

Thermococcus kodakarensis:
the key to affordable biohydrogen
production

Sebastiaan K. Spaans

Thesis committee

Promotor

Prof. Dr J. van der Oost

Personal chair at the Laboratory of Microbiology

Wageningen University

Co-promotor

Dr S.W.M. Kengen

Assistant professor, Laboratory of Microbiology

Wageningen University

Other members

Prof. Dr G. Eggink, Wageningen University

Prof. Dr W.R. Hagen, Delft University of Technology

Dr E.W.J. van Niel, Lund University, Sweden

Dr W. Kuit, MicCell Bioservices BV, Didam

This research was conducted under the auspices of the Graduate School VLAG (Advanced studies in Food Technology, Agrobiotechnology, Nutrition and Health Sciences).

Thermococcus kodakarensis:
the key to affordable biohydrogen
production

Sebastiaan K. Spaans

Thesis

submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. Dr A.P.J. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Thursday 9 June 2016
at 11 a.m. in the Aula.

Sebastiaan K. Spaans

Thermococcus kodakarensis: the key to affordable hydrogen production
246 pages.

PhD thesis, Wageningen University, Wageningen, NL (2016)

With references, with summary in English

ISBN 978-94-6257-772-5

Table of contents

Chapter 1 General introduction and thesis outline	7
Chapter 2 NADPH-generating systems in bacteria and archaea	29
Chapter 3 Introduction of the pentose phosphate pathway in <i>Thermococcus kodakarensis</i> improves growth and hydrogen formation	67
Chapter 4 The chromosome copy number of the hyperthermophilic archaeon <i>Thermococcus kodakarensis</i> KOD1	135
Chapter 5 Identification of <i>Thermococcus kodakarensis</i> genes involved in GDGT synthesis	155
Chapter 6 Thesis summary and General Discussion	195
Appendices	
References	221
Acknowledgements - Dankwoord	240
About the author	244
Overview of completed training activities	245

Chapter 1

General introduction
and thesis outline

A brief history about the importance of energy

Energy is the ability to do work. Ever since the existence of mankind, humans have searched for ways of harvesting energy and putting it to work for them. The more we succeeded, the more time and power became available to us to alter the environment around us, eventually enabling growth and development into the complex society of today. In other words, there is a matching rise in energy use and cultural development or complexity (**Figure 1.1**), or as the anthropologist Leslie White once said: “Energy = civilization” [1]. For example, up until the discovery of fire, human civilization was very basic and the population was very small as the only source of energy available was human muscle (actually metabolism). The discovery of fire and the burning of wood, however, provided an extra source of available energy enabling things like cooking and the development of metal tools. With domestication, mankind also learned to harvest the power of animals to plough their fields and increase crop yields. Harvesting of wind and water energy, through the development of sails, windmills and water-wheels, allowed further population growth and cultural complexity [2].

Together, these sources of energy remained the most powerful and reliable means to utilize energy for thousands of years. However, since the energy harvested by these methods is only relatively low, societies were restricted in size and complexity. For example, the development of the Roman Empire started with their ability to efficiently organize and process relatively big amounts of energy, which led to a surplus of available energy (such as food or man- and animal-power). Consumption of this surplus made it possible to develop into a complex society with a growing non-producing population and a complex system of social facilities. However, to sustain this complex society and ever growing population, huge amounts of energy were required. At first, these requirements could be met by expanding their Empire to obtain new farmland, natural resources and man- and animal-power. However, as energy requirements kept increasing, while expansion became ever more difficult, their energy reserves and systems slowly became insufficient. As no alternatives were available, this led to even more pressure on existing energy systems, which led to even faster exhaustion and eventually social unrest and the collapse of their society [3].

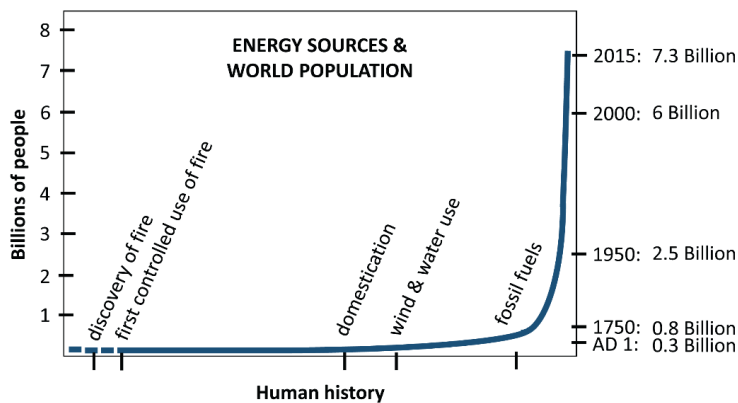


Figure 1.1 Overview of historical energy sources and their effect on the world population

The industrialized civilization of today has reached an unprecedented (global) complexity and the world population has grown far beyond numbers ever seen before. So how is this possible? The answer is: fossil fuels [4, 5]. For thousands of years energy use has been relatively low and the world population has only gradually increased to ~800 million in 1750 (**Figure 1.1**). At the beginning of the industrial revolution, however, Europeans began to unlock the potential of fossil fuels (coal, natural gas and oil), providing a new source of available energy. This led to an energy surplus that absolutely dwarfed anything mankind had ever used before. Consumption of this surplus caused a wave of unprecedented innovation and advancement, as well as an almost unrestricted growth of population. However, like the Romans, sustaining this complex civilization and ever growing population requires huge amounts of energy. Unfortunately, the majority of this energy is still provided by fossil fuels and although large reserves still exist, it is generally accepted that the end of the cheap fossil energy era will be reached within the next few decades. This urges the search for alternative energy scenarios, in which the world ultimately becomes independent of fossil fuels. Growing concerns with respect to greenhouse gas emissions, economic dependence, and political stability will only speed up this development.

Fossil fuels are a non-renewable primary energy source. Primary energy sources are sources of energy that occur naturally. They need to be transformed in energy conversion processes to more convenient forms of energy, called secondary energy sources or energy carriers, that can be used in everyday life to perform work. For example, coal and natural gas (primary energy) can be converted in a fossil fuel power station (energy conversion process) to electricity (energy carrier). Similarly, crude oil (primary energy) can be converted in an oil refinery (energy conversion process) to fuels such as gasoline and diesel (energy carrier). Future energy scenarios will mainly be based on a mix of renewable primary energy sources, such as solar-, wind-, water-, or geothermal energy, and biomass, in combination with energy carriers that are able to address environmental and security issues posed by current fuels. One of the promising future energy carriers in this respect is hydrogen gas (H_2).

Hydrogen

What is it and where can it be used for?

Hydrogen gas, or molecular hydrogen, was discovered in the 18th century [6]. Its monoatomic form is the most abundant element in the universe and molecular clouds of it are associated with the formation of stars [7]. Moreover, it is the main fuel that powers the sun (and many other stars), through a special chemical reaction called nuclear fusion. Nuclear fusion is special as it is the only process in which hydrogen is a primary energy source rather than an energy carrier. Through the sun, hydrogen is thus responsible for the majority of the available energy on earth [8]. Under ordinary conditions on earth, elemental hydrogen exists as the diatomic gas, H_2 . Although elemental hydrogen is the most abundant element in the universe, H_2 is very scarce in the earth's atmosphere. This is because of its light weight, which enables it to escape from earth's gravity more easily than heavier gasses [9]. Moreover, contrary to natural gas (methane, CH_4), no natural accumulation sites of H_2 exist. However, hydrogen is commonly present on the earth's surface, but it is mostly bound in chemical compounds like hydrocarbons (C_nH_m), carbohydrates ($C_n(H_2O)_m$) or water (H_2O).

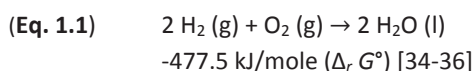
To produce H_2 , hydrogen must be freed from these compounds. This requires quite some work (energy), but some bacteria, archaea and algae are naturally able to do this (**Table 1.1**). In addition, H_2 can be produced industrially via various methods such as electrolysis or

thermolysis [10-14]. Since its discovery, H₂ has become an important chemical commodity. In 2013, its global demand was estimated to be 254.5 billion cubic meters, representing tens of billions of U.S. dollars, and is predicted to grow to an annual demand of ~300 billion cubic meters in 2018 [15]. H₂ is commonly used in numerous industrial processes. Particularly, in the food, metal, electronics and glass-processing industries, where it is for instance used for welding, as coolant, O₂-scavenger, or reactant in various hydrogenation processes. Moreover, it is used in petroleum refining and for the synthesis of other basic chemicals such methanol and ammonia. Since H₂ is an energy carrier, it can also be implemented as an energy storage, transportation and delivery medium and has thus attracted widespread attention as a promising alternative to fossil fuels [16-22].

Table 1.1 Overview of mesophilic, thermophilic and extreme/hyperthermophilic H₂-fermenting bacteria (B) and archaea (A)

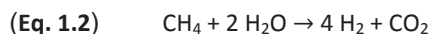
Organism	Domain	Optimal growth temperature	Substrate	H ₂ /hexose	Reference
Mesophiles (~20-45°C)					
<i>Enterobacter aerogenes</i> HU-101	B	37	glucose	0.7-1.1	[23]
<i>Clostridium acetobutylicum</i> ATCC 4259	B	34	glucose	1.0	[24]
<i>Klebsiella oxytoca</i> HP1	B	38	glucose	1.0	[25]
Thermophiles (~45-65°C)					
<i>Clostridium thermosaccharolyticum</i> LMG 6564	B	55	glucose/ fructose	1.6	[26]
<i>Clostridium thermocellum</i> ATCC 27405	B	60	cellulose	1.7	[27]
<i>Thermoanaerobacterium</i> <i>thermosaccharolyticum</i> PSU-2	B	60	starch	2.8	[28]
Extreme/hyperthermophiles (~65-121°C)					
<i>Pyrococcus furiosus</i> DSM3638	A	90	cellobiose	2.8	[29]
<i>Thermococcus kodakarensis</i> KOD1	A	85	maltose	3.5	[29]
			starch	3.3	[30]
<i>Caldicellulosiruptor saccharolyticus</i> DSM8903	B	72	glucose	3.6	[31]
<i>Thermotoga maritima</i> MSB8	B	80	glucose	4.0	[32]

When used as a fuel, H_2 can be oxidized in a combustion engine or in an H_2 -fueled fuel cell. Like conventional combustion engines, a H_2 -fueled one is powered by the energy (heat) released by the oxidation of the fuel, in this case H_2 (**Eq. 1.1**). However, in contrast to conventional fuels, H_2 oxidation with normal air is relatively clean, producing water (H_2O), no carbon dioxide (CO_2), and only a fraction of the normal levels of other environmental pollutants (NO_x). Moreover, when pure oxygen (O_2) is used instead of normal air, H_2 oxidation is completely emission free and produces only H_2O . Similarly, the energy released by the oxidative conversion of H_2 to water can relatively easily be converted to electricity using an H_2 -fuelled fuel cell. H_2 -fueled fuel cells are considered very promising as no pollutants are produced and the conversion of the chemical energy of H_2 into electricity is far more efficient than combustion processes [16-22, 33].



Biological hydrogen production

Since H_2 is an energy carrier, it must be produced from primary energy resources. Unfortunately, most of the current H_2 -generating techniques still rely on fossil fuels in one way or another and can hence not be considered sustainable. For example, the extraction of H_2 by steam reforming of CH_4 (**Eq. 1.2**), which is one of the most applied methods today, requires the direct input of natural gas. Moreover, even when methods are used that do not require the direct input of fossil fuels, such as electrolysis of water (**Eq. 1.3**), the energy (electricity) needed to extract H_2 often is produced from fossil fuels. Therefore, to shift to a more sustainable production, H_2 must be derived from renewable resources, like biomass [37-39]. This so-called biological hydrogen ($BioH_2$) would achieve nearly zero carbon emissions in a whole life cycle and allows for easy combination with carbon sequestration (i.e. carbon capture at the site of production and long term storage in for example biomass). Subsequently using this captured CO_2 for the autotrophic growth of microalgae, for instance, would even allow carbon-neutral production of other biofuels or value-added products [40, 41].



BioH₂ can be produced by a number of different processes including biophotolysis, microbial photo-fermentation, microbial electrolysis, or microbial dark fermentation [11, 16, 42]. Though the exact mechanism and energy source may differ, the principle reaction behind all of these methods is the same: two protons (H⁺) and two electrons (e⁻) are brought together to form H₂ (Eq. 1.4). During microbial dark fermentation, biomass-derived carbohydrates, such as glucose, are anaerobically digested to organic acids and/or alcohols, CO₂ and H₂ [42]. Generating BioH₂ by microbial dark fermentation has several advantages. It requires only relatively simple reactor technology, no direct energy input is needed, many (facultative) anaerobic bacteria and archaea are naturally able to perform it (Table 1.1), and the hydrogen evolution rate is higher in contrast to other methods [42, 43]. Moreover, it allows using (complex) organic wastes, such as agro-industrial residues, for BioH₂ production, which are abundant, cheap, renewable and biodegradable [44-46]. Unfortunately, some major known drawbacks of this method are that large amounts of side products (acids, alcohols) are formed and that H₂ production yields are still too low to make the process economically competitive [15, 43, 44, 47]. The latter is mainly due to the fact that BioH₂ is a low-value product, for which costs of feedstock carbohydrates often account for the majority of the price of the products [47]. This implies that, to enhance the economy of the microbial dark fermentation it is required to improve the sugar-to-H₂ yield.

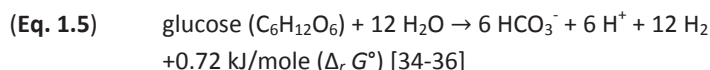


Microbial dark fermentation

Current sugar-to-H₂ yields by microbial dark fermentation vary from less than 1 to almost 4 mole of H₂/mole of hexose consumed (Table 1.1) [13, 14, 43, 48, 49]. The differences in yield are mainly due to the types and relative proportions of the products generated by fermentation. Besides H₂, hydrogen-fermenting microbes make other products to satisfy their metabolic needs and to allow growth to continue. For example, during the anaerobic

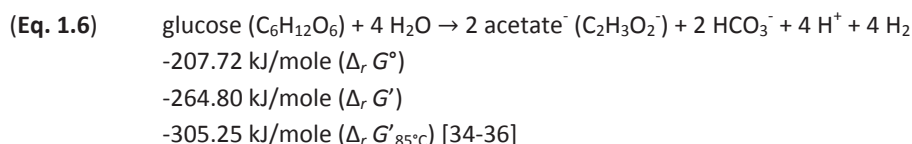
digestion of biomass-derived carbohydrates, electrons are released which are typically transferred to electron carriers like NAD(P)⁺ or ferredoxin (Fd_{ox}). The reduced electron carriers generated, NAD(P)H and Fd_{red}, need to get re-oxidized (recycled) to keep sugar oxidation going and conserve energy for growth. Depending on their genome and environmental (culture) conditions, (facultative) anaerobic microorganisms may have different options to do this; these include the subsequent transfer of electrons to protons (H⁺), generating H₂, or to catabolic intermediates, producing reduced products such as lactate, ethanol or butanol (at the expense of H₂). Moreover, during anaerobic growth, electron transport phosphorylation (ETP), which acts as an electron sink and ATP source, is absent and cells mainly depend on substrate level phosphorylation (SLP) for their ATP (energy) production. Like reductant recycling, microbes may have different options for generating ATP by SLP, which may vary in efficiency or effect on H₂ production. Overall, carbon flow, reductant recycling (including H₂ production) and ATP generation are thus tightly coupled.

To maximize H_2 production, dark fermentation should only lead to fermentation end-products that are coupled to H_2 production. Ideally, the common biomass-derived carbohydrate glucose is completely oxidized to CO_2 and H_2 (Eq. 1.5). However, complete anaerobic oxidation of glucose is thermodynamically difficult, which is indicated by the positive value for the Gibbs energy ($\Delta_r G^\circ$) associated with this reaction, and is only possible at very low partial H_2 pressures [48]. Moreover, since dark fermentation is performed by microorganisms, fermentations have been optimized by evolution to produce biomass (ATP) and not H_2 . According to the known metabolic pathways, complete anaerobic glucose oxidation generates no ATP by SLP and hence no single organism exists that is capable of catalyzing this conversion.



In nature, anaerobic digestion of glucose generates the highest ATP and H₂ yield (4 ATP and H₂ per glucose) when two moles of acetate and bicarbonate (HCO₃⁻) are concomitantly formed as fermentation end-products (**Eq. 1.6**) [50, 51]. **Table 1.1** shows that

hyperthermophilic bacteria and archaea, especially, appear to perform best in this respect. This is because the conversion of glucose to acetate, CO₂ and H₂ is thermodynamically difficult as sufficient Gibbs energy must be obtained to generate the 4 ATP coupled to this fermentation-pathway. Under physiological conditions, the Gibbs energy change needed for ATP synthesis must amount to a minimum of ~70 kJ/mole of ATP [52]. Thus, the minimum Gibbs energy quantum that can generate 4 mole of ATP would amount to ~280 kJ/mole of glucose. However, under standard conditions, the Gibbs energy released by the acetate fermentation-pathway is only -207.72 kJ/mole ($\Delta_r G^\circ$) (**Eq. 1.6**), which means that the conversion of glucose to acetate, CO₂ and H₂ is not possible under these conditions. Even under physiological conditions (10 mM) and a low partial H₂ pressure (10⁻¹ atm), the conversion does not generate sufficient Gibbs energy to synthesize 4 mole ATP (**Eq. 1.6**, -264.80 kJ/mole ($\Delta_r G'$)). Only when the temperature is increased as well, the conversion of glucose to acetate, CO₂ and H₂ generates enough Gibbs energy to synthesize 4 mole of ATP (**Eq. 1.6**, -305.25 kJ/mole ($\Delta_r G'_{85^\circ\text{C}}$)).



However, the production of 4 mole H₂/glucose that can be reached by some hyperthermophiles is still only a third of the theoretical maximal H₂ yield of 12 mole H₂/glucose. This means that about two thirds of the hydrogen and energy bound in glucose is excreted as acetate. Nevertheless, due to the metabolic and thermodynamic constraints explained above, the production of 4 mole H₂/glucose is generally regarded as the upper limit that can be reached by microbial dark fermentation [51]. Especially, since a subsequent conversion of the fermentation end-product acetate to H₂ and CO₂ requires an input of additional energy (**Eq. 1.7**), and can hence not be performed by dark fermentation.

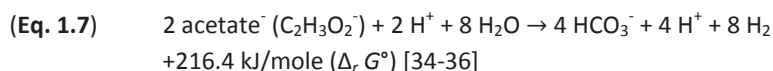


Photo-fermentative organisms on the other hand, can overcome this energy barrier by utilizing the energy from solar light obtained via their light-driven proton pumps [11, 42]. Similarly, acetate can be converted to H_2 and CO_2 using microbial electrolysis, in which the additional energy required can be provided by an externally applied potential [53]. So the overall H_2 /glucose yield can be increased by using dark fermentation end-products as a substrate in a subsequent photo-fermentative or microbial electrolytic process [41]. However, such combined systems have lost the benefits associated with dark fermentation (i.e. relatively simple reactor technique, no energy input, high H_2 evolution rate, etc.), and are therefore too expensive.

Therefore, it would be desirable to combine the beneficial aspects of dark and photo-fermentation in one easy to cultivate microorganism. However, as explained, it is very unlikely that such an organism has been created by evolution. Nevertheless, the increasing number of microbial genomes being sequenced, as well as the increasing availability and ever improving quality of genetic tools, might enable us to genetically engineer such an organism. Proof of principle that extra substrate oxidation steps can improve H_2 yields from glucose, has already been shown in vitro. For example, near stoichiometric amounts were reached by using enzymatic, cell-free, synthetic pathways [54-56]. Although this so-called cell-free synthetic pathway biotransformation (SyPaB) definitely has its advantages [47], it is also heavily challenged by for example its constant and costly dependence on replenishment of labile enzymes and coenzymes, which would not be needed in an in vivo system. To engineer an organism that is able to convert glucose to higher H_2 yields than currently obtained, it would make sense to start with a hyperthermophilic organism that is already able to produce H_2 near the natural limit of 4 mole H_2 /glucose and that is genetically accessible, such as *Thermococcus kodakarensis*.

Thermococcus kodakarensis

General information and description

Thermococcus kodakarensis is a marine hyperthermophilic, heterotrophic, obligate anaerobic archaeon that can grow at temperatures ranging from 60°C to 100°C, with an optimum at 85°C. *T. kodakarensis* can grow on a variety of organic substrates in the presence of elemental sulfur, producing hydrogen sulfide (H₂S), or when pyruvate or starch are present, in the absence of elemental sulfur producing hydrogen gas (H₂). Under optimal conditions, *T. kodakarensis* grows rapidly to high cell densities, having an estimated doubling time of ~40 minutes. Moreover, it can be grown on solid media, forming defined but small colonies [57, 58]. With its genome being fully sequenced [59], its natural competence and the establishment of various genetic engineering techniques [60-63], *T. kodakarensis* rapidly became one of the model organisms for studying anaerobic Archaea.

The type strain *T. kodakarensis* KOD1 was isolated from a sulfatara (a volcanic area that releases hot vapors and sulfurous gases into the environment, 102°C, pH 5.8) on the shore of Kodakara Island, Kagoshima, Japan [58]. Its cells are irregular cocci, having a diameter of 1-2 µm, which are motile due to several polar flagella (**Figure 1.2**). *T. kodakarensis* has a single ether lipid membrane, consisting of both tetra-ether and di-ether type lipids, and has an outer glycoprotein coat which is also known as the S-layer. It has a circular chromosome

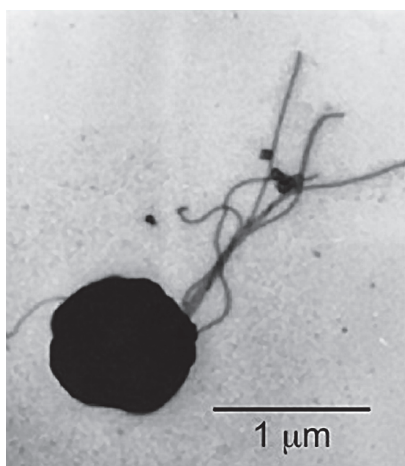


Figure 1.2 *Thermococcus kodakarensis* KOD1
(Image by F. Fukuri, Kyoto University)

containing a small 2.09 Mb genome, which is about half of the average size of microbial genomes (4 Mb), and is closely related to *Pyrococcus* spp. [59]. Originally, the strain was reported as *Pyrococcus* sp. KOD1 [58], but was later reclassified as a species of the euryarchaeal genus *Thermococcus*, based on its 16S rRNA sequence [57]. The members of the genus *Thermococcus* are ubiquitously present in natural high-temperature environments, and are therefore considered to play a major role in the

ecology and metabolic activity of microbial consortia within hot-water ecosystems [64, 65]. After its reclassification, the strain was designated *Thermococcus kodakar-a-ensis* KOD1 sp. nov., which was later corrected on validation to Rule 61 of the Bacteriological Code to *T. kodakarensis* KOD1 [66].

Energy metabolism and H₂ production

T. kodakarensis is an anaerobic heterotroph that uses environmentally abundant polymers, such as peptides, chitin or starch, as its carbon and energy source. The strain grows optimally at 85°C, pH 6.5, using sulfur (S⁰) as a terminal electron acceptor (producing H₂S) [57, 58]. When growing on proteins and (large) peptides, *T. kodakarensis* first converts these polymers extracellularly to smaller peptides and amino acids which can be taken up by the cell. Once inside, the peptides are further degraded to amino acids, which are subsequently catabolized via three steps: (i) oxidative deamination of amino acids by glutamate dehydrogenase (GDH: **Figure 1.3**) or aminotransferases, (ii) oxidative decarboxylation by 2-oxoacid:ferredoxin oxidoreductases (POR, VOR, IOR and OGOR: **Figure 1.3**), and (iii) hydrolysis of acyl-coenzyme A (CoA) by ADP-forming acyl-CoA synthetases (ACS I-III, SCS and ICS: **Figure 1.3**). During this catabolic process, reductant (Fd_{red} and NADPH), ATP and CO₂ are produced. To keep amino acid oxidation going and to conserve additional energy for growth, the generated reductant needs to get re-oxidized, which ultimately results in H₂S and some H₂ (**Figure 1.3, box D & F**) [67-71].

In the absence of S⁰, *T. kodakarensis* can grow on pyruvate and starch, using protons (H⁺) as a terminal electron acceptor (producing hydrogen gas (H₂)) [57, 58]. Before taken up by the cell, starch polymers have to be hydrolyzed extracellularly to malto-oligosaccharides like maltodextrin, maltotriose or cyclomaltodextrin, but preferably not to mono- and disaccharides as these cannot be taken up by the cell [30, 59, 72]. The extracellular hydrolysis of starch is realized by secreting various enzymes that degrade the available starch polymers synergistically, such as amylases, pullulanase, amylopullulanase and cyclomaltodextrin glucanotransferase [59, 72-76]. Once imported in the cell, these malto-oligosaccharides are further hydrolyzed intracellularly to either glucose or, via glucose-1-

phosphate, to glucose-6-phosphate were they enter a modified Embden-Meyerhof (EM) pathway (**Figure 1.3, box A-C**) [59, 76, 77].

This modified EM pathway for glucose degradation is also present in other *Thermococcales* and differs from the classical EM pathway found in, for example, *Escherichia coli* in some key aspects. For instance, the phosphorylation reactions of glucose and fructose-6-phosphate are catalyzed by ADP-dependent glucokinase and phosphofructokinase, respectively, rather than by the usual ATP-dependent versions [78-80]. Moreover, the conversion of glyceraldehyde-3-phosphate (GAP) to 3-phosphoglycerate (3PGA), normally catalyzed by the consecutive action of NAD⁺-dependent GAP dehydrogenase (GAPDH) and the ATP-generating phosphoglycerate kinase (PGK), is catalyzed by both ferredoxin-dependent GAP:ferredoxin oxidoreductase (GAPOR) and NADP⁺-dependent non-phosphorylating glyceraldehyde 3-phosphate dehydrogenase (GAPN), in an irreversible, phosphate-independent, one-step reaction [77]. This is an important difference as the direct oxidation of GAP to 3PGA by GAPOR/GAPN omits ATP production via substrate level phosphorylation by PGK. Finally, the last step of glycolysis (i.e. the ATP-generating conversion of phosphoenolpyruvate (PEP) to pyruvate), normally catalyzed by ADP-dependent pyruvate kinase (PYK), is additionally catalyzed by AMP-dependent phosphoenolpyruvate synthase, which in contrast to PYK has been proven to be essential in *T. kodakarensis* [81]. This PEP synthase catalyzed reaction enables the conservation of 2 energetic phosphate bonds instead of just 1 in the pyruvate kinase reaction. Thus the PEP synthase reaction compensates for the absence of ATP synthesis in GAP oxidation.

Reduced ferredoxins (Fd_{red}), generated by glycolysis, must be re-oxidized to keep sugar oxidation going and conserve energy for growth. Although *T. kodakarensis* has several different options for the regeneration of Fd_{ox} from Fd_{red}, the main system used during sugar oxidation is the production of H₂ by a membrane bound hydrogenase. Part of the generated H₂ can subsequently be re-consumed by the cell and re-oxidized by a cytosolic or soluble hydrogenase to generate NADPH when required (**Figure 1.3, box F**) [68, 82, 83]. Any excess NADPH may be disposed of through alanine production, mediated by glutamate

dehydrogenase and alanine aminotransferase, which has been shown to be a potential electron sink (**Figure 1.3 box C**) [30, 59, 68, 83, 84].

T. kodakarensis has previously been described to ferment starch to acetate, CO₂ and H₂, having a H₂-production yield of 3.3 mole/mole of hexose [30]. Since this is already relatively close to the generally considered upper limit of 4 mole/mole of hexose, it is a good starting point for improvement. In addition, its relatively simple (energy) metabolism and the availability of a genetic toolbox, might allow for the introduction of additional substrate oxidation steps.

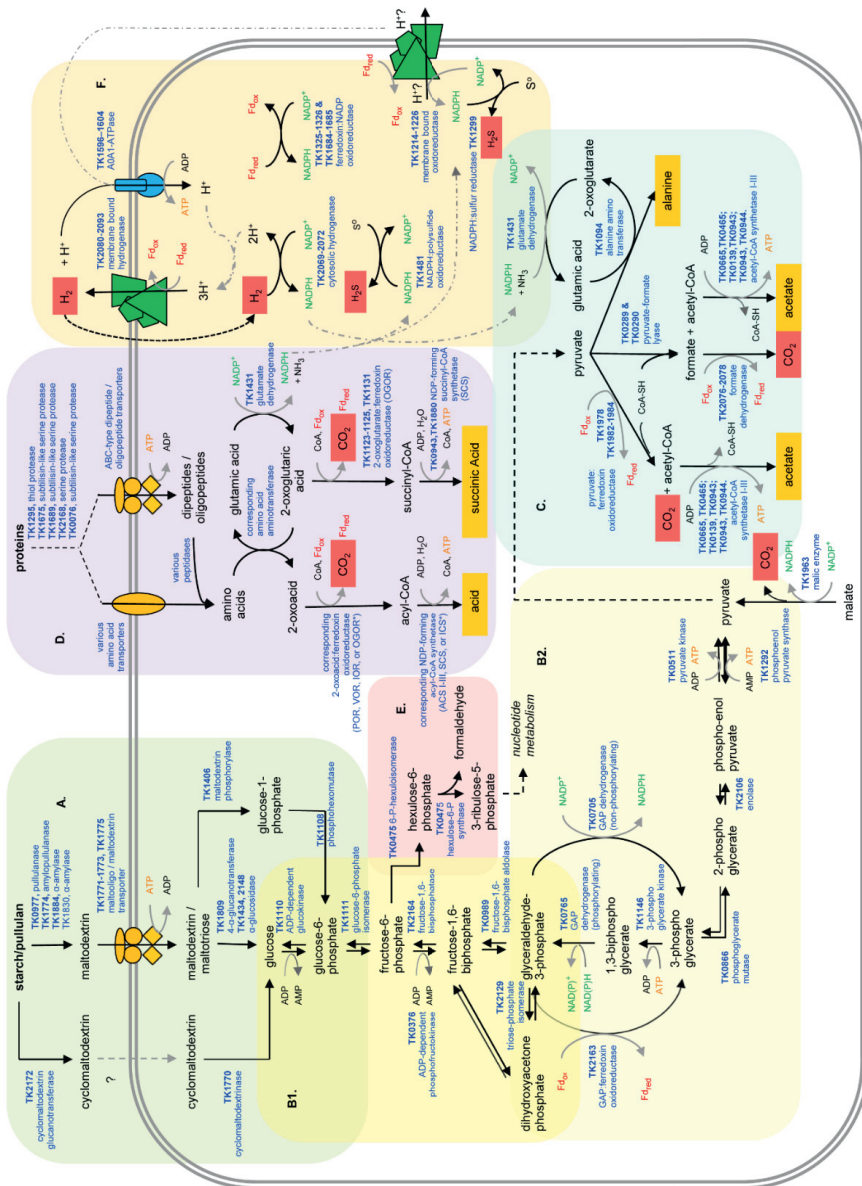


Figure 1.3 Overview of energy and hydrogen metabolism in *Thermococcus kodakarensis*. Substrates (in bold), metabolites (normal text), and fermentation end-products (gases in red boxes, acids and alanine in yellow boxes) are in black. Enzymes/proteins (normal text) and their corresponding locus tags (in bold) are in blue. The energy-rich intermediate compounds reduced Ferredoxin (Fd_{red}), NADPH and ATP are in red, green and orange, respectively. Panels: A (green), starch/maltodextrin metabolism; B1 (yellow), preparation phase modified Embden-Meyerhof pathway; B2 (light yellow), pay-off phase modified Embden-Meyerhof pathway; C (blue), pyruvate metabolism; D (purple), amino acid catabolism; E (pink), ribulose monophosphate pathway (substitutes for missing PPP); F (orange), major pathways for reduced cofactor recycling. * = POR, pyruvate:ferredoxin oxidoreductase (TK1978, 1982-1984); VOR, 2-ketoisovalerate:ferredoxin oxidoreductase (TK1978-1981); IOR, indolepyruvate:ferredoxin oxidoreductase (TK0135-0136); OGOR, 2-oxoglutarate:ferredoxin oxidoreductase (TK1123-1125, 1131); ACS I, acetyl-CoA synthetase I (TK0465, 0665); ACS II, acetyl-CoA synthetase II (TK0139, 0943); ACS III, acetyl-CoA synthetase III (TK0943, 0944); SCS, succinyl-CoA synthetase (TK0943, 1880); ICS, 2-(imidazol-4-yl) acetyl-CoA synthase (TK0943, 2127). The figure is based on a review of the following references [59, 67-77, 81-83, 85-96].

Thesis Outline

Many substrate oxidation steps result in the generation of NADPH. However, H₂ formation from reduced pyridine nucleotides (NAD(P)H) is unfavorable as the Gibbs energy under standard conditions is positive. Only at relatively high temperatures and at low H₂ partial pressures, the reaction becomes exergonic [48]. Nevertheless, H₂ formation from NAD(P)H is possible, which can also be concluded from the H₂/hexose ratios found for certain (hyper)thermophiles. For instance, *Thermoanaerobacter tengcongensis* and *Caldicellulosiruptor saccharolyticus* both produce reducing equivalents as NADH and Fd_{red} (50/50%), and based on the observed high H₂/hexose ratios, both types of electron carrier are apparently able to transfer their reducing equivalents to H₂ [97, 98].

For the direct production of H₂ from NADPH, an NADPH-dependent hydrogenase is required. Although this sub-class of [NiFe]-hydrogenases is rare [99-101], *T. kodakarensis* does contain such a cytosolic hydrogenase (**Figure 1.3, box F** (TK2069-2072)). However, this hydrogenase is normally utilized for oxidizing H₂ to generate NADPH when NADPH levels are low [82, 83], but various studies have shown that these [NiFe]-hydrogenases are readily reversible in vitro [56, 101-103]. Increasing intracellular NADPH levels by adding additional substrate oxidation steps could therefore shift the catalytic direction of the reaction towards H₂ evolution. **Chapter 2** provides an overview of the major canonical and non-canonical reactions involved in the production and regeneration of NADPH in prokaryotes and discusses their key enzymes. In addition, an overview of how different enzymes have been applied to increase NADPH availability and thereby enhance productivity of many biotechnological products is provided.

Chapter 3 builds upon knowledge gained in chapter 2 and describes the introduction of the (bacterial) pentose phosphate pathway (PPP) in the hyperthermophilic archaeon *T. kodakarensis*. The PPP is a major NADPH source in many organisms and is able to provide additional glucose oxidation steps in *T. kodakarensis*. Especially, since the species does not normally contain a complete version of this pathway. Instead, it relies on the ribulose monophosphate pathway for the synthesis of its pentoses (**Figure 1.3, box**) [91]. Therefore,

to make sure that sugars taken up by the cell are (partially) directed through the introduced PPP, the ribulose monophosphate pathway was disrupted simultaneously. Together, these metabolic changes may improve the sugar-to-H₂ yield by microbial dark fermentation, while ensuring sufficient Gibbs energy for ATP synthesis by SLP. The effects of the metabolic changes on the growth performance, metabolism, and the fermentation profile of *T. kodakarensis* are evaluated.

Genomic modifications are generally relatively easily obtained in *T. kodakarensis*, and are therefore routinely applied (e.g [60-62, 71, 81-83, 104-106]). However, while creating the strains discussed in chapter 3, some puzzling and contradicting results were observed; it appeared that different chromosome versions co-existed within one strain. Although such observations may potentially be explained by cross-contamination, the difficulty in obtaining homozygous recombinant cultures, as well as the repeated observation of this phenomenon among different engineering attempts, suggested that *T. kodakarensis* may possess multiple copies of its chromosome. This prompted us to investigate the chromosome copy number of *T. kodakarensis* KOD1, which is discussed in **Chapter 4**.

In **Chapter 5**, a quest for a glycerol-dialkyl-glycerol tetraether (GDGT) synthase in *T. kodakarensis* is described. As shown in **Table 1.1**, hyperthermophilic bacteria and archaea are better H₂ producers than thermophilic or mesophilic ones. In the introduction, it has already been explained that this is mainly due to thermodynamic constraints, favoring H₂ production at higher temperatures. Therefore, it is interesting to understand how these hyperthermophilic microbes are able to thrive at such high temperatures, so that these characteristics might be transferred to other H₂ producers as well. Naturally, the main focus of how these microbes withstand the perturbations of cellular integrity has been on the plasma membrane. Hyperthermophilic archaea have ether lipid membranes (as opposed to the bacterial ester lipids ones), which consists of either di-ether type lipids (GDGD), tetra-ether type lipids (GDGT), or a combination of both [107, 108]. GDGTs especially, are thought to confer enhanced thermal stability to the lipid membrane [109]. However, despite two alternative biochemical pathways have been proposed, it is currently unclear how GDGTs are synthesized and particularly which enzyme(s) catalyze the tail-to-tail condensations between

Chapter 1

isoprenoids [110]. *T. kodakarensis* offers an unique opportunity to identify this missing link in archaeal lipid biosynthesis, as it meets a couple of important criteria; (i) it contains both GDGD and GDGT type lipids, (ii) its genome has been sequenced and contains only a relatively small number of genes, (iii) the strain can be cultured in a laboratory, and (iv) a genetic toolbox is available.

Finally, **Chapter 6** summarizes and discusses the research presented in this thesis and concludes with recommendations for future research.

Chapter 2

NADPH-generating systems in bacteria and archaea

This chapter has been published as:

Spaans, S.K., R.A. Weusthuis, J. van der Oost, S.W.M. Kengen. *NADPH-generating systems in bacteria and archaea*. *Frontiers in Microbiology*, 2015. **6**(742)

Abstract

Reduced nicotinamide adenine dinucleotide phosphate (NADPH) is an essential electron donor in all organisms. It provides the reducing power that drives numerous anabolic reactions, including those responsible for the biosynthesis of all major cell components and many products in biotechnology. The efficient synthesis of many of these products, however, is limited by the rate of NADPH regeneration. Hence, a thorough understanding of the reactions involved in the generation of NADPH is required to increase its turnover through rational strain improvement. Traditionally, the main engineering targets for increasing NADPH availability have included the dehydrogenase reactions of the oxidative pentose phosphate pathway and the isocitrate dehydrogenase step of the tricarboxylic acid (TCA) cycle. However, the importance of alternative NADPH-generating reactions has recently become evident. In the current review, the major canonical and non-canonical reactions involved in the production and regeneration of NADPH in prokaryotes are described, and their key enzymes are discussed. In addition, an overview of how different enzymes have been applied to increase NADPH availability and thereby enhance productivity is provided.

Introduction

Reduced nicotinamide adenine dinucleotide phosphate (NADPH) is an essential electron donor in all eukaryotes, bacteria, and archaea. Not only is NADPH vital for the generation of reactive oxygen species (ROS) [111-114] and the anti-oxidative defense mechanisms of most organisms [115-117], most importantly, it is also the driving force of most biosynthetic enzymatic reactions, including those responsible for the biosynthesis of all major cell components, such as DNA and lipids [115-121]. Owing to this essential role in biosynthesis, NADPH availability has been of major interest to industry [122, 123]. Many natural products of industrial importance are complex secondary metabolites, the production of which often involves NADPH-dependent enzymes. To synthesize such products using purified enzyme systems *in vitro* would require the addition of huge amounts of NADPH in order to sustain production. From an industrial point of view, this would be too expensive. Hence, *in situ* NADPH regeneration from its oxidized counterpart (NADP⁺) is required [124-128].

Regeneration of NADPH *in situ* can be achieved by various strategies, including chemical, electrochemical, photochemical, and enzyme-based methods [125, 127-131]. NADPH can be regenerated enzymatically by complementing the *in vitro* system with additional enzymatic reactions or by using substrate-coupled reaction systems. The latter system employs enzymes that use both NADP⁺ and NADPH and that are able to catalyze the synthesis of the desired product from one substrate and cofactor regeneration with another substrate [124, 130, 132]. However, reduced productivity compared to systems without *in situ* regeneration and problems associated with enzyme stability make these options unattractive.

Microbial *in vivo* production systems also provide *in situ* NADPH regeneration and have several advantages when compared to *in vitro* systems. For example, microbes are able to grow on inexpensive renewable feedstocks that provide the organisms with reductant for the regeneration of NADPH. They also contain numerous pathways, involving stable and highly specific enzymes, thus obviating the need for expensive enzyme purification. In addition, our knowledge of natural metabolic pathways is rapidly advancing, allowing for rational design towards product formation [122, 123, 133, 134]. Therefore, it is not

surprising that microbial conversion is the preferred method for the synthesis of a range of products.

With the possibility of engineering microbial metabolism to facilitate product formation, it became clear that NADPH availability remains a major hurdle in the efficient generation of many products. These products range from medicinal compounds [133-135] and (essential) amino acids [123, 136] to molecules used as biofuels [137-139] and building-blocks for biodegradable plastic [140]. Given its involvement in a multitude of crucial biological functions and its importance in biosynthesis, NADPH is without question an essential molecule. Hence, a key question arises: what are the major NADPH-generating reactions and systems?

Traditionally, the dehydrogenase reactions of the oxidative pentose phosphate pathway (oxPPP), the Entner-Doudoroff (ED) pathway, and the isocitrate dehydrogenase step of the tricarboxylic acid (TCA) cycle have been considered the major sources of NADPH. However, the importance of other NADPH-generating enzymes, such as transhydrogenases, glucose dehydrogenases, and non-phosphorylating glyceraldehyde 3-phosphate dehydrogenase (GAPN), is becoming clear, indicating that the traditional view is over-simplistic [77, 78, 141].

In this review, we describe the major canonical and noncanonical biochemical mechanisms that are involved in the production and regeneration of NADPH in prokaryotes and discuss the key enzymes involved. We have divided the mechanisms into those that are directly coupled to central carbon metabolism and those that are not (**Table 2.1**). In addition, we briefly address how different enzymes have been applied to increase NADPH availability and thereby enhance NADPH-dependent biotransformation processes.

Table 2.1 Overview of major canonical and non-canonical NADPH-generating enzymes

Enzyme	EC: number	Pathway	Bacteria* (%)	Archaea* (%)	Additional cofactors	Applied**	$\Delta_r G^m$ *** (kJ/mol)
NADPH-generating enzymes coupled to central carbon metabolism	G6PDH	oxPPP, ED	66	0	n.a.	yes	-2.3 ± 2.6
	6PGDH	oxPPP	62	27	n.a.	yes	-6.0 ± 6.3
	IDH	TCA cycle	82	59	n.a.	yes	-10.7 ± 6.3
	ME	Anaplerotic node	47	25	n.a.	no	-3.1 ± 6.2
	GAPN	EMP	12	31	n.a.	yes	-36.1 ± 1.1
	NADP ⁺ -GAPDH	EMP	n.d.	n.d.	n.a.	yes	25.9 ± 1.0
	GDHs	Modified EDs	18	10	n.a.	yes	-2.4 ± 2.2
NADPH-generating enzymes not coupled to central carbon metabolism	STH	n.a.	n.d.	n.d.	n.a.		
	H ⁺ -TH	n.a.	19	0	NADH, FAD	yes	1.0 ± 0.7
	FNR	n.a.	50	5	NADH	yes	1.0 ± 0.7
	SH	n.a.	63	29	FAD or FMN, Fd _{red}	no	-15.6 ± 11.7
		n.a.	10	6	FAD or FMN	no	-16.5 ± 5.9
	NADK	n.a.	96	98	NTP or poly(P), NAD ⁺ or NADH	yes	n.a.

* Distribution of the enzyme in completely sequenced prokaryotic genomes. All available prokaryotic genomes were downloaded to Excel from the IMG database on 6/16/2015. The analysis included only finished genomes (198 archaeal and 3380 bacterial genomes) and only one representative genome from each genus (83 archaeal and 636 bacterial genomes), to correct for the many draft genomes and for the many sequences in closely related genomes, respectively. The obtained distribution is merely an indication of the true distribution of the respective enzymes among prokaryotes because the database is incomplete and gene functions are often automatically annotated.

** Use of the enzyme in metabolic engineering to increase NADPH availability and enhance biotransformation processes.

*** Gibbs energies ($\Delta_r G^m$), calculated using the biochemical thermodynamics calculator eQuilibrator 2.0 [34-36], which calculates $\Delta_r G^m$. This term describes the change in Gibbs energy due to a chemical reaction at a particular pH and ionic strength (subscript r), using 1 mM concentrations for all reactants (superscript m). Default settings (pH 7 and ionic strength 0.1 M) were used for all reactions.

Abbreviations: n.a., not applicable; n.d., no data available in the IMG database; G6PDH, glucose-6-phosphate dehydrogenase; 6PGDH, 6-phosphogluconate dehydrogenase; IDH, isocitrate dehydrogenase; ME, malic enzyme; GAPN, non-phosphorylating glyceraldehyde 3-phosphate dehydrogenase; NADP⁺-GAPDH, NADP⁺-dependent phosphorylating glyceraldehyde 3-phosphate dehydrogenase; GDH, glucose dehydrogenase; STH, energy-independent soluble transhydrogenase; H⁺-TH, energy-dependent or proton-translocating, membrane-bound transhydrogenase; FNR, ferredoxin: NADP⁺ oxidoreductase; SH, cytosolic NADP⁺-reducing hydrogenase; NADK, NAD⁺ kinase

Biosynthesis of NADP⁺

To maintain a sufficient quantity of NADP⁺ for the generation of NADPH, NADP⁺ biosynthesis is essential. While the necessity for NADP⁺ synthesis during cell proliferation is clear, the requirement for NADP⁺ synthesis in non-dividing cells might not be obvious. NADP(H) is generally thought of solely as a redox carrier that facilitates the transfer of electrons between two redox couples, a role that does not account for the need for constant resynthesis. The constant need to resynthesize NADP⁺ arises from its participation in other crucial biological processes [142]. In eukaryotes, for example, NADP⁺ serves as the substrate in the synthesis of nicotinic acid-adenine dinucleotide phosphate (NAADP), an important intracellular Ca²⁺-mobilizing messenger [143-146]. Although NAADP has also been found in bacteria [143], its physiological relevance in these organisms remains to be established.

The enzymes responsible for NADP⁺ resynthesis, best studied in eukaryotes, mainly include NAD(P)⁺ nucleosidase [147, 148], NADP⁺ phosphatase, and NADPH phosphatase. The enzymes in prokaryotes are less well studied, but similar activities and proteins have been found in this domain [149-155]. NADP⁺ phosphatase and NADPH phosphatase, together with NAD(H) kinase, regulate the intracellular balance of NAD(H) and NADP(H) [152]. Like NADPH, NADH mainly serves to transfer electrons from one molecule to another. However, unlike NADPH, NADH is primarily involved in catabolic reactions [148]. In addition to its redox function, NAD⁺ serves as a substrate for mono- and poly-ADP ribosylation, participates in histone deacetylation, and contributes to the production of the signaling molecule cyclic ADP-ribose. Most of these reactions have been characterized in eukaryotes, but the ribosylation reactions also play a role in toxin production by pathogenic bacteria [155, 156]. The consumption of NADP⁺ is thus connected to the consumption of NAD⁺ and to the regulation of various major biological activities such as DNA repair, gene expression, apoptosis, nitrogen fixation, and calcium homeostasis [142, 148, 152, 155, 156].

Given that the intracellular balance between NAD(H) and NADP(H) is regulated via the addition or removal of a phosphate group by NAD(H) kinase or NADP(H) phosphatase, respectively, it is clear that the biosynthesis of NAD⁺ plays a crucial role in the metabolism of

NADP⁺. Two principal NAD⁺ biosynthesis pathways have been characterized: 1) the *de novo* pathway and 2) the salvage pathway [157-159]. Both pathways have been reviewed recently [148, 155, 160, 161]. In the *de novo* pathway, NAD⁺ is generated from quinolinic acid, which in prokaryotes is produced from either L-aspartate or L-tryptophan [159, 162, 163]. In the salvage pathway, degradation products containing a pyridine ring, namely nicotinic acid and nicotinamide, are utilized to regenerate NAD⁺ (**Figure 2.1(a)**) [159, 160].

A portion of the synthesized NAD⁺ can be converted into NADP⁺ by NAD kinase (NADK) (**Figure 2.1(b)**). In the NAD⁺ synthesis pathways, several variations exist and multiple enzymes are involved [161, 164]. In contrast, NADK is the sole enzyme able to generate NADP⁺ *de novo*. NAD kinase is found in archaea [165], bacteria [166-168], and eukaryotes [169, 170]) and has been proven to be essential in prokaryotes [171, 172]. According to the literature, only one species not able to synthesize NADP⁺ from NAD⁺ has been identified: *Chlamydia trachomatis* [152, 173]. The intracellular parasite appears to lack NADK and hence relies completely on the metabolism of its host cell.

According to Kawai and Murata [152], NADK orthologs can be classified into three types according to their substrate specificity: I) NADKs that utilize both ATP and inorganic polyphosphate (poly(P)) as a phosphoryl donor and phosphorylate both NAD⁺ and NADH. This type of NADK has been identified in gram-positive bacteria and archaea. II) NADKs that utilize ATP but not poly(P) and phosphorylate both NAD⁺ and NADH. This type has been identified in eukaryotes. III) NADKs that utilize ATP but not poly(P) and phosphorylate NAD⁺ but not NADH. The latter type of NADK has been identified in gram-negative bacteria. Because of its vital role in NADP⁺ synthesis, NADK has received much attention since its discovery in 1950 [174]. Hence, many papers about its structure, function, and application are available, including reviews by Kawai and Murata [152], Shi *et al.* [175], and Agledal *et al.* [142].

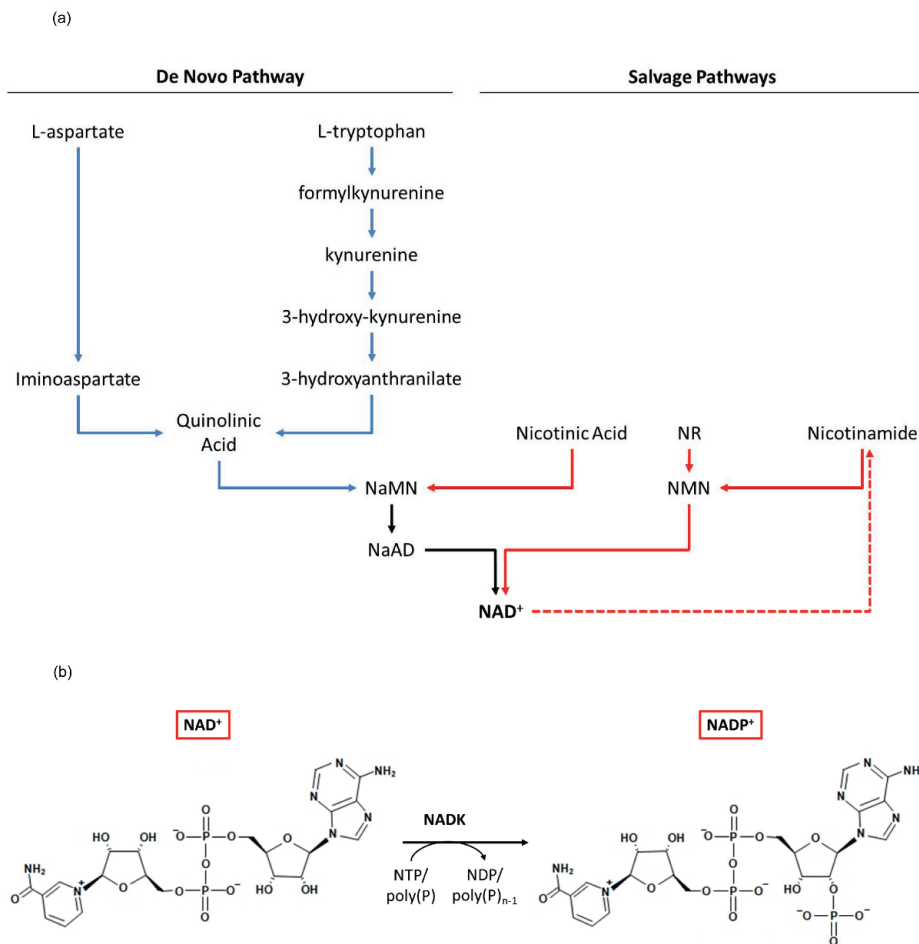


Figure 2.1 NAD⁺ and NADP⁺ biosynthetic pathways in prokaryotes. (a) *De novo* and salvage pathways for NAD⁺ biosynthesis. (b) NADP⁺ biosynthesis by NAD kinase. Abbreviations: NaMN, nicotinic acid mononucleotide; NaAD, nicotinic acid adenine dinucleotide; NR, nicotinamide riboside; NMN, nicotinamide mononucleotide; NAD, nicotinamide adenine dinucleotide; NADP, nicotinamide adenine dinucleotide phosphate; NADK, NAD kinase.

Redox potential of $\text{NADP}^+/\text{NADPH}$

The redox potential of $\text{NADP}^+/\text{NADPH}$ at standard physiological conditions is identical to that of NAD^+/NADH (E_0' : -320 mV). However, the function of the cofactors is different. $\text{NADP}^+/\text{NADPH}$ is used for anabolic redox reactions, whereas NAD^+/NADH is used for oxidation reactions. This is possible because $\text{NADP}^+/\text{NADPH}$ is generally maintained in a reduced state and NAD^+/NADH in an oxidized state [176]. The ratios for the electron carriers have been reported, and the values can differ by orders of magnitude, depending on the organism and growth conditions. For example, reported NAD^+/NADH values range from 3.74 to 31.3 for bacterial cells [52, 177-179] and reach 1820 for mammalian cells [180], whereas $\text{NADP}^+/\text{NADPH}$ values range from 0.017 to 0.95 [52, 177-179]. This indicates that the actual redox potential of both redox couples can deviate significantly from the standard potential (i.e., generally more negative/positive than -320 mV for $(\text{NADP}^+/\text{NADPH})/(\text{NAD}^+/\text{NADH})$, respectively). Such deviation affects the Gibbs energy of the NADPH-generating reactions and the feasibility of converting NADH into NADPH. To assess the feasibility of various NADPH-generating reactions, we calculated the Gibbs energies ($\Delta_r G'^m$) using the biochemical thermodynamics calculator eQuilibrator 2.0 [34-36], setting all reactants at a physiological concentration of 1 mM (**Table 2.1**). However, even under these conditions, several reactions appeared to be at equilibrium or slightly endergonic. At $\text{NADP}^+/\text{NADPH}$ ratios lower than 1, the reaction becomes even less feasible, indicating that the actual concentration of the reaction components must be substantially different for NADPH generation to proceed.

Systems for NADPH generation

NADP^+ synthesized by NADK or generated by NADPH-oxidizing reactions is eventually reduced to NADPH. To ensure that the cellular redox balance is maintained in the absence of *de novo* NADP^+ synthesis and NADP^+ consumption, the catabolic fluxes through the NADPH-regenerating reactions must be matched to the anabolic demand. However, this is generally not the case [181]. In addition, the precise NADPH formation rate depends on fluxes through the generating pathways, which in turn vary with different growth conditions [82, 182, 183].

Therefore, prokaryotes must have other network-wide biochemical mechanisms that maintain the cellular redox balance [181]. The exact mechanisms are not fully understood and are beyond the scope of this review, but papers about the topic are available [117, 181].

Although numerous reactions reduce NADP^+ to NADPH (at least 143 reactions according to the MetaCyc database, accessed on 03/30/15), only a few have been thought to contribute significantly. However, new studies have demonstrated the importance of other NADPH-generating reactions. The present review provides a general overview of the known major NADPH-generating reactions and discusses the key enzymes involved. We have divided the reactions into I) those that are directly coupled to central carbon metabolism and II) those that are not (**Table 2.1**). The enzymes that comprise the first group are the oxPPP enzymes glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase; isocitrate dehydrogenase of the TCA cycle; malic enzyme; and three enzymes involved in noncanonical NADPH-generating reactions: phosphorylating GAP dehydrogenase (NADP^+ -GAPDH), non-phosphorylating GAP dehydrogenase (GAPN), and glucose dehydrogenase. The enzymes in the second group are transhydrogenases (NADH:NADP^+), ferredoxin: NADP^+ oxidoreductase, hydrogenases ($\text{H}_2:\text{NADP}^+$), and NAD(H) kinase, although the latter is a *de novo* NADP^+ /NADPH-synthesizing enzyme rather than an NADPH-regenerating enzyme. NAD(H) kinase is included because it plays an essential role in NADP^+ biosynthesis and the regulation of the NAD(H)/NADP(H) balance offers a potential strategy for improving the biosynthesis of industrially valuable metabolites.

NADPH-generating reactions coupled to central carbon metabolism

Glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase

The oxidative pentose phosphate pathway (oxPPP) enzymes glucose-6-phosphate dehydrogenase (G6PDH, EC:1.1.1.49) and 6-phosphogluconate dehydrogenase (6PGDH, EC:1.1.1.44) have a central role in metabolism in many microbes (**Table 2.1**). The enzymes are involved in the conversion of glucose-6-phosphate into ribulose-5-phosphate (**Figure 2.2**), a precursor of important molecules such as nucleic acids, and are generally considered a major source of NADPH. The importance of the oxPPP enzymes as a source of NADPH has been demonstrated in various organisms [141, 184, 185]. Flux through these enzymes

reportedly increases when NADPH requirements are high [184, 186-188] and decreases when NADPH requirements are low [189, 190]. This is consistent with the anticipated anabolic role of the oxPPP as a supplier of pentoses and NADPH for biosynthesis. The key enzyme with regard to the control of PPP flux is thought to be G6PDH [191-193]. However, in addition to its role in the oxPPP, G6PDH can also participate in the classical phosphorylated version of the Entner-Doudoroff (ED) pathway (**Figure 2.2**). The ED pathway can be considered an alternative to the Embden-Meyerhof-Parnas (EMP) glycolytic pathway, because both catabolize glucose to pyruvate [94, 194]. The ED pathway is mainly present in prokaryotes, although some eukaryotes possess a functional ED pathway as well [195]. A generally held view is that the ED pathway is less important than the EMP pathway with respect to glucose catabolism. However, Fuhrer *et al.* have shown that this might be a misconception and that the ED pathway might be a major pathway for glucose catabolism even in species that possess both pathways [196].

