCl₂, MnCl₂, CaCl₂, CoCl₂ or NiCl₂. For both reactions' highest activity (301 mU/mg (polyP formation) and 4.5 U/mg (NTP formation)) was observed in presence of MgCl₂. The MgCl₂ dependence was further studied for polyP (B) and ATP (D) formation, in the presence of 0 - 12.5 mM and 100 mM MgCl₂ (polyP formation (B) only) revealing a maximal activity at 10 mM MgCl₂ (147 mU/mg (polyP formation) and 1.03 U/mg (NTP formation)) for both reactions. For all experiments three independent measurements (n=3) were performed and error bars indicate the standard error of the mean (SEM).

3.1. Polyphosphate metabolism

A

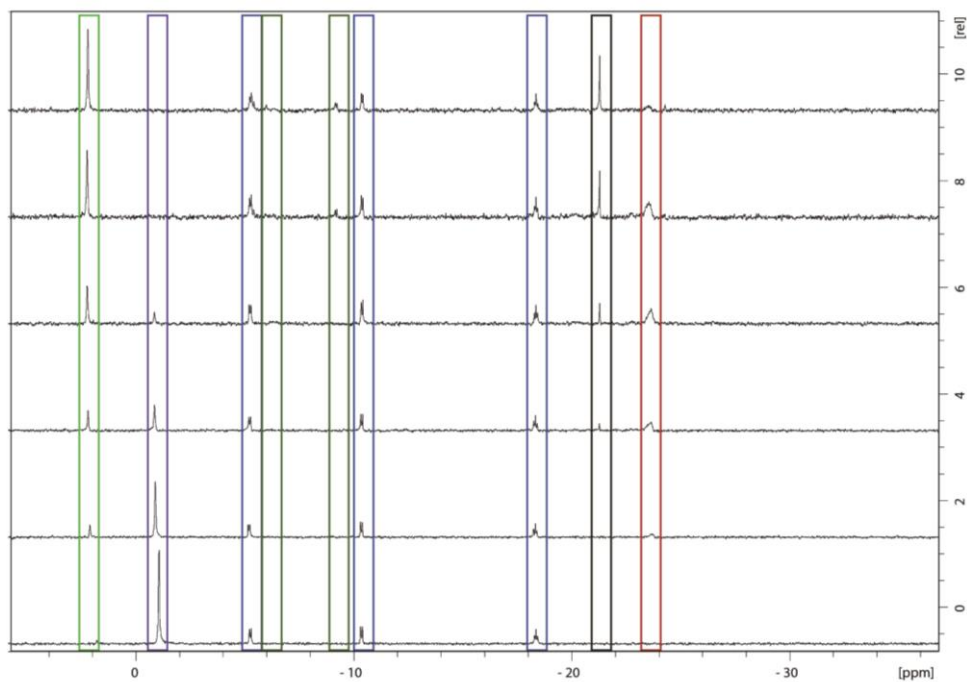


B

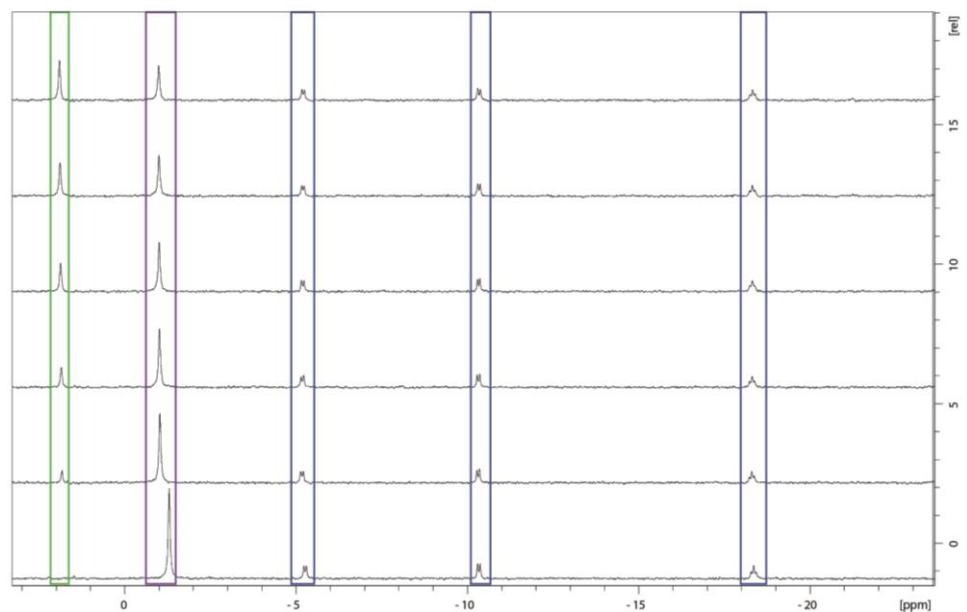


3.1. Polyphosphate metabolism

C



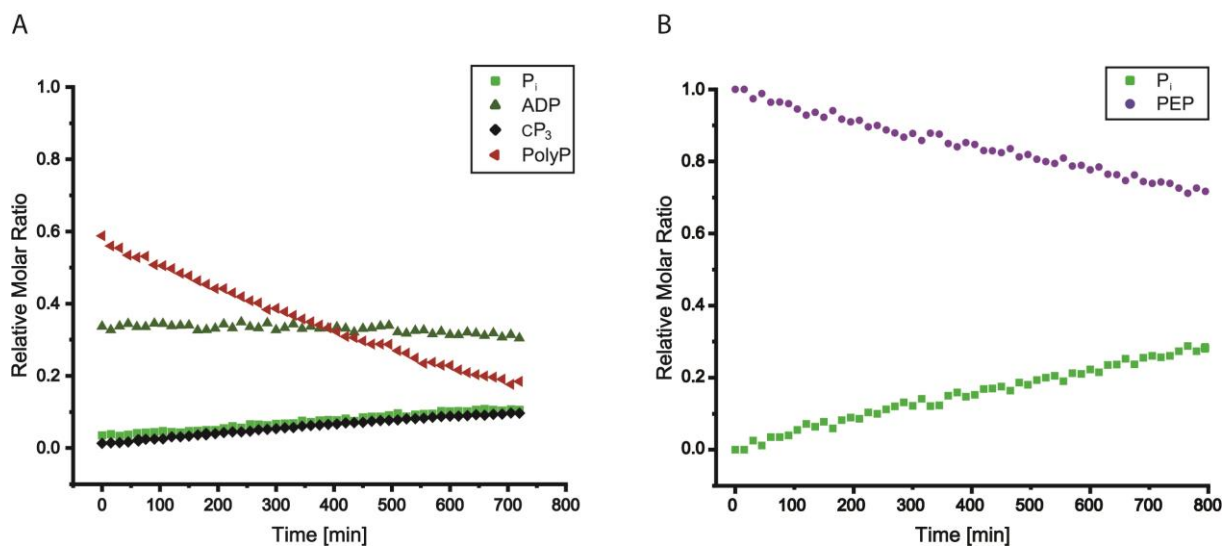
D



SI Figure 12: Stacked ³¹P NMR spectra of PPK activity in NTP (A) and polyP formation without (B) and with ATP recycling system with heteromeric (C) and homomeric PPK (D).

All assays were performed in 50 mM TRIS/HCl pH, 7.4 with 50 mM KCl in 10 % D₂O (1 mL total volume). (A) The polyP-dependent ATP formation was performed with 20 mM ADP, 500 μM polyP₄₅, 10 mM MgCl₂ with 20 μg heteromeric PPK. (B) The ATP-dependent polyP formation was performed with 2 mM ATP and 4 mM MgCl₂ with 20 μg heteromeric PPK. To the PPK reaction an ATP recycling system was added containing 20 mM MgCl₂, 10 mM PEP, 8 μg SsoPK and 2 mM ATP 20 μg heteromeric PPK (C) and 20 μg homomeric PPK (Saci_2019) (D). The signals obtained in the ³¹P NMR spectra were identified as polyP (red), cP₃ (black), ATP (blue), ADP (dark green), PEP (purple), P_i (light green). The reaction was followed from 0 min (bottom spectrum) to 330 min (top spectrum) at 70 °C.

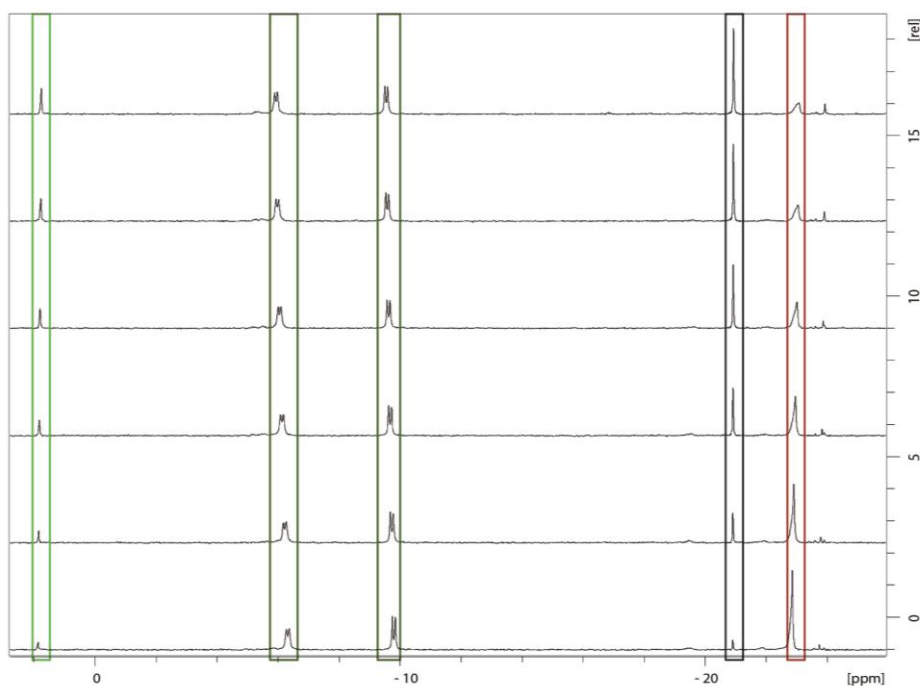
3.1. Polyphosphate metabolism



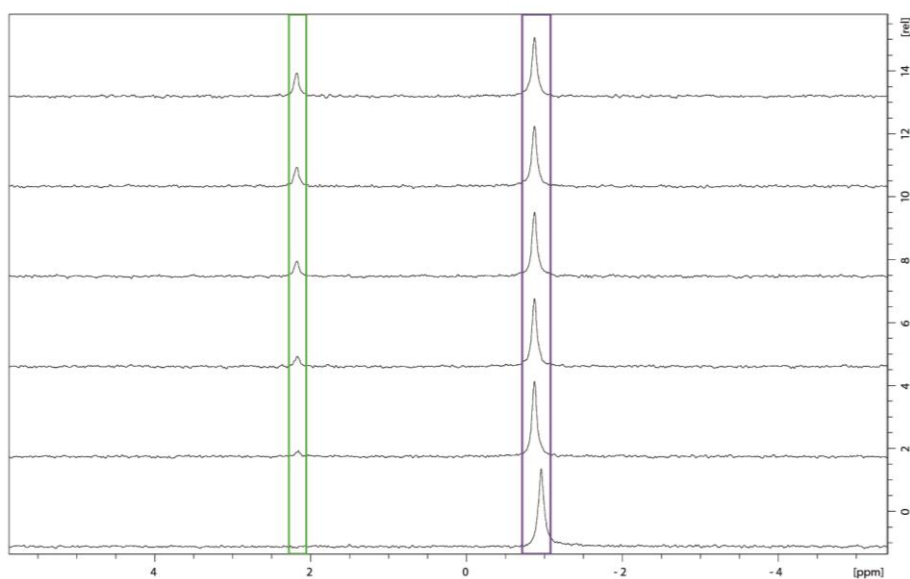
SI Figure 13: Analysis of polyP₄₅, ADP and PEP hydrolysis using ³¹P NMR spectroscopy. (A) The reaction mixture for polyP-dependent ATP formation (Fig. 5 A) was analyzed in absence of heteromeric PPK at 70 °C to follow the non-enzymatic degradation of polyP and formation of cP_3 and P_i . (B) The PEP hydrolysis was followed at 70 °C in the ATP-dependent polyP formation assay in absence of SsoPK and ATP to follow the non-enzymatic of P_i liberation from PEP. The obtained signals in the ³¹P NMR spectra were identified as ATP (blue triangles), ADP (green triangles), polyP (red triangles), P_i (green squares), PEP (purple dots) and cP_3 (black squares) (for NMR spectra see SI Fig. 14).

3.1. Polyphosphate metabolism

A



B



SI Figure 14: Stacked ^{31}P NMR spectra of polyP, ADP (A) and PEP hydrolysis (SI Fig. 13)

All assays were performed in 50 mM TRIS/HCl pH, 7.4 with 50 mM KCl in 10 % D_2O (1 mL total volume). (A) The reaction mixture contained 20 mM ADP, 500 μM polyP₄₅, 10 mM MgCl_2 to follow the non-enzymatic degradation of polyP and formation of cP_3 and P_i . (B) The PEP hydrolysis reaction mixture contained 20 mM MgCl_2 , 10 mM PEP and 20 μg heteromeric PPK and to follow the non-enzymatic of P_i liberation from PEP at 70 °C. The obtained signals in the ^{31}P NMR spectra were identified as ADP (dark green), polyP (red), P_i (light green), PEP (purple) and cP_3 (black). The reaction was followed from 0 min (bottom spectrum) to 330 min (top spectrum) at 70 °C.

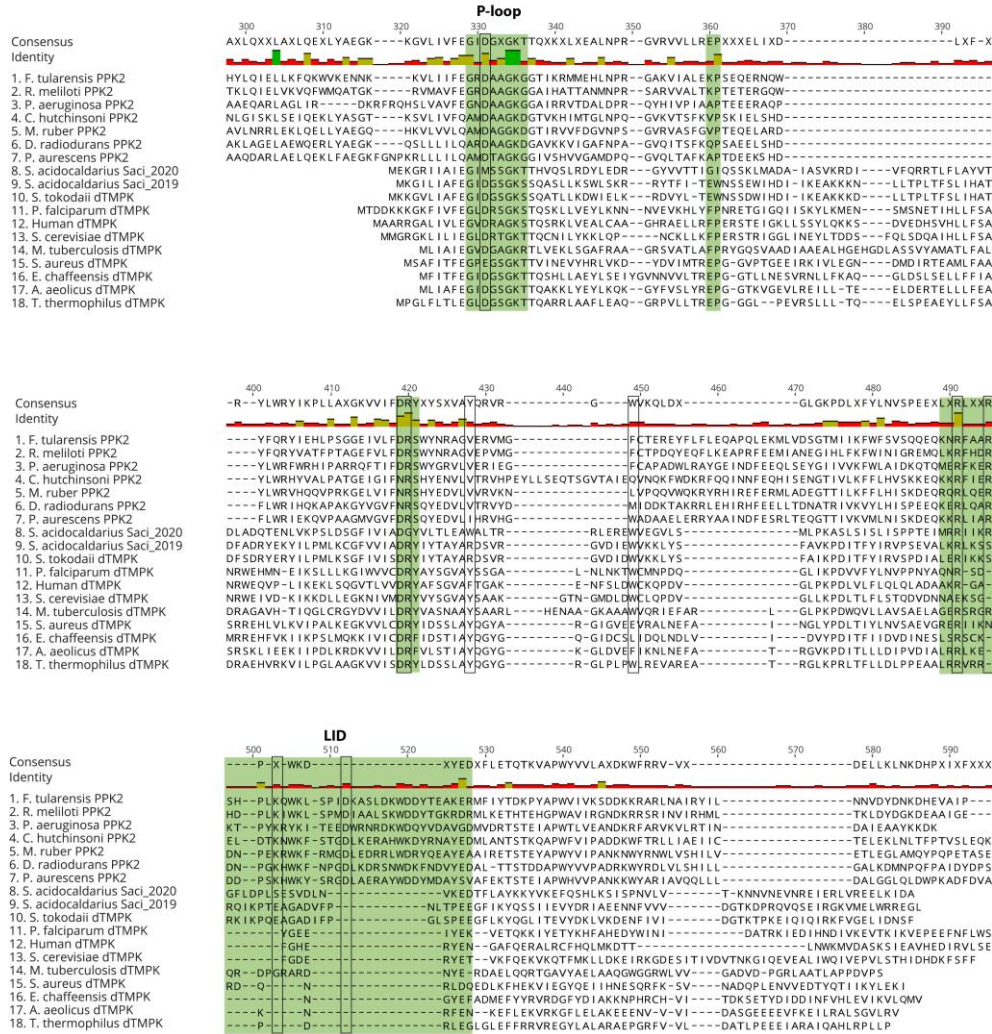
3.1. Polyphosphate metabolism



SI Figure 15: Sequence alignment of PPXs including some with solved crystal structure and *SaciPPX*.

The sequences of PPX from *Vibrio cholerae* [231], *Escherichia coli* [200, 201], *Salmonella typhimurium* [201], *Pectobacterium atrosepticum* [201], *Yersinia pestis* [201], *S. acidocaldarius*, *Porphyromonas gingivalis* [232], *Aquifex aeolicus* [203] and *Helicobacter pylori* [201] are presented. The three P-loop regions (P1, P2, P3) are present in PPX sequences. P1 and the glycine rich P3-loop are conserved in the active site, where Asp and Glu residues inside the DXGS[N/Y]S motif (motif 1) and [D/E]XGG[G/A]SXE (motif 2) and the catalytic Glu (between loop 2 and 3) interact with the Mg²⁺ to coordinate water and product molecule [202]. These residues known to be conserved are highlighted with black boxes. The P2 loop is responsible to recognize the phosphate molecule, coordinated by two metal ions in the product-bound state. The four domains of *E.coli* PPX are marked with boxes in the last line. Where domain I and II contain catalytic important residues, domain III was mainly involved in dimerization (*A. aeolicus* only comprises domain I and II and forms a monomer), and domain IV resembles a cold-shock ribosome associated RNA-binding protein [201].

3.1. Polyphosphate metabolism



SI Figure 16: Sequence alignment of dTMPKs and PPKs including some with solved crystal structure and Saci_2019 and Saci_2020.

The sequences of PPK2 from *Francisella tularensis* [220], *Rhizobium meliloti* [184], *Pseudomonas aeruginosa* [184], *Cytophaga hutchinsoni* [56], *Meiothermus ruber* [183], *Deinococcus radiodurans* [183], *Paenarthrobacter aureescens* [56], Saci_2019, Saci_2020 and dTMPKs from *Sulfurisphaera tokodaii* [228], *Plasmodium falciparum* [233], human [234, 235], *Saccharomyces cerevisiae* [223, 224], *Mycobacterium tuberculosis* [221], *Staphylococcus aureus* [225], *Ehrlichia chaffeensis* [236], *Aquifex aeolicus* [237] and *Thermus thermophilus* [222] are presented. The P-loop is reported to be involved in ATP binding in PPK2 and dTMPK enzymes, moreover the Asp residue found in dTMPKs forms a hydrogen bond with the ribose of dTMP [221, 233]. Usually, dTMPK enzymes exhibit a characteristic (E/F)P-motif forming the base of the TMP binding cavity (in *S. tokodaii* replaced by WN) [228]. The Arg residue inside the DRX motif interacts with TMP reported for *P. falciparum* dTMPK [233], whereas it is described to be involved in ATP and polyP binding in PPK2 enzymes [56, 220]. For PPK2 enzymes, the LID region contains multiple residues interacting with the polyP chain (Arg, Lys) [56], whereas in dTMPKs only one Arg residue is described to form stacking interactions with dTMP [233].

3.1. Polyphosphate metabolism

SI Table 1: Optimization of the ^{31}P NMR PPK assay to detect polyP formation using an ATP recycling.

The SsoPK catalyzes the ADP-dependent (also requires Mg^{2+}) conversion of PEP to pyruvate and thereby releases ATP. For the optimal ATP regeneration, different PEP, MgCl_2 and SsoPK concentrations were tested and the PPK specific activity was determined following ADP formation using the discontinuous PK-LDH assay. PolyP concentration was determined after 90 minutes using polyP assay.

	Without ATP recycling	With ATP Recycling				
PEP [mM]		2.5	2.5	2.5	10	10
SsoPK [$\mu\text{g/ml}$]		2	2	8	8	8
MgCl_2 [mM]		10	10	20	10	20
Specific activity [mU/mg]	0.399	0.365	0.561	0.297	0.721	0.697
PolyP concentration after 90 minutes [μM]	306.1	584.1	682.9	479.9	796.4	919.8

SI Table 2: Characteristics and kinetic parameters of selected PPX and PPK enzymes.

Organism	Enzyme class	Oligomerization	Reaction	Substrate	K_M [μM]	V_{max} [$\mu\text{mol/min/mg}$]	k_{cat} [s^{-1}]	K_{cat}/K_M [$\mu\text{M}^{-1}\text{s}^{-1}$]	Reference
<i>E. coli</i>	PPX1	dimer	polyP hydrolysis	polyP ₁₄	19.3 \pm 1.8	8.2 \pm 2.3	16.7 \pm 1.6	2037 \pm 765	[238]
<i>A. aeolicus</i>	PPX/GPPA	monomer	pppGpp hydrolysis	pppGpp	N.D.	N.D.	N.D.	N.D.	[202, 203]
<i>H. pylori</i>	PPX/GPPA	dimer	polyP hydrolysis	polyP750	49.72	1.3	1.345		[201]
<i>C. tepidum</i>	PPX1	monomer	polyP hydrolysis	polyP ₃	97.7	590	398.7	4.08	[197]
<i>S. solfataricus</i>	PPX		polyP hydrolysis	polyP ₇₀₀₋₈₀₀		0.6			[101]
<i>E. coli</i>	PPK1	dimer	polyP synthesis	ATP	2000	51	59	0.0295	[74]
			NTP synthesis	ADP	250	3.7	10.5	0.042	
				GDP	630	57 x 10 ⁻³	0.088	0.14 x 10 ⁻³	
				ppppG synthesis	460	9.0 x 10 ⁻³	0.027	0.04 x 10 ⁻³	
<i>P. aeruginosa</i>	PPK2C class II	tetramer	NTP synthesis	AMP	820	40		0.05	[239]
				polyP	62			0.61	
<i>M. ruber</i>	PPK2 class III	tetramer	NTP synthesis	AMP	33	1.	7.1	214x10 ³	[220]
				ADP	246	2.7	0.246	100 x10 ³	
<i>D. radiodurans</i>	PPK2 class III	Homo-tetramer	NTP synthesis	AMP	270		94.5	0.35	[56]
				ADP	200	0.6	0.4	2.0x10 ⁻³	
				polyP	13		132	10.2	

3.1. Polyphosphate metabolism

Cloning

SI Table 3: List of primers for cloning, plasmids used and constructed in this work, and strains employed

Primers	Sequence (5'-3')	
saci_0893 fwd <i>NdeI</i>	GCGCCATATGTTTATAATATCATTTGAAGGAATAGATGG	
saci_0893 rev <i>XhoI</i>	GCGGCCTCGAGTTAGAAATGGTTTAGAAAAGAATAAAC	
saci_2018 <i>ppx</i> fwd <i>NdeI</i>	CCCCATATGCGATATGCGGTAATAG	
saci_2018 <i>ppx</i> rev <i>XhoI</i>	GCGCCTCGAGGTCATAGCACACCAGCCACTGAC	
saci_2019 <i>dtmpk</i> rev <i>NcoI</i>	GCGCCATGGTGCCGTCACAACAACAAAG	
saci_2019 <i>dtmpk</i> fwd <i>BamHI</i>	GCCGGATCCATATAAGGGCAATAGCGAAG	
saci_2020 <i>dtmpk</i> fwd <i>NdeI</i>	CCCCCATATGGATGGAGAAGGGAAGGATTATAGC	
saci_2020 <i>dtmpk</i> fwd <i>BamHI</i>	CCCCGGATCCCCTCAGGCGTCAATTTTCAGCTC	
Plasmids		
pET15b	<i>E. coli</i> expression plasmid carrying an N-terminal His Tag	Novagen, USA
pET28b	<i>E. coli</i> expression plasmid carrying an C- or N-terminal His-Tag	Novagen, USA
pETDuet	<i>E. coli</i> co-expression plasmid carrying an N-terminal His Tag (MCS1) and C-terminal S-Tag	Novagen, USA
pET15b-Saci_2018	<i>E. coli</i> expression plasmid of <i>ppx</i> (<i>saci_2018</i>) cloned into pET15b	This work
pET15b-Saci_2019	<i>E. coli</i> expression plasmid of <i>dtmpk</i> (<i>saci_2019</i>) cloned into pET15b	This work
pET28b-Saci_2020	<i>E. coli</i> expression plasmid of <i>dtmpk</i> (<i>saci_2020</i>) cloned into pET28b	This work
Strains		Source/Reference
<i>E. coli</i> DH5 α	Hanahan, USA	
<i>E. coli</i> Rosetta (DE3)	Stratagene, USA	

3.1. Polyphosphate metabolism

NMR spectroscopy: ^1H - and ^{31}P NMR spectra were recorded on an DRX 300, DRX 500 or AVANCE NEO 400 MHz nuclear magnetic resonance spectrometer (Bruker, Billerica, USA) at ambient temperature in 90 % H_2O and 10 % D_2O . The ^{31}P NMR chemical shifts are given in ppm (for AVANCE NEO 400 MHz); splitting patterns are given as singlet (s), doublet (d), triplet (t), plus coupling constants (J) are reported in Hz.

polyP_n

^{31}P NMR(400MHz, 90 % H_2O and 10 % D_2O) δ [ppm] = -0.71 (s, 2P, 1-P; n-P), -8.43 (t, $^2J_{2,1}=14.34$ Hz, 2P, 2-P; (n-2)-P), -21.67 (s, (n-4)P, (3 to (n-2))-P).