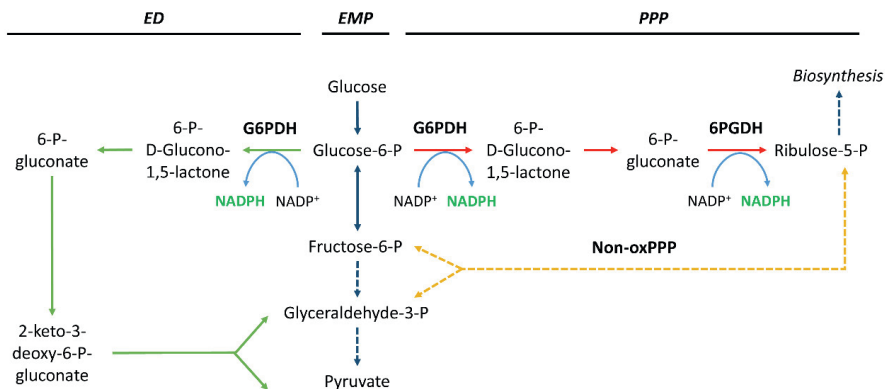
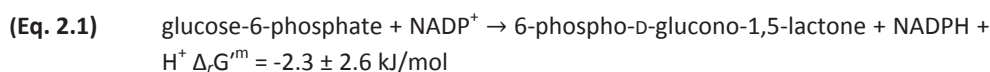
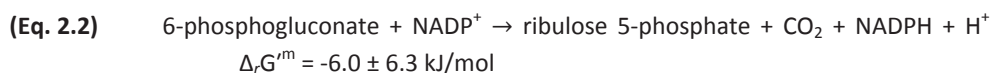


Figure 2.2 NADPH generation by glucose-6-phosphate dehydrogenase (G6PDH) and 6-phosphogluconate dehydrogenase (6PGDH). Solid and dashed lines indicate single and lumped reactions, respectively. Pathway abbreviations: ED, Entner-Doudoroff pathway (in green); EMP, Embden-Meyerhof-Parnas pathway (in blue); PPP, pentose phosphate pathway (oxidative phase in red, non-oxidative phase in yellow).

G6PDHs and 6PGDHs have been widely studied with respect to cofactor specificity. Although some bacterial G6PDHs and 6PGDHs show dual cofactor specificity [181, 197, 198] or have a strong preference for NAD^+ [197, 198] or other cofactors such as the flavin derivative F_{420} [199], most are NADP^+ -dependent [181, 200, 201]. In addition, the dual cofactor specificity that some G6PDHs show is often observed *in vitro* under saturated conditions. However, physiological *in vivo* conditions often result in higher specificities for NADP^+ [181]. Therefore, G6PDHs and 6PGDHs are generally considered to catalyze the following reactions:



and



Like eukaryotes, some prokaryotes possess multiple G6PDH [198, 202, 203] or 6PGDH isozymes [204]. In general, these isozymes are similar, with the same cofactor preferences, but exceptions exist. *Mycobacterium smegmatis*, for example, has two G6PDH isozymes, one of which is NADP^+ -dependent. The other is F_{420} -dependent and does not utilize NAD^+ or NADP^+ [205]. In addition, cell-free extracts of several other *Mycobacterium* and *Nocardia* species contain both F_{420} -dependent and NADP^+ -dependent G6PDH activity [199], indicating that the genomes of these species likely encode both variants. Similarly, exceptions among 6PGDH isozymes have been identified. The genome of *Bacillus subtilis*, for example, encodes three functional 6PGDH isozymes; two prefer NAD^+ , and one prefers NADP^+ [204]. As discussed by Zamboni *et al.*, the presence of isozymes with different cofactor preferences would, in principle, enable NADP^+ , NAD^+ , and F_{420} reduction in the oxPPP to adjust to the overall metabolic requirements of the cell [204]. However, the exact functional roles of these isozymes remains to be determined.

The availability of NADPH is of principal importance for various industrially important classes of products, including amino acids, proteins, antibiotics, organic acids, and high-value metabolites. The apparent relationship between the rate of NADPH generation and actual carbon fluxes through the oxPPP suggests that the oxPPP can be a target for metabolic

engineering for overproduction of NADPH. The oxPPP has been engineered to increase the NADPH/NADP⁺ ratio, through overexpression of oxPPP enzymes [136, 206-208] and through redirection of the carbon flux from glycolysis to the oxPPP, by disrupting phosphoglucose isomerase [133, 140, 209] or phosphofructokinase [210, 211], overexpressing fructose 1,6-bisphosphatase [212], or introducing glucose dehydrogenase [213]. Both strategies have been applied successfully in various prokaryotes, but reduction in growth is a common side effect [208, 209, 214]. However, the opposite effect was observed with an archaeal strain recently developed in our lab: introduction of the oxPPP increased the hydrogen yield and significantly improved the growth rate (unpublished data).

Although both strategies have been applied successfully in various prokaryotes, they are not effective in every organism [215, 216]. Moreover, as mentioned above, G6PDH is a key enzyme not only for control of the PPP flux but also for control of the ED pathway. Overexpressing G6PDH could therefore affect the flux of both pathways simultaneously. However, this is generally not observed. Cells apparently possess the ability to regulate fluxes to ensure a network-wide balancing of NADPH supply and demand [181, 217].

The effects of G6PDH and 6PGDH disruption have also been investigated. The exact effects are strongly dependent on the organism, the genetic background of the parent strain, and the environmental growth conditions [185, 202, 204, 218-220]. However, in general, G6PDH-deficient and 6PGDH-deficient prokaryotes exhibit remarkably mild phenotypes. Their growth rates are similar to or somewhat lower than those of the parent strain, and they are more sensitive to oxidative stress. Flux through the oxPPP and ED pathways is lower in G6PDH-deficient strains. To compensate for this and to ensure a sufficient supply of anabolic precursors and NADPH, flux rerouting via the EMP pathway and the non-oxPPP is generally observed (**Figure 2.2**). In addition, glucose-grown mutants generally display enhanced TCA cycle activity. In contrast, acetate- or pyruvate-grown cells generally display decreased TCA cycle activity [182, 220]. Like G6PDH-deficient strains, 6PGDH-deficient strains exhibit a strong reduction in oxPPP flux, but not in ED pathway flux. In contrast, prokaryotes with a functional ED pathway generally respond to 6PGDH knockout by rerouting through the ED pathway and by reversing the direction of the non-oxPPP, slightly increasing the flux through

the EMP pathway, and activating malic enzyme [220, 221]. Prokaryotes that lack a functional ED pathway display a similar response, but instead of rerouting the flux through the ED pathway, they increase flux through the EMP pathway [204]. Thus, in general, prokaryotes can compensate for disruption of G6PDH or 6PGDH.

Isocitrate dehydrogenase

Isocitrate dehydrogenase (IDH, EC:1.1.1.42) is a component of the TCA cycle that catalyzes the decarboxylation of isocitrate to 2-oxoglutarate (also known as α -ketoglutarate), with the release of CO₂ and NADPH (**Figure 2.3**). The IDH reaction is important for the generation of reducing power. Through the generation of 2-oxoglutarate, the reaction also links nitrogen and carbon metabolism and plays an important role in the cellular defense against oxidative damage and detoxification of ROS [222-224]. Together with isocitrate lyase, IDH is a branching point between the TCA cycle and the glyoxylate shunt (**Figure 2.3**), a pathway needed for growth on non-fermentative carbon sources such as acetate and ethanol. IDH activity is important in controlling the metabolic flux between both pathways and is affected by various regulatory factors, such as metal ions [225, 226], PEP concentration [227], and phosphorylation/dephosphorylation [228, 229]. In addition, IDH is a component of the reductive tricarboxylic acid (RTCA) cycle. The RTCA cycle is a CO₂ fixation pathway, present in some bacteria and archaea, in which four molecules of CO₂ are fixed to produce one molecule of oxaloacetate [230-232]. Obviously, IDH cannot act as a source of NADPH under these conditions.

In contrast to eukaryotes, prokaryotes are generally thought to possess only NADP⁺-dependent IDHs. Although this is true for most prokaryotes, a small but increasing number of reports have described bacterial and archaeal NAD⁺-dependent IDHs, such as the one found in *Hydrogenobacter thermophilus* [233-241]. Moreover, a survey of sequenced prokaryotic genomes by Zhu *et al.* identified nine additional species that likely contain NAD⁺-dependent IDHs [242]. However, although the IDH of *H. thermophilus* was originally described as an NAD⁺-dependent IDH, Aoshima and Igarashi later showed it was not a conventional decarboxylating IDH but a novel non-decarboxylating enzyme that catalyzes

the conversion between isocitrate and oxalosuccinate. They therefore proposed that the enzyme should not be categorized as an IDH [243].

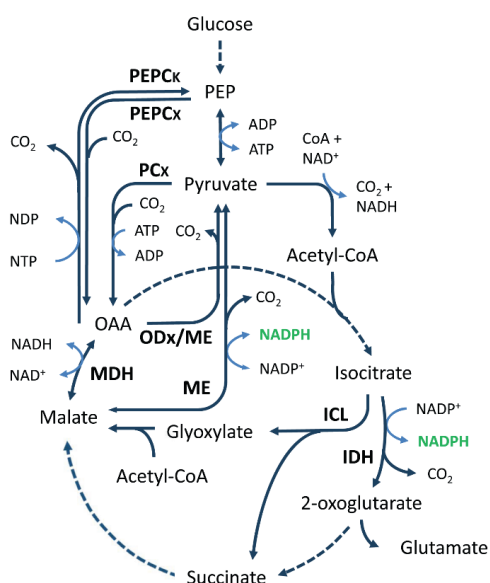
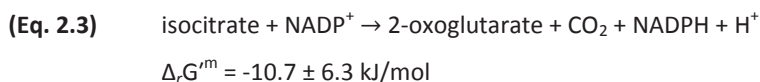


Figure 2.3 General overview of the TCA cycle and anaplerotic node. Solid and dashed arrows represent single and lumped enzymatic reactions, respectively. Malic enzyme is able to catalyze the decarboxylation of malate and OAA. Abbreviations for metabolites (normal text) and enzymes (bold text): PEP, phosphoenolpyruvate; OAA, oxaloacetate; PEPCx, phosphoenolpyruvate carboxylase; PCx, pyruvate carboxylase; ODx, oxaloacetate decarboxylase; MDH, malate dehydrogenase; ME, malic enzyme; ICL, isocitrate lyase; IDH, isocitrate dehydrogenase.

With the exception of the few NAD^+ -preferring IDHs, prokaryotic IDHs generally catalyze the following decarboxylating reaction:



In addition to the cofactor NADP^+ , IDH also needs divalent cations, particularly Mg^{2+} or Mn^{2+} . Although most IDHs have a strong requirement for either Mg^{2+} or Mn^{2+} , substitution of other divalent cations, or even univalent ions such as Na^+ , is sometimes possible. However, in most cases, this leads to little or no activity [225, 226, 244, 245].

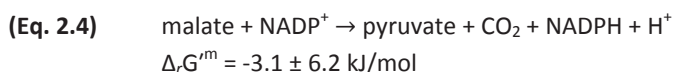
Like G6PDH and 6PGDH, IDH is an important source of NADPH [141, 218]. In wild-type *Escherichia coli* aerobically grown in chemostats under glucose- or ammonia-limited conditions, the IDH reaction was the major producer of NADPH, accounting for more than 60% of the total NADPH production [218]. Accordingly, overexpression of IDH resulted in a 31% increase in the production of GDP-L-fucose, whose formation is NADPH-dependent. While this is a significant improvement, the biosynthesis of GDP-L-fucose in the IDH overexpression strain was not better than that in a G6PDH overexpression strain [246]. This suggests that, under the applied conditions, IDH is not a better option than G6PDH for NADPH regeneration in biotransformation reactions [122]. The ability of IDH overexpression to enhance glutamate formation in *Corynebacterium glutamicum* has also been tested [247]. *C. glutamicum* secretes glutamate under certain conditions. Because IDH supplies 2-oxoglutarate, the direct precursor of glutamate, overexpression of IDH was an obvious choice. However, overexpression of IDH, even in combination with glutamate dehydrogenase, did not enhance glutamate production.

Disruption of IDH has also been investigated. The exact effects of IDH disruption were dependent on the specific conditions and species used, but some general trends were apparent [247-251]. First, all IDH mutants investigated were glutamate auxotrophs. Second, although glutamate-containing cultivation media sustained the growth of IDH mutants, their growth was slower than that of the parental strains. The reduced specific growth rate was most likely due to a lower ATP/ADP ratio and a lower NADPH/NADP⁺ ratio, as shown in an *E. coli* IDH disruption strain grown on glucose under aerobic conditions [249]. Moreover, other significant effects observed in the mutant *E. coli* strain included upregulation of oxPPP enzymes and the anaplerotic glyoxylate pathway. Similarly, redirection from the TCA cycle to anaplerosis has also been reported for an IDH knockdown of *C. glutamicum* [248]. However, with respect to citrate synthase, contradictory results have been obtained. Kabir and Shimizu [249] found upregulation of citrate synthase, but an earlier study of *E. coli* IDH mutants [252] and a study of *Rhizobium meliloti* [250] found that IDH mutants spontaneously gave rise to IDH-citrate synthase double mutants.

Malic enzyme

The TCA cycle-associated malic enzyme (ME, EC:1.1.1.40) catalyzes the NAD(P)⁺-dependent oxidative decarboxylation of malate to pyruvate (**Figure 2.3**). Confusingly, malic enzyme is also known as “malate dehydrogenase (oxaloacetate-decarboxylating),” while true malate dehydrogenases, enzymes that catalyze the non-decarboxylating reaction between malate and oxaloacetate (**Figure 2.3**), are sometimes referred to as malic enzyme. In addition, various classes of ME that differ in their preference for NAD⁺ or NADP⁺ and in their ability to decarboxylate oxaloacetate have also been characterized [253]. Enzymes in the first ME class (EC:1.1.1.38) use NAD⁺ and can decarboxylate oxaloacetate as well as malate. Enzymes in the second class (EC:1.1.1.39) also prefer NAD⁺, but they are unable to decarboxylate oxaloacetate. Enzymes in the third class (EC:1.1.1.40) are NADP⁺-dependent; they can catalyze the decarboxylation of malate and oxaloacetate. Moreover, MEs that do not fit the general classification scheme exactly, such as those from the gram-negative bacterium *Rhizobium meliloti* and the hyperthermophilic archaeon *Thermococcus kodakarensis*, have been described as well [86, 254]. The latter, for example, prefers NADP⁺ but, unlike normal NADP⁺-dependent MEs, is not able to decarboxylate oxaloacetate.

Although malic enzymes (MEs) are widely distributed amongst prokaryotes (**Table 2.1**), most ME-containing species possess an ME that belongs to the third class [255, 256] and catalyzes the following reaction:



However, prokaryotes without an ME or with only an NAD⁺-dependent ME, such as lactic acid bacteria [257-260], or prokaryotes with both NAD⁺- and NADP⁺-dependent MEs have also been characterized [254-256, 261, 262]. Moreover, Morimoto *et al.* recently showed that directed evolution could easily alter the cofactor preference of an ME from *T. kodakarensis* [263]. Thus, NAD⁺/NADP⁺ specificity seems to depend only on a few specific, highly conserved residues [86, 255, 264-267].

All MEs require a divalent cation as a cofactor. Maximum activity is generally observed in the presence of Mn²⁺, followed closely by Mg²⁺. Divalent cations such as Co²⁺ and Ni²⁺ are usually

only partially able to replace Mn^{2+} , while others such as Ca^{2+} , Cu^{2+} , or Sr^{2+} cannot replace Mn^{2+} at all [86, 268, 269]. In addition, the MEs of various prokaryotes also require a monovalent cation for activity, preferably NH_4^+ and K^+ [258, 268, 270-272].

In many aerobic and facultative anaerobic organisms, the TCA cycle-associated ME is part of the PEP-pyruvate-oxaloacetate node, also referred to as the anaplerotic node (**Figure 2.3**). This node represents the metabolic link between glycolysis/gluconeogenesis and the TCA cycle [256, 273]. As such, ME is involved in the interconversion of C4 and C3 compounds, which is important for maintaining the levels of TCA cycle intermediates (anaplerotic reactions) and for growth on C4 and C3 compounds and substrates that enter central metabolism via acetyl-CoA, such as acetate, fatty acids, and ethanol. Although ME can catalyze C3-carboxylation and C4-decarboxylation reactions, it is generally involved in the latter (decarboxylation of malate to pyruvate with the concomitant formation of NADPH) [254-256, 268, 269]. However, examples of prokaryotic MEs without a preference for either [86] and MEs with a preference for the C3-carboxylation reaction have been identified [274]. To determine the exact contribution of ME to the cellular NADPH pool, the physiological direction and flux through the enzyme needs to be established. However, regulation of the carbon flux at the anaplerotic node involves a complex interplay between the enzymes involved and depends on the growth conditions and species [183, 275-277]. A detailed discussion of this complex regulation is beyond the scope of this review, but comprehensive reviews, such as those by Owen *et al.* [273] and Sauer and Eikmanns [256], are available.

To understand the physiological role of ME in the metabolism of prokaryotes, the disruption of ME and alternative pathways has been studied. The significance of the latter is clearly demonstrated by the absence of a detectable phenotype in *Campylobacter jejuni* and *E. coli* ME disruption strains [278, 279]. The lack of a clear phenotype can be explained by the presence of alternative pathways and enzymes, such as the malate dehydrogenase-phosphoenolpyruvate carboxykinase pathway (mdh-PEPCK) and isocitrate dehydrogenase, that can fulfill the pyruvate/PEP and NADPH requirements, respectively (**Figure 2.3**) [269, 279-281]. Because of the wide distribution of such alternatives in prokaryotes, it is likely that ME gene disruption in other species will yield similar results. Nevertheless, species in which

ME gene disruption yields a phenotype, such as *Corynebacterium glutamicum*, *Bacillus subtilis*, and various lactic acid bacteria, have been identified [255, 257, 259, 268].

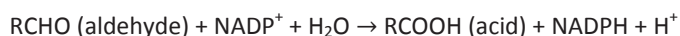
Analysis of various disruption strains created to understand the physiological role of ME in prokaryotes has shown that it is crucial to take the genetic background of a strain into consideration because the exact physiological role of ME differs from species to species [259, 278, 279, 282, 283]. Analysis has also shown that the effect of ME disruption can be subtle and/or heavily dependent on the cultivation conditions [255, 268]. A *C. glutamicum* ME disruption strain, for example, showed no detectable phenotype during growth on either acetate or glucose, but showed significantly modified growth behavior during lactate metabolism [268]. Moreover, the physiological role of NAD⁺- and NADP⁺-dependent MEs appears to be different [255, 257, 259, 261, 269, 282]. Whereas NAD⁺-dependent MEs generally play a role in malate catabolism, NADP⁺-dependent MEs function either as gluconeogenic enzymes, by supplying pyruvate from C₄-dicarboxylic acids, or as NADPH-generating systems needed for various biosynthetic purposes. It is interesting in this respect that the concerted action of ME, MDH, and pyruvate carboxylase (PCx) can act as a kind of transhydrogenase. The combined action of these enzymes enables the ATP-dependent conversion of pyruvate to OAA and malate and back to pyruvate, while converting NADH into NADPH (**Figure 2.3**) [256]. A similar reaction sequence can involve PEP carboxylase (PEPCx) instead of pyruvate carboxylase, which also enables conversion of NADH into NADPH. This so-called malate shunt was recently described in *Clostridium thermocellum* [284].

NAD⁺- and NADP⁺-dependent MEs have also been used in metabolic engineering approaches to improve specific biotechnological applications. However, to our knowledge, no prokaryotic ME has been used to increase intracellular NADPH levels, as has been done in the yeast *Saccharomyces cerevisiae* [285, 286]. Instead, prokaryotic MEs have been overexpressed for the formation of C₄-dicarboxylic acids from glucose or pyruvate [287-290]. These acids are important intermediates in the production of tetrapyrroles [288] and amino acids [291], for example, and they can be chemically converted into common intermediates in the petrochemical industry. Hence, C₄-dicarboxylic acids are important intermediates in

the production of commodity chemicals from renewable carbohydrate feedstocks [292, 293]. However, to produce these acids from glucose or pyruvate, one has to reverse the physiological direction of ME. This has been accomplished in different ways [287, 289, 290].

NADP⁺-dependent glyceraldehyde-3-phosphate dehydrogenases

NADP⁺-dependent glyceraldehyde-3-phosphate dehydrogenases belong to the aldehyde dehydrogenase (ALDH) superfamily, a divergently related group of enzymes widely distributed among all three domains of life [294, 295]. Enzymes in this group metabolize a broad spectrum of endogenous and exogenous aldehydes, which are oxidized to the corresponding carboxylic acid, using NAD⁺ or NADP⁺ as a cofactor [296, 297]:



As such, the enzymes play essential roles in processes such as intermediary metabolism and detoxification [298, 299]. Although many ALDHs are NADP⁺-dependent, only the NADP⁺-dependent glyceraldehyde-3-phosphate dehydrogenases will be addressed here, because these enzymes contribute significantly to NADPH levels in certain prokaryotes. Moreover, extensive reviews of the ALDH superfamily are available [294, 295, 300, 301].

While most glyceraldehyde-3-phosphate dehydrogenases (GAPDH, EC:1.2.1.12) are NAD⁺-dependent enzymes responsible for the reversible oxidation of GAP into 1,3-bisphosphoglycerate (1,3-BPG) as part of the classical EMP pathway, NADP⁺-dependent variants exist as well. In fact, there are two different types of NADP⁺-dependent GAPDHs, a phosphorylating type (NADP⁺-GAPDH, EC:1.2.1.13), producing 1,3-BPG, and a non-phosphorylating type (GAPN, EC:1.2.1.9), unidirectionally producing 3-phosphoglycerate (3-PG) (**Figure 2.4**). However, the term NADP⁺-GAPDH is used inconsistently in different studies to denote either the phosphorylating or non-phosphorylating type or both types, thus generating confusion.

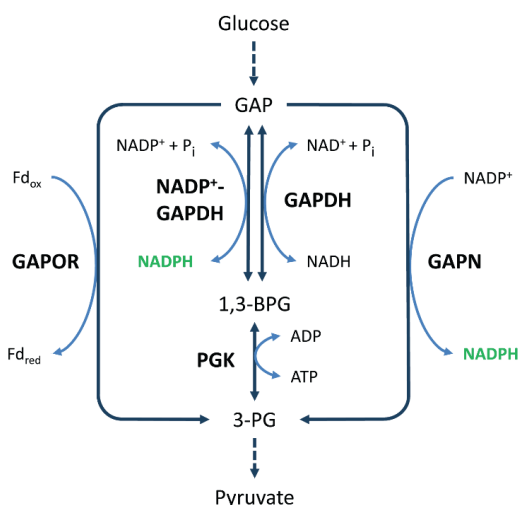
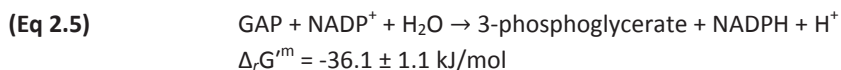


Figure 2.4 NADPH generation by GAP dehydrogenases. Solid and dashed arrows represent single and lumped enzymatic reactions, respectively. Abbreviations for metabolites (normal text) and enzymes (bold text): GAP, glyceraldehyde 3-phosphate; 1,3-BPG, 1,3-bisphosphoglycerate; 3-PG, 3-phosphoglycerate; GAPOR, GAP:ferredoxin oxidoreductase; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; GAPN, non-phosphorylating GAPDH; PGK, phosphoglycerate kinase.

2

GAPN was originally thought to be present only in photosynthetic eukaryotes [302, 303]. However, early reports of GAPN activity in *Streptococcus mutans* were later confirmed by the characterization of the first prokaryotic GAPN [304, 305]. Since then, GAPN has also been found in various (hyper)thermophilic archaea, such as *Thermococcus kodakarensis*, *Thermoproteus tenax*, and *Sulfolobus solfataricus* [77, 306, 307], and several gram-positive bacteria, such as *Clostridium acetobutylicum*, *Streptococcus pyogenes*, and *Bacillus halodurans* (**Table 2.1**) [142, 295, 308-310]. GAPN catalyzes the irreversible, non-phosphorylating oxidation of GAP to 3-PG with the concurrent reduction of NADP^+ to NADPH:

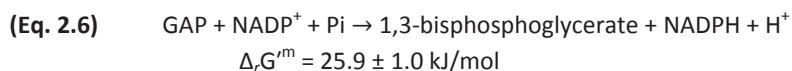


In this reaction, 1,3-BPG is not formed, and the oxidation of GAP is not coupled to the generation of ATP (**Figure 2.4**). Such modifications of the classical EMP or ED pathway are regularly found in archaea [77, 78, 307, 311, 312]. Interestingly, GAPN has been reported to have a key function in the regulation of these modified pathways [306, 307, 313, 314].

With the exception of the GAPN in *T. tenax*, which has a preference for NAD^+ , prokaryotic GAPNs, in general, have a clear preference for NADP^+ . However, although the GAPN of *T. tenax* was shown to be NAD^+ -dependent, it was later shown to be allosterically regulated, shifting its cofactor preference from NAD^+ to NADP^+ in the presence of activators such as F6P and G1P [314]. Based on sequence similarity, similar allosteric regulation might also exist for GAPNs in a number of other hyperthermophiles [314, 315]. Aside from NADP^+ , GAPNs do not require other cofactors or metals.

While the physiological significance of GAPN might vary between organisms, the enzyme is an important and essential source of NADPH in the bacterium *S. mutans* and the archaeon *T. kodakarensis*, respectively [77, 304, 316]. Moreover, the ability to use GAPN as a source of NADPH has been demonstrated in a mutant *E. coli* strain in which the endogenous NAD^+ -dependent glyceraldehyde-3-phosphate dehydrogenase (NAD^+ -GAPDH) was replaced by the GAPN of *S. mutans*. Transcriptional analysis of this strain revealed upregulation of the transhydrogenase and downregulation of the PPP and TCA cycle, which together indicate a response to avoid NADPH excess [317]. In addition, GAPN has been used successfully in metabolic engineering efforts intended to improve the production of metabolites whose biosynthesis depends on NADPH, such as L-lysine or poly-3-hydroxybutyrate, or to decrease the formation of undesired by-products such as glycerol in industrial ethanol production strains [318-320].

In contrast to GAPN, NADP^+ -GAPDH produces 1,3-BPG. It is therefore more similar to NAD^+ -GAPDH (**Figure 2.4**).



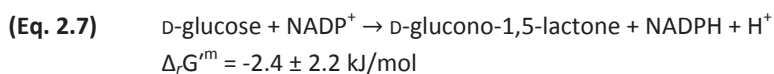
The generation of 1,3-BPG is an important difference between the reactions catalyzed by NADP^+ -GAPDH and GAPN, because further oxidation of 1,3-BPG to 3-PG generates ATP. Moreover, while the oxidation of GAP by NADP^+ -GAPDH is reversible, the oxidation by GAPN is not. Hence, NADP^+ -GAPDH, in contrast to GAPN, is active in both glycolysis and gluconeogenesis. However, in prokaryotes that also contain alternative GAP oxidation enzymes, such as NAD^+ -GAPDH, GAPN, or GAP:ferredoxin oxidoreductase (GAPOR) (**Figure**

2.4), NADP⁺-GAPDH is primarily involved in gluconeogenesis and does not appear to play a role in generating NADPH [77, 315, 321-324].

As with GAPN, the significance of NADP⁺-GAPDH as a source of NADPH has been demonstrated. Overexpression of NADP⁺-GAPDH or replacement of the native NAD⁺-GAPDH with the NADP-variant improved the yield of NADPH-dependent products such as lycopene, ϵ -caprolactone, L-ornithine, or coenzyme Q₁₀ [325-328]. Moreover, metabolic flux analysis of the exchange mutant revealed that the oxPPP branch and TCA fluxes were significantly reduced, presumably to avoid NADPH excess [327]. In addition, an attempt in *C. glutamicum* to alter the cofactor specificity of native NAD⁺-GAPDH to NADP⁺ through protein engineering resulted in a 60% increase in lysine production [329].

NADP⁺-dependent glucose dehydrogenase

Similar to GAPN and GAPDH, other non-canonical NADP⁺-dependent enzymes, such as NADP⁺-dependent glucose dehydrogenase (NADP⁺-GDH), might make a larger contribution to NADPH generation than previously thought. Glucose dehydrogenases (GDHs, EC:1.1.1.47 and EC:1.1.1.119) are responsible for the first step of the modified ED pathways found in some archaea and bacteria (**Table 2.1**), i.e., the semiphosphorylated ED, the nonphosphorylated ED, and the branched ED pathways (**Figure 2.5**) [78]. GDHs are responsible for a main difference between modified ED pathways and the classical ED pathway. GDHs catalyze the direct oxidation of glucose into gluconolactone. In contrast, in the classical ED, glucose is phosphorylated before its oxidation into 6-P-gluconolactone [78].



Two different classes of GDHs are capable of using NADP⁺ as cofactor: EC:1.1.1.47, which shows dual cosubstrate specificity with a preference for NADP⁺, and EC:1.1.1.119, which is strictly dependent on NADP⁺. GDHs belonging to these groups have been found in extremely halophilic and thermoacidophilic archaea, as well as in bacteria such as *Bacillus subtilis*, *Bacillus megaterium*, and *Gluconobacter oxydans* [78, 330-332].

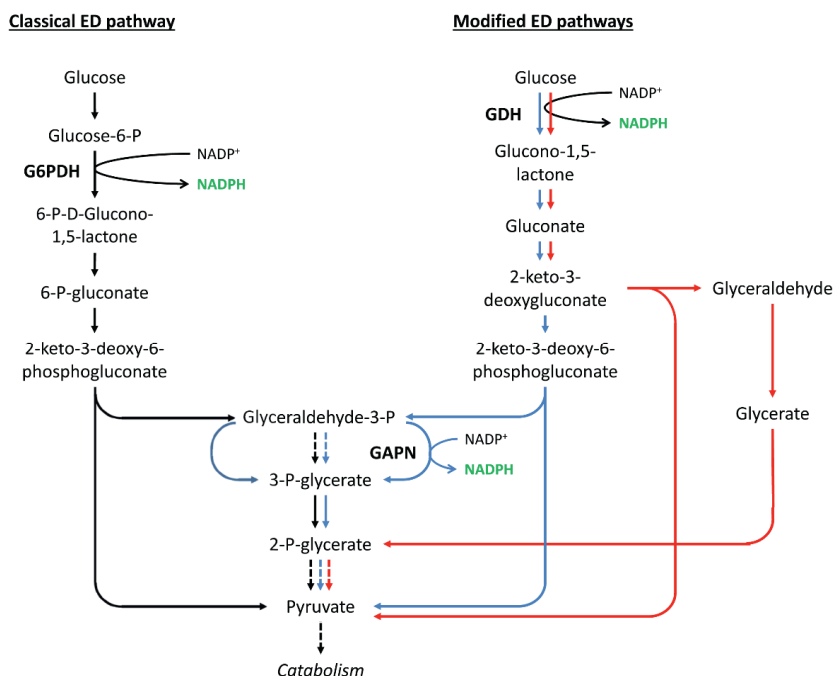


Figure 2.5 Classical ED pathway compared to modified ED pathways. Solid and dashed arrows represent single and lumped enzymatic reactions, respectively. Shown are the classical ED pathway (in black), the semiphosphorylated ED pathway (in blue), the nonphosphorylated ED pathway (in red), and the branched ED pathway that includes the semi- and nonphosphorylated ED pathways. Only the glucose degradation directions are shown. Abbreviated enzymes (in bold text): G6PDH, glucose-6-phosphate dehydrogenase; GAPN, non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase; GDH, glucose dehydrogenase.

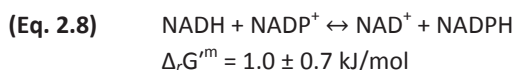
NADP⁺-GDHs have been used for NADPH regeneration in various production systems, such as the *in vitro* synthesis of poly(3-hydroxybutyrate) or L-leucovorin [333, 334]. In addition, NADP⁺-GDHs are important components of many commercial glucose-sensing assays. As such, they have been utilized for blood glucose monitoring [335]. Moreover, NADP⁺-GDHs have also been used as NADPH regenerators in various whole-cell, or *in vivo*, production systems. The GDHs of *B. megaterium* and *B. subtilis*, for example, have been expressed in several *E. coli* strains for the synthesis of various products, including indigo, S-sulfoxide, and ethyl (R/S)-4-chloro-3-hydroxybutanoate [213, 336-342].

NADPH-generating reactions not coupled to carbon metabolism

Transhydrogenases

Pyridine nucleotide transhydrogenase (TH) directly catalyzes the reversible hydride transfer between NAD(H) and NADP(H). There are two different isoforms: the energy-independent soluble transhydrogenase (STH, EC:1.6.1.1) and the energy-dependent, or proton-translocating, membrane-bound transhydrogenase (H^+ -TH, EC:1.6.1.2) (**Figure 2.6**). Soluble transhydrogenases from *Gammaproteobacteria* such as *Escherichia coli*, *Azotobacter vinelandii*, *Pseudomonas fluorescens*, and *Pseudomonas aeruginosa* have been studied in some detail [343-346]. In contrast, proton-translocating, membrane-bound transhydrogenases have been investigated extensively, reviewed in [122, 347-352]. They are widely distributed in the mitochondria of eukaryotes and in certain bacteria, but are very rare in archaea (**Table 2.1**) [350, 351]. Bacteria with multiple copies of the transhydrogenase genes are not uncommon, and some species, particularly among the *Enterobacteriaceae*, contain both isoforms [141].

The two transhydrogenase isoforms transfer electrons between NAD(H) and NADP(H) in the following reaction:



However, in contrast to STHs, H^+ -THs couple the electron transfer reaction to the translocation of a proton across the membrane, according to the reaction:



Where “out” and “in” denote the periplasmic space and cytosol, respectively. Although the reaction is in principle reversible, the electrochemical proton gradient (Δp) across the membrane under most physiological conditions strongly favors NADPH formation [348, 349, 352, 353].

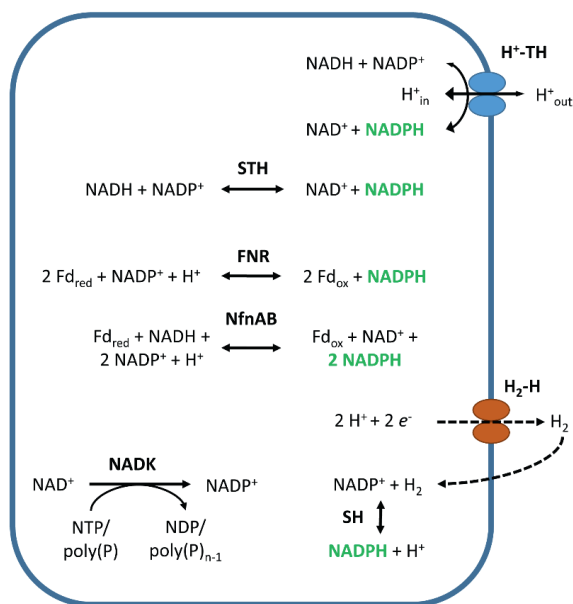


Figure 2.6 NADPH-generating reactions not coupled to carbon metabolism.

Abbreviated enzymes (in bold): H₂-H, hydrogen-evolving hydrogenase; SH, cytosolic NADP⁺-reducing hydrogenase; NfnAB, electron-bifurcating NADH-dependent reduced ferredoxin:NADP⁺ oxidoreductase; FNR, ferredoxin:NADP⁺ oxidoreductase; STH, energy-independent soluble transhydrogenase; H⁺-TH, energy-dependent or proton-translocating, membrane-bound transhydrogenase.

Since the discovery of transhydrogenases, there has been much debate about their physiological role [176, 352-356]). Given the type of reaction they catalyze, transhydrogenases might play an important role in maintaining the redox balance within microorganisms. Various studies have shown that several bacterial species require STH for growth under metabolic conditions with excess NADPH formation [141, 181, 218, 357, 358]. Moreover, mutant bacteria suffering from a disturbed redox balance due to an increased NADPH concentration often benefit from the heterologous expression of STH [359-362]. These observations support the idea that the physiological role of STH is to convert NADPH into NADH to prevent an excess of NADPH [356], a notion that is generally accepted and is supported by actual NAD⁺/NADH and NADP⁺/NADPH ratios (see above), which indicate that the formation of NADPH from NADH will generally require energy input.

However, Sanchez *et al.* have reported that the NADPH-requiring production of poly(3-hydroxybutyrate) in *E. coli* benefitted from the overexpression of STH [363]. This result is contrary to the generally accepted physiological role of STH and suggests that the enzyme catalyzes the opposite reaction. Moreover, recent reports have shown that STH is able to catalyze the opposite reaction and that overexpression of the enzyme can improve NADPH-

dependent production yields [364-367]. Additionally, although STH disruption and overexpression mutants often have disturbed or improved redox balances, respectively, similar mutants that do not show any effect have also been described [368, 369]. Together, these examples show that we have yet to achieve a thorough understanding of the physiological role of STH and that its role is strongly dependent on species, culture conditions, and genotypes.

Similarly, the physiological role of H^+ -TH has been debated in the literature. It has often been reported that the H^+ -TH isoform is a major source of NADPH, needed for the biosynthesis of amino acids, for example, or the reduction of glutathione (required to minimize oxidative damage caused by free radicals generated in the respiratory chain) [141, 181, 218, 353, 354, 356, 365, 370, 371]. In *E. coli*, H^+ -TH has been shown to provide about 40% of the total NADPH during growth on glucose in batch cultures [141]. Moreover, its expression is induced when there is demand for NADPH, and it is required for optimal growth on carbon sources whose metabolism does not directly generate NADPH [141, 371]. Additionally, Fuhrer and Sauer calculated that it is very unlikely, from a thermodynamic point of view, that H^+ -THs catalyze NADPH oxidation. They concluded that the $NADP^+$ reduction catalyzed by H^+ -THs is irreversible under physiological conditions in various bacterial species [181]. Reports such as these contribute to the general view that the physiological role of H^+ -TH is reduction of $NADP^+$ at the expense of NADH. However, other reports suggest that H^+ -TH catalyzes the opposite reaction or show that disruption or overexpression has no measurable effect [366, 368, 369].

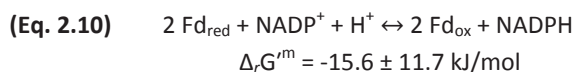
THs require the cofactors NAD(H) and NADP(H) for activity. In addition, and in contrast to H^+ -THs, STHs also require a flavin (flavin adenine dinucleotide, FAD) cofactor [343, 372, 373]. THs generally do not require other cofactors or metal ions. In fact, metal ions have even been reported to inhibit TH activity [372, 374, 375]. Nevertheless, submillimolar concentrations of Ca^{2+} and Mg^{2+} have been reported to stimulate the H^+ -THs of *Rhodobacter capsulatus* and *E. coli* [376-378]. Similarly, saturating concentrations of Ca^{2+} have been reported to release the phosphate-induced inhibition of the $NADP^+$ reducing reaction and to

cause full activation of the enzyme in *Azotobacter vinelandii* and *Pseudomonas aeruginosa*, respectively [379, 380].

Because of their ability to generate NADPH, STHs and H⁺-THs have been overexpressed to improve the production yields of NADPH-dependent metabolites. For example, overexpression of STH increased NADPH-dependent poly(3-hydroxybutyrate) production by 82% in an engineered *E. coli* strain [363]. Moreover, the overexpression of STH in two other engineered *E. coli* strains increased (S)-2-chloropropionate and thymidine production 150% and 100%, respectively [366, 367]. Similarly, overexpression of H⁺-TH has significantly improved the yields of various NADPH-dependent products, such as chiral alcohols, isobutanol, and 3-hydroxypropionic acid in *E. coli* [381-384] and L-valine and L-lysine in *C. glutamicum* [187, 385].

Ferredoxin:NADP⁺ oxidoreductase

Ferredoxin:NADP⁺ oxidoreductases (FNRs, EC:1.18.1.2) are ubiquitous flavoenzymes (**Table 2.1**) that catalyze the reversible transfer of reducing equivalents between the one-electron carrier ferredoxin (Fd) and the two-electron carrying NADP(H), according to the following reaction (**Figure 2.6**):



FNRs are present in all three domains of life [83, 386-388]. They are mainly known for their essential role in photosynthetic organisms (plants, algae, and cyanobacteria), in which they connect the Fd_{red}-generating reactions of photosynthesis to NADPH-requiring carbon assimilation [389, 390]. In addition to enabling the flux of electrons from the photosystem into other metabolic pathways, FNRs might also avert this efflux of electrons and mediate the return of Fd-bound electrons to the plastoquinone pool of the photosynthetic system. By doing so, FNRs contribute to the so-called cyclic electron flow, which is essential for balancing the ATP/NADPH ratio generated by the photosystems [389, 391-394].

In non-photosynthetic tissues and organisms, FNRs participate in the electron transfer chains of various other metabolic processes such as nitrogen fixation, sulfate assimilation,

isoprenoid biosynthesis, osmotic and oxidative stress responses, and iron-sulfur cluster biogenesis [395-402]. However, in contrast to the FNRs involved in photosynthesis, the physiological direction of FNRs involved in these processes is toward the production of Fd_{red} . As a result, FNRs are sometimes classified as autotrophic (all photosynthetic FNRs) and heterotrophic (all other FNRs) [395], which, from a protein-family point of view, is not entirely correct.

FNRs are generally classified into two phylogenetically and structurally unrelated protein families that are subdivided into four subclasses: the plant-type FNRs comprising the plastidic-type and bacterial-type subclasses and the glutathione (GR)-type FNRs comprising the mitochondrial adrenodoxin reductases (AdR-like) and oxygenase-coupled NADH:ferredoxin reductases (ONFR-like) [386, 403-406]. In addition, a third type of FNR has been reported, the so-called thioredoxin reductase-like FNR (TRLF), which functions as an FNR but has structural homology to NADPH-dependent thioredoxin reductases [388, 407-410]. Interestingly, a novel type of FNR has recently been reported, the electron-bifurcating NADH-dependent reduced ferredoxin:NADP⁺ oxidoreductase (NfnAB), which couples the exergonic reduction of NADP⁺ with Fd_{red} and the endergonic reduction of NADP⁺ with NADH in a reversible reaction (**Eq. 2.11** and **Figure 2.6**) [411-413]. Genome analyses have revealed the *NfnAB* genes are present in many anaerobic bacteria and archaea [411, 414], a finding that might reflect the importance of a postulated NfnAB function: NfnAB and other electron-bifurcating reactions have been implicated as a third mode of energy conservation, in addition to substrate level phosphorylation and electron transport phosphorylation [414].



Despite what the names imply, both subdivisions of the plant-type FNRs (plastidic-type and bacterial-type) as well as FNRs of the GR-type, TRLF-type, and “NfnAB-type” are found in prokaryotes [386, 403, 407]. Unfortunately, the literature does not always make clear the class to which an FNR belongs, generating confusion and making it difficult to judge the true NADPH-generating potential of FNRs found in different prokaryotes. Nevertheless, under normal physiological conditions, it appears that only the plastidic-type FNRs, found in

cyanobacteria, for example, and the novel “NfnAB-type” have significant effects on NADPH availability in prokaryotes.

The reaction catalyzed by the different FNR types is in principle reversible, as demonstrated by various *in vitro* studies, as well as by *in vivo* studies in cyanobacteria and the thermophilic bacterium *Hydrogenobacter thermophilus* TK-6. Cyanobacteria contain a single FNR that is responsible for NADP⁺ reduction in vegetative cells as well as for Fd reduction in heterocysts [415]. A similar situation exists in *H. thermophilus* TK-6, where a single FNR catalyzes either the forward or reverse reaction depending on the specific type of Fd involved in the reaction. A [4Fe-4S]-containing Fd is involved in the forward (NADP⁺ reducing) reaction, whereas a [2Fe-2S]-containing Fd is involved in the reverse reaction [416]. Such insights might facilitate the development of rational engineering approaches aimed at gaining directional control of the reaction. Elucidating the reaction mechanisms of various FNR-types is a topic of many studies and has been reviewed extensively [395, 403, 404, 406, 417-421].

FNRs, in general, contain a noncovalently bound flavin (FAD) cofactor as a prosthetic group, but a flavin mononucleotide (FMN)-containing FNR has also been characterized [416]. The FAD group can alternate between an oxidized, one-electron-reduced semiquinone and a fully reduced hydroquinone state, thereby mediating the reversible electron change between the one-electron carrier Fd and the two-electron carrying NADP(H) [404]. The iron-sulfur-containing Fd can sometimes be substituted by the flavin mononucleotide (FMN)-containing flavodoxin (under iron-depleted conditions in some bacteria) or, in the case of the GR-type FNRs, by the mitochondrial iron-sulfur-containing adrenodoxin [386, 422-424]. Moreover, a novel rubrerythrin-like protein, named ferriperoxin (Fpx), has been described that exhibits NADPH- and FNR-dependent peroxidase activity in the thermophilic hydrogen-oxidizing bacterium *H. thermophilus* [425]. Whereas most FNRs are NADP(H)-dependent, an NAD(H)-dependent variant also exists, the ONFR-like subtype of the GR-type FNRs (EC:1.18.1.3) [403]. However, because this group of FNRs is not involved in NADPH availability, it will not be discussed further in this review.

NADP⁺-dependent hydrogenase

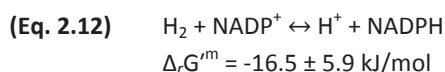
Hydrogenases (H₂ases) are metalloenzymes that catalyze the reversible oxidation of hydrogen gas into two protons and two electrons:



H₂ases enable organisms to utilize H₂ as a source of reducing power or to use protons as terminal electron acceptors, generating H₂. According to the metal atoms at their active site, H₂ases can be classified into three types: [NiFe]-H₂ases, [FeFe]-H₂ases, and [Fe]-H₂ases. The latter were formerly known as metal-free or iron-sulfur-free H₂ases; they are involved in CO₂ reduction with H₂ to methane and are only found in methanogenic archaea [101, 426-430]. [NiFe]-H₂ases are generally responsible for H₂ oxidation and are widely distributed among both aerobic and anaerobic bacteria and archaea; they have been identified in 28.9% and 57.8% of the available complete bacterial and archaeal genomes, respectively. [FeFe]-H₂ases, in contrast, are generally responsible for H₂ evolution and are only found in anaerobic bacteria; they have been identified in 9.8% of the available complete bacterial genomes [427]. Alignment of the sequence motifs that coordinate the active site separates [NiFe]-H₂ases into at least four, and possibly five, functional groups that are, in general, consistent with the physiological roles of the enzymes [100, 101, 427, 431]. Thus, H₂ases are a diverse group of enzymes. They are cytosolic or membrane bound, involved in H₂ oxidation and/or evolution, and able to use a variety of electron acceptors and donors. It is beyond the scope of this review to address all of the different types of H₂ases. We will focus on those involved in NADPH generation in prokaryotes. For more information about H₂ases in general, the reader is referred to various excellent reviews, such as those by Vignais and Colbeau [101], Shima and Thauer [428] and Peters *et al.* [427].

H₂ases that use the cofactor NADP⁺ as an electron acceptor are not common, but have been identified in some bacteria and archaea (**Table 2.1**). Cytosolic NADP⁺-reducing hydrogenases (EC:1.12.1.3, SH) are present in the sulfate-reducing bacterium *Desulfovibrio fructosovorans* and the hyperthermophilic archaea *Pyrococcus furiosus* and *Thermococcus kodakarensis* [102, 432-434]. Moreover, cytosolic bidirectional NAD(P)⁺-H₂ases and a membrane-bound NAD(P)⁺-H₂ase, known to use both NAD⁺ and NADP⁺ as substrates, have been found in

cyanobacteria and *Klebsiella pneumonia*, respectively [435-437]. Many of these H₂ases recycle the H₂ produced by H₂-evolving H₂ases by coupling the oxidation of H₂ to the reduction of NADP⁺ (**Figure 6**) [82, 83, 434, 436, 438, 439], according to the following reaction:



H₂ recycling might have a distinct energetic advantage because the NADP⁺-reducing H₂ase could provide reductant in the form of NADPH without interfering with the energy balance through electron transport phosphorylation [438]. However, various studies have shown that disruption of NADP⁺-reducing H₂ases has minor effects, suggesting that the role of the enzymes is not that important for overall growth [82, 83, 438, 440-442].

Like most H₂ases, NADP⁺-reducing H₂ases require metal ions and cofactors for their activity. They contain multiple iron-sulfur clusters ([Fe-S]), which are coordinated in the active site by carbon monoxide (CO) and cyanide (CN⁻) groups. Moreover, various H₂ases capable of reducing NADP⁺ belong to the [NiFe]-H₂ases, making the availability of Ni essential as well [102, 103, 427, 443]. In addition, the enzymes have been reported to contain flavin cofactors: FAD in the case of *T. kodakarensis* and *P. furiosus* [102, 103, 444] and FMN in the case of *D. fructosovorans* [432].

The requirement for these different ligands, metals, and cofactors makes the biosynthesis of multisubunit H₂ases a complex and dynamic process, involving various maturation proteins [427, 445-447]. It is thus no surprise that reports about the practical applications of NADP⁺-reducing H₂ases for NADPH generation are limited. Nevertheless, the potential of these enzymes to improve the production yields of NADPH-dependent metabolites has been demonstrated. For example, two prochiral model substrates, acetophenone and (2S)-hydroxy-1-phenyl-propanone, were quantitatively reduced to the corresponding (S)-alcohol and (1R,2S)-diol by an NADPH-dependent alcohol dehydrogenase in an *in vitro* system containing H₂ and the partially purified NADP⁺-reducing H₂ase of *P. furiosus* [448]. Although *in vitro* examples like this demonstrate the potential of NADP⁺-H₂ases, expression of heterologous [NiFe]-H₂ases outside of a closely related host is still challenging [449]. It often

results in an inactive enzyme or activity that is detectable only *in vitro* [450-452]. Nevertheless, successful expression of heterologous [NiFe]-H₂ase has been reported [437].

NAD⁺ kinase and NADH kinase

NAD⁺ kinase (NADK, EC:2.7.1.23) catalyzes the conversion of NAD⁺ to NADP⁺. It is the sole enzyme leading to *de novo* NADP⁺ biosynthesis (**Figures 2.1 and 2.6**). As such, NADK is essential for NADP⁺ and NADPH availability. Together with NADP⁺ phosphatase (NADPase), which is thought to catalyze the opposite reaction (the conversion of NADP⁺ to NAD⁺), NADK is crucial for the regulation of the intracellular balance of NAD(H) and NADP(H) [152, 175]. Although NADPase activity has been detected in various organisms [453, 454]), to our knowledge, no enzyme that functions solely as an NADPase has been isolated to date. However, some enzymes that exhibit NADPase activity in addition to another activity have been isolated, such as the bifunctional NADP⁺ phosphatase/NAD⁺ kinase, discovered in the archaeon *Methanococcus jannaschii* [455, 456].

NADK activity was enriched in extracts from the yeast *Saccharomyces cerevisiae* for the first time in 1950 [174] and has since been purified from various organisms [159, 457]. However, it was not until 2000 that the gene encoding NADK was identified by Kawai *et al.* in *Micrococcus flavus* and *Mycobacterium tuberculosis* H37Rv [167]. Since then, NADKs from various organisms have been studied in detail, and although the exact physiological functions, related pathways, and regulatory mechanisms have not always clearly been elucidated, some reviews of NADK structure, function, and potential applications are available [142, 152, 155, 175, 458].

NADKs are ubiquitous enzymes (**Table 2.1**). Homologs, according to literature, can be found in almost all eukaryotes and prokaryotes, with the exception of the intracellular parasite *Chlamydia trachomatis* [152, 173, 459]. The discrepancy between the distribution shown in **Table 2.1** and that described in the literature might reflect the rapidly increasing availability of prokaryotic genomes and errors inherent to automated genome annotation [460]. Functionally and structurally, NADKs are weakly related to other kinases, suggesting that they belong to a kinase superfamily [461]. With the exception of the NADKs of gram-negative bacteria, the enzymes accept NAD⁺ and its reduced form, NADH [152].

Nevertheless, they are generally referred to as NAD⁺ kinases (EC:2.7.1.23) because they usually have a strong preference for the oxidized form. NADKs with a preference for the reduced form exist as well (EC:2.7.1.86) [462-464], but have not to our knowledge been found in prokaryotes.

Aside from NAD(H), NADKs also need a phosphoryl donor for catalysis. The preferred phosphoryl donor often is a nucleoside triphosphate, particularly ATP, but many bacterial and archaeal NADKs can also utilize the ancient inorganic energy carrier poly(P) [165, 167, 456, 465]. In addition, some bacterial enzymes can utilize ADP or glucose-6-phosphate [466, 467]. However, the use of glucose-6-phosphate as a phosphoryl donor has not been investigated further. Finally, NADKs often also require divalent cations such as Mg²⁺, Mn²⁺, Zn²⁺, or Ca²⁺ for their activity [165-167, 465]. The NADK of *Corynebacterium glutamicum*, however, does not require divalent cations when using ATP as phosphoryl donor, but the enzyme is stimulated in their presence. When using poly(P) as phosphoryl donor, divalent cations are required for activity [467].

According to the phosphoryl donor specificity, NADK orthologs are subdivided into several types. The classification typically applied divides the NADK orthologs into: I) poly(P)/ATP NADKs, which can use poly(P) and ATP as phosphoryl donors and II) ATP NADKs, which are specific for ATP [468, 469]. Poly(P)/ATP NADKs have been identified in gram-positive bacteria such as *Mycobacterium tuberculosis*, *Micrococcus flatus*, and *Bacillus subtilis* [167, 465] and in archaea such as *Methanococcus jannaschii* and *Pyrococcus horikoshii* [165, 456]. ATP NADKs have been found in gram-negative α - and γ -Proteobacteria such as *Sphingomonas* sp. A1, *Salmonella enterica*, and *E. coli* and in eukaryotes [166, 168, 459, 469]. A slightly different classification system is sometimes applied, dividing the NADK orthologs into: I) ATP NADKs, which are strictly dependent on ATP, II) NTP NADKs, which use ATP in addition to other nucleoside triphosphates, and III) poly(P)/NTP NADKs, which use nucleoside triphosphates and poly(P) [175]. The existence of these slightly different classification systems can make it difficult to ascertain the type of NADK being discussed. However, because no strict ATP NADKs (as defined in the latter system) have been identified in

prokaryotes to date, it is not an issue in this review. The reaction catalyzed by NADKs can be described as either:



or



NADK is the sole enzyme leading to *de novo* NADP^+ biosynthesis. It has a central role in the regulation of the $\text{NAD(H)}/\text{NADP(H)}$ balance because it can directly phosphorylate NADH to form NADPH or cooperate with various NADP^+ reducing enzymes to form $\text{NAD(H)}/\text{NADP(H)}$ metabolic networks that ultimately convert NADH into NADPH [117, 175, 470]. In view of the importance of NADK, it is not surprising that disruption of NADK is lethal in various prokaryotes [171, 172, 459, 471-474].

Because of its essential role in NADP^+ biosynthesis and the regulation of the $\text{NAD(H)}/\text{NADP(H)}$ balance, NADKs have attracted attention as a potential strategy by which to improve the biosynthesis of industrially valuable metabolites. Several studies have shown that overexpression of NADK-encoding genes can increase the production yields of various NADPH-dependent biosynthetic pathways, such as poly(3-hydroxybutyrate) production and thymidine production in *E. coli* [367, 475, 476], lysine and isoleucine production in *C. glutamicum* [467, 477], and arginine production in *Corynebacterium crenatum* [478]. These results suggest that increased NADP^+ availability, resulting from NADK overexpression, disturbs the intracellular redox balance, subsequently affecting the expression of NADP^+ -reducing enzymes [367, 475, 476]. However, a comparable attempt to improve isobutanol production in *E. coli* showed that NADK overexpression alone is not always sufficient and that the co-(over)expression of an NADP^+ reducing enzyme might be required to increase NADPH availability [383]. Similarly, combined overexpression of the NADK gene from *E. coli* and the NADP^+ -GAPDH gene from *Bacillus subtilis* in *E. coli* increased 2-chloropropionic acid production yield [328]. In addition to increasing NADPH availability indirectly by overexpressing NADK, a more direct approach has also been applied. For example, the

production yields of GDP-L-fucose and ϵ -caprolactone in *E. coli* were increased by introducing the NADH kinase (EC:2.7.1.86) gene from the yeast *S. cerevisiae* [479].

Final remarks and conclusion

NADPH is an essential reducing agent in all organisms. It plays an important role in various biological processes and is the driving force behind numerous biosynthetic enzymatic reactions of importance to industry. However, efficient NADPH regeneration is often a bottleneck that limits the productivity of such biotransformation processes. A thorough understanding of the reactions involved in the generation of NADPH is thus essential to increase the availability of NADPH and enhance productivity in engineered production strains.

This review has provided an overview of the major canonical NADPH-generating reactions as well as recently described NADPH-generating reactions (**Table 2.1**), and it has summarized how these different reactions have been applied to improve NADPH availability. However, additional NADPH-generating enzymes, such as gluconate 5-dehydrogenase or the recently discovered NADP⁺-specific electron-bifurcating [FeFe]-hydrogenase might also contribute to the overall NADPH supply [480, 481]. Similarly, enzymes generally considered to consume NADPH or enzymes newly developed to regenerate NADPH, such as glutamate dehydrogenase and F420-dependent NADP⁺ reductases or phosphite dehydrogenase, respectively, can occasionally become important NADPH suppliers [71, 122, 482].

Although various NADPH-generating enzymes have been overexpressed successfully, no general strategy that guarantees improved NADPH supply exists. Cells usually contain various integrated, network-wide, redox balancing systems, such that engineering of NADPH-generating pathways can result in unexpected metabolic perturbations. Thus, strategies often need to be optimized individually, taking the culture conditions and genetic background of the strains into consideration. Future research involving a systems biology approach is thus required to improve our understanding of the individual NADPH-generating

enzymes and their effect on various metabolic backgrounds, in order to predict which approaches will increase NADPH availability in various biotransformation processes.

Chapter 3

Introduction of the pentose phosphate pathway in *Thermococcus kodakarensis* improves growth and hydrogen formation

Sebastiaan K. Spaans¹, Thomas J. Santangelo², Haruyuki Atomi³, John van der Oost¹, Servé W. M. Kengen¹

¹ Laboratory of Microbiology, Wageningen University, Wageningen, the Netherlands

² Department of Biochemistry and Molecular Biology, Colorado State University, Fort Collins, CO 80523, USA.

³ Department of Synthetic Chemistry and Biological Chemistry, Graduate School of Engineering, Kyoto University, Katsura, Nishikyo-ku, Kyoto 615-8510, and JST, CREST, 7, Gobancho, Chiyoda-ku, Tokyo 102-0076, Japan

Abstract

Background: Hydrogen (H_2) production by microbial fermentation has attracted widespread attention as a promising alternative to fossil fuels. Unfortunately, however, H_2 yields are not yet economically competitive. Herein, we set out to achieve a proof of principle that H_2 yields can be increased by introduction of heterologous enzymes providing additional substrate oxidation steps. In the absence of alternative reduced end products such as alanine, the PPP-generated additional reductant was predicted to lead to higher H_2 levels. Herein, we introduced bacterial-derived Pentose Phosphate Pathway (PPP) genes into the genome of the hyperthermophilic archaeon *Thermococcus kodakarensis* in an attempt to add extra reductant-generating steps. To force glucose catabolism through the introduced PPP, the endogenous ribulose monophosphate pathway (RuMP) was disrupted simultaneously.

Results: Strain SKS001 and SKS002 were complemented with only the oxidative-PPP (oxPPP) enzymes glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase or with the entire set of missing PPP enzymes, respectively. In-gel activity staining and proteomics confirmed the functional expression of the heterologous genes. Growth and fermentation profile analyses showed that introduction of merely the oxPPP enzymes is sufficient to replace RuMP as a source of pentose phosphates. H_2 production levels were significantly higher in both strains, especially in the early exponential phase where a 96% and 50% increase in the H_2/OD_{600} , and 42% and 33% increase in the H_2 /acetate ratio was observed for strains SKS001 and SKS002, respectively. Moreover, the elevated H_2 /acetate ratios in the early exponential phase suggested that NADPH was responsible for (part of) the increase in H_2 formation. Unexpectedly, significantly higher growth rates and final OD_{600} -values were observed for the mutant strains. This could be a result of an increased supply of NADPH, but improved growth of the parental strain upon addition of nucleosides and proteomics suggested that this could also be due to an enhanced supply of ribulose-5-phosphate.

Conclusions: Introduction of bacterial PPP genes into the archaeon *T. kodakarensis* was successful and resulted in higher H₂ production levels during the early exponential phase. Surprisingly, introduction of the PPP also resulted in significantly higher growth rates and final OD₆₀₀-values. Due to the effect on both growth rate and H₂ yield, H₂ productivity was substantially higher in the oxPPP-containing strains.

Introduction

Hydrogen (H_2) is a commodity chemical, commonly used in various industries as a coolant, O_2 -scavenger or reactant in hydrogenation processes. In addition, hydrogen gas has attracted widespread attention as a promising alternative to fossil fuels [16, 18-20, 22]. However, in contrast to fossil fuels, H_2 is not a primary energy resource but an energy carrier, which implies that it must be produced from other energy sources. Unfortunately, most of the current techniques for H_2 production still rely on fossil fuels in one way or another and can hence not be considered sustainable. Therefore, to shift to a more sustainable production, H_2 must be derived from renewable resources, such as biomass [37-39].

Biological H_2 (Bio H_2) can be produced by microbial dark fermentation. However, some major known drawbacks of this method are that large amounts of side products, such as acids or alcohols, are formed and that H_2 production yields are still too low to make the process economically competitive [43, 44, 47]. This is primarily due to the fact that Bio H_2 is a low-value product, for which costs of feedstock carbohydrates often account for the majority of the price of the products [44, 47]. To enhance the economy of the microbial dark fermentation, it is necessary to improve the sugar-to- H_2 yield.

Current sugar-to- H_2 yields by microbial dark fermentation vary from less than 1 to almost 4 moles of H_2 /mole of hexose consumed [13, 14, 43, 48, 49], the latter corresponding to about a third of the theoretical yield of 12 moles. The production of 4 H_2 /hexose is regarded as the upper limit ("Thauer limit", [51]), of standard metabolisms, and increased yields require further oxidation of substrates to CO_2 . Complete anaerobic oxidation of glucose, however, is thermodynamically difficult and is only possible at very low partial H_2 pressures [48]. Furthermore, dark fermentation has to rely on substrate level phosphorylation (SLP) for its energy conservation, and thus requires sufficient free energy for the ATP coupled reactions (~ 70 kJ/mol) [50, 52, 97]. The consequence of the latter is twofold: i) evolution has optimized fermentation pathways to produce biomass (ATP) and not H_2 , i.e. evolution has optimized for incomplete anaerobic glucose oxidation, and ii) even if an organism containing

the necessary machinery for complete oxidation of sugars was known, complete oxidation of glucose to CO_2 would not be energetically valuable as, according to the known metabolic pathways, no ATP is obtained by SLP. Nevertheless, bioenergetic analyses show that a sugar-to- H_2 yield of 8 H_2 /hexose should be possible by microbial dark fermentation as sufficient ATP by SLP might be obtained under non-standard conditions, especially at elevated temperatures and when partial H_2 pressures can be kept low (**Table 3.1**).

In this study, we set out to achieve a proof of principle that H_2 production yields, *in vivo*, can be increased above the current upper limit of 4 moles/mole of hexose in a single organism. To do this, we selected the previously well characterized and genetically accessible archaeon *Thermococcus kodakarensis* [59, 60, 62, 77, 81, 83, 91, 483]. This heterotrophic hyperthermophile has an optimal growth temperature of 85°C, and is able to degrade environmentally abundant polymers (i.e. starch) to sugars that are oxidized via a modified Embden-Meyerhof pathway (**Figure 3.1**). Reduced ferredoxins (Fd_{red}), generated by glycolysis, must be re-oxidized to keep sugar oxidation going and conserve energy for growth. Although *T. kodakarensis* has several different options for the regeneration of Fd_{ox} from Fd_{red} , the main system used during sugar oxidation is the production of H_2 by a membrane-bound hydrogenase (MBH). Part of the generated H_2 can subsequently be re-consumed by the cell and re-oxidized by a cytosolic or soluble hydrogenase (Hyh) to generate NADPH when required [68, 82, 83]. Any excess NADPH may be disposed of through alanine production, mediated by glutamate dehydrogenase (GDH) and alanine amino transferase (ALT), which has been shown to be a potential electron sink [30, 59, 68, 83, 84].

T. kodakarensis has previously been described to ferment starch to acetate, CO_2 and H_2 , having a H_2 -production yield of 3.3 moles/mole of hexose [30]. Since this is already relatively close to the generally considered upper limit of 4 moles/mole of hexose, it is a good starting point for improvement. In addition, its relatively simple (energy) metabolism and the availability of a genetic toolbox, allow for the introduction of additional oxidation steps which can be realized by complementing its incomplete Pentose Phosphate Pathway (PPP) [59, 91, 99, 484]. Introducing the missing PPP-enzymes will result in the production of extra

NADPH, which may subsequently be converted to Fd_{red} by the action of ferredoxin:NADP oxidoreductase (FNR), or act as electron donor for direct H_2 formation (**Figure 3.1**) [99].

Table 3.1 Thermodynamic data of anaerobic glucose oxidation

1) Glucose + 4 H_2O	→	2 Acetate ⁻ + 2 HCO_3^- + 4 H^+ + 4 H_2	($\text{H}_2/\text{acetate} = 2$)
2) Glucose + 8 H_2O	→	1 Acetate + 4 HCO_3^- + 5 H^+ + 8 H_2	($\text{H}_2/\text{acetate} = 8$)
3) Glucose + 12 H_2O	→	6 HCO_3^- + 6 H^+ + 12 H_2	($\text{H}_2/\text{acetate} = \infty$)

Reaction #	<i>n</i> ATP by SLP*	Energy required for ATP production (KJ)	$\Delta G_0'$ (KJ)	$\Delta G'$ (KJ)***	$\Delta G'_{85^\circ\text{C}}$ (KJ)***
1	4	-280	-207.72	-264.80	-305.25
2	2	-140	-103.50	-194.82	-264.77
3	0	n.a.**	+0.72	-124.85	-224.29

*= conversion AMP to ATP (PEP to PYR) is counted as 2 ATP, ATP consumption for sugar uptake and ATP production as a result of MBH activity not included. **= this reaction is biologically not relevant as no ATP by SLP is formed. ***= concentration of substrates and products 10 mM and partial H_2 pressure of 10^{-1} atm.

H_2 formation from reduced pyridine nucleotides is, however, unfavorable as the Gibbs energy under standard conditions is positive. Only at relatively high temperatures and at low H_2 partial pressures does the reaction becomes exergonic [48]. That H_2 formation from NAD(P)H is possible, *in vivo*, can also be concluded from the H_2 /hexose ratios found for certain (hyper)thermophiles. For instance, *Thermoanaerobacter tengcongensis* and *Caldicellulosiruptor saccharolyticus* both produce reducing equivalents as NADH and Fd_{red} (50/50%), and based on the observed high H_2 /hexose ratios, both types of electron carrier are apparently able to transfer their reducing equivalents to H_2 [97, 98].

For the direct production of H_2 from NADPH, an NADPH-dependent hydrogenase is required. Although this sub-class of [NiFe]-hydrogenases is rare [99-101], *T. kodakarensis* does contain such a cytosolic hydrogenase (Hyh). Although Hyh is normally utilized for oxidizing H_2 to generate NADPH when NADPH levels are low [82, 83], various studies have shown that these [NiFe]-hydrogenases are readily reversible *in vitro* [56, 101-103]. Increasing the intracellular NADPH/ NADP^+ ratio of *T. kodakarensis* by introducing the PPP might, therefore, shift the catalytic direction of Hyh towards H_2 evolution.

Figure 3.1 Simplified overview of metabolism of *T. kodakarensis*. Solid and dashed arrows represent single and lumped enzymatic reactions, respectively. Shown are enzymes naturally present in *T. kodakarensis* TS559 (in black), enzymes introduced to create strain SKS001 (in green), and enzymes additionally introduced to create strain SKS002 (in orange). Abbreviations for metabolites (normal text) and enzymes (bold text): G6P, glucose 6-phosphate; F6P, fructose 6-phosphate; GAP, glyceraldehyde 3-phosphate; 3PGA, 3-phosphoglycerate; Pyr, pyruvate; Ace, acetate; Ala, alanine; 6PGL, 6-phospho-gluconolactone; 6PG, 6-phosphogluconate; Ru5P, ribulose 5-phosphate; HCHO, formaldehyde; R5P, ribose 5-phosphate; X5P, xylulose 5-phosphate; S7P, sedoheptulose 7-phosphate; E4P, erythrose 4-phosphate; PGI, glucose-6-phosphate isomerase; GAPOR, glyceraldehyde-3-phosphate:ferredoxin oxidoreductase; GAPN, NADP-dependent glyceraldehyde-3-phosphate dehydrogenase; ALT, alanine amino transferase; GDH, glutamate dehydrogenase; ATs, amino acid aminotransferases; G6PDH, glucose-6-phosphate dehydrogenase; 6PGLase, 6-phospho-gluconolactonase; 6PGDH, 6-phospho-gluconate dehydrogenase; HPS-PHI, 3-hexulose-6-phosphate synthase (hps) - 6-phospho-3-hexuloisomerase (phi) fusion enzyme; RPI, ribulose-5-phosphate isomerase; Ru5PE, ribulose-phosphate 3-epimerase; TK, transketolase; TAL, transaldolase; MBH, membrane-bound hydrogenase; Hyh, NADPH-dependent soluble hydrogenase; FNR, ferredoxin:NADP⁺ oxidoreductase

Results

Construction of the mutant strains

The genes needed to complement the missing PPP enzymes were derived from the thermophilic bacterium *Thermotoga maritima* (Table 3.2), which has a similar optimal growth temperature as *T. kodakarensis* [485]. To make sure that sugars taken up by the cell are directed through the desired H₂-producing pathway, we made use of the capacity of the PPP to generate ribulose-5-phosphate (Ru5P). By simultaneously disrupting the Ribulose Monophosphate Pathway (RuMP) (Figure 3.2), normally responsible for the synthesis of pentoses in *T. kodakarensis* [91], we forced the organism to use the introduced, bacterial-derived PPP as a source of pentoses. *T. kodakarensis* strain TS559 (Δ pyrF; Δ trpE::pyrF; Δ TK0664; Δ TK0149) was chosen as a parent strain to exploit the previously developed selection and counter-selection procedure for use in rich media [60, 61, 486]. An important feature of this selection and counter-selection procedure is that each strain so constructed has not only the desired modification, but also retains the Δ pyrF; Δ trpE::pyrF; Δ TK0664; Δ TK0149 genotype and can hence be used as a parental strain for another round of the procedure to construct strains having additional modifications.

Table 3.2 Genes introduced and disrupted in this study

Enzyme	EC-number	Source/host	Locus-tag*	Strain
Gene knock-ins				
Glucose-6-phosphate dehydrogenase	EC 1.1.1.49	<i>T. maritima</i>	TM1155	SKS001 + SKS002
6-phospho-gluconate dehydrogenase	EC 1.1.1.44	<i>T. maritima</i>	TM0438	SKS001 + SKS002
6-phospho-gluconolactonase	EC 3.1.1.31	<i>T. maritima</i>	TM1154	SKS002
Transaldolase	EC 2.2.1.2	<i>T. maritima</i>	TM0295	SKS002
Ribulose-5-phosphate 3-epimerase	EC 5.1.3.1	<i>T. maritima</i>	TM1718	SKS002
Gene knock-out				
Hexulose-6-phosphate synthase / 6-phospho-3-hexuloseisomerase	EC 4.1.2.43 / 5.3.1.27	<i>T. kodakarensis</i>	TK0475	SKS001 + SKS002

*= JCVI gene annotation, release date 1999/06/01, modified 2014/12/22

The metabolism of *T. kodakarensis* (**Figure 3.1**) suggests that the introduction of only the oxidative PPP (oxPPP) enzymes may already be enough to complement the cell with all the intermediates associated with the PPP, two different mutant strains were created. The first strain, designated SKS001, was created by introducing glucose-6-phosphate dehydrogenase (G6PDH) and 6-phosphogluconate dehydrogenase (6PGDH), while simultaneously disrupting the RuMP. The oxPPP enzyme 6-phosphogluconolactonase (6PGLase) was not included as it was assumed that the labile intermediate 6-phosphogluconolactone, normally converted to 6-phosphogluconate by 6PGLase, decomposes spontaneously because of the high temperature [487-489]. The second strain, designated SKS002, was created by using strain SKS001 as the parent and was additionally complemented with the lacking oxPPP enzyme, 6PGLase, as well as the missing non-oxPPP enzymes ribulose-5-phosphate 3-epimerase (Ru5PE) and transaldolase (TAL) (**Figure 3.2**). This resulted in a *T. kodakarensis* strain containing a complete PPP; to the best of our knowledge, this is the first archaeal strain with this metabolic feature.

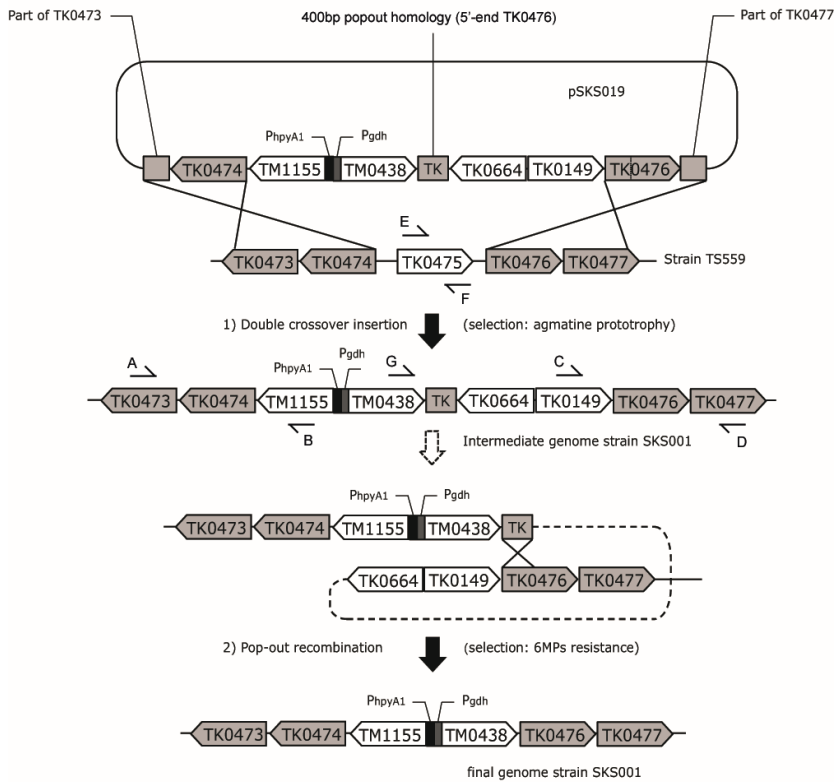
Transformants were selected by colony growth on medium lacking agmatine [60, 486]. Diagnostic PCRs and sequencing of the PCR fragments were performed to discriminate between transformants that were the result of a single homologous recombination (resulting in the integration of the entire donor plasmid DNA integrated into the genome of *T. kodakarensis*) and transformants that were the result of two crossover events (resulting in the desired intermediate genomes as depicted in **Figure 3.2**). Representative intermediates, resulting from two cross-overs, were confirmed to be sensitive to 6-methyl purine (6MP_s), a phenotype conferred by the expression of *TK0664*. Next, precise and markerless deletion of the selection and counter-selection cassette was stimulated by plating the transformants on plates containing 6MP and agmatine. Genomic DNA of spontaneous 6MP resistant (6MP_R) clones growing on these plates was analyzed by diagnostic PCR, confirming that the selection and counter-selection cassette had been deleted by recombination in the flanking, duplicated chromosomal region. Representative clones, having the desired final genome orientations (**Figure 3.2**), were designated *T. kodakarensis* strain SKS001 (TS559 +

$\Delta TK0475::TM1155$, $TM0438$) and *T. kodakarensis* strain SKS002 (SKS001 + $\Delta TK1765::TM0295$, $TM1718$, $TM1154$), respectively.

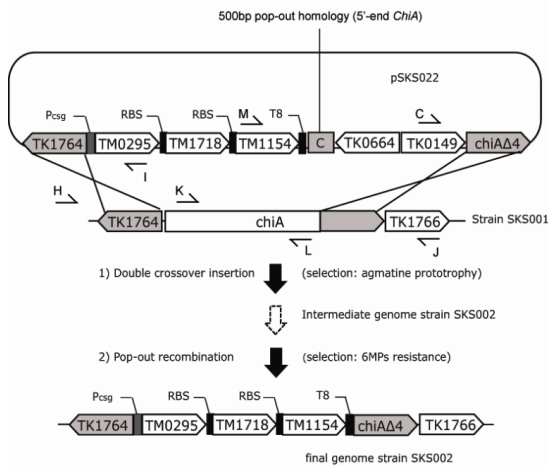
The diagnostic PCRs performed showed that during the procedure, most transformants growing on media lacking agmatine (designated intermediates) turned out to be the result of a single homologous recombination. Moreover, transformants generating desired PCR fragments for the first two reactions (see Materials and Methods), suggesting an intermediate genotype being the result of two cross-over events and hence the loss of the sequence that is normally present in between the two homologous recombinations (*hps-phi* and *chiA44* in strains SKS001 and SKS002, respectively), often also generated a representative PCR product in the third PCR (**Figure S3.1**). These confusing results can be explained by the recent finding that *T. kodakarensis* is polyploid (7-19 chromosomes/cell) (see **chapter 4**) [490], which entails that the desired strains are only obtained when the modifications are incorporated in all chromosome copies. Nevertheless, repetitive and intensive re-plating and screening, eventually led to the desired, homozygous, genotypes for both strain SKS001 and strain SKS002.

As a final diagnostic test, total RNA of both strain SKS001 and SKS002 was isolated and used in a Reverse Transcriptase-PCR (RT-PCR) to confirm the presence or absence of transcription of the genes of interest. Primers used in PCR were developed so that transcription of a particular gene could be monitored by the production of a PCR product of ~300 bp. The results shown in **Figure 3.2C** confirm both the transcription of the introduced codon optimized PPP genes as well as the loss of transcription of the *hps-phi*-gene. The gene coding for the α -subunit of the elongation factor-1 complex, *puf* (*TK0308*), was used for normalization as it is considered (at least under the used conditions) a stably expressed housekeeping gene [491].

A.



B.



C.

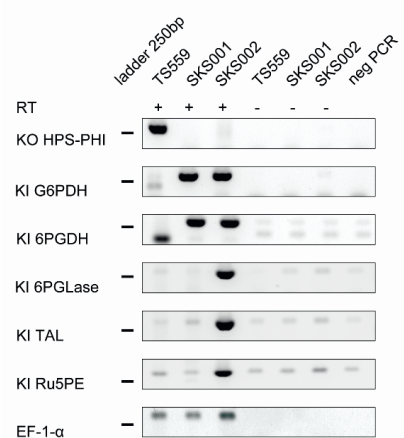


Figure 3.2 Schematic diagram of PPP introduction and RuMP disruption in *Thermococcus kodakarensis*.

(A) Plasmid pSKS019 was constructed for the simultaneous integration of the oxPPP enzymes glucose-6-phosphate dehydrogenase (*TM1155*) and 6-phosphogluconate dehydrogenase (*TM10438*), and the disruption of the RuMP (*TK0475*). It was constructed, as illustrated, with the PCR-amplified regions of the *T. kodakarensis* genome flanking the [*TM1155*; *TM0438-TK0664*; *TK0149*] expression cassette and used to transform *T. kodakarensis* TS559 (Δ *pyrF*; Δ *trpE::pyrF*; Δ *TK0664*; Δ *TK0149*). The genome organization of transformants (intermediate strain), selected by growth in the absence of agmatine, was confirmed by diagnostic PCR (data not shown). A verified intermediate strain was plated on a medium containing 6MP, and 6MP_R clones were isolated. A second round of diagnostic PCRs confirmed the deletion of the selection cassette (*TK0664*; *TK0149*) as well as the desired final genome organization (TS559 + Δ *TK0475::TM1155*, *TM0438*). The verified clone was designated *T. kodakarensis* SKS001.

(B) Plasmid pSKS022 was constructed, as illustrated, with the PCR-amplified regions of the *T. kodakarensis* genome flanking the [*TM0295*; *TM1718*; *TM1154 - TK0664*; *TK0149*] expression cassette and used to transform *T. kodakarensis* SKS001. Selection and verification was performed as described above. A clone having the desired final genome organization (SKS001 + Δ *TK1765::TM0295*, *TM1718*, *TM1154*) was designated *T. kodakarensis* SKS002.

(C) RT-PCR analysis on total RNA isolated from strain TS559, SKS001 and SKS002, confirming the desired integrations and disruption. Total RNA from the aforementioned strains was isolated and exposed to a DNase I treatment to remove any traces of isolated chromosomal DNA. cDNA was created via a reverse transcriptase reaction and subsequently used as template in a diagnostic PCR to confirm expression of the introduced TM-genes as well as loss of expression of the RuMP. All RT-PCR products had the appropriate size on SYBR safe stained agarose gels. The housekeeping gene for elongation factor 1- α was used for normalization.

Abbreviations: T8, synthetic transcription terminator; P_{hpyA1}, histone gene *hpyA1* promoter from *Pyrococcus furiosus* [492]; P_{gdh}, glutamate dehydrogenase promoter [493]; P_{csg}, cell surface glycoprotein promoter [106]; KO, knockout; KI, knock-in; RT, Reverse transcriptase; HPS-PHI, 3-hexulose-6-phosphate synthase (hps) - 6-phospho-3-hexuloisomerase (phi) fusion enzyme; G6PDH, glucose-6-phosphate dehydrogenase; 6PGDH, 6-phospho-gluconate dehydrogenase; 6PGLase, 6-phospho-gluconolactonase; TAL, transaldolase; Ru5PE, ribulose-phosphate 3-epimerase; EF-1 α , elongation factor 1- α . “+”, RT present; “-”, RT absent

Confirmation of enzyme activity

To confirm that the genes introduced into *T. kodakarensis* to generate strains SKS001 and SKS002 are functionally expressed, various bulk assays were performed. We initially attempted to monitor enzyme activities using cell extracts of both the mutant strains, the parental strain TS559 (negative control) and *Thermotoga maritima* MSB8 (positive control). The activities of G6PDH and 6PGDH are traditionally monitored by following NADPH production spectrophotometrically, but our assays employing total cell extracts were then subject to the competing activities of the strong NADPH-oxidase activity observed in the controls [494]. In an attempt to circumvent these competing activities, the G6PDH and 6PGDH activity assays were repeated anaerobically such that NADPH-oxidase activity was now severely hampered by the lack of oxygen. Unfortunately strong NADPH converting background activities also made these attempts inconclusive. Similarly, the coupled anaerobic enzyme assays used to test the functionality of the TAL and Ru5PE, were not successful.

To minimize the interfering background activities, the cell extracts were analyzed by an in-gel activity assay. The benefit of this method is that the soluble dehydrogenases are separated from interfering NADPH oxidases and that functional enzymes can be visualized very specifically by using nitro blue tetrazolium (NBT) as a final electron acceptor. Cell extracts of the mutant strain SKS001, the parental strain TS559 (negative control) and *T. maritima* MSB8 (positive control), were loaded and resolved on a native polyacrylamide gel, which was then used for the in-gel activity assay. The results obtained showed clear, single bands for the lanes containing cell extract of the mutant strain SKS001 or cell extract of *T. maritima* MSB8, but not for the lanes containing cell extract of the parental strain TS559. Moreover, control gels loaded with the same cell extracts placed in similar activity buffers, but without the addition of either glucose-6-phosphate (G6P) or 6-phospho-gluconate (6PG) showed no bands (**Figure S3.2**). Together, these results confirm the functionality of G6PDH and 6PGDH when introduced into *T. kodakarensis*.

Finally, a growth assay on plates containing minimal medium, ASW-AA, supplemented with agmatine but without nucleosides, was performed. A previous study demonstrated that the

ribulose monophosphate pathway (RuMP) is the only relevant pathway for the biosynthesis of Ru5P in *T. kodakarensis* and that a *hps-phi* deletion mutant did not exhibit any growth in minimal medium, while growth could be recovered by the addition of nucleosides to the medium [91]. The *hps-phi* deletion mutants created in this study, however, were predicted to grow on minimal medium as the introduction of the oxPPP should provide an alternative route for Ru5P biosynthesis. Moreover, the presence of a complete PPP in the mutant strain SKS002 may even result in better growth results than those of SKS001. To test this, growth of the mutant strains, SKS001 and SKS002, on ASW-AA plates was compared to that of the previously constructed *hps-phi* deletion mutant, $\Delta hps-phi$ 7A, as well as to that of the wild-type strain KOD1. As expected, no colonies formed on the ASW-AA plates containing the $\Delta hps-phi$ 7A strain, while the ones containing the wild-type strain KOD1 showed many. The plates containing the mutant strains SKS001 and SKS002 also showed many colonies, indicating that the oxPPP is functionally expressed in both strains and that the presence of 6PGLase is indeed not required. Unfortunately, the functionality of TAL and Ru5PE remained unclear, as no difference in growth performance between the two mutant strains created in this study (SKS001 and SKS002) was observed. However, functional expression of TAL and Ru5PE may not necessarily result in better growth, as all intermediates for anabolism can already be synthesized by SKS001.

Growth and fermentation data

The consequences of the modifications made to generate strains SKS001 and SKS002 on growth and alanine, acetate, CO₂ and H₂ production were measured and compared to the parental strain TS559 for cells growing in media containing artificial seawater (ASW) salts, yeast extract and tryptone (YT), and modified potato starch (SA2) (**Figure 3.3**). For each strain, a series of pre-cultures was prepared to allow the strains to fully adapt to the carbon source. Data was collected from multiple independent cultures for each strain, and remarkably, growth of the mutant strains SKS001 and SKS002, measured by the increase in OD₆₀₀ in time, was significantly faster than that of the parental strain TS559 (**Figure 3.3A**). Growth rates obtained were $0.24 \pm 0.011 \text{ h}^{-1}$, $0.19 \pm 0.009 \text{ h}^{-1}$ and $0.14 \pm 0.008 \text{ h}^{-1}$ for SKS001, SKS002 and TS559, respectively (**Figure S3.3**). Final OD₆₀₀-values reached by the

mutants were also significantly higher than those of the parental strain. Growth of both the mutant strains and the parental strain collapsed after ~18-20 h and ~33 h, respectively. Together, these observations show that the growth characteristics of the mutant strains were severely affected by the genetic modifications. To test whether the difference in growth rate was caused by a limited availability of nucleosides, growth of the parental strain TS559 in the absence and presence of nucleosides was compared. We applied nucleosides instead of nucleotides, since the phosphate group may hamper transport over the membrane [91]. Although growth of the cultures was initially very similar, the ones supplemented with nucleosides eventually approached the growth rates observed for the mutants, whereas the ones without nucleosides did not (**Figure S3.4**).

The production of the fermentation end-products alanine, acetate, CO₂ and H₂ by the three strains was simultaneously monitored to investigate their fermentation characteristics. As for the OD₆₀₀ measurements, samples were taken at regular intervals throughout the experiment for each of the replicates for every strain. Average production values (concentration of the product at a certain time point minus the concentration of that product at time point zero) were determined for each time point (**Figure 3.3B**). Similar to the OD₆₀₀ values, the results obtained show that the absolute production of acetate, alanine, CO₂ and H₂ was significantly higher in the mutant strains than in the parental strain (**Figure 3B**). Although these higher product levels may be simply a result of the higher OD₆₀₀ levels, calculation of the specific total reductant production (total reductant production/OD₆₀₀) suggests they are not. For example, a 109% and 63% increase in the specific total reductant production was obtained in the early exponential phase for strains SKS001 and SKS002, respectively (**Figure 3.4A**). Similarly, a 96% and 50% increase in the specific H₂ production (H₂/OD₆₀₀) was obtained (**Figure 3.4B**). For a complete time course of the H₂/OD₆₀₀ ratio see **Figure S3.5A/B**.

To further investigate the differences in the fermentation characteristics of the strains, it was decided to compare the various fermentation end-product ratios. Preferably, the end-products should have been normalized to the amount of hexose that is converted, but the

outcome of the glucose consumption analysis (soluble starch as C-source) was not consistently accurate enough to allow precise calculation of such yields.

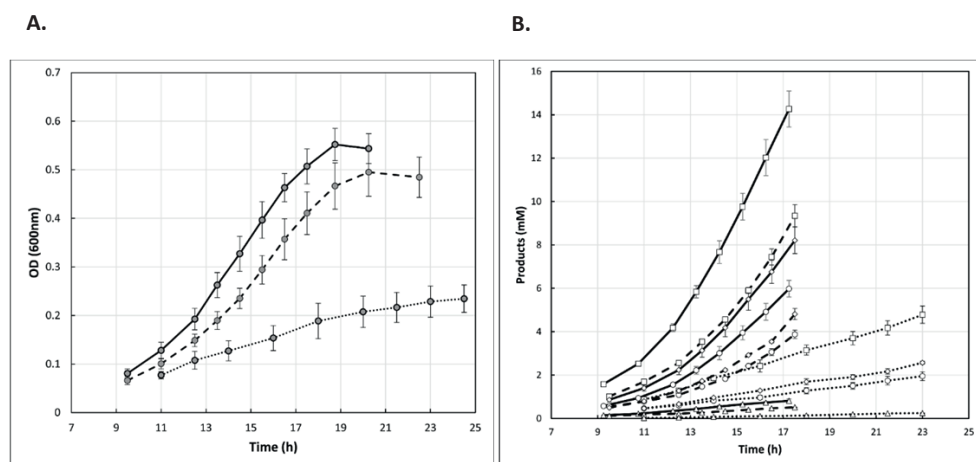


Figure 3.3 Growth and fermentation profile of *T. kodakarensis* strains TS559 (dotted lines), SKS001 (solid lines) and SKS002 (dashed lines) in ASW-YT-SA2 media supplemented with 500 μ M agmatine and 75 μ M 6MPs. **(A)** Increases in the optical density (OD₆₀₀) were measured over time. The results shown for each strain is the average value from the data collected from three independent cultures. **(B)** Production of metabolites was measured over time. Production values were determined by subtracting the concentration of each metabolite at the time of inoculation from the measured concentration of the metabolite at the individual time points. The results shown are the average value from the data collected from three independent cultures. The metabolites analyzed in this study are Hydrogen (squares), CO₂ (diamonds), Acetate (circles) and Alanine (triangles).

Based on the metabolism of *T. kodakarensis* (**Figure 3.1**), certain results are to be expected when cells are grown on sugars. Oxidation of sugars to acetate and CO₂ along the modified Embden-Meyerhof pathway generates Fd_{red} that must be re-oxidized (Fd_{ox}) to maintain metabolic flux and conserve energy for growth. In the absence of S⁰, disposal of reductant is accomplished by a membrane-bound hydrogenase complex (Mbh) that combines the oxidation of Fd_{red} with the reduction of protons, generating H₂ (**Figure 3.1**) [68, 82, 83]. This implies that when *T. kodakarensis* (parent strain) is growing on sugars exclusively, a constant value of 2 is to be expected for both the H₂/acetate and H₂/CO₂ ratios, irrespective of

possible alanine formation. Moreover, a constant value of 1 is expected for the $\text{CO}_2/\text{acetate}$ ratio. In contrast, the alanine/acetate ratio may increase in time, as the partial H_2 -pressure also increases in time, which may force Fd_{red} to transfer electrons, via NADPH, to alanine, or stimulate H_2 uptake by Hyh, also resulting in more NADPH and hence alanine production [30, 59, 68, 82-84]. However, if the oxPPP contributes to the glucose conversion, the $\text{H}_2/\text{acetate}$ and $\text{CO}_2/\text{acetate}$ ratios will increase, whereas the H_2/CO_2 will remain 2. Moreover, the alanine values for the mutant strains may increase, as these strains are expected to produce extra NADPH in the oxPPP, which may be used to produce more H_2 , but also more alanine.

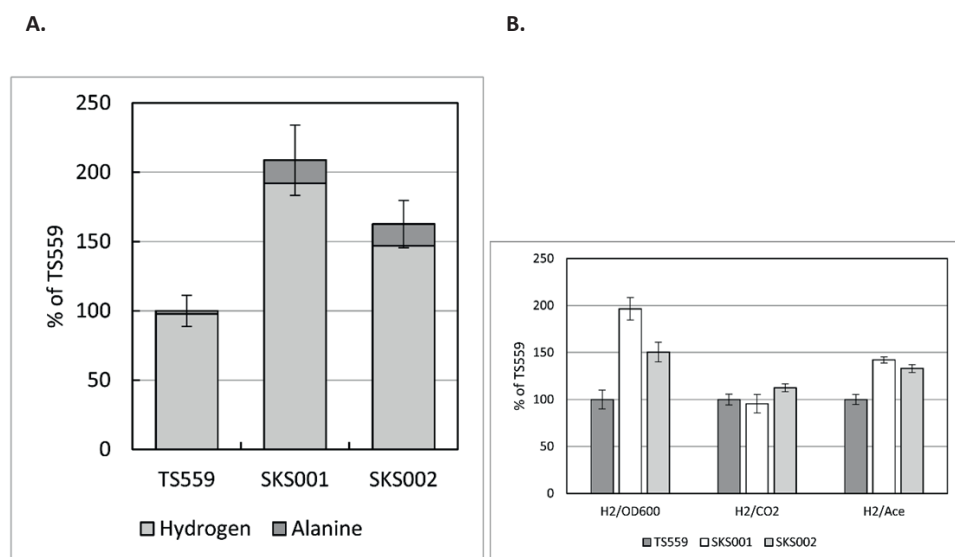


Figure 3.4 Relative reductant and hydrogen production after 11 h of growth. (A) Relative total reductant production per OD_{600} , expressed as a percentage of TS559 **(B)** relative hydrogen production ratios, expressed as a percentage of TS559

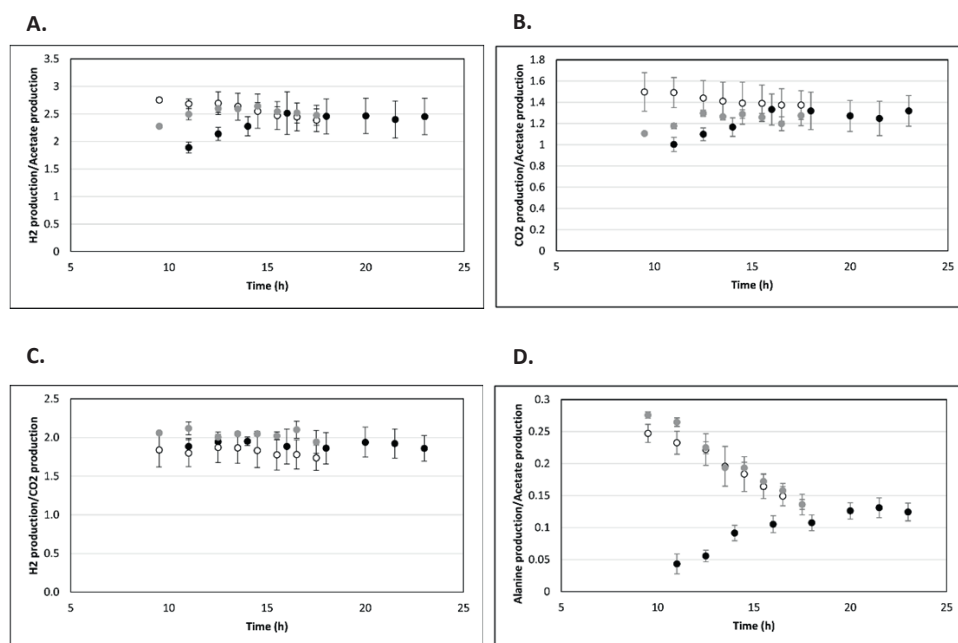


Figure 3.5 Fermentation results (product/product) of *T. kodakarensis* strain TS559 (black circles), SKS001 (white circles) and SKS002 (grey circles) in ASW-YT-SA2.

Figure 3.5A shows that the H₂/acetate ratios obtained for the mutant strains SKS001 and SKS002 are very similar to each other and that the values obtained at early logarithmic phase are significantly higher than those obtained for the parental strain TS559. For example, after 11 h of growth a 42% and 33% increase in the H₂/acetate ratio was observed for SKS001 and SKS002, respectively (**Figure 3.4B**). Moreover, the ratios obtained for the mutants at early logarithmic phase are also well above the expected wildtype ratio of 2 (~2.3-2.7). Remarkably, while the H₂/acetate ratio of both mutant strains remained rather constant in time or slightly decreased, the ratio for TS559 increased in time (from ~1.9 towards ~2.5; **Figure 3.5A**). Thus, all three strains eventually appear to level off around 2.5. The effect of the mutations on the CO₂/acetate ratios appear to be somewhat similar to the effect observed on the H₂/acetate ratios. **Figure 3.5B** shows that the CO₂/acetate ratios obtained for the mutant strains at early logarithmic phase are significantly higher than those obtained for the parental strain TS559 and well above the expected wildtype ratio of 1, especially for

strain SKS001, but that all eventually appear to level off around 1.3. Concerning the H_2/CO_2 ratios, no significant changes were observed. Like TS559, the H_2/CO_2 ratios obtained for the mutants are relatively constant and show a value around the expected value of 2 (**Figure 3.4B** and **3.5C**). The alanine/acetate ratios of the mutant strains SKS001 and SKS002 are very comparable to each other and start significantly higher than the ratio obtained for TS559 (**Figure 3.5D**). Interestingly, alanine/acetate ratios for the mutants show a fast decline in time, whereas TS559 shows the opposite and increases from ~0.05 towards ~0.13 (**Figure 3.5D**).

Differential proteome analyses

Proteomic analysis was used to investigate the metabolic response of *T. kodakarensis* to the changes made in the strains SKS001 and SKS002. The analysis resulted in the identification of 569 proteins (corresponding to 24.67% of the identified coding DNA sequences [59]), 403 of which were found in TS559, 460 of which were found in SKS001 and 412 of which were found in SKS002. The relative abundance of the differentially expressed proteins in *T. kodakarensis* TS559, SKS001, and SKS002 grown on starch, is shown in **Table S3.1** and a short list of relevant proteins which show differential expression is shown in **Table 3.3**. The results confirm the loss of expression of *TK0475* (HPS-PHI) in the mutants, as well as the expression of *TM1155* (G6PDH), and additionally *TM0295* (TAL), *TM1154* (6PGLase) and *TM1718* (Ru5PE), in strains SKS001 and SKS002, respectively. Proteins associated with DNA replication (such as DNA-polymerases, DNA topoisomerases or DNA primases), transcription (RNA-polymerases), translation (aminoacyl-tRNA synthetases, tRNA/rRNA/mRNA modification and processing proteins) and protein folding (chaperonins) are all slightly upregulated in the mutants as compared to the parental strain TS559. This is likely related to the higher growth rate of strain SKS001 and SKS002. Contrary to these results, cell division proteins are generally slightly downregulated in the mutant strains. With respect to proteins involved in electron transfer or NADP(H) conversion, no big differences were observed. The alanine aminotransferase (*TK1094*) and the glutamate dehydrogenase (*TK1431*), which are responsible for the disposal of reductant as alanine, were both downregulated. A subunit of

the soluble hydrogenase Hyh (*TK2072*) also showed a 4-fold downregulation, whereas protein levels of the membrane bound hydrogenase (*TK2080-2093*) were unaltered. The enzyme responsible for the first oxidation step of the glycolysis, viz. glyceraldehyde-3-phosphate:ferredoxin oxidoreductase (GAPOR; *TK2163*) is slightly upregulated, while its NADP-dependent counterpart (GAPN; *TK0705*) remained unaltered. Protein levels of the PPP enzyme ribose-5-phosphate isomerase (*TK1426*) are slightly higher or doubled in SKS001 and SKS002, respectively. The latter suggests that the function of 6PGLase, i.e. converting 6PGL into 6PG, is not completely redundant, as has previously also been reported by Miclet et al. [488]

Table 3.3 Relative abundance of the differentially expressed proteins related to this study, in *T. kodakarensis* TS559, SKS001 and SKS002

Protein	Gene (Locus tag)	SKS001/TS559	SKS002/TS559	SKS002/SKS001
Heterologously expressed proteins [<i>Thermotoga maritima</i> MSB8]				
translaldolase (TAL)	TM0295	n.d.	SKS002	SKS002
6-phosphogluconate dehydrogenase (6PGDH)	TM0438	SKS001	n.d.*	SKS001
6-phospho-gluconolactonase (6PGLase)	TM1154	n.d.	SKS002	SKS002
glucose-6-phosphate 1-dehydrogenase (GGPDH)	TM1155	SKS001	SKS002	0.44
ribulose-5-phosphate 3-epimerase (Ru5PE)	TM1718	n.d.	SKS002	SKS002
Native <i>Thermococcus kodakarensis</i> proteins				
Hexulose-6-phosphate synthase/6-phospho-3-hexuloseisomerase (HPS-PHI)	TK0475	TS559	TS559	n.d.
NADP-dependent glyceraldehyde-3-phosphate dehydrogenase (GAPN)	TK0705	1.07	0.94	0.88
alanine aminotransferase (ALT)	TK1094	0.35	0.43	1.24
ribose-5-phosphate isomerase A	TK1426	1.30	2.06	1.58
glutamate dehydrogenase (GDH)	TK1431	0.90	0.72	0.80
cytosolic NiFe-hydrogenase subunit beta (Hyh)	TK2072	0.22	0.24	1.06
membrane bound hydrogenase subunit MbHE	TK2084	SKS001	n.d.	SKS001
membrane bound hydrogenase subunit MbHl	TK2088	n.d.	SKS002	SKS002
membrane bound hydrogenase, NiFe-hydrogenase small subunit	TK2089	1.10	0.91	0.82
membrane bound hydrogenase, NiFe-hydrogenase large subunit 1	TK2090	1.06	1.11	1.05
membrane bound hydrogenase, NiFe-hydrogenase large subunit 2	TK2091	1.02	1.05	1.03
membrane bound hydrogenase, MbHm subunit	TK2092	n.d.	SKS002	SKS002
glyceraldehyde-3-phosphate:ferredoxin oxidoreductase (GAPOR)	TK2163	1.67	1.72	1.03

* = The expression of *TM0438* (controlled by the *gdh* promoter) was considerably lower than that of the other heterologously expressed genes and was in strain SKS002 only detected below the cutoff value. Nonetheless, the results of the in-gel activity assay already confirmed its functional expression (**Figure S3.2**).

Discussion

In this study, we set out to achieve a proof of principle that the H₂ production yield in the archaeon *T. kodakarensis* could be increased by introducing extra substrate oxidation (or reducing equivalent-generating) steps. To do this, bacterial PPP genes were successfully introduced, yielding strains SKS001 and SKS002, while simultaneously disrupting the endogenous RuMP to ensure partial rerouting of the sugar metabolism through the oxPPP. *T. kodakarensis*, or archaea in general, usually do not contain a functional PPP [59, 91, 99, 484]. Instead, pentose phosphates (Ru5P and R5P) are generated by the RuMP [91, 484]. As expected, introduction of the oxPPP enzymes G6PDH and 6PGDH was sufficient to substitute for the loss of the RuMP and to provide the cell with all the intermediates normally associated with the PPP. This is clearly shown by the successful cultivation of both the SKS001 and SKS002 strain on ASW-AA plates and is also suggested by the similar growth behavior of these strains in ASW-YT-SA2 medium (**Figure 3.3A**).

Introduction of oxPPP leads to higher hydrogen yields

The introduction of extra substrate oxidation steps in the sugar metabolism of *T. kodakarensis* may lead to higher H₂ yields if the generated NADPH can be used for proton reduction. Alternatively, NADPH can be used to produce alanine, at the expense of pyruvate. Thus, increased production of NADPH by introduction of the oxPPP enzymes, may lead to higher levels of H₂ and/or alanine. Other reduced end-products are not produced by *T. kodakarensis*. If we compare the H₂ and alanine levels of the mutant and the parental strains we see that reductant levels are indeed significantly higher for SKS001 and SKS002 (**Figure 3.3**). Even when we normalize these data for the difference in growth (OD₆₀₀), both PPP mutants show a significantly higher level of specific reductant, both hydrogen and alanine (**Figure 3.4A**).

Although this is a considerable improvement, it does not necessarily mean that the H₂ production per hexose has increased as well. This is because the higher specific reductant level might, for example, also be due to an improved overall sugar oxidation rate. To determine whether the engineered strains actually produce H₂ beyond the threshold of 4

H₂/hexose, we need to determine the exact amount of converted sugar. Unfortunately, the starch-based medium used in this study does not allow for the precise quantification of such yields. Another way to determine whether the oxPPP contributes to the generated H₂, is by calculating the H₂/acetate ratio. In the EM pathway that is naturally present in *T. kodakarensis*, the H₂/acetate ratio is always predicted to be ~2. Even when part of the generated H₂ is used for alanine formation (re-oxidation of H₂ by Hyh), this does not affect the H₂/acetate ratio; as for each alanine produced two H₂ and one acetate (actually pyruvate) is needed. However, if part of the sugar is converted along the oxPPP, more reductant and less acetate is produced, leading to an increase in the H₂/acetate ratio. If all sugar would be converted via the oxPPP the H₂/acetate ratio would even become 8 (**Table 3.1**).

In the early exponential phase, the H₂/acetate ratios obtained in this study are ~1.9 and ~2.7/2.5 for the parental strain TS559 and the mutants SKS001/SKS002, respectively (**Figure 3.5A**). This suggests that part of the sugar in the mutants is indeed metabolized via the oxPPP and that the extra reductant associated with this is (partially) converted to H₂. However, when growth ceases, the parental strain also reaches a value of ~2.5 (**Figure 3.5A**), which is higher than expected based on the natural sugar fermentation pathway. A similar pattern is observed for the CO₂/acetate ratios, i.e. lower values for the parent strain at the early exponential phase (~1), reaching similar values as the mutants (~1.3) when growth ceases (**Figure 3.5B**). These observations cannot easily be explained.

Possibly, the contribution of amino acid fermentation increases. This is because the medium used contains a substantial amount of amino acids and it has been shown previously that under the conditions used, sugars and amino acids can be oxidized simultaneously [68]. Like sugar oxidation, amino acid oxidation in the absence of S⁰ results in the production of acids (other than acetate), H₂ and CO₂. This implies that in the case of simultaneous sugar and amino acid oxidation a higher than expected H₂/acetate- and CO₂/acetate ratio might indeed be obtained, i.e. above the expected value of 2 and 1, respectively. However, amino acid oxidation in the absence of S⁰ is thermodynamically difficult and seems to be only possible at very low H₂ concentrations [495, 496]. Moreover, an increasing contribution of amino acid

fermentation would also imply a decreasing H_2/CO_2 ratio (exclusive sugar oxidation results in a H_2/CO_2 ratio of 2, exclusive amino acid oxidation results in a H_2/CO_2 ratio of ~ 1 [68, 497]). The fact that the H_2/CO_2 ratios obtained are constant in time and show values around the expected value for sugar oxidation (~ 2), argues against any contribution of amino acid oxidation (**Figure 3.5C**).

Interestingly, also the alanine/acetate ratio is significantly different in the early phase of growth, but reaches similar values when growth rates decline (**Figure 3.5D**). Higher alanine values for the mutant strains are plausible, as these strains are expected to produce extra NADPH in the oxPPP, which may be used to produce more H_2 , but also more alanine. Again, why these differences disappear near the end of growth is not clear.

Altogether the results show that the introduction of the oxPPP genes results in a significantly different (hydrogen) metabolism during exponential growth, but that the mutant strains shift to a similar type of metabolism as observed in the parent strain near the stationary phase.

Introduction of oxPPP leads to higher growth rates

Surprisingly, substituting the RuMP with the PPP not only results in higher H_2 levels, but, when growing in ASW-YT-SA2 medium, also results in a significant increase in both the growth rate and the final OD_{600} -values reached (**Figure 3.3A**). These improved growth results were unexpected as the introduction of extra oxidation steps, to convert some of the substrate into CO_2 at the expense of acetate, implies that less ATP is formed by substrate level phosphorylation (SLP) (**Figure 3.1**). It might be, however, that the loss of ATP production by SLP is compensated for by extra ATP production via the respiratory system in the following way. The introduction of the oxPPP enzymes G6PDH and 6PGDH is supposed to increase the NADPH supply, which, if successful, may allow an increase in the Fd-reducing GAPOR flux at the expense of NADP-reducing GAPN, which is no longer needed for NADPH generation (**Figure 3.1**). The extra Fd_{red} generation that is associated with the increased GAPOR flux eventually also leads to an increased MBH activity and hence proton gradient. The proton gradient can then be used by the membrane-bound ATP synthase, to eventually produce extra ATP and hence conserve energy for growth. Interestingly, the proteome

analysis indeed shows an 67% and 72% increase in the level of GAPOR, whereas the expression level of GAPN only increased 7% or even decreased by 6% in strains SKS001 and SKS002, respectively (**Table 3.3 and S3.1**). Moreover, the proteome analysis also showed a decrease in the levels of several ATP consuming enzymes, such as flagella related proteins or various ABC transporters (**Table S3.1**). Though the rationale behind this is unclear, these results suggests that the improved growth performance might also be due to a decreased ATP demand.

Another possible explanation for the improved growth might be that the extra NADPH generation associated with the oxPPP has a positive effect on biosynthesis. *T. kodakarensis* does not contain many sites for NADPH generation, especially not under glycolytic growth conditions [77, 99]. Extra NADPH generation by the oxPPP might therefore be beneficial for various biosynthetic pathways in the cell. Confirmation that the introduction of the oxPPP indeed led to higher NADPH availability seems to be provided by the higher specific alanine production in the early exponential phase (**Figure 3.4A**), as well as by the lower Hyh levels and the relative increase in the GAPOR/GAPN levels observed in the proteome analysis (**Table 3.3**). Similarly, introduction of the oxPPP might have a positive effect on the availability of pentose phosphates (Ru5P and R5P), which are used to build the pentose moiety of nucleotides (**Figure 3.1**) [91]. We could confirm this hypothesis, by showing that addition of nucleosides indeed caused an increase in the growth rate of the parental strain TS559 (**Figure S3.4**). These remarkable results suggest that pentose phosphate availability is limiting in TS559 under glycolytic conditions, which may be resolved by substituting the RuMP for the oxPPP. Moreover, an increased availability of Ru5P also seems to be suggested by the proteome analysis, where an 30% and 110% increase in the level of ribose-5-phosphate isomerase was observed for strains SKS001 and SKS002, respectively.

Conclusions

The results presented demonstrate that bacterial PPP genes were successfully introduced into the archaeon *T. kodakarensis* and that introduction of just the oxPPP enzymes G6PDH and 6PGDH is sufficient to replace the RuMP. Introduction of the PPP resulted in significantly higher H₂ production levels, and the elevated H₂/acetate ratios in the early exponential phase suggested that increased NADPH production may ultimately be responsible for the increase in H₂ formation. Further research using quantifiable sugars is needed to obtain H₂/hexose data and to exclude the contribution of amino acid fermentation to increased H₂ production. Surprisingly, introduction of the PPP also resulted in significantly higher growth rates and final OD₆₀₀-values, which, together with the higher specific H₂ production values (H₂/OD₆₀₀), results in a substantially higher overall H₂ productivity. The proteome analysis suggested that the higher growth rate could be due to an improved supply of ribulose-5-phosphate, which was later confirmed by the improved growth characteristics of the parental strain in ASW-YT medium supplemented with nucleosides. In addition, an increase in NADPH availability, and possibly a decrease in ATP consumption might be responsible. Finally, the data presented suggests that some oxidation of amino acids may occur as well and that its relative contribution may increase in time. Although a proof of principle was obtained that introduction of extra oxidation steps can lead to higher H₂ levels, this is only the beginning. Even higher yields are necessary to make biohydrogen production by dark fermentation economically competitive. Thermodynamics predict that H₂/acetate ratios of 8 should be thermodynamically possible. Future research will, therefore, focus on further improvement of the H₂ yield, among others by blocking the competing glycolytic pathway.

Materials and Methods

Microorganisms and culture conditions

E. coli DH5 α was used for general DNA manipulation, and *E. coli* strains were cultivated at 37°C in Luria-Bertani (LB) medium supplemented with either 100 μ g/ml ampicillin or 50 μ g/ml kanamycin. *T. maritima* MSB8, used as a positive control in enzyme activity assays, was cultivated anaerobically at 80°C in M3-medium containing 5 g/L glucose as described previously [498]. *T. kodakarensis* strains and plasmids used in this study are listed in **Table 3.4**. *T. kodakarensis* strains were cultivated anaerobically in stoppered serum bottles at 85°C in a nutrient-rich medium (ASW-YT), composed of 0.8x artificial seawater (0.8x ASW) [62], 5.0 g/L yeast extract, 5.0 g/L tryptone, 1 ml/L KOD vitamin mixture (1000x) and 0.8 mg/L resazurine. The headspace of all cultures was exchanged for N₂ or, when an internal gas standard was required, for N₂ and CH₄ (see below). Before inoculation, either 2.0 g/L S⁰ (ASW-YT-S⁰) or 5.0 g/L of the modified potato starch “Paselli SA2” (ASW-YT-SA2) was added to the medium, as well as Na₂S₉H₂O to reduce the dissolved oxygen, until the medium became colorless. The modified potato starch “Paselli SA2” has a degree of polymerization of ~60 and was a gift of AVEBE FOOD, Veendam, The Netherlands. When required, the media were furthermore supplemented with 500 μ M agmatine and/or 75 μ M 6-methyl purine (6MP) [60, 61], and in the case of the ASW-YT-SA2 medium with 10 mM PIPES buffer (the pH was adjusted to pH 7 at RT). For the growth experiment with supplemented nucleosides, 0.25% of adenosine and uridine was added to the ASW-YT-SA2 medium. Cultures were continuously shaken (120 rpm) during growth. The minimal medium (ASW-AA) used in the growth assay on plates was prepared as described previously [62]. In the case of plate culture, 1% (wt/vol) Gelrite was added to solidify the medium together with 2 ml/L polysulfide solution [62] instead of S⁰. Cells inoculated on plates were incubated at 85°C in an air-tight anaerobic jar containing an Anaerocult A mini (Merck Millipore, Darmstadt, Germany) to ensure anaerobic conditions.