P₃

^{31}P NMR(400MHz, 90 % H_2O and 10 % D_2O) δ [ppm] = -5.41 (d, $^2J_{1,3}=19.4$ Hz, 2P, 1-P, 3-P), -20.15 (t, $^2J_{2,1}=19.4$ Hz, 1P, 2-P).

cP₃

^{31}P NMR(400MHz, 90 % H_2O and 10 % D_2O) δ [ppm] = -20.15 (s, 1P, 1-P, 2-P, 3-P).

PP_i

^{31}P NMR(400MHz, 90 % H_2O and 10 % D_2O) δ [ppm] = -6.00 (s, 2P, 1-P, 2-P).

P_i

^{31}P NMR(400MHz, 90 % H_2O and 10 % D_2O) δ [ppm] = 2.42 (s, 1P, 1-P).

ATP

^{31}P NMR(400MHz, 90 % H_2O and 10 % D_2O) δ [ppm] = -5.47 (d, $^2J_{\gamma,\beta}=15.2$ Hz 1P, γ -P), -10.63 (d, $^2J_{\alpha,\beta}=15.2$ Hz, 1P, α -P), -19.05 (t, 1P, β -P).

ADP

^{31}P NMR(400MHz, 90 % H_2O and 10 % D_2O) δ [ppm] = -6.08. (d, $^2J_{\beta,\alpha}=18.5$ Hz, 1P, β -P), -10.14. (d, $^2J_{\alpha,\beta}=18.5$ Hz, 1P, α -P).

ADP

^{31}P NMR(400MHz, 90 % H_2O and 10 % D_2O) δ [ppm] = -6.08. (d, $^2J_{\beta,\alpha}=18.5$ Hz, 1P, β -P), -10.14. (d, $^2J_{\alpha,\beta}=18.5$ Hz, 1P, α -P).

AMP

^{31}P NMR(400MHz, 90 % H_2O and 10 % D_2O) δ [ppm] = 4.20 (s, 1P, α -P)

dTMP

^{31}P NMR(400MHz, 90 % H_2O and 10 % D_2O) δ [ppm] = 3.52 (s, 1P, α -P)

dTDP

^{31}P NMR(400MHz, 90 % H_2O and 10 % D_2O) δ [ppm] = -10.09 (d, $^2J_{\alpha,\beta}=21.2$, 1P, β -P), -11.45 (d, $^2J_{\alpha,\beta}=21.2$ Hz, 1P, α -P)

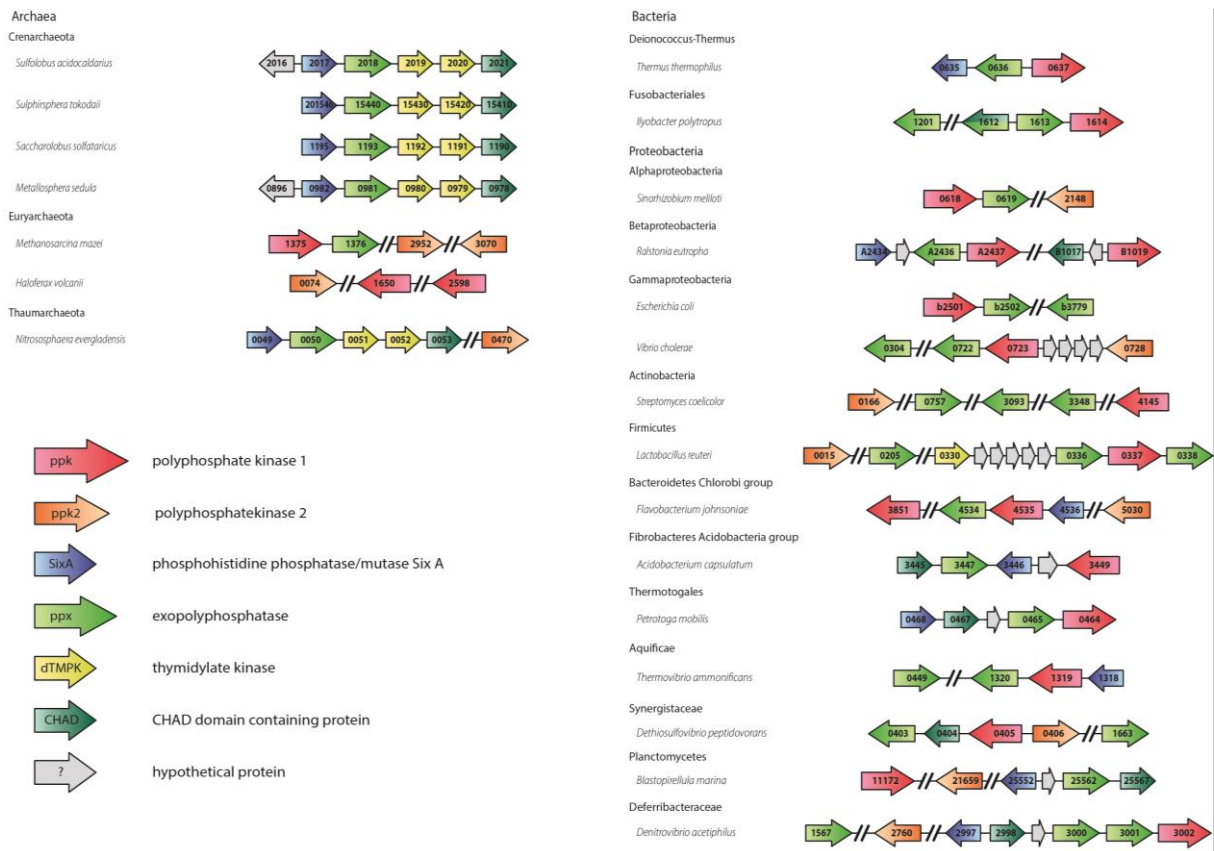
PEP

^{31}P NMR(400MHz, 90 % H_2O and 10 % D_2O) δ [ppm] = -1.08 (s, 1P, α -P).

3.1. Polyphosphate metabolism

Appendix

Gene Neighborhood



SI Figure 17: Gene neighborhood analysis of polyP metabolizing enzymes in bacteria and archaea. The operon and neighborhood relations were investigated using the KEGG and STRING databases. The scheme shows genes encoding the polyphosphate kinase family 1 (red), 2 (orange), the phosphohistidine phosphatase/mutase SixA (blue), the exopolyphosphatase (green), the thymidylate kinase (yellow), the CHAD domain containing protein (turquoise) and hypothetical proteins (grey). An arrow in the respective color represents each gene with the associated gene number.

Among bacteria and some archaeal species, the *ppk* gene is occurring in the immediate neighborhood of the *ppx* gene. Only in crenarchaeal organisms, the gene organization did not present a PPK homolog. However, the *ppx* gene cluster within Crenarchaeota seems to be conserved, containing *sixA*, *ppx*, two *dtmpk* and *chad* genes, thus suggesting one of these genes has an effect on polyp metabolism.

Chapter 3.2
Proteomic analysis of myristoylation in
Sulfolobus acidocaldarius* by *in vivo
metabolic labeling

Proteomic analysis of myristoylation in *Sulfolobus acidocaldarius* using *in vivo* metabolic labeling

Authors

Svenja Höfmann¹⁺, Sabrina Ninck²⁺, Xiaoxiao Zhou¹, Farnusch Kaschani², Christopher Bräsen¹, Markus Kaiser^{2*}, Bettina Siebers^{1*}

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

²Department of Chemical Biology, Center of Medical Biotechnology, Faculty of Biology, University of Duisburg-Essen, Essen, Germany

*Corresponding authors: Markus Kaiser (Chemical Biology) markus.kaiser@uni-due.de, Prof. Dr. Bettina Siebers (Biochemistry): bettina.siebers@uni-due.de

+These authors are contributed equally to this work.

Keywords

Archaea – *Sulfolobus acidocaldarius* - Bioorthogonal labeling - Metabolic labeling – Acylation - Myristoylation – Fatty acids

3.2. Myristoylome profiling

Abbreviations

(FA) fatty acids, (\equiv MyA) myristic acid alkyne, (NMT) N-myristoyl transferase, GCN5-related N-acetyltransferase (GNAT), (alkyl-DHAP synthase) alkyl dihydroxyacetone phosphate synthase, (ACS) acyl-CoA synthetase, (FAD) flavin adenine dinucleotide, (DTT) dithiothreitol, (ACN) acetonitrile

Abstract

Metabolic labeling has become a popular method for myristoylome profiling in various organisms. Myristoylation, the attachment of a myristoyl moiety onto a protein, is known to influence protein interactions, membrane associations, and enzyme catalytic function. To date, myristoylation has been studied mainly in eukaryotes such as mammals, fungi, and plants. So far, only a few studies address this lipid modification in archaea, which is probably related to the absence of fatty acids (FAs) in the lipid composition in this domain. Nevertheless, the presence and utilization of FAs has been described for some archaea, but the function is still not well understood.

Here we present a first myristoylome profile of the thermoacidophilic Crenarchaeon *Sulfolobus acidocaldarius*. It was already demonstrated that this organism comprises a β -oxidation pathway for fatty acid degradation as well as fatty acid synthesis could be reconstructed *in vitro*. Here, we describe *in vivo* metabolic labeling with a C14 fatty acid derivative (\equiv MyA) to identify target proteins for myristoylation in *S. acidocaldarius*. First, an in-gel fluorescence analysis demonstrated the functional method of the metabolic labeling. In addition, we identified the proteins that were abundant in each experiment by comparing the results from labeled versus unlabeled samples. Based on the proteomic data obtained from four biological replicates, we mainly observed proteins that are predicted to be related to metabolism and genetic information processing. Considering the metabolism, the most identified proteins were classified be involved in metabolism take mainly part in lipid transport and metabolism as well as energy production and conversion following the arCOG code. The myristoylated proteins, abundant in each experiment, are annotated as chromatin protein Alba (Saci_1322), β -ketothiolase (Saci_1114) and alkylidihydroxyacetonephosphate synthase (alkyl-DHAP synthase, Saci_1188). Notably, the acetylation of the Alba protein in *Saccharolobus solfataricus* (homolog of Saci_1322) has already been described and the β -ketothiolase was described to take part in the β -oxidation pathway of *S. acidocaldarius*. In addition, compared to recently published archaeal acetylome data, we could identify a certain overlap of identified candidates, which suggests that the targets of protein acetylation may also comprise acylation modifications. To the best of our knowledge, this is the first myristoylome study for archaea providing an initial insight into the presence and function of myristoylation. It offers a good starting point for further studies by presenting reproducible protein targets that can be used to further investigate the potential function of fatty acids in archaea.

Introduction

Archaea differ significantly from eukaryotes and bacteria with respect to their membrane lipids, by comprising isoprenoid chains that are ether-linked to *sn*-glycerol-1-phosphate [21, 240]. Since archaeal membrane lipids thus do not contain fatty acids (FAs), in contrast to eukaryotes and bacteria, the function of FAs in archaea is not well understood. Some studies provide evidence that some archaea can use long-chain FAs and moreover, a previous study demonstrated that *S. acidocaldarius* can grow on butyrate as well as on hexanoate as sole carbon source [27, 162, 241, 242]. In accordance, some archaea comprise all genes for a β -oxidation pathway to degrade FAs, thus suggesting that at least some archaea are able to use FA as carbon and energy source [26]. Notably, the β -oxidation pathway from *S. acidocaldarius* could be reconstructed *in vitro*, resulting in a complete oxidation of C8 FAs to acetyl-CoA [28]. Moreover, the presence of FAs in archaea has been demonstrated [26, 160] suggesting that they are able to synthesize FAs. However, since FAs are neither membrane lipids nor storage lipids, which so far have not been described in archaea, the reason for their existence is obscure. One possible function would be the acylation of proteins, one of the most important post-translational modifications (PTMs) mainly described in Eukaryotes, which is defined as attachment of fatty acyl groups onto proteins [111]. So far, many types of acylation have been described, the most prominent being acetylation, formylation, propionylation, butyrylation, malonylation, succinylation, glutarylation, palmitoylation, myristoylation, benzoylation and crotonylation [111]. Moreover, acylation is reported to comprise several regulatory properties on physiological activities including, transcription, metabolism, subcellular targeting, membrane association as well as protein stability and folding [111]. In bacteria, acetylation was mainly studied and was reported to play a role in central and secondary metabolism, virulence, transcription, translation and virulence [117]. In general, protein acetylation can occur either on the N-terminus ($N\alpha$ -acetylation) or on a lysine residue ($N\epsilon$ -acetylation). The reversible $N\epsilon$ -acetylation was described to be enzymatically catalyzed by acetyltransferases using acetyl-CoA as donor and is removed by deacetylases [117]. In contrast to several studies that describe protein acetylation, there are a few reporting acylation in bacteria. However, it has been shown, for example, that an NAD^+ -dependent deacetylase (Sir2) catalyzes lysine-demalonylation of different peptides [243]. In archaea, acetylation has also been described for some organisms, including members of the *Sulfolobaceae* and Haloarchaea, including some proteomic analyses as well as *in vitro* acetylation studies [137, 140, 141]. Focusing on myristoylation, defining the attachment of a myristoyl-moiety covalently linked to glycine, cysteine or lysine of the protein, some studies are described for bacteria in addition to those in eukaryotes [31], however for archaeal organisms myristoylation has only been reported for *Methanobacterium thermoautotrophicum* and *Halobacterium cutirubrum* [143]. Myristoylation of proteins was reported to influence the

3.2. Myristoylome profiling

membrane association and thereby functions as molecular anchor [244]. Furthermore, it is known to be involved in protein-protein interactions, signaling and protein stability [245-248].

In Eukaryotes the N-myristoyl transferase (NMT) is the main enzyme known to catalyze N-terminal myristoylation and is a member of the Gcn5-related N-acetyltransferase (GNAT) superfamily [145]. This enzyme is well characterized in Eukaryotes like fungi, mammals and plants [249-251]. The demyristoylation is known to be mainly catalyzed by NAD⁺-dependent deacetylases, also known as sirtuins [31]. Specifically for archaea, a Sir2 deacylase from *Archaeoglobus fulgidus* has been described to catalyze demyristoylation [252]. However, despite the rarely reported presence of myristoylation in archaea [143], Sir2 homologs were demonstrated to be present in *Sulfolobaceae* and catalyze the deacetylation of the DNA/RNA binding protein Alba [137]. Moreover, a protein acetyltransferase (Pat) from *S. solfataricus* was demonstrated to acetylate Alba at K16. In addition, the structure of SsoPat was solved and comprises homology to prokaryotic GNATs as well as eukaryotic histone acetyltransferases (HATs) with an additional “bent” helix [253]. It was suggested that due to its structural similarity to eukaryotic HATs as well as its role in chromatin regulation SsoPat is evolutionary related to the eukaryotic enzymes. The SsoPat displays a smaller protein acetyltransferase scaffold with lower efficiency compared to the eukaryotic HATs [253]. Considering the chromatin protein Alba, a recently published study showed that Alba serves as a primary driver of chromatin compaction by mediating the concentration-dependent formation of folded DNA structures [254]. The enzyme SsArd1 was demonstrated to catalyze N-terminal acetylation and is a member of the NatA family, conserved in eukaryotes [120]. Whereas the catalytic subunit of NatAs is conserved from archaea to higher eukaryotes, the structural comparison of SsArd1 showed an altered substrate specificity dependent from a Glu35 residue suggested to be involved in the catalytic mechanism [120]. The SsArd1 was demonstrated to acetylate a broader range of N-terminal sequences including Ser, Ala and Met-Glu termini compared to the resembling eukaryotic NATs [254]. In addition, N α -acetylation has also been reported to occur on subunits of the thermosome, proteome as well as the ribosome of *S. solfataricus* [254]. The mechanism of NAT enzymes comprises an ordered water molecule for mediating deprotonation. For the thermostable SsArd1 structural analysis led to the suggestion that the binding of the water molecule takes place with a higher number of hydrogen bonds to trap the water molecule even at high temperatures and allow stable binding in the active site [255].

To date, acetylome studies for *Sulfolobus islandicus* demonstrated that 26 % and 44 % of proteins were acetylated on N ϵ -Lys and the N-terminus, respectively [141]. In addition, the acetyltransferase SisPat has been described to mainly acetylate acyl-CoA synthetases (ACS), whereas SisArd1 primarily acetylates the α -amino group of proteins. The study of deletion

3.2. Myristoylome profiling

strains demonstrated only for the $\Delta ard1$ strain a growth deficiency phenotype on Zillig's medium [141]. Previously, an lysine-specific acetylome of *T. gammatolerans*, was described, comprising 338 acetylated sites as well as acetylation of histones has been reported to affect their assembly and affinity to DNA [140]. In addition, proteins involved in metabolic processes, information processing and storage mechanisms were predominantly identified to be acetylated and thus suggested a functional role of protein acetylation in these organisms [140].

The first studies dealing with N-myristoylation used the effect of the attached myristoyl moiety covering the positive charge of the N-terminus of proteins and thus changing the electrophilic mobility upon gel electrophoresis [256]. Moreover, another method was described, where both the N-myristoyl transferase (NMT) and its target protein were coexpression in *E. coli*. However, this method required an additional confirmation, which was demonstrated by a variant harboring G2A substitution, which prevented N-terminal myristoylation of the target protein [257]. For *in silico* analysis, a prediction program has been developed of co- or posttranslational modifications of proteins [258, 259]. In 2006, metabolic labeling was firstly described as high-throughput method to study co- and post-translational N-myristoylation of proteins [260]. Briefly, a cell-free protein synthesis system derived from insect cells was used for *in vitro* metabolic labeling using radioactive [^3H]myristic acid, followed by SDS-PAGE and fluorography detection of myristoylated proteins. This method was already described for some eukaryotic organisms, including human and plants (for review see [31]). Later on, large myristoylome data sets were generated by combining metabolic labeling with natural product-inspired probes and mass spectrometry (MS), which became a common method to study myristoylation in living cells. Specifically for the study of acylation, FA azide or alkyne probes are used to allow *in vivo* metabolic labeling with subsequent copper-catalyzed click chemistry (CuAAC) enabling the identification of labeled proteins via LC-MS/MS analysis (described in 1.9) [165, 261, 262]. Myristoylome data resulting from metabolic labeling approaches using myristic acid azide or alkyne have been published for several organisms including mammals and plants [31, 263, 264].

Here we present the first myristoylome profiling of a crenarchaeal organism, *S. acidocaldarius* MW001. We used *in vivo* metabolic labeling with a C14 fatty acid alkyne ($\equiv\text{MyA}$) to identify target proteins for myristoylation. Briefly, *S. acidocaldarius* MW001 cells were grown Brock's basal medium at 76 °C and pH 3.0 supplemented with 0.3 % (w/v) dextrin, 0.1 % (w/v) N-Z-Amine and 20 $\mu\text{g}/\text{mL}$ of uracil until an OD_{600} of 1.0. The $\equiv\text{MyA}$ probe was added to the cultures with a final concentration of 20 μM and cells were incubated for 6 h at 76 °C to enable metabolic labeling. Subsequently, cells were harvested and protein extracts were used for copper-catalyzed click chemistry, attaching a Cy3 fluorophore enabling in-gel fluorescence detection. For proteomic analysis, click chemistry was performed using a TAMRA biotin tag to allow affinity enrichment with ongoing tryptic on-bead digest and following

3.2. Myristoylome profiling

LC-MS/MS analysis. Statistical evaluation resulted in several abundant proteins obtained from each experiment when comparing probe and non-probe samples (\log_2 fold change > 1). Between four independent experiments containing four biological replicates each, an overlay of abundant proteins was obtained indicating a functional approach for investigating the myristoylome profile of *S. acidocaldarius* MW001.

Results

In vivo metabolic labeling of *S. acidocaldarius* MW001 using myristic acid alkyne probe

To identify target proteins for myristoylation in *S. acidocaldarius* MW001 we used the method of *in vivo* metabolic labeling with myristic acid alkyne (\equiv MyA). The probe was added to the growing culture during exponential phase, while ethanol was used for the control samples since the probe was solved in ethanol (76 °C). During *in vivo* metabolic labeling, the cultures usually showed almost no growth. However, this effect was not observed in all experiments, as sometimes minimal growth or even a decrease in optical density was detected (data not shown). After 6 hours of incubation, the protein extracts were prepared and copper-catalyzed click chemistry was used to attach the Cy3 fluorophore for in-gel visualization. Additionally, the TAMRA biotin tag was clicked to a larger scale protein extracts of labelled cells followed by affinity purification via avidin agarose beads. After a tryptic on-bead digest we were able to identify the labelled proteins using LC-MS/MS-analysis (Fig. 1). The experiment was carried out in four biological replicates comprising four technical replicates, each.

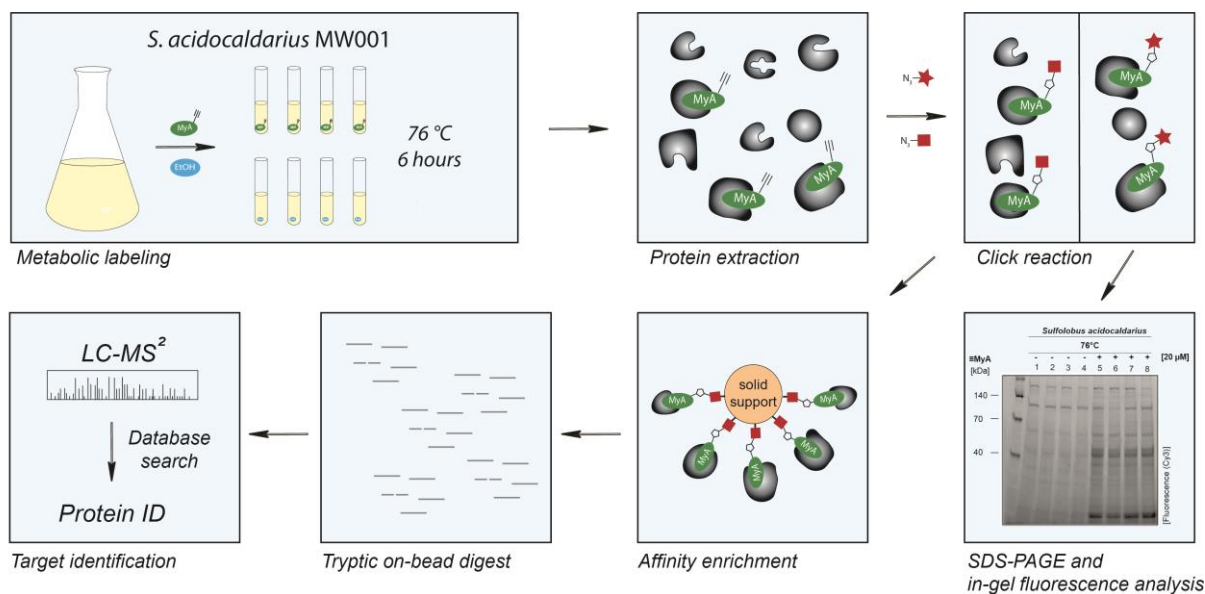


Figure 1: Workflow of metabolic labeling using myristic acid alkyne probe (\equiv MyA) in *S. acidocaldarius* MW001.