Table 3.4 Strains and plasmids used in this study

Strain or plasmid	Relevant characteristics	Source or reference
<i>T. kodakarensis</i> strains		
KOD1	Wild type	[57]
TS559	$\Delta pyrF$; $\Delta trpE::pyrF$; $\Delta TK0664$; $\Delta TK0149$	[486]
Δhps -phi 7A	$\Delta pyrF$; $\Delta trpE::pyrF$; Δhps -phi:: <i>trpE</i>	[91]
SKS001	TS559 derivative, $\Delta TK0475$ (<i>hps</i> -phi):: <i>TM1155</i> , <i>TM0438</i>	This study
SKS002	SKS001 derivative, $\Delta TK1765::TM0295$, <i>TM1718</i> , <i>TM1154</i>	This study
Plasmids		
pUDhps-phi	pUC118 derivative, Δhps -phi:: <i>trpE</i>	[91]
pSKS019	<i>TK0664</i> ; <i>TK0149</i> ; $\Delta TK0475$ (<i>hps</i> -phi):: <i>TM1155</i> , <i>TM0438</i>	This study
pSecPyrF	<i>PyrF</i> overexpression and secretion system, integration at <i>TK1765</i> locus ($\Delta TK1765::trpE$, <i>PyrF</i>)	[106]
pSKS022	<i>TK0664</i> ; <i>TK0149</i> ; $\Delta TK1765::TM0295$, <i>TM1718</i> , <i>TM1154</i>	This study

Construction of plasmids and mutant strains

Plasmids incapable of autonomous replication in *T. kodakarensis* were constructed and amplified in *Escherichia coli*, carrying the desired genes adjacent to the previous developed selection and counter-selection cassette (*TK0149*; *TK0664*). The PPP genes and selection and counter-selection cassette were flanked by a ~1-kb region homologous to *T. kodakarensis* genomic sequences that naturally were located immediately up- and downstream of the target locus (**Figure 3.2**). Standard molecular biology procedures were used for PCR amplifications, plasmid construction, *E. coli* DH5 α transformation, transformants selection, and plasmid isolation. The sequences of oligonucleotide primers used in this study are listed in **Table S3.2**.

Two different vectors, designated pSKS019 and pSKS022, were constructed to create strain SKS001 and SKS002, respectively. Plasmid pSKS019, designed to introduce the oxPPP enzymes G6PDH and 6PGDH while simultaneously disrupting the HPS-PHI fusion enzyme and hence the RuMP, was constructed as follows. DNA fragments containing either the genes of interest, *TM1155*; *TM0438*, codon optimized for expression in *T. kodakarensis*, or the selection and counter-selection cassette, *TK0664*; *TK0149* [60, 61], were synthesized by

GeneArt (Life technologies, Regensburg, Germany), ligated together and digested from the resulting plasmid using PvuII and EcoRI, to create a linear blunt ended DNA fragment containing both the TM-genes and the selection and counter-selection cassette. Next, an inverse-PCR was carried out using the previously constructed *hps-phi* disruption vector pUDhps-phi [91], as template with the primer-pair BG3884-BG3885, to amplify the *hps-phi* flanking regions and the pUC118 backbone, thereby removing the *trpE* marker. Finally, the created vector backbone containing the *hps-phi* flanking regions, and the linear DNA fragment containing the TM-genes and selection and counter-selection cassette (*TM1155-TK0149*), were ligated together to form pSKS019 as shown in **Figure 3.2**. Similarly, plasmid pSKS022, designed to introduce the remaining missing PPP enzymes into SKS001, was constructed by carrying out an inverse PCR using the previously constructed overexpression and secretion vector pSecPyrF [106], as template with the primer-pair BG4718-BG4720, to amplify the desired flanking regions and backbone, thereby removing the *csg*-promoter, the *TK1675* secretion signal, the PyrF coding sequence, the STREP-tag sequence, and the *gdh*-terminator. pSecPyrF was chosen as template as integration in the genome at the *ChiA* locus targeted by this plasmid has shown to be efficient and non-disturbing. The linear vector backbone created by this PCR was flanked by BamHI and EcoRI restriction sites, which were incorporated in the primers used. A linear DNA fragment containing a synthetic operon composed of *TM0295*, *TM1718* and *TM1154* codon optimized for expression in *T. kodakarensis*, as well as the *TK0664*; *TK0149* cassette was synthesized by GeneArt (Life technologies, Regensburg, Germany). The entire fragment was flanked by BamHI and EcoRI restriction sites, allowing for easy ligation with the amplified vector backbone, creating pSKS022 as shown in **Figure 3.2**. Transcription of the synthetic operon was designed to be driven by the strong constitutive cell surface glycoprotein (*csg*) promoter [106] and to be inhibited by the presence of a synthetic transcription terminator (T_8) [499].

T. kodakarensis SKS001 was generated by transforming pSKS019 into *T. kodakarensis* TS559, and following the previously described selection and counter-selection procedure (**Figure 3.2**) [60, 61]. *T. kodakarensis* SKS002 was generated by transforming pSKS022 into the newly created strain SKS001, and repeating the selection and counter-selection procedure.

Diagnostic PCRs were performed to select and confirm genotypes of both intermediate and final strains. For the selection of intermediates, at least three diagnostic PCRs were performed. In one PCR, the primers used were complementary to the sequence upstream of the desired integration locus as well as to the sequence inside one of the codon optimized PPP-genes (primer-pair BG4131-BG4133 and BG4871-BG4872, illustrated by A-B and H-I in **Figure 3.2**, for SKS001 and SKS002 respectively). Similarly, a second PCR was included using primers that were complementary to the sequence inside the selection and counter-selection cassette as well as to the sequence downstream of the desired integration locus (primer-pair BG4126-BG4129 (C-D) and BG4126-BG4873 (C-J) for SKS001 and SKS002, respectively). A third PCR was performed to confirm the loss of *hps-phi* (primer-pair BG4658-4659 (E-F)) or *chiAΔ4* (BG5059-BG5060 (K-L)) in strain SKS001i and SKS002i, respectively. A generated PCR product having the correct size for the first two reactions, in combination with a failure to generate a PCR product for the third reaction, suggested the desired intermediate genotype. To discriminate between final strains and intermediate strains, an additional fourth PCR was performed using the genomic DNA of spontaneous 6MP resistant (6MP_R) clones as template and primers that were complementary to the sequence downstream of the desired integration locus as well as to the sequence inside one of the codon optimized genes (BG4129-3952 (D-G) and BG4873-BG4874 (J-M)). A generated PCR product having the correct size for the first and fourth reaction, in combination with a failure to generate a PCR product for the second and third reaction, suggested the desired final genotype. Sequences of the primers are listed in **Table S3.2**.

RNA isolation and RT-PCR

To confirm transcription of the introduced PPP-genes as well as the loss of transcription of the *hps-phi* gene, Reverse Transcriptase PCR (RT-PCR) was performed. To do this, the parental strain TS559 and the mutant strains SKS001 and SKS002 were cultivated in 120-ml serum bottles containing 40 ml ASW-YT-SA2 medium, supplemented with 500 μM agmatine and 75 μM 6MP. Samples of 10 ml were taken and cells were harvested by centrifugation (4700 g, 30 min, 4°C). Total RNA was isolated by the following method. Cell pellets were resuspended in 1 ml Trizol and 200 μl chloroform was added. The mixtures were shaken by

hand for 15 seconds, followed by an incubation at RT for 3 min. Next, the samples were centrifuged (13,000 g, 15 min, 4°C) and the supernatants were transferred to fresh Eppendorf tubes. 5 µg glycogen and 500 µl isopropanol were added and the mixtures were incubated at RT for 10 min. The samples were centrifuged (13,000 g, 15 min, 4°C) and the pellets were washed in 75% ethanol, air-dried for 10 min and resuspended in MQ water. RNA concentrations were determined with a Nanodrop ND-1000 and the total RNA samples were stored at -20°C until further processing.

Before RT-PCR, the total RNA samples (1 µg each) were treated with DNaseI, to remove residual genomic DNA. cDNA was prepared by following the protocol for first strand cDNA synthesis of the Thermo Scientific RevertAid RT kit (20 µl reactions), using random hexamers as primers. Gene-specific cDNA fragments were amplified by conventional PCR using the freshly prepared cDNA as template and primer pairs developed to specifically target the genes of interest (**Table S3.2**). The genes of interest were: *TK0475* (*hps-phi*), *TM0295* (TAL), *TM0438* (6PGDH), *TM1154* (6PGLase), *TM1155* (G6PDH), *TM1718* (Ru5PE). The housekeeping gene *TK0308* (*puf*), encoding the α -subunit of the elongation factor-1 complex (EF-1- α), was included for normalization. All primer pairs were designed to produce a specific product of ~300 bp. All reactions were run for 35 cycles with the exception of the reaction analyzing *TK0308*, which was run for only 18 cycles. PCR amplification products were analyzed by agarose gel electrophoresis. Control experiments were included by following the same procedure, but without addition of RevertAid Reverse Transcriptase enzyme to the RT reaction mix.

Cell lysis and enzyme activity assays

To test if the introduced PPP-genes are functionally expressed, various assays were performed. For the enzyme activity assays, including the in-gel activity assay, cell-free extracts were prepared. To do this, both the mutant strains SKS001 and SKS002, and the parental strain TS559 (negative control) were cultivated anaerobically in 200 ml ASW-YT-SA2 (500-ml anaerobic bottles) as described above. In addition, a 40-ml culture of *T. maritima* MSB8 (positive control) in M3 medium containing glucose was prepared. Cells were harvested by centrifugation (5000 g, 30 min, 4°C) and the pellets were resuspended in 2 ml

and 1 ml lysis buffer, for *T. kodakarensis* and *T. maritima* cultures respectively. The lysis buffer was composed of 50 mM Tris-HCl buffer (pH adjusted to pH 8.0 by adding HCl), 1 mM EDTA, 150 mM NaCl, 1 mM DTT and 0.1 mM PMSF. The cells were lysed by sonication, and the cellular debris was removed by centrifugation (13,000 *g*, 30 min, 4°C). The protein concentration was determined using a Pierce BCA Protein Assay Kit (Thermo Scientific). The cell-free extracts were stored at -20°C until further processing.

The enzyme assays were performed in pre-heated 1-ml quartz cuvettes. The cuvettes containing the assay mixture were allowed to reach the desired temperature inside a temperature controlled Hitachi U-2001 spectrophotometer. The reaction was started by either adding cell-free extract or an essential substrate. In case of the anaerobic assays, stoppered 1-ml quartz cuvettes were used, that were flushed with N₂. In addition, all cell-free extracts, stock solutions and the buffer were flushed with N₂.

Glucose-6-phosphate dehydrogenase (EC 1.1.1.49) and 6-phospho-gluconate dehydrogenase (EC 1.1.1.44) were assayed by measuring the formation of NADPH at 334 nm. The assays were performed at 80°C, at which both the *T. kodakarensis* and *T. maritima* enzymes are active. The assay mixtures contained: 100 mM Tris-HCl (pH 7.8), 10 mM MgCl₂, 0.5 mM NADP⁺, 2 mM of either glucose-6-phosphate or 6-phospho-gluconate, and at least 20 µl extract (60-120 µg protein). Transaldolase (EC 2.2.1.2) and Ribulose-5-phosphate 3-epimerase (EC 5.1.3.1) were assayed spectrophotometrically at 80°C in a coupled enzyme assay with GAPOR of *T. kodakarensis*, using benzyl viologen as artificial electron acceptor. The reaction mixtures in stoppered anaerobic quartz cuvettes contained 50 mM Tris-HCl (pH 8.0) and 1 mM benzyl viologen. In the cuvette, the buffer was slightly prereduced with sodium dithionite until a light blue color was maintained, before addition of the extracts. The reactions were started by either adding 2 mM of fructose-6-phosphate and erythrose-4-phosphate, or 2 mM of ribulose-5-phosphate for the TAL and Ru5PE-assay, respectively. The reduction of benzyl viologen was followed at 560 nm. In addition, a positive control reaction was included using 2 mM of GAP to start the reaction.

Glucose-6-phosphate dehydrogenase (EC 1.1.1.49) and 6-phospho-gluconate dehydrogenase (EC 1.1.1.44) were also assayed by an in-gel activity assay. Using this method, soluble dehydrogenases can be visualized very specifically. The assay was performed by loading 100 µg of cell-free extracts of the mutant strain SKS001, the parental strain TS559 (negative control), and *T. maritima* MSB8 (positive control) on a native polyacrylamide gel (5% stacking gel, pH 6.8 and 10% resolving gel, pH 8.8). The gels were resolved by running them at RT at 15 mA per gel for ~1.5 h. Following electrophoresis, the gels were stained for dehydrogenase activity with phenazine methosulfate (PMS) and nitro blue tetrazolium chloride (NBT) according to the method of Benoist and Schwencke (1990) [500]. PMS acts as an intermediate electron carrier for coupling the production of NADPH to the reduction of NBT to insoluble precipitates. When the indicated dehydrogenases are actively expressed precipitates will be formed at their position in the gel. The assay was performed by placing the gels in 20 ml activity staining buffer, composed of 100 mM NaH₂PO₄ (pH 8.0), 200 mM NaCl, 5 mM MgCl₂, 0.03% NBT and 0.001% PMS. NADP⁺ (1 mM) and either glucose-6-phosphate or 6-phospho-gluconate (100 mM each) were used as substrates. The activity staining was performed by incubating the gels for 10-30 minutes in an anaerobic container at 60°C and 37°C, for the glucose-6-phosphate and 6-phospho-gluconate containing gel respectively. The gas phase of the anaerobic container consisted of 96% N₂ and 4% H₂. A palladium catalyst (BASF, Arnhem, The Netherlands) was used to remove traces of oxygen. Control experiments were done under the same conditions in the absence of either glucose-6-phosphate or 6-phospho-gluconate.

Growth assay on ASW-AA plates

In addition to the activity assays, a growth assay on ASW-AA plates was performed. Plates were prepared by adding 1% (wt/vol) Gelrite to ASW-AA medium, supplemented with 2 ml/L polysulfide solution [62] instead of S⁰, 500 µM agmatine and in the case of SKS001 and SKS002, 10 µM 6MP. Cells inoculated on plates were incubated at 85°C for at least 3 days in an air-tight anaerobic jar containing a N₂/H₂ gas mixture and an Anaerocult A mini (Merck Millipore, Darmstadt, Germany) to ensure anaerobic conditions. *T. kodakarensis* strains used

for inoculation were: SKS001, SKS002, KOD1 (positive control), and Δ hps-phi 7A (negative control).

Determination of growth properties

T. kodakarensis strains were pre-cultured in 120-ml serum bottles containing 40 ml ASW-YT-S⁰ with N₂ as headspace, at 85°C for 15 h (o/n). Next, the cells (1% vol/vol) were transferred to similar bottles containing ASW-YT-SA2 and the cultures were incubated at 85°C until growth was clearly visible. A final round of pre-culturing was performed by inoculating fresh ASW-YT-SA2 containing 120-ml serum bottles until the optical density at 600 nm (OD₆₀₀) reached 0.3 and 0.4, for TS559 and the mutant strains SKS001 and SKS002 respectively. The OD₆₀₀ was measured using an U-1500 spectrophotometer (Hitachi). Final cultures were inoculated with a 1% (vol/vol) inoculum for the mutant strains (OD₆₀₀ = 0.4), and the equivalent of that (assuming a linear relationship between OD₆₀₀ and absolute cell count) for TS559. The growth properties of the mutant strains were compared to that of the parental strain by following the increase in OD₆₀₀ in time in 500-ml anaerobic bottles (average of three independent cultures for each strain) containing 150-ml ASW-YT-SA2 with 10 mM PIPES buffer (pH 7.0 at RT). For the growth experiment with added nucleosides, 120-ml serum bottles (average of 5 independent cultures for each strain) containing 40 ml ASW-YT-SA2 with 10 mM PIPES buffer (pH 7.0 at RT) were used. The gas phase in all bottles was a mixture of N₂ and a fixed amount of CH₄ that was used as an internal standard for the gas measurements described below. Sigma plot application software version 12.3 was used for fitting the growth data and a modified Gompertz equation (**Figure S3.3**) was used to estimate the maximum growth rate of each strain [501]. The samples for the OD measurements, 1 ml and 1.5 ml for the 120-ml and 500-ml bottles respectively, were taken by syringe. After the OD measurements, the 1.5 ml samples were immediately centrifuged (13,000 *g*, 15 min, 4°C), and the supernatant was stored at -20°C for further processing (see section product analysis).

Analysis of substrates and fermentation end products

Gas samples to analyze H₂ and CO₂ production were taken at the same time points as the culture medium samples. Gas samples (500 μ l) were taken using a 2.5-ml air-tight syringe,

and the concentration of both gasses was determined by a gas chromatograph (G-A-S CompactGC), using CH_4 as internal standard. Final gas concentrations are expressed per volume culture medium, and are corrected for sample withdrawal. The GC was equipped with a capillary column (15 m x 0.32 mm; carboxen 1010; Sigma Aldrich; carrier gas, helium) and a TCD detector.

The 1.5-ml supernatant samples stored at -20°C (see above) were used to determine the concentrations of the fermentation products alanine and acetate, as well as the substrate glucose. The concentration of alanine in the supernatant was determined by HPLC after a pre-column derivatization with dabsyl chloride. The derivatization was performed by adding 100 μl NaHCO_3 (0.1 M, pH 9.0), 5 μl internal standard (4 mM norleucine in 200 mM HCl), 5 μl supernatant and 200 μl freshly prepared derivatization reagent (saturated solution of dabsyl chloride in acetonitrile) to glass 1.5-ml HPLC screwcap vials. The mixtures were heated for 20 min at 70°C in an oven. Next, the mixtures were cooled down and 690 μl dilution buffer (25 mM H_3PO_4 in water/acetonitrile 90:10 (v/v), pH 7.0) was added. The derivatives were analyzed using a HPLC apparatus (Thermo) equipped with a Polaris C18-A column (200 x 2.0 mm; Agilent). The column was thermostated at 45°C . Mobile phase A consisted of 24 mM sodium acetate (pH 6.6) – acetonitrile (82:18 v/v). Mobile phase B consisted of isopropanol – acetonitrile (60:40 v/v). The elution was performed as follows: from 0 to 3 min, 30% B; from 3 min to 19 min, a linear gradient from 30 to 45 % B; from 19 min to 21 min, a linear gradient from 45 to 95 % B; from 21 to 25.5 min, 95% B; from 25.5 min to 26 min, a linear gradient from 95 to 20 % B. The flow-rate 0.6 ml/min. The UV-absorbance was measured at 436 nm.

The concentration of acetate was determined using a HPLC (Thermo Scientific, Dionex, ICS5000) apparatus equipped with a metacarb 67H column (300 x 6.5 mm; Agilent) and a refractive index detector. The column was thermostated at 60°C . The mobile phase was 5 mM H_2SO_4 , with a flow-rate of 0.8 ml/min. HPLC-samples were prepared by adding 0.5 ml supernatant and 0.5 ml H_2SO_4 (50 mM) containing 10 mM crotonate as internal standard, to 1.5-ml HPLC screwcap vials.

The concentration of glucose consumed was determined by subtracting the glucose concentration in the supernatant of the samples (t_x), from the glucose concentration present at the start of cultivation (t_0). The glucose concentration present in the supernatant was determined by HPLC after a pre-column enzymatic digestion of the added substrate starch. The enzymatic digestion of starch into glucose was performed by mixing 0.5 ml supernatant and 0.5 ml acetic acid (200 mM, pH 5.5) containing 5 mM CaCl_2 in a 1.5-ml Eppendorf tube, and adding ~50 Units of heat-stable versions of both α -amylglucosidase and α -amylase. The reaction mixtures were incubated at 55°C for 4 h, while shaking (750 rpm). The Eppendorf tubes containing the reaction mixtures were weighed before and after the incubation, so that the increase in concentrations due to water evaporation could be corrected for. Control reactions with similar reaction mixtures containing only one of the enzymes, or control reactions in which the addition of supernatant was replaced by water, were included as well.

Differential proteome analyses

T. kodakarensis strains TS559, SKS001 and SKS002 were cultivated anaerobically in duplicate, in 120-ml stoppered serum bottles containing 40 ml ASW-YT-SA2 as described above. Cells were harvested (OD_{600} of 0.45 and 0.3 for the mutants and TS559, respectively) by centrifugation (5000 *g*, 30 min, 4°C) and the pellets were resuspended in 500 μl 20 mM Tris-HCl (pH 7.8). The cells were lysed by sonication, and the cellular debris was removed by centrifugation (13,000 *g*, 30 min, 4°C). Proteome analyses was performed by the Radboud Proteomics Centre of the Radboud University Nijmegen Medical Centre, Nijmegen, the Netherlands. 5 μg of each of the provided samples (protein concentration determination by use of the 2D-Quant kit, GE Healthcare) was in-solution digested including reduction and alkylation. After digestion the samples were loaded on Omix C18 tips (Agilent) for desalting and concentration and analyzed by nanoflow C18 reversed phase liquid chromatography coupled online to a Maxis 4G +ETD QTOF (Bruker Daltonics). The samples were measured with CID (Collision Induced Dissociation) as fragmentation technique and chromatographic separation was achieved via a linear 90 minutes gradient of 5 to 32% acetonitrile using 0.1% formic acid as ion pair reagent. For data analysis a *T. kodakarensis* KOD1 database, downloaded from NCBI, with added sequences for TM0295, TM0438, TM1154, TM1155 and

TM1718 was used. For database searches, validation and relative quantification the MaxQuant software package was used (version 1.5.0.0).

Abbreviations

PPP: Pentose Phosphate Pathway; oxPPP: oxidative PPP; RuMP: ribulose monophosphate pathway; G6PDH: glucose-6-phosphate dehydrogenase; 6PGDH: 6-phosphogluconate dehydrogenase; SLP: substrate level phosphorylation; Fd_{red} : reduced ferredoxin; Fd_{ox} : oxidized ferredoxin; GDH: glutamate dehydrogenase; ALT: alanine amino transferase; Hyh: soluble hydrogenase; MBH: membrane-bound hydrogenase; 6PGLase: 6-phosphogluconolactonase; Ru5PE: ribulose-5-phosphate 3-epimerase; TAL: transaldolase; 6MP: 6-methyl purine; NBT: nitro blue tetrazolium chloride; G6P: glucose-6-phosphate; 6PG: 6-phospho-gluconate; Ru5P: ribulose-5-phosphate; R5P: ribose-5-phosphate; ASW: artificial seawater; ASW-YT: ASW with yeast extract and tryptone; ASW-AA: ASW with amino acids; SA2; modified potato starch; GAPOR: glyceraldehyde-3-phosphate:ferredoxin oxidoreductase; GAPN: NADP-dependent glyceraldehyde-3-phosphate dehydrogenase; PMS: phenazine methosulfate

Authors' contributions

SKS and SWMK conceived and designed the study. SKS performed all experiments and analyzed the data. SKS and SWMK interpreted the data and wrote the manuscript. TJS and HA provided materials for metabolic engineering. All authors read, revised and approved the final manuscript.

Acknowledgements

This work is part of the Open Programme (“Biological hydrogen production beyond current limits” file number 820.02.011), which is financed by the Netherlands Organisation for Scientific Research (NWO). TJS is supported by funding from the USA Department of Energy, Basic Energy Sciences Division, grant DE-SC0014597. The authors thank Hans Leemhuis of AVEBE FOOD, Veendam, The Netherlands, for kindly providing us with the modified potato starch “Paselli SA2”.

Supplementary figures and tables

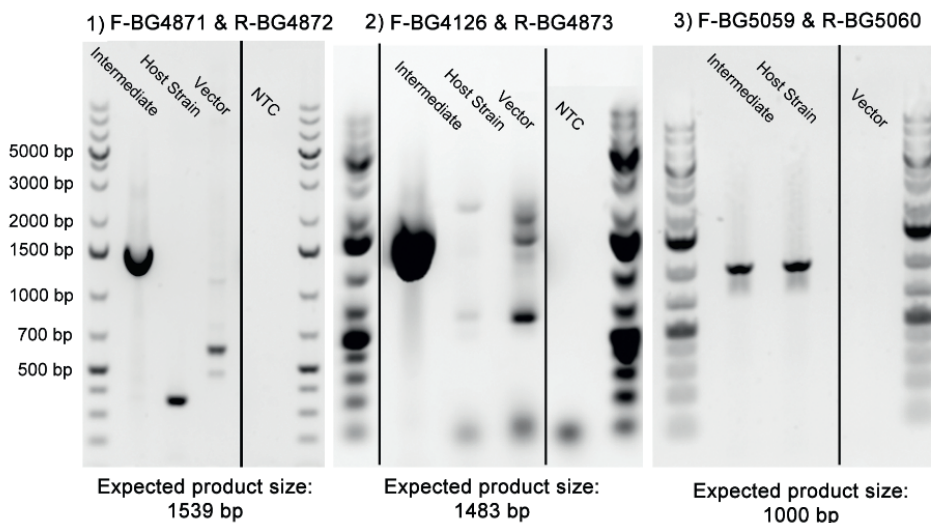


Figure S3.1 Example of diagnostic PCR results of a transformant. Results of diagnostic PCR analysis on isolated genomic DNA of an SKS002 transformant, i.e. an agmatine prototroph. Three different PCR analyses were performed, the results of which are visible in the gels above. On top of each gel, the primer-pair used in the specific reaction is indicated (BG-numbers). As described in the material and methods section, the BG-numbers correspond to the primers that are illustrated in Figure 3.2B by the letters H-I, C-J, and K-L. Below each gel, the expected product size (in bp) of the specific PCR reaction is shown. Besides the isolated genomic DNA of one of the agmatine prototrophs, each analysis also included isolated genomic DNA of the host strain, the original modification vector, and a No Template Control (NTC). The only exception to this is reaction number 3, in which no NTC was included. However, the absence of a product in the reaction containing the modification vector as template DNA, confirms that none of the other ingredients of the reaction was contaminated by genomic DNA. The solid lines visible in some of the gels indicate non-adjacent lanes of the same gel. Gel image #1 confirms integration of the vector (either by single or double recombination) downstream of TK1764. Gel image #2 confirms integration of the vector upstream of TK1765. Together these results suggest that the desired "intermediate recombinant genome" as shown Figure 3.2B, was obtained. Contrary, gel image #3 suggest that the isolated genomic DNA is either that of the host strain or that of a transformant that is the result of a single crossover event (leading to the integration of the entire modification vector). Together, gel images #1-3 suggest that there is a mixture of chromosomes in the preparation labeled "intermediate".

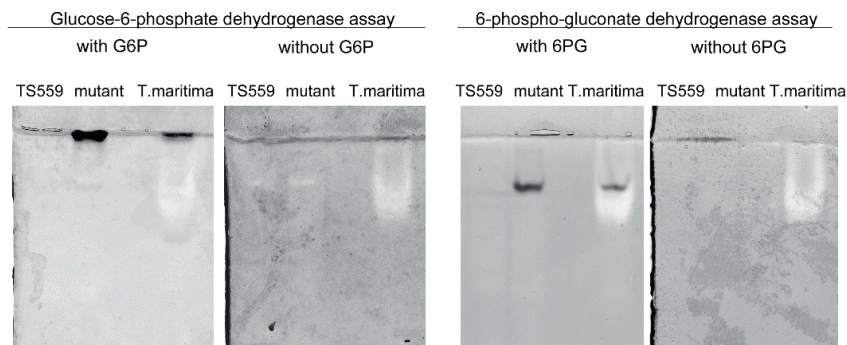


Figure S3.2 In-gel G6PDH and 6PGDH activity assays. cell extracts of the mutant strain SKS001, the parental strain TS559 (negative control) and *T. maritima* MSB8 (positive control), were loaded and resolved on a native polyacrylamide gel. Next, the gel was placed in an activity staining buffer containing NADP^+ , NBT and phenazine methosulfate (PMS) and was analyzed for G6PDH and 6PGDH activity by adding either glucose-6-phosphate (G6P) or 6-phospho-gluconate (6PG), respectively. PMS acts as an intermediate electron carrier for coupling the production of NADPH to the reduction of NBT to insoluble precipitates. When the indicated dehydrogenases are actively expressed precipitates will be formed at their position in the gel

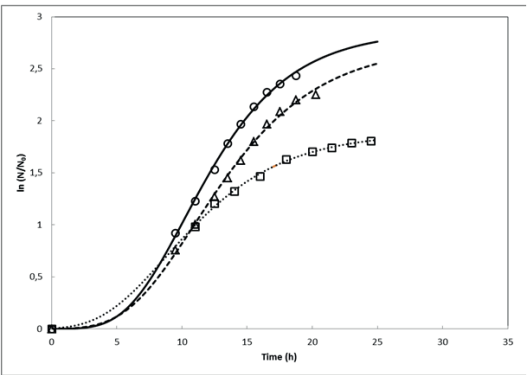


Figure S3.3 Logarithmic plot of growth data and curve fitting of SKS001 (circles), SKS002 (triangles) and TS559 (squares). Data were fitted using a modified Gompertz equation: $y = A \text{EXP}\{-\text{EXP}[\mu_{\text{max}}.e/A(\lambda-t) + 1]\}$ where $y = \ln(N/N_0)$, A = maximum OD, μ = maximum growth rate, λ = lag phase [501]

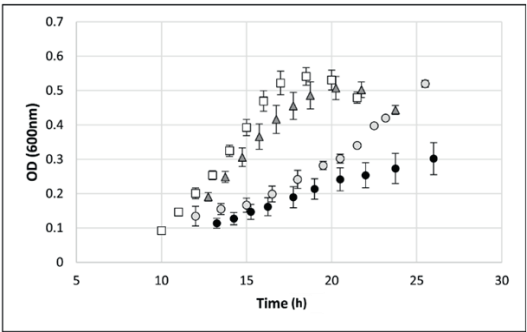
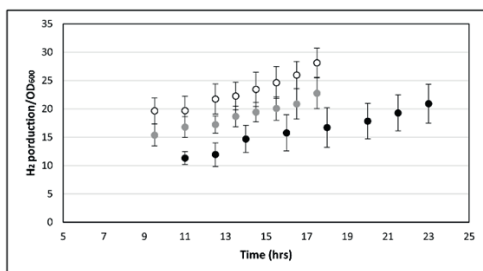


Figure S3.4 Growth properties of *T. kodakarensis* strains TS559 (black circles), SKS001 (white squares) and SKS002 (grey triangles) in ASW-YT-SA2, or of TS559 (light grey circles) in ASW-YT-SA2 supplemented with the nucleosides A (purine) and U (pyrimidine)

A.



B.

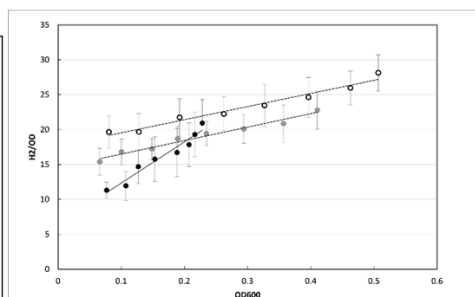


Figure S3.5 Normalized H₂ production by *T. kodakarensis* TS559 (black), SKS001 (white) and SKS002 (grey) in ASW-YT-SA2 (A) Specific H₂ production (H₂/OD₆₀₀) in time (B) Specific H₂ production (H₂/OD₆₀₀) per OD. The data presented in this graph is corrected for the different growth rates of the various strains

Table S3.1 Relative abundance of the differentially expressed proteins in *T. kodakarensis* TS559, SKS001 and SKS002 grown on starch

Protein	Gene identification	SKS001/TS559	SKS002/TS559	SKS002/SKS001
Heterologously expressed proteins				
Translaldolase	[<i>Thermotoga maritima</i> MSB8]	TM0295	n.d.	SKS002
6-phosphogluconate dehydrogenase	[<i>Thermotoga maritima</i> MSB8]	TM0438	SKS001	n.d.*
Sol/DevB family oxidoreductase (6PGLase)	[<i>Thermotoga maritima</i> MSB8]	TM1154	n.d.	SKS002
glucose-6-phosphate 1-dehydrogenase	[<i>Thermotoga maritima</i> MSB8]	TM1155	SKS001	SKS002
ribulose-phosphate 3-epimerase	[<i>Thermotoga maritima</i> MSB8]	TM1718	n.d.	SKS002
DNA metabolism				
<i>Chromosome-associated proteins</i>				
DNA/RNA-binding protein alba		TK0560	2.36	1.76
Histone A		TK1413	51.77	2.45
<i>Replication</i>				
DNA-dependent DNA polymerase family B protein		TK0001	1.43	0.96
DNA polymerase sliding clamp		TK0535	1.76	2.07
DNA topoisomerase VI subunit A		TK0798	1.45	1.87
DNA topoisomerase VI subunit B		TK0799	1.22	1.45
DNA polymerase sliding clamp		TK0535	1.76	2.07
DNA topoisomerase VI subunit A		TK0798	1.45	1.87
DNA topoisomerase VI subunit B		TK0799	1.22	1.45
flap endonuclease-1		TK1281	n.d.	SKS002
DNA primase		TK1410	1.08	1.31
DNA primase large subunit		TK1790	1.52	0.82
DNA polymerase II large subunit		TK1903	n.d.	SKS002
replication factor A complex, RPA32 subunit		TK1959	0.80	1.27
replication factor C small subunit		TK2218	1.49	1.57

replication factor C large subunit	TK2219	n.d.	SKS002	SKS002
<i>Recombination and repair</i>				
endonuclease V	TK0906	n.d.	SKS002	SKS002
AP endonuclease	TK1165	SKS001	n.d.	SKS001
DNA repair and recombination protein RadA	TK1899	1.36	0.87	0.64
bipolar DNA helicase	TK2213	1.15	0.99	0.86
<i>DNA degradation and restriction</i>				
RecJ-like exonuclease	TK0155	1.36	1.33	0.98
putative 5-methylcytosine restriction system, catalytic subunit	TK0794	TS559	0.65	SKS002
putative 5-methylcytosine restriction system, GTPase subunit	TK0795	0.58	0.70	1.21
type II restriction endonuclease	TK1147	SKS001	n.d.	SKS001
endonuclease-methyltransferase fusion protein	TK1460	SKS001	SKS002	1.79
Transcription				
<i>RNA polymerase</i>				
DNA-directed RNA polymerase subunit M	TK0533	SKS001	n.d.	SKS001
DNA-directed RNA polymerase subunit P	TK0616	SKS001	n.d.	SKS001
DNA-directed RNA polymerase subunit A	TK1081	1.12	1.30	1.16
DNA-directed RNA polymerase subunit A	TK1082	1.30	1.61	1.24
DNA-directed RNA polymerase subunit B	TK1083	1.27	1.22	0.96
<i>Transcription factors</i>				
transcription factor	TK0126	0.65	0.78	1.20
transcription factor	TK0132	1.54	1.67	1.08
transcription elongation factor NusA-like protein	TK0481	SKS001	n.d.	SKS001
transcription elongation factor NusA-like protein	TK1079	SKS001	SKS002	0.71
transcription elongation factor NusA	TK1112	1.04	0.85	0.82
TBP-interacting protein Tip49-like protein	TK1199	1.14	1.45	1.28
transcription initiation factor IIB	TK1280	0.92	1.09	1.18

transcription antitermination protein NusG	TK1419	1.05	1.06	1.00
nascent polypeptide-associated complex protein	TK1973	0.76	0.97	1.27
transcription initiation factor E subunit alpha	TK2024	TS559	TS559	n.d.
Translation				
<i>Ribosomal proteins</i>				
30S ribosomal protein S10P	TK0307	1.18	1.28	1.08
50S ribosomal protein L21e	TK0902	TS559	TS559	n.d.
50S ribosomal protein L35	TK0969	n.d.	SKS002	SKS002
30S ribosomal protein S7	TK1077	0.63	0.73	1.16
30S ribosomal protein S12	TK1078	0.95	1.11	1.17
50S ribosomal protein L30e	TK1080	SKS001	n.d.	SKS001
30S ribosomal protein S8e	TK1191	0.99	0.90	0.91
30S ribosomal protein S3Ae	TK1254	0.53	0.42	0.79
30S ribosomal protein S19e	TK1276	1.25	1.62	1.30
30S ribosomal protein S28e	TK1310	0.92	0.93	1.01
50S ribosomal protein L7Ae	TK1311	1.21	1.25	1.04
50S ribosomal protein L31e	TK1320	SKS001	SKS002	1.13
50S ribosomal protein LX	TK1322	0.97	0.91	0.93
50S ribosomal protein L12	TK1415	1.88	1.31	0.70
acidic ribosomal protein P0	TK1416	1.16	1.13	0.98
50S ribosomal protein L1P	TK1417	1.23	1.08	0.88
50S ribosomal protein L11	TK1418	1.15	1.15	1.00
30S ribosomal protein S2	TK1496	1.06	1.14	1.07
30S ribosomal protein S9	TK1500	0.68	0.67	0.97
50S ribosomal protein L13	TK1501	1.32	1.51	1.14
30S ribosomal protein S11	TK1504	TS559	TS559	n.d.
30S ribosomal protein S4	TK1505	SKS001	n.d.	SKS001

50S ribosomal protein L15	TK1519	SKS001	SKS002	0.99
50S ribosomal protein L30	TK1520	SKS001	SKS002	1.40
30S ribosomal protein S5	TK1521	1.08	1.04	0.96
50S ribosomal protein L19e	TK1523	SKS001	n.d.	SKS001
50S ribosomal protein L32e	TK1524	SKS001	SKS002	1.14
50S ribosomal protein L6	TK1525	1.30	1.74	1.34
30S ribosomal protein S8	TK1526	n.d.	SKS002	SKS002
30S ribosomal protein S14	TK1527	n.d.	SKS002	SKS002
50S ribosomal protein L5	TK1528	0.69	1.10	1.60
50S ribosomal protein L14	TK1531	1.22	1.15	0.94
30S ribosomal protein S3	TK1536	1.21	1.20	1.00
30S ribosomal protein S19	TK1538	1.09	TS559	SKS001
50S ribosomal protein L2	TK1539	1.41	1.99	1.41
50S ribosomal protein L4P	TK1541	SKS001	n.d.	SKS001
50S ribosomal protein L3P	TK1542	1.26	1.34	1.06
50S ribosomal protein L10e	TK1546	SKS001	n.d.	SKS001
30S ribosomal protein S24e	TK1696	1.31	1.57	1.20
30S ribosomal protein S17e	TK2292	TS559	TS559	n.d.
<i>tRNA aminoacylation</i>				
cysteinyI-tRNA synthetase	TK0444	SKS001	n.d.	SKS001
aspartyl-tRNA synthetase	TK0492	1.45	1.00	0.69
tyrosyl-tRNA synthetase	TK0568	1.61	1.26	0.78
asparaginyl-tRNA synthetase	TK0759	2.17	2.06	0.95
glutamyl-tRNA(Gln) amidotransferase subunit D	TK0908	SKS001	SKS002	0.99
glutamyl-tRNA(Gln) amidotransferase subunit E	TK0911	1.27	1.32	1.04
phenylalanyl-tRNA synthetase subunit alpha	TK0921	TS559	1.05	SKS002
glycyl-tRNA synthetase	TK0978	1.46	1.31	0.90

seryl-tRNA synthetase	TK1140	SKS001	n.d.	SKS001
threonyl-tRNA synthetase	TK1170	1.12	1.41	1.26
glutamyl-tRNA synthetase	TK1408	1.27	0.90	0.70
alanyl-tRNA synthetase	TK1567	1.03	1.11	1.08
isoleucyl-tRNA synthetase	TK1748	SKS001	SKS002	0.86
tryptophanyl-tRNA synthetase	TK1874	1.62	1.71	1.05
<i>tRNA/rRNA modification and processing</i>				
fibrillarin	TK0183	1.71	1.34	0.78
C/D box methylation guide ribonucleoprotein complex aNOP56 subunit	TK0184	1.03	1.22	1.19
N(2),N(2)-dimethylguanosine tRNA methyltransferase	TK0970	1.51	1.10	0.73
exosome complex RNA-binding protein Csl4	TK1169	SKS001	n.d.	SKS001
cleavage and polyadenylation specificity factor subunit-like protein	TK1428	1.14	1.28	1.12
exosome complex exonuclease Rrp41	TK1634	1.02	1.43	1.40
exosome complex RNA-binding protein Rrp4	TK1635	1.51	1.70	1.12
H/ACA RNA-protein complex component Gar1	TK2286	0.93	0.78	0.84
<i>mRNA processing</i>				
small nuclear ribonucleoprotein	TK0976	0.93	1.57	1.68
nonsense-mediated mRNA decay protein NMD3-like protein	TK1297	1.21	1.29	1.06
<i>Translation factors</i>				
zinc-ribbon RNA-binding protein involved in translation	TK0111	n.d.	SKS002	SKS002
translation initiation factor IF-2	TK0185	1.35	1.09	0.81
elongation factor 1-alpha	TK0308	1.68	1.56	0.93
elongation factor EF-2	TK0309	1.65	1.05	0.64
translation initiation factor IF-2	TK0556	1.72	TS559	SKS001
translation initiation factor IF-1	TK0802	SKS001	n.d.	SKS001
translation initiation factor IF-5A	TK0878	1.43	1.07	0.75
translation initiation factor IF-2	TK1100	1.54	1.42	0.92

peptide chain release factor 1	TK1239	SKS001	n.d.	SKS001
translation initiation factor Sui1	TK1534	1.49	1.12	0.75
translation initiation factor IF-2	TK1621	0.76	1.86	2.43
translation initiation factor IF-2	TK1946	1.57	1.68	1.07
<i>General prediction only</i>				
translation-associated GTPase	TK0506	1.00	0.99	0.99
Cellular processes				
<i>Cell division</i>				
ParA/Mind family ATPase	TK0349	0.89	1.23	1.38
cell division protein CDC48	TK0669	0.82	1.04	1.27
cell division protein pelota-like protein	TK0964	SKS001	n.d.	SKS001
chromosome segregation ATPase	TK1017	1.35	1.16	0.86
cell division protein CDC48	TK1157	0.66	0.94	1.43
cell division protein FtsZ	TK1421	0.96	0.83	0.86
ParA/Mind family ATPase	TK2007	1.20	0.89	0.74
ParA/Mind family ATPase	TK2021	SKS001	n.d.	SKS001
ParA/Mind family ATPase	TK2137	TS559	TS559	n.d.
chromosome segregation protein	TK2211	TS559	TS559	n.d.
cell division protein FtsZ	TK2271	0.36	0.33	0.90
ParA/Mind family ATPase	TK2272	SKS001	SKS002	1.13
<i>Signal transduction</i>				
signaling protein	TK1113	SKS001	SKS002	1.85
<i>Cell motility</i>				
flagellin	TK0038	TS559	TS559	n.d.
flagellin	TK0039	0.65	0.74	1.14
flagella-related protein C	TK0043	0.35	0.33	0.94
flagella-related protein D	TK0044	0.54	0.60	1.11

flagella-related protein I, secretion ATPase	TK0048	0.55	0.44	0.80
flagellar assembly protein J	TK0049	TS559	0.37	SKS002
methyl-accepting chemotaxis protein	TK0156	0.49	0.45	0.93
chemotaxis signal transduction protein	TK0629	0.60	0.67	1.11
methyl-accepting chemotaxis protein	TK0630	0.68	0.83	1.23
chemotaxis response regulator	TK0632	0.89	0.63	0.71
chemotaxis-specific methyltransferase	TK0633	SKS001	n.d.	SKS001
chemotaxis histidine kinase, frame shift, partial	TK0634	0.53	0.62	1.17
methyl-accepting chemotaxis protein	TK0638	1.04	0.82	0.79
methyl-accepting chemotaxis protein	TK1606	0.32	0.37	1.19
methyl-accepting chemotaxis protein	TK2147	0.96	0.76	0.79
<i>Detoxification</i>				
superoxide reductase	TK0525	0.34	0.58	1.69
peroxiredoxin	TK0537	0.68	0.58	0.85
NAD(P)H:rubredoxin oxidoreductase	TK0828	3.83	TS559	SKS001
Protein fate				
<i>Protein and peptide degradation</i>				
membrane protease subunit stomatin/prohibitin-like protein	TK0348	0.86	0.70	0.81
endopeptidase IV	TK1164	1.32	0.96	0.73
M50 family metallopeptidase	TK1247	SKS001	n.d.	SKS001
ATP-dependent protease Lon	TK1264	0.65	0.69	1.05
membrane protease subunit stomatin/prohibitin-like protein	TK1316	1.08	1.11	1.03
proteasome subunit beta	TK1429	1.49	1.37	0.91
proteasome subunit alpha	TK1637	1.42	1.70	1.20
membrane-associated metalloprotease	TK1820	0.99	1.05	1.06
aminopeptidase family M28 protein	TK1912	1.12	1.33	1.19
proteasome subunit beta 2	TK2207	1.87	1.72	0.92

<i>Protein and peptide secretion and trafficking</i>				
signal peptidase	TK0053	1.11	1.19	1.07
signal recognition particle GTPase	TK0653	SKS001	n.d.	SKS001
signal recognition particle protein Srp54	TK1486	0.75	0.84	1.12
preprotein translocase subunit SecY	TK1518	1.20	1.16	0.97
preprotein translocase subunit SecF	TK1593	SKS001	SKS002	0.84
preprotein translocase subunit SecD	TK1594	1.26	1.04	0.83
type II/IV secretion system ATPase	TK1853	SKS001	n.d.	SKS001
<i>Protein folding and stabilization</i>				
fused FKBP-type peptidyl-prolyl cis-trans isomerase/uncharacterized protein	TK0441	2.05	1.96	0.96
prefoldin subunit beta	TK0643	0.90	2.18	2.43
heat shock protein HtpX	TK0677	1.30	2.09	1.61
chaperonin subunit alpha	TK0678	2.06	1.30	0.63
prefoldin subunit alpha	TK1005	0.89	1.96	2.19
small heat shock protein	TK1155	0.52	0.81	1.56
chaperonin beta subunit	TK2303	2.40	1.60	0.67
<i>Protein modification and repair</i>				
protein disulfide oxidoreductase	TK1085	SKS001	n.d.	SKS001
hydrogenase maturation protein HypF	TK1997	n.d.	SKS002	SKS002
hydrogenase expression/formation protein hypD	TK2000	SKS001	n.d.	SKS001
hydrogenase nickel incorporation protein	TK2008	SKS001	SKS002	1.19
Regulatory functions				
<i>Transcription regulators</i>				
cell wall-binding repeat 2-/DNA-binding HTH domain-containing transcription regulator	TK0142	0.68	0.75	1.11
ArsR family transcriptional regulator	TK0225	n.d.	SKS002	SKS002
transcriptional regulator	TK0471	2.32	2.74	1.18
ArsR family transcriptional regulator	TK0559	SKS001	n.d.	SKS001

transcriptional regulator	TK1272	0.93	0.97	1.04
Lrp/AsnC family transcriptional regulator	TK1331	SKS001	n.d.	SKS001
transcriptional regulator	TK1591	1.15	1.36	1.18
transcriptional regulator	TK1688	SKS001	n.d.	SKS001
ArsR family transcriptional regulator	TK1881	1.00	TS559	SKS001
Lrp/AsnC family transcriptional regulator	TK1955	1.88	TS559	SKS001
YpuH-like family transcriptional regulator	TK1962	1.09	1.12	1.02
Lrp/AsnC family transcriptional regulator	TK2052	0.99	1.07	1.08
Lrp/AsnC family transcriptional regulator	TK2110	SKS001	n.d.	SKS001
ExsB family transcriptional regulator	TK2263	n.d.	SKS002	SKS002
<i>Others</i>				
bis(5-adenosyl)-triphosphatase	TK0868	1.69	1.08	0.64
putative ATPase RIL	TK1031	1.53	0.90	0.59
amino acid metabolism regulator	TK1227	3.28	1.91	0.58
phosphate transport regulator	TK1457	SKS001	n.d.	SKS001
proteasome-activating nucleotidase	TK2252	SKS001	SKS002	0.64
Cell envelope				
<i>Lipid metabolism</i>				
geranylgeranyl hydrogenase	TK1088	SKS001	n.d.	SKS001
undecaprenyl diphosphate synthase	TK1173	1.20	1.15	0.95
bifunctional short-chain isoprenyl diphosphate synthase	TK1468	1.26	0.91	0.72
<i>Surface polysaccharides and protein glycosylation</i>				
glucosamine--fructose-6-phosphate aminotransferase	TK0809	TS559	TS559	n.d.
oligosaccharyl transferase	TK0810	1.15	TS559	SKS001
glycosyl transferase family protein	TK0949	TS559	TS559	n.d.
sugar-phosphate nucleotidyltransferase	TK0955	2.19	1.52	0.69
phosphohexomutase	TK1108	0.98	1.58	1.61

bifunctional mannose-6-phosphate isomerase/mannose-1-phosphate guanylyl transferase	TK1109	0.96	TS559	SKS001
UDP-N-acetylglucosamine 2-epimerase	TK1230	1.36	1.07	0.79
UDP-N-acetyl-D-mannosaminuronate dehydrogenase	TK1231	1.42	1.22	0.86
glycosyl transferase family protein	TK1714	0.97	1.43	1.47
glycosyl transferase family protein	TK1721	SKS001	n.d.	SKS001
glycosyl transferase family protein	TK1723	0.49	0.43	0.89
fused dolichol-phosphate mannosyltransferase/uncharacterized protein	TK1732	SKS001	SKS002	1.25
fused dolichol-phosphate mannosyltransferase/GtrA-like domain-containing protein	TK1909	1.55	1.94	1.25
UDP-N-acetylglucosamine--dolichyl-phosphate N-acetylglucosaminephosphotransferase	TK2254	TS559	TS559	n.d.
bifunctional phosphatase/dolichol-phosphate glucosyltransferase	TK2255	0.91	1.02	1.12
<i>Surface structure</i>				
S-layer protein	TK0895	1.04	1.08	1.04
Central intermediary metabolism				
<i>Extracellular protein degradation</i>				
thiol protease	TK1295	SKS001	SKS002	1.23
serine protease	TK2168	1.27	1.40	1.11
<i>Polysaccharide degradation</i>				
maltodextrin phosphorylase	TK1406	SKS001	n.d.	SKS001
alpha-glucosidase	TK1434	TS559	TS559	n.d.
alpha-amylase	TK1436	SKS001	n.d.	SKS001
amylopullulanase	TK1774	1.23	1.23	1.00
4-alpha-glucanotransferase	TK1809	SKS001	n.d.	SKS001
<i>Polyamine</i>				
agmatinase	TK0882	SKS001	n.d.	SKS001
<i>Compatible solute</i>				
myo-inositol-1-phosphate synthase	TK2278	8.67	7.12	0.82

<i>Others</i>				
anaerobic glycerol 3-phosphate dehydrogenase	TK1393	SKS001	n.d.	SKS001
Amino acid metabolism				
<i>Aromatic amino acid family</i>				
3-phosphoshikimate 1-carboxyvinyltransferase	TK0263	TS559	TS559	n.d.
3-deoxy-7-phosphoheptulonate synthase	TK0268	SKS001	SKS002	0.34
<i>Aspartate family</i>				
S-adenosylmethionine synthetase	TK0545	3.07	1.15	0.37
adenylosuccinate synthetase	TK1002	1.72	TS559	SKS001
L-asparaginase	TK2246	1.96	1.72	0.88
<i>Glutamate family</i>				
proline dehydrogenase subunit beta	TK0122	TS559	TS559	n.d.
glutamine amidotransferase, class I	TK0187	1.14	0.97	0.85
arginase	TK0474	1.98	1.75	0.88
adenylosuccinate synthetase	TK1002	1.72	TS559	SKS001
glutamate dehydrogenase	TK1431	0.90	0.72	0.80
glutamine synthetase	TK1796	7.25	0.85	0.12
<i>Serine family</i>				
2-amino-3-ketobutyrate coenzyme A ligase	TK2217	0.70	0.98	1.39
Nucleotide metabolism				
<i>Purines</i>				
inosine 5-monophosphate dehydrogenase	TK0194	0.60	0.52	0.86
phosphoribosylglycinamide formyltransferase 2	TK0207	0.56	TS559	SKS001
adenylosuccinate synthetase	TK1002	1.72	TS559	SKS001
<i>Pyrimidines</i>				
FAD-dependent thymidylate synthase	TK0870	SKS001	n.d.	SKS001
CTP synthetase	TK1193	0.43	TS559	SKS001

orotate phosphoribosyltransferase	TK2138	1.25	0.56	0.45
aspartate carbamoyltransferase regulatory subunit	TK2195	TS559	TS559	n.d.
<i>2-Deoxyribonucleotide metabolism</i>				
ribonucleotide-diphosphate reductase subunit alpha	TK1736	1.39	1.00	0.72
<i>Nucleotide and nucleoside interconversion</i>				
uridylate kinase	TK0305	0.76	0.49	0.65
putative kinase	SKS001	n.d.	SKS001	
nucleoside diphosphate kinase	TK1307	n.d.	SKS002	SKS002
cytidylate kinase	TK1514	TS559	TS559	n.d.
adenylate kinase	TK1517	n.d.	SKS002	SKS002
<i>Purine and pyrimidine salvage</i>				
thymidine phosphorylase	TK0352	SKS001	SKS002	1.29
adenine phosphoribosyltransferase	TK1489	1.08	0.83	0.77
xanthine/guanine phosphoribosyltransferase	TK1737	SKS001	SKS002	1.04
Cofactor and prosthetic group metabolism				
<i>Molybdopterin</i>				
molybdenum cofactor biosynthesis protein B	TK0544	1.07	1.79	1.67
molybdenum cofactor biosynthesis protein A	TK1228	SKS001	n.d.	SKS001
molybdopterin converting factor subunit 2	TK2115	TS559	TS559	n.d.
molybdopterin converting factor subunit 1	TK2118	1.83	1.21	0.66
molybdenum cofactor biosynthesis protein A	TK2225	1.31	1.42	1.08
<i>Pantothenate and CoA</i>				
bifunctional phosphopantothienoylcysteine decarboxylase/ phosphopantothenate synthase	TK0517	SKS001	SKS002	1.02
<i>Pyridine nucleotide</i>				
ATP-NAD kinase	TK1305	1.22	1.13	0.92
<i>Pyridoxine</i>				

pyridoxal biosynthesis lyase PdxS	TK0217	0.54	0.53	0.98
<i>Others</i>				
6-pyruvoyl-tetrahydropterin synthase	TK0563	SKS001	n.d.	SKS001
biotin biosynthesis protein BioY	TK0748	SKS001	n.d.	SKS001
3-octaprenyl-4-hydroxybenzoate carboxy-lyase	TK0929	1.77	TS559	SKS001
lipoate-protein ligase A, C-terminal section, partial	TK1234	1.01	TS559	SKS001
cysteine desulfurase	TK1990	1.33	1.22	0.92
Energy metabolism				
<i>Amino acid degradation</i>				
indolepyruvate oxidoreductase subunit beta	TK0135	1.13	1.74	1.53
indolepyruvate: ferredoxin oxidoreductase subunit alpha	TK0136	1.33	1.53	1.15
multiple substrate aminotransferase	TK0186	3.09	2.59	0.84
acetyl-CoA synthetase I subunit beta	TK0465	0.76	0.77	1.01
aspartate aminotransferase	TK0548	1.04	0.92	0.88
acetyl-CoA synthetase I subunit alpha	TK0665	1.26	1.06	0.84
acetyl-CoA synthetase II subunit beta	TK0943	0.91	0.62	0.68
acyl-CoA synthetase large subunit	TK0944	0.93	0.91	0.97
alanine aminotransferase	TK1094	0.35	0.43	1.24
2-oxoglutarate ferredoxin oxidoreductase subunit gamma	TK1123	TS559	TS559	n.d.
2-oxoglutarate ferredoxin oxidoreductase subunit beta	TK1124	TS559	TS559	n.d.
2-oxoglutarate ferredoxin oxidoreductase subunit alpha	TK1125	0.24	TS559	SKS001
2-oxoglutarate ferredoxin oxidoreductase subunit beta	TK1129	TS559	TS559	n.d.
2-oxoglutarate ferredoxin oxidoreductase subunit alpha	TK1130	TS559	TS559	n.d.
2-oxoacid:ferredoxin oxidoreductase subunit delta	TK1131	0.22	TS559	SKS001
indolepyruvate: ferredoxin oxidoreductase subunit alpha	TK1643	0.45	0.33	0.74
succinyl-CoA synthetase large subunit	TK1880	0.97	0.86	0.89
pyruvate/ketoisovalerate ferredoxin oxidoreductase subunit gamma	TK1978	0.98	1.18	1.21

2-ketoisovalerate ferredoxin oxidoreductase subunit delta	TK1979	1.32	0.90	0.68
2-ketoisovalerate ferredoxin oxidoreductase subunit alpha	TK1980	1.88	1.04	0.56
2-ketoisovalerate ferredoxin oxidoreductase subunit beta	TK1981	1.19	1.23	1.03
pyruvate ferredoxin oxidoreductase subunit delta	TK1982	0.85	1.49	1.74
pyruvate ferredoxin oxidoreductase subunit alpha	TK1983	1.23	1.15	0.94
pyruvate ferredoxin oxidoreductase subunit beta	TK1984	1.29	1.52	1.18
indolepyruvate oxidoreductase subunit B	TK2244	n.d.	SKS002	SKS002
aspartate aminotransferase	TK2268	0.01	TS559	SKS001
<i>Glycolysis/gluconeogenesis</i>				
ADP-specific phosphofructokinase	TK0376	n.d.	SKS002	SKS002
NADP-dependent glyceraldehyde-3-phosphate dehydrogenase	TK0705	1.07	0.94	0.88
cofactor-independent phosphoglycerate mutase	TK0866	TS559	TS559	n.d.
fructose-bisphosphate aldolase	TK0989	SKS001	n.d.	SKS001
phosphohexomutase	TK1108	0.98	1.58	1.61
glucose-6-phosphate isomerase	TK1111	SKS001	n.d.	SKS001
phosphoenolpyruvate synthase	TK1292	1.20	1.17	0.98
triosephosphate isomerase	TK2129	2.31	1.46	0.63
tungsten-containing glyceraldehyde-3-phosphate:ferredoxin oxidoreductase	TK2163	1.67	1.72	1.03
fructose-1,6-bisphosphatase	TK2164	SKS001	SKS002	1.01
<i>Modified ED pathway</i>				
glucose-1-dehydrogenase	TK0673	TS559	TS559	n.d.
<i>Pentose metabolism</i>				
bifunctional D-arabino 3-hexulose-6-phosphate formaldehyde lyase/phosphohexuloisomerase	TK0475	TS559	TS559	n.d.
ribose-5-phosphate isomerase A	TK1426	1.30	2.06	1.58
<i>Pyruvate/oxaloacetate metabolism</i>				
pyruvate carboxylase subunit B	TK0990	0.58	0.66	1.14

<i>Propanoate metabolism</i>				
methyilmalonyl-CoA mutase, C-terminus of large subunit, partial	TK0328	TS559	TS559	n.d.
methyilmalonyl-CoA epimerase	TK0330		0.40	0.30
methyilmalonyl-CoA decarboxylase subunit alpha	TK1622		0.33	0.34
methyilmalonyl-CoA decarboxylase subunit delta	TK1623		0.41	0.70
putative acetyl-CoA carboxylase biotin carboxyl carrier protein subunit	TK1624		0.38	0.42
<i>ATP-<i>proton motive force</i> interconversion **</i>				
V-type ATP synthase subunit H	TK1596		0.54	0.48
V-type ATP synthase subunit I	TK1597		0.52	0.56
V-type ATP synthase subunit K	TK1598		TS559	TS559
V-type ATP synthase subunit E	TK1599		0.69	0.70
V-type ATP synthase subunit C	TK1600		0.81	0.63
V-type ATP synthase subunit F	TK1601		0.64	0.71
V-type ATP synthase subunit A	TK1602		0.74	0.74
V-type ATP synthase subunit B	TK1603		0.67	0.97
V-type ATP synthase subunit D	TK1604		0.46	0.68
<i>Electron transport</i>				
tungsten-containing oxidoreductase	TK0072		TS559	TS559
NAD(P)H-flavin oxidoreductase	TK0284		n.d.	SKS002
rubrerythrin-like protein	TK0650		1.15	0.75
short chain dehydrogenase	TK0736		TS559	TS559
type A flavoprotein	TK0814		0.86	0.86
tungsten-containing oxidoreductase (aldehyde ferredoxin oxidoreductase)	TK0844		0.59	0.73
rubrerythrin-like protein	TK1056		0.84	TS559
glutaredoxin-like protein	TK1208		n.d.	SKS001
NADH:polysulfide oxidoreductase	TK1481		0.88	SKS002
putative oxidoreductase	TK1684		0.62	0.80
				0.57
				0.93

ferredoxin-NADP(+) reductase subunit alpha	TK1685	1.74	TS559	SKS001
ferredoxin 1	TK1694	n.d.	SKS002	SKS002
cytosolic NiFe-hydrogenase subunit beta	TK2072	0.22	0.24	1.06
4Fe-4S binding protein	TK2073	n.d.	SKS002	SKS002
putative glutamate synthase subunit beta	TK2074	0.43	0.41	0.94
4Fe-4S binding protein	TK2075	SKS001	SKS002	0.93
formate dehydrogenase subunit alpha	TK2076	1.13	1.13	0.99
4Fe-4S binding protein	TK2077	1.13	1.05	0.93
putative monovalent cation/H ⁺ antiporter subunit E	TK2080	TS559	TS559	n.d.
putative monovalent cation/H ⁺ antiporter subunit G	TK2082	TS559	TS559	n.d.
membrane bound hydrogenase subunit MbHc	TK2084	SKS001	n.d.	SKS001
membrane bound hydrogenase subunit MbHl	TK2088	n.d.	SKS002	SKS002
membrane bound hydrogenase, NiFe-hydrogenase small subunit	TK2089	1.10	0.91	0.82
membrane bound hydrogenase, NiFe-hydrogenase large subunit 1	TK2090	1.06	1.11	1.05
membrane bound hydrogenase, NiFe-hydrogenase large subunit 2	TK2091	1.02	1.05	1.03
membrane bound hydrogenase, MbHm subunit	TK2092	n.d.	SKS002	SKS002
NADH-plastoquinone oxidoreductase subunit	TK2093	n.d.	SKS002	SKS002
rubrerythrin-like protein	TK2095	SKS001	n.d.	SKS001
<i>Others</i>				
inorganic pyrophosphatase	TK1700	1.31	1.29	0.99
Transport and binding proteins				
<i>Anions</i>				
ABC-type sulfate transport system, ATPase component	TK0021	0.78	TS559	SKS001
phosphate ABC transporter ATP-binding protein	TK1868	TS559	TS559	n.d.
sodium/phosphate symporter	TK2061	TS559	TS559	n.d.
<i>Cations</i>				
molybdate ABC transporter periplasmic substrate-binding protein	TK0015	1.20	1.13	0.94

iron(III) ABC transporter periplasmic protein	TK0570	SKS001	n.d.	SKS001
iron(III) ABC transporter permease	TK0571	1.89	1.04	0.55
ABC-type iron(III) transport system, ATPase component	TK0572	1.62	1.47	0.91
ABC-type manganese/zinc transport system periplasmic protein	TK0661	0.83	1.04	1.26
iron(III)-siderophore ABC transporter periplasmic subunit	TK0706	1.62	1.67	1.03
ABC-type iron(III)-siderophore transport system, ATPase component	TK0708	n.d.	SKS002	SKS002
molybdate ABC transporter permease	TK0718	TS559	0.67	SKS002
molybdate ABC transporter ATPase	TK0719	0.65	0.80	1.22
ABC-type manganese/zinc transport system, ATPase component	TK0803	1.43	1.04	0.72
heavy-metal transporting P-type ATPase	TK0837	0.75	0.94	1.25
ABC-type sodium efflux pump system permease	TK1336	TS559	TS559	n.d.
ABC-type sodium efflux pump system, ATPase component	TK1337	0.69	0.61	0.87
<i>Carbohydrates and organic compounds</i>				
xanthine/uracilpermease	TK0157	0.97	TS559	SKS001
SNF family sodium-dependent transporter	TK1432	TS559	TS559	n.d.
ABC-type maltodextrin transport system maltodextrin-binding periplasmic protein	TK1771	1.61	1.41	0.88
ABC-type maltodextrin transport system permease	TK1773	1.44	1.04	0.72
ABC-type maltodextrin transport system, ATPase component	TK1775	1.03	0.78	0.75
formate transporter	TK2079	n.d.	SKS002	SKS002
<i>Peptides and amino acids</i>				
proton/glutamate symporter	TK0986	3.77	2.83	0.75
dipeptide/oligopeptide ABC transporter ATPase	TK1800	0.65	0.49	0.75
dipeptide/oligopeptide ABC transporter ATPase	TK1801	0.65	0.63	0.96
dipeptide/oligopeptide ABC transporter permease	TK1802	0.69	0.48	0.69
dipeptide/oligopeptide ABC transporter permease	TK1803	0.77	0.64	0.83
peptide ABC transporter periplasmic protein	TK1804	0.29	0.27	0.93
lipoprotein releasing system, ATP-binding protein	TK1859	TS559	TS559	n.d.