As natural product-inspired probe we used myristic acid alkyne (\equiv MyA, green circle) and performed *in vivo* labeling by adding 20 μ M \equiv MyA to *S. acidocaldarius* MW001 cells during exponential phase. Ethanol was used for the control samples with the same volume (blue circle). After cell lysis and protein extraction, the covalently labelled proteins were clicked by a copper-catalyzed reaction with N_3 -rhodamine (Cy3, red star) for fluorescence visualization or N_3 -biotin (red square) for affinity enrichment. The enriched proteins were digested with trypsin and used for MS-based proteomics. The experiment was carried out in four biological replicates with four technical replicates for probe and control samples, respectively.

The visualization of the \equiv MyA labeling experiments was enabled via a Cy3 fluorophore clicked onto the probe exemplarily shown for one biological replicate (Fig. 2). A slight background was observed for the samples not containing the fatty acid alkyne probe, whereas a significant increase in fluorescent protein bands were obtained for the samples containing the probe.

3.2. Myristoylome profiling

Some distinct fluorescence signals were obtained for proteins with a molecular mass of 60, 40 and 15 kDa, whereas the labelled bands at about 130 and above 140 kDa were also present in the control samples indicating a background labeling. Since a significant difference between labeling results of non-probe and probe samples were observed, the in-gel fluorescence analysis already shows the functional approach that the metabolic labeling worked out well. For the identification of labelled proteins, Strep-tag based affinity enrichment followed by tryptic on-bead digestion and LC-MS²-analysis was performed.

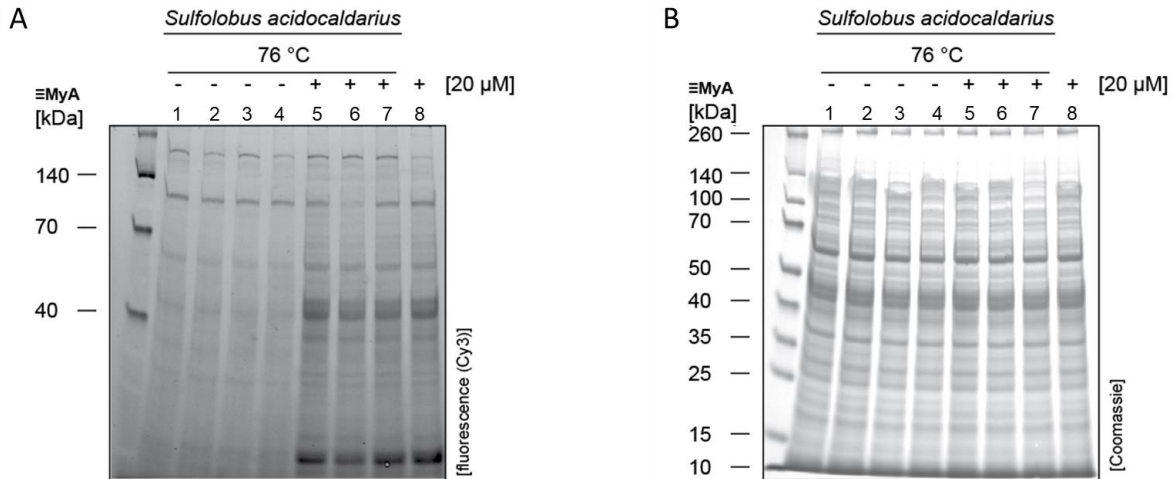


Figure 2: In-gel fluorescence detection of metabolic labeling samples from *S. acidocaldarius* MW001 using \equiv MyA.

S. acidocaldarius MW001 cells were labeled with 20 μ M \equiv MyA (samples 5-8) where control cultures were treated with ethanol (samples 1-4) at 76 °C. The protein extracts were clicked onto a Cy3 fluorophore. The fluorescence detection (A) and the Coomassie staining of an 11 % Bis-Tris-gel is shown. [ACE_0593]

The proteomics approach was used to identify differentially abundant proteins in the probe samples compared to the control samples, based on LFQ. Statistical evaluation of identified proteins revealed for the four experiments the abundance of 37 - 14 proteins, respectively, with a \log_2 -fold change > 1 (Fig. 3). The volcano plots from four biological replicates are depicted (in Fig. 3A – D, SI Tab. 1 - 4) and already show that the biological variance between the replicates is rather high.

3.2. Myristoylome profiling

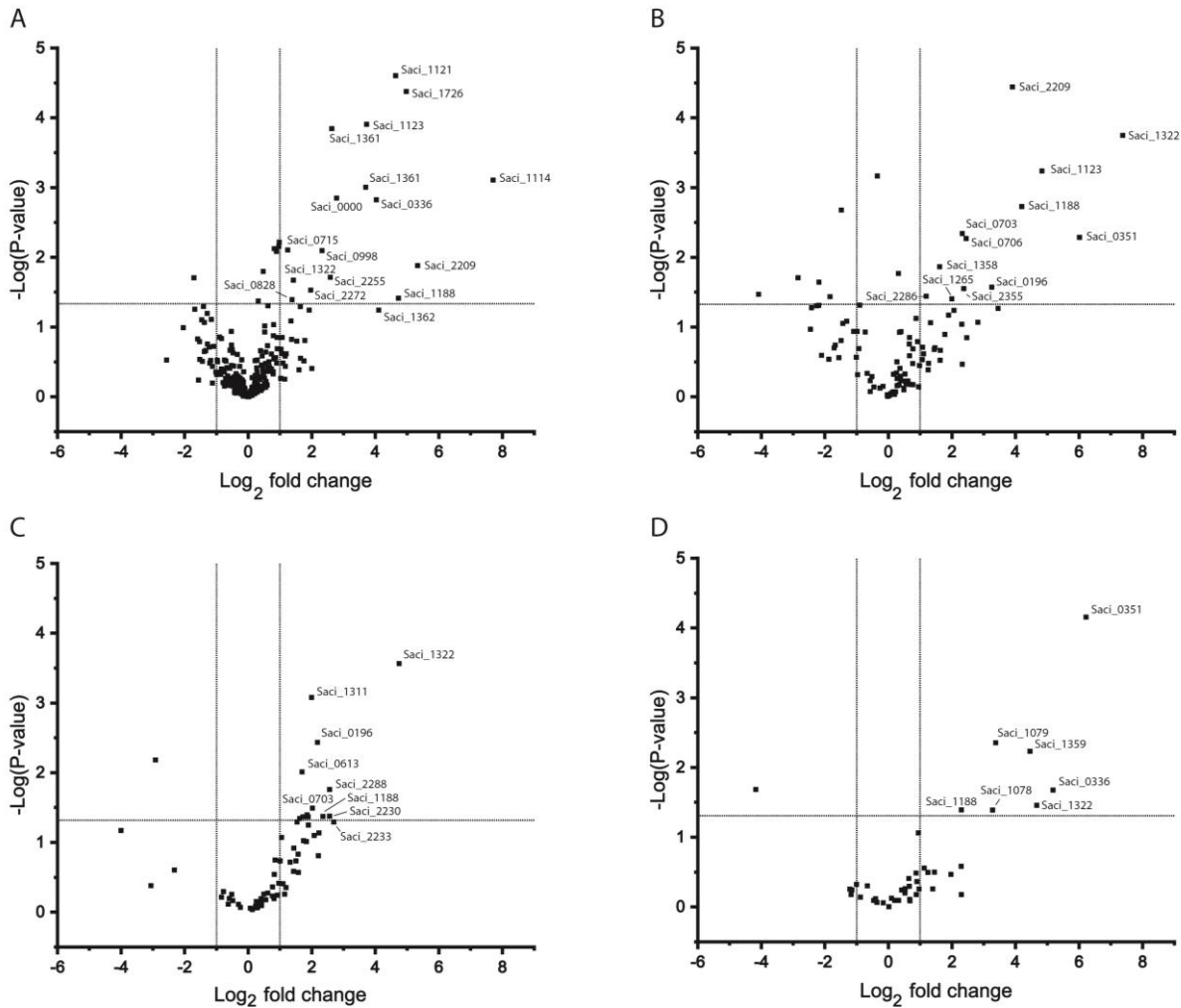


Figure 3: Differentially abundant proteins resulting from metabolic labeling of *S. acidocaldarius* MW001 cells at 76 °C using \equiv MyA.

The protein abundance profile of cells containing the \equiv MyA probe compared to non-probe cells. *S. acidocaldarius* MW001 cells were labeled with 20 μ M myristic acid alkyne, where control cultures were treated with ethanol at 76 °C. The proteins showing a \log_2 -fold change > 1 were abundant in the \equiv MyA containing samples compared to the control ones. The volcano plots (A – D biological replicates) show the p -value against the \log_2 -fold change in protein abundance. Statistical abundance was calculated based on a two-sample Student's test. The protein-IDs of abundant proteins were assigned.

This is further indicated by the Venn diagram (Fig. 4) which shows the number of abundant proteins detected in each experiment (ACE_0408, ACE_0589, ACE_0593, ACE_0604) and compare them to those found in each of other experiments. The numbers in the overlapping regions represent the candidate proteins that were detected and shared by the respective experiments. We obtained 3 candidates in each experiment and 2 and 1 proteins to be abundant in almost three of four experiments. In two experiments we identified, 2, 2, 2 and 3 proteins, respectively, representing the overlap of four biological replicates (SI Tab. 5). Thus, the majority of proteins were found in only one of the experiments (24, 15, 18 and 16), whereas the overlap between the experiments is comparably low.

3.2. Myristoylome profiling

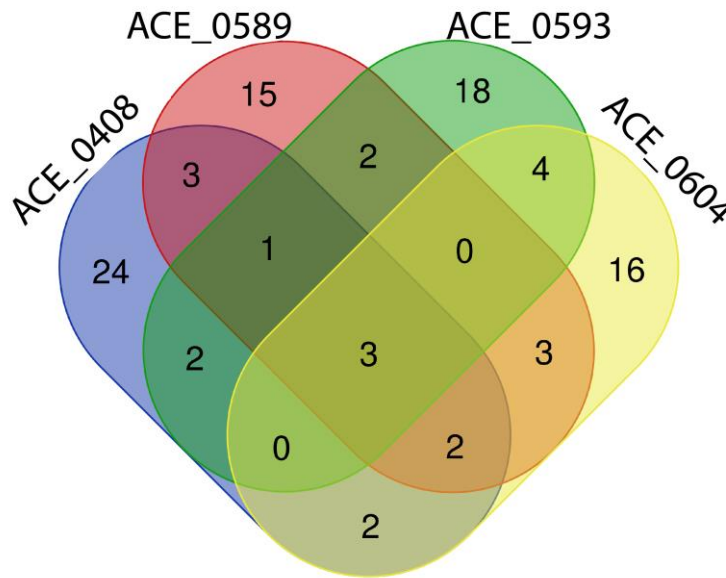


Figure 4: \equiv MyA labeled proteins in four biological replicates of metabolic labeling experiments in *S. acidocaldarius* MW001 represented by different colors.

Each color represents the abundant proteins from one *in vivo* metabolic labeling experiment of *S. acidocaldarius* MW001 using \equiv MyA probe. The overlapping regions represent the number of candidates obtained as abundant in the respective myristoylome data with a \log_2 -fold change > 1 , except ACE-0604 with a positive \log_2 -fold change. Each experiment contained four technical replicates.

The identified myristoylated proteins were classified according to the arCOG functional code [170] (Fig. 5, SI Tab. 5). Most proteins were identified to be involved in metabolism (arCOG category: C, E-I, P, Q), whereas a smaller proportion of abundant proteins play a role in processing genetic information (arCOG classification: J-L) and only a few proteins have an unknown function (arCOG category: S). The majority of proteins with predicted metabolic function relate to energy production and conversion (C), lipid transport and metabolism (I) and amino acid metabolism (E). The remaining hits were predicted to be involved in transport and metabolism of nucleotides, coenzymes, or inorganic ions, as well as biosynthesis of secondary metabolites, transport, and catabolism.

3.2. Myristoylome profiling

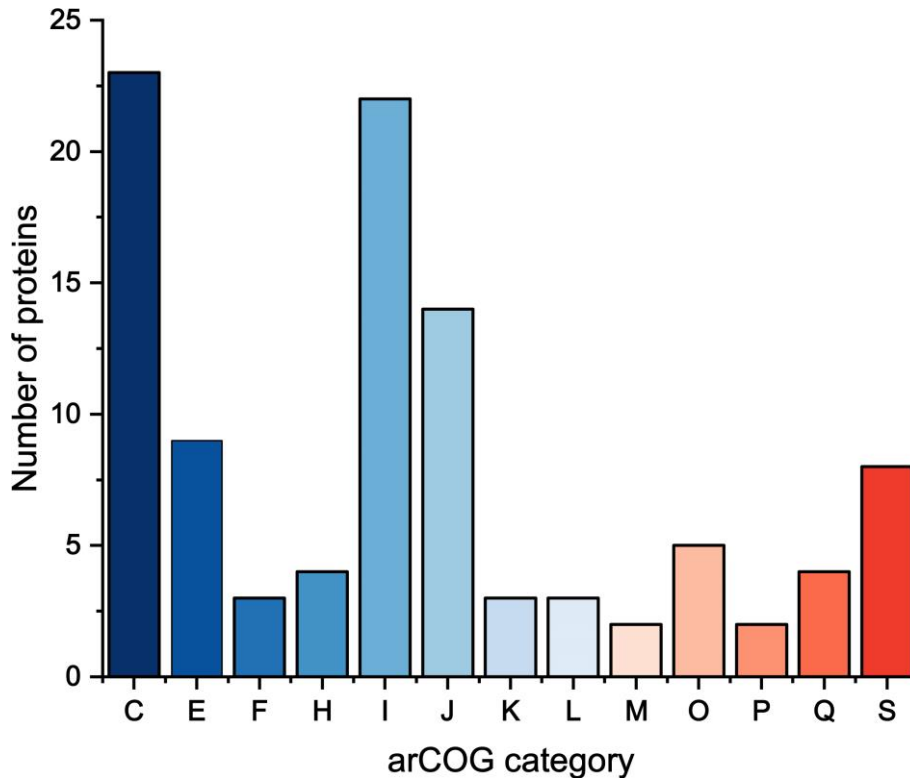


Figure 2: Identified labeled proteins categorized by arCOG functional code.

The abundant proteins from the myristoylome data were classified using arCOG database [170] and the number of proteins categorized in the respective arCOG class were depicted. The proteins categorized into two arCOG classes were counted for both categories.

C: Energy production and conversion, E: Amino acid transport and metabolism, F: Nucleotide transport and metabolism, H: Coenzyme transport and metabolism, I: Lipid transport and metabolism, J: Translation, ribosomal structure and biogenesis, K: Transcription, L: Replication, recombination and repair, M: Cell wall/membrane/envelope biogenesis, O: Post-translational modification, protein turnover, and chaperones, P: Inorganic ion transport and metabolism, Q: Secondary metabolites biosynthesis, transport, and catabolism, S: Function unknown.

From four biological replicates, three proteins were obtained, namely Saci_1322, Saci_1188 and Saci_1114 and thus represent the core of the myristoylome profile of *S. acidocaldarius* MW001. The candidates are annotated as DNA/RNA binding protein Alba (Saci_1322), alkyldihydroxyacetonephosphate synthase (alkyl-DHAP synthase; Saci_1188) and thiolase (Saci_1114). Additionally, three more candidates were present in three of four myristoylome data, namely Saci_0196, Saci_0336 and Saci_2209. Saci_0196 is annotated as putative mannose-1-phosphate guanylyltransferase. Saci_0336 is a conserved protein with unknown function and additionally, Saci_2209 is annotated as an acetyl-CoA thiolase. In addition, some abundant proteins identified in two experiments were classified into the arCOG category of lipid transport and metabolism, including an annotated acetyl-CoA synthetase (ACS, Saci_0306) and a acyl-CoA dehydrogenase (ACAD, Saci_1123).

Discussion

The function of FAs in archaea is still under discussion and not well understood. There are only a few studies that provide some evidence why archaea contain FAs what their function is. One suggestion is that archaea may use FAs for acylation modifications of proteins. To study the presence of myristoylation we determined the myristoylome data of *S. acidocaldarius* MW001, to our knowledge the first one of an archaeal organism.

So far, myristoylome profiling has mainly been reported for eukaryotes. In human, 83 myristoylated proteins were identified using the metabolic labeling approach to enable quantification of myristoylation [265]. Moreover, this method was also used to analyze the protein myristoylation in the zebrafish *Danio rerio* [265] as well as the parasite *Trypanosoma brucei* [266]. In addition, a large myristoylome was identified for *Arabidopsis thaliana*, where 72 proteins were N-terminally myristoylated on a glycine residue [267]. In addition to numerous studies in eukaryotes, myristoylome profiles have also been described from some bacteria [31]. In bacteria, protein myristoylation has been described to occur on the N-terminus as well as on N-terminal cysteine residues of inner and outer membrane proteins. It was also demonstrated that these cysteine residues can undergo triacylation, where the first step of triacylation is the alkylation of the prolipoprotein catalyzed by phosphatidylglycerol:prolipoprotein diacylglyceryltransferase. Subsequently, the prolipoprotein signal peptidase cleaves the upstream signal peptide and finally N-acylation can be catalyzed by the (apo)lipoprotein N-acyltransferases [31, 268-270]. The resulting triacylated lipoprotein can be translocated by a special machinery to the outer membrane and inserts there [157]. In addition, proteins transported to the cytoplasmic membrane via the secretory pathway (involving the SecYEG translocon) and the twin arginine transport proteins have been reported to be myristoylated [31]. In *Chlostridium difficile* a proteomic analysis identified 187 lipoproteins to be enriched by using a \equiv MyA probe [271]. In summary, the myristoylation mechanism of bacteria is very different to that in eukaryotes, which probably explains the low amount of myristoylation studies reported for bacteria.

In addition to the bioorthogonal labeling method, the use of radioactive labeled myristate for the identification of myristoylated proteins has been described. However, this method exhibited low sensitivity and represents a time-consuming process due to the very long exposure time required in the detection of radioactive myristate in proteins [272]. In addition, high throughput could hardly be achieved by this method, resulting in the establishment of the bioorthogonal labeling method [272]. The coupled metabolic labeling together with mass spectrometry analysis provides a high-throughput approach to generate acylome data. Despite C14 alkyne probes, a multitude of FA derivatives, usually azide or alkynes, can be used to allow click chemistry and subsequently, proteome analysis [31].

3.2. Myristoylome profiling

In our experiments, we obtained a relatively high biological variance. It has already been described that replicates are required to obtain reliable results from metabolic labeling, thus excluding false positive events [31]. However, this was ensured by the control samples treated without a probe and by the preparation of four biological replicates. Acetylome studies have also reported that reproducibility was not possible in some cases, e.g., *S. enterica* [135]. Therefore, we focused on the proteins identified in multiple experiments, to circumvent false positive candidates.

Our results obtained several proteins involved in lipid transport and metabolism as well as energy production and conversion (arCOG category: C and I). Some enzymes involved in lipid metabolism were already characterized and demonstrated to take part in β -oxidation pathway of *S. acidocaldarius* [28]. In particular, Saci_1123 was identified as abundant protein in the myristoylome data and was reported to comprises medium-chain acyl-CoA dehydrogenase (ACAD) activity during fatty acid degradation and thus catalyzes the first step of the bacterial-like β -oxidation pathway, containing a FAD cofactor transporting electrons to the respiratory chain via the electron transfer flavoprotein (ETF, Saci_0315) [28]. However, due to the high sequence similarity of Saci_1123 and Saci_2217, the differentiation of the two proteins was not possible. Saci_1114, a β -ketothiolase, was shown to cleave acetoacetyl-CoA and catalyzed the reversed reaction by condensing two acetyl-CoA molecules [28]. Moreover, Saci_2209 shares an amino acid sequence identity of 77.3% with the aforementioned Saci_1114 and therefore a similar function is suggested. In *S. acidocaldarius*, the FA metabolism is regulated by a TetR-family transcriptional regulator (FadR_{sa}), by acting as repressor of a 30-kb gene cluster encoding lipases and β -oxidation enzymes (*saci_1103* - *saci_1126*) [27]. Moreover, deletion of *fadR_{sa}* resulted in upregulation of the lipase/fatty acid gene cluster [27]. In the native FadR_{sa} structure a short-chain acyl-CoA was observed in contrast to bacterial counterparts where medium chain acyl-CoAs were observed [27, 273, 274]. In agreement with the observations in *S. acidocaldarius*, long-chain acyl-CoAs were demonstrated to act as metabolic signal for FA degradation by antagonizing the FadR protein in bacteria [275].

The Alba protein was labeled in each experiment of our myristoylome data. Saci_1322 shares high sequence similarity to the Alba protein from *S. solfataricus* (92.8 % sequence identity to SSO0962, BlastX, $1e^{-121}$), which was demonstrated to occur acetylated and non-acetylated and comprises N-terminal as well as N ϵ -lysine acetylation [137]. The acetylation modification was demonstrated to reduce DNA-binding affinity of the chromatin protein [137]. The SsoPat was demonstrated to catalyze the N ϵ -lysine acetylation with low catalytic efficiency compared to other acetyltransferases [138] and more over the structure was solved resembling eukaryotic HATs [253]. Moreover, the characterized N-acetyltransferase SsArd1 contains the

3.2. Myristoylome profiling

eukaryotic NatA catalytic subunit, conserved from archaea to higher eukaryotes [120]. The NatA is mainly known to acetylate Ser, Ala, Thr, Gly and Val at the N-terminal region, followed by a Pro residue to promote protein degradation [111]. In comparison, SsoArd1 was reported to comprise a broader substrate spectrum compared to the human NatA [120]. When considering the two studies acetyltransferases (SsoPat and ssArd1) of *S. solfataricus*, notably both enzymes resemble structurally their eukaryotic counterparts [121, 141, 253]. The enzymes described are assigned to the GNAT superfamily, which includes the NMTs described in eukaryotes that primarily catalyze N-myristoylation. However, it is not clear whether they can also use acyl-CoA as substrates or another enzyme is able to catalyze acylation in archaea. By investigation of the genome of *S. acidocaldarius* several genes were identified encoding N-terminal acetyltransferases belonging to the GNAT superfamily, which was also reported for bacterial genomes [117]. However, whether one of the acetyltransferases from *S. acidocaldarius* can modify proteins via acylation, has to be investigated, but the identification of the Alba protein leads to the assumption that this protein could function as target protein for myristoylation since it was already shown to be acetylated in *S. solfataricus* [137].

In addition, the deacetylation of K16 of Alba was reported to be catalyzed by Sir2 from *S. solfataricus* [138, 276]. Notably, archaeal sirtuins were already reported to catalyze long-chain deacylation in *A. fulgidus* [252]. Two sirtuins, namely Sir2Af1 and Sir2Af2 were demonstrated catalyze desuccinylation and demyristoylation, respectively. Moreover, the crystal structure of SirAf1 revealed an attached succinylated peptide and Sir2Af2 a myristoylated peptide. The Sir2Af2 structure in complex with an acetylated peptide comprised a conformational change, thus indicating biochemical dynamics of the active site. These findings lead to the suggestion that archaeal sirtuins can also catalyze deacylation, as in eukaryotes e.g. the human sirtuin 6 was described to catalyze N ϵ -lysine demyristoylation of K19 and K20 of tumour necrosis factor- α [155, 156, 252]. The genome of *S. acidocaldarius* contains a Sir2 homolog (*saci_0381*), which, according to its homology to Sir2Af2 from *A. fulgidus*, (49 % sequence identity, BlastX, $1e^{-70}$), seems to be a promising candidate to study deacylation of proteins.

The identified Saci_1188 is annotated as alkyl-DHAP synthase (EC 2.5.1.26). This enzyme class is known to catalyze the cleavage of an ester-linked fatty acid of the substrate and to replace it by a long-chain alcohol in an ether-linkage and thereby releasing an carboxylic acid via a ping-pong mechanism [277, 278]. The alkyl-DHAP synthase was reported to be involved in eukaryotic ether phospholipid synthesis. Although archaeal membranes are composed of ether lipids, there is no evidence that this enzyme is involved in the synthesis of archaeal membrane lipids. Nevertheless, it should be mentioned that this protein is among the most abundant in our myristoylome data, but no further information for its function can be provided.

3.2. Myristoylome profiling

However, despite its unknown function there could be a potential for myristoylation and thereby could serve as target for *in vitro* myristoylation studies.

Saci_0196 is annotated as putative mannose-1-phosphate guanyltransferase (Mpg1), which is classified to be involved in translation, ribosomal structure and biogenesis as well as cell wall/membrane/envelope biogenesis (arCOG: J and M). It is predicted to be involved in lipopolysaccharide biosynthesis, but archaea were reported to lack lipopolysaccharide-containing outer membrane [18]. Thus, the function of this candidate remains unsolved. So far, regulation of this enzyme by acylation has not yet been demonstrated and the enzyme was not observed in myristoylome data published to date.

In addition an ACS (Saci_0306) was identified which catalyzes the formation of acetyl-CoA from acetate by AMP-formation. This enzyme is known to be regulated by inactivation via N ϵ -lysine acetylation [135]. In *Salmonella enterica*, acetylation of ACS by the protein acetyltransferase SePat was the first time shown to inactivate the enzyme, followed by several other studies on reversible acetylation of ACS [131, 135, 279]. Moreover, the propionyl-CoA synthetase of *S. enterica* was reported to be controlled by propionylation [135] indicating a regulation of ACS by their products. In bacteria, structural analysis demonstrated the acetylation of a lysine residue in the active site, which blocks the acetyl-CoA synthesis. Moreover, the cleavage of the acetyl-moiety by a deacetylase was described, resulting in the activation of ACS and thereby leads to accumulation of AMP and depletion of ATP [117]. Thereby the ACS is one example how protein acetylation can be involved in regulation of the energy charge of the cell.

A proteomic analysis of lysine acetylation has already been reported for *Haloflex mediterranei*, where 17 % of total proteins were observed to be acetylated, whereas the identified proteins were mainly predicted to be involved in metabolic and cellular processes [280]. Most proteins were observed to be distributed in the cytoplasm, whereas only a few proteins were located in the cytoplasmic membrane [280]. Moreover, the acetylome study of *S. islandicus* identified about 26 % of total proteins to be N ϵ -acetylated [141]. Several abundant proteins were obtained to be involved in information storage and processing, mainly translation, ribosomal structure and biogenesis (J). Only a few proteins were identified to play a role in cellular processes and signaling. Notably, a high number of proteins was predicted to be involved in metabolism, predominantly amino acid transport and metabolism (E), energy production and conversion (C), carbohydrate transport and metabolism (G) as well as lipid transport and metabolism (I). Moreover, the lysine-specific acetylome data from *T. gammaolerans* identified mainly proteins involved in metabolism followed by information storage and processing [140]. In particular, the classification of enzymes comprised a clear consensus with the *S. islandicus* acetylome by identification of mainly proteins predicted to be

3.2. Myristoylome profiling

catogrized in arCOG class C, E and G. The predicted functions of identified candidates resemble our myristoylome data, where several proteins were identified to belong to arCOG class C, E and I. Considering the single candidates, ACS homologs were predominantly present in the above mentioned acetylomes as well as in our myristoylome data [140, 141]. In addition, an acyl-CoA dehydrogenase, two thiolases as well as one homolog of the Alba protein were identified to be N ϵ -lysine acetylated in *S. islandicus* overlapping with our myristoylome data [141].

Comparing the eukaryotic myristoylome profiling and the one from *S. acidocaldarius*, the results show a slight similarity. Firstly, the number of myristoylated eukaryotic proteins identified shows a wide range from 238 to 4,469 candidates including plants, animals and unicellular organisms like fungi [281]. The eukaryotic studies revealed, that most myristoylated proteins identified in plants and animals comprise a catalytic function followed by involvement in cell organization and sensing, which partially resembles our results. Regarding the protein family most proteins comprise a kinase, proteostasis, hydrolase and oxidoreductase function [281]. Several targets for eukaryotic NMTs were identified including ADP-ribosylation factor, 26S proteasome, the Vps20 subunit of the ESCRT-III complex (endosomal sorting complex required for transport), an EF-hand calcium-binding protein of the recoverin family as well as a calcium-dependent protein kinase were identified as G-myristoylation targets. However, these candidates were not found in our myristoylome data.

The results obtained from metabolic labeling resemble the archaeal lysine-acetylome data with respect to the protein classification as well as showing certain candidates that were already reported to be regulated by acetylation (ACS, Alba, thiolase) [140, 141]. Thus, it can be assumed, that our myristoylome is a functional approach providing some promising candidates for myristoylation. Since the archaeal acetylomes addressed lysine modifications, it is possible that the same candidates found in the myristoylome of *S. acidocaldarius* are N ϵ -lysine myristoylated. Since reversible lysine acetylation has already been described for the Alba homolog of *S. solfataricus*, an investigation of long-chain modification e.g. myristoylation by using the corresponding homologs of *S. acidocaldarius* would could serve as an additional verification to confirm our results.

In conclusion, the *in vivo* metabolic labeling approach provided some potential candidates for protein myristoylation in *S. acidocaldarius*, thus offering important insights into possible functions of FAs in archaea. From the results obtained, it could be suggested that a regulation on protein level via N ϵ -lysine myristoylation could be present in archaea. Therefore, the identified proteins offer some potential candidates to further investigate protein acylation and will thereby provide some important insights into the hardly understood functions of FAs in archaea.

Materials and Methods

Strains and growth conditions

S. acidocaldarius MW001 (uracil auxotroph mutant [97]) was grown in Brock's basal medium at 76 °C and pH 3.0 supplemented with 0.3 % (w/v) dextrin, 0.1% (w/v) N-Z-Amine and 10 µg/mL of uracil (180 rpm).

Metabolic labeling with myristic acid alkyne

S. acidocaldarius MW001 strain was grown in Brock's basal medium at 76 °C and pH 3.0 supplemented with 0.3 % (w/v) dextrin, 0.1% (w/v) N-Z-Amine and 10 µg/mL of uracil. For metabolic labelling, the cultures of OD₆₀₀ = 1.0 - 1.2 were diluted to an OD₆₀₀ of 1.0 and labelled with 20 µM myristic acid alkyne probe (≡MyA) (from 25 mM stock in ethanol/DMSO). As negative control, an equal volume of ethanol or DMSO was added to the culture. The OD₆₀₀ was measured before and after 6 hours of incubation at 76 °C and 180 rpm, respectively. The experiment was performed in quadruplicates. The cells were harvested at room temperature (20 °C) and supernatant was removed. The cell pellets were flash frozen in liquid nitrogen and stored at -80 °C overnight.

Protein extraction and click reactions for in-gel analysis

The proteins were extracted by firstly resuspending the cell pellets in 50 mM HNa₂PO₄ pH 8.0 supplemented with 1× MS-SAFE Protease and Phosphatase Inhibitor (Sigma-Aldrich, United States). After transfer into 1.5 mL reaction tubes, cell disruption took place via sonication using a bioruptor (Diagenode, Belgium) with 10 cycles (1 min pulse, 30 sec break). The cleared lysate was obtained after two centrifugation steps (21,000 x g, 20 min and 5 min, 4 °C) and protein concentration was determined via ROTI[®]Nanoquant (Carl Roth, Germany).

To confirm successful labeling, samples were analyzed via in-gel fluorescence analysis. Therefore, 25 µg of protein were used for click reaction with 0.01 mM of the fluorescent dye Cy3-N₃, 0.1 mM tris((1-benzyl-4-triazolyl)methyl)amine (TBTA; Sigma-Aldrich, United States), 2 mM tris(2-carboxyethyl)phosphine (TCEP; Sigma-Aldrich, United States) and 2 mM CuSO₄. The samples were incubated for 1 hour in the dark and afterwards 4x loading buffer supplemented with 0.1 mM dithiothreitol (DTT; Sigma-Aldrich, United States) was added and mixed for 15 minutes at 70 °C with 750 rpm of shaking. 20 µg of protein were loaded on a 11 % Bis-Tris gel (51.33 mL 30 % acrylamide/bisacrylamide solution, 40.04 mL 1.25 M Bis-Tris buffer, 0.57 mL 10 % (w/v) APS, 0.116 mL TEMED, 48 mL MS-grade ultrapure water) and the gel was run for 1.5 hours at 120 V and 55 mA using

3.2. Myristoylome profiling

3-(*N*-morpholino)propanesulfonic acid (MOPS) buffer. The fluorescence of the gel was scanned using a Typhoon laser scanner with the green laser emitting light at a wavelength of 532 nm. The gel was stained with colloidal Coomassie and destained in 10 % acetate overnight.

Click chemistry for on-bead digestion

For affinity purification, a large scale reaction containing 500 µg of protein was used for click chemistry as described, but with 5/6-TAMRA- N₃-biotin (Jena Bioscience; CLK-1048) instead of Cy3-N₃. Afterwards the proteins were precipitated using a 4 equivalents of methanol and incubated at -20 °C overnight. After an additional step of washing with methanol, the pellets were air dried and resuspended in 1x PBS buffer (155 mM NaCl, 3 mM Na₂HPO₄, 1.06 mM KH₂PO₄ in MS grade ultrapure water; pH 7.4) containing 2 % SDS at 37 °C with gentle agitation. After centrifugation, the supernatant was used and diluted 10-fold with 1x PBS. The protein solution was added to avidin beads (Pierce™ Avidin Agarose, Thermo Scientific, United States; 20219) and incubated for one hour rotating at room temperature (20 °C). After centrifugation (400 × *g*, 5 min), the supernatant was removed and the beads were washed five and four times with 1 % SDS and MS-pure water, respectively.

For on-bead tryptic digestion, the beads were taken up in 50 mM ammonium bicarbonate buffer (ABC) containing 0.8 M urea and 10 mM DTT followed by incubation for one hour at room temperature with shaking at 1,500 rpm. Afterwards, 25 mM iodoacetamide (IAM) was added and the beads were again incubated for one hour. DTT was supplemented to reach a final concentration of 35 mM and the samples were incubated for another 10 minutes. In total, 1 µg trypsin (Pierce Trypsin, Thermo Scientific, United States; 90057) in 50 mM acetic acid was added per sample and proteins were digested overnight at 37 °C. The peptide mix was recovered after centrifugation (3000 × *g*, 1 min) and formic acid (FA) was added to a final concentration of 5 %. Next, the beads were washed with 1 % (v/v) FA and the solution was combined with the recovered peptide mix. To remove remaining beads, the peptide mix was passed over a two-disc glass fiber (pore size 1.2 µm, thickness 0.26 mm, Cytiva Lifesciences, United States) tip. The samples were desalted using two-disc C18 Stage tips [282]. The Stage tips were activated with 50 µL methanol (600 × *g*, 2 min) and equilibrated with 80 % (v/v) acetonitrile (ACN) with 0.5 % (v/v) FA (600 × *g*, 2 min) and 0.5 % (v/v) FA (800 × *g*, 3 min). The tips were loaded twice with the digested peptides (800 × *g*, 3-4 min). Peptides were washed with 0.5 % (vol/vol) FA solution and for elution a 0.5% (vol/vol) FA solution in 80 % (vol/vol) ACN/H₂O was used. The eluted peptides were dried using a vacuum concentrator (Eppendorf, Germany) and stored at -20 °C. Prior to MS analysis, the peptides were resuspended in 10 µL 0.1 % FA.

3.2. Myristoylome profiling

LC-MS/MS Analysis

For the experiments an Orbitrap Elite or Orbitrap Fusion Lumos mass spectrometer (Thermo Fischer Scientific, USA) was used coupled to an EASY-nLC 1000 or 1200 liquid chromatography (LC) system (Thermo Fischer scientific, USA) operating in the one-column mode. A fused silica capillary (inner diameter 75 μm \times 37 cm or 45 cm) with an integrated PicoFrit emitter (New Objective, USA or ESI Source Solutions) packed in-house with Reprosil-Pur 120 C18-AQ 1.9 μm (Dr. Maisch, Germany) was used as analytical column. A column oven (Sonation, Germany) encased the column and oven temperature was adjusted to 45 $^{\circ}\text{C}$ or 50 $^{\circ}\text{C}$ during data acquisition. The analytical column was also attached to a nanospray flex ion source (Thermo Fischer Scientific, USA). The two mobile phases of LC were: solvent A (0.1 % FA, in water) and solvent B (0.1 % FA, 20 % H_2O , in ACN); only ACE_0408: 0.1 % FA in ACN). All solvents were of UHPLC (ultra-high-performance liquid chromatography) grade (Honeywell, Germany). The peptides were directly loaded onto the analytical column, which exhibited a maximum flow rate not exceeding the set pressure limit of 980 bar (usually around 0.5 – 0.8 $\mu\text{l min}^{-1}$). Subsequently, peptide solutions were separated on the analytical column using different gradients (SI Fig. 6).

The mass spectrometers were operated using Xcalibur software (Elite: v2.2 SP1.48; Lumos: v4.3.7.3.11) and set in the positive ion mode. In the Orbitrap analyzer (FTMS; Fourier Transform Mass Spectrometry with the internal lock mass option turned on (lock mass was 445.120025 m/z , polysiloxane) precursor ion scanning (MS¹) was performed [283]. MS² Product ion spectra were recorded only from ions with a charge $> +1$ and in a data dependent fashion in the ITMS. All relevant MS settings (Resolution, scan range, AGC, ion acquisition time, charge states isolation window, fragmentation type and details, cycle time, number of scans performed, and various other settings) for the individual experiments can be found in the (SI Fig. 7).

Peptide and protein identification using MaxQuant and Perseus

RAW spectra were submitted to an Andromeda search [284] in MaxQuant (version 1.6.2.6 or 2.0.1.0) with the default settings [285], where Label-free quantification (LFQ) [286] and match between runs were activated. MS/MS spectra data were searched against the UniProt *S. acidocaldarius* (DSM639) database (UP000001018-330779; 2221 resp. 2222 entries) with all searches including a contaminants database (as implemented in MaxQuant, 246 sequences), which contains known MS contaminants and was included to estimate the level of contamination. Andromeda searches allowed dynamic modifications viz. oxidation of methionine residues (16 Da) and acetylation of the protein N-terminus (42 Da) while

3.2. Myristoylome profiling

carbamidomethylation of cysteine residues (57 Da, alkylation with IAM) was selected as static modification. Enzyme specificity was set to “Trypsin/P” and the instrument type was set to Orbitrap in Andromeda searches. The precursor mass tolerance was set to ± 20 ppm (first search) and ± 4.5 ppm (main search), where the FTMS MS/MS match tolerance was set to ± 20 ppm. The peptide spectrum match FDR and the protein FDR were set to 0.01 (based on target-decoy approach). Minimum peptide length was 7 amino acids. Unique and razor peptides were allowed for protein quantification. For quantification, modified peptides with dynamic modifications were additionally allowed to unmodified peptides, where the minimum score for modified peptides was set to 40 [287].

Perseus software (version 1.6.2.1 or 1.6.15.0) was used for further data analysis and filtering of the MaxQuant [288]. Label-free quantification (LFQ) intensities were loaded into the matrix from the proteinGroups.txt file, removing potential contaminants, reverse hits from the reverse database and hits only identified based on peptides with a modification site. The data was transformed to the \log_2 -scale or all MS-based proteomics experiments and to allow comparison of the different treatment groups biological replicates were combined into categorical groups. For further analysis only hits with a valid LFQ intensity for a minimum of two samples in total were retained. Missing values were imputed from a normal distribution (width 0.3, down shift 1.8) before quantification. Between different MS runs, comparison of normalized protein group quantities (relative quantification) was solely based on the LFQ intensities as calculated by MaxQuant (MaxLFQ algorithm) [286]. Briefly, LFQ was switched on and unique and razor peptides with a minimum ratio count of 2 were considered for quantification. The built-in nonlinear time-rescaling algorithm allowed recalibration of retention times. MS/MS identifications were transferred between LC-MS/MS runs with the “Match between runs” option using a match time window of 0.7 min and an alignment time window of 20 min. The quantification was based on the “value at maximum” of the extracted ion current. At least two quantification events were required for a quantifiable protein. The \log_2 -fold enrichment of protein groups with myristoyl-alkyne was calculated based on the mean LFQ intensity compared to the control group lacking the probe. Statistical significance was calculated based on a two-sided Student’s t-test [287]. The results were reported in Fig. 3. Proteins with a \log_2 -fold change < 1 were defined as abundant compared to the non-probe samples and termed as hits for evaluation of proteomic data. For ACE_0604 all proteins with a \log_2 fold change > 0 were considered, due to the low number of hits in the experiment. The remaining proteins were reported in SI Tab.1 – 4. The remaining hits from all four experiments (SI Tab. 5) were compared, resulting in a partial overlap presented in a Venn diagram Fig. 4. The hits obtained from all experiments were used for classification using arCOG database [170].

Supplementary Materials

Proteomic analysis of myristoylation in *Sulfolobus acidocaldarius* using *in vivo* metabolic labeling

Authors

Svenja Höfmann¹⁺, Sabrina Ninck²⁺, Xiaoxiao Zhou¹, Farnusch Kaschani², Christopher Bräsen¹, Markus Kaiser^{2*}, Bettina Siebers^{1*}

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

²Department of Chemical Biology, Center of Medical Biotechnology, Faculty of Biology, University of Duisburg-Essen, Essen, Germany

*Corresponding authors: Markus Kaiser (Chemical Biology) markus.kaiser@uni-due.de, Prof. Dr. Bettina Siebers (Biochemistry): bettina.siebers@uni-due.de

+These authors are contributed equally to this work.

3.2. Myristoylome profiling

SI Table 1: Differentially abundant proteins from *S. acidocaldarius* MW001 identified from *in vivo* metabolic labeling at 76 °C (ACE_0408).

The protein abundance profile of cells containing the \equiv MyA probe compared to non-probe cells.

The proteins showing a \log_2 fold change > 1 were abundant the \equiv MyA containing samples compared to the non-probe samples. The statistical significance was calculated based on a two-sided Student's test (p -value > 0.05 corresponds to $-\log(p\text{-value}) > 1.301$). The annotations of identified abundant proteins were assigned.

Protein-ID	Annotation	$-\log(p\text{-value})$	\log_2 -fold change
Saci_1114	Thiolase	3.107	7.708
Saci_2209	Acetyl-CoA thiolase	1.881	5.332
Saci_1726	CoA-transferase family III	4.377	4.974
Saci_1188	Alkyldihydroxyacetonephosphate synthase	1.414	4.730
Saci_1121	Acetyl-CoA thiolase	4.604	4.639
Saci_1362	UPF0219 protein Saci_1362	1.244	4.110
Saci_0336	Conserved protein	2.823	4.031
Saci_1123;Saci_2217	Acyl-CoA dehydrogenase	3.909	3.725
Saci_1241	Thioesterase superfamily protein	3.006	3.699
Saci_1116	Acetyl esterase	2.848	2.781
Saci_1361	Acetyl-CoA acetyltransferases	3.846	2.631
Saci_2255	2-oxoacid oxidoreductase (ferredoxin)	1.713	2.581
Saci_0998	Purine phosphoribosyltransferase	2.095	2.320
Saci_0064;Saci_0362	DNA-binding protein 7d/e	0.405	2.005
Saci_2272	Rubrerhythrin	1.530	1.968
Saci_0854	Thiamine thiazole synthase	1.243	1.923
Saci_2154	NAD ⁺ -dependent malic enzyme	0.807	1.786
Saci_1718	Conserved protein	0.515	1.756
Saci_0646	Adenosylhomocysteinase	0.551	1.644
Saci_1021	Conserved Crenarchaeal protein	1.293	1.643
Saci_0768	Aspartate--tRNA(Asp/Asn) ligase	0.387	1.602
Saci_0208	2-oxoacid oxidoreductase (ferredoxin)	0.798	1.530
Saci_1322	DNA/RNA-binding protein Alba	1.674	1.420
Saci_0828	Methionine synthase	1.392	1.378
Saci_1331	S-adenosyl-L-methionine-dependent tRNA 4-demethylwyosine synthase	0.822	1.366
Saci_1359	3-hydroxy-3-methylglutaryl coenzyme A reductase	1.087	1.346
Saci_0715	DNA repair and recombination protein RadA	2.104	1.243
Saci_2233	Acetyl-CoA C-acetyltransferase	0.617	1.193
Saci_1109	Enoyl-CoA hydratase	0.436	1.168
Saci_0813	50S ribosomal protein L14e	0.582	1.165
Saci_1548	V-type ATP synthase α chain	0.255	1.146
Saci_0196	Putative mannose-1-phosphate guanyltransferase	0.611	1.145
Saci_0329	Conserved protein	0.485	1.094
Saci_1033	Acyl-CoA dehydrogenase	0.850	1.068
Saci_1358	Serine hydroxymethyltransferase	0.621	1.051

3.2. Myristoylome profiling

Saci_0306	Acetyl-coenzyme A synthetase	0.683	1.031
Saci_1585	Conserved Archaeal protein	0.267	1.024

3.2. Myristoylome profiling

SI Table 2: Differentially abundant proteins from *S. acidocaldarius* MW001 identified from *in vivo* metabolic labeling at 76 °C (ACE_0598).

The protein abundance profile of cells containing the \equiv MyA probe compared to non-probe cells.

The proteins showing a \log_2 fold change > 1 were abundant the \equiv MyA containing samples compared to the non-probe samples. The statistical significance was calculated based on a two-sided Student's test (p -value > 0.05 corresponds to $-\log(p\text{-value}) > 1.301$). The annotations of identified abundant proteins were assigned.

Protein-ID	Annotation	$-\log(p\text{-value})$	\log_2 -fold change
Saci_1322	DNA/RNA-binding protein Alba	3.750	7.372
Saci_0351	Biotin/lipoate A/B protein ligase	2.286	6.012
Saci_1123;Saci_2217	Acyl-CoA dehydrogenase	3.238	4.828
Saci_1188	Alkyldihydroxyacetonephosphate synthase	2.727	4.197
Saci_2209	Acetyl-CoA thiolase	4.442	3.901
Saci_0594	50S ribosomal protein L2	1.268	3.451
Saci_1358	Serine hydroxymethyltransferase	1.572	3.255
Saci_1114	Thiolase	1.068	2.813
Saci_1727	Conserved Archaeal protein	0.849	2.468
Saci_0706	Adenylosuccinate lyase	2.271	2.449
Saci_2355	S-layer protein A	1.550	2.373
Saci_0828	Methionine synthase	0.468	2.327
Saci_0703	Conserved Archaeal radical SAM superfamily protein	2.342	2.320
Saci_0336	Conserved protein	1.041	2.311
Saci_0076	Tyrosine--tRNA ligase	1.240	2.065
Saci_1265	Succinyl-CoA synthetase β chain	1.406	1.998
Saci_2208	3-hydroxybutyryl-CoA dehydrogenase	1.171	1.890
Saci_2231	Conserved Archaeal protein	0.894	1.775
Saci_1214	Aconitate hydratase	0.669	1.635
Saci_0176	50S ribosomal protein L10e	0.523	1.625
Saci_0196	Putative mannose-1-phosphate guanyltransferase	1.866	1.613
Saci_0349	Probable glycine cleavage system H protein 3	0.706	1.461
Saci_2344	NADH-ubiquinone oxidoreductase formyltransferase	0.678	1.440
Saci_0911	Alcohol dehydrogenase	1.062	1.328
Saci_1421	Tryptophan synthase β chain	0.481	1.266
Saci_2235	Acyl-CoA synthetase	0.387	1.248
Saci_2286	Enoyl-CoA hydratase/isomerase	1.442	1.192
Saci_1393	Alcohol dehydrogenase	0.616	1.110
Saci_2225	Medium-chain-fatty-acid-CoA ligase	0.710	1.076
Saci_0325	Heterodisulfide reductase B	0.534	1.063

3.2. Myristoylome profiling

SI Table 3: Differentially abundant proteins from *S. acidocaldarius* MW001 identified from *in vivo* metabolic labeling at 76 °C (ACE_0593).