<i>General</i>				
ABC transporter ATPase	TK0019	SKS001	n.d.	SKS001
ABC transporter periplasmic protein	TK0657	1.78	1.52	0.85
ABC transporter ATPase	TK0658	1.25	0.86	0.68
ABC transporter permease	TK0659	TS559	TS559	n.d.
ABC transporter permease	TK0660	2.89	TS559	SKS001
multidrug ABC transporter ATPase and permease	TK0676	0.52	0.47	0.90
multidrug ABC transporter ATPase	TK0694	0.51	0.43	0.85
multidrug ABC transporter ATPase	TK0727	TS559	0.40	SKS002
multidrug ABC transporter ATPase	TK1579	0.48	0.62	1.29
multidrug ABC transporter ATPase	TK2053	0.97	1.07	1.10
major facilitator superfamily permease	TK2266	TS559	TS559	n.d.
multidrug ABC transporter ATPase	TK2281	TS559	0.53	SKS002
General function predicted only/Unknown				
nucleic acid-binding protein	TK0137	1.42	0.89	0.63
metal-dependent hydrolase	TK0141	1.51	1.38	0.91
HAD superfamily hydrolase	TK0477	SKS001	SKS002	0.96
HD superfamily metal-dependent phosphohydrolase	TK0540	SKS001	n.d.	SKS001
calcineurin superfamily metallophosphoesterase	TK0574	1.18	1.12	0.94
DHH family phosphohydrolase	TK0645	n.d.	SKS002	SKS002
ATPase	TK0754	0.80	1.01	1.27
putative RNA-processing protein	TK0800	1.53	1.66	1.09
ATPase	TK0953	1.17	1.93	1.66
calcineurin superfamily metallophosphoesterase	TK1037	0.70	0.66	0.95
acetyltransferase	TK1174	SKS001	SKS002	3.66
GTPase	TK1190	1.41	2.15	1.53
RNA-binding protein	TK1257	1.24	TS559	SKS001

HAD superfamily hydrolase	TK1383	0.90	0.87	0.97
phosphomannomutase	TK1404	0.72	0.53	0.73
fibronectin-binding protein	TK1438	SKS001	SKS002	1.14
ATPase	TK1549	TS559	TS559	n.d.
dehydrogenase	TK1557	1.46	1.27	0.86
metallo-beta-lactamase superfamily hydrolase	TK1605	SKS001	SKS002	0.97
putative RNA-associated protein	TK1636	SKS001	SKS002	1.54
ATPase	TK1789	0.40	0.57	1.43
HAD superfamily hydrolase	TK1811	SKS001	SKS002	1.21
MoxR-like AAA family ATPase	TK1846	0.04	TS559	SKS001
MMR1/HSR1 family GTPase	TK1857	n.d.	SKS002	SKS002
carboxymuconolactone decarboxylase-related protein	TK1974	SKS001	n.d.	SKS001
ATPase	TK2042	1.65	1.19	0.72
ATPase	TK2098	1.79	1.35	0.75
RNA-binding protein FAU-1	TK2227	2.34	1.87	0.80
metallo-beta-lactamase superfamily hydrolase	TK2232	1.42	1.01	0.71
GHMP kinase	TK2242	SKS001	n.d.	SKS001
putative RNA-binding protein	TK2275	n.d.	SKS002	SKS002
Hypothetical proteins				
hypothetical protein	TK0009	1.28	1.31	1.02
hypothetical protein	TK0022	SKS001	n.d.	SKS001
hypothetical protein	TK0023	SKS001	SKS002	0.79
hypothetical protein	TK0024	0.72	0.61	0.84
hypothetical protein	TK0025	0.97	0.72	0.74
hypothetical protein	TK0033	0.56	0.63	1.11
hypothetical protein	TK0050	TS559	TS559	n.d.
hypothetical protein	TK0062	TS559	TS559	n.d.

hypothetical protein	TK0124	0.64	TS559	SKS001
hypothetical protein	TK0129	SKS001	n.d.	SKS001
hypothetical protein	TK0130	1.15	0.73	0.63
hypothetical protein	TK0133	0.48	0.57	1.19
hypothetical protein	TK0138	n.d.	SKS002	SKS002
hypothetical protein	TK0144	0.90	1.18	1.32
hypothetical protein	TK0145	0.55	0.65	1.18
hypothetical protein	TK0173	SKS001	n.d.	SKS001
hypothetical protein	TK0195	TS559	0.48	SKS002
hypothetical protein	TK0205	SKS001	n.d.	SKS001
hypothetical protein	TK0221	SKS001	n.d.	SKS001
hypothetical protein	TK0226	0.17	0.45	2.69
hypothetical protein	TK0346	n.d.	SKS002	SKS002
hypothetical protein	TK0350	0.52	0.94	1.81
hypothetical protein	TK0399	1.04	0.83	0.80
hypothetical protein	TK0400	SKS001	n.d.	SKS001
hypothetical protein	TK0404	0.44	0.60	1.36
hypothetical protein	TK0438	0.77	0.52	0.67
hypothetical protein	TK0446	TS559	1.13	SKS002
hypothetical protein	TK0453	SKS001	SKS002	1.17
hypothetical protein	TK0467	0.77	0.94	1.21
hypothetical protein	TK0478	SKS001	SKS002	0.79
hypothetical protein	TK0493	1.77	1.43	0.81
hypothetical protein	TK0558	SKS001	n.d.	SKS001
hypothetical protein	TK0569	2.18	1.94	0.89
hypothetical protein	TK0592	SKS001	SKS002	1.32
hypothetical protein	TK0628	0.62	0.83	1.34

hypothetical protein	TK0640	0.52	0.66	1.28
hypothetical protein	TK0648	1.89	0.73	0.39
hypothetical protein	TK0662	SKS001	SKS002	1.01
hypothetical protein	TK0668	SKS001	n.d.	SKS001
hypothetical protein	TK0710	1.39	0.57	0.41
hypothetical protein	TK0743	TS559	TS559	n.d.
hypothetical protein	TK0756	SKS001	SKS002	1.29
hypothetical protein	TK0778	1.27	1.51	1.19
hypothetical protein	TK0790	1.36	1.68	1.24
hypothetical protein	TK0826	1.16	1.98	1.71
hypothetical protein	TK0855	0.85	0.60	0.71
hypothetical protein	TK0894	SKS001	SKS002	0.88
hypothetical protein	TK0896	1.18	TS559	SKS001
hypothetical protein	TK0923	TS559	2.47	SKS002
hypothetical protein	TK0930	TS559	TS559	n.d.
hypothetical protein	TK0946	SKS001	SKS002	0.74
hypothetical protein	TK0947	1.63	1.66	1.02
hypothetical protein	TK0959	0.70	0.58	0.82
hypothetical protein	TK0971	0.90	1.06	1.18
hypothetical protein	TK0985	1.55	TS559	SKS001
hypothetical protein	TK1003	SKS001	SKS002	0.69
hypothetical protein	TK1018	SKS001	SKS002	1.06
hypothetical protein	TK1023	TS559	TS559	n.d.
hypothetical protein	TK1030	1.08	TS559	SKS001
hypothetical protein	TK1040	n.d.	SKS002	SKS002
hypothetical protein	TK1042	1.17	2.98	2.54
hypothetical protein	TK1046	1.52	1.92	1.26

hypothetical protein	TK1050	0.68	1.14	1.68
hypothetical protein	TK1051	1.64	1.73	1.05
hypothetical protein	TK1060	SKS001	SKS002	1.94
hypothetical protein	TK1097	SKS001	SKS002	1.61
hypothetical protein	TK1102	1.17	1.24	1.06
hypothetical protein	TK1115	0.82	1.15	1.40
hypothetical protein	TK1136	1.22	0.83	0.68
hypothetical protein	TK1142	TS559	TS559	n.d.
hypothetical protein	TK1148	n.d.	SKS002	SKS002
hypothetical protein	TK1163	SKS001	n.d.	SKS001
hypothetical protein	TK1168	n.d.	SKS002	SKS002
hypothetical protein	TK1186	1.55	1.33	0.86
hypothetical protein	TK1187	0.76	0.70	0.93
hypothetical protein	TK1192	TS559	TS559	n.d.
hypothetical protein	TK1241	SKS001	SKS002	1.16
hypothetical protein	TK1245	3.09	2.27	0.73
hypothetical protein	TK1249	SKS001	SKS002	0.65
hypothetical protein	TK1268	SKS001	SKS002	0.78
hypothetical protein	TK1278	n.d.	SKS002	SKS002
hypothetical protein	TK1312	2.32	1.30	0.56
hypothetical protein	TK1313	1.20	1.40	1.17
hypothetical protein	TK1329	SKS001	n.d.	SKS001
hypothetical protein	TK1334	SKS001	n.d.	SKS001
hypothetical protein	TK1335	SKS001	SKS002	0.83
hypothetical protein	TK1362	TS559	0.77	SKS002
hypothetical protein	TK1373	TS559	TS559	n.d.
hypothetical protein	TK1374	0.81	0.63	0.78

hypothetical protein	TK1384	TS559	TS559	n.d.
hypothetical protein	TK1389	0.84	1.07	1.27
hypothetical protein	TK1427	SKS001	n.d.	SKS001
hypothetical protein	TK1451	SKS001	n.d.	SKS001
hypothetical protein	TK1459	SKS001	SKS002	1.25
hypothetical protein	TK1475	SKS001	SKS002	0.78
hypothetical protein	TK1483	0.61	0.54	0.90
hypothetical protein	TK1493	TS559	1.19	SKS002
hypothetical protein	TK1497	n.d.	SKS002	SKS002
hypothetical protein	TK1580	SKS001	n.d.	SKS001
hypothetical protein	TK1607	SKS001	n.d.	SKS001
hypothetical protein	TK1608	0.92	0.99	1.07
hypothetical protein	TK1609	n.d.	SKS002	SKS002
hypothetical protein	TK1619	n.d.	SKS002	SKS002
hypothetical protein	TK1626	1.74	1.88	1.08
hypothetical protein	TK1630	SKS001	n.d.	SKS001
hypothetical protein	TK1631	0.96	1.05	1.10
hypothetical protein	TK1641	0.91	0.79	0.87
hypothetical protein	TK1642	TS559	TS559	n.d.
hypothetical protein	TK1657	TS559	TS559	n.d.
hypothetical protein	TK1661	TS559	0.81	SKS002
hypothetical protein	TK1664	SKS001	SKS002	1.50
hypothetical protein	TK1668	0.82	1.13	1.38
hypothetical protein	TK1670	2.16	1.64	0.76
hypothetical protein	TK1686	SKS001	SKS002	0.73
hypothetical protein	TK1724	SKS001	SKS002	1.15
hypothetical protein	TK1753	n.d.	SKS002	SKS002

hypothetical protein	TK1786	1.92	1.39	0.73
hypothetical protein	TK1834	0.59	0.56	0.96
hypothetical protein	TK1849	SKS001	n.d.	SKS001
hypothetical protein	TK1851	1.54	1.18	0.76
hypothetical protein	TK1852	0.87	0.79	0.91
hypothetical protein	TK1854	SKS001	n.d.	SKS001
hypothetical protein	TK1855	1.70	1.29	0.75
hypothetical protein	TK1897	0.91	1.39	1.53
hypothetical protein	TK1953	0.65	0.70	1.09
hypothetical protein	TK1967	0.96	0.81	0.85
hypothetical protein	TK1972	0.72	0.81	1.13
hypothetical protein	TK1975	0.73	TS559	SKS001
hypothetical protein	TK1991	n.d.	SKS002	SKS002
hypothetical protein	TK2025	1.59	1.09	0.69
hypothetical protein	TK2046	SKS001	n.d.	SKS001
hypothetical protein	TK2048	SKS001	n.d.	SKS001
hypothetical protein	TK2060	0.11	TS559	SKS001
hypothetical protein	TK2102	0.61	TS559	SKS001
hypothetical protein	TK2108	n.d.	SKS002	SKS002
hypothetical protein	TK2116	SKS001	n.d.	SKS001
hypothetical protein	TK2120	1.01	0.91	0.90
hypothetical protein	TK2145	1.43	1.15	0.80
hypothetical protein	TK2157	1.30	1.83	1.41
hypothetical protein	TK2171	SKS001	n.d.	SKS001
hypothetical protein	TK2200	0.17	TS559	SKS001
hypothetical protein	TK2201	TS559	TS559	n.d.
hypothetical protein	TK2224	TS559	TS559	n.d.

hypothetical protein	TK2226	TS559	TS559	n.d.
hypothetical protein	TK2237	SKS001	SKS002	0.91
hypothetical protein	TK2256	TS559	TS559	n.d.
hypothetical protein	TK2297	1.07	0.79	0.73

* = The expression of *TM0438* (controlled by the *gdh* promoter) was considerably lower than that of the other heterologously expressed genes and was in strain SKS002 only detected below the cutoff value. Nonetheless, the results of the in-gel activity assay already confirmed its functional expression (**Figure S3.2**). ** = in the database downloaded from NCBI, the membrane-bound ATP synthase was annotated as a V-type ATP synthase, which should be an A_0A_1 -type [59, 502].

Table S3.2 sequences of oligonucleotide primers used in this study

Oligonucleotide	Sequence (5'-3')	Application
BG3884	GACACGACCTCCGGAGATTTTCAGTTGAAGGTTT	Creation pSKS019: inverse PCR on pUDhps-phi (R)
BG3885	TGAGTGAGGGTCGCGATGTTCTCCTTTACCCATATT	Creation pSKS019: inverse PCR on pUDhps-phi (F)
BG3952	TGAAGAGCCACATAGGCC	Diagnostic PCR: binds <i>TM0438</i> sequence (F)
BG4126	GCATATGAGCTGGACAACCC	Diagnostic PCR: binds <i>TK0149</i> sequence (F)
BG4129	GTATCCAGATCCTTAAGCC	Diagnostic PCR: binds <i>TK0477</i> sequence (R)
BG4131	CTCAAGTTCTTCAATGTCCTCTGG	Diagnostic PCR: binds <i>TK0473</i> sequence (F)
BG4133	ACCTCTTCATGAGGAG	Diagnostic PCR: binds <i>TM1155</i> sequence (R)
BG4654	ACGCCAACCCGGAAGTTCCTC	RT-PCR: <i>TM1155</i> (F)
BG4655	GCCGAGGTAGTGGTCTATCC	RT-PCR: <i>TM1155</i> (R)
BG4656	CGCCAGGAAGGGCTACAAGG	RT-PCR: <i>TM0438</i> (F)
BG4657	CTAACGCCCATGCCGAGG	RT-PCR: <i>TM0438</i> (R)
BG4658	CCCACTGGCTTGAAGTAGGC	Diagnostic PCR: binds <i>TK0475</i> sequence (F); RT-PCR: <i>TK0475</i> (F)
BG4659	TGTGGACGAGGATGTAGTGG	Diagnostic PCR: binds <i>TK0475</i> sequence (R); RT-PCR: <i>TK0475</i> (R)
BG4662	GCTTGGGTCATGGACAGGCTC	RT-PCR: <i>TK0308</i> (F)
BG4663	TCTTAACCTGCTCGGCGACC	RT-PCR: <i>TK0308</i> (R)
BG4718	GCGCGATCCCTTCCCGAGCACTTCTTCGC	Creation pSKS022: inverse PCR on pSecPyrF (F)
BG4720	CCGGAATTCCGGACAACCCCTTGAGCTTTG	Creation pSKS022: inverse PCR on pSecPyrF (R)
BG4869	CCCGACCCTCATAAGCAAAG	RT-PCR: <i>TM0295</i> (F)
BG4871	CATCCTCACAGGGACGTTTG	Diagnostic PCR: binds <i>TK1763</i> sequence (F)
BG4872	TGAACTCGGCACCCTCTTTG	Diagnostic PCR: binds <i>TM0295</i> sequence (R)
BG4873	ACAGACTTGAGGTGGAAGG	Diagnostic PCR: binds <i>TK1766</i> sequence (R)
BG4874	ACCTTCAGGGCACTCAACAC	Diagnostic PCR: binds <i>TM1154</i> sequence (F)
BG4982	GTTGCTGAGGTCGTCCATCC	RT-PCR: <i>TM0295</i> (R)
BG4983	CCCGAGGACTACATAGACAG	RT-PCR: <i>TM1718</i> (F)
BG4984	CTCCGCCATCAACCATATC	RT-PCR: <i>TM1718</i> (R)
BG4985	AGCGACCAGAGCAACTTTAG	RT-PCR: <i>TM1154</i> (F)
BG4986	TTGAGTGCCCTGAAGGTGAG	RT-PCR: <i>TM1154</i> (R)
BG5059	AAGCTACCTCGACCAAGTATC	Diagnostic PCR: binds <i>TK1765</i> sequence (F)
BG5060	GTCCAGTAGTCCATAACTC	Diagnostic PCR: binds <i>TK1765</i> sequence (R)

Chapter 4

The chromosome copy number of the hyperthermophilic archaeon *Thermococcus* *kodakarensis* KOD1

This chapter has been published as:

Spaans, S.K., J. van der Oost, S.W.M. Kengen. *The chromosome copy number of the hyperthermophilic archaeon Thermococcus kodakarensis*. *Extremophiles*, 2015. **19**(4): 741-750

Abstract

The euryarchaeon *Thermococcus kodakarensis* is a well characterized anaerobic hyperthermophilic heterotroph and due to the availability of genetic engineering systems it has become one of the model organisms for studying Archaea. Despite this prominent role among the Euryarchaeota, no data about the ploidy level of this species is available. While polyploidy has been shown to exist in various Euryarchaeota, especially *Halobacteria*, the chromosome copy number of species belonging to one of the major orders within that phylum, i.e. the Thermococcales (including *Thermococcus* spp. and *Pyrococcus* spp.), has never been determined. This prompted us to investigate the chromosome copy number of *T. kodakarensis*. In this study, we demonstrate that *T. kodakarensis* is polyploid with a chromosome copy number that varies between 7 to 19 copies, depending on the growth phase. An apparent correlation between the presence of histones and polyploidy in Archaea is observed.

Introduction

Thermococcus kodakarensis is a marine hyperthermophilic, heterotrophic, obligate anaerobic euryarchaeon that can grow at temperatures ranging from 60°C to 100°C, with an optimum at 85°C. *T. kodakarensis* can grow on a variety of organic substrates in the presence of elemental sulfur, producing hydrogen sulfide, or when pyruvate or starch are present, in the absence of elemental sulfur producing hydrogen. Under optimal conditions, *T. kodakarensis* grows rapidly to high cell densities, having an estimated doubling time of ~40 minutes. Moreover, it can be grown on solid media, forming defined but small colonies [57, 58]. With its genome being fully sequenced [59], its natural competence and the establishment of various genetic engineering techniques [60-63], *T. kodakarensis* rapidly became one of the model organisms for studying anaerobic Archaea.

While genome modifications are generally relatively easily obtained in this organism, and thus routinely applied (e.g. [62, 71, 81-83, 104-106]), occasionally, we experienced difficulties in obtaining transformants with two homologous recombinations. Without going into too much detail, those data suggested that the desired double crossover modification co-existed with single crossover insertions, as well as with the wild type chromosome. The difficulty in obtaining homozygous recombinant cultures by repetitive cycles of re-streaking and screening, as well as the repeated observation of this phenomenon among different transformation attempts, suggested that *T. kodakarensis* may possess multiple copies of its chromosome.

In contrast to the general idea of prokaryotes being monoploid, polyploidy has in recent years been demonstrated for various species including many Euryarchaeota (**Table 4.1**), the phylum that *T. kodakarensis* belongs to. Although monoploid and polyploid species can co-exist within one phylum [503], none of the Euryarchaeota investigated so far was monoploid. It has therefore been suggested that polyploidy might be a common feature in Euryarchaeota [504]. In this study, we determined the chromosome copy number of the euryarchaeon *T. kodakarensis*. The copy number was determined for three different growth-phases; the early exponential phase, the late exponential / linear phase and the early

stationary phase. Our results confirm that *T. kodakarensis* is truly polyploid and that the exact ploidy level is dependent on the growth phase. In addition, a potential correlation between the presence of histones and polyploidy in Archaea is suggested.

Results

Growth analysis and preliminary quantification by “agarose gel method”

To make growth curves, five independent cultures of *T. kodakarensis* were grown on pyruvate, and their average OD₆₀₀ values and their variances were used to construct **Figure 4.1**. The results show that growth is highly reproducible, which is a prerequisite for reliable chromosome copy number quantification methods. Before an exact chromosome copy number quantification was performed, however, a preliminary in-gel quantification was applied first to determine the probability of polyploidy in *T. kodakarensis*, and hence if a more exact quantification method was worth applying. This was done by preparing cell lysate samples, in duplicate, of three independent cultures having an OD₆₀₀ of ~0.76. Corresponding cell densities were determined and aliquots of the 6 samples obtained were loaded on an agarose gel flanked by a known amount of an 1 kb GeneRuler DNA ladder (Thermo Scientific) as shown in **Figure 4.2A**. The DNA concentration and chromosome copy number of each sample was quantified by measuring the fluorescence intensity of the corresponding gel bands and comparing it to the ladder containing known concentrations of DNA. Together with the original cell density of each sample, the number of chromosome copies per cell was calculated (supplementary file). The results are shown in **Figure 4.2B**. These results clearly indicate polyploidy in *T. kodakarensis*, suggesting an average chromosome copy number of 17.8 ± 3.8 . It was therefore decided to determine the chromosome copy number by a more exact quantification method as well.

Table 4.1 Chromosome copy number of archaeal species and the distribution of archaeal histones

PHYLUM Order <i>Species</i>	Doubling time (h)	Genome copy number		Ploidy	Histones	Reference
		Exp. phase	Stat. phase			
EURYARCHAEOTA						
Halobacteriales						
<i>Halobacterium cutirubrum</i>	6.7	10.6		Polyploid	yes ^a	[505]
<i>Halobacterium cutirubrum</i>	13.3	6.3		Polyploid	yes ^a	[505]
<i>Halobacterium salinarum</i>	4	25	1	Polyploid	yes	[506]
<i>Halobacterium salinarum</i>	8	25	15	Polyploid	yes	[506]
<i>Haloferax volcanii</i>	4	17	10	Polyploid	yes	[506]
<i>Haloferax mediterranei</i>		30-55	20-30	Polyploid	yes	[507]
Methanosarcinales						
<i>Methanosarcina acetivorans</i>	6	18	16	Polyploid	yes	[504]
<i>Methanosarcina acetivorans</i>	49	3	5	Polyploid	yes	[504]
Methanobacteriales						
<i>Methanothermobacter thermautotrophicus</i>		2	1-2	Diploid ^b	yes	[508]
Methanomicrobiales						
<i>Methanoculleus marisnigri</i>					yes	
Thermococcales						
<i>Thermococcus kodakarensis</i>	1.25	19.4	7.5	Polyploid	yes	This study
<i>Pyrococcus furiosus</i>		multiple			yes	[509]
<i>Pyrococcus abyssi</i>		multiple			yes	[510]
Methanococcales						
<i>Methanocaldococcus jannaschii</i>	0.5	10-15	1-5	Polyploid	yes	[509]
<i>Methanococcus maripaludis</i>	2	55	30	Polyploid	yes	[504]
Thermoplasmatales						
<i>Thermoplasma acidophilum</i>					no ^c	
<i>Thermoplasma volcanium</i>					no ^c	
<i>Ferroplasma acidarmanus</i>					no ^c	
Archaeoglobales						
<i>Archaeoglobus fulgidus</i>		multiple			yes	[509]
Methanopyrales						
<i>Methanopyrus kandleri</i>					yes	
Methanocellales						
<i>Methanocella paludicola</i>					Yes	
CRENARCHAEOTA						
Acidilobales						
<i>Acidianus hospitalis</i>		1-2	2	Monoploid	no	[511]
Desulfurococcales						
<i>Aeropyrum pernix</i>	3.3	1-2	2	Monoploid	no	[511]
Fervidicoccales						
<i>Fervidicoccus fontis</i>					no	
Sulfolobales						
<i>Sulfolobus acidocaldarius</i>	3.5	1-2	2	Monoploid	no	[512]
<i>Sulfolobus tokodaii</i>	8	1-2	2	Monoploid	no	[511]
<i>Sulfolobus solfataricus</i>	7	1-2	2	Monoploid	no	[512]
Thermoproteales						
<i>Pyrobaculum aerophilum</i>		1-2	2	Monoploid	no	[511]

<i>Pyrobaculum calidifontis</i>	3.4	1-2	2	Monoploid	no	[511]
<i>Thermofilum pendens</i>					yes	
<i>Caldivirga maquilingensis</i>					yes	
<i>Vulcanisaeta distributa</i>					Yes	
THAUMARCHAEOTA						
Cenarchaeales						
<i>Cenarchaeum symbiosum</i>					yes	
Nitrosopumilales						
<i>Nitrosopumilus maritimus</i>					yes	
Nitrososphaerales						
<i>Nitrososphaera viennensis</i>					Yes	
KORARCHAEOTA						
<i>Candidatus Korarchaeum cryptofilum</i>					Yes	
NANOARCHAEOTA						
<i>Nanoarchaeum equitans</i>					yes	

Table 1. Chromosome copy number of archaeal species and the distribution of archaeal histones The table gives an overview of all studies in which chromosome copy numbers of archaeal species have been determined. The table was adapted from a table presented by Hillenbrand et al. [504], by including chromosome copy numbers of several additional species, grouping the results per order and including an extra column showing the presence or absence of archaeal histone homologs. The organisms that were not present in the histone database were additionally checked by BLAST for histone homologs. For the orders for which no ploidy level data is available, representative species, whose genome is available and could be screened for the presence of archaeal histone homologs, were randomly picked.

^aThe species *Halobacterium salinarum*, *Halobacterium halobium*, and *Halobacterium cutirubrum* are so similar that they should be regarded as strains of one species named *Halobacterium salinarum*. *H. cutirubrum* is therefore, as such, not present in the histone database [513].

^bThe cells grow in filaments; the numbers given are thus genome copies per cell, not per filament

^cThe Thermoplasmatales do not encode archaeal histones, however, they do encode homologs of HU-proteins (bacterial histone-like proteins)

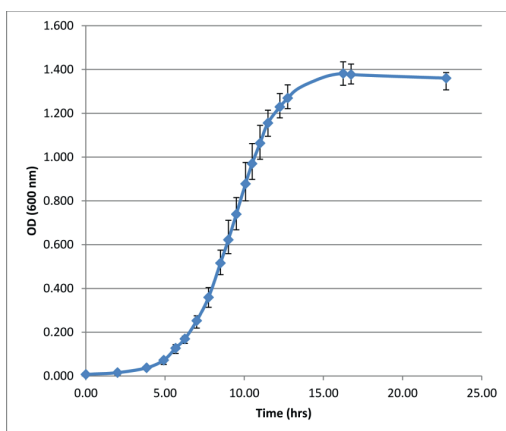


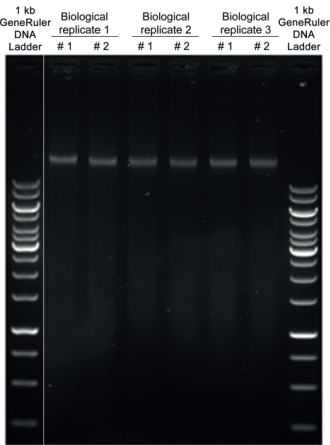
Figure 4.1. Growth curve. Increase in optical density of *Thermococcus kodakarensis* KOD1 (WT) in time on ASW-YT-Pyr medium. The graph represents the average optical density at 600 nm (OD₆₀₀) of five independent cultures

Quantification by “real-time PCR method”

To certify polyploidy in *T. kodakarensis* and to quantify chromosome copy numbers more accurately, a previously developed real-time PCR method for quantification of chromosome copy numbers was optimized for *T. kodakarensis*. This chromosome quantification method has been established for haloarchaea in 2006 [506], and has since then been applied to various other prokaryotes. The method has been validated against several independent techniques and has shown to be very reliable [503, 504, 514]. A schematic overview of the method can be found in the previous report [504].

To give a short overview, a fragment of about 1 kb is amplified first, using standard PCR with genomic DNA as template. A dilution series of this fragment is made and used to generate a standard curve in a real-time PCR analysis. This is done by amplifying a fragment of about 300 bp, internal to the standard fragment, and by plotting the threshold value (C_q) against the corresponding template concentration (in copies / μ l). To quantify the chromosome copy number of the species of interest, cells are lysed and a dilution series of the resulting cell lysates is analysed by real-time PCR in parallel to the standards. By comparing the C_q values of the cell lysate samples to the standard curve, the number of chromosome copies per sample can be determined, which in combination with the cell density of the corresponding culture, allows for calculating the ploidy level.

A.



B.

Biological replicate	OD ₆₀₀	Copy number based on left DNA ladder \pm SD	Copy number based on right DNA ladder \pm SD	Mean chromosome copy number \pm SD
1	~0.76	22.8	16.1	
2	~0.76	18.3	13.0	
3	~0.76	21.4	15.2	
Mean	0.76	20.8 \pm 2.3	14.8 \pm 1.6	17.8 \pm 3.8

Figure 4.2. Chromosome copy number quantification by “agarose gel method” (A) Agarose gel containing cell lysate samples of three independent cultures, in duplicate, flanked by a 1 kb GeneRuler DNA ladder. The white line indicates non-adjacent lanes of the same gel. The DNA concentration and chromosome copy number of each sample was quantified by comparing the fluorescence intensity of the gel bands to that of the left and right ladder. (B) The number of chromosome copies per cell. This was calculated by dividing the chromosome copy number of the lysate by the original cell density of each replicate

Cultures for chromosome copy number determination were grown to the optical densities as shown in **Figure 4.3**, representing the early exponential phase, late exponential / linear phase and early stationary phase. For each growth phase, samples were taken from three independent cultures. When the cultures reached the desired optical density, 1.5-ml samples were taken and cells were harvested by centrifugation and the supernatant was checked microscopically to verify that it was free of cells; which was the case. The cell numbers of the cultures at the moment of sampling were determined (see **Figure 4.3C**) and cells were lysed by adding SDS, which was found to be the most effective (see material and methods).

A tenfold dilution series of the cell lysates was prepared and used for real-time PCR analysis in parallel with the standard fragment. To verify that the added SDS or other components present in the cell lysate did not negatively influence the reaction itself, the standard dilution series was also added to 4 selected dilutions of the cell lysate sample as an internal control. No inhibiting effects were observed for any of the selected samples. The final real-time PCR was performed by including technical duplicates, in addition to the biological triplicates, and had an efficiency of 90.3%, an R^2 of 0.999 and a slope of -3.5. The results of the real-time PCR are shown in **Figure 4.3**. It was found that cells in the early exponential phase contain about 19.4 ± 4.9 copies of their chromosome, and that this value drops to 7.5 ± 2.8 copies in early stationary phase. The two independent quantification methods thus both confirm polyploidy in *T. kodakarensis*.

Discussion

The results obtained in this study show that the euryarchaeon *T. kodakarensis* contains multiple chromosomes, with numbers varying from ~7 to 19 copies per cell. Although polyploidy has been demonstrated for several Euryarchaeota, the chromosome copy number of species belonging to one of the major orders within that phylum, i.e. the Thermococcales (including *Thermococcus* spp. and *Pyrococcus* spp.), has never been determined. The existence of polyploidy in *T. kodakarensis* supports the suggestion of Hildenbrand, et al. (2011) that polyploidy might be a common trait of all Euryarchaeota [504]. The only exception known so far is *Methanothermobacter thermautotrophicus*, which was reported to be diploid instead [508]. *T. kodakarensis* was found to have a growth-phase-dependent ploidy level having as many as 19.4 ± 4.9 chromosome copies at the early exponential phase, which is strongly decreasing towards the stationary phase. The observation of a growth-phase-dependent ploidy level has been reported for many other prokaryotes as well, including the majority of the investigated euryarchaeal species [503, 504, 506, 514]. Unfortunately, no common pattern has emerged from the characterized prokaryotic species [506]. Where some species have a higher chromosomal copy number in

the exponential phase (e.g. *T. kodakarensis* or some halophilic Archaea [506]), others have a higher chromosomal copy number in the stationary phase (e.g. *Azotobacter vinelandii* [515]).

Although a chromosome copy number of around 19 might seem to be rather high, this result is comparable to copy numbers (2-55) found in many other Euryarchaeota (see **Table 4.1**). The euryarchaea *Methanococcus maripaludis* and *Haloferax mediterranei* were even found to have up to 55 chromosome copies per cell, which is the highest copy number for Archaea described so far. While this is more than double the number found for *T. kodakarensis*, it is still low compared to the copy numbers reported for the cyanobacterium *Synechocystis* sp. PCC 6803 or the firmicute *Epulopiscium* spp., having up to 218 or even tens of thousands of copies, respectively [514, 516]. The bacterium *Epulopiscium* spp., however, is quite extraordinary in that it is one of the largest known bacteria (200-700 μm), having sufficient intracellular space to contain an exceptionally high number of chromosomes [517]. *Synechocystis* sp. PCC 6803, in contrast, is similar in size (2 μm) to *T. kodakarensis*, and one may wonder how it is able to accommodate up to 10 times more chromosomes [57, 518].

The finding that *T. kodakarensis* is polyploid is also in good agreement with results obtained in chromosome quantification studies of the genetically accessible extremophiles *D. radiodurans* and *T. thermophilus* [519, 520]. Both bacteria were reported to have similar engineering difficulties as the ones occasionally observed for *T. kodakarensis* and were subsequently found to have multiple copies (6-10 and 4-5, respectively) of their chromosome. The existence of multiple chromosome copies can readily explain the occasionally observed difficulty in obtaining homozygous recombinant cultures, as it entails that the desired strain is only obtained when the modification is incorporated in all chromosome copies

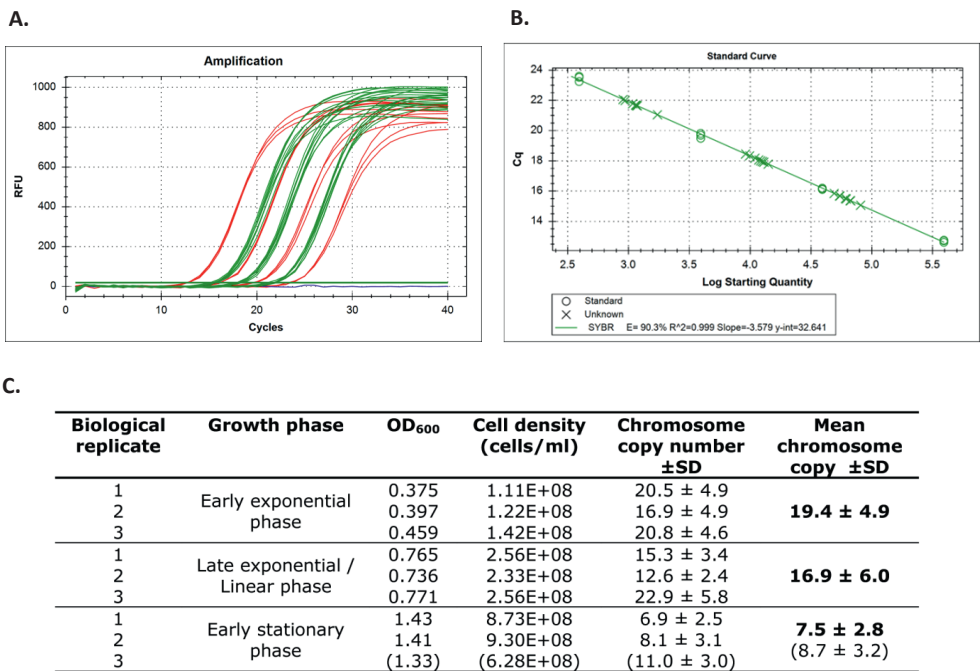


Figure 4.3. Chromosome copy number quantification by real time PCR method **a.** Selected real time PCR results. The fluorescent intensity curves from four standard dilutions (in red), three sample dilutions (in green), and a No Template Control (NTC, in blue) are shown. The reactions containing the standard dilutions were performed in triplicate and the ones containing the samples in duplicate (these technical duplicates were in addition to the biological triplicates). The curves shown are the result of serial tenfold dilutions of the templates. In addition to the selected reactions shown here, each experiment included more standard and more sample dilutions. **b.** Standard curve. The standard curve was created by plotting the threshold values (C_t) of the individual reactions against the corresponding template concentration. **c.** Average chromosome copy number values and their standard variations of *T. kodakarensis* in three different growth phases. The third sample of the early stationary phase (in parentheses) could be argued not to be in the early stationary phase. Therefore, the average chromosome copy number and its standard deviation in this phase was calculated by both including (in parentheses) and excluding the third biological replicate

Physical implications of polyploidy

Based on the results obtained, several questions arise: (1) how are polyploid prokaryotes able to fit that many copies of its chromosome in a cell, and related to that, how is spontaneous aggregation due to the high DNA concentrations prevented, and (2) what possible evolutionary advantages could polyploidy have? Since the latter aspect has been thoroughly covered in recent papers [521-523], describing potential advantages such as conferring resistance to DNA damage or allowing heterozygosity to occur (both of which

might be beneficial for survival in extreme conditions), only the former question will be briefly addressed here.

It is well known that all living things have to extensively condense their genomic DNA to fit it in the physically small space of a cell. In addition, this so-called DNA packaging also prevents spontaneous aggregation due to the high concentration imposed by the confinement [524]. The key mechanisms employed for DNA packaging are DNA supercoiling, macromolecular crowding and the association of DNA-binding proteins that fold the genome into a more condensed structure. The interactions between DNA-binding proteins and the genomic DNA can have different structural effects on DNA, such as bending, bridging or wrapping, each of which mediates DNA-packaging in a different way (reviewed in e.g. [525, 526]).

DNA-binding proteins are generally small basic proteins that are abundantly present. The best examples are the eukaryotic histones that wrap genomic DNA by folding it around their surface forming so-called nucleosomes, the primary unit of DNA compaction. Archaeal homologs of histone proteins have been found in almost all Euryarchaeota (including *T. kodakarensis*), Nanoarchaeota and Thaumarchaeota, but are generally not encoded by Crenarchaeota [527-532]. Interestingly, a similar dichotomy between Euryarchaeota and Crenarchaeota can be observed regarding their ploidy levels: almost all tested Euryarchaeal species have been found to be polyploid, while, all the Crenarchaeal species tested to date were found to be monoploid [504] (**Table 4.1**). Whether there is any correlation between these observations is currently unknown, however it is tempting to speculate that histones play an important role in enabling polyploidy in Archaea.

While DNA packaging explains how a single genome is sufficiently condensed to fit the physically small space of a cell, it does not clarify how many chromosome copies actually fit in. To answer this question we compared the situation in *T. kodakarensis* with the situation in *E. coli*, for which it is known that on average two chromosome copies (~ 9.2 Mbp in total, ~ 0.136 - $0.16 \mu\text{m}^3$ [533]) occupy about a quarter of the total cell volume (total cell volume of *E. coli*: 0.58 - $0.69 \mu\text{m}^3$ [534]) [535]. Assuming that DNA packaging in *T. kodakarensis* is as efficient as in *E. coli* and that the irregular cocci of *T. kodakarensis* are perfect spheres

(having a cell diameter of 1-2 μm and hence a cell volume of $\sim 0.52\text{-}4.18 \mu\text{m}^3$), fitting up to ~ 19 copies of its chromosome (~ 39.7 Mbp in total, $\sim 0.587\text{-}0.690 \mu\text{m}^3$), should physically, thus, be possible.

Conclusion

We clearly showed that *T. kodakarensis* has multiple chromosome copies and that this result supports earlier reports on the presence of polyploidy in Euryarchaeota. The existence of multiple chromosome copies can also readily explain the occasionally observed difficulty in obtaining homozygous recombinant cultures, as it entails that the desired strain is only acquired when the modification is incorporated in all chromosome copies. Moreover, an apparent correlation between the presence of histones and polyploidy in Archaea is observed. The histones may assist in the efficient DNA packaging which is required for fitting multiple chromosomes in a single cell.

Materials and Methods

Culture conditions and growth analysis

The hyperthermophilic archaeon *Thermococcus kodakarensis* KOD1 wild-type strain was grown anaerobically at 85°C in ASW-YT-Pyr medium. The ASW-YT medium was composed of 0.8x artificial seawater (0.8x ASW), 5.0 g/L yeast extract, 5.0 g/L tryptone and 0.8 mg/L resazurine. Before inoculation, 5.0 g/L sodium pyruvate (ASW-YT-Pyr medium) was added to the medium, as well as Na₂S·9H₂O, until it became colorless. For all cultivations, 120-ml serum bottles containing 40 ml culture with N₂ as headspace were used. Cultures were continuously shaken (120 rpm) during growth. Growth was monitored by measuring the optical density at 600 nm (OD₆₀₀) using an U-1500 spectrophotometer (Hitachi). Five independent cultures were grown, and their average OD₆₀₀ value and their variances were calculated (**Figure 4.1**).

Sample extraction, cell lysis and determination of cell number

Cultures for chromosome copy number determination were inoculated from fresh pre-cultures in the late exponential growth phase (1% inoculum) and grown to the desired optical density. For each growth phase, samples were taken from three independent cultures. When the cultures reached the desired optical density, 1.0 ml and 1.5 ml-samples were taken by syringe and cells were harvested by centrifugation (5000 g, 30 min, RT), for the “agarose gel method” and “real time PCR method” respectively. The supernatant was checked microscopically to verify that it was free of cells. Cell pellets were stored at -20°C until further processing. The cell number of the cultures at the moment of sampling was determined by directly counting the cells using a Neubauer counting chamber. The pellets were resuspended in 100 µl TE buffer (10 mM Tris; 0.1 mM EDTA, pH8.5). Cell lysis was performed by adding SDS to a final concentration of 0.3%. Complete cell lysis was verified microscopically and the integrity of genomic DNA was verified by agarose gel electrophoresis.

Preliminary quantification of chromosome copy number by “agarose gel method”

To determine if *T. kodakarensis* could indeed possess multiple chromosomes per cell, and hence if it would be worth pursuing quantification by a more accurate method, a preliminary quantification experiment was performed first. This was done by preparing cell lysates of three independent cultures having an OD₆₀₀ of ~0.76 (see above). To minimize the disturbing effect of SDS during agarose gel electrophoresis, the cell lysate was diluted 10-fold by adding TE buffer to a final volume of 1 ml. Aliquots of 5 µl of each sample were loaded, in duplicate, on a 1% agarose gel containing the SYBR Safe DNA gel stain (Life technologies), flanked by a known amount of a 1 kb GeneRuler DNA ladder (Thermo scientific). Since the SYBR Safe DNA gel stain moves away from the migrating DNA, the stain was also added to the running buffer to ensure an equal concentration of the stain in the entire gel. After electrophoresis, the DNA concentration of each sample was quantified by measuring the fluorescence intensity of the corresponding bands using a G:box imager and GeneSys analysis software (Syngene) and comparing it to the ladder with known concentrations of DNA. The molecular weight (in Daltons) of one chromosome copy was computed with the “DNA molecular weight calculator” (<http://www.currentprotocols.com/WileyCDA>), which was then converted to its mass (in grams) by using a “Dalton to grams conversion calculator” (<http://www.unitconversion.org>). Next, the chromosome copy number in the cell lysate was calculated by dividing the DNA concentration of the lysate by the mass of one chromosome. Finally, the number of chromosome copies per cell was calculated by combining this with the original cell density of the sample (an overview of these calculations is available online: DOI 10.1007/s00792-015-0750-5; supplementary file).

Quantification of chromosome copy number by “real-time PCR method”

To determine chromosome copy numbers, a previously developed real-time PCR approach was used [506]. A schematic overview of the method can be found in Hildenbrand, et al. [504]. First, to make a standard curve, a 1 kb fragment was amplified using conventional PCR with genomic DNA as template. The genomic DNA was isolated by using a genomic DNA purification kit (Thermo scientific, protocol for Gram positive bacteria). The 1 kb fragment generated in this study was internal to the chitinase gene (TK1765) as the amplification of this fragment had previously been shown to be successful, resulting in a clear, single

product-band on gel. The primers used are shown in **Table S4.1**. The fragment was purified by using preparative agarose gel electrophoresis (0.8%) and an GeneJET Gel extraction kit (Thermo Scientific), and concentrated by an DNA Clean and Concentrator kit (ZYMO RESEARCH). The DNA concentration was determined with a Nanodrop ND-1000, and the number of DNA molecules was calculated using the molecular weight computed with “oligo calc” (www.basic.northwestern.edu/biotools/oligocalc.html). A serial tenfold dilution containing defined numbers of standard molecules was generated in triplicate, and used for real-time PCR analysis in parallel with a dilution series of cell lysate samples.

To minimize the inhibiting effect of SDS on real-time PCR, the cell lysate (see above) was first diluted to a final SDS concentration of 0.03% by adding TE buffer to a final volume of 1 ml. This diluted cell lysate sample was used for making the dilution series that was used for real-time PCR. A 5- μ l aliquot of the dilution series was used as template in the real-time PCR analyses for quantification of chromosome copy numbers (see below). To ensure that the PCR efficiency of the cell lysate dilution series and the standard dilutions series was identical, the standard dilution series was also added to 4 selected dilutions of the cell lysate sample as an internal control, yielding 6 different dilution series in total. The fragment targeted in the real-time PCR analysis was 293 bp and was internal to the standard fragment.

Real-time PCR was performed in 25- μ l reactions containing 5 μ l of template (cell lysate, standard, or cell lysate with added standard), 100 mM of primer BG5331 and BG5332 each and 2x qPCR Master Mix (iQ SYBR Green Supermix, BIO RAD). The master mix contained antibody-mediated hot start iTaq DNA polymerase, dNTPs, SYBR GREEN I, $MgCl_2$, enhancers, stabilizers and fluorescein (concentrations not released by BIO RAD). The PCR reaction conditions were 10 min at 95°C, 40 cycles of amplification (30 sec at 95°C, 30 sec at 59°C, 30 sec at 72°C), and an final incubation of 5 min at 72°C. The real-time PCRs were performed in a CFX96 thermal cycler (BioRad). At the end of the PCR, a melting curve was generated using the following settings: 65°C to 95°C with increments of 0.5°C per 5 sec. For each sample the numbers of cycles was determined until its fluorescence intensity reached the threshold (Cq value). By comparison of the threshold cycle (Cq) differences of the different dilutions it was verified that the PCR was exponential at least up to the threshold DNA concentration used

for the analysis (i.e. a 10-fold dilution corresponds to a C_q difference of 3.6 > C_q > 3.1, having an efficiency of 90%-110%) The C_q values of the standards were used to construct a standard curve (**Figure 4.3**) that was used to determine the chromosome copy number in the cell lysates and to check the PCR efficiency in the cell lysates including internal standards. In combination with the cell density, the number of chromosome copies per cell was calculated.

Method Optimization

The protocol as outlined above was established by optimizing different steps individually. First, various genomic DNA isolation methods, such as sonication, French press, osmotic shock, CTAB/NaCl, lysozyme, SDS or repetitive freeze-thaw cycles were tested for their effectiveness (extent of cell lysis and gDNA recovery) and ability to leave the genomic DNA intact. The only method meeting both standards, and yielding reproducible results, was the addition of SDS to the cell suspensions. Since SDS is known to have disturbing effects on both PCR and agarose gel electrophoresis, a series of SDS concentrations was tested to determine the minimal concentration needed for complete lysis and yielding reproducible results. This was found to be 0.3%. chromosome copy number quantification by the “agarose gel method” was improved by addition of the SYBR Safe DNA gel stain to the running buffer to ensure an equal concentration of the stain in the entire gel. Real-time PCR was optimized by testing three different primer sets for the amplification of a small internal fragment (all ~300 bp). For each primer set, the optimal annealing temperature was determined by gradient PCR. The primer set BG5331-BG5332 (see **Table S4.1**), resulting in the amplification of a 293 bp fragment, turned out to be the most reproducible, without any artificial by-products. The latter was confirmed by visualizing the amplified fragments on a 1% agarose gel as well as by including a melting curve in the real-time PCR protocol.

Acknowledgements

We thank Jörg Soppa from the Goethe-University for valuable advice concerning the chromosome quantification by the “real-time PCR method” and Remus Dame from Leiden University for his valuable input on discussing the physical implications of polyploidy.

This work is part of the Open Programme (“Biological hydrogen production beyond current limits” file number 820.02.011), which is financed by the Netherlands Organisation for Scientific Research (NWO).