The protein abundance profile of cells containing the \equiv MyA probe compared to non-probe cells.

The proteins showing a \log_2 fold change > 1 were abundant the \equiv MyA containing samples compared to the non-probe samples. The statistical significance was calculated based on a two-sided Student's test (p -value > 0.05 corresponds to $-\log(p\text{-value}) > 1.301$). The annotations of identified abundant proteins were assigned..

Protein-ID	Annotation	$-\log(p\text{-value})$	\log_2 -fold change
Saci_1322	DNA/RNA-binding protein Alba	3.566	4.751
Saci_2233	Acetyl-CoA C-acetyltransferase	1.293	2.694
Saci_2288	Nonspecific lipid-transfer protein	1.760	2.563
Saci_2230	Conserved protein	1.378	2.561
Saci_1188	Alkyldihydroxyacetonephosphate synthase	1.375	2.352
Saci_1114	Thiolase	1.136	2.230
Saci_0155	Glutamate dehydrogenase	0.809	2.209
Saci_0196	Putative mannose-1-phosphate guanyltransferase	2.436	2.183
Saci_2244	Conserved amidohydrolase	1.101	2.076
Saci_0703	Conserved Archaeal radical SAM superfamily protein	1.493	2.021
Saci_1311	Translation initiation factor 5A	3.081	1.995
Saci_0981	Succinate dehydrogenase subunit B	1.250	1.891
Saci_0980	Succinate dehydrogenase subunit C	1.367	1.886
Saci_0917	Carbon monoxide dehydrogenase small chain	1.395	1.849
Saci_2298	Acetyl-coenzyme A synthetase	1.012	1.831
Saci_2295	Catechol 2,3-dioxygenase	1.024	1.744
Saci_1122	Acyl-CoA synthetase	1.367	1.721
Saci_0613	Proteasome subunit α	2.011	1.693
Saci_1549	V-type ATP synthase β chain	1.341	1.621
Saci_0306	Acetyl-coenzyme A synthetase	0.572	1.581
Saci_0666	Thermosome subunit β	0.831	1.575
Saci_0315	Electron transfer flavoprotein	1.292	1.536
Saci_0214	Formate-dependent phosphoribosylglycinamide formyltransferase	0.737	1.501
Saci_2140	Parathion hydrolase	0.918	1.435
Saci_0599	Isoleucine--tRNA ligase	0.586	1.430
Saci_0688	30S ribosomal protein S12	0.716	1.316
Saci_1572	Leucine--tRNA ligase 2	0.353	1.192
Saci_2271	Glyceraldehyde dehydrogenase large chain	0.258	1.146
Saci_2198	Sulfurtransferase	0.408	1.094
Saci_1214	Aconitate hydratase	1.072	1.053
Saci_2209	Acetyl-CoA thiolase	0.735	1,006

3.2. Myristoylome profiling

SI Table 4: Differentially abundant proteins from *S. acidocaldarius* MW001 identified from *in vivo* metabolic labeling at 76 °C (ACE_0604).

The protein abundance profile of cells containing the \equiv MyA probe compared to non-probe cells.

The proteins showing a \log_2 fold change > 1 were abundant the \equiv MyA containing samples, for overlap evaluation for ACE_0604 all proteins with a \log_2 fold change > 0 were considered, due to the low number of hits. The statistical significance was calculated based on a two sided Student's test (p value > 0.05 corresponds to $-\log(p\text{-value}) > 1.301$). The annotations of identified abundant proteins were assigned.

Protein-ID	Annotation	$-\log(p\text{-value})$	\log_2 -fold change
Saci_0351	Biotin/lipoate A/B protein ligase	4.158	6.217
Saci_0336	Conserved protein	1.674	5.179
Saci_1322	DNA/RNA-binding protein Alba	1.458	4.667
Saci_1359	3-hydroxy-3-methylglutaryl coenzyme A reductase	2.232	4.455
Saci_1079	Glucose 1-dehydrogenase	2.354	3.374
Saci_1078	3-hydroxybutyryl-CoA dehydrogenase	1.389	3.277
Saci_1188	Alkyldihydroxyacetonephosphate synthase	1.392	2.297
Saci_0001	ORC1-type DNA replication protein 3	0.178	2.295
Saci_2339	Membrane linked NADH-ubiquinone oxidoreductase	0.582	2.288
Saci_2209	Acetyl-CoA thiolase	0.469	1.969
Saci_0763	2-oxoacid oxidoreductase (ferredoxin)	0.500	1.446
Saci_0617	50S ribosomal protein L15e	0.260	1.392
Saci_0349	Probable glycine cleavage system H protein 3	0.497	1.243
Saci_2230	Conserved protein	0.557	1.126
Saci_1549	V-type ATP synthase β chain	0.259	0.965
Saci_0331	Pyridine nucleotide-disulphide oxidoreductase	1.064	0.946
Saci_2255	2-oxoacid oxidoreductase (ferredoxin)	0.364	0.898
Saci_2271	Glyceraldehyde dehydrogenase large chain	0.180	0.873
Saci_1114	Thiolase	0.487	0.871
Saci_2375	Isocitrate dehydrogenase (NADP ⁺)	0.088	0.676
Saci_0243	Citrate synthase	0.110	0.669
Saci_1401	Thermosome subunit α	0.298	0.655
Saci_1199	Glucose-1-phosphate adenylyltransferase	0.409	0.644
Saci_0685	Elongation factor 1- α	0.263	0.523
Saci_0911	Alcohol dehydrogenase	0.199	0.517
Saci_0214	Formate-dependent phosphoribosylglycinamide formyltransferase	0.243	0.420
Saci_1629	Conserved protein	0.096	0.319
Saci_0261	Conserved protein	0.094	0.197
Saci_0944	Conserved Archaeal protein	0.125	0.093
Saci_0692	DNA-directed RNA polymerase subunit Rpo1N	0.004	0.015

3.2. Myristoylome profiling

SI Table 5: Overlap of differentially abundant proteins from *S. acidocaldarius* MW001 identified from four *in vivo* metabolic labeling experiments (76 °C).

The protein abundance profile of cells containing the \equiv MyA probe compared to non-probe cells.

The proteins showing a \log_2 fold change > 1 were abundant the \equiv MyA containing samples and were used for overlap evaluation. For ACE_0604 all proteins with a \log_2 fold change > 0 were considered, due to the low number of hits. The statistical significance was calculated based on a two sided Student's test (p value > 0.05 corresponds to $-\log(p\text{-value}) > 1.301$). The annotations of identified abundant proteins were assigned.

Protein-ID	Annotation	ACE_0408	ACE_0589	ACE_0593	ACE_0604
Saci_1322	DNA/RNA-binding protein Alba	+	+	+	+
Saci_1188	Alkyldihydroxyacetonephosphate synthase	+	+	+	+
Saci_1114	Thiolase	+	+	+	+
Saci_0196	Putative mannose-1-phosphate guanyltransferase	+	+	+	
Saci_0336	Conserved protein	+	+		+
Saci_2209	Acetyl-CoA thiolase	+	+		+
Saci_0828	Methionine synthase	+	+		
Saci_1123;Saci_2217	Acyl-CoA dehydrogenase	+	+		
Saci_1358	Serine hydroxymethyltransferase	+	+		
Saci_0306	Acetyl-coenzyme A synthetase	+		+	
Saci_2233	Acetyl-CoA C-acetyltransferase	+		+	
Saci_2255	2-oxoacid oxidoreductase (ferredoxin)	+			+
Saci_1359	3-hydroxy-3-methylglutaryl coenzyme A reductase	+			+
Saci_0703	Conserved Archaeal radical SAM superfamily protein		+	+	
Saci_1214	Aconitate hydratase		+	+	
Saci_0349	Probable glycine cleavage system H protein 3		+		+
Saci_0911	Alcohol dehydrogenase		+		+
Saci_0351	Biotin/lipoate A/B protein ligase		+		+
Saci_0214	Formate-dependent phosphoribosylglycinamide formyltransferase			+	+
Saci_1549	V-type ATP synthase β chain			+	+
Saci_2230	Conserved protein			+	+
Saci_2271	Glyceraldehyde dehydrogenase large chain			+	+

3.2. Myristoylome profiling

SI Table 6: Gradients

ACE_0408	Thermo Orbitrap Elite/ Thermo Easy-nLC 1000	Solvent A: 0.1% FA in water Solvent B: 0.1% FA in ACN	140 min gradient (gradient: start with 7% B; gradient 7-35% B for 120 min; gradient 35-80% B for 10 min and 80% B for 10 min)
ACE_0589	Thermo Orbitrap Elite/ Thermo Easy-nLC 1000	Solvent A: 0.1% FA in water Solvent B: 0.1% FA in 80% ACN	140 min gradient (gradient: start with 7% B; gradient 7-35% B for 120 min; gradient 35-80% B for 10 min and 80% B for 10 min)
ACE_0593	Thermo Orbitrap Elite/ Thermo Easy-nLC 1000	Solvent A: 0.1% FA in water Solvent B: 0.1% FA in 80% ACN	140 min gradient (gradient: start with 9% B; gradient 9-44% B for 120 min; gradient 44-100% B for 10 min and 100% B for 10 min)
ACE_0604	Thermo Orbitrap Fusion Lumos/ Thermo Easy- nLC 1200	Solvent A: 0.1% FA in water Solvent B: 0.1% FA in 80% ACN	105 min gradient (gradient: start with 9% B; gradient 9-44% B for 90 min; gradient 44-100% B for 10 min and 100% B for 5 min)

3.2. Myristoylome profiling

SI Table 7: MS_Settings

Project	MS1	MS2	Comments; special settings
ACE_0408	Analyzer: FT Res.: 60000 SR: 300 - 1800 AGC: 3x10 ⁶ AcT: 50 ms RF: 30 SF: -- DDM: NS/15	Analyzer: IT Res./ScR: -/Rapid SR: Auto AGC: 1x10 ⁴ AcT: 60 ms CS: +2 and higher IsM: IT IsW: 2.0 Frag.: CID NCE: 35	classic orbitrap experiment: MS1 in Orbitrap at high resolution and data dependent MS2 in lontrap at rapid scan rate. Dynamic exclusion enabled (exclude after n times=1; Exclusion duration (s)= 120; mass tolerance= ± 10ppm)
ACE_0589	Analyzer: FT Res.: 60000 SR: 300 - 1800 AGC: 3x10 ⁶ AcT: 50 ms RF: 30 SF: -- DDM: NS/15	Analyzer: IT Res./ScR: -/Rapid SR: Auto AGC: 1x10 ⁴ AcT: 60 ms CS: +2 and higher IsM: IT IsW: 2.0 Frag.: CID NCE: 35	classic orbitrap experiment: MS1 in Orbitrap at high resolution and data dependent MS2 in lontrap at rapid scan rate. Dynamic exclusion enabled (exclude after n times=1; Exclusion duration (s)= 120; mass tolerance= ± 10ppm)
ACE_0593	Analyzer: FT Res.: 60000 SR: 300 - 1800 AGC: 3x10 ⁶ AcT: 50 ms RF: 30 SF: -- DDM: NS/15	Analyzer: IT Res./ScR: -/Rapid SR: Auto AGC: 1x10 ⁴ AcT: 60 ms CS: +2 and higher IsM: IT IsW: 2.0 Frag.: CID NCE: 35	classic orbitrap experiment: MS1 in Orbitrap at high resolution and data dependent MS2 in lontrap at rapid scan rate. Dynamic exclusion enabled (exclude after n times=1; Exclusion duration (s)= 120; mass tolerance= ± 10ppm)
ACE_0604	Analyzer: FT Res.: 240000 SR: 375 - 1500 AGC: Standard AcT: 50 ms RF: 30 SF: -- DDM: CT/3sec	Analyzer: IT Res./ScR: -/Turbo SR: Auto AGC: 300% AcT: 60 ms CS: +2 to +7 IsM: Q IsW: 1.6 Frag.: HCD NCE: 30	classic orbitrap experiment: MS1 in Orbitrap at high resolution and data dependent MS2 in lontrap at turbo scan rate. Dynamic exclusion enabled (exclude after n times=1; Exclusion duration (s)= 120; mass tolerance= ± 10ppm)

Note: **FT**= Fourier Transform (Orbitrap); **IT**= Iontrap; **Q**= Quadrupol; **Res.**= max. Resolution at 200 m/z (Lumos) or 400 m/z (Elite) [FWHM (full width at half maximum)]; **ScR**= scan rate for measurements in the IT; **SR**= scan range [m/z]; **AGC**= automatic gain control, max number of acquired ions per measurement; **AcT**= max. Ion acquisition time [ms]; **CS**= charge states used for fragmentation; **IsM**= Isolation mode (Q or IT), MS2 isolation and further is only done in IT; **IsW**= Isolation window [m/z], value followed by scan mode the isolation is based on (MS1, MS2 ...); **Frag.**= Fragmentation method; **HCD**= Higher-energy collisional dissociation; **CID**= Collision-induced dissociation; **ETD**= Electron-transfer dissociation; **ETHcd**= Electron-Transfer/Higher-Energy Collision Dissociation; **sHCD**= stepped HCD; **NCE**= normalized collision energy; **cycles**: number of MSn recorded or max cycle time; RF= RF Lens [%]; **SF**= Source Fragmentation [V]; **DDM**: Data dependent Mode (cycle time in seconds, CT/[s] or number of scans, NS); **NS**= Number of data dependent scans.

4 Summary

Archaea were introduced as the third domain of life in the late 1970s. Nevertheless, both chapters of this thesis illustrate that essential metabolic pathways and cellular processes are still not fully understood. An important feature of the archaeal domain is the great variance of organisms, containing some members which are able to adapt to diverse extreme habitats. Previous studies already demonstrated the function of polyP in Archaea, acting in a variety of stress responses and thus helping the cell to adapt to new environmental conditions. Likewise, this polymer was described to provide energy through high-energy phosphoanhydride bonds, as well as having many other functions such as phosphate storage, influence on motility, biofilm formation and copper tolerance in Archaea.

With respect to polyP metabolism studied in chapter 3.1, it has been determined for Archaea by bioinformatic analysis that many of these organisms do not comprise a complete metabolic pathway. This includes the Crenarchaeon *Sulfolobus acidocaldarius*, where no enzyme responsible for polyP synthesis could be ascertained. However, the presence of polyP, as well as the presence of an exopolyphosphatase (PPX), the enzyme responsible for polyP hydrolysis, have been demonstrated in this organism. Accordingly, PPX was biochemically characterized and showed high specific activity, favoring long-chain polyP (P₇₀₀) degradation. Using ³¹P NMR analysis, enzymatic degradation of polyP₄₅ was demonstrated, yielding the enzymatically catalyzed formation of orthophosphate and the heat-stable by-product cyclic-P₃ (non-enzymatically). In the search for the polyphosphate kinase, the examination of the genome organization revealed two thymidylate kinase (dTMPK) genes (*saci_2019*, *saci_2020*) in the gene neighborhood of the *ppx* gene, with dTMPKs sharing high structural similarity to PPK2 enzymes. Moreover, a third *dtmpk* gene (*saci_0893*) was examined, which is differently located in the genome and is the only one of the three dTMPKs comprising dTMPK function. Moreover, *Saci_2019* was demonstrated to function as PPK, with *Saci_2020* stimulating PPK activity, resulting in an increased specific activity for the heteromeric enzyme. Heteromeric PPK catalyzed polyP formation directly from ATP without the use of a precursor, meanwhile ADP inhibition and a clear preference for ATP as a substrate were observed. Consistent with PPK2 features, the enzyme also catalyzed the reverse reaction and showed a clear preference in the polyP-dependent nucleotide-forming direction (ATP over GTP). In addition, the reaction equilibrium could be influenced in favor of polyP synthesis by circumventing ADP inhibition using an ATP recycling system, thus allowing detection of polyP formation by ³¹P NMR. In summary, the first chapter of this thesis presents the discovery of the first crenarchaeal PPK, which we describe as a dTMPK-like heteromeric PPK.

4. Summary

The second part of this thesis addresses the study of myristoylation in *S. acidocaldarius*. To date, myristoylation has been predominantly studied in Eukaryotes. Focusing on Archaea, the myristoylation has rarely been studied in this domain and has only been reported for some members of the Euryarchaeota. However, some arguments are present suggesting that myristoylation plays a role in Archaea, although FAs are not part of their membrane lipids. For instance, some archaeal organisms are able to utilize FAs and the presence of FA metabolism has also been demonstrated in *S. acidocaldarius*. Furthermore, a NAD⁺-dependent deacetylase (Sir2) was described to deacetylate Alba *in vitro*. This class of enzymes was also reported to catalyze the cleavage of long-chain acyl-modifications, including myristoylation. Since *S. acidocaldarius* also exhibits a Sir2 homolog and has been shown to synthesize and degrade FAs, this organism was used to study myristoylation of proteins. To identify the target proteins a bioorthogonal labeling method using a myristic acid alkyne probe was performed. Through cellular processes, the corresponding proteins were labeled *in vivo* and after subsequent cell disruption, the probe could be bound to a fluorophore using copper-catalyzed azide-alkyne cycloaddition (CuAAC). In-gel visualization illustrated the results of this method and enabled direct comparison with the negative control. On a larger scale, cell lysates were clicked with a biotin tag and subsequently enriched via an avidin pulldown. These samples were further used for LC-MS/MS analysis, thus identifying the labeled proteins. Three top hits Saci_1322, Saci_1188 and Saci_1114 were identified by the overlap of four biological replicates. Saci_1322 is a homolog of the studied Alba protein from *S. solfataricus*, whereas Saci_1114 has previously been identified to be a β -ketothiolase involved in FA metabolism. Considering classification of all labeled proteins obtained the predicted functions were related to metabolism and genetic information processing. According to the representation by the three candidates, most proteins are involved in lipid transport and metabolism as well as energy production and conversion. Moreover, the myristoylome data resembles archaeal N ϵ -lysine acetylome data, which leads to the suggestion that lysine myristoylation may be present in Archaea and thus comprises a regulatory function of metabolic and cellular processes. In order to conclude concrete statements about these properties, further analyses would have to follow up on these initial experiments, in which, for example, enzymatic myristoylation will be confirmed *in vitro*. For this, the potential target proteins identified here may be used. In summary, the first archaeal myristoylome data was described providing some initial insights in the presence and function of protein myristoylation in *S. acidocaldarius*.

5 Zusammenfassung

Archaeen wurden in den späten 1970er Jahren als dritte Domäne des Lebens eingeführt. Dennoch veranschaulichen beide Kapitel dieser Arbeit, dass bis heute wesentliche Stoffwechselwege und zelluläre Prozesse von Archaeen noch nicht aufgeklärt sind. Ein wesentlicher Aspekt ist dabei auch die große Varianz der Organismen innerhalb dieser Domäne, die in der Lage sind, sich unterschiedlichsten extremen Lebensräumen anzupassen. In einigen Arbeiten wurde bereits die Funktion von Polyphosphaten (PolyP) demonstriert, die als Reaktion auf eine Vielzahl von Stressfaktoren synthetisiert oder abgebaut werden und so der Zelle helfen, sich an neue Umweltbedingungen anzupassen. Ebenfalls werden diesem Polymer aber auch die Bereitstellung von Energie durch die hoch-energetische Phosphoanhydrid-Bindung zugesprochen, sowie viele weitere Funktionen, wie Phosphatspeicherung, Einfluss auf Zellmotilität, Biofilmbildung und Kupferresistenz.

In Bezug auf den PolyP-Metabolismus, der in Kapitel 3.1 untersucht wurde, ist für Archaeen durch bioinformatische Analysen festgestellt worden, dass viele archaelle Organismen keinen vollständigen Stoffwechselweg aufweisen. Dazu gehört auch das Crenarchaeon *Sulfolobus acidocaldarius*, dem das Enzym fehlt, welches verantwortlich für die PolyP-Synthese ist. In den bisher verfügbaren Forschungsstudien wurde die Anwesenheit des Polymers in *S. acidocaldarius* demonstriert, sowie die Anwesenheit der Exopolyphosphatase (PPX), des Enzyms, welches für die PolyP-Hydrolyse verantwortlich ist. Darauf aufbauend wurde die PPX biochemisch charakterisiert und zeigte eine hohe spezifische Aktivität, wobei langkettige polyP (P_{700}) favorisiert umgesetzt wurden. Mithilfe von ^{31}P -NMR Messungen konnte der enzymatische Abbau von PolyP₄₅ gezeigt werden, wobei das hitzestabile Nebenprodukt Trimetaphosphat (cP_3) entstanden ist. Durch Anwesenheit von Kalium- oder Magnesium-Ionen wird diese Zyklisierung bei hohen Temperaturen begünstigt, sodass dieses Nebenprodukt neben dem enzymatisch freigesetzten Phosphat (P_i) gebildet wurde. Bei der Suche nach der Polyphosphatkinase (PPK) wurde zuerst die Genomorganisation näher angeschaut. Neben dem *ppx* Gen wurden zwei Thymidylatkinase (dTMPK)-Gene (*saci_2019* und *saci_2020*) identifiziert. Da dTMPKs für ihre strukturelle Ähnlichkeit zu den PPKs der Familie 2 bekannt sind, wurden diese zunächst auf ihre enzymatische Funktion untersucht, darin eingeschlossen eine dTMPK, dessen Gen an anderer Stelle im Genom lokalisiert ist (*saci_0893*). Dabei wurde die dTMPK-Funktion von *Saci_0893* nachgewiesen, sowie die PPK-Funktion der übrigen annotierten dTMPKs (*Saci_2019* und *Saci_2020*). Genauere Untersuchungen ergaben, dass *Saci_2019* die aktive Untereinheit darstellte, aber *Saci_2020* die PPK-Aktivität von *Saci_2019* stimulierte, sodass das heteromere Enzym eine deutlich höhere spezifische Aktivität zeigte. Entsprechend der

5. Zusammenfassung

Eigenschaft einer PPK der Familie 2 zeigte das Enzym eine Präferenz in die Nukleotid-bildende Richtung, wobei PolyP verbraucht wird. Zusätzlich konnte durch ein ATP-Recycling System die ADP-Hemmung umgangen werden zugunsten der PolyP-Synthese. Dieses System ermöglichte den Nachweis der PolyP-Bildung mithilfe von ^{31}P -NMR Messungen. Somit präsentiert das erste Kapitel dieser Arbeit eine neuartige PPK, die wir als dTMPK-artige heteromere PPK beschreiben, welche die erste identifizierte und charakterisierte crenarchaeelle PPK darstellt.

Der zweite Teil dieser Arbeit thematisiert die Untersuchung von Protein-Myristoylierung in *S. acidocaldarius*. Bis heute wurde die Modifikation von Proteinen durch Myristoylierung überwiegend in Eukaryonten untersucht, es ist aber auch eine Studie aus Euryarchaeen publiziert, die Myristoylierung beschreibt. Dennoch ist diese Art der Protein-Acylierung in Archaeen kaum untersucht worden. Es sprechen aber einige Argumente dafür, dass die Myristoylierung von Proteinen in Archaeen eine Rolle spielt, obwohl Fettsäuren nicht in ihren Membranlipiden vorkommen. Neben der Möglichkeit von einigen archaellen Organismen Fettsäuren verbrauchen zu können, wurde auch die Anwesenheit eines Fettsäurestoffwechsels in *S. acidocaldarius* nachgewiesen. Außerdem wurde für *S. solfataricus*, eine NAD^+ abhängige Deacetylase (Sir2) beschrieben, welche das Alba Protein deacetylieren konnte. Diese Enzymklasse ist auch dafür bekannt, langkettige Acyl-Modifikationen abspalten zu können, darunter auch Protein-Myristoylierungen. Da auch *S. acidocaldarius* ein Sir2 Homolog aufweist und nachweislich Fettsäuren synthetisieren und abbauen kann, wurde dieser Organismus verwendet, um die Myristoylierung von Proteinen zu untersuchen. Mithilfe einer bioorthogonalen Markierungsmethode durch eine Myristinsäure-Alkyn-Sonde wurden die entsprechenden Proteine *in vivo* durch zelluläre Prozesse markiert. Nach dem anschließenden Zellaufschluss, konnte die Sonde an ein Fluorophor mithilfe Kupfer-katalysierter Azid-Alkyn Cycloaddition (CuAAC) gebunden werden. Die in-Gel-Visualisierung veranschaulichte die Ergebnisse dieser Methode und ermöglichte einen ersten Einblick in die Funktionalität dieser Methode. In einem größeren Maßstab wurden die Zellysate an einen Biotin-Tag gebunden und anschließend mit einem „Avidin-Pulldown“ angereichert. Diese Proben wurden weiter mittels LC-MS/MS Methode analysiert, um somit die markierten Proteine zu identifizieren. Dabei wurden durch die Überlappung der angereicherten Kandidaten drei Proteine Saci_1322, Saci_1188 und Saci_1114 identifiziert. Saci_1322 ist ein Alba Protein, dessen Homolog nachweislich in *S. solfataricus* als Zielprotein für Protein-Acetylierung und -Deacetylierung nachgewiesen wurde. Saci_1114 ist eine bereits charakterisierte β -Ketothiolase, die am Fettsäurestoffwechsel von *S. acidocaldarius* beteiligt ist. Die Klassifizierung aller markierten Proteine ergab, dass die vorhergesagten Funktionen mit dem Stoffwechsel und der Verarbeitung genetischer Informationen zusammenhängen. Dabei sind die meisten identifizierten Proteine am Lipidtransport und –Stoffwechsel sowie

5. Zusammenfassung

an der Energieproduktion und –Umwandlung beteiligt. Darüber hinaus ähneln die Myristoylom-Daten den veröffentlichten Lysin-Acetylom-Daten von Archaeen, was die Vermutung nahelegt, dass Lysin-Myristoylierung in Archaeen vorkommen könnte und somit auch eine mögliche regulatorische Funktion von Stoffwechsel- und zellulären Prozessen umfasst. Um konkrete Aussagen treffen zu können, müssten auf diese Experimente erste Analysen folgen, in denen z.B. die enzymatische Myristoylierung *in vitro* bestätigt wird. Hierfür können die hier identifizierten Proteine verwendet werden. Zusammenfassend lässt sich festhalten, dass die ersten archaellen Myristoylom-Daten beschrieben wurden, die erste Einblicke in das Vorkommen und die Funktion von Protein-Myristoylierung als mögliche Form der Acylierung in *S. acidocaldarius* geben.

6 References

1. Farlow, W.G., *On the nature of the peculiar reddening of salted codfish during the summer season.* 1880.
2. Kocur, M. and W. Hodgkiss, *Taxonomic Status of the Genus Halococcus Schoop.* International Journal of Systematic and Evolutionary Microbiology, 1973. **23**(2): p. 151-156.
3. Woese, C.R. and G.E. Fox, *Phylogenetic structure of the prokaryotic domain: the primary kingdoms.* Proceedings of the National Academy of Sciences, 1977. **74**(11): p. 5088-5090.
4. Woese, C.R., O. Kandler, and M.L. Wheelis, *Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya.* Proceedings of the National Academy of Sciences, 1990. **87**(12): p. 4576-4579.
5. Bang, C. and R.A. Schmitz, *Archaea associated with human surfaces: not to be underestimated.* FEMS Microbiology Reviews, 2015. **39**(5): p. 631-648.
6. Eckburg, P.B., P.W. Lepp, and D.A. Relman, *Archaea and Their Potential Role in Human Disease.* Infection and Immunity, 2003. **71**(2): p. 591-596.
7. Guy, L. and T.J. Ettema, *The archaeal 'TACK' superphylum and the origin of eukaryotes.* Trends in microbiology, 2011. **19**(12): p. 580-587.
8. Rinke, C., et al., *Insights into the phylogeny and coding potential of microbial dark matter.* Nature, 2013. **499**(7459): p. 431-437.
9. Zaremba-Niedzwiedzka, K., et al., *Asgard archaea illuminate the origin of eukaryotic cellular complexity.* Nature, 2017. **541**(7637): p. 353-358.
10. Zaremba-Niedzwiedzka, K., et al., *Asgard archaea illuminate the origin of eukaryotic cellular complexity.* Nature, 2017. **541**(7637): p. 353-358.
11. Spang, A., et al., *Complex archaea that bridge the gap between prokaryotes and eukaryotes.* Nature, 2015. **521**(7551): p. 173-179.
12. Eme, L., et al., *Archaea and the origin of eukaryotes.* Nature Reviews Microbiology, 2017. **15**(12): p. 711-723.
13. Williams, T.A., et al., *Phylogenomics provides robust support for a two-domains tree of life.* Nature Ecology & Evolution, 2020. **4**(1): p. 138-147.
14. Lindås, A.-C. and R. Bernander, *The cell cycle of archaea.* Nature Reviews Microbiology, 2013. **11**(9): p. 627-638.
15. Grohmann, D. and F. Werner, *Recent advances in the understanding of archaeal transcription.* Current opinion in microbiology, 2011. **14**(3): p. 328-334.
16. Wenck, B.R. and T.J. Santangelo, *Archaeal transcription.* Transcription, 2020. **11**(5): p. 199-210.
17. Woese, C.R., L.J. Magrum, and G.E. Fox, *Archaeobacteria.* Journal of Molecular Evolution, 1978. **11**(3): p. 245-252.
18. Albers, S.-V. and B.H. Meyer, *The archaeal cell envelope.* Nature Reviews Microbiology, 2011. **9**(6): p. 414-426.

6. References

19. Jarrell, K.F. and S.-V. Albers, *The archaeellum: an old motility structure with a new name*. Trends in Microbiology, 2012. **20**(7): p. 307-312.
20. Albers, S.-V. and K.F. Jarrell, *The archaeellum: an update on the unique archaeal motility structure*. Trends in microbiology, 2018. **26**(4): p. 351-362.
21. Caforio, A. and A.J. Driessen, *Archaeal phospholipids: Structural properties and biosynthesis*. Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids, 2017. **1862**(11): p. 1325-1339.
22. Chong, P.L.-G., *Archaeobacterial bipolar tetraether lipids: Physico-chemical and membrane properties*. Chemistry and Physics of Lipids, 2010. **163**(3): p. 253-265.
23. Scholte, A., et al., *First Isolation and Structure Elucidation of GDNT- β -Glu – Tetraether Lipid Fragment from Archaeal Sulfolobus Strains*. ChemistryOpen, 2021. **10**(9): p. 889-895.
24. Jensen, S.M., et al., *The effects of temperature and growth phase on the lipidomes of Sulfolobus islandicus and Sulfolobus tokodaii*. Life, 2015. **5**(3): p. 1539-1566.
25. Gabriel, J.L. and P. Lee Gau Chong, *Molecular modeling of archaeobacterial bipolar tetraether lipid membranes*. Chemistry and Physics of Lipids, 2000. **105**(2): p. 193-200.
26. Dibrova, D.V., M.Y. Galperin, and A.Y. Mulkidjanian, *Phylogenomic reconstruction of archaeal fatty acid metabolism*. Environmental microbiology, 2014. **16**(4): p. 907-918.
27. Wang, K., et al., *A TetR-family transcription factor regulates fatty acid metabolism in the archaeal model organism Sulfolobus acidocaldarius*. Nature Communications, 2019. **10**(1): p. 1542.
28. Zhou, X., *Fatty acid metabolism in Sulfolobus acidocaldarius and its potential as platform organism in biotechnology*. 2021.
29. Lewis, A.M., et al., *The biology of thermoacidophilic archaea from the order Sulfolobales*. FEMS Microbiology Reviews, 2021. **45**(4).
30. Esser, D., et al., *Protein phosphorylation and its role in archaeal signal transduction*. FEMS microbiology reviews, 2016. **40**(5): p. 625-647.
31. Giglione, C. and T. Meinel, *Mapping the myristoylome through a complete understanding of protein myristoylation biochemistry*. Progress in Lipid Research, 2022. **85**: p. 101139.
32. Bräsen, C., et al., *Carbohydrate metabolism in Archaea: current insights into unusual enzymes and pathways and their regulation*. Microbiology and Molecular Biology Reviews, 2014. **78**(1): p. 89-175.
33. Thauer, R.K., et al., *Methanogenic archaea: ecologically relevant differences in energy conservation*. Nature Reviews Microbiology, 2008. **6**(8): p. 579-591.
34. Siebers, B. and P. Schönheit, *Unusual pathways and enzymes of central carbohydrate metabolism in Archaea*. Current opinion in microbiology, 2005. **8**(6): p. 695-705.

6. References

35. Verhees, C.H., et al., *ADP-dependent phosphofructokinases in mesophilic and thermophilic methanogenic archaea*. *Journal of Bacteriology*, 2001. **183**(24): p. 7145-7153.
36. Kengen, S.M., et al., *ADP-dependent glucokinase and phosphofructokinase from Pyrococcus furiosus*, in *Methods in enzymology*. 2001, Elsevier. p. 41-53.
37. Siebers, B., H.-P. Klenk, and R. Hensel, *PPI-dependent phosphofructokinase from Thermoproteus tenax, an archaeal descendant of an ancient line in phosphofructokinase evolution*. *Journal of bacteriology*, 1998. **180**(8): p. 2137-2143.
38. Lorentzen, E., et al., *Structural basis of allosteric regulation and substrate specificity of the non-phosphorylating glyceraldehyde 3-phosphate dehydrogenase from Thermoproteus tenax*. *Journal of molecular biology*, 2004. **341**(3): p. 815-828.
39. Reher, M., S. Gebhard, and P. Schönheit, *Glyceraldehyde-3-phosphate ferredoxin oxidoreductase (GAPOR) and nonphosphorylating glyceraldehyde-3-phosphate dehydrogenase (GAPN), key enzymes of the respective modified Embden–Meyerhof pathways in the hyperthermophilic crenarchaeota Pyrobaculum aerophilum and Aeropyrum pernix*. *FEMS microbiology letters*, 2007. **273**(2): p. 196-205.
40. Ettema, T.J.G., et al., *The non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase (GAPN) of Sulfolobus solfataricus: a key-enzyme of the semi-phosphorylative branch of the Entner–Doudoroff pathway*. *Extremophiles*, 2008. **12**(1): p. 75-88.
41. Mitchell, P., *Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism*. *Nature*, 1961. **191**: p. 144-8.
42. Mitchell, P., *Chemiosmotic coupling in oxidative and photosynthetic phosphorylation*. *Biol Rev Camb Philos Soc*, 1966. **41**(3): p. 445-502.
43. Wang, L., et al., *Bioinformatics analysis of metabolism pathways of archaeal energy reserves*. *Scientific reports*, 2019. **9**(1): p. 1034.
44. Seufferheld, M.J., H.M. Alvarez, and M.E.J.A.E.M. Farias, *Role of polyphosphates in microbial adaptation to extreme environments*. 2008. **74**(19): p. 5867-5874.
45. Toso, D.B., et al., *Discovery and Characterization of Iron Sulfide and Polyphosphate Bodies Coexisting in Archaeoglobus fulgidus Cells*. *Archaea*, 2016. **2016**: p. 4706532.
46. Orell, A., et al., *Inorganic polyphosphates in extremophiles and their possible functions*. *Extremophiles*, 2012. **16**(4): p. 573-583.
47. Remonsellez, F., A. Orell, and C.A. Jerez, *Copper tolerance of the thermoacidophilic archaeon Sulfolobus metallicus: possible role of polyphosphate metabolism*. *Microbiology*, 2006. **152**(1): p. 59-66.
48. Rivero, M., et al., *Inorganic Polyphosphate, Exopolyphosphatase, and Pho84-Like Transporters May Be Involved in Copper Resistance in Metallosphaera sedula DSM 5348(T)*. *Archaea*, 2018. **2018**: p. 5251061.

6. References

49. Paula, F.S., et al., *The potential for polyphosphate metabolism in Archaea and anaerobic polyphosphate formation in Methanosarcina mazei*. bioRxiv, 2019: p. 689885.
50. Wang, L., et al., *Distribution patterns of polyphosphate metabolism pathway and its relationships with bacterial durability and virulence*. *Frontiers in microbiology*, 2018. **9**: p. 782.
51. Rao, N.N., M.R. Gómez-García, and A.J.A.r.o.b. Kornberg, *Inorganic polyphosphate: essential for growth and survival*. 2009. **78**: p. 605-647.
52. Neville, N., N. Roberge, and Z. Jia, *Polyphosphate Kinase 2 (PPK2) Enzymes: Structure, Function, and Roles in Bacterial Physiology and Virulence*. *Int J Mol Sci*, 2022. **23**(2).
53. Kumble, K.D., K. Ahn, and A. Kornberg, *Phosphohistidyl active sites in polyphosphate kinase of Escherichia coli*. *Proceedings of the National Academy of Sciences*, 1996. **93**(25): p. 14391-14395.
54. Keppler, M., et al., *Make or break: the thermodynamic equilibrium of polyphosphate kinase-catalysed reactions*. *Beilstein Journal of Organic Chemistry*, 2022. **18**: p. 1278-1288.
55. Leipe, D.D., E.V. Koonin, and L. Aravind, *Evolution and classification of P-loop kinases and related proteins*. *Journal of molecular biology*, 2003. **333**(4): p. 781-815.
56. Nocek, B.P., et al., *Structural Insights into Substrate Selectivity and Activity of Bacterial Polyphosphate Kinases*. *ACS Catalysis*, 2018. **8**(11): p. 10746-10760.
57. Brown, M.R. and A. Kornberg, *Inorganic polyphosphate in the origin and survival of species*. *Proc Natl Acad Sci U S A*, 2004. **101**(46): p. 16085-7.
58. Moreno, S.N. and R. Docampo, *Polyphosphate and its diverse functions in host cells and pathogens*. *PLoS pathogens*, 2013. **9**(5): p. e1003230.
59. Gómez-García, M.R. and A. Kornberg, *Formation of an actin-like filament concurrent with the enzymatic synthesis of inorganic polyphosphate*. *Proceedings of the National Academy of Sciences*, 2004. **101**(45): p. 15876-15880.
60. Boyce, K.J., M. Kretschmer, and J.W. Kronstad, *The vtc4 Gene Influences Polyphosphate Storage, Morphogenesis, and Virulence in the Maize Pathogen Ustilago maydis*. *Eukaryotic Cell*, 2006. **5**(8): p. 1399-1409.
61. Hothorn, M., et al., *Catalytic core of a membrane-associated eukaryotic polyphosphate polymerase*. *Science*, 2009. **324**(5926): p. 513-516.
62. Akiyama, M., E. Crooke, and A. Kornberg, *An exopolyphosphatase of Escherichia coli. The enzyme and its ppx gene in a polyphosphate operon*. *Journal of Biological Chemistry*, 1993. **268**(1): p. 633-639.
63. Aravind, L. and E.V. Koonin, *A novel family of predicted phosphoesterases includes Drosophila prune protein and bacterial RecJ exonuclease*. *Trends in biochemical sciences*, 1998. **23**(1): p. 17-19.
64. Lonetti, A., et al., *Identification of an evolutionarily conserved family of inorganic polyphosphate endopolyphosphatases*. *Journal of Biological Chemistry*, 2011. **286**(37): p. 31966-31974.

6. References

65. Sharfstein, S.T. and J. Keasling, *Polyphosphate metabolism in Escherichia coli*. Annals of the New York Academy of Sciences, 1994. **745**(1): p. 77-91.
66. Ghorbel, S., et al., *Regulation of ppk expression and in vivo function of Ppk in Streptomyces lividans TK24*. Journal of bacteriology, 2006. **188**(17): p. 6269-6276.
67. Schulte, J.E. and M. Goulian, *The Phosphohistidine Phosphatase SixA Targets a Phosphotransferase System*. 2018. **9**(6): p. e01666-18.
68. Tumlirsch, T. and D.J.A.E.M. Jendrossek, *Proteins with CHADs (conserved histidine α -helical domains) are attached to polyphosphate granules in vivo and constitute a novel family of polyphosphate-associated proteins (phosins)*. 2017. **83**(7): p. e03399-16.
69. Lorenzo-Orts, L., et al., *Molecular characterization of CHAD domains as inorganic polyphosphate-binding modules*. 2019. **2**(3).
70. Iyer, L.M. and L. Aravind, *The catalytic domains of thiamine triphosphatase and CyaB-like adenylyl cyclase define a novel superfamily of domains that bind organic phosphates*. BMC genomics, 2002. **3**(1): p. 33-33.
71. Vogt, M.S., et al., *The archaeal triphosphate tunnel metalloenzyme SaTTM defines structural determinants for the diverse activities in the CYTH protein family*. Journal of Biological Chemistry, 2021. **297**(1).
72. Keppetipola, N., R. Jain, and S. Shuman, *Novel triphosphate phosphohydrolase activity of Clostridium thermocellum TTM, a member of the triphosphate tunnel metalloenzyme superfamily*. Journal of Biological Chemistry, 2007. **282**(16): p. 11941-11949.
73. Martinez, J., V. Truffault, and M. Hothorn, *Structural determinants for substrate binding and catalysis in triphosphate tunnel metalloenzymes*. Journal of Biological Chemistry, 2015. **290**(38): p. 23348-23360.
74. Ahn, K. and A. Kornberg, *Polyphosphate kinase from Escherichia coli. Purification and demonstration of a phosphoenzyme intermediate*. Journal of Biological Chemistry, 1990. **265**(20): p. 11734-11739.
75. Kornberg, A., *Inorganic polyphosphate: toward making a forgotten polymer unforgettable*. Journal of bacteriology, 1995. **177**(3): p. 491-496.
76. Yamagata, Y., et al., *Volcanic production of polyphosphates and its relevance to prebiotic evolution*. Nature, 1991. **352**(6335): p. 516-519.
77. Keefe, A.D. and S.L. Miller, *Are polyphosphates or phosphate esters prebiotic reagents?* Journal of molecular evolution, 1995. **41**(6): p. 693-702.
78. Reusch, R., *Transmembrane ion transport by polyphosphate/poly-(R)-3-hydroxybutyrate complexes*. BIOCHEMISTRY C/C OF BIOKHMIIA, 2000. **65**(3): p. 280-295.
79. Green, J., *Precambrian Lunar Volcanic Protolife*. International Journal of Molecular Sciences, 2009. **10**(6): p. 2681-2721.
80. Desfougères, Y., A. Saiardi, and C. Azevedo, *Inorganic polyphosphate in mammals: where's Wally?* Biochem Soc Trans, 2020. **48**(1): p. 95-101.

6. References

81. Jimenez-Nuñez, M.D., et al., *Myeloma cells contain high levels of inorganic polyphosphate which is associated with nucleolar transcription*. *haematologica*, 2012. **97**(8): p. 1264.
82. Azevedo, C. and A. Saiardi, *Functions of inorganic polyphosphates in eukaryotic cells: a coat of many colours*. *Biochemical Society Transactions*, 2014. **42**(1): p. 98-102.
83. Gray, M.J., et al., *Polyphosphate is a primordial chaperone*. *Molecular cell*, 2014. **53**(5): p. 689-699.
84. Gray, M.J. and U. Jakob, *Oxidative stress protection by polyphosphate—new roles for an old player*. *Current opinion in microbiology*, 2015. **24**: p. 1-6.
85. Azevedo, C., T. Livermore, and A. Saiardi, *Protein Polyphosphorylation of Lysine Residues by Inorganic Polyphosphate*. *Molecular Cell*, 2015. **58**(1): p. 71-82.
86. Jasso-Chávez, R., et al., *Air-adapted Methanosarcina acetivorans shows high methane production and develops resistance against oxygen stress*. *PLoS One*, 2015. **10**(2): p. e0117331.
87. Li, L., et al., *Quantitative Proteomic and Microarray Analysis of the Archaeon Methanosarcina acetivorans Grown with Acetate versus Methanol*. *Journal of proteome research*, 2007. **6**(2): p. 759-771.
88. Recalde, A., et al., *The role of polyphosphate in motility, adhesion, and biofilm formation in Sulfolobales*. *Microorganisms*, 2021. **9**(1): p. 193.
89. Xie, L. and U. Jakob, *Inorganic polyphosphate, a multifunctional polyanionic protein scaffold*. *J Biol Chem*, 2019. **294**(6): p. 2180-2190.
90. Lorenzo-Orts, L., D. Couto, and M. Hothorn, *Identity and functions of inorganic and inositol polyphosphates in plants*. *New Phytologist*, 2020. **225**(2): p. 637-652.
91. Brock, T.D., et al., *Sulfolobus: a new genus of sulfur-oxidizing bacteria living at low pH and high temperature*. *Archiv für Mikrobiologie*, 1972. **84**(1): p. 54-68.
92. Quehenberger, J., et al., *Sulfolobus—a potential key organism in future biotechnology*. *Frontiers in microbiology*, 2017. **8**: p. 2474.
93. Wakagi, T. and T. Oshima, *Energy metabolism of a thermoacidophilic archaeobacterium, Sulfolobus acidocaldarius*. *Origins of life and evolution of the biosphere*, 1987. **17**(3): p. 391-399.
94. Nunn, C.E., et al., *Metabolism of pentose sugars in the hyperthermophilic archaea Sulfolobus solfataricus and Sulfolobus acidocaldarius*. *Journal of Biological Chemistry*, 2010. **285**(44): p. 33701-33709.
95. Albers, S.-V. and B. Siebers, *The Family Sulfolobaceae*. 2014, Springer: Berlin, Germany. p. 323-346.
96. Esser, D., et al., *Functional curation of the Sulfolobus solfataricus P2 and S. acidocaldarius 98-3 complete genome sequences*. *Extremophiles*, 2011. **15**(6): p. 711.
97. Wagner, M., et al., *Versatile genetic tool box for the crenarchaeote Sulfolobus acidocaldarius*. *Frontiers in microbiology*, 2012. **3**: p. 214.

6. References

98. Chen, L., et al., *The genome of Sulfolobus acidocaldarius, a model organism of the Crenarchaeota*. Journal of bacteriology, 2005. **187**(14): p. 4992-4999.
99. Skorko, R., J. Osipiuk, and K. Stetter, *Glycogen-bound polyphosphate kinase from the archaeobacterium Sulfolobus acidocaldarius*. Journal of bacteriology, 1989. **171**(9): p. 5162-5164.
100. Cardona, S., et al., *The Glycogen-Bound Polyphosphate Kinase from Sulfolobus acidocaldarius Is Actually a Glycogen Synthase*. Applied and environmental microbiology, 2001. **67**(10): p. 4773-4780.
101. Cardona, S.T., F.P. Chávez, and C.A. Jerez, *The exopolyphosphatase gene from Sulfolobus solfataricus: characterization of the first gene found to be involved in polyphosphate metabolism in archaea*. Appl. Environ. Microbiol., 2002. **68**(10): p. 4812-4819.
102. Walsh, C., *Posttranslational modification of proteins: expanding nature's inventory*. 2006: Roberts and Company Publishers.
103. Kennelly, P.J., *Archaeal protein kinases and protein phosphatases: insights from genomics and biochemistry*. Biochemical Journal, 2003. **370**(2): p. 373-389.
104. Dwek, R.A., *Glycobiology: Toward Understanding the Function of Sugars*. Chemical Reviews, 1996. **96**(2): p. 683-720.
105. Mescher, M.F. and J.L. Strominger, *Structural (shape-maintaining) role of the cell surface glycoprotein of Halobacterium salinarium*. Proceedings of the National Academy of Sciences, 1976. **73**(8): p. 2687-2691.
106. Jarrell, K.F., et al., *N-Linked Glycosylation in Archaea: a Structural, Functional, and Genetic Analysis*. Microbiology and Molecular Biology Reviews, 2014. **78**(2): p. 304-341.
107. Meyer, B.H., et al., *Sulfoquinovose synthase – an important enzyme in the N-glycosylation pathway of Sulfolobus acidocaldarius*. Molecular Microbiology, 2011. **82**(5): p. 1150-1163.
108. Helsen, K., et al., *Bioinformatics analysis of a Saccharomyces cerevisiae N-terminal proteome provides evidence of alternative translation initiation and post-translational N-terminal acetylation*. Journal of proteome research, 2011. **10**(8): p. 3578-3589.
109. Varland, S., C. Osberg, and T. Arnesen, *N-terminal modifications of cellular proteins: The enzymes involved, their substrate specificities and biological effects*. PROTEOMICS, 2015. **15**(14): p. 2385-2401.
110. Vetting, M.W., et al., *Structure and functions of the GNAT superfamily of acetyltransferases*. Archives of biochemistry and biophysics, 2005. **433**(1): p. 212-226.
111. Xu, Y., Z. Shi, and L. Bao, *An Expanding Repertoire of Protein Acylations*. Molecular & Cellular Proteomics, 2022. **21**(3): p. 100193.
112. Cha-Molstad, H., et al., *Amino-terminal arginylation targets endoplasmic reticulum chaperone BiP for autophagy through p62 binding*. Nature cell biology, 2015. **17**(7): p. 917-929.