Table S4.1. Sequences and applications of oligonucleotides

Oligonucleotide	Sequence (5'-3')	Application
BG4126	GCATATGAGCTGGACAACCC	Diagnostic PCR analysis Figure S1
BG4869	CCCGACCCTCATAAGCAAAG	Diagnostic PCR analysis Figure S1
BG4870	CGGCGAGCTTCTCGTAAACC	Diagnostic PCR analysis Figure S1
BG4871	CATCCTCACAGGGACGTTTG	Diagnostic PCR analysis Figure S1
BG4872	TGAACTCGGCACCCTCTTTG	Diagnostic PCR analysis Figure S1
BG4873	ACAGACTTGAGGTGGAAAGG	Diagnostic PCR analysis Figure S1
BG4874	ACCTTCAGGGCACTCAACAC	Diagnostic PCR analysis Figure S1
BG5059	AAGCTACCTCGACCAGTATC	Diagnostic PCR analysis Figure S1 & Synthesis of real-time PCR standard; 1 kb fragment internal to TK1765
BG5060	GTCCCAGTAGTCCATAACTC	Diagnostic PCR analysis Figure S1 & Synthesis of real-time PCR standard; 1 kb fragment internal to TK1765
BG5061	TGCCGGGAAGCTAGAGTAAG	Diagnostic PCR analysis Figure S1
BG5062	CCCAGAAACGCTGGTGAAAG	Diagnostic PCR analysis Figure S1
BG5329	ACAAGGACTACCTCCTAACG	Detection of genome copies in <i>T. kodakarensis</i> , primer set 1 (300 bp fragment)
BG5330	CAAAGCTCCTGCTGTAGAAG	Detection of genome copies in <i>T. kodakarensis</i> , primer set 1 (300 bp fragment)
BG5331	ACTACGCCTTCTCTGACCTC	Detection of genome copies in <i>T. kodakarensis</i> , primer set 2 (293 bp fragment)
BG5332	GTAGTTGCCC GCCATACCTC	Detection of genome copies in <i>T. kodakarensis</i> , primer set 2 (293 bp fragment)
BG5333	AGAGTCCC GACGACGTAAG	Detection of genome copies in <i>T. kodakarensis</i> , primer set 3 (306 bp fragment)
BG5334	CGGTGTAGTTGACGCAGAAG	Detection of genome copies in <i>T. kodakarensis</i> , primer set 3 (306 bp fragment)

Chapter 5

Identification of *Thermococcus kodakarensis* genes involved in GDGT synthesis

Melvin F. Siliakus^{1†}, Sebastiaan K. Spaans^{1†}, Kira Makarova⁴, Rie Matsumi⁵, Michel Koenen³, Ellen C. Hopmans³, Laura Villanueva³, Antonella Caforio², Servé W.M. Kengen¹, Arnold J.M. Driessen², Jaap Sinninghe Damsté³, John van der Oost¹

¹Laboratory of Microbiology, Wageningen University, Dreijenplein 10, 6703 HB Wageningen, The Netherlands

²Department of Molecular Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, 9747 AG Groningen, The Netherlands

³Department of Marine Organic Biogeochemistry, Royal Netherlands Institute for Sea Research (NIOZ), PO Box 59, NL-1790 AB Den Burg, The Netherlands

⁴National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, 8600 Rockville Pike, Bethesda, MD, 20894, USA

⁵Department of Microbiology, The Ohio State University, 105 Biological Sciences Building, 484 West 12th Avenue, Columbus, OH 43210, USA

[†]Authors contributed equally

Abstract

In this study, we set out to identify the enzyme(s) involved in the synthesis of tetraether lipids in *Thermococcus kodakarensis*, more specifically the one(s) responsible for the condensation of two C₂₀ isoprenyl chains, i.e. the putative tetraether synthase(s). To do this, a list of genes that are potentially involved in tetraether lipid synthesis was composed, and prioritized based on a couple of characteristics assumed to be important for a tetraether synthase. This resulted in the selection of 12 genes, which were targeted for overexpression and disruption. Subsequent diether/tetraether lipid analysis of the mutant strains showed that at least three of these genes, i.e. *TK1548*, *TK1836* and *TK1957* play a role in controlling the amount of GDGTs in the archaeal membrane. Whereas *TK1548*, annotated as serine-pyruvate aminotransferase, has no obvious link with lipid biosynthesis, both *TK1836* and *TK1957* have putative roles as geranylgeranyl hydrogenase and polyprenyltransferase, respectively. Their role in GDGT synthesis is discussed. Disruption of various other candidate genes turned out to be non-viable, which is an interesting result as disruption of a tetraether synthase may very well be lethal. However, the data do not allow us to identify the key tetraether synthase, and more research is needed to get a better understanding of the tetraether synthesis pathway and to identify the tetraether synthase(s).

Introduction

Micro-organisms are able to colonize and survive an extraordinary range of environments. Most strikingly, some not only survive, but even thrive under extreme conditions such as very high temperatures (thermophiles), salinities (halophiles), or extreme pH levels (acidophilic and alkaliphiles). Though the majority of these so-called extremophiles belong to the archaea, they have also been found among the bacterial domain. Since their discovery, extremophiles have attracted much attention because of both their biotechnological potential and their remarkable physiological capabilities. Obviously, the main focus of how these microbes withstand the perturbations of cellular integrity has been on the plasma membrane, as the membrane shields the inner homeostatic conditions of the cell from the hostile environment. Before the discovery of extremophiles, the membrane composition and adaptation strategies of mesophilic bacteria were well described [536-539].

The cell membrane generally consists primarily of a phospholipid bilayer. In bacteria, the framework molecule of these phospholipids is phosphatidic acid. Phosphatidic acid is composed of two acyl-chains which are esterified to the backbone *sn*-glycerol-3-phosphate. Phosphatidic acid serves as a basis for a high variety of phospholipids, which are adjusted to the needs of the membrane. The main membrane parameters which are subject to change in response to varying physicochemical conditions are fluidity and permeability [540, 541]. In response to physicochemical stressors, bacteria regulate the saturation and length of the acyl chains to keep the membrane parameters within acceptable boundaries.

Archaea have a different lipid architecture and thus also differ in their homeoviscous adaptation strategies [108]. Instead of phosphatidic acid, archaea are generally thought to use archaetidic acid as the framework molecule of their phospholipids. Archaetidic acid, which is also known as (S)-2,3-di-*O*-geranylgeranyl-*sn*-1-glycerol phosphate (DGGGP), differs from phosphatidic acid in several key aspects. For example, the hydrophobic moiety is composed of highly branched C₂₀-isoprenoid chains, rather than the bacterial acyl chains. Moreover, these chains are bound by ether bonds (as opposed to the bacterial ester bonds) to *sn*-glycerol-1-phosphate, the stereochemical opposite of the bacterial *sn*-glycerol-3-

phosphate [542, 543]. These key differences are unique to both domains and have therefore been termed “the lipid divide” [544]. Hence, phosphatidic acid and archaetidic acid are regarded as marker molecules between prokaryotic lineages.

Archaetidic acid is generally considered to be the framework molecule of a broad range of archaeal membrane lipids [545]. Nevertheless, archaeal membranes typically consist of a variation of two main core structures: *sn*-2,3-diphytanyl glycerol diether, which in this study will simply be referred to as diether (and is also known as archaeol), and *sn*-2,3-dibiphytanyl diglycerol tetraether, which in this study will simply be referred to as tetraether (and is also known as glycerol-dialkyl-glycerol tetraether (GDGT), or caldarchaeol). The diether lipids are generated from archaetidic acid by reducing the unsaturated isoprenoid chains, forming C₂₀-phytanyl chains, and by subsequent addition of a polar headgroup to the glycerol moiety. Tetraether lipids are bipolar, consisting of saturated C₄₀-biphytanyl chains (a molecule that is composed of two condensed C₂₀-phytanyl units) and a polar head group at each end of that chain. Tetraether lipids form a monolayer membrane, whereas the diether lipids form a bilayer structure that closely resembles the traditional double leaflet bilayer of bacteria [110, 542, 543].

The variation observed in tetraether lipids involves the incorporation of one or more pentacycles, the incorporation of a single hexamer, the addition of methyl groups, or the incorporation of bridges between the isoprenoids [108, 546-548]. Tetraether lipids are believed to contribute considerably to an increased stability, permeability and resistance of the membrane [540, 549]. Moreover, though the diether:tetraether ratio appears to vary substantially between archaea at comparable conditions, for some hyperthermophilic archaea the amount of tetraethers has been shown to increase steadily as the growth temperature rises [546, 550, 551]. This suggests that the diether:tetraether ratio might be regulated to adapt membrane fluidity and permeability. Additionally, for some hyperthermophilic archaea it has been shown that in response to increasing temperatures, the number of incorporated pentacycles also increases [540, 551, 552]. However, as thermal adaptation has frequently been studied in archaea, it soon became clear that not all archaea

form tetraethers and that different archaeal species respond to temperature fluctuations in different ways.

In general, tetraethers have been shown to be the dominant lipid species in Crenarchaeota and Thaumarchaeota, whereas members of the Euryarchaeota synthesize diethers, tetraethers, or both [110]. However, the exact route leading to tetraether synthesis is still unknown. Nevertheless, two different pathways have been postulated (**Figure 5.1**) [110, 542, 543]. In the first pathway (to which we will refer to as the traditional route), archaetidic acid is first converted to diether lipids, after which two of these molecules are coupled via a yet unknown route by a tail-to-tail condensation of the isoprenoid chains to form one tetraether molecule [542, 543, 553-555]. Evidence for this traditional route comes from pulse chase and labelling experiments using cell extracts of the euryarchaeon *Thermoplasma acidophilum* incubated with ^{14}C -mevalonate, which is a precursor molecule for the synthesis of the isoprenoid building blocks IPP and DMAPP [542, 553]. These experiments showed that the label first incorporates into diethers until saturation and only then into tetraethers. Moreover, when an inhibitor of squalene epoxidase (known as terbinafine) is used, an inhibition of tetraether formation was observed. In addition, pulse labelling in the presence of terbinafine showed an accumulation of radioactivity in the diether fraction and a decrease in radioactivity in the tetraether fraction, which was reversible by removal of the inhibitor [553, 554]. These observations are interesting as, in the traditional route, the putative tetraether synthase is predicted to be an analog of squalene or phytoene epoxidase [553].

However, no incorporation of radiolabelled diether lipids (i.e. archaeol) into tetraether lipids (i.e. GDGT-0) was observed *in vivo* in the euryarchaeon *Methanospirillum hungatei* [556]. Moreover, from a chemical point of view, the reaction required for the formation of tetraethers from diethers (i.e. the condensation of two saturated phytanyl chains of archaeol) is considered quite unusual and some studies indeed suggest that the presence of double bonds is a prerequisite for the formation of tetraethers [110, 557, 558]. These observations are contradicting the ones described above and suggest that diether lipids (i.e. archaeol) are not the precursor of tetraether lipids (i.e. GDGT). This discrepancy might, for example, be explained by the possibility that the diether accumulation upon terbinafine

exposure is the result of growth retention by toxicity rather than the selective inhibition of an enzyme. Especially, as it has been demonstrated several times that the ratio of diether-to-tetraether is highly dependent on growth phase as well [546, 547, 559, 560].

Recently, an alternative route for the formation of tetraether lipids has been proposed in which the tail-to-tail condensation is thought to take place directly after the synthesis of various forms of isoprenoid chains (i.e. with or without ring moieties) [110]. Thus, the tail-to-tail condensation of two C_{20} isoprenyl molecules is suggested to occur before attachment to the glycerol-1-phosphate unit and the subsequent reduction of the unsaturated isoprenoid chains (**Figure 5.1**). The advantage of this is that the proposed pathway eliminates the need for the unusual condensation of the two saturated phytanyl chains present in diether lipids (i.e. archaeol). This alternative route implies that diether lipids are not the precursor of tetraether lipids and that the ether-bond forming enzymes, GGGP synthase and DGGGP synthase (responsible for the coupling of the isoprenyl chains to the glycerol unit), can utilize not one but various prenyl species as substrate. These substrates include both C_{20} and C_{40} isoprenoid chains (for the formation of diether lipids and tetraether lipids, respectively) with or without ring moieties. Thus, in this alternative pathway, tetraether lipids are formed by creating C_{40} isoprenoid chains first, which are subsequently used as a substrate for the GGGP synthase and DGGGP synthase to attach both the first and second glycerol units. However, like the traditional route, the enzyme(s) responsible for the condensation of the C_{20} isoprenyl chains are still unknown.

Despite the high interest in the biochemical mechanism, nothing is known about the enzymes involved in the tail-to-tail condensation. This might be due to the lack of archaeal genetic toolboxes. Here, the anaerobic, sulphur reducing, hyperthermophilic Euryarchaeon *Thermococcus kodakarensis* was used as a model organism that meets a couple of important criteria that enable such a study: i) *T. kodakarensis* has previously been confirmed to have both diether lipids and tetraether lipids in its cytoplasmic membrane [58] ii) *T. kodakarensis* has a genetic toolbox available [60-63] iii) It has a relatively small and fully sequenced genome of an estimated 2306 coding DNA sequences [59], many of which have been functionally characterized already, and iv) *T. kodakarensis* is fairly easily grown in a

laboratory setting. By performing comparative bioinformatics analyses we selected several candidate genes that might code for enzyme(s) involved in the tail-to-tail condensation reaction, i.e. the putative tetraether synthase. Moreover, we composed a list of characteristics likely to encounter in a putative tetraether synthase. By scoring the candidate genes according to these characteristics, we ultimately selected 12 candidate genes that were targeted for overexpression and deletion. The generated mutant strains were analysed for their diether/tetraether lipid composition.

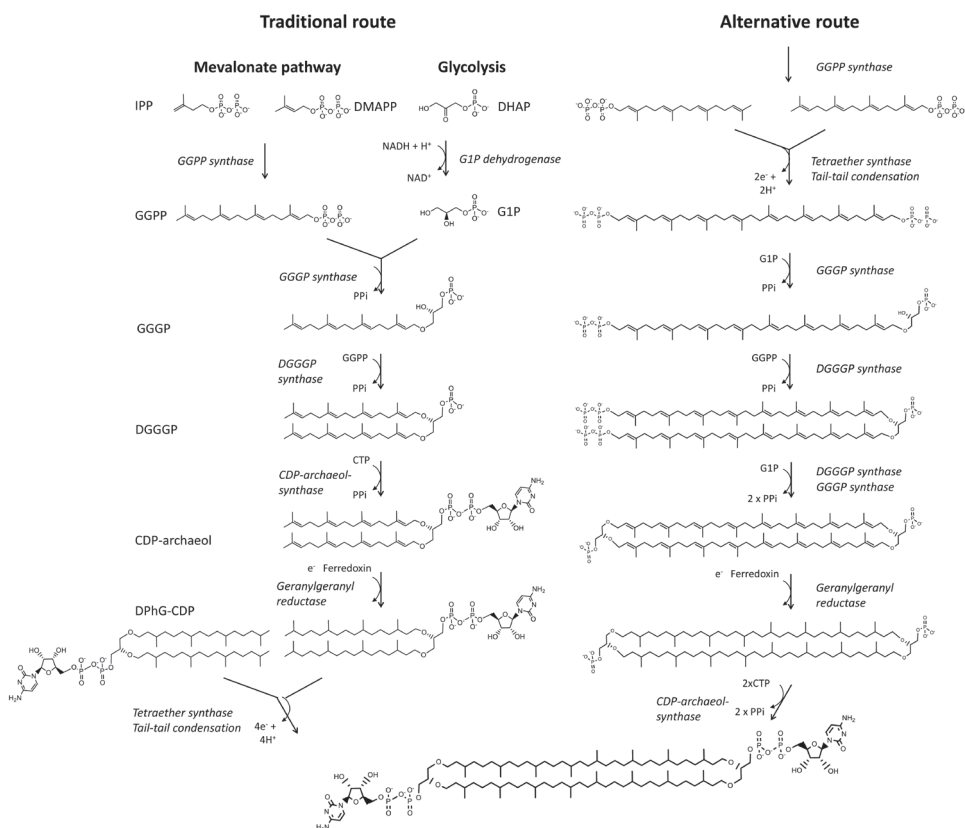


Figure 5.1. Two proposed pathways for tetraether synthesis. The first pathway is represented by the “**traditional route**”. This pathway proposes a substrate product relationship between archaeol and glycerol-dialkyl-glycerol-tetraethers (GDGT). Initial steps are completed by full synthesis of DGGGP. From this stage, DGGGP can either bifurcate towards full synthesis of saturated diethers or towards synthesis of GDGT’s by the successive condensation of the two tail ends of the aligned archaeols. This reaction is here represented by two fully saturated diphytanylglycerol-CDP (DPhG-CDP) and the accompanying release of two electrons and two protons. Although the saturated diether was the first proposed substrate (as depicted here), some studies suggest the unsaturated DGGGP as a more likely substrate. The second pathway that has been proposed is represented by the “**alternative route**”. In this pathway a C₄₀ hydrocarbon chain is synthesized first by condensation of two molecules of GGGP. The two subsequent reactions are considered similar to the formation of GGGP and DGGGP except for the chain length of the substrates. The GDGT is finalized by the attachment of the glycerol moiety to the tail ends of the C₄₀ chains by ether bond formations by either GGGP or DGGGP synthesis.

Results

Selection of candidate genes

To select gene(s) that are potentially involved in tetraether synthesis, two independent bioinformatics approaches were performed. Both analyses were resting on the assumption that the gene(s) involved and the presence of tetraether lipids exhibit a genotype-phenotype correlation. In addition to archaea producing both diether and tetraether lipids, there are also a number of archaea within the phylum of Euryarchaeota that completely lack tetraethers [108]. In the first analysis, we made use of these phenotypes by selecting a set of reference species that are phylogenetically closely related to *T. kodakarensis* and that have a confirmed presence or absence of tetraethers in their membranes. The selected reference genomes belonging to the first group (i.e. confirmed presence of tetraethers) were those of the hyperthermophiles *Pyrococcus furiosus* and *Thermoplasma acidophilum* [553, 561-565] and those of the second group (i.e. confirmed absence) were those of the haloalkaliphilic *Natronomonas pharaonis* and the hyperthermophilic *Methanopyrus kandleri* [566-568]. The genome of *T. kodakarensis* was screened for homologs present in the first two, but absent in the latter two, by performing a phylogenetic profiler single gene search (using the IMG system of JGI). This resulted in the identification of 57 *T. kodakarensis* genes that may potentially be involved in tetraether synthesis (**Table S5.1a**).

In the second comparative genome analysis we attempted to correlate the presence of specific arCOGs (archaeal Clusters of Orthologous Genes [569] in 149 sequenced archaeal genomes to species with a confirmed tetraether lipid membrane. The arCOGs included in the analysis (**Table S5.2**) were related to the following categories: phospholipid biosynthesis, isoprenoid mevalonate biosynthesis, uncharacterized proteins associated with mevalonate biosynthesis, terpene cyclases, fatty acid synthesis, and phytoene/squalene synthesis or saturases. However, none of the arCOGs analysed showed a strong correlation with species having a confirmed tetraether lipid membrane. Assuming correct membrane profiling, this analysis suggests that tetraether synthesis is either catalysed by different enzymes over different archaeal lineages or that a single enzyme is present in all archaea whose synthesis

or activity is tightly regulated (e.g. by post-translational modification, temperature activation, etc.).

Since no correlation was observed, it was decided to select all uncharacterized *T. kodakarensis* genes that are present in the mentioned arCOGs categories instead. This led to the selection of 32 additional genes. To exclude other members of the archaeal lipid biosynthetic pathway (e.g. IPP synthase, GGPP synthase, GGGP synthase, etc.), experimentally characterized genes, that have shown to be responsible for the reactions leading to isoprenoid and CDP-archaeol synthesis in various archaea, were BLASTed against the genome of *T. kodakarensis* [570-577]. *T. kodakarensis* genes that showed a high sequence similarity to these characterized genes were excluded from the list of 32 genes (**Figure S5.1**). However, the candidate genes *TK1324* and *TK1957* both showed sequence similarity to known DGGGP synthases, making it, based on this information only, impossible to decide if both or only one (and if the latter, which one) encode for a DGGGP synthase. Moreover, both are annotated as 4-hydroxybenzoate polyprenyltransferase belonging to the UbiA-like prenyltransferase COG0382, and it was therefore decided to retain and prioritize both of them. This ultimately led to the selection of 26 additional genes that may be involved in tetraether lipid biosynthesis (**Table S5.1b**).

Subsequently, it was decided to prioritize the resulting 81 genes based on a couple of characteristics assumed to be important for a putative tetraether synthase. These characteristics were (in a non-specific order): i) predicted prenyl binding. ii) the presence of an FAD or Flavin mononucleotide (FMN) binding domain, which is potentially important for tetraether synthases because carbon-carbon bond formations are occasionally aided by FAD cofactors as dual electron acceptor [578]. iii) the presence of aspartic acid rich DDxxD motifs. These motifs contribute to the substrate binding of prenyltransferases, which are responsible for the transfer of prenyl groups and chain elongation of isoprenoids [579-581]. iv) Homology to squalene epoxidase. This has been suggested by Kon et al. (2002), who, upon addition of a squalene epoxidase inhibitor (i.e. terbinafine) to the growth medium of *Thermoplasma acidophilum*, observed an accumulation of diether lipids at the expense of tetraether lipid biosynthesis. [553]. v) Transcriptional regulation by the *Thermococcales* heat

shock protein Phr. Phr is a transcriptional regulator that represses transcription under normal growth conditions by binding to the promoters of certain genes and that is released from them upon heat shock [582-584]. Since previous studies showed an increase in tetraether lipid content upon thermal treatment, transcription of the putative tetraether synthase may also be Phr-dependent [546, 550, 551]. Altogether, this scoring resulted in a ranking of the candidate genes (**Table S5.1**). In case the candidate genes had an equal score, priority was given to the presence of prenyl binding or -transferase domains. By doing this, we selected 10 priority genes (in addition to *TK1324* & *TK1957*) that were targeted for overexpression and disruption (**Table 5.1**). These genes were: *TK1088*, *TK0929*, *TK2159*, *TK1836*, *TK0509*, *TK0513*, *TK1173*, *TK1477*, *TK1548* & *TK1732*.

Moreover, an additional column was included to **Table 5.1** showing the presence or absence of transmembrane spanning helices. This information may become valuable as it can demonstrate the validity of either the classical or alternative synthesis pathway. This is because in the classical biosynthesis pathway, the tetraether lipid biosynthesis is proposed to occur by condensing two diether lipids, the synthesis of which already takes place in the membrane [542]. In contrast, however, in the alternative pathway, the condensation reaction leading to the C₄₀ isoprenoid chains takes place in the early steps of the ether lipid biosynthesis and is therefore more likely to be catalyzed by a cytosolic enzyme [110].

Table 5.1. Top-10 ranking candidate genes

Gene locus	Gene annotation	Predicted Prenyl binding	FAD/FMN binding domain	DDxxD motif	SQE homology	Phr-regulation	TMH	Total score
TK1088	Geranygeranyl hydrogenase	+	+	+	+	+	0	5
TK0929	3-octaprenyl-4-hydroxybenzoate decarboxylase	+	+	+		+	0	4
TK2159	Predicted oxidoreductase		+		+	+	1	3
TK1836	Geranygeranyl hydrogenase	+	+	+			0	3
TK0509	3-octaprenyl-4-hydroxybenzoate decarboxylase UbiX	+	+				0	2
TK0513	Hypothetical protein	+				+	0	2
TK1173	Undecaprenyl diphosphate synthase	+				+	0	2
TK1477	Hypothetical protein			+		+	0	2
TK1548	Probable serine—glyoxylate aminotransferase			+		+	0	2
TK1732	Dolichol-phosphate mannosyltransferase	+		+			4	2

Abbreviations: b/t, binding or transferase; FAD, Flavin adenine dinucleotide; FMN, Flavin mononucleotide; SQE, squalene epoxidase; Phr, the Phr-transcriptional regulator; TMH, transmembrane spanning helices

Construction of mutant strains

To screen for enzymes involved in tetraether lipid formation, 12 genes of interest (GOIs) were selected. For each of these genes, we attempted to create a disruption strain and a strain overexpressing the GOI. *T. kodakarensis* strain TS559 was chosen as a parent strain to utilize the previously developed selection and counter-selection procedure for use in rich media (**Figure S5.2a**) [60, 61, 486]. Disruption strains were created as described previously [60]. Transformation of the disruption vectors (**Table 5.2**) resulted in agmatine prototrophs for 11 out of the 12 GOIs. *TK2159* is the only GOI for which no transformants were obtained, even not after 3 independent transformation attempts. Transformants of the remaining 11 GOIs were confirmed to be 6-methyl purine sensitive (6MP_s), a phenotype conferred by the

expression of *TK0664*. Next, disruption of the GOIs as well as precise and markerless deletion of the selection and counter-selection cassette was stimulated by plating the transformants on plates containing 6MP and agmatine. For each of the remaining GOIs, genomic DNA of 30 spontaneous 6MP resistant (6MP_R) colonies was analyzed by diagnostic PCR. In this analysis, we screened for the deletion of both the GOIs and of the selection and counter-selection cassette. Ultimately, 6 of the 12 selected GOIs were successfully deleted. For the remaining 5 (*TK2159* did not result in transformants), only host strains were obtained. For the 6 successful GOI disruptions, representative clones having the desired final genome configurations were designated *T. kodakarensis* strain Δ *TK0513*, Δ *TK1324*, Δ *TK1477*, Δ *TK1548*, Δ *TK1732*, and Δ *TK1836*, in which the Δ denotes disruption and the number indicates which TK-locus has been deleted (**Table 5.2**).

Overexpression strains were obtained in the following way. First, a universal vector backbone was created suitable for integration of an additional copy of the GOIs into the chromosome of *T. kodakarensis*. Next, the 12 selected GOIs were individually cloned into this vector backbone, placing them under control of the strong constitutive *hmtB*-promoter [486], resulting in 12 unique overexpression vectors (**Table 5.2**). Subsequent transformation of these vectors in *T. kodakarensis* TS559 ultimately resulted in agmatine prototrophs for all GOIs. For each GOI, genomic DNA of 5 agmatine prototrophic clones was analyzed by diagnostic PCR to confirm integration of the various overexpression vectors. Moreover, the diagnostic PCR was also used to discriminate between transformants that were the result of a single homologous recombination (resulting in the integration of the entire donor plasmid DNA integrated into the genome of *T. kodakarensis*) and transformants that were the result of two crossover events (resulting in the integration of only the GOI and selection marker (*TK0149*)). The latter was done since only the transformants resulting from two crossover events can be considered stable (and are thus preferred), while the transformants resulting from a single homologous recombination may ultimately revert to the original genome. Nevertheless, both variants are considered sufficient for this study as long as the agmatine prototrophy selection pressure is maintained. The diagnostic PCRs showed that preferred double crossover mutants were obtained for *TK1324*, *TK1477*, *TK1548*, *TK1732*, *TK1088*,

TK1957 and *TK2159*, but not for *TK0509*, *TK0513*, *TK1836*, *TK1173* and *TK0929*. Next, the entire integration locus was PCR amplified and subsequent sequencing confirmed that the integrated additional copy of the GOIs did not contain any undesired mutations. Representative clones, containing correct sequences of the additional copies of the GOIs were designated *T. kodakarensis* strains: OE0509, OE0513, OE0929, OE1088, OE1173, OE1324, OE1468, OE1477, OE1548, OE1732, OE1836, OE1957 and OE2159, in which the OE denotes 'overexpression' and the number specifies the introduced TK gene (**Table 5.2**).

As a final check, total RNA of all OE- and Δ -strains was isolated and used in a semi quantitative Reverse Transcriptase-PCR (RT-PCR) to confirm increased transcription or loss of transcription of the GOIs, respectively. The primers used for the diagnostic PCRs were designed such that expression of a particular gene ultimately resulted in a PCR product of ~250 bp. The results shown in **Figure 5.2** confirm increased transcription levels of the GOIs in their corresponding OE-strain and the absence of transcription in their corresponding Δ -strains. The housekeeping gene *puf* (*TK0308*) [491], encoding the α -subunit of the elongation factor-1 complex, was used for normalization and as positive control in the RT-PCRs for the OE- and Δ -strains, respectively.

Table 5.2. *T. kodakarensis* strains and plasmids used in this study

Strain or plasmid	Relevant characteristics	Source or reference
Strains		
TS559	Δ pyrF; Δ trpE::pyrF; Δ TK0664; Δ TK0149	[486]
Disruption strains		
Δ TK0513	TS559 derivative, Δ TK0513	This study
Δ TK1324	TS559 derivative, Δ TK1324	This study
Δ TK1477	TS559 derivative, Δ TK1477	This study
Δ TK1548	TS559 derivative, Δ TK1548	This study
Δ TK1732	TS559 derivative, Δ TK1732	This study
Δ TK1836	TS559 derivative, Δ TK1836	This study
Overexpression strains		
OE0509	TS559 derivative, Δ TK1765::TK0509, TK0149	This study
OE0513	TS559 derivative, Δ TK1765::TK0513, TK0149	This study
OE0929	TS559 derivative, Δ TK1765::TK0929, TK0149	This study
OE1088	TS559 derivative, Δ TK1765::TK1088, TK0149	This study
OE1173	TS559 derivative, Δ TK1765::TK1173, TK0149	This study
OE1324	TS559 derivative, Δ TK1765::TK1324, TK0149	This study
OE1477	TS559 derivative, Δ TK1765::TK1477, TK0149	This study
OE1548	TS559 derivative, Δ TK1765::TK1548, TK0149	This study
OE1732	TS559 derivative, Δ TK1765::TK1732, TK0149	This study
OE1836	TS559 derivative, Δ TK1765::TK1836, TK0149	This study
OE1957	TS559 derivative, Δ TK1765::TK1957, TK0149	This study
OE2159	TS559 derivative, Δ TK1765::TK2159, TK0149	This study
Plasmids		
pTS700	Basis for the creation of "A"-plasmids, Harboring TK0664; TK0149	[60]
pSKS022	Derivative of pSecPyrF [106], enabling integration at TK1765 locus. Basis for the creation of vector backbone of OE plasmids, harboring TK0149	Chapter 3 and [106]
Disruption vectors*		
p Δ "GOI" (12x)	TK0664; TK0149; Δ "GOI"	This study
Integration vectors*		
pOE"GOI" (12x)	TK0149; Δ TK1765::"GOI"	This study

* = for each gene of interest (GOI) a disruption (Δ) vector and integration (or overexpression; OE) vector was created. The GOIs were TK0509, TK0513, TK0929, TK1088, TK1173, TK1324, TK1477, TK1548, TK1732, TK1836, TK1957, and TK2159.

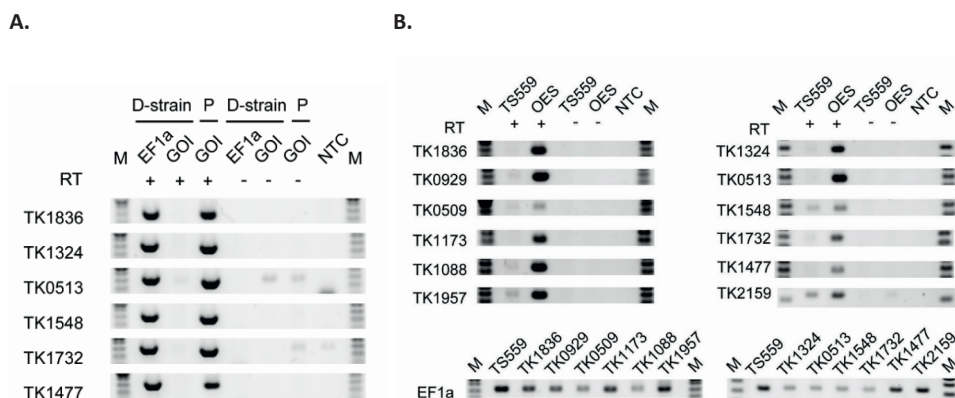


Figure 5.2. Semi quantitative Reverse Transcriptase-PCR (RT-PCR) analysis. Total RNA was isolated from the parental and mutant strains and subjected to a DNase I treatment to degrade any traces of isolated chromosomal DNA. cDNA was created via a reverse transcriptase reaction and subsequently used as template in a diagnostic PCR to confirm (A) loss of expression or (B) overexpression of the genes of interest in the disruption- and overexpression strains, respectively. All RT-PCR products had the appropriate size on SYBR safe stained agarose gels. The housekeeping gene elongation factor 1- α was used for normalization. Abbreviations: D-strain, disruption strain; P, parent strain TS559; M, marker; EF1a, elongation factor 1- α ; GOI, gene of interest; NTC, no template control; RT, Reverse transcriptase; TS559, parent strain; OES, overexpression strain; TK#, gene locus of GOI. “+”, RT present; “-”, RT absent

Phenotypical changes elicited by the OE- and Δ -strains

To assess the involvement of the corresponding GOIs in tetraether lipid biosynthesis, the total diether and tetraether lipid content of the OE- and Δ -strains was determined and compared to that of the parental strain TS559. To do this, overnight cultures were prepared in triplicate and the cells harvested. The membrane lipids were then hydrolysed to cleave off polar head groups and subjected to a total lipid extraction to isolate core lipids. These isolates were subsequently characterized for their diether and tetraether lipid content by UHPLC-API/MS by monitoring ions at m/z 652.6 and 1301.3. However, the total amounts of diether and tetraether lipids (in μg) turned out to be highly variable between the different strains, even after normalization per mg dry cell weight (DCW) (Figure 5.3, in w/w%). For example, the total amount of diether and tetraether lipids obtained for strain OE0929 was nearly twice as high as that obtained for the parental strain TS559, while hardly any diether and tetraether lipids were obtained from strain OE1324.

To allow comparison of the membrane composition of the different strains, the average diether and tetraether lipid content per strain was expressed as a percentage of their cumulative amount (**Figure 5.4**). This showed that the parental strain TS559 has a tetraether lipid content of $93.6 \pm 0.9 \%$, which roughly agrees with the results obtained by Matsuno Y. *et al.* (2009) [546]. Concerning the possible involvement of the GOIs in tetraether lipid biosynthesis, the most striking effects were observed for the strains OE1957 and $\Delta TK1548$, for which a tetraether lipid content of $98.3 \pm 0.5 \%$ and $23.8 \pm 3.9 \%$ was observed, respectively. Moreover, an increased tetraether lipid content was also observed for the OE-strains of *TK0509*, *TK1088*, and *TK1732*, having a tetraether lipid content of $96.3 \pm 1.1 \%$, $95.5 \pm 0.6 \%$, and $96.4 \pm 0.4 \%$, respectively. In addition, a significant decrease in the tetraether lipid content of the OE-strain of *TK1836* was observed. Interestingly, though disruption of *TK1548* led to a significant decrease in the tetraether lipid content, no notable increase was observed upon its overexpression. With the exception of the disruption of *TK1548*, no significant changes in the tetraether lipid content of the obtained disruption strains was observed.

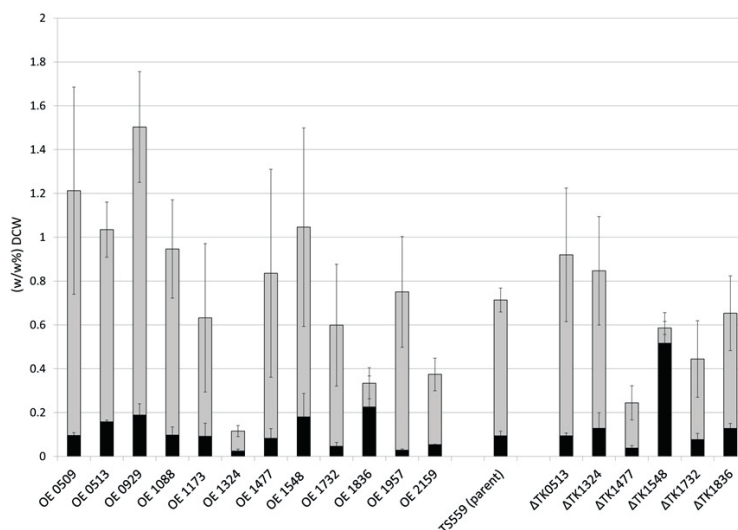


Figure 5.3. Average diether and tetraether lipid content per strain. Light grey: average tetraether lipid content (i.e. GDGT-0), black: average diether lipid content (i.e. archaeol).

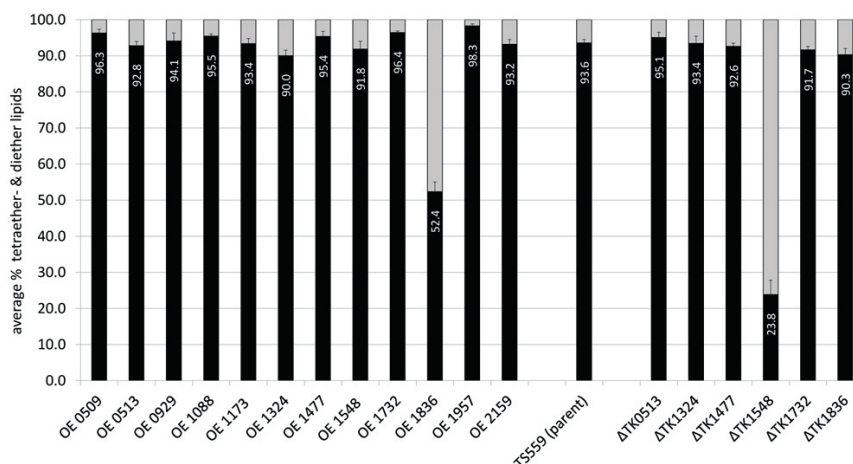


Figure 5.4. Average % of tetraether and diether lipids per strain. The tetraether lipid contents (in black) and diether lipid contents (in light grey) of three biological replicates were normalized to the total amount of lipids measured per strain (i.e. to the sum of GDGT-0 and archaeol). The vertical numbers (in white) show the average fraction (in %) of tetraether lipids per strain.

To investigate if the membrane composition (i.e. the diether:tetraether lipid ratio) has an effect on the physical appearance of *T. kodakarensis*, samples of the parental strain and mutant strains showing the biggest extremes in their membrane lipid composition (i.e. strain OE1957 and $\Delta TK1548$) were studied by transmission electron microscopy (TEM). In addition, samples of the mutant strain OE1324 were included to check its cellular integrity as this strain showed a very low overall lipid content (**Figure 5.3**). The images obtained showed no clear effect of the membrane lipid composition on the physical appearance of the cell nor the membrane of *T. kodakarensis* **Figure 5.5**.

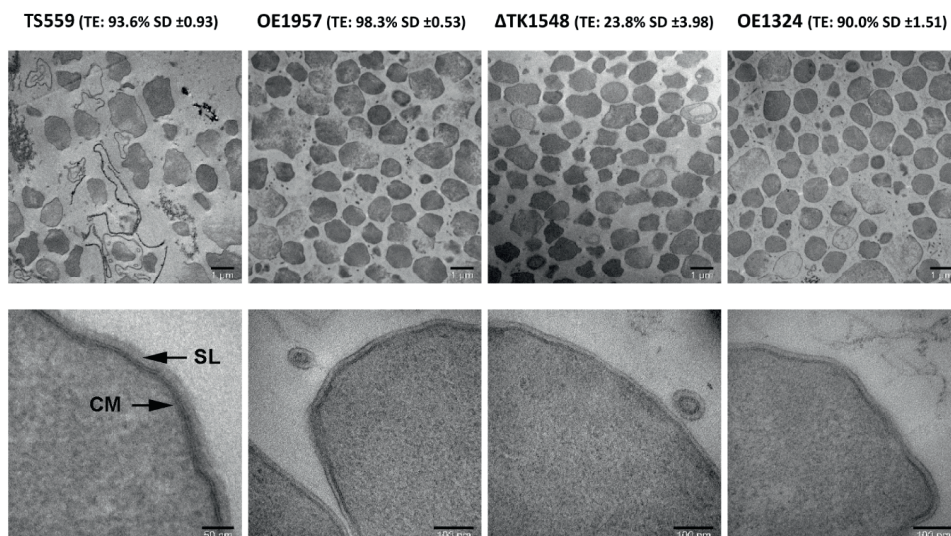


Figure 5.5. Thin cross-sections of *T. kodakarensis* strains TS559, OE1957, Δ TK1548 & OE1324 observed by Transmission Electron Microscopy (TEM). Strain OE1957 and Δ TK1548 were selected for TEM as these strains showed the highest and lowest tetraether lipid content (TE), respectively. Strain OE1324 was included to check its cellular integrity as hardly any diether or tetraether lipids were obtained for this strain in the membrane characterization experiment. The upper micrographs show multiple cells of each strain (bars 1 μ m). The lower micrographs show the cytoplasmic membrane (CM) and S-layer (SL) of individual cells (bar TS559 50 nm, bars OE1957, Δ TK1548 & OE1324 100 nm). The tetraether lipid content (TE) of each strain is shown in brackets.

Discussion

In this study, we set out to identify the enzyme(s) responsible for the synthesis of tetraether lipids. Especially, the one(s) responsible for the condensation of two C_{20} isoprenyl chains, i.e. the putative tetraether synthase(s). To do this, a list of genes that are potentially involved in tetraether synthesis was created, that was prioritized based on a couple of characteristics assumed to be important for a tetraether synthase. This resulted in the selection of 12 genes, which were targeted for overexpression and disruption. Interestingly, subsequent lipid analysis of the mutant strains, strongly suggests that at least three of these genes indeed play a role in the synthesis of tetraether lipids. These genes are *TK1548*, *TK1836* and *TK1957*. In addition, the genes that appear to be non-viable upon disruption may be

interesting in this respect as well. Especially, the ones for which no disruption strain was obtained, but a slight increase in the tetraether lipid content was observed upon its overexpression, such as *TK0509* and *TK1088*. This is because the membrane of the parental strain primarily exists of tetraether lipids (i.e. $93.6\% \pm 0.93$). Losing the ability to synthesize tetraether lipids due to the disruption of the tetraether synthase may thus very well be lethal.

One of the genes that seems play a role in the synthesis of tetraether lipids is *TK1548*, which is annotated as a serine-glyoxylate aminotransferase. This gene was selected from the list of candidate genes as a priority gene since it contains a DDxxD motif, which is associated with substrate binding of prenyltransferases, and it has been shown to be slightly upregulated upon heat shock (**Table 5.1**). Disruption of *TK1548* resulted in a strain with a significantly lower tetraether lipid content (**Figure 5.4**), which is a strong indication that the enzyme encoded by this gene plays an important role in tetraether synthesis. However, despite the disruption of *TK1548*, a considerable amount of tetraethers could still be detected (i.e. $23.8\% \pm 3.9$). This either suggests it is not the sole tetraether synthase encoding gene, and hence that there might be multiple isozymes, or that *TK1548* is encoding a different enzyme of importance to tetraether synthesis. Moreover, since this gene has been annotated as an aminotransferase it is possible that the enzyme it encodes shows a broad substrate specificity, meaning that it could be involved in many pathways and that the observed effect may be indirect. To obtain clarity about its involvement in tetraether lipid synthesis more experimental data is needed. Unfortunately, overexpression of *TK1548* did not result in a significant effect (**Figure 5.4**), which may suggest it is indeed not a tetraether synthase, but might also simply imply it is not a rate limiting step in the pathway.

A second gene that has an influence on the tetraether lipid content is *TK1836*. This gene has been annotated as a geranylgeranyl hydrogenase and was selected as a priority gene for its predicted prenyl binding, and the presence of a FAD/FMN binding domain and DDxxD motif (**Table 5.1**). Though disruption of *TK1836* does not seem to have a detrimental effect on the tetraether lipid content of the membrane, overexpression surprisingly does (**Figure 5.4**). These results imply that *TK1836* is involved in lipid biosynthesis, but that it is very unlikely to

be a tetraether synthase. Instead, subsequent phylogenetic analysis suggests that *TK1836* probably belongs to a dehydrogenase family that has been shown to be prenyl reductases (data not shown) [585, 586]. If true, the results obtained for this mutant suggest that the presence of double bonds in the precursor molecule of tetraether lipids is a prerequisite to the formation of tetraether lipids. This would be an interesting finding as the absence or presence of double bonds in the C₂₀ isoprenyl chains (i.e. before the condensation reaction) is a major difference between the classical and alternative tetraether biosynthesis pathway, respectively [110].

A third gene that seems to be involved in tetraether lipid synthesis is *TK1957*, which has been annotated as a 4-hydroxybenzoate polyprenyltransferase. Though this gene shows sequence similarity to experimentally characterized DGGGP synthases [570, 572] (which should lead to the exclusion of this gene from the list of candidate genes), it was decided to retain it and to select it as a priority gene for further analysis. The reason for this was that another candidate gene also annotated as a 4-hydroxybenzoate polyprenyltransferase, i.e. *TK1324*, showed sequence similarity to DGGGP synthases as well, making it difficult to decide which one is a DGGGP synthase and which one may potentially be a tetraether synthase. It was therefore decided to retain both and to select them as priority genes for further analysis. Interestingly, overexpression of *TK1957* resulted in a strain that has a membrane that almost exclusively exists of tetraether lipids (i.e. 98.3% ± 5; **figure 5.4**). This is a strong indication that *TK1957* is involved in the tetraether lipid synthesis and may perhaps even imply it is the tetraether synthase. Moreover, the observation that disruption of *TK1957* appears to be non-viable is quite interesting in this respect as well. However, though these results may imply that *TK1957* is a tetraether synthase, a similar reasoning as described above would also apply to a DGGGP synthase. This is because, in both the classical and alternative biosynthesis pathways proposed, DGGGP synthases are key enzymes in the formation of archaeal membrane lipids (**Figure 5.1**). It therefore seems likely that the annotation of *TK1957* is correct and that *TK1957* is not a tetraether synthase, but a DGGGP synthase.

Contrary to *TK1957*, overexpression and disruption of *TK1324* (the other gene annotated as a DGGGP synthase) did not result in a significant change in the diether/tetraether lipid ratio (**Figure 5.4**). This implies that *TK1324* is most likely neither a DGGGP nor a tetraether synthase. However, the significant decrease in the total amount of diether and tetraether lipids obtained upon its overexpression (**Figure 5.3**), combined with the observation that the physical appearance of these cells nevertheless appears to be normal (**figure 5.5**), suggests that *TK1324* is involved in the biosynthesis of lipids other than the ones measured in this study. The analysis performed in this study encompasses a selected ion monitoring rather than a total ion monitoring, and as such was aimed to detect only the most prevalent lipids in the membrane which are represented by the core lipid species archaeol and GDGT-0. Therefore, before any conclusions can be drawn, more research is needed in which effects on lipid species other than archaeol and GDGT-0 are also investigated.

In conclusion, though interesting results have been obtained future research is needed to get a better understanding of the tetraether synthesis pathway and to identify the tetraether synthase(s). With respect to the latter, *TK1548* especially seems to be an interesting candidate, though the observed effects may be indirect. Moreover, the candidate genes for which no disruption strains were obtained might be worth investigating further as disruption of a tetraether synthase may very well be lethal. In addition, more of the genes present on the candidate list (**Table S5.1**) may be targeted for overexpression and disruption, after which the subsequent analysis of the membrane composition should include multiple lipid species. All together, these analyses could provide a better understanding of the tetraether lipid biosynthesis pathway and may ultimately lead to the identification of the tetraether synthase(s).

Materials and Methods

Selection of candidate genes

Two independent bioinformatics approaches were performed. In the first analysis, a set of reference species that are phylogenetically closely related to *T. kodakarensis* and that have a confirmed presence (*Pyrococcus furiosus* and *Thermoplasma acidophilum*) or absence (*Natronomonas pharaonis* and *Methanopyrus kandleri*) of tetraethers in their membranes

was selected [553, 564-568]. The genome of *T. kodakarensis* was screened for homologs present in the first two, but which are absent in the latter two by performing a phylogenetic profiler single gene search using the IMG system of JGI (<https://img.jgi.doe.gov/cgi-bin/m/main.cgi>). In the second analysis, it was attempted to correlate the presence of specific arCOGs (archaeal Clusters of Orthologous Genes [569] in 149 sequenced archaeal genomes to species with a confirmed tetraether lipid membrane.

The list of genes resulting from two bioinformatics approaches was prioritized based on the following characteristics: i) predicted prenyl binding, ii) the presence of a FAD or Flavin mononucleotide (FMN) binding domain or Rosmann Fold [587], iii) the presence of aspartic acid rich DDxxD motifs, iv) homology to squalene epoxidase, and v) transcriptional regulation by the *Thermococcales* heat shock protein Phr. Screening for i and ii was done by scanning the protein sequences against the InterPro collection of protein signature databases (<http://www.ebi.ac.uk/Tools/pfa/iprscan5/>). The screening for iii was done by manually checking the protein sequences for presence of a DDxxD motif. To identify possible homologs of squalene epoxidase (iv), the selected genes were included in a BLAST search using the *Saccharomyces cerevisiae* ERG1 squalene epoxidase sequence (EC1.14.99.7) as a query. The *S. cerevisiae* squalene epoxidase was selected, as a recently created three-dimensional model of this protein allowed for the identification of crucial residues [588]. To determine if the transcription of the candidate genes may be influenced by Phr (v), we made use of the results of a previously published microarray experiment that was performed to identify the Phr-dependent heat shock regulon in *T. kodakarensis* [582]. The presence or absence of transmembrane spanning helices was determined by constructing hydrophobicity plots using the online Transmembrane Hidden Markov Model (TMHMM <http://www.cbs.dtu.dk/services/TMHMM-2.0/>) tool to predict probability of transmembrane spanning segments. From this analysis it was shown that the following gene products (*TK1324*, *TK1732*, *TK1957* and *TK2159*) are integral membrane proteins or membrane associated.

Microorganisms and culture conditions

E. coli DH5 α was used for general DNA manipulation, and *E. coli* strains were cultivated at 37°C in LB medium supplemented with 100 μ g/ml ampicillin. *T. kodakarensis* strains and plasmids used in this study are listed in **Table 5.2**. *T. kodakarensis* strains were cultivated anaerobically in stoppered serum bottles at 85°C in a nutrient-rich medium (ASW-YT) [62], composed of 0.8x artificial seawater (0.8x ASW), 5.0 g/L yeast extract, 5.0 g/L tryptone, 1 ml/L KOD vitamin mixture (1000x) and 0.8 mg/L resazurine. The headspace of all cultures was exchanged for N₂. Before inoculation, either 2.0 g/L sulfur (ASW-YT-S⁰) or 5.0 g/L pyruvate (ASW-YT-Pyr) was added to the medium, as well as Na₂S \cdot 9H₂O to reduce the dissolved oxygen, until the resazurin became colorless. When required, the media were supplemented with 500 μ M agmatine and/or 75 μ M 6-methyl purine (6MP) [60, 61]. Cultures were continuously shaken (50 rpm) during growth. In the case of plate culture, 1% (wt/vol) Gelrite was added to solidify the medium together with 2 ml/L polysulfide solution [62] instead of S⁰. Cells inoculated on plates were incubated at 85°C for at least 48 h in an air-tight anaerobic jar containing an Anaerocult A mini (Merck Millipore, Darmstadt, Germany) to maintain anaerobic conditions.

Construction of plasmids and mutant strains

Plasmids incapable of autonomous replication in *T. kodakarensis* were constructed and amplified in *Escherichia coli*, carrying the genes of interest (GOIs) or knockout constructs adjacent to the previous developed selection and, in the case of the disruption vectors, counter-selection cassette (*TK0149*; *TK0664*) [60, 61]. The GOIs and selection cassette (in case of the OE vectors) and the selection/counter-selection cassette (in case of the disruption vectors) were flanked by a ~500-1000 bp region homologous to *T. kodakarensis* genome that naturally was located immediately up- and downstream of the target locus (**Figure S5.2**). Standard molecular biology procedures were used for PCR amplifications, plasmid construction, *E. coli* DH5 α transformation, transformants selection, plasmid isolation, and sequence verification. The sequences of oligonucleotide primers used in this study are listed in **Table S5.3**.

The vectors created to disrupt or overexpress the GOIs are shown in **Table 5.2**. The disruption vectors were constructed as described before [60]; first, so-called “A”-plasmids were created by amplifying each GOI along with ~500-700 bp of flanking DNA and cloning the corresponding amplicon into pTS700 in a ligation-independent mechanism. Next, the “A”-plasmids served as a foundation plasmid from which so-called “B”-plasmids, or disruption vectors, were generated. This was done by a single-step PCR-based procedure wherein the original GOI was deleted from the plasmid, while leaving the flanking DNA that will target integration of the entire B-plasmid to the TS559 genome. Two separate initial integration events are possible for incorporation of the entire B-plasmid to the genome, and each generates an intermediate strain that, when 6MP-based counter-selective pressure is applied, can undergo an internal recombination event to either restore the original genome or produce a strain containing a genome with the desired, targeted deletion (**Figure S5.2a**).

The overexpression vectors were constructed as follows. First an inverse PCR using the primer-pair BG4881-BG4882 and the previously constructed integration vector pSKS022 (**Chapter 3**) as template was performed. pSKS022 is a derivative of pSecPyrF [106] and was chosen as template, as integration in the genome at the *ChiA* locus targeted by this plasmid has shown to be efficient and non-disturbing. The PCR resulted in the amplification of a vector backbone for the overexpression vectors, that contained the flanking regions needed for integration as well as the strong constitutive *hmtB*-promoter. The PCR thus removed the *csg*-promoter, the *TM*-gene expression cassette, the synthetic transcription terminator (T_8), the 500-bp pop-out homology, and the counter-selection marker *TK0664* from pSKS022. The linear vector backbone created by this PCR was extended with a unique *AgeI* restriction site, which was incorporated in the primer BG4882. Next, the GOIs were amplified by Phusion polymerase PCR (Thermo Fisher Scientific) using the corresponding primers listed in **Table S5.3**. The reverse primers were extended with a unique *AgeI* restriction site, allowing for easy and directional ligation in the amplified vector backbone, creating the desired vectors as shown in **Figure S5.2b**. Transcription of the GOIs was designed to be controlled by the strong constitutive *hmtB*-promoter [486].

The *T. kodakarensis* Δ - and OE-strains shown in **Table 5.2** were generated by transforming the corresponding vectors into *T. kodakarensis* TS559 (Δ pyrF; Δ trpE::pyrF; Δ TK0664; Δ TK0149) [60, 61, 486]. Disruption strains were created as described previously [60]. Overexpression strains were created by a slightly adapted approach in which the counter-selection marker (*TK0664*) was not involved and desired mutants were thus simply selected by growing colonies on medium lacking agmatine. An important difference between these two methods is that, as a result of excluding *TK0664*, the selection marker (*TK0149*) cannot be removed from the genome after obtaining initial transformants. This means that, in contrast to the disruption strains, the created overexpression strains do not allow another round of the procedure to make additional genome modifications. Diagnostic PCRs were performed to select and confirm genotypes of desired strains. To do this, 6MP-resistant colonies growing on ASW-YT plates supplemented with agmatine (in the case of the disruption strains) or agmatine-prototrophic colonies growing on ASW-YT plates without agmatine (in the case of the overexpression strains) were picked and used to inoculate liquid ASW-YT- S⁰ cultures (with and without agmatine, respectively), which were incubated o/n at 85°C. Next, 1.5-ml culture was sampled using a syringe and genomic DNA was isolated according to the protocol for Gram-positive bacteria of the GeneJET genomic DNA purification kit (Thermo Scientific). Purified DNA was then used as template in three diagnostic PCRs. In the first PCR, the primers used were complementary to the sequence upstream of the desired integration locus as well as to the C-terminus of the integrated GOI (illustrated by primers A and B in **Figure S5.2**). Similarly, a second PCR was included using primers that were complementary to the sequence downstream of the desired integration locus as well as to the N-terminus of the integrated GOI (illustrated by primers C and D). A third PCR was performed to identify double crossover mutants, using primers that were complementary to the sequence downstream and upstream of the integration locus (illustrated by primers A and D). A generated PCR product having the correct size for all three reactions suggested the desired double crossover genotype. In addition, fragments containing the additional copy of the GOI were PCR amplified from the genome of the desired clones and sent for sequencing to confirm the absence of undesired mutations. Sequences of the primers are listed in **Table S5.3**.

RNA isolation and RT-PCR

To confirm loss of transcription or increased mRNA levels of the GOIs in the disruption strains or overexpression strains, respectively, reverse transcriptase PCR (RT-PCR) was performed. To do this, the parental strain TS559 and the various mutant strains created in this study were cultivated in 120-ml serum bottles containing 40 ml ASW-YT-S⁰ medium, which in the case of the disruption strains were also supplemented with 500 μ M agmatine and 75 μ M 6MP. Samples of 10 ml were removed and cells were harvested by centrifugation (4700x *g*, 30 min, 4°C). Total RNA was isolated by the following method. Cell pellets were resuspended in 1 ml Trizol and 200 μ l chloroform was added. The mixtures were shaken by hand for 15 seconds, followed by an incubation at RT for 3 min. Next, the samples were centrifuged (13,000x *g*, 15 min, 4°C) and the supernatants were transferred to fresh Eppendorf tubes. 5 μ g glycogen and 500 μ l isopropanol were added and the mixtures were incubated at RT for 10 min. The samples were centrifuged (13,000x *g*, 15 min, 4°C) and the pellets were washed in 75% ethanol, air-dried for 10 min and resuspended in 20 μ l MQ water. RNA concentrations were determined with a Nanodrop ND-1000 and the total RNA samples were stored at -20°C until further processing.

Before RT-PCR, the total RNA samples (1 μ g each) were treated with DNaseI, to remove residual genomic DNA. cDNA was prepared by following the protocol for first strand cDNA synthesis of the Thermo Scientific RevertAid RT kit (20 μ l reactions), using random hexamers as primers. Gene-specific cDNA fragments were amplified by conventional PCR using the freshly prepared cDNA as template and primer pairs developed to specifically target the genes of interest (**Table S5.3**). The housekeeping gene *TKO308 (puf)*, encoding the α -subunit of the elongation factor-1 complex (EF-1- α), was included for normalization [491]. All primer pairs were designed to produce a specific product of ~250 bp. All reactions were run for the amount of cycles needed to make the respective product in the reactions containing cDNA of the parental strain TS559 just visible. The only exception to this are the reactions used for normalization (*TKO308*), which were run for 18 cycles. PCR amplification products were analyzed by agarose gel electrophoresis. Control experiments were included by following the

same procedure, but without addition of RevertAid Reverse Transcriptase enzyme to the RT reaction mixture.

Core lipid extraction and analysis by UHP-APCI/MS

Triplicate cultures of the various mutant strains and parental strain TS559 were cultured on ASW-YT-Pyr. When the cultures reached the stationary phase, cells were harvested by 10 min centrifugation at 4700x *g* and lyophilized overnight. To extract the core lipids of the various strains, lyophilized cells were hydrolysed with 1 N 5% HCl in methanol by reflux at 130 °C for 3 h. The hydrolysate was adjusted to pH 7 with 2 N KOH-methanol (MeOH) (1:1, vol/vol) and, after addition of water to a final 1:1 ratio of H₂O-MeOH, extracted three times with dichloromethane (DCM). The DCM fractions were collected and dried over sodium sulfate. The resulting extract was dried under nitrogen and dissolved in hexane–2-propanol (99:1, vol/vol), filtered over a 0.45-µm polytetrafluoroethylene filter, and analysed by high-performance liquid chromatography (HPLC)–atmospheric pressure chemical ionization (APCI) mass spectrometry (MS) for archaeol and GDGTs. The UHPLC-APCI/MS analysis was done according to Hopmans *et al.* [589], with some modifications. Briefly, analysis was performed on an Agilent 1260 UHPLC coupled to a 6130 quadrupole MSD in selected ion monitoring (SIM) mode. Separation was achieved on two UHPLC silica columns (BEH HILIC columns, 2.1 x 150 mm, 1.7 µm; Waters) in series, fitted with a 2.1 x 5 mm pre-column of the same material (Waters) and maintained at 30°C. Archaeol and GDGTs were eluted isocratically for 10 min with 10% B, followed by a linear gradient to 18% B in 20 min, then a linear gradient to 100% B in 20 min, where A is hexane and B is hexane: isopropanol (9:1). Flow rate was 0.2 ml/min. Total run time is 60 min with a 20 min re-equilibration. Source settings were identical to Schouten *et al.* (2007) [590]. Typical injection volume was 5 µl of a 1 mg/ml solution. *M/z* of the protonated molecules of archaeol and isoprenoid GDGTs were monitored. Archaeol and GDGTs were quantified by comparing the response in the appropriate mass chromatogram to the response of known amounts of standards (archaeol obtained from Avanti Polar Lipids, Inc. Alabaster, Alabama, USA; GDGT-0 prepared according to Sinninghe Damste *et al.*, 2002 [591])

Transmission electron microscopy (TEM) of thin sections.