6. References

113. Foyn, H., et al., *Protein N-terminal acetyltransferases act as N-terminal propionyltransferases in vitro and in vivo*. *Molecular & Cellular Proteomics*, 2013. **12**(1): p. 42-54.
114. Polevoda, B., et al., *Identification and specificities of N-terminal acetyltransferases from Saccharomyces cerevisiae*. *The EMBO journal*, 1999. **18**(21): p. 6155-6168.
115. Soppa, J., *Protein Acetylation in Archaea, Bacteria, and Eukaryotes*. *Archaea*, 2010. **2010**: p. 820681.
116. Gordiyenko, Y., et al., *Acetylation of L12 increases interactions in the Escherichia coli ribosomal stalk complex*. *Journal of molecular biology*, 2008. **380**(2): p. 404-414.
117. VanDrise, C.M. and J.C. Escalante-Semerena, *Protein Acetylation in Bacteria*. *Annual review of microbiology*, 2019. **73**.
118. Aivaliotis, M., et al., *Large-Scale Identification of N-Terminal Peptides in the Halophilic Archaea Halobacterium salinarum and Natronomonas pharaonis*. *Journal of Proteome Research*, 2007. **6**(6): p. 2195-2204.
119. Falb, M., et al., *Archaeal N-terminal protein maturation commonly involves N-terminal acetylation: a large-scale proteomics survey*. *Journal of molecular biology*, 2006. **362**(5): p. 915-924.
120. Chang, Y.-Y. and C.-H. Hsu, *Structural Basis for Substrate-specific Acetylation of N α -acetyltransferase Ard1 from Sulfolobus solfataricus*. *Scientific Reports*, 2015. **5**(1): p. 8673.
121. Mackay, D.T., et al., *An acetylase with relaxed specificity catalyses protein N-terminal acetylation in Sulfolobus solfataricus*. *Molecular Microbiology*, 2007. **64**(6): p. 1540-1548.
122. Hildmann, C., D. Riester, and A. Schwienerhorst, *Histone deacetylases—an important class of cellular regulators with a variety of functions*. *Applied microbiology and biotechnology*, 2007. **75**(3): p. 487-497.
123. Lee, K.K. and J.L. Workman, *Histone acetyltransferase complexes: one size doesn't fit all*. *Nature reviews Molecular cell biology*, 2007. **8**(4): p. 284-295.
124. Kimura, A., K. Matsubara, and M. Horikoshi, *A Decade of Histone Acetylation: Marking Eukaryotic Chromosomes with Specific Codes*. *The Journal of Biochemistry*, 2005. **138**(6): p. 647-662.
125. Avvakumov, N. and J. Côté, *The MYST family of histone acetyltransferases and their intimate links to cancer*. *Oncogene*, 2007. **26**(37): p. 5395-5407.
126. Yamamoto, T. and M. Horikoshi, *Novel substrate specificity of the histone acetyltransferase activity of HIV-1-Tat interactive protein Tip60*. *Journal of Biological Chemistry*, 1997. **272**(49): p. 30595-30598.
127. Lu, L., K.A. Berkey, and R.A. Casero, *RGFGIGS is an amino acid sequence required for acetyl coenzyme A binding and activity of human spermidine/spermine N1acetyltransferase*. *Journal of Biological Chemistry*, 1996. **271**(31): p. 18920-18924.
128. Gregoretto, I., Y.-M. Lee, and H.V. Goodson, *Molecular Evolution of the Histone Deacetylase Family: Functional Implications of Phylogenetic Analysis*. *Journal of Molecular Biology*, 2004. **338**(1): p. 17-31.

6. References

129. Preyat, N. and O. Leo, *Sirtuin deacylases: a molecular link between metabolism and immunity*. Journal of Leukocyte Biology, 2013. **93**(5): p. 669-680.
130. Finkel, T., C.-X. Deng, and R. Mostoslavsky, *Recent progress in the biology and physiology of sirtuins*. Nature, 2009. **460**(7255): p. 587-591.
131. Gardner, J.G., et al., *Control of acetyl-coenzyme A synthetase (AcsA) activity by acetylation/deacetylation without NAD⁺ involvement in Bacillus subtilis*. Journal of bacteriology, 2006. **188**(15): p. 5460-5468.
132. Starai, V., et al., *Sir2-dependent activation of acetyl-CoA synthetase by deacetylation of active lysine*. Science, 2002. **298**(5602): p. 2390-2392.
133. Yan, J., et al., *In Vivo Acetylation of CheY, a Response Regulator in Chemotaxis of Escherichia coli*. Journal of Molecular Biology, 2008. **376**(5): p. 1260-1271.
134. Seidel, J., C. Klockenbusch, and D. Schwarzer, *Investigating Deformylase and Deacylase Activity of Mammalian and Bacterial Sirtuins*. ChemBioChem, 2016. **17**(5): p. 398-402.
135. Hentchel, K.L. and J.C. Escalante-Semerena, *Acylation of Biomolecules in Prokaryotes: a Widespread Strategy for the Control of Biological Function and Metabolic Stress*. Microbiology and Molecular Biology Reviews, 2015. **79**(3): p. 321-346.
136. Altman-Price, N. and M. Mevarech, *Genetic Evidence for the Importance of Protein Acetylation and Protein Deacetylation in the Halophilic Archaeon Haloferax volcanii*. Journal of Bacteriology, 2009. **191**(5): p. 1610-1617.
137. Bell, S.D., et al., *The interaction of Alba, a conserved archaeal chromatin protein, with Sir2 and its regulation by acetylation*. Science, 2002. **296**(5565): p. 148-151.
138. Marsh, V.L., S.Y. Peak-Chew, and S.D. Bell, *Sir2 and the acetyltransferase, Pat, regulate the archaeal chromatin protein, Alba*. Journal of Biological Chemistry, 2005. **280**(22): p. 21122-21128.
139. Cao, J., et al., *Insights into the post-translational modifications of archaeal Sis10b (Alba): lysine-16 is methylated, not acetylated, and this does not regulate transcription or growth*. Molecular microbiology, 2018. **109**(2): p. 192-208.
140. Alpha-Bazin, B., et al., *Lysine-specific acetylated proteome from the archaeon Thermococcus gammatolerans reveals the presence of acetylated histones*. Journal of Proteomics, 2021. **232**: p. 104044.
141. Cao, J., et al., *Functional Insights Into Protein Acetylation in the Hyperthermophilic Archaeon Sulfolobus islandicus*. Molecular & Cellular Proteomics, 2019. **18**(8): p. 1572-1587.
142. Jiang, H., et al., *Protein Lipidation: Occurrence, Mechanisms, Biological Functions, and Enabling Technologies*. Chemical Reviews, 2018. **118**(3): p. 919-988.
143. Pugh, E.L. and M. Kates, *Acylation of proteins of the archaebacteria Halobacterium cutirubrum and Methanobacterium thermoautotrophicum*. Biochimica et Biophysica Acta (BBA) - Biomembranes, 1994. **1196**(1): p. 38-44.
144. Johnson, D.R., et al., *Genetic and biochemical studies of protein N-myristoylation*. Annual review of biochemistry, 1994. **63**(1): p. 869-914.

6. References

145. Dyda, F., D.C. Klein, and A.B. Hickman, *GCN5-Related N-Acetyltransferases: A Structural Overview*. Annual Review of Biophysics and Biomolecular Structure, 2000. **29**(1): p. 81-103.
146. Rudnick, D., et al., *Kinetic and structural evidence for a sequential ordered Bi Bi mechanism of catalysis by Saccharomyces cerevisiae myristoyl-CoA: protein N-myristoyltransferase*. Journal of Biological Chemistry, 1991. **266**(15): p. 9732-9739.
147. Bhatnagar, R.S., et al., *Structure of N-myristoyltransferase with bound myristoylCoA and peptide substrate analogs*. Nature structural biology, 1998. **5**(12): p. 1091-1097.
148. Bhatnagar, R.S., et al., *Titration calorimetric analysis of AcylCoA recognition by myristoylCoA: protein N-myristoyltransferase*. Biochemistry, 1997. **36**(22): p. 6700-6708.
149. Farazi, T.A., G. Waksman, and J.I. Gordon, *Structures of Saccharomyces cerevisiae N-myristoyltransferase with bound myristoylCoA and peptide provide insights about substrate recognition and catalysis*. Biochemistry, 2001. **40**(21): p. 6335-6343.
150. Rampoldi, F., et al., *Immunosuppression and aberrant T cell development in the absence of N-myristoylation*. The Journal of Immunology, 2015. **195**(9): p. 4228-4243.
151. Takasaki, A., et al., *Identification of the Calmodulin-binding Domain of Neuron-specific Protein Kinase C Substrate Protein CAP-22/NAP-22*. Journal of Biological Chemistry, 1999. **274**(17): p. 11848-11853.
152. Nagar, B., et al., *Structural basis for the autoinhibition of c-Abl tyrosine kinase*. Cell, 2003. **112**(6): p. 859-871.
153. DeMar, J.C. and R.E. Anderson, *Identification and Quantitation of the Fatty Acids Composing the CoA Ester Pool of Bovine Retina, Heart, and Liver*. Journal of Biological Chemistry, 1997. **272**(50): p. 31362-31368.
154. Komander, D., *The emerging complexity of protein ubiquitination*. Biochemical Society Transactions, 2009. **37**(5): p. 937-953.
155. Jiang, H., et al., *SIRT6 regulates TNF- α secretion through hydrolysis of long-chain fatty acyl lysine*. Nature, 2013. **496**(7443): p. 110-113.
156. Jiang, H., X. Zhang, and H. Lin, *Lysine fatty acylation promotes lysosomal targeting of TNF- α* . Scientific reports, 2016. **6**(1): p. 1-10.
157. Sharma, S., et al., *Mechanism of LoICDE as a molecular extruder of bacterial triacylated lipoproteins*. Nature Communications, 2021. **12**(1): p. 4687.
158. Issartel, J.-P., V. Koronakis, and C. Hughes, *Activation of Escherichia coli prohaemolysin to the mature toxin by acyl carrier protein-dependent fatty acylation*. Nature, 1991. **351**(6329): p. 759-761.
159. Benz, R., *RTX-Toxins*. Toxins, 2020. **12**(6): p. 359.
160. Hamerly, T., et al., *Characterization of Fatty Acids in Crenarchaeota by GC-MS and NMR*. Archaea, 2015. **2015**.
161. Gochnauer, M. and D. Kushner, *Growth and nutrition of extremely halophilic bacteria*. Canadian Journal of Microbiology, 1969. **15**(10): p. 1157-1165.

6. References

162. Falb, M., et al., *Metabolism of halophilic archaea*. *Extremophiles*, 2008. **12**(2): p. 177-196.
163. Tornøe, C.W., C. Christensen, and M. Meldal, *Peptidotriazoles on Solid Phase: [1,2,3]-Triazoles by Regiospecific Copper(I)-Catalyzed 1,3-Dipolar Cycloadditions of Terminal Alkynes to Azides*. *The Journal of Organic Chemistry*, 2002. **67**(9): p. 3057-3064.
164. Rostovtsev, V.V., et al., *A Stepwise Huisgen Cycloaddition Process: Copper(I)-Catalyzed Regioselective "Ligation" of Azides and Terminal Alkynes*. *Angewandte Chemie International Edition*, 2002. **41**(14): p. 2596-2599.
165. Wright, M.H. and S.A. Sieber, *Chemical proteomics approaches for identifying the cellular targets of natural products*. *Natural Product Reports*, 2016. **33**(5): p. 681-708.
166. Martin, B.R. and B.F. Cravatt, *Large-scale profiling of protein palmitoylation in mammalian cells*. *Nature Methods*, 2009. **6**(2): p. 135-138.
167. Krysiak, J.M., et al., *Activity-Based Probes for Studying the Activity of Flavin-Dependent Oxidases and for the Protein Target Profiling of Monoamine Oxidase Inhibitors*. *Angewandte Chemie International Edition*, 2012. **51**(28): p. 7035-7040.
168. Wright, A.T., J.D. Song, and B.F. Cravatt, *A Suite of Activity-Based Probes for Human Cytochrome P450 Enzymes*. *Journal of the American Chemical Society*, 2009. **131**(30): p. 10692-10700.
169. Wright, A.T. and B.F. Cravatt, *Chemical Proteomic Probes for Profiling Cytochrome P450 Activities and Drug Interactions In Vivo*. *Chemistry & Biology*, 2007. **14**(9): p. 1043-1051.
170. Makarova, K.S., Y.I. Wolf, and E.V. Koonin, *Archaeal Clusters of Orthologous Genes (arCOGs): An Update and Application for Analysis of Shared Features between Thermococcales, Methanococcales, and Methanobacteriales*. *Life (Basel)*, 2015. **5**(1): p. 818-40.
171. Kornberg, A., N.N. Rao, and D. Ault-Riche, *Inorganic polyphosphate: a molecule of many functions*. *Annual review of biochemistry*, 1999. **68**(1): p. 89-125.
172. Liebermann, L., *Über das Nuclein der Hefe und künstliche Darstellung eines Nucleins aus Eiweiss und Metaphosphorsäure*. *Berichte der deutschen chemischen Gesellschaft*, 1888. **21**(1): p. 598-600.
173. Kornberg, A., S. Kornberg, and E.S. Simms, *Metaphosphate synthesis by an enzyme from Escherichia coli*. *Biochimica et biophysica acta*, 1956. **20**: p. 215-227.
174. Bondy-Chorney, E., et al., *A Broad Response to Intracellular Long-Chain Polyphosphate in Human Cells*. *Cell Rep*, 2020. **33**(4): p. 108318.
175. Lichko, L., et al., *Inorganic polyphosphates and exopolyphosphatases in cell compartments of the yeast Saccharomyces cerevisiae under inactivation of PPX1 and PPN1 genes*. *Bioscience reports*, 2006. **26**(1): p. 45-54.
176. Eldarov, M.A., et al., *Polyphosphates and exopolyphosphatase activities in the yeast Saccharomyces cerevisiae under overexpression of homologous and heterologous PPN1 genes*. *Biochemistry (Moscow)*, 2013. **78**(8): p. 946-953.

6. References

177. Trilisenko, L., et al., *Polyphosphates and polyphosphatase activity in the yeast Saccharomyces cerevisiae during overexpression of the DDP1 gene*. Biochemistry (Moscow), 2015. **80**(10): p. 1312-1317.
178. Ryazanova, L.P., et al., *Inorganic Polyphosphate and Physiological Properties of Saccharomyces cerevisiae Yeast Overexpressing Ppn2*. Biochemistry (Mosc), 2020. **85**(4): p. 516-522.
179. Soto, D.F., et al., *Global effect of the lack of inorganic polyphosphate in the extremophilic archaeon Sulfolobus solfataricus: A proteomic approach*. Journal of proteomics, 2019. **191**: p. 143-152.
180. Miyachi, S., et al., *Metabolic roles of inorganic polyphosphates in Chlorella cells*. Biochimica et Biophysica Acta (BBA)-General Subjects, 1964. **93**(3): p. 625-634.
181. Jiménez, J., et al., *Polyphosphate: popping up from oblivion*. Current Genetics, 2017. **63**(1): p. 15-18.
182. Zhang, H., et al., *Polyphosphate kinase 1, a conserved bacterial enzyme, in a eukaryote, Dictyostelium discoideum, with a role in cytokinesis*. Proc Natl Acad Sci U S A, 2007. **104**(42): p. 16486-91.
183. Motomura, K., et al., *A New Subfamily of Polyphosphate Kinase 2 (Class III PPK2) Catalyzes both Nucleoside Monophosphate Phosphorylation and Nucleoside Diphosphate Phosphorylation*. Applied and Environmental Microbiology, 2014. **80**(8): p. 2602-2608.
184. Nocek, B., et al., *Polyphosphate-dependent synthesis of ATP and ADP by the family-2 polyphosphate kinases in bacteria*. Proceedings of the National Academy of Sciences, 2008.
185. Ishige, K., H. Zhang, and A. Kornberg, *Polyphosphate kinase (PPK2), a potent, polyphosphate-driven generator of GTP*. Proceedings of the National Academy of Sciences, 2002. **99**(26): p. 16684-16688.
186. Tammenkoski, M., et al., *Kinetic and mutational analyses of the major cytosolic exopolyphosphatase from Saccharomyces cerevisiae*. Journal of Biological Chemistry, 2007. **282**(13): p. 9302-9311.
187. Keasling, J., L. Bertsch, and A. Kornberg, *Guanosine pentaphosphate phosphohydrolase of Escherichia coli is a long-chain exopolyphosphatase*. Proceedings of the National Academy of Sciences, 1993. **90**(15): p. 7029-7033.
188. Blank, L.M., *The cell and P: from cellular function to biotechnological application*. Current opinion in biotechnology, 2012. **23**(6): p. 846-851.
189. Sethuraman, A., N.N. Rao, and A. Kornberg, *The endopolyphosphatase gene: essential in Saccharomyces cerevisiae*. Proceedings of the National Academy of Sciences, 2001. **98**(15): p. 8542-8547.
190. Sakuraba, H., R. Kawakami, and T. Ohshima, *First archaeal inorganic polyphosphate/ATP-dependent NAD kinase, from hyperthermophilic archaeon Pyrococcus horikoshii: cloning, expression, and characterization*. Applied and environmental microbiology, 2005. **71**(8): p. 4352-4358.
191. Scherer, P. and H.-P. Bochem, *Ultrastructural investigation of 12 Methanosarcinae and related species grown on methanol for occurrence of*

6. References

- polyphosphatelike inclusions*. Canadian journal of microbiology, 1983. **29**(9): p. 1190-1199.
192. Rulliere, C., et al., *Heat treatment effect on polyphosphate chain length in aqueous and calcium solutions*. Food chemistry, 2012. **134**(2): p. 712-716.
 193. Chuang, Y.-M., et al., *Deficiency of the Novel Exopolyphosphatase Rv1026/PPX2 Leads to Metabolic Downshift and Altered Cell Wall Permeability in Mycobacterium tuberculosis*. mBio, 2015. **6**(2): p. e02428-14.
 194. Braun, F., et al., *Putative Nucleotide-Based Second Messengers in the Archaeal Model Organisms Haloferax volcanii and Sulfolobus acidocaldarius*. Frontiers in microbiology, 2021: p. 3604.
 195. Cellini, A., et al., *Stringent control in the archaeal genus Sulfolobus*. Research in microbiology, 2004. **155**(2): p. 98-104.
 196. Martucci, N.M., et al., *The magic spot ppGpp influences in vitro the molecular and functional properties of the elongation factor 1 α from the archaeon Sulfolobus solfataricus*. Extremophiles, 2012. **16**(5): p. 743-749.
 197. Albi, T. and A. Serrano, *Two exopolyphosphatases with distinct molecular architectures and substrate specificities from the thermophilic green-sulfur bacterium Chlorobium tepidum TLS*. Microbiology (Reading), 2014. **160**(Pt 9): p. 2067-2078.
 198. Beassoni, P.R., et al., *Pseudomonas aeruginosa Exopolyphosphatase Is Also a Polyphosphate: ADP Phosphotransferase*. Enzyme Res, 2015. **2015**: p. 404607.
 199. Bolesch, D.G. and J. Keasling, *The effect of monovalent ions on polyphosphate binding to Escherichia coli exopolyphosphatase*. Biochemical and biophysical research communications, 2000. **274**(1): p. 236-241.
 200. Rangarajan, E.S., et al., *The Structure of the Exopolyphosphatase (PPX) from Escherichia coli O157:H7 Suggests a Binding Mode for Long Polyphosphate Chains*. Journal of Molecular Biology, 2006. **359**(5): p. 1249-1260.
 201. Song, H., et al., *Structure and activity of PPX/GppA homologs from Escherichia coli and Helicobacter pylori*. The FEBS Journal, 2020. **287**(9): p. 1865-1885.
 202. Kristensen, O., et al., *Structural Characterization of the Stringent Response Related Exopolyphosphatase/Guanosine Pentaphosphate Phosphohydrolase Protein Family*. Biochemistry, 2004. **43**(28): p. 8894-8900.
 203. Kristensen, O., B. Ross, and M. Gajhede, *Structure of the PPX/GPPA Phosphatase from Aquifex aeolicus in Complex with the Alarmone ppGpp*. Journal of Molecular Biology, 2008. **375**(5): p. 1469-1476.
 204. Meyer, W., et al., *Purification, Cloning, and Sequencing of Archaeobacterial Pyrophosphatase from the Extreme Thermoacidophile Sulfolobus acidocaldarius*. Archives of Biochemistry and Biophysics, 1995. **319**(1): p. 149-156.
 205. Hansen, T., et al., *The Extreme Thermostable Pyrophosphatase from Sulfolobus acidocaldarius: Enzymatic and Comparative Biophysical Characterization*. Archives of Biochemistry and Biophysics, 1999. **363**(1): p. 135-147.

6. References

206. Oliver, E.B., et al., *Characterization of an archaeal inorganic pyrophosphatase from Sulfolobus islandicus using a [31P]-NMR-based assay*. Biochemical and Biophysical Research Communications, 2021. **585**: p. 8-14.
207. Gardner, S.G. and W.R. McCleary, *Control of the phoBR Regulon in Escherichia coli*. EcoSal Plus, 2019. **8**(2).
208. Hsieh, Y.-J. and B.L. Wanner, *Global regulation by the seven-component Pi signaling system*. Current opinion in microbiology, 2010. **13**(2): p. 198-203.
209. Santos-Beneit, F., *The Pho regulon: a huge regulatory network in bacteria*. Frontiers in microbiology, 2015. **6**: p. 402.
210. Mouillon, J.-M. and B.L. Persson, *New aspects on phosphate sensing and signalling in Saccharomyces cerevisiae*. FEMS yeast research, 2006. **6**(2): p. 171-176.
211. Secco, D., et al., *Phosphate homeostasis in the yeast Saccharomyces cerevisiae, the key role of the SPX domain-containing proteins*. FEBS letters, 2012. **586**(4): p. 289-295.
212. Pedersen, B.P., et al., *Crystal structure of a eukaryotic phosphate transporter*. Nature, 2013. **496**(7446): p. 533-536.
213. Orell, A., et al., *Life in blue: copper resistance mechanisms of bacteria and archaea used in industrial biomining of minerals*. Biotechnology advances, 2010. **28**(6): p. 839-848.
214. Auernik, K.S., et al., *The genome sequence of the metal-mobilizing, extremely thermoacidophilic archaeon Metallosphaera sedula provides insights into bioleaching-associated metabolism*. Applied and environmental microbiology, 2008. **74**(3): p. 682-692.
215. Neville, N., et al., *A Dual-Specificity Inhibitor Targets Polyphosphate Kinase 1 and 2 Enzymes To Attenuate Virulence of Pseudomonas aeruginosa*. mBio, 2021. **12**(3): p. e00592-21.
216. Itoh, H. and T. Shiba, *Polyphosphate Synthetic Activity of Polyphosphate:AMP Phosphotransferase in Acinetobacter johnsonii 210A*. Journal of Bacteriology, 2004. **186**(15): p. 5178-5181.
217. Lindner, S.N., et al., *NCgl2620 encodes a class II polyphosphate kinase in Corynebacterium glutamicum*. Applied and environmental microbiology, 2007. **73**(15): p. 5026-5033.
218. Hildenbrand, J.C., A. Teleki, and D. Jendrossek, *A universal polyphosphate kinase: PPK2c of Ralstonia eutropha accepts purine and pyrimidine nucleotides including uridine diphosphate*. Applied microbiology and biotechnology, 2020. **104**(15): p. 6659-6667.
219. Shiba, T., et al., *Polyphosphate:AMP phosphotransferase as a polyphosphate-dependent nucleoside monophosphate kinase in Acinetobacter johnsonii 210A*. J Bacteriol, 2005. **187**(5): p. 1859-65.
220. Parnell, A.E., et al., *Substrate recognition and mechanism revealed by ligand-bound polyphosphate kinase 2 structures*. Proc Natl Acad Sci U S A, 2018. **115**(13): p. 3350-3355.

6. References

221. Li de la Sierra, I., et al., *Crystallization and preliminary X-ray analysis of the thymidylate kinase from Mycobacterium tuberculosis*. Acta Crystallogr D Biol Crystallogr, 2000. **56**(Pt 2): p. 226-8.
222. Chaudhary, S.K., J. Jeyakanthan, and K. Sekar, *Structural and functional roles of dynamically correlated residues in thymidylate kinase*. Acta Crystallographica Section D: Structural Biology, 2018. **74**(4): p. 341-354.
223. Lavie, A., et al., *Structure of thymidylate kinase reveals the cause behind the limiting step in AZT activation*. Nat Struct Biol, 1997. **4**(8): p. 601-4.
224. Lavie, A., et al., *Crystal structure of yeast thymidylate kinase complexed with the bisubstrate inhibitor P1-(5'-adenosyl) P5-(5'-thymidyl) pentaphosphate (TP5A) at 2.0 Å resolution: implications for catalysis and AZT activation*. Biochemistry, 1998. **37**(11): p. 3677-86.
225. Kotaka, M., et al., *Structures of S. aureus thymidylate kinase reveal an atypical active site configuration and an intermediate conformational state upon substrate binding*. Protein Sci, 2006. **15**(4): p. 774-84.
226. Andexer, J.N. and M. Richter, *Emerging enzymes for ATP regeneration in biocatalytic processes*. ChemBioChem, 2015. **16**(3): p. 380-386.
227. Andreeva, N., et al., *Ppn2 endopolyphosphatase overexpressed in Saccharomyces cerevisiae: Comparison with Ppn1, Ppx1, and Ddp1 polyphosphatases*. 2019. **163**: p. 101-107.
228. Biswas, A., et al., *Crystal structures of an archaeal thymidylate kinase from Sulfolobus tokodaii provide insights into the role of a conserved active site Arginine residue*. Journal of structural biology, 2017. **197**(3): p. 236-249.
229. Keasling, J. and G.A. Hupf, *Genetic manipulation of polyphosphate metabolism affects cadmium tolerance in Escherichia coli*. Applied and environmental microbiology, 1996. **62**(2): p. 743-746.
230. Haferkamp, P., et al., *The carbon switch at the level of pyruvate and phosphoenolpyruvate in Sulfolobus solfataricus P2*. Frontiers in microbiology, 2019. **10**: p. 757.
231. Ogawa, N., et al., *Inorganic polyphosphate in Vibrio cholerae: genetic, biochemical, and physiologic features*. Journal of Bacteriology, 2000. **182**(23): p. 6687-6693.
232. Zhang, A., et al., *The structure of exopolyphosphatase (PPX) from Porphyromonas gingivalis in complex with substrate analogs and magnesium ions reveals the basis for polyphosphate processivity*. Journal of Structural Biology, 2021. **213**(3): p. 107767.
233. Whittingham, Jean L., et al., *Structural basis for the efficient phosphorylation of AZT-MP (3'-azido-3'-deoxythymidine monophosphate) and dGMP by Plasmodium falciparum type I thymidylate kinase*. Biochemical Journal, 2010. **428**(3): p. 499-509.
234. Ostermann, N., et al., *Potentiating AZT activation: structures of wild-type and mutant human thymidylate kinase suggest reasons for the mutants' improved kinetics with the HIV prodrug metabolite AZTMP*. J Mol Biol, 2000. **304**(1): p. 43-53.

6. References

235. Ostermann, N., et al., *Insights into the phosphoryltransfer mechanism of human thymidylate kinase gained from crystal structures of enzyme complexes along the reaction coordinate*. Structure, 2000. **8**(6): p. 629-42.
236. Leibly, D.J., et al., *Structure of thymidylate kinase from Ehrlichia chaffeensis*. Acta Crystallogr Sect F Struct Biol Cryst Commun, 2011. **67**(Pt 9): p. 1090-4.
237. Biswas, A., et al., *Structural studies of a hyperthermophilic thymidylate kinase enzyme reveal conformational substates along the reaction coordinate*. Febs j, 2017. **284**(15): p. 2527-2544.
238. Choi, M.Y., et al., *The two PPX-GppA homologues from Mycobacterium tuberculosis have distinct biochemical activities*. PLoS One, 2012. **7**(8): p. e42561.
239. Zhang, H., K. Ishige, and A. Kornberg, *A polyphosphate kinase (PPK2) widely conserved in bacteria*. Proceedings of the National Academy of Sciences, 2002. **99**(26): p. 16678-16683.
240. Boucher, Y., M. Kamekura, and W.F. Doolittle, *Origins and evolution of isoprenoid lipid biosynthesis in archaea*. Molecular microbiology, 2004. **52**(2): p. 515-527.
241. Khelifi, N., et al., *Anaerobic Oxidation of Fatty Acids and Alkenes by the Hyperthermophilic Sulfate-Reducing Archaeon Archaeoglobus fulgidus*. Applied and Environmental Microbiology, 2010. **76**(9): p. 3057-3060.
242. Slobodkina, G.B., et al., *Geoglobus acetivorans sp. nov., an iron(III)-reducing archaeon from a deep-sea hydrothermal vent*. International Journal of Systematic and Evolutionary Microbiology, 2009. **59**(11): p. 2880-2883.
243. Peng, C., et al., *The First Identification of Lysine Malonylation Substrates and Its Regulatory Enzyme*. Molecular & Cellular Proteomics, 2011. **10**(12).
244. Adam, R.M., et al., *Cholesterol Sensitivity of Endogenous and Myristoylated Akt*. Cancer Research, 2007. **67**(13): p. 6238-6246.
245. Hu, T., et al., *Myristoylated Naked2 antagonizes Wnt- β -catenin activity by degrading Dishevelled-1 at the plasma membrane*. Journal of Biological Chemistry, 2010. **285**(18): p. 13561-13568.
246. Wright, M.H., et al., *Protein myristoylation in health and disease*. Journal of Chemical Biology, 2010. **3**(1): p. 19-35.
247. Tan, Y.W., W.J. Hong, and J.J.H. Chu, *Inhibition of enterovirus VP4 myristoylation is a potential antiviral strategy for hand, foot and mouth disease*. Antiviral Research, 2016. **133**: p. 191-195.
248. Schwertassek, U., et al., *Myristoylation of the dual-specificity phosphatase c-JUN N-terminal kinase (JNK) stimulatory phosphatase 1 is necessary for its activation of JNK signaling and apoptosis*. The FEBS Journal, 2010. **277**(11): p. 2463-2473.
249. Rudnick, D.A., et al., *Structural and functional studies of Saccharomyces cerevisiae myristoyl-CoA:protein N-myristoyltransferase produced in Escherichia coli. Evidence for an acyl-enzyme intermediate*. Journal of Biological Chemistry, 1990. **265**(22): p. 13370-13378.

6. References

250. Raju, R.V., et al., *Molecular Cloning and Biochemical Characterization of Bovine Spleen Myristoyl CoA: Protein N-Myristoyltransferase*. Archives of biochemistry and biophysics, 1997. **348**(1): p. 134-142.
251. Qi, Q., et al., *Molecular cloning, genomic organization, and biochemical characterization of myristoyl-CoA: Protein N-Myristoyltransferase from Arabidopsis thaliana*. Journal of Biological Chemistry, 2000. **275**(13): p. 9673-9683.
252. Ringel, A.E., C. Roman, and C. Wolberger, *Alternate deacylating specificities of the archaeal sirtuins Sir2Af1 and Sir2Af2*. Protein Science, 2014. **23**(12): p. 1686-1697.
253. Brent, M.M., et al., *Structure and biochemical characterization of protein acetyltransferase from Sulfolobus solfataricus*. Journal of Biological Chemistry, 2009. **284**(29): p. 19412-19419.
254. Cajili, M.K.M. and E.I. Prieto, *Interplay between Alba and Cren7 Regulates Chromatin Compaction in Sulfolobus solfataricus*. Biomolecules, 2022. **12**(4): p. 481.
255. Chang, Y.-Y., S. Hagawa, and C.-H. Hsu, *Adaptation of thermophilic acetyltransferase to a water-mediated catalytic mechanism*. Chemical Communications, 2020. **56**(72): p. 10537-10540.
256. Lodge, J.K., et al., *N-Myristoylation of Arf proteins in Candida albicans: an in vivo assay for evaluating antifungal inhibitors of myristoyl-CoA:protein N-myristoyltransferase*. Microbiology, 1997. **143**(2): p. 357-366.
257. Duronio, R.J., et al., *Protein N-myristoylation in Escherichia coli: reconstitution of a eukaryotic protein modification in bacteria*. Proc Natl Acad Sci U S A, 1990. **87**(4): p. 1506-10.
258. Eisenhaber, B., et al., *Prediction of sequence signals for lipid post-translational modifications: Insights from case studies*. PROTEOMICS, 2004. **4**(6): p. 1614-1625.
259. Maurer-Stroh, S., et al., *MYRbase: analysis of genome-wide glycine myristoylation enlarges the functional spectrum of eukaryotic myristoylated proteins*. Genome Biology, 2004. **5**(3): p. R21.
260. Sakurai, N., et al., *Detection of co- and posttranslational protein N-myristoylation by metabolic labeling in an insect cell-free protein synthesis system*. Analytical Biochemistry, 2007. **362**(2): p. 236-244.
261. Wright, M.H., et al., *Validation of N-myristoyltransferase as an antimalarial drug target using an integrated chemical biology approach*. Nature Chemistry, 2014. **6**(2): p. 112-121.
262. Kallemeijn, W.W., et al., *Proteome-wide analysis of protein lipidation using chemical probes: in-gel fluorescence visualization, identification and quantification of N-myristoylation, N- and S-acylation, O-cholesterylation, S-farnesylation and S-geranylgeranylation*. Nature Protocols, 2021. **16**(11): p. 5083-5122.
263. Lin, V.S., *Interrogating Plant-Microbe Interactions with Chemical Tools: Click Chemistry Reagents for Metabolic Labeling and Activity-Based Probes*. Molecules, 2021. **26**(1): p. 243.

6. References

264. Charron, G., et al., *Robust Fluorescent Detection of Protein Fatty-Acylation with Chemical Reporters*. Journal of the American Chemical Society, 2009. **131**(13): p. 4967-4975.
265. Broncel, M., et al., *Multifunctional Reagents for Quantitative Proteome-Wide Analysis of Protein Modification in Human Cells and Dynamic Profiling of Protein Lipidation During Vertebrate Development*. Angewandte Chemie International Edition, 2015. **54**(20): p. 5948-5951.
266. Wright, M.H., et al., *Global Profiling and Inhibition of Protein Lipidation in Vector and Host Stages of the Sleeping Sickness Parasite Trypanosoma brucei*. ACS Infectious Diseases, 2016. **2**(6): p. 427-441.
267. Majeran, W., et al., *Targeted Profiling of Arabidopsis thaliana Subproteomes Illuminates Co- and Posttranslationally N-Terminal Myristoylated Proteins*. The Plant Cell, 2018. **30**(3): p. 543-562.
268. Gupta, S.D. and H.C. Wu, *Identification and subcellular localization of apolipoprotein N-acyltransferase in Escherichia coli*. FEMS Microbiology Letters, 1991. **78**(1): p. 37-41.
269. Wiktor, M., et al., *Structural insights into the mechanism of the membrane integral N-acyltransferase step in bacterial lipoprotein synthesis*. Nature Communications, 2017. **8**(1): p. 15952.
270. Buddelmeijer, N., *The molecular mechanism of bacterial lipoprotein modification—How, when and why?* FEMS Microbiology Reviews, 2015. **39**(2): p. 246-261.
271. Charlton, T.M., et al., *Quantitative lipoproteomics in Clostridium difficile reveals a role for lipoproteins in sporulation*. Chemistry & Biology, 2015. **22**(11): p. 1562-1573.
272. Martin, D.D.O., et al., *Rapid detection, discovery, and identification of post-translationally myristoylated proteins during apoptosis using a bio-orthogonal azidomyristate analog*. The FASEB Journal, 2008. **22**(3): p. 797-806.
273. Agari, Y., et al., *TetR-family transcriptional repressor Thermus thermophilus FadR controls fatty acid degradation*. Microbiology, 2011. **157**(6): p. 1589-1601.
274. Fujihashi, M., et al., *Structural characterization of a ligand-bound form of Bacillus subtilis FadR involved in the regulation of fatty acid degradation*. Proteins: Structure, Function, and Bioinformatics, 2014. **82**(7): p. 1301-1310.
275. Fujita, Y., H. Matsuoka, and K. Hirooka, *Regulation of fatty acid metabolism in bacteria*. Molecular Microbiology, 2007. **66**(4): p. 829-839.
276. Chang, J.-H., et al., *Structural basis for the NAD-dependent deacetylase mechanism of Sir2*. Journal of Biological Chemistry, 2002. **277**(37): p. 34489-34498.
277. Brown, A.J. and F. Snyder, *Alkyldihydroxyacetone-P synthase. Solubilization, partial purification, new assay method, and evidence for a ping-pong mechanism*. Journal of Biological Chemistry, 1982. **257**(15): p. 8835-8839.
278. Ridgway, N. and R. McLeod, *Biochemistry of lipids, lipoproteins and membranes*. 2008: Elsevier.

6. References

279. Christensen, D.G., et al., *Post-translational protein acetylation: an elegant mechanism for bacteria to dynamically regulate metabolic functions*. *Frontiers in Microbiology*, 2019. **10**: p. 1604.
280. Liu, J., et al., *Systematic Analysis of Lysine Acetylation in the Halophilic Archaeon *Haloferax mediterranei**. *Journal of Proteome Research*, 2017. **16**(9): p. 3229-3241.
281. Meinnel, T., C. Dian, and C. Giglione, *Myristoylation, an ancient protein modification mirroring eukaryogenesis and evolution*. *Trends in biochemical sciences*, 2020. **45**(7): p. 619-632.
282. Rappsilber, J., M. Mann, and Y. Ishihama, *Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips*. *Nature Protocols*, 2007. **2**(8): p. 1896-1906.
283. Olsen, J.V., et al., *Parts per Million Mass Accuracy on an Orbitrap Mass Spectrometer via Lock Mass Injection into a C-trap*. *Molecular & Cellular Proteomics*, 2005. **4**(12): p. 2010-2021.
284. Cox, J., et al., *Andromeda: A Peptide Search Engine Integrated into the MaxQuant Environment*. *Journal of Proteome Research*, 2011. **10**(4): p. 1794-1805.
285. Cox, J. and M. Mann, *MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification*. *Nature Biotechnology*, 2008. **26**(12): p. 1367-1372.
286. Cox, J., et al., *Accurate Proteome-wide Label-free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed MaxLFQ**. *Molecular & Cellular Proteomics*, 2014. **13**(9): p. 2513-2526.
287. Klaus, T., et al., *Activity-based protein profiling for the identification of novel CAZymes involved in xylan degradation in the hyperthermophilic Euryarchaeon *Thermococcus* sp. strain 2319x1E*. *Frontiers in microbiology*, 2021: p. 3868.
288. Tyanova, S., T. Temu, and J. Cox, *The MaxQuant computational platform for mass spectrometry-based shotgun proteomics*. *Nature Protocols*, 2016. **11**(12): p. 2301-2319.

I. Acknowledgements

Lastly, I would like to thank everyone who made this possible.

My biggest thanks go to my PhD mother Prof. Dr. Bettina Siebers, who gave me the opportunity to work in her lab, as well as for her helpful support and kind motivation. I am very grateful for the many opportunities to conduct my scientific work and for the experiences, I was able to take with me along the way.

My special thanks go to Dr. Christopher Bräsen, who always provided me with scientific advice and helped me to analyze experimental problems. Thanks for the kindness, trust and support you have given me.

I would like to thank Prof. Dr. Markus Kaiser for being my co-referee as well as co-supervisor of the Lipid Divide project. Additionally, for the opportunity to work in his laboratory in a nice atmosphere.

I would especially like to thank my colleagues for their day-to-day support. For always being there for me, laughing with me, cursing with me and always having an open ear for me! Even though certainly my only permanent office partner quickly recognized my equal figure in Greek mythology. Thank you, Christian Schmerling, for your ever-faithful advice and your scientific and emotional support at all times. I would also like to thank Laura Kuschmierz for her always kind and encouraging words, they have always motivated me and given me strength. Also I would like to thank Thomas Klaus for his unique humor and cheering up the group at all times. A big thank you also goes to my other office partners Carmen Peraglie and Carsten Schroeder for the fun and never boring time in the office. And again: Thank you Carmen for seeing many things the way I do, we women just have to support each other! I would also like to thank Dr. Xiaoxiao Zhou, Promise Akua Dziwornu, Larissa Schocke, Dr. Sabine Krevet, Dr. Christina Stracke, Dr. Benjamin Meyer, Dr. Lu Shen and Ravi Shankar Ojha, it has been a pleasure to work with you.

For the technical assistance and helpful instructions, I would like to thank Thomas Knura and Sabine Dietl. You have helped me many times and taken work off my hands whenever I needed it.

Thanks to my students Svenja Parge, Reinhold Nazareus, Jordi Weisz, Vaishnavi Shrirang Wagh, Aline-Kathrin Andert and of course also my sister Lea for their help and greatful work.

My special thanks go to Dr. Sabrina Ninck for the nice and friendly collaboration and nice time together in the lab.

I. Acknowledgements

Special thanks to Dr. Torsten Schaller and Dr. Felix Niemeyer for their great help with the NMR analysis. Only with your help could we gain such a good understanding of the story!

My great gratitude also goes to my family primary to my parents, who supported me and were always by my side. Thanks especially to my siblings who were always there for me, I am very grateful for our deep connection.

I would also like to thank my friends for the great support and fun times, and the stable community for the enjoyable evenings. In the course of this, I would also like to thank my four-legged friends, who always distract me and have given me a lot of joy all the time. And here you have now made it in after all, maybe not quite deserved, but anyway... thank you Laurèl for always bringing me to other thoughts and motivating me anew.

II. Curriculum vitae

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

III. Verfassungserklärung

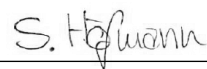
Hiermit erkläre ich, dass ich die vorliegende Arbeit:

Polyphosphate metabolism and profiling of myristoylation in *Sulfolobus acidocaldarius*

ohne fremde Hilfe und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus fremden Quellen (einschließlich elektronischer Quellen) direkt oder indirekt übernommenen Gedanken sind ausnahmslos als solche kenntlich gemacht. Ich bestätige, dass die eingereichten Arbeiten meine eigenen sind, es sei denn, die Arbeiten waren ein Teil von gemeinsam verfassten Veröffentlichungen.

Ich versichere außerdem, dass ich die vorliegende Dissertation nur in diesem und keinem anderen Promotionsverfahren eingereicht habe und, dass diesem Promotionsverfahren keine endgültig gescheiterten Promotionsverfahren vorausgegangen sind.

Essen, den 31.10.2022



Ort, Datum

(Svenja Höfmann)

IV. List of abbreviations

≡MyA	Myristic acid alkyne
°C	Degrees Celsius
³¹ P-NMR	Phosphorus NMR
A	Ampere
aa	Amino acid
ACAD	Acyl-CoA dehydrogenase
ACE	Experiment number by Analytics Core Facility Essen
ACN	Acetonitrile
ACS	Acetyl-CoA synthetase
ADP	Adenosine diphosphate
Alba	Acetylation lowers binding affinity
alkyl DHAP synthase	Alkyldihydroxyacetonephosphate synthase
AMP	Adenosine monophosphate
APS	Ammonium Persulfate
arCOG	Archaeal Clusters of Orthologous Genes
ATP	Adenosine triphosphate
BLAST	Basic Local Alignment Search Tool
bp	base pairs
CoA	Coenzyme A
Cy3	Cyanine3
Da	Dalton
DMAO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
dTDP	Deoxythymidine diphosphate
dTMP	Deoxythymidine monophosphate
dTMPK	thymidylate kinase
dTNP	Deoxynucleotide triphosphates
DTT	Dithiothreitol
E	Extinction coefficient
e.g.	for example
ED	Entner-Doudoroff pathway
EMP	Embden-Meyerhof-Parnas pathway
EtOH	Ethanol
FA	Fatty acid
FA	Formic acid
FAD	Flavin adenine dinucleotide
g	grams
G6P	Glucose-6-phosphate
G6PDH	Glucose-6-phosphate dehydrogenase
GDP	Guanosine diphosphate
GF	Gelfiltration
GMP	Guanosine monophosphate
GNAT	Gcn5-related N-acetyltransferase
GTP	Guanosine triphosphate
HDAC	Histone deacetylase
HEPES	4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid
HK	Hexokinase
HP	Heat precipitation
Hz	hertz
IAM	iodoacetamide
IMAC	Immobilized metal ion affinity chromatography
IPTG	Isopropyl-β-D-thiogalactopyranoside
Kan	kanamycin
L	Liter
λ	Lambda, wavelength
LB	Luria-Bertani
LDH	Lactate dehydrogenase
μ	micro (10 ⁻⁶)

IV. List of abbreviations

M	Molarity
min	Minute
MOPS	3-(<i>N</i> -morpholino)propanesulfonic acid
MS	mass spectrometry
MW	molecular weight
Myr	Myristic acid
NAD ⁺	Nicotinamide adenine dinucleotide (oxidized)
NADH	Nicotinamide adenine dinucleotide (reduced)
NADP ⁺	Nicotinamide adenine dinucleotide phosphate (oxidized)
NADPH	Nicotinamide adenine dinucleotide phosphate (reduced)
NAT	N-terminal acetyltransferase
NCBI	National Centre for Biotechnology Information
NDP	Nucleoside diphosphate
NMR	Nuclear magnetic resonance
nt	nucleotides
NTP	Nucleoside triphosphate
OD	optical density
ORF	open reading frame
PAGE	Polyacrylamide gel electrophoresis
PAP	Polyphosphate:AMP phosphotransferase
PAT	Protein acetyltransferase
PCR	Polymerase chain reaction
PEP	Phosphoenolpyruvate
pH	negative logarithm of the hydrogen ion concentration
PHA	Polyhydroxyalkanoates
P _i	Inorganic phosphate
PK	Pyruvate kinase
polyP	polyphosphate
ppGpp	Guanosine-5'diphosphate-3'-diphosphate
PP _i	Pyrophosphate
PPK	Polyphosphate Kinase
PPN	Endopolyphosphatase
pppGpp	Guanosine-5'triphosphate-3'-diphosphate
PPX	Exopolyphosphatase
Psi	Pounds per square inch
<i>p</i> -value	probability value in null-hypothesis significance testing
RNA	Ribonucleic acid
rpm	revolutions per minute
rRNA	ribosomal ribonucleic acid
RT	Room temperature
Saci	<i>Sulfolobus acidocaldarius</i>
SDS	Sodium dodecyl sulfate
sec	seconds
Sir	Sitruin
Sso	<i>Saccharolobus solfataricus</i>
TB	Terrific Broth
TBTA	Tris((1-benzyl-4-triazolyl)methyl)amine
TCEP	Tris(2-carboxyethyl)phosphine
TED	Tris(carboxymethyl)ethylene diamine
TEMED	<i>N,N,N',N'</i> -Tetramethylethane-1,2-diamine
UV	Ultraviolet
V	Volt
v/v	volume per volume
w/v	weight per volume
x <i>g</i>	Earth acceleration / gravity