Cells of *T. kodakarensis* strains TS559, OE1324, OE1957 and Δ 1548 cultured in ASW-YT-Pyr and harvested in the stationary phase were fixed by incubation in 3% glutaraldehyde in 0.1 M sodium cacodylate buffer pH 7.2 for 1.5 h at room temperature. After three washes in the same buffer, a post fixation was done in 1% (w/v) OsO₄ and 1.5% (w/v) potassium ferricyanide in the same buffer for 1 h at room temperature and subsequently rinsed three times in water. One volume of the sample suspensions were mixed with equal volumes of 3% low melting point agarose. Between each step the samples were centrifuged for 5 min at 8,000 rpm (5220x *g*) and the pellets were suspended again in the next liquid. After centrifugation the agarose was solidified at 0°C and the parts containing the pelleted sample were cut in very small pieces. The sample pieces were dehydrated in a graded ethanol series (10, 30, 50, 70, 90, 100, 100%), infiltrated with modified Spurr resin mixture (Serva) and polymerized for 8 h at 70°C. The samples were ultra-thin (70-80 nm) sectioned with an ultra-microtome (Reichert, Ultracut S), post-stained with 5 times diluted Reynold's lead citrate for 5 minutes and analyzed with a transmission electron microscope (JEOL, JEM 1011). Micrographs were made with a digital camera (Olympus, Veleta). Sample preparation and imaging were performed by the Wageningen Electron Microscopy Centre (WEMC) facility.

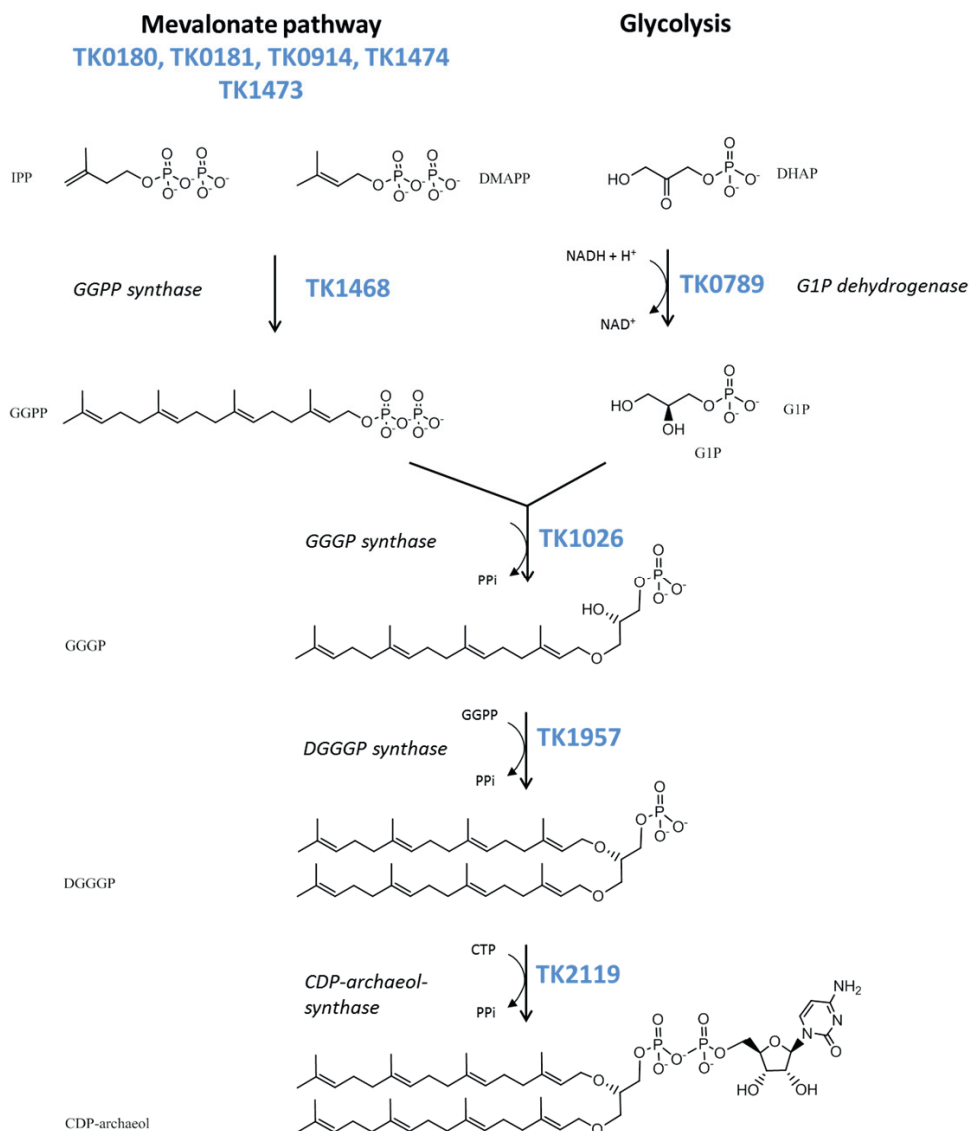
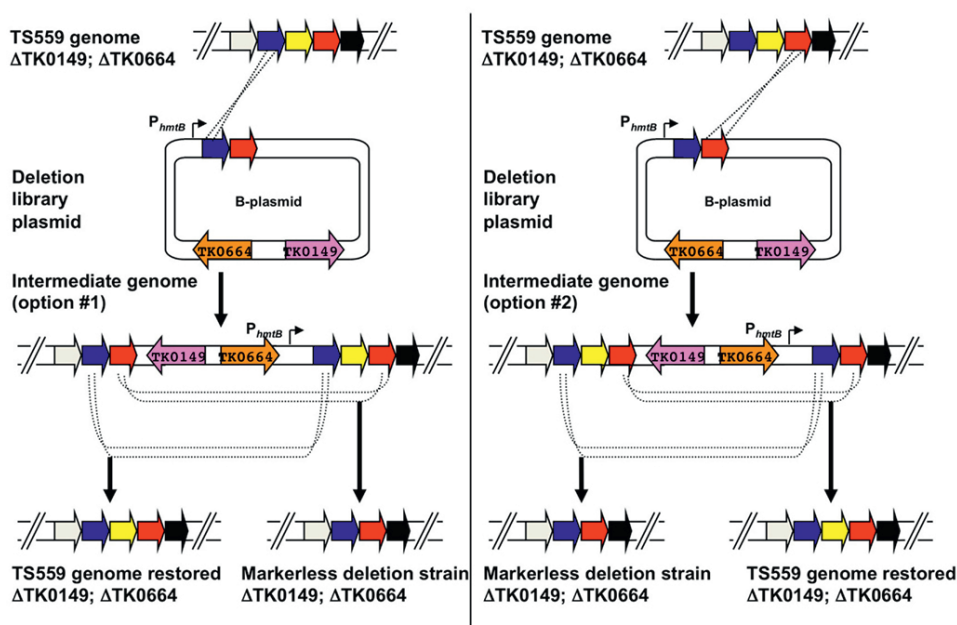


Figure S5.1 Hypothetical pathway responsible for the biosynthesis of CDP-archaeol in *T. kodakarensis*. CDP-archaeol is a CDP-activated precursor of the final diether lipid (i.e. archaeol). The pathway was constructed by BLASTing the sequences of experimentally characterized genes, that have shown to be responsible for the reactions leading to isoprenoid and CDP-archaeol synthesis in various archaea, against the genome of *T. kodakarensis* [570-577]. *TK1957* was analysed in this study and has shown to be involved in GDGT formation.

A.



B.

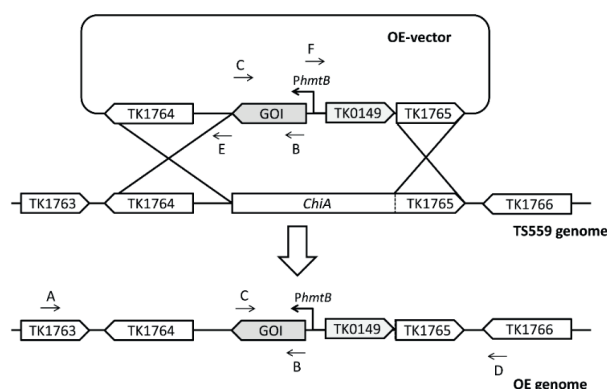


Figure S5.2 Schematic diagram of disruption and overexpression procedure in *T. kodakarensis*. a) Use of disruption vectors (B-plasmids) to generate markerless knockouts in the genome of *T. kodakarensis* TS559. The image has been reused from Hileman, T.H., and Santangelo, T.J. (2012) [60]. A hypothetical region of the genome of *T. kodakarensis* strain TS559 is shown at top, with the left and right panels depicting the two possible integration events yielding agmatine-prototrophic intermediate strains

from the diagrammed B-plasmid. Both intermediate strains #1 and #2 contain direct repeats flanking the target locus, and dependent on recombination responsible for excision upon counter-selection with 6MP, either the original TS559 genome can be restored or the desired deletion genome generated. b) Use of overexpression vectors (OE-vector) to generate *T. kodakarensis* strains overexpressing the gene of interest (GOI). The GOI is integrated at the *ChiA*-locus of *T. kodakarensis* strain TS559 via a double homologous recombination event yielding agmatine-prototrophic OE-strains. The letters A-D represent the binding sites of the primers used in the diagnostic PCRs. The letters E & F represent the binding sites of the primers used to amplify the vector backbone using pSKS022 as a template. Abbreviations: TK0149, agmatine selection marker; TK0664, counter-selection marker; P_{hmtB} , strong constitutive promoter

Table S5.1 genes that are potentially involved in tetraether synthesis

a) Results of phylogenetic profiler single gene search

Gene locus	Gene annotation	prenyl binc	FAD FMN	DDxD	SOE hom	Phr	TMH	+score	selected
TK0038	archaeal flagellin B1 precursor			+		+	1	2	
TK0039	archaeal flagellin B2 precursor			+		+	1	2	
TK0040	archaeal flagellin B3 precursor					+	1	1	
TK0041	archaeal flagellin B4 precursor					+	1	1	
TK0042	archaeal flagellin B5 precursor					+	1	1	
TK0049	archaeal flagella-related membrane protein J						9	0	
TK0135	indolepyruvate: ferredoxin oxidoreductase, beta subunit					+	1	1	
TK0136	indolepyruvate: ferredoxin oxidoreductase, alpha subunit					+	0	1	
TK0139	acetyl-CoA synthetase II (NDP forming), alpha subunit		+			+	0	2	
TK0147	spermidine synthase					+	1	1	
TK0172	hypothetical membrane protein			+			6	1	
TK0182	hypothetical protein						0	0	
TK0188	hypothetical protein						0	0	
TK0306	DEAD/DEAH box RNA helicase					+	0	1	
TK0314	hydrolase, metallo-beta-lactamase superfamily					+	0	1	
TK0428	riboflavin biosynthesis protein Riba		+				0	1	
TK0537	peroxiredoxin, AhpC/TSA family					+	0	1	
TK0624	multisubunit sodium/hydrogen antiporter, MnhC subunit						3	0	
TK0711	hypothetical protein						0	0	
TK0810	oligosaccharyl transferase, STT3 subunit					+	13	1	
TK0817	2-oxoacid:ferredoxin oxidoreductase, alpha subunit						0	0	
TK0834	transcription regulator, Lrp/AsnC family		+				0	1	
TK0882	Agmatinase						0	0	
TK0893	pyruvate formate- lyase activating enzyme-related protein					+	0	1	
TK0929	3-octaprenyl-4-hydroxybenzoate decarboxylase	+	+	+		+	0	4	x
TK0950	asparagine synthase (glutamine-hydrolyzing)						0	0	
TK0951	XPA-binding protein 1 homolog						0	0	
TK0971	hypothetical membrane protein						4	0	
TK1008	Fe-containing alcohol dehydrogenase						0	0	
TK1045	cobalamin adenosyltransferase					+	0	1	
TK1057	iron(II) uptake regulation protein						0	0	
TK1219	membrane bound hydrogenase, MbxH' subunit					+	14	1	
TK1220	membrane bound hydrogenase, MbxH subunit					+	13	1	
TK1227	predicted regulator of amino acid metabolism, containing ACT domain					+	0	1	
TK1258	hypothetical membrane protein					+	7	1	x
TK1285	transcription regulator, Lrp/AsnC family					+	0	1	
TK1296	hypothetical protein						0	0	

TK1386	metallophosphoesterase, calcineurin superfamily						+	0	1
TK1393	anaerobic glycerol 3-phosphate dehydrogenase				+			0	1
TK1396	glycerol kinase						+	0	2
TK1561	hypothetical membrane protein							+	1
TK1592	S-adenosylmethionine decarboxylase proenzyme							+	0
TK1659	small neutral amino acid transporter							+	0
TK1718	oligosaccharyl transferase, STT3 subunit							+	11
TK1732	dolichol-phosphate mannosyltransferase				+			+	4
TK1806	probable dihydroorotate dehydrogenase, electron transfer subunit				+			0	1
TK1812	hypothetical membrane protein							+	2
TK1820	membrane-associated metalloprotease							+	6
TK1851	hypothetical membrane protein, conserved							+	5
TK1852	hypothetical membrane protein, conserved							+	4
TK1890	histone macroH2A-related protein (C-terminus)							+	0
TK2052	predicted transcription regulator, Lrp/AsnC family							+	0
TK2087	membrane bound hydrogenase, MbHf subunit							+	13
TK2112	hypothetical protein							+	0
TK2257	deoxycytidylate deaminase							+	0
TK2275	RNA-binding protein, containing PUA domain							+	0
TK2296	hypothetical protein, conserved, radical SAM superfamily							+	0

b) Uncharacterized *T. kodakarensis* genes present in the selected arCOGs

Gene locus	COG	COG annotation	Gene annotation	prenyl bind	FAD FMN	DBXD	SQE hom	Phr	TMH	+score	selected
TK0059	COG0558	Phosphatidylglycerophosphate synthase	CDP-alcohol phosphatidyltransferase					+	3	1	
TK0179	COG1545	Predicted nucleic-acid-binding protein containing a Zn-ribbon	Predicted nucleic-acid-binding protein containing a Zn-ribbon					+	0	1	
TK0180	COG0183	Acetyl-CoA acetyltransferase	Acetyl-CoA acetyltransferase			+		+	0	2	
TK0217	COG0214	Pyridoxine biosynthesis enzyme	Pyridoxine biosynthesis enzyme					+	0	1	
TK0283	COG0119	isopropylmalate/homocitrate/citramalate synthases	homocitrate synthase						0	0	
TK0509	COG0163	3-polyprenyl-4-hydroxybenzoate decarboxylase	3-octaprenyl-4-hydroxybenzoate decarboxylase UbiX	+	+				0	2	x
TK0513	COG1689	Uncharacterized protein conserved in archaea	hypothetical protein	+				+	0	2	x
TK0551	COG1052	Lactate dehydrogenase and related dehydrogenases	D-isomer specific 2-hydroxyacid dehydrogenase					+	0	1	
TK0683	COG1052	Lactate dehydrogenase and related dehydrogenases	glyoxylate reductase						0	0	
TK0789	COG0371	Glycerol dehydrogenase and related enzymes	glycerol-1-phosphate dehydrogenase						0	0	
TK0858	COG1267	Phosphatidylglycerophosphatase A and related proteins	hypothetical protein						0	0	
TK0939	COG0043	3-polyprenyl-4-hydroxybenzoate decarboxylase and related decarboxylases	3-octaprenyl-4-hydroxybenzoate decarboxylase	+	+	+		+	0	4	x
TK1088	COG0644	Dehydrogenases (flavoproteins)	geranylgeranyl hydrogenase	+	+	+		+	0	5	x
TK1173	COG0020	Undecaprenyl pyrophosphate synthase	undecaprenyl diphosphate synthase	+				+	0	2	x
TK1324	COG0382	4-hydroxybenzoate poly(prenyl)transferase and related prenyltransferases	4-hydroxybenzoate octaprenyltransferase	+					6	1	x
TK1444	COG0083	Homoserine kinase	Homoserine kinase						0	0	
TK1473	COG1608	Predicted archaeal kinase	amino acid kinase						0	0	
TK1477	COG1355	Predicted dioxygenase	hypothetical protein			+		+	0	2	x
TK1548	COG0075	Serine-pyruvate aminotransferase/archaeal aspartate aminotransferase	probable serine-glyoxylate aminotransferase			+		+	0	2	x
TK1806	COG0167	Dihydroorotate dehydrogenase	probable dihydroorotate dehydrogenase		+				0	1	
TK1831	COG0153	Galactokinase	Galactokinase						0	0	
TK1836	COG0644	Uncharacterized protein conserved in archaea	geranylgeranyl hydrogenase	+	+	+			0	3	x
TK1849	COG4072	Uncharacterized protein conserved in archaea	hypothetical protein					+	0	1	
TK1957	COG0382	4-hydroxybenzoate poly(prenyl)transferase and related prenyltransferases	4-hydroxybenzoate octaprenyltransferase	+					8	1	x
TK2119	COG0575	CDP-diglyceride synthetase	hypothetical protein						5	0	
TK2159	COG1233		predicted oxidoreductase					+	1	3	x
TK2279	COG0558	Phosphatidylglycerophosphate synthase	bifunctional sugar nucleotidyltransferase/CDP-alcohol phosphatidyltransferase synthase						3	1	

Table S5.2. arCOGs included in the second comparative genome analysis.

COG0020	Undecaprenyl pyrophosphate synthase
COG0039	Malate/lactate dehydrogenases
COG0043	3-polyprenyl-4-hydroxybenzoate decarboxylase and related decarboxylases
COG0075	Serine-pyruvate aminotransferase/archaeal aspartate aminotransferase
COG0083	Homoserine kinase
COG0119	Isopropylmalate/homocitrate/citramalate synthases
COG0142	Geranylgeranyl pyrophosphate synthase
COG0153	Galactokinase
COG0163	3-polyprenyl-4-hydroxybenzoate decarboxylase
COG0167	Dihydroorotate dehydrogenase
COG0183	Acetyl-CoA acetyltransferase
COG0210	Superfamily I DNA and RNA helicases
COG0214	Pyridoxine biosynthesis enzyme
COG0240	Glycerol-3-phosphate dehydrogenase
COG0280	Phosphotransacetylase
COG0282	Acetate kinase
COG0318	Acyl-CoA synthetases (AMP-forming)/AMP-acid ligases II
COG0331	(acyl-carrier-protein) S-malonyltransferase
COG0371	Glycerol dehydrogenase and related enzymes
COG0382	4-hydroxybenzoate polyprenyltransferase and related prenyltransferases
COG0407	Uroporphyrinogen-III decarboxylase
COG0417	DNA polymerase elongation subunit (family B)
COG0439	Biotin carboxylase
COG0469	Pyruvate kinase
COG0558	Phosphatidylglycerophosphate synthase
COG0575	CDP-diglyceride synthetase
COG0578	Glycerol-3-phosphate dehydrogenase
COG0644	Dehydrogenases (flavoproteins)
COG0688	Phosphatidylserine decarboxylase
COG0761	Penicillin tolerance protein
COG1022	Long-chain acyl-CoA synthetases (AMP-forming)
COG1024	Enoyl-CoA hydratase/carnithine racemase
COG1052	Lactate dehydrogenase and related dehydrogenases
COG1154	Deoxyxylulose-5-phosphate synthase
COG1183	Phosphatidylserine synthase
COG1233	Phytoene dehydrogenase-related protein
COG1239	Mg-chelatase subunit ChII
COG1250	3-hydroxyacyl-CoA dehydrogenase
COG1257	Hydroxymethylglutaryl-CoA reductase
COG1267	Phosphatidylglycerophosphatase A and related proteins
COG1304	L-lactate dehydrogenase (FMN-dependent) and related alpha-hydroxyl acid dehydrogenases
COG1331	Highly conserved protein containing a thioredoxin domain
COG1348	Nitrogenase subunit NifH (ATPase)
COG1355	Predicted dioxygenase
COG1429	Cobalamin biosynthesis protein CoBN and related Mg-chelataes
COG1443	Isopentenylidiphosphate isomerase
COG1502	Phosphatidylserine/Phosphatidylglycerophosphate/cardiolipin synthases and related enzymes
COG1545	Predicted nucleic-acid-binding protein containing a Zn-ribbon
COG1562	Phytoene/squalene synthetase

Chapter 5

COG1577	Mevalonate kinase
COG1608	Predicted archaeal kinase
COG1646	Predicted phosphate-binding enzymes, TIM-barrel fold
COG1657	Squalene cyclase
COG1689	Uncharacterized protein conserved in archaea
COG1960	Acyl-CoA dehydrogenases
COG2025	Electron transfer flavoprotein, alpha subunit
COG2057	Acyl-CoA acetate/3-ketoacid CoA transferase, beta subunit
COG2086	Electron transfer flavoprotein, beta subunit
COG2098	Uncharacterized protein conserved in archaea
COG2227	2-polyprenyl-3-methyl-5-hydroxy-6-methoxy-1,4-benzoquinol methylase
COG2324	Predicted membrane protein
COG2440	Ferredoxin-like protein
COG2710	Nitrogenase molybdenum-iron protein, alpha and beta chains
COG3407	Mevalonate pyrophosphate decarboxylase
COG3425	3-hydroxy-3-methylglutaryl CoA synthase
COG3890	Phosphomevalonate kinase
COG4072	Uncharacterized protein conserved in archaea
COG4851	Protein involved in sex pheromone biosynthesis
COG5190	TFIIF-interacting CTD phosphatases, including NLI-interacting factor
COG5243	HRD ubiquitin ligase complex, ER membrane component
COG5272	Ubiquitin

Table S5.3. Sequences of oligonucleotide primers used in this study

Oligonucleotide	Sequence (5'-3')	Application
Disruption vectors		
0010509	GGAGGTGAATTCGTATGGACTTATACGTCAGAAAAGC	<i>TK0509</i> initial amplification
0020509	GGTGAAGGATTCCTCCTTATTATTCCTTGCTCTCC	<i>TK0509</i> initial amplification
0160509	AGCCCTATCAAGGTCGTCCAGCTGTACCATCCAATCACC TAACGTTTCGTAGGGACGCTT	<i>TK0509</i> inverse PCR to construct B plasmids
0170509	AAGCGTCCCTAGCGAACGTTAGGTGATTGGATGGTACA GCTGGACGACCTTGATAGGGCT	<i>TK0509</i> inverse PCR to construct B plasmids
0010513	GGAGGTGAATTCAGGGGATTATCGAGTATCTTCAGG	<i>TK0513</i> initial amplification
0020513	GGTGAAGGATTCCTTCTCTATCTTATGCCCCCTATCC	<i>TK0513</i> initial amplification
0160513	CGCCTGTTCTACTCCATAGGAGCGGTAGAACTTTTAA ACAATCCATCCTTTCGAAAGC	<i>TK0513</i> inverse PCR to construct B plasmids
0170513	GCTTTCGAAAGGATGGATTGTTTAAAGATTCTACCGCC TCCTATGGAGTAGAACAGGCG	<i>TK0513</i> inverse PCR to construct B plasmids
0010929	GGAGGTGAATTCGCGCATTCATCGAAGAGGACTATCG	<i>TK0929</i> initial amplification
0020929	GGTGAAGGATTCATGTCGTTTAAAGAAAGTGTGACC	<i>TK0929</i> initial amplification
0160929	AGACCTTTTGAAGTTGGTATGGTGATAGCTCACTCCCC CATTTCCCTAATCTTCCTGAC	<i>TK0929</i> inverse PCR to construct B plasmids
0170929	GTCAGGAAGATTAGGGAAATGGGGGAGTGAGCTATCA CCATACCAACTCCAAAAGGTCT	<i>TK0929</i> inverse PCR to construct B plasmids
0011088	GGAGGTGAATTCAGCGAGGAAGTACTATAACATCG	<i>TK1088</i> initial amplification
0021088	GGTGAAGGATTCACCCCTGATCTTGATCTTCAGTTAGG	<i>TK1088</i> initial amplification
0161088	AGCGCTTAAACCCCTGGAGGTGAGCGAATGATCAAGGA GCTTCTTTAACCAGTCTGAACCC	<i>TK1088</i> inverse PCR to construct B plasmids
0171088	GGGTTTCAGACTGGTTAAAGAAGCTCCTTGATCATTGCT CACCTCCAGGGGTTAGCGCT	<i>TK1088</i> inverse PCR to construct B plasmids
0011173	GGAGGTGAATTCAGCTCAAAGACCTCATAGTGTTCC	<i>TK1173</i> initial amplification
0021173	GGTGAAGGATTCAGGACTGCGAATCCTTTACTTTTCG	<i>TK1173</i> initial amplification
0161173	TTAATAAGCTTTAGCGGAAAAATTTAAAAATTGCCCTCA CTCTATCTTCTCTCTTGAG	<i>TK1173</i> inverse PCR to construct B plasmids
0171173	CTCAAGAGGAAGAAGATAGAGTGAGGGCAATTTTTAA ATTTTCCGCTAAAGCTTATTAA	<i>TK1173</i> inverse PCR to construct B plasmids
0011324	GGAGGTGAATTCCTCACCAGTTCTCTCTTTATTGC	<i>TK1324</i> initial amplification
0021324	GGTGAAGGATTCATACGACACCCGGTTACGTTGG	<i>TK1324</i> initial amplification
0161324	CTATTTTCATTTCTGCACCCCTAAGACAGGTATCACTC AATAGCTCCCAAATAGGAC	<i>TK1324</i> inverse PCR to construct B plasmids
0171324	GTCCTATTTGGGAGCTATTGAGTGATACCTGTCTTAGG GGTGCAGGAAAATGAAAATAG	<i>TK1324</i> inverse PCR to construct B plasmids
0011477	GGAGGTGAATTCCTCTAACAAGTCCTTCAAATACG	<i>TK1477</i> initial amplification
0021477	GGTGAAGGATTCGTTTACAAAGAGGCTATAGAGTTTCG	<i>TK1477</i> initial amplification
0161477	AAATATGGAAAAGGGCTTATCACACTTCGCATCTCCAT CACCCCAATGGATTTTAGCT	<i>TK1477</i> inverse PCR to construct B plasmids
0171477	AGCTAAAATCCATTGGGGGTGATGGAGATGCGAAAGT GTGATAAGCCCTTTTCCATATTT	<i>TK1477</i> inverse PCR to construct B plasmids
0011548	GGAGGTGAATTCGAAAGTCCCAATCTATTGTGTAGG	<i>TK1548</i> initial amplification
0021548	GGTGAAGGATTTCTATCATGGGAGTCTACATCTTTACC	<i>TK1548</i> initial amplification
0161548	CCTTGAGCAACTAAGGGCGAGGTGATAACTTTCCAATT CTCCCTCTTTCCCTTCTTAT	<i>TK1548</i> inverse PCR to construct B plasmids
0171548	ATAAGAAGGGGAAAAGAGGGAGAATTGGAAAGTTATC ACCTCGCCCTTAGTTGCTCAAGG	<i>TK1548</i> inverse PCR to construct B plasmids

0011732	GGAGGTGAATTTCAAAGAGAACAGTATTTCCCATCACC	<i>TK1732</i> initial amplification
0021732	GGTGAAGGATTTTCGGAAGGAGAGTTGAGAGACTGACC	<i>TK1732</i> initial amplification
0161732	AAAATAGGGAGCAACGAGCAGAACCTCATTCAACTGT CATCTGCGACCTCTCGTAGAA	<i>TK1732</i> inverse PCR to construct B plasmids
0171732	TTCTACGAGAGGGTCGCAGATGACAGTTGAATGAGGG TTCTGCTCGTTGCTCCCTATTTT	<i>TK1732</i> inverse PCR to construct B plasmids
0011836	GGAGGTGAATTTCCAACGAGCTTTAATAACCTTCTCG	<i>TK1836</i> initial amplification
0021836	GGTGAAGGATTTTCGCAAGTAAAAGCTCTCGTTGTCC	<i>TK1836</i> initial amplification
0161836	TTTTAAACCACCCCATCCGAGGGGTGAAGGGTGAGAA GATGGCCAGATGGAAGATGGGC	<i>TK1836</i> inverse PCR to construct B plasmids
0171836	GCCCATCTTCCATCTGGCCATCTTCTCACCTTCACCCCT CGGATGGGGTGGTTTTAAAA	<i>TK1836</i> inverse PCR to construct B plasmids
0011957	GGAGGTGAATTTAGCTGGTAAACCTGACATTATTGC	<i>TK1957</i> initial amplification
0021957	GGTGAAGGATTTCAAGAAGAGGAGAGAAAAGTGAAG G	<i>TK1957</i> initial amplification
0161957	CTTCCGACAGAATTTAGTTAGGTGCCCCATGGATATGG ATTTTGAGAAAGAACTGATTA	<i>TK1957</i> inverse PCR to construct B plasmids
0171957	TAATCAGTTCTTTCTCAAAATCCATATCCATAGGGGCAC CTAACTAAATTCTGTCGGAAG	<i>TK1957</i> inverse PCR to construct B plasmids
0012159	GGAGGTGAATTTTCGTAACCTATGACCGGAGCCTCTACG	<i>TK2159</i> initial amplification
0022159	GGTGAAGGATTTTCGAAGGTATCAAAAATCCCAAAGTGG	<i>TK2159</i> initial amplification
0162159	CCACTCTACAAATGTAAAGAAAAGTGCCGCTCTACCACC TGAAATGCCAAGGGTGAGGAA	<i>TK2159</i> inverse PCR to construct B plasmids
0172159	TTCCTCACCTTGGCATTTTCTAGGTGGTAGAGCGCACTT TTCTTTACATTTGTAGAGTGG	<i>TK2159</i> inverse PCR to construct B plasmids

Overexpression vectors

BG4881	CACCTCCGTGATATTATCTATTAC	vbb amplification (pSKS022)
BG4882	CGC <u>ACCGGT</u> CCGGACAACACCCCTTGAGCTTTG	vbb amplification (pSKS022)
BG4934	ATGCTATACAGATTGGTCTCCCATATCCC	<i>TK1173</i> amplification
BG4935	CGC <u>ACCGGT</u> CTACCTTCCAAGCGCCGCTGTCTC	<i>TK1173</i> amplification
BG4936	ATGCTGAGGGAGATACTCTCAC	<i>TK0929</i> amplification
BG4937	CGC <u>ACCGGT</u> TCACACCTCGCCCTCTCAAAC	<i>TK0929</i> amplification
BG4938	ATGAAGTACGATGTCCTCATAATCGGCGGC	<i>TK1836</i> amplification
BG4939	GCG <u>ACCGGT</u> TACTCCTCTGGAGGCCATCTTCCAT	<i>TK1836</i> amplification
BG4940	ATGAAAGTCATTGTCGAATAACGGGGGC	<i>TK0509</i> amplification
BG4941	GCG <u>ACCGGT</u> TCAAGGTCTCCAGCTGTAC	<i>TK0509</i> amplification
BG4942	ATGGCATCTGTTTCAGCTGTTATAAGG	<i>TK1324</i> amplification
BG4943	GCG <u>ACCGGT</u> CGCTGACGAGGGCTATTTTCAT	<i>TK1324</i> amplification
BG5196	ATGACCTGGAAGTACGACGTCGTTG	<i>TK1088</i> amplification
BG5197	GCG <u>ACCGGT</u> GAAGAAGGGGTTGAGACTGGTT	<i>TK1088</i> amplification
BG5200	ATGGAATTCAGGCAATTCATCGAG	<i>TK1957</i> amplification
BG5201	GCG <u>ACCGGT</u> GGCGTCTTGAACGTTATGACAG	<i>TK1957</i> amplification
BG5202	ATGGGCTCGAAGCTTGATGAG	<i>TK0513</i> amplification
BG5203	GCG <u>ACCGGT</u> TCCCGCACGTTAGTCTACAC	<i>TK0513</i> amplification
BG5204	ATGGAGCTCCACTTTGACATGACC	<i>TK1548</i> amplification
BG5205	GCG <u>ACCGGT</u> TGGGGAAAAGAGGAGAAATTG	<i>TK1548</i> amplification
BG5584	ATGACAGTTGAAGTCTCCGTG	<i>TK1732</i> amplification
BG5585	GCG <u>ACCGGT</u> AACGAGCAGAACCCCTCATTC	<i>TK1732</i> amplification
BG5586	ATGGAGATGGTGGTAAGGTATCCC	<i>TK1477</i> amplification
BG5587	GCG <u>ACCGGT</u> CTAGCGCCTGAAGACTATG	<i>TK1477</i> amplification
BG5588	ATGAGTAAAGGAGTTGTACATCC	<i>TK2159</i> amplification
BG5589	GCG <u>ACCGGT</u> GCCGCTCAGAGGTACCATTCGG	<i>TK2159</i> amplification

Diagnostic primers

BG4871	CATCCTCACAGGGACGTTTG	OExxxx diagnosis & sequencing (primer A in Figure S2b)
BG4873	ACAGACTTGAGGTGGAAAGG	OExxxx diagnosis & sequencing (primer D in Figure S2b)
BG4662	GCTTGGGTCATGGACAGGCTC	RT-PCR: <i>TK0308</i> (F)
BG4663	TCTTAACCTGCTCGGCGACC	RT-PCR: <i>TK0308</i> (R)
BG5570	TAGAAGGATAACCGCAGGTG	RT-PCR: <i>TK1477</i> (F)
BG5571	CGATGGAGTGCTCGTATTTG	RT-PCR: <i>TK1477</i> (R)
BG5572	TGGAAACTACCTGGCGTCTC	RT-PCR: <i>TK1836</i> (F)
BG5573	AGTCGGCTCCCTTCCTTACC	RT-PCR: <i>TK1836</i> (R)
BG5574	CTCACTGAACCCAGATAAGG	RT-PCR: <i>TK1324</i> (F)
BG5575	GATTCTCCCGTCCATTGAAG	RT-PCR: <i>TK1324</i> (R)
BG5576	AAGACGGCGGCTACTGCTTC	RT-PCR: <i>TK0513</i> (F)
BG5577	CGCCAAGACCAACGGCAAAC	RT-PCR: <i>TK0513</i> (R)
BG5578	TGGTGAAGCCGAAGTACAAG	RT-PCR: <i>TK1548</i> (F)
BG5579	AACGCTCCAATAACCGTGAC	RT-PCR: <i>TK1548</i> (R)
BG5580	GGTTGGGCGTTTCCTATGAG	RT-PCR: <i>TK1732</i> (F)
BG5581	TGCACCGGGTTCGATCATTC	RT-PCR: <i>TK1732</i> (R)
BG5723	GCTATCTCCAGCTCGCTAAC	RT-PCR: <i>TK0929</i> (F)
BG5724	CTCAAGAAGCTCCCGATTCC	RT-PCR: <i>TK0929</i> (R)
BG5725	GGATGCGGGCATAATCACAG	RT-PCR: <i>TK0509</i> (F)
BG5726	GCTTGGGCATGAGGTGTGTG	RT-PCR: <i>TK0509</i> (R)
BG5727	CGCTGTAGGCGATCTGATAG	RT-PCR: <i>TK1173</i> (F)
BG5728	AACATAGCCCTCGCATACGG	RT-PCR: <i>TK1173</i> (R)
BG5729	GCGGAGCGTAGGTGTTTATG	RT-PCR: <i>TK1088</i> (F)
BG5730	ACAAGGAAGTCAGCGGAGTG	RT-PCR: <i>TK1088</i> (R)
BG5733	ATGAGGAACGCAGAGTACAG	RT-PCR: <i>TK1957</i> (F)
BG5734	ATAGCCGTTGGTGAGATTGG	RT-PCR: <i>TK1957</i> (R)
BG5737	TCTCCTTGATGCCGATGTTC	RT-PCR: <i>TK2159</i> (F)
BG5738	AGATAGGCGACAACGAGTTC	RT-PCR: <i>TK2159</i> (R)

Chapter 6

Thesis summary
and general discussion

Summary

Hydrogen (H_2) production by microbial dark fermentation has attracted widespread attention as a promising alternative to fossil fuels. However, a major known drawback of this method is that H_2 production yields are still too low to make the process economically competitive. Therefore, to enhance the economics of the microbial dark fermentation it is necessary to improve the sugar-to- H_2 yield. Current sugar-to- H_2 yields by microbial dark fermentation vary from less than 1 to almost 4 mole of H_2 /mole of hexose consumed, the latter corresponding to only about a third of the theoretical yield of 12 mole. Nevertheless, the production of 4 H_2 /hexose is generally regarded as the upper limit that can be reached naturally by dark fermentation. This is due to both thermodynamic constraints and the fact that evolution has optimized fermentations to produce biomass (ATP) and not H_2 . To reach H_2 production yields beyond current limits, extra substrate oxidation steps need to be introduced into the catabolism of current top H_2 producers via genetic engineering approaches, while maintaining sufficient ATP synthesis by SLP for microbial survival. To show that this is possible, we selected the hyperthermophilic archaeon *Thermococcus kodakarensis*, which is genetically accessible and has been reported to ferment starch to acetate, CO_2 and H_2 , having a H_2 -production yield of 3.3 mole/mole of hexose (**chapter 1**).

Many substrate oxidation steps that can potentially be added to *T. kodakarensis* result in the generation of NADPH (e.g. TCA cycle or Wood-Ljungdahl pathway). Although H_2 formation from NADPH is thermodynamically difficult, the high optimal growth temperature of *T. kodakarensis* (85°C) should allow the conversion to proceed under physiological conditions. Moreover, in contrast to most organisms, *T. kodakarensis* naturally possesses an NADPH-dependent hydrogenase. However, this cytosolic hydrogenase is normally utilized for oxidizing H_2 rather than for H_2 production, but various studies have shown that these [NiFe]-hydrogenases are readily reversible in vitro. Increasing intracellular NADPH levels by adding additional substrate oxidation steps could therefore shift the catalytic direction of the reaction towards H_2 evolution.

Chapter 2 provides an overview of the major canonical and non-canonical reactions involved in the production and regeneration of NADPH in prokaryotes and discusses their key enzymes. In addition, an overview of how different enzymes have been applied to increase NADPH availability and thereby enhance productivity of many biotechnological products is provided. The reactions discussed were divided into I) those that are directly coupled to central carbon metabolism and II) those that are not. The enzymes that comprise the first group are the oxPPP enzymes glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase; isocitrate dehydrogenase of the TCA cycle; malic enzyme; and three enzymes involved in non-canonical NADPH-generating reactions, namely phosphorylating GAP dehydrogenase (NADP⁺-GAPDH), non-phosphorylating GAP dehydrogenase (GAPN), and glucose dehydrogenase. The enzymes in the second group are transhydrogenases (NADH:NADP⁺), ferredoxin:NADP⁺ oxidoreductase, hydrogenases (H₂:NADP⁺), and NAD(H) kinase, although the latter is a *de novo* NADP⁺/NADPH-synthesizing enzyme rather than an NADPH-regenerating enzyme. The overview showed that although various NADPH-generating enzymes have been successfully overexpressed to increase NADPH availability, no general strategy that guarantees improved NADPH supply exists. Cells usually contain various integrated, network-wide, redox balancing systems, such that engineering of NADPH-generating pathways can result in unexpected metabolic perturbations. Thus, strategies often need to be optimized individually, taking the culture conditions and genetic background of the strains into consideration.

Chapter 3 builds upon knowledge provided in chapter 2 and describes the successful introduction of bacterial-derived Pentose Phosphate Pathway (PPP) genes into the genome of the hyperthermophilic archaeon *Thermococcus kodakarensis* in an attempt to add extra reductant-generating steps. To force glucose catabolism through the introduced PPP, the endogenous ribulose monophosphate pathway (RuMP) was disrupted simultaneously. Two different strains were created and their growth- and fermentation profile and proteome, when cultured in starch containing medium, were compared to that of the parental strain. The two strains created, designated strain SKS001 and SKS002, were complemented with only the oxidative-PPP (oxPPP) enzymes glucose-6-phosphate dehydrogenase and 6-

phosphogluconate dehydrogenase or with the entire set of missing PPP enzymes, respectively. In-gel activity staining and proteomics confirmed the functional expression of the heterologous genes. Growth and fermentation profile analyses showed that introduction of merely the oxPPP enzymes is sufficient to replace RuMP as a source of pentose phosphates. H_2 production levels were significantly higher in both strains, especially in the early exponential phase where a 96% and 50% increase in the H_2/OD_{600} , and 42% and 33% increase in the H_2 /acetate ratio was observed for strains SKS001 and SKS002, respectively. Moreover, the elevated H_2 /acetate ratios in the early exponential phase suggested that NADPH was responsible for the increase in H_2 formation. Unexpectedly, significantly higher growth rates and final OD_{600} -values were observed for the mutant strains, which together with the higher specific H_2 production values (H_2/OD_{600}), resulted in a substantially higher overall H_2 productivity. The higher growth rates could be a result of an increased supply of NADPH, but improved growth of the parental strain upon addition of nucleosides and proteomics suggested that this could also be due to an enhanced supply of ribulose-5-phosphate.

In **chapter 4**, some puzzling and contradicting results observed during the engineering of the strains discussed in chapter 3 were further investigated. While genomic modifications are generally relatively easily obtained in *T. kodakarensis*, obtaining strain SKS001 and SKS002 proved rather difficult. Not only was it very difficult to obtain desired “double crossover mutants”, but it appeared that different chromosome versions co-existed within the same culture. Although the latter observation may potentially be explained by cross-contamination, the difficulty in obtaining homozygous recombinant cultures, as well as the repeated observation of this phenomenon among different engineering attempts, suggested that *T. kodakarensis* may possess multiple copies of its chromosome. This prompted us to investigate the chromosome copy number of *T. kodakarensis* KOD1. In this chapter, we clearly showed that *T. kodakarensis* is indeed polyploid and that its chromosome copy number varies between 7 and 19 copies, depending on the growth phase. The finding of polyploidy in *T. kodakarensis* explains the observed difficulty in obtaining homozygous recombinant cultures of strain SKS001 and SKS002, as it entails that the desired strain is only

acquired when the modification is incorporated in all chromosome copies. Moreover, an apparent correlation between the presence of histones and polyploidy in Archaea was observed. The histones may assist in efficient DNA packaging which is required for fitting multiple chromosomes in a single cell.

As shown in **Table 1.1**, hyperthermophilic bacteria and archaea are better H₂ producers than thermophilic or mesophilic ones. Therefore, it is interesting to understand how these hyperthermophilic microbes are able to thrive at such high temperatures, so that these characteristics might be transferred to other H₂ producers as well. One of the factors thought to contribute to this thermal resistance in archaea is the incorporation of tetraether (GDGT) lipids in their membrane. However, it is currently unclear how these lipids are synthesized. In **Chapter 5**, we set out to identify the enzyme(s) involved in the synthesis of tetraether lipids in *Thermococcus kodakarensis*, more specifically the one(s) responsible for the condensation of two C₂₀ isoprenyl chains, i.e. the putative tetraether synthase(s). To do this, a list of genes that are potentially involved in tetraether lipid synthesis was composed, and prioritized based on a couple of characteristics assumed to be important for a tetraether synthase. This ultimately resulted in the selection of 12 genes, which were targeted for overexpression and disruption. Subsequent diether/tetraether lipid analysis of the mutant strains strongly suggested that at least three of these genes are indeed involved in the tetraether synthesis pathway. Moreover, disruption of various candidate genes turned out to be non-viable, which is an interesting result as disruption of a tetraether synthase may very well be lethal. However, though interesting results have been obtained, the data did not allow for the establishment of the exact functions of the candidate genes. Future research is thus needed to get a better understanding of the tetraether synthesis pathway and to identify the tetraether synthase(s).

General discussion

Hydrogen production beyond current limits, continued

To enhance the economy of the microbial dark fermentation it is necessary to improve the sugar-to-H₂ yield beyond the current limit of 4 mole H₂/hexose (i.e. the “Thauer limit”). To show that this is possible, we introduced extra substrate oxidation (or reducing-equivalent generating) steps in the hyperthermophilic archaeon *Thermococcus kodakarensis*, creating strain SKS001 and SKS002. As shown in **chapter 3**, H₂ production levels were significantly higher in both strains. Moreover, the elevated H₂/acetate ratios in the early exponential phase (~1.9 versus ~2.7/2.5 for the parental strain TS559 and the mutants SKS001/SKS002, respectively) suggested that NADPH was indeed responsible for (part of) the increase in H₂ formation. Although these results are very promising, the obtained increases in the H₂/acetate ratios are not large enough to be absolutely certain that the H₂/hexose yield has broken the “Thauer limit”, especially since the parental strain TS559 also reached a H₂/acetate ratio of ~2.5 when growth approached the stationary phase. Therefore, to obtain absolute prove, we need to reach either significantly higher H₂/acetate yields or, ideally, determine the exact amount of converted sugar to ultimately obtain H₂/glucose yields. This might be realized in several ways, which is discussed in the following sections.

Improving the H₂/acetate ratio further

Bioenergetic analyses showed that a H₂/acetate ratio of 8 should be possible by microbial dark fermentation as sufficient ATP might be obtained by SLP under non-standard conditions, especially at elevated temperatures and when partial H₂ pressures can be kept low (**Table 3.1**). This implies that there is still sufficient room for improving the H₂/acetate ratio further. To attempt this, several options are available, which can be divided into those that require extra engineering steps and those that do not. For example, one might try to reach higher H₂/acetate ratios by simply repeating the fermentation analysis performed in **chapter 3**, but by changing one important feature. The experiments described in **chapter 3** were performed in batch, without any control over the fermentation characteristics. This implies that partial H₂ pressures rapidly increased along with increasing cell densities, which

from a H_2 -production point of view is thermodynamically unfavorable. Especially, in the case of H_2 production from NAD(P)H (**Figure 6.1**). Thus, by repeating the fermentation analysis and implementing control over the partial H_2 pressures, to keep the H_2 concentration experienced by the cells as low as possible, H_2 /acetate (and H_2 /glucose) ratios may be increased further. This can be realized by using either batch or continuous cultures, in which evolved gas metabolites are continuously flushed out and measured, using N_2 and continuous gas chromatography, respectively.

In addition, one might try to increase the H_2 /acetate ratio further by increasing the fraction of the sugar catabolism directed through the introduced Pentose Phosphate Pathway (PPP). This might be realized, without the need of extra engineering steps, by culturing the cells in a pentose-free medium. The ASW-YT medium used in this study is a very rich medium containing large amounts of tryptone and yeast extract. It is thus likely that the medium contains considerable amounts of pentoses or pentose-containing molecules, such as AMP, CMP, UMP or nucleosides. These molecules may be taken up by the cell [592], which decreases the need for novel pentose biosynthesis via the introduced PPP. Therefore, by removing pentoses and pentose-containing molecules from the medium, cells might be forced to direct a larger fraction of their sugar catabolism via the PPP.

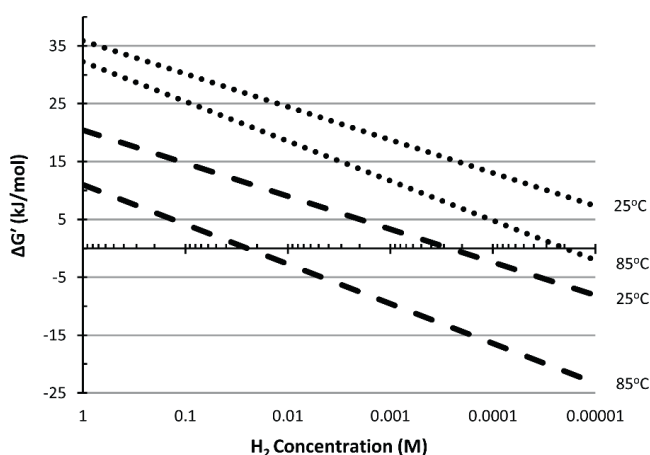


Figure 6.1 Effect of H_2 concentration and temperature on the Gibbs energy of H_2 formation from NAD(P)H (dotted line) and reduced ferredoxin (dashed line) at 25°C and 85°C. The figure was adapted from [48]

Similarly, one could go even further and attempt to direct sugar catabolism completely through the introduced PPP. Theoretically, this should result in a H_2 /acetate (and H_2 /glucose) ratio of 8 (**Table 3.1**). To realize this, hexoses need to be prevented to follow the modified EM pathway naturally present in *T. kodakarensis*, while keeping the phosphoglucose isomerase (PGI; *TK1111*) and the carbon-three part of the EM-pathway intact to enable conversion of the PPP products fructose-6-phosphate and glyceraldehyde-3-phosphate. This can be done by disrupting the ADP-dependent phosphofructokinase (PFK; *TK0376*) which, in contrast to the examples given above, thus requires some additional metabolic engineering (**Figure 6.2**). However, since the EM-pathway is normally required for survival of *T. kodakarensis* under glycolytic conditions, disruption of PFK might be extremely challenging or even lethal, despite the introduction of the PPP. Therefore, it would be desirable to additionally gain control over the expression of PFK-encoding gene *TK0376*. To do this, its natural promoter should be exchanged for a controllable one that can be either switched on/off via specific activators or repressors, or that allows transcriptional tuning via specific inducers in a dose-response fashion [593-596]. Unfortunately, to the best of our knowledge, such controllable promoters are not (yet) available for *T. kodakarensis*.

Interestingly, the transcription of the PFK-encoding gene, as well as that of other glycolytic and gluconeogenic enzyme-encoding genes, is naturally regulated in a similar manner. Both glycolytic and gluconeogenic promoters contain a palindromic sequence, called the Thermococcales-Glycolytic Motif (TGM), that is present in many sugar-metabolizing Thermococcales [597]. This motif is recognized by a global transcriptional regulator, called the Thermococcales-Glycolytic regulator (Tgr), that specifically binds the TGM and that is only released when maltotriose is present. In glycolytic promoters, the TGM is positioned downstream of the Transcription factor B-responsive element (BRE)/TATA sequence, which means that binding of the Tgr to the TGM (i.e. gluconeogenic conditions) structurally blocks RNA polymerase recruitment and thus inhibits transcription. When maltodextrins are available in the medium (i.e. glycolytic conditions), the intermediate maltotriose binds to the Tgr protein, leading to displacement of Tgr from TGM, enabling RNA polymerase recruitment and hence transcription. In gluconeogenic promoters, the TGM is positioned upstream of the

BRE/TATA sequence and binding of the Tgr to the TGM (i.e. gluconeogenic conditions) is thought to stimulate RNA polymerase recruitment and hence transcription. Displacement of the Tgr from the TGM upon the presence of maltotriose (i.e. glycolytic conditions), diminishes this activating effect and hence inhibits transcription. Thus, Tgr simultaneously controls, in response to sugar availability, both glycolytic and gluconeogenic metabolism in *T. kodakarensis* via its direct binding to the TGM [88], i.e. transcription of glycolytic enzyme-encoding genes is “switched off” under gluconeogenic conditions and “switched on” under glycolytic conditions, while transcription of gluconeogenic enzyme-encoding genes is “switched off” under glycolytic conditions and “switched on” under gluconeogenic conditions. This suggests that it should be possible to obtain a PFK-disruption mutant under gluconeogenic conditions.

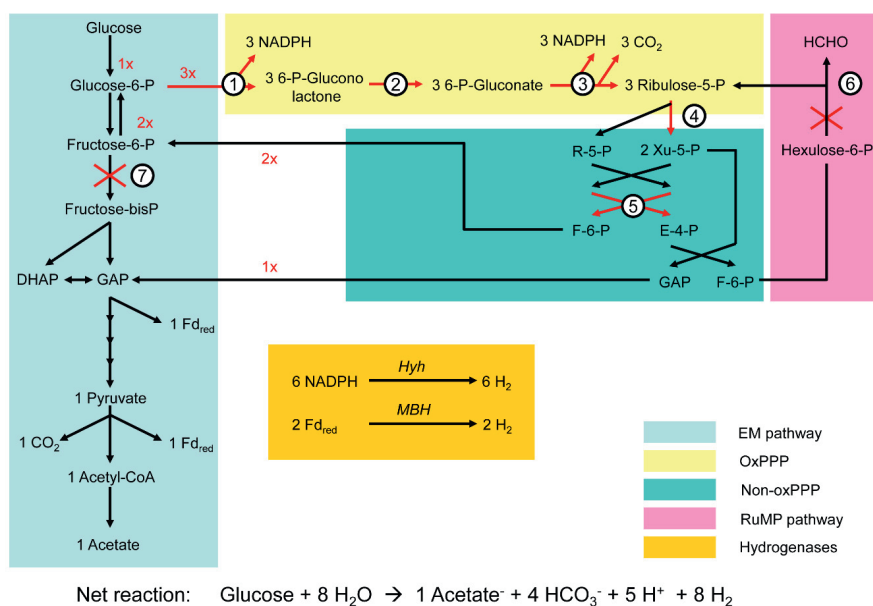


Figure 6.2 Schematic representation of the proposed hydrogen production metabolism. The enzymes already introduced (red arrows) and disrupted in strain SKS002 are numbered 1-6. The proposed additional PFK-disruption is shown as number 7. The numbers represent: 1, glucose-6-phosphate dehydrogenase; 2, 6-phosphogluconolactonase; 3, 6-phosphogluconate dehydrogenase; 4, ribulose-phosphate 3-epimerase; 5, transaldolase; 6, 3-hexulose-6-phosphate synthase (hps) - 6-phospho-3-hexuloisomerase (phi) fusion enzyme; 7, phosphofructokinase

To do this, a similar engineering strategy as used for the creation of strain SKS001 and SKS002 can be applied (**Figure 6.3**). Especially, since the selection/counter-selection procedure utilized was designed to be performed under gluconeogenic conditions, i.e. on ASW-YT plates supplemented with polysulfide [60, 61, 486]. To create a PFK-disruption, strain SKS002 should be used as a parent as a complete PPP is required to enable growth under glycolytic conditions. Transforming strain SKS002 with the disruption vector shown in **Figure 6.3**, should result in agmatine prototrophs that can easily be selected for on ASW-YT plates containing polysulfide [60, 61, 486]. These agmatine prototrophs are the result of separate integration events at either the up- or downstream flanking region, resulting in the incorporation of the entire vector to the chromosome of *T. kodakarensis*. Next, an internal recombination event can be stimulated by plating the transformants on plates containing 6MP and agmatine, to either restore the original genome or produce the desired PFK-disruption strain (**Figure 6.3**).

Once a PFK-disruption mutant has been obtained, the strain can be tested for growth under glycolytic conditions by using the starch-containing ASW-YT-SA2 medium. If successful, the fermentation analysis, as performed on strain SKS001 and SKS002, should be repeated using the newly developed mutant strain SKS003. Ideally, the analysis should be performed in such a way that partial H_2 pressures can be controlled, to keep the H_2 concentration experienced by the cells as low as possible (see above). All of this may ultimately result in higher H_2 and alanine levels as observed for strain SKS001 and SKS002. In addition to H_2 and alanine, the fermentation analysis may reveal other unwanted reduced end products (e.g. formate or lactate), which could suggest additional gene-disruption targets. Moreover, one may attempt to reveal bottlenecks in the metabolism by using metabolome measurements and ^{13}C -metabolic flux analysis [598, 599]. Finally, as growth of the engineered SKS003 strain is dependent on a functional PPP, a random optimization approach can be conducted in chemostats to improve the fermentation kinetics under sugar limited conditions. Such an evolutionary engineering approach has for example successfully been applied for the optimization of xylose utilization in yeast [600].

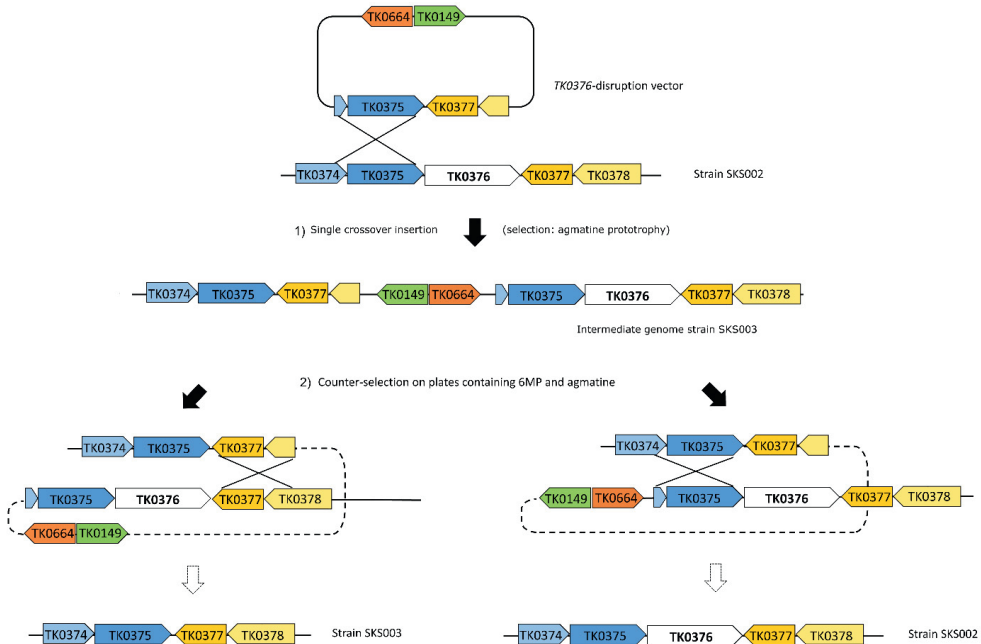


Figure 6.3 Disruption strategy for the PFK-encoding gene *TK0376* of *T. kodakarensis* strain SKS002. The initial single crossover insertion (step 1) can occur at either the upstream homologous region (in blue, shown above) or at the downstream homologous region (in yellow, not shown). Both integration events yields agmatine-prototrophic intermediate strains, that contain direct repeats flanking the target locus. Dependent on recombination responsible for excision upon counter-selection with 6MP (step 2), either the original SKS002 genome can be restored or the desired deletion genome SKS003 can be generated.

Determining the exact H_2 /glucose ratio

Another way to determine if the H_2 /glucose yield has breached the “Thauer limit”, is by determining the exact amount of converted sugar. Unfortunately, the outcome of the glucose consumption analysis performed in **chapter 3**, with soluble starch as C-source, was not consistently accurate enough to allow precise calculation of H_2 /glucose yields. Therefore, It would be desirable to repeat the fermentation analysis and to replace starch with sugars that can be quantified by HPLC directly, such as glucose, maltose or maltotriose. Unfortunately, however, *T. kodakarensis* is not able to grow on mono- and disaccharides as their respective transporters are missing [59]. Growth on maltotriose, on the other hand, should be possible, as has also been shown for the wild-type strain KOD1 [72]. However,

when we tried to cultivate strain SKS001 and SKS002 in ASW-YT medium supplemented with maltotriose, no growth was observed. Moreover, growth was also not observed in a control inoculated with the parent strain TS559, but was observed in a control containing the wild-type strain KOD1 (data not shown). These results show that the wild-type strain KOD1 is indeed able to grow on maltotriose, but that this ability has somehow been lost by strain TS559, and hence strain SKS001 and SKS002.

Therefore, it would be desirable to either restore the ability to grow on maltotriose in strain SKS001 and SKS002, or to introduce transporters that enable growth on mono- and/or disaccharides. To restore the ability to grow on maltotriose, it is needed to establish the cause of the perturbation. At first, an unwanted mutation in the maltooligo/maltodextrin transporter (*MalEFGK*, *TK1771-1773*, *TK1775*, **Figure 1.3**) seems likely as this might disable maltotriose uptake. However, since this transporter is the only sugar transporter identified in *T. kodakarensis* [59, 72] and growth on starch is still possible, a disrupted maltooligo/maltodextrin transporter seems unlikely. Alternatively, a mutation in the amylopullulanase (*TK1774*) might be responsible as a previously developed *TK1774*-disruption strain retained the ability to grow on larger maltooligosaccharides, but was unable to grow on maltotriose [72]. To establish if undesired mutations in the genes described above are indeed responsible for the loss of growth in maltotriose-containing medium, the corresponding chromosomal regions could be PCR amplified (using a high fidelity polymerase) and analyzed by sequencing.

Alternatively, a transport system that enables the uptake of mono- and/or disaccharides and that is functional at 85°C can be introduced into *T. kodakarensis*. Although most enzyme complexes are not thermostable, such a transporter system can be found in the closely related species *T. sibiricus* (optimal growth temperature of 78°C), *T. litoralis* (85-88°C), and *Pyrococcus furiosus* (100°C). All of these species contain a maltose/trehalose-specific transporter that is closely related to the maltooligo/maltodextrin transporter present in *T. kodakarensis* [601-604]. Like the maltooligo/maltodextrin transporter, the maltose/trehalose transporter is encoded by four genes designated *MalEFGK*. However, where the *malE*-gene of the maltooligo/maltodextrin transporter encodes a periplasmic

binding protein specific for maltooligosaccharides longer than a trimer, the one of the maltose/trehalose transporter encodes a binding protein specific for maltose and trehalose. Therefore, replacing the natural *MalE*-gene of *T. kodakarensis* with the one encoding the maltose/trehalose-specific binding protein, might enable growth on these quantifiable sugars. However, as this will result in the loss of the ability to grow on maltooligosaccharides longer than a trimer, it might be more desirable to just introduce all four of the heterologous *MalEFGK* genes. This will most likely result in a *T. kodakarensis* strain that, in addition to growth on longer maltooligosaccharides, is also able to grow on the disaccharides maltose and trehalose.

Another, more radical, approach to determine if the proposed hydrogen production pathway (**Figure 6.2**) is able to break the “Thauer limit”, is by starting all over and selecting a H₂-producing hyperthermophilic organism that is already able to grow on mono- and/or disaccharides. At the start of this thesis, *T. kodakarensis* was the only H₂-producing hyperthermophilic organism that was genetically accessible. However, today, other (hyper)thermophilic organisms have become genetically accessible as well. For example, the closely related and excellent H₂-producer *Pyrococcus furiosus* now has an efficient toolbox and has become one of the model organisms in archaeal research [442, 605]. Like *T. kodakarensis*, *P. furiosus* already produces H₂ near the current limit of 4 mole/hexose (**Table 1.1**), contains NADPH-dependent hydrogenases (in contrast to *T. kodakarensis*, *P. furiosus* contains not one but two NADPH-dependent hydrogenases; SHI and SHII), and is missing many of the classical PPP-enzymes [29, 78, 103]. However, in contrast to *T. kodakarensis*, *P. furiosus* is naturally able to utilize the disaccharides maltose and trehalose. In addition, *P. furiosus* has an even higher optimal growth temperature (i.e. 100°C instead of 85°C), which, provided that the heterologous PPP enzymes are still stable, might be beneficial for H₂-production (**Figure 6.1**).

Similarly, the hyperthermophilic bacterium *Thermotoga maritima* (optimum growth temperature 80°C) has recently been reported to be transformable, though the efficiency was extremely low (~0.25 transformants per µg of DNA) [606]. Moreover, the *Thermotoga* species RQ2 and RQ7, which are both closely related to *T. maritima*, have been reported to

be transformable as well and strain RQ2 has even been used for the heterologous expression of cellulases from *Caldicellulosiruptor saccharolyticus* [607, 608]. This is interesting as *T. maritima* is known as one of the best H₂-producers, capable of producing 4 mole H₂ per mole of glucose (**Table 1.1**), and naturally contains the PPP, but lacks a NADPH-dependent hydrogenase. Thus, Introducing the NADPH-dependent hydrogenase from either *T. kodakarensis* or *P. furiosus*, might enable H₂ production above current limits.

When successful, the various options described above may result in a H₂/glucose yield of maximally 8 (**Table 3.1**). This means that in addition to the 8 moles of H₂ and 4 moles of CO₂, 1 mole of acetate is still produced per glucose. As explained in **chapter 1 and 3**, further oxidation of this fermentation-end product to CO₂ and H₂ by dark fermentation is not possible. Especially, as this conversion requires a substantial input of energy (**Eq. 1.7**). However, to increase the H₂/glucose yield to its theoretical limit of 12, and thus to extract the remaining 4 moles of H₂ that is still “locked” in the acetate, a subsequent photo-fermentation or microbial electrolysis may be performed [11, 42, 53]. In these methods, the energy barrier associated with the conversion of acetate (**Eq. 1.7**) can be overcome by utilizing the energy from solar light or by applying an external potential, respectively. Nevertheless, although technically feasible, such combined systems are too expensive from an economical point of view.

Confirming the physiological role of non-phosphorylating glyceraldehyde 3-phosphate dehydrogenase (GAPN)

As discussed in **chapter 1**, *T. kodakarensis* contains a modified Embden-Meyerhof (EM) pathway (**Figure 1.3, box A and B**). This modified EM pathway is also present in other *Thermococcales* and differs from the classical EM pathway present in, for example, *Escherichia coli* in some key aspects. One of these aspects is the mode of conversion of glyceraldehyde-3-phosphate (GAP) to 3-phosphoglycerate (3-PG). In the classical EM pathway, GAP is converted to 3-PG via 1,3-bisphosphoglycerate by the consecutive functions of phosphorylating NAD⁺-dependent GAP dehydrogenase (GAPDH) and ATP-generating 3-PG kinase (PGK) (**Figure 6.4**). Together, these two enzymes play a crucial role in conserving

energy in the classical glycolysis by generating NADH and ATP. Although GAPDH and PGK homologs are also present in *T. kodakarensis*, they are not involved in glycolysis. Instead, the conversion of GAP to 3-PG is catalyzed by both ferredoxin-dependent GAP:ferredoxin oxidoreductase (GAPOR) and NADP⁺-dependent non-phosphorylating glyceraldehyde 3-phosphate dehydrogenase (GAPN), in an irreversible, phosphate-independent, one-step reaction (**Figure 6.4**) [77]. This is an important difference as the direct oxidation of GAP to 3-PG by GAPOR or GAPN omits ATP production via substrate level phosphorylation by PGK.

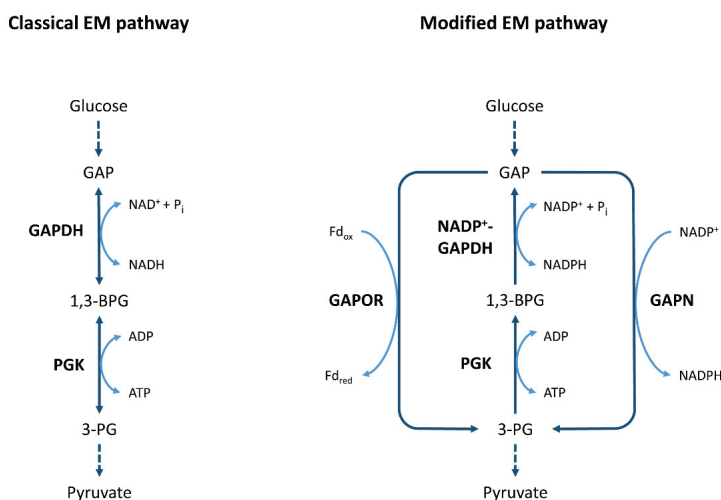


Figure 6.4 Schematic overview of the conversion of glyceraldehyde-3-phosphate (GAP) to 3-phosphoglycerate (3-PG), in the classical EM pathway (left) and the modified EM pathway of *T. kodakarensis* (right). Abbreviations for metabolites (normal text) and enzymes (in bold): 1,3-BPG; 1,3-bisphosphoglycerate, GAPDH; GAP dehydrogenase, PGK; 3-PGA kinase, GAPOR; GAP:ferredoxin oxidoreductase, and GAPN; non-phosphorylating glyceraldehyde 3-phosphate dehydrogenase

Surprisingly, both GAPOR and GAPN have been shown by gene disruption experiments to be essential under glycolytic conditions [77]. These results were rather unexpected as both catalyze the conversion of GAP to 3-PG, which suggest that they should be able to complement each other in terms of carbon conversion. Moreover, reduced ferredoxin (Fd_{red}) associated with the GAPOR reaction can also be generated by pyruvate:ferredoxin

oxidoreductase (POR) (**Figure 1.3**). The authors discussed that a feasible explanation for the necessity of GAPOR under glycolytic conditions may have to do with the metabolic flux through both enzymes. The capacity of GAPOR in terms of metabolic flux can be considered relatively high as its co-substrate Fd_{ox} is readily regenerated by the function of the membrane-bound hydrogenase (**Figure 1.3**) [30, 82, 83]. In contrast, however, the regeneration rate of $NADP^+$ by anabolic processes is expected to be relatively low and might therefore not allow a sufficient turnover of GAPN to complement the absence of GAPOR.

At present, the reason why GAPN is essential under glycolytic conditions is most likely due to its role in generating NADPH [77]. *T. kodakarensis* does not contain many sites for NADPH generation, especially not under glycolytic growth conditions. For example, the wild-type strain does not possess the oxidative PPP (**chapter 2 and 3**) and the transcription of the glutamate dehydrogenase (GDH) gene is strongly repressed under glycolytic conditions [88]. Alternatively, the authors suggested that the disruption of GAPN may lead to an accumulation of intracellular GAP, which in turn may lead to a shutdown of the GAPOR route. This alternative explanation to why GAPN is essential under glycolytic conditions is based on the observation that the GAPOR of *P. furiosus* is significantly inhibited by GAP concentrations of 0.5 mM or higher [77, 609].

Interestingly, the strains created in **chapter 3**, i.e. SKS001 and SKS002, can be used to clarify why GAPN is essential under glycolytic conditions. This is because the introduction of the oxidative PPP in these strains provides two additional NADPH generation sites (i.e. G6PDH and 6PGDH) which, in case the first theory is correct (essential role in NADPH generation), should enable a GAPN-disruption strain to survive under glycolytic conditions. Thus, a viable GAPN-disruption mutant, in an SKS001 or SKS002 background, would confirm GAPN as an essential NADPH generator, whereas a non-viable mutant would suggest that the alternative theory (inhibitory effect of GAP on GAPOR) might be true. To test this, various approaches are possible, but the most straight-forward one would be to use either strain SKS001 or SKS002 as a parent and to slightly adapt the previously developed pUD2- Δ gapN disruption vector to make the selection procedure compatible with the genetic background of these strains; selection of recombinants after transformation with the pUD2- Δ gapN disruption

vector is based on the integration of the PyrF marker, that confers uracil prototrophy [62, 63, 77], whereas the genetic background of SKS001/SKS002 only allows for selection based on tryptophan and/or agmatine prototrophy, and not on uracil prototrophy [60, 61, 486].

To make the pUD2-ΔgapN disruption vector compatible with the strains developed in **chapter 3**, the PyrF marker can be replaced by the *TK0149-TK0664* selection/counter-selection cassette (selection based on agmatine prototrophy and counterselection based on 6-methyl purine resistance), which would allow for a GAPN-disruption strategy similar to the one shown in **Figure 6.3**. Alternatively, a disruption strategy based on tryptophan selection can be applied. For practical reasons not further discussed here, it was decided to attempt to obtain a GAPN-disruption strain using the second strategy. To do this, the disruption vector was adapted such that it directly targets the GAPN gene locus, so that its integration into the genome of strain SKS001/SKS002 will result in a GAPN disruption. Unfortunately, though the procedure was performed under gluconeogenic conditions, no tryptophan prototrophs were obtained, even not after several transformation attempts. Therefore, a rather different approach was attempted as well, in which the strategy was basically turned around: instead of starting with the SKS001/SKS002 strain and trying to disrupt GAPN, it was decided to start with the previously developed GAPN disruption mutant [77] and to try to introduce the PPP. Like done in **chapter 3**, two different vectors were created that should ultimately result in the simultaneous disruption of the ribulose monophosphate pathway (RuMP) and the introduction of either the oxPPP enzymes glucose-6-phosphate dehydrogenase (G6PDH), 6-phosphogluconolactonase (6PGLase), and 6-phosphogluconate dehydrogenase (6PGDH), or of all five of the missing PPP enzymes. The difference with the vectors created in **chapter 3** is that the selection of transformants is based on uracil prototrophy rather than agmatine prototrophy, and that the first vector introduces all three of the oxPPP enzymes instead of just G6PDH and 6PGDH. In theory, the simultaneous disruption of the RuMP and the introduction of just the oxPPP should be enough to complement the loss of GAPN. Nevertheless, in accordance with **chapter 3**, it was decided to include the introduction of the entire PPP as well. Unfortunately, however, no desired mutants have been obtained yet.

Polyploidy in *Thermococcus kodakarensis*

Correlation does not imply causation

In **chapter 4**, we showed that *T. kodakarensis* is polyploid and that its chromosome copy number is growth-phase-dependent. (i.e. a chromosome copy number as high as 19.4 ± 4.9 was observed in the early exponential phase, whereas the ploidy level in the stationary phase was only 7.5 ± 2.8). Moreover, we observed an apparent correlation between the presence of histones and polyploidy in archaea (**Table 4.1**). Though it is tempting to speculate that histones play an important role in enabling polyploidy, statistics tells us that correlation does not necessarily imply causation. Therefore, to confirm a causal relationship, further research is needed.

To test if there is an actual causal relationship between the presence of histones and polyploidy in archaea, more species need to be tested for their ploidy level. Especially, those that diverge from the apparent dichotomy between Euryarchaeota/Crenarchaeota: all Euryarchaeal species tested to date encode archaeal histone homologs and are polyploid, while all Crenarchaeal species tested do not encode archaeal histone homologs and were found to be monoploid. In this respect, **Table 4.1** clearly reveals several candidate species. For instance, *Thermoplasma acidophilum*, *T. volcanium* and *Ferroplasma acidarmanus*, are all Euryarchaeotes, and therefore believed to be polyploid, but none of these species encodes archaeal histones. It should be mentioned, however, that these species do encode homologs of the bacterial histone-like proteins (HU-proteins) that have been proposed to have a similar function [525]. Similarly, the Crenarchaeota *Thermofilum pendens*, *Caldivirga maquilensis* and *Vulcanisaeta distributa* are interesting as they are among the few Crenarchaeota known to encode archaeal histones. Moreover, determining the chromosome copy number of species belonging to the archaeal phyla Thaumarchaeota, Korarchaeota and Nanoarchaeota, which are all known to encode histones, may be valuable as well in strengthening the polyploidy/histone dichotomy hypothesis.

Another, perhaps more direct way to test if there is a causal relationship between the presence of histones and polyploidy in archaea is to disrupt the archaeal histone encoding gene(s) and to determine its effect on the ploidy level. However, since histones can play an

important role in DNA packaging and the regulation of gene expression, disruption of histone encoding genes might be expected to perturb cell growth and/or metabolism. Yet, examples of Euryarchaeal species in which histone encoding genes were successfully deleted have been reported. For instance, *Methanococcus voltae* and *T. kodakarensis* both contain two archaeal histones and though the respective double mutants could not be obtained, single disruption mutants for each of the genes have shown to be viable [529, 610]. Moreover, a viable histone disruption strain of *Methanosarcina mazei* Gö1 has been reported as well [611]. The latter is of special interest as this species contains only one histone encoding gene and the disruption strain thus lacks any histone proteins. Comparing the ploidy level of this mutant strain to that of the wild-type would thus be very interesting.

Although the polyploidy/histone correlation may hold for archaea, it is for sure not a general trait. For example, polyploidy has been observed in many bacteria as well, while this domain is well known not to encode any histones. Instead, bacteria generally encode different DNA-binding proteins for DNA packaging, such as HU, IHF, Fis or Lrp, some of which have also been found in Archaea [524, 525, 529]. Thus, other systems or proteins may exist that contribute to polyploidy. Moreover, it might be that it is not histones per se, but rather their function (i.e. condensating DNA by wrapping) that enables polyploidy in archaea. If so, a correlation between polyploidy and other wrapping proteins might exist in bacteria and/or archaea as well.

Another interesting observation done in **chapter 4** was that the chromosome copy number of *T. kodakarensis* is growth-phase-dependent. The observation of a growth-phase-dependent ploidy level has been reported for many other prokaryotes as well, including the majority of the investigated Euryarchaeal species [503, 504, 506, 514]. Unfortunately, no common pattern has emerged from the characterized prokaryotic species. Where some species have a higher chromosomal copy number in the exponential phase (including *T. kodakarensis*), others have a higher chromosomal copy number in the stationary phase [506, 515]. Nevertheless, the observation of a growth-phase-dependent ploidy level in itself is quite interesting and one may wonder what is causing this effect.

A possible explanation for the observation that the ploidy level of *T. kodakarensis* is decreasing towards the stationary phase might be that chromosomal replication cannot keep up with cell division. If this is the case, it would suggest that there is a phase of cell division independent chromosomal replication, to prepare the cells for rapid growth. This would imply that the ploidy levels observed in the exponential phase are an overestimation of the true ploidy level of *T. kodakarensis* (of non-growing cells) and that the ones observed in the stationary phase are more accurate. Moreover, it also implies that, through chromosomal replication and cell division, ploidy levels are somehow regulated. In this regard, it is interesting that in species that contain more than one histone encoding gene, differences in their expression levels have been reported to correlate with differences in laboratory growth rates and/or growth phases [610, 612, 613]. Nevertheless, although interesting, there is currently no data in support of any of the above and other explanations may exist as well.

Potential solutions for polyploidy based engineering problems in *T. kodakarensis*

As discussed in **chapter 4**, the existence of multiple chromosome copies can readily explain the difficulty observed in obtaining strain SKS001 and SKS002, as it entails that the desired strain is only obtained when the modification is incorporated in all chromosome copies. The co-existence of different chromosome variants, i.e. double crossover recombinants, single crossover integrations or wild type chromosomes, in the majority of the transformants obtained in this chapter may have indicated that the desired modification is harmful and possibly lethal. Especially, since genetic modifications are normally relatively easily obtained. However, since desired homozygous recombinants were eventually obtained by successive cycles of re-plating and screening, this is not necessarily always the case. Nevertheless, acquiring homozygous recombinants via this approach can be rather laborious and time-consuming. The knowledge that *T. kodakarensis* is polyploid may therefore be quite valuable as it allows for the development of more efficient approaches. For example, the observation that the chromosome copy number of *T. kodakarensis* is growth-phase dependent, suggests that re-plating cells that are in the stationary phase might be more effective than the typically used exponentially growing cells. Moreover, using cells that are in the late

stationary phase might be even better as it is likely that their copy number is even lower than the one described in **chapter 4** (cells in this chapter were harvested in the early stationary phase).

Similarly, other approaches resulting in a reduced ploidy level might also be applied. One could for example try to lower the chromosome copy number by reducing the growth rate by using suboptimal growth conditions such as a reduced incubation temperature or the usage of suboptimal growth media (see effect of doubling time in **Table 4.1**). Interestingly, we indeed observed that genome modifications involving two independent homologous recombinations were often more successful when using plates containing defined ASW-AA medium rather than nutrient-rich ASW-YT medium. In addition, the recent discovery that *Haloferax volcanii* uses DNA as a phosphate storage polymer and that its ploidy level is severely reduced under phosphate-limiting conditions might also be interesting for *T. kodakarensis* [522]. Even if *T. kodakarensis* does not use DNA as phosphate storage, culturing under phosphate-limited conditions is likely to be effective.

All of these alternative approaches aiming at improving double recombination efficiency have in common that they aim at a reduction of the ploidy level of the cell. However, strategies for improving the success rate that are independent of the reduction of ploidy levels might also be possible. One could for example apply a selection based on antibiotic resistance. To promote the presence of the desired genome only, one could place the antibiotic resistance marker under the control of a weak promoter. In theory, gradually increasing the concentration of the antibiotic in the selection medium entails that only those cells that increase the copy number of the resistance marker accordingly, are able to survive. The availability of thermostable antibiotics, such as mevinolin, is, however, crucial for this strategy [60, 486].

Perhaps the most interesting strategy would be one that eliminates the need for optimization procedures, such as the recently developed CRISPR-Cas9 system. The CRISPR-Cas system is regarded as an immune system of bacteria and archaea. It has, however, been shown that it can be easily adapted and applied as an efficient, marker-independent,

genome editing system. CRISPR assisted editing induces recombination at a desired target locus and kills non-edited cells leading to a recovery of close to a 100% of edited cells [614]. In theory, the same principle should be true for non-edited chromosomes. Only those chromosome copies that contain the desired modification should be able to survive the constant CRISPR attack. Unfortunately, applying this system in *T. kodakarensis* is not possible as a thermostable CRISPR-Cas9 system is not (yet) available.

Continuing the search for a tetraether synthase

In **chapter 5** we set out to identify the enzyme(s) responsible for the synthesis of tetraether lipids. Especially, the one(s) responsible for the condensation of two C₂₀ isoprenyl chains, i.e. the putative tetraether synthase(s). Though interesting results have been obtained, the data did not allow for the establishment of the exact functions of the candidate genes. Future research is thus needed to get a better understanding of the tetraether synthesis pathway and to identify the tetraether synthase(s). To do this, several approaches may be followed.

First of all, one may repeat the lipid analysis. The reason for this is that in the analysis performed in **chapter 5**, we used HPLC–APCI/MS with selected ion monitoring, aimed to detect only the most prevalent lipids in the membrane which are represented by the core lipid species archaeol (diether) and GDGT-0 (tetraether). However, the amount of diether and tetraether lipids obtained for some of the mutant strains was so low (**Figure 5.3**) that either the cellular integrity of the mutant strains must have been compromised or that lipids other than archaeol and GDGT-0 must have been produced. To test this, thin cross-sections of the mutant strain showing the lowest amount of diether and tetraether lipids, i.e. OE1324, were prepared and analyzed by Transmission Electron Microscopy (TEM). The results obtained (**Figure 5.5**) clearly showed that the cellular integrity of this mutant strain has not been compromised, strongly suggesting that other lipids must have been produced. Repeating the lipid analysis with a total ion monitoring, or with a selected ion monitoring aimed to detect additional lipid species, could therefore provide new insights into the lipid biosynthesis pathway and the membrane composition of *T. kodakarensis*.

To help the identification of the tetraether synthase, some additional experiments may be performed. For example, from literature it seems that as a response to various stress conditions, many archaea increase the tetraether lipid content of their membrane [546, 550, 551, 553]. Therefore, exposing *T. kodakarensis* to various stress conditions (e.g. high/low pH, temperature, or salinity) and subsequently analyzing its membrane composition, might reveal growth conditions that lead to an increased tetraether lipid content. Once such a growth condition has been identified, a differential proteome analysis may be performed which might help in identifying the tetraether synthase.

Moreover, to establish the exact functions of various candidate genes, and to identify the tetraether synthase, it would be desirable to be able to perform in vitro enzyme activity assays. However, given the complicated nature of these reactions, performing such in vitro assays may not be easy to realize. For example, not only is the exact substrate of the tetraether synthase still unknown, but the enzyme itself may need to be embedded in a membrane to be active. Moreover, real-time monitoring of the reaction, as for instance is done in a colorimetric assay in a spectrophotometer, might not be possible. Nevertheless, developing an in vitro enzyme assay may ultimately be required to confirm promising candidate genes as the tetraether synthase.

Hydrogen: a clean fuel of the future?

Future energy scenarios will mainly be based on a mix of renewable primary energy sources (e.g. solar-, wind-, or geothermal energy, and biomass), in combination with energy carriers that are able to address environmental and security issues posed by current fuels. As discussed in **chapter 1 and 3**, one of the promising fuels in this respect is H₂. Hydrogen is readily available (though it may take some work to access), can be produced locally, is non-toxic, can be burned without harmful emissions (no CO₂ and little or no NO_x), has the highest energy density per unit fuel mass (which is why H₂ is already used as a propellant in space industry), and can be renewable. Moreover, the energy stored in H₂ can relatively easily be converted to electricity using H₂ fuel cells, that can also be applied to power vehicles and are in fact more efficient than combustion processes (e.g. H₂ fuel cells can reach efficiencies up

to 60% whereas conventional gasoline powered combustion engines only have an efficiency of less than 20% [615]). In addition, H_2 fuel cell powered vehicles (FCVs) have several benefits compared to the already more general available battery electric vehicles (BEVs). For instance, FCVs show a better mileage, can be refueled a lot faster, and are more scalable, which makes them more suitable for heavy transportation. Finally, when compared to electricity (the other clean energy carrier), H_2 has the major advantage that it can be used for long-term renewable energy storage.

Nevertheless, though very promising, several barriers still need to be overcome to realize the huge potential of H_2 as a true clean fuel [17]. For example, as H_2 is only a means for transmitting energy from primary energy sources to end-users, it has to be produced from primary energy sources. Unfortunately, the far majority of the H_2 produced today still relies on fossil fuels. Therefore, to become a true clean fuel, H_2 has to be produced exclusively from clean renewable primary energy sources, which from an economically point of view is still too expensive. Moreover, in contrast to the current fuels, the required H_2 infrastructure to make it a practical fuel is virtually non-existing. This means that huge investments are needed to build the necessary pipelines and fueling stations, and that, until this is realized, H_2 -powered vehicles remain unattractive for consumers. In addition, though FCVs recently became available on the market, they are currently still too expensive to be attractive to the general public. Finally, another major barrier that needs to be overcome is storage.

Though H_2 has the highest energy density per unit fuel mass, it also has the lowest energy density per unit fuel volume. This means that in order to store sufficient H_2 to power FCVs, extremely large storage tanks would be required under standard conditions. Therefore, to store H_2 more efficiently and compactly, it must be compressed or cryogenically liquefied. However, even in a compressed form, H_2 still has a very low energy density per unit fuel volume, which severely limits its practicality as a fuel for FCVs. Similarly, though better than pressurized H_2 , liquid H_2 still only has about a third of the volumetric value compared to gasoline [616]. Moreover, liquid H_2 also has some challenges in terms of efficiency. This is because a substantial part of its energy content is lost during the liquefaction process and, when finally stored as a liquid, H_2 must be allowed to bleed off to avoid pressure build-up

[17]. So for now, though major improvements have been accomplished in the last decade, H_2 still remains just a very promising alternative to current fuels. Future research and development is required to unlock its potential and to implement H_2 as a widespread and practical transportation fuel.

References

1. White, L.A., *The Science of Culture: A Study of Man and Civilization*, 1949, Michigan: Farrar, Straus.
2. Smil, V., *Energy in World History*, 1994, Boulder, Colorado: Westview Press.
3. Rifkin, J., *The Hydrogen Economy*, 2003, New York: Penguin Publishing Group.
4. Fröling, M., *Energy use, population and growth, 1800–1970*. Journal of Population Economics, 2011. **24**(3): p. 1133–1163.
5. Kostic, M., *Energy: global and historical background* in *Encyclopedia of energy engineering and technology*, 2007, New York: Taylor & Francis. p. 601–615
6. Cavendish, H., *Three Papers, Containing Experiments on Factitious Air*. Philosophical Transactions (1683–1775), 1766. **56**: p. 141–184.
7. Heiserman, D., *Exploring Chemical Elements and Their Compounds*, 1991, New York: McGraw-Hill Companies, Incorporated.
8. Züttel, A., A. Borgschulte, and L. Schlapbach, *Hydrogen as a future energy carrier*, 2011, Weinheim: John Wiley & Sons.
9. Hart, D. and S. Hydrogen. 2011; Available from: <http://www.eoearth.org/view/article/153620>.
10. Chou, C.J., et al., *Hydrogenesis in hyperthermophilic microorganisms: Implications for biofuels*. Metabolic Engineering, 2008. **10**(6): p. 394–404.
11. Das, D. and T.N. Veziroğlu, *Hydrogen production by biological processes: a survey of literature*. International Journal of Hydrogen Energy, 2001. **26**(1): p. 13–28.
12. Dincer, I. and C. Acar, *Review and evaluation of hydrogen production methods for better sustainability*. International Journal of Hydrogen Energy, 2014. **40**(34): p. 11094–11111.
13. Kengen, S.W.M., et al., *biological hydrogen production by anaerobic microorganisms*. In Biofuels Edited by: Soetaert W, Vandamme E.J. Chichester: John Wiley & Sons, 2009: p. 197–221.
14. Pawar, S.S. and E.W.J. van Niel, *Thermophilic biohydrogen production: how far are we?* Applied Microbiology and Biotechnology, 2013. **97**(18): p. 7999–8009.
15. The Freedonia Group Inc., *World hydrogen*, 2014: Cleveland.
16. Armaroli, N. and V. Balzani, *The Hydrogen Issue*. ChemSusChem, 2011. **4**(1): p. 21–36.
17. Bennaceur, K., et al., *Hydrogen: a future energy carrier?* Oilfield Review, 2005. **17**(1): p. 30–41.
18. Marbán, G. and T. Valdés-Solis, *Towards the hydrogen economy?* International Journal of Hydrogen Energy, 2007. **32**(12): p. 1625–1637.
19. Ramachandran, R. and R.K. Menon, *An overview of industrial uses of hydrogen*. International Journal of Hydrogen Energy, 1998. **23**(7): p. 593–598.
20. Rosen, M.A., *Advances in hydrogen production by thermochemical water decomposition: A review*. Energy, 2010. **35**(2): p. 1068–1076.
21. Winter, C.-J., *Hydrogen energy—Abundant, efficient, clean: A debate over the energy-system-of-change*. International Journal of Hydrogen Energy, 2009. **34**(14): p. S1–S52.
22. Yilanci, A., I. Dincer, and H.K. Ozturk, *A review on solar-hydrogen/fuel cell hybrid energy systems for stationary applications*. Progress in Energy and Combustion Science, 2009. **35**(3): p. 231–244.
23. Rachman, M.A., et al., *Hydrogen production with high yield and high evolution rate by self-flocculated cells of Enterobacter aerogenes in a packed-bed reactor*. Applied Microbiology and Biotechnology, 1998. **49**(4): p. 450–454.
24. Kim, B.H., et al., *Control of Carbon and Electron Flow in Clostridium acetobutylicum Fermentations: Utilization of Carbon Monoxide to Inhibit Hydrogen Production and to Enhance Butanol Yields*. Applied and Environmental Microbiology, 1984. **48**(4): p. 764–770.
25. Minnan, L., et al., *Isolation and characterization of a high H₂-producing strain Klebsiella oxytoca HP1 from a hot spring*. Research in microbiology, 2005. **156**(1): p. 76–81.
26. Vancanneyt, M., et al., *Ethanol Production in Batch and Continuous Culture from Some Carbohydrates with Clostridium thermosaccharolyticum LMG 6564*. Systematic and Applied Microbiology, 1990. **13**(4): p. 382–387.
27. Magnusson, L., et al., *Continuous hydrogen production during fermentation of α-cellulose by the thermophilic bacterium Clostridium thermocellum*. Biotechnology and Bioengineering, 2009. **102**(3): p. 759–766.
28. O-Thong, S., et al., *Thermophilic fermentative hydrogen production by the newly isolated Thermoanaerobacterium thermosaccharolyticum PSU-2*. International Journal of Hydrogen Energy, 2008. **33**(4): p. 1204–1214.
29. Kengen, S.W.M. and A.J.M. Stams, *Growth and energy conservation in batch cultures of Pyrococcus furiosus*. Fems Microbiology Letters, 1994. **117**(3): p. 305–309.
30. Kanai, T., et al., *Continuous hydrogen production by the hyperthermophilic archaeon, Thermococcus kodakaraensis KOD1*. Journal of Biotechnology, 2005. **116**(3): p. 271–282.
31. de Vrije, T., et al., *Glycolytic pathway and hydrogen yield studies of the extreme thermophile Caldicellulosiruptor saccharolyticus*. Applied Microbiology and Biotechnology, 2007. **74**(6): p. 1358–1367.
32. Schröder, C., M. Selig, and P. Schönheit, *Glucose fermentation to acetate, CO₂ and H₂ in the anaerobic hyperthermophilic eubacterium Thermotoga maritima: involvement of the Embden-Meyerhof pathway*. Archives of Microbiology, 1994. **161**(6): p. 460–470.
33. Heffel, J.W., *NO_x emission and performance data for a hydrogen fueled internal combustion engine at 1500rpm using exhaust gas recirculation*. International Journal of Hydrogen Energy, 2003. **28**(8): p. 901–908.
34. Flamholz, A., et al., *eQuilibrator—the biochemical thermodynamics calculator*. Nucleic Acids Research, 2012. **40**(D1): p. D770–D775.
35. Noor, E., et al., *An integrated open framework for thermodynamics of reactions that combines accuracy and coverage*. Bioinformatics, 2012. **28**(15): p. 2037–2044.

Appendices

36. Noor, E., et al., *Consistent Estimation of Gibbs Energy Using Component Contributions*. PLoS Comput Biol, 2013. **9**(7): p. e1003098.
37. Holladay, J.D., et al., *An overview of hydrogen production technologies*. Catalysis Today, 2009. **139**(4): p. 244-260.
38. Rostrup-Nielsen, J.R., *Conversion of hydrocarbons and alcohols for fuel cells*. Physical Chemistry Chemical Physics, 2001. **3**(3): p. 283-288.
39. Zhang, Y.H.P., *Hydrogen Production from Carbohydrates: A Mini-Review*, in *Sustainable Production of Fuels, Chemicals, and Fibers from Forest Biomass*, 2011, American Chemical Society. p. 203-216.
40. Hsueh, H.T., H. Chu, and S.T. Yu, *A batch study on the bio-fixation of carbon dioxide in the absorbed solution from a chemical wet scrubber by hot spring and marine algae*. Chemosphere, 2007. **66**(5): p. 878-886.
41. Lo, Y.-C., et al., *Sequential dark-photo fermentation and autotrophic microalgal growth for high-yield and CO₂-free biohydrogen production*. International Journal of Hydrogen Energy, 2010. **35**(20): p. 10944-10953.
42. Hallenbeck, P.C. and D. Ghosh, *Advances in fermentative biohydrogen production: the way forward?* Trends in Biotechnology, 2009. **27**(5): p. 287-297.
43. Rittmann, S. and C. Herwig, *A comprehensive and quantitative review of dark fermentative biohydrogen production*. Microb Cell Fact, 2012. **11**(1): p. 115.
44. Claassen, P.A.M., et al., *Non-thermal production of pure hydrogen from biomass: HYVOLUTION*. Journal of Cleaner Production, 2010. **18**, **Supplement 1**: p. S4-S8.
45. Guo, X.M., et al., *Hydrogen production from agricultural waste by dark fermentation: A review*. International Journal of Hydrogen Energy, 2010. **35**(19): p. 10660-10673.
46. Sinha, P. and A. Pandey, *An evaluative report and challenges for fermentative biohydrogen production*. International Journal of Hydrogen Energy, 2011. **36**(13): p. 7460-7478.
47. Zhang, Y.H.P., *Production of biocommodities and bioelectricity by cell-free synthetic enzymatic pathway biotransformations: Challenges and opportunities*. Biotechnology and Bioengineering, 2010. **105**(4): p. 663-677.
48. Bielen, A.A.M., et al., *Biohydrogen Production by the Thermophilic Bacterium Caldicellulosiruptor saccharolyticus: Current Status and Perspectives*. Life : Open Access Journal, 2013. **3**(1): p. 52-85.
49. Verhaart, M.R., et al., *Hydrogen production by hyperthermophilic and extremely thermophilic bacteria and archaea: mechanisms for reductant disposal*. Environ Technol, 2010. **31**(8-9): p. 993-1003.
50. Hallenbeck, P.C. and J.R. Benemann, *Biological hydrogen production; fundamentals and limiting processes*. International Journal of Hydrogen Energy, 2002. **27**(11-12): p. 1185-1193.
51. Thauer, R., *Limitation of microbial hydrogen formation via fermentation*, in *Microbial energy conversion*, H.G. Schlegel and J. Barnea(ed), Editors. 1976, Erich Goltze KG Göttingen, Germany. p. 201-294.
52. Thauer, R.K., K. Jungermann, and K. Decker, *Energy conservation in chemotrophic anaerobic bacteria*. Bacteriol Rev, 1977. **41**(1): p. 100-80.
53. Liu, H., S. Grot, and B.E. Logan, *Electrochemically Assisted Microbial Production of Hydrogen from Acetate*. Environmental Science & Technology, 2005. **39**(11): p. 4317-4320.
54. Rollin, J.A., et al., *High-yield hydrogen production from biomass by in vitro metabolic engineering: Mixed sugars coutilization and kinetic modeling*. Proceedings of the National Academy of Sciences, 2015. **112**(16): p. 4964-4969.
55. Ye, X., et al., *Spontaneous high-yield production of hydrogen from cellulosic materials and water catalyzed by enzyme cocktails*. ChemSusChem, 2009. **2**(2): p. 149-152.
56. Zhang, Y.-H.P., et al., *High-yield hydrogen production from starch and water by a synthetic enzymatic pathway*. Plos One, 2007. **2**(5): p. e456.
57. Atomi, H., et al., *Description of Thermococcus kodakaraensis sp. nov., a well studied hyperthermophilic archaeon previously reported as Pyrococcus sp. KOD1*. Archaea, 2004. **1**: p. 263-267.
58. Morikawa, M., et al., *Purification and characterization of a thermostable thiol protease from a newly isolated hyperthermophilic Pyrococcus sp.* Applied and Environmental Microbiology, 1994. **60**(12): p. 4559-4566.
59. Fukui, T., et al., *Complete genome sequence of the hyperthermophilic archaeon Thermococcus kodakaraensis KOD1 and comparison with Pyrococcus genomes*. Genome Research, 2005. **15**(3): p. 352-363.
60. Hileman, T.H. and T.J. Santangelo, *Genetics techniques for Thermococcus kodakaraensis*. Frontiers in Microbiology, 2012. **3**: p. 195.
61. Santangelo, T.J. and J.N. Reeve, *Genetic tools and manipulations of the hyperthermophilic heterotrophic archaeon Thermococcus kodakaraensis*, in *Extremophiles Handbook*, 2011, Springer. p. 567-582.
62. Sato, T., et al., *Targeted gene disruption by homologous recombination in the hyperthermophilic archaeon Thermococcus kodakaraensis KOD1*. Journal of Bacteriology, 2003. **185**(1): p. 210-220.
63. Sato, T., et al., *Improved and versatile transformation system allowing multiple genetic manipulations of the hyperthermophilic archaeon Thermococcus kodakaraensis*. Applied and Environmental Microbiology, 2005. **71**(7): p. 3889-3899.
64. Holden, J.F., et al., *Diversity among three novel groups of hyperthermophilic deep-sea Thermococcus species from three sites in the northeastern Pacific Ocean*. FEMS Microbiology Ecology, 2001. **36**(1): p. 51-60.
65. Orphan, V.J., et al., *Culture-Dependent and Culture-Independent Characterization of Microbial Assemblages Associated with High-Temperature Petroleum Reservoirs*. Applied and Environmental Microbiology, 2000. **66**(2): p. 700-711.
66. Euzéby, J., *Validation of publication of new names and new combinations previously effectively published outside the IUSEM*. International Journal of Systematic and Evolutionary Microbiology, 2005. **55**(3): p. 983-985.
67. Awano, T., et al., *Characterization of two members among the five ADP-forming acyl-CoA synthetases reveals the presence of a 2-(imidazol-4-yl)acetyl-CoA synthetase in Thermococcus kodakaraensis*. Journal of Bacteriology, 2014. **196**(1): p. 140-147.
68. Nohara, K., et al., *Genetic Examination and Mass Balance Analysis of Pyruvate/Amino Acid Oxidation Pathways in the Hyperthermophilic Archaeon Thermococcus kodakaraensis*. Journal of Bacteriology, 2014. **196**(22): p. 3831-3839.

69. Shikata, K., et al., *A Novel ADP-forming Succinyl-CoA Synthetase in Thermococcus kodakaraensis Structurally Related to the Archaeal Nucleoside Diphosphate-forming Acetyl-CoA Synthetases*. Journal of Biological Chemistry, 2007. **282**(37): p. 26963-26970.
70. Siddiqui, M., S. Fujiwara, and T. Imanaka, *Indolepyruvate ferredoxin oxidoreductase from Pyrococcus sp. KOD1 possesses a mosaic structure showing features of various oxidoreductases*. Molecular and General Genetics MGG, 1997. **254**(4): p. 433-439.
71. Yokooji, Y., et al., *Genetic Examination of Initial Amino Acid Oxidation and Glutamate Catabolism in the Hyperthermophilic Archaeon Thermococcus kodakaraensis*. Journal of Bacteriology, 2013. **195**(9): p. 1940-1948.
72. Matsumi, R., et al., *Disruption of a sugar transporter gene cluster in a hyperthermophilic archaeon using a host-marker system based on antibiotic resistance*. Journal of Bacteriology, 2007. **189**(7): p. 2683-2691.
73. Ahmad, N., et al., *Novel Maltotriose-Hydrolyzing Thermoacidophilic Type III Pullulan Hydrolase from Thermococcus kodakaraensis*. Applied and Environmental Microbiology, 2014. **80**(3): p. 1108-1115.
74. Guan, Q., et al., *Cloning, purification and biochemical characterisation of an organic solvent-, detergent-, and thermo-stable amylopullulanase from Thermococcus kodakaraensis KOD1*. Process Biochemistry, 2013. **48**(5-6): p. 878-884.
75. Rashid, N., et al., *Characterization of an archaeal cyclodextrin glucanotransferase with a novel C-terminal domain*. Journal of Bacteriology, 2002. **184**(3): p. 777-84.
76. Sun, Y., et al., *Recombinant Cyclodextrinase from Thermococcus kodakaraensis KOD1: Expression, Purification, and Enzymatic Characterization*. Archaea, 2015. **2015**: p. 397924.
77. Matsubara, K., et al., *Biochemical and genetic characterization of the three metabolic routes in Thermococcus kodakaraensis linking glyceraldehyde 3-phosphate and 3-phosphoglycerate*. Molecular Microbiology, 2011. **81**(5): p. 1300-1312.
78. Brasen, C., et al., *Carbohydrate metabolism in Archaea: current insights into unusual enzymes and pathways and their regulation*. Microbiology and molecular biology reviews : MMBR, 2014. **78**(1): p. 89-175.
79. Kengen, S.W., et al., *Evidence for the operation of a novel Embden-Meyerhof pathway that involves ADP-dependent kinases during sugar fermentation by Pyrococcus furiosus*. Journal of Biological Chemistry, 1994. **269**(26): p. 17537-41.
80. Verhees, C.H., et al., *The unique features of glycolytic pathways in Archaea*. Vol. 375. 2003. 231-246.
81. Imanaka, H., et al., *Phosphoenolpyruvate synthase plays an essential role for glycolysis in the modified Embden-Meyerhof pathway in Thermococcus kodakaraensis*. Molecular Microbiology, 2006. **61**(4): p. 898-909.
82. Kanai, T., et al., *Distinct Physiological Roles of the Three [NiFe]-Hydrogenase Orthologs in the Hyperthermophilic Archaeon Thermococcus kodakaraensis*. Journal of Bacteriology, 2011. **193**(12): p. 3109-3116.
83. Santangelo, T.J., L. Cubonova, and J.N. Reeve, *Deletion of alternative pathways for reductant recycling in Thermococcus kodakaraensis increases hydrogen production*. Molecular Microbiology, 2011. **81**(4): p. 897-911.
84. Kengen, S.W. and A.J. Stams, *Formation of L-alanine as a reduced end product in carbohydrate fermentation by the hyperthermophilic archaeon Pyrococcus furiosus*. Archives of Microbiology, 1994. **161**(2): p. 168-175.
85. Abbas, S.N., et al., *Molecular cloning and characterization of TK1111, a cupin-type phosphoglucose isomerase from Thermococcus kodakaraensis*. Türk Biyokimya Dergisi [Turkish Journal of Biochemistry-Türk J Biochem], 2013. **38**(4): p. 438-444.
86. Fukuda, W., et al., *Characterization of an archaeal malic enzyme from the hyperthermophilic archaeon Thermococcus kodakaraensis KOD1*. Archaea (Vancouver, B C), 2005. **1**(5): p. 293-301.
87. Imanaka, H., et al., *Gene cloning and characterization of fructose-1,6-bisphosphate aldolase from the hyperthermophilic archaeon Thermococcus kodakaraensis KOD1*. Journal of Bioscience and Bioengineering, 2002. **94**(3): p. 237-243.
88. Kanai, T., et al., *A global transcriptional regulator in Thermococcus kodakaraensis controls the expression levels of both glycolytic and gluconeogenic enzyme-encoding genes*. Journal of Biological Chemistry, 2007. **282**(46): p. 3139-3150.
89. Kobori, H., et al., *Characterization of NADH oxidase/NADPH polysulfide oxidoreductase and its unexpected participation in oxygen sensitivity in an anaerobic hyperthermophilic archaeon*. Journal of Bacteriology, 2010. **192**(19): p. 5192-202.
90. Mizanur, R.M., A.K. Griffin, and N.L. Pohl, *Recombinant production and biochemical characterization of a hyperthermostable α -glucan/maltodextrin phosphorylase from Pyrococcus furiosus*. Archaea, 2008. **2**(3): p. 169-176.
91. Orita, I., et al., *The ribulose monophosphate pathway substitutes for the missing pentose phosphate pathway in the archaeon Thermococcus kodakaraensis*. Journal of Bacteriology, 2006. **188**(13): p. 4698-4704.
92. Rashid, N., et al., *Among Multiple Phosphomannomutase Gene Orthologues, Only One Gene Encodes a Protein with Phosphoglucomutase and Phosphomannomutase Activities in Thermococcus kodakaraensis*. Journal of Bacteriology, 2004. **186**(18): p. 6070-6076.
93. Rasool, N., et al., *Proteolytic inventory of Thermococcus kodakaraensis*. Afr. J. Microbiol. Res, 2013. **7**: p. 3139-3150.
94. Sato, T. and H. Atom, *Novel metabolic pathways in Archaea*. Current Opinion in Microbiology, 2011. **14**(3): p. 307-314.
95. Tachibana, Y., et al., *Cloning and expression of the α -glucanotransferase gene from the hyperthermophilic archaeon Pyrococcus sp. KOD1, and characterization of the enzyme*. Journal of Fermentation and Bioengineering, 1997. **83**(6): p. 540-548.
96. Tachibana, Y., et al., *Cloning and expression of the α -amylase gene from the hyperthermophilic archaeon Pyrococcus sp. KOD1, and characterization of the enzyme*. Journal of Fermentation and Bioengineering, 1996. **82**(3): p. 224-232.
97. Soboh, B., D. Linder, and R. Hedderich, *A multisubunit membrane-bound [NiFe] hydrogenase and an NADH-dependent Fe-only hydrogenase in the fermenting bacterium Thermoanaerobacter tengcongensis*. Microbiology, 2004. **150**(7): p. 2451-2463.
98. van de Werken, H.J., et al., *Hydrogenomics of the extremely thermophilic bacterium Caldicellulosiruptor saccharolyticus*. Appl Environ Microbiol, 2008. **74**(21): p. 6720-9.
99. Spaans, S.K., et al., *NADPH-generating systems in bacteria and archaea*. Frontiers in Microbiology, 2015. **6**(742): p. 159-88.
100. Vignais, P.M. and B. Billoud, *Occurrence, classification, and biological function of hydrogenases: an overview*. Chemical reviews, 2007. **107**(10): p. 4206-4272.
101. Vignais, P.M. and A. Colbeau, *Molecular biology of microbial hydrogenases*. Current issues in molecular biology, 2004. **6**(2): p. 159-88.
102. Kanai, T., S. Ito, and T. Imanaka, *Characterization of a Cytosolic NiFe-Hydrogenase from the Hyperthermophilic Archaeon Thermococcus kodakaraensis KOD1*. Journal of Bacteriology, 2003. **185**(5): p. 1705-1711.

Appendices

103. Ma, K. and M.W.W. Adams, *Hydrogenases I and II from Pyrococcus furiosus*, in *Methods in enzymology*, R.M.K. Michael W. W. Adams, Editor 2001, Academic Press. p. 208-216.
104. Fukuda, W., et al., *Agmatine is essential for the cell growth of Thermococcus kodakaraensis*. *Fems Microbiology Letters*, 2008. **287**(1): p. 113-120.
105. Sato, T., et al., *Genetic evidence identifying the true gluconeogenic fructose-1,6-bisphosphatase in Thermococcus kodakaraensis and other hyperthermophiles*. *Journal of Bacteriology*, 2004. **186**(17): p. 5799-5807.
106. Takemasa, R., et al., *Thermococcus kodakarensis as a Host for Gene Expression and Protein Secretion*. *Applied and Environmental Microbiology*, 2011. **77**(7): p. 2392-2398.
107. Hanford, M. and T. Peeples, *Archaeal tetraether lipids*. *Appl Biochem Biotechnol*, 2002. **97**(1): p. 45-62.
108. Oger, P.M. and A. Cario, *Adaptation of the membrane in Archaea*. *Biophysical Chemistry*, 2013. **183**: p. 42-56.
109. Boyd, E., et al., *The Role of Tetraether Lipid Composition in the Adaptation of Thermophilic Archaea to Acidity*. *Frontiers in Microbiology*, 2013. **4**(62).
110. Villanueva, L., J.S.S. Damste, and S. Schouten, *A re-evaluation of the archaeal membrane lipid biosynthetic pathway*. *Nat Rev Micro*, 2014. **12**(6): p. 438-448.
111. Brown, D.I. and K.K. Griendling, *Nox proteins in signal transduction*. *Free Radical Biology and Medicine*, 2009. **47**(9): p. 1239-1253.
112. Bylund, J., et al., *Intracellular generation of superoxide by the phagocyte NADPH oxidase: How, where, and what for?* *Free Radical Biology and Medicine*, 2010. **49**(12): p. 1834-1845.
113. Foreman, J., et al., *Reactive oxygen species produced by NADPH oxidase regulate plant cell growth*. *Nature*, 2003. **422**(6930): p. 442-446.
114. Nakamura, M., A. Bhatnagar, and J. Sadoshima, *Overview of Pyridine Nucleotides Review Series*. *Circulation Research*, 2012. **111**(5): p. 604-610.
115. Minard, K.I. and L. McAlister-Henn, *Sources of NADPH in Yeast Vary with Carbon Source*. *Journal of Biological Chemistry*, 2005. **280**(48): p. 39890-39896.
116. Nordberg, J. and E.S.J. Arnér, *Reactive oxygen species, antioxidants, and the mammalian thioredoxin system*. *Free Radical Biology and Medicine*, 2001. **31**(11): p. 1287-1312.
117. Singh, R., et al., *A Novel Strategy Involved Anti-Oxidative Defense: The Conversion of NADH into NADPH by a Metabolic Network*. *Plos One*, 2008. **3**(7): p. e2682.
118. Arnér, E.S.J. and A. Holmgren, *Physiological functions of thioredoxin and thioredoxin reductase*. *European Journal of Biochemistry*, 2000. **267**(20): p. 6102-6109.
119. Hügler, M., et al., *Malonyl-Coenzyme A Reductase from Chloroflexus aurantiacus, a Key Enzyme of the 3-Hydroxypropionate Cycle for Autotrophic CO₂ Fixation*. *Journal of Bacteriology*, 2002. **184**(9): p. 2404-2410.
120. Koh, H.-J., et al., *Cytosolic NADP⁺-dependent Isocitrate Dehydrogenase Plays a Key Role in Lipid Metabolism*. *Journal of Biological Chemistry*, 2004. **279**(38): p. 39968-39974.
121. Miller, W.L. and R.J. Auchus, *The Molecular Biology, Biochemistry, and Physiology of Human Steroidogenesis and Its Disorders*. *Endocrine Reviews*, 2011. **32**(1): p. 81-151.
122. Lee, W.H., et al., *Engineering of NADPH regenerators in Escherichia coli for enhanced biotransformation*. *Applied Microbiology and Biotechnology*, 2013. **97**(7): p. 1-12.
123. Papagianni, M., *Recent advances in engineering the central carbon metabolism of industrially important bacteria*. *Microbial Cell Factories*, 2012. **11**: p. 50-63.
124. Chenault, H.K., E.S. Simon, and G.M. Whitesides, *Cofactor Regeneration for Enzyme-Catalysed Synthesis*. *Biotechnology and Genetic Engineering Reviews*, 1988. **6**(1): p. 221-270.
125. Uppada, V., S. Bhaduri, and S.B. Noronha, *Cofactor regeneration—an important aspect of biocatalysis*. *Current Science*, 2014. **106**(7): p. 946.
126. Wandrey, C., *Biochemical reaction engineering for redox reactions*. *Chem Rec*, 2004. **4**(4): p. 254-65.
127. Wichmann, R. and D. Vasic-Racki, *Cofactor regeneration at the lab scale*. *Adv Biochem Eng Biotechnol*, 2005. **92**: p. 225-60.
128. Zhang, J. and J.H. Xu, *Biocatalysis, Cofactor Regeneration*. *Encyclopedia of Industrial Biotechnology*, 2010.
129. Chenault, H.K. and G.M. Whitesides, *Regeneration of nicotinamide cofactors for use in organic synthesis*. *Appl Biochem Biotechnol*, 1987. **14**(2): p. 147-97.
130. Liu, W. and P. Wang, *Cofactor regeneration for sustainable enzymatic biosynthesis*. *Biotechnol Adv*, 2007. **25**(4): p. 369-84.
131. Zhao, H. and W.A. van der Donk, *Regeneration of cofactors for use in biocatalysis*. *Curr Opin Biotechnol*, 2003. **14**(6): p. 583-9.
132. van der Donk, W.A. and H. Zhao, *Recent developments in pyridine nucleotide regeneration*. *Curr Opin Biotechnol*, 2003. **14**(4): p. 421-6.
133. Chemler, J.A., et al., *Improving NADPH availability for natural product biosynthesis in Escherichia coli by metabolic engineering*. *Metabolic Engineering*, 2010. **12**(2): p. 96-104.
134. Siedler, S., S. Bringer, and M. Bott, *Increased NADPH availability in Escherichia coli: improvement of the product per glucose ratio in reductive whole-cell biotransformation*. *Applied Microbiology and Biotechnology*, 2011. **92**(5): p. 929-937.
135. Zhao, Z., et al., *Cytosolic NADPH balancing in Penicillium chrysogenum cultivated on mixtures of glucose and ethanol*. *Applied Microbiology and Biotechnology*, 2011. **89**(1): p. 63-72.
136. Becker, J., et al., *Metabolic flux engineering of L-lysine production in Corynebacterium glutamicum—over expression and modification of G6P dehydrogenase*. *Journal of Biotechnology*, 2007. **132**(2): p. 99-109.
137. Asadollahi, M.A., et al., *Enhancing sesquiterpene production in Saccharomyces cerevisiae through in silico driven metabolic engineering*. *Metabolic Engineering*, 2009. **11**(6): p. 328-334.
138. Peralta-Yahya, P.P., et al., *Microbial engineering for the production of advanced biofuels*. *Nature*, 2012. **488**(7411): p. 320-328.
139. Kim, Y.M., et al., *Engineering the pentose phosphate pathway to improve hydrogen yield in recombinant Escherichia coli*. *Biotechnology and Bioengineering*, 2011. **108**(12): p. 2941-2946.

140. Kabir, M.M. and K. Shimizu, *Fermentation characteristics and protein expression patterns in a recombinant Escherichia coli mutant lacking phosphoglucose isomerase for poly(3-hydroxybutyrate) production*. Applied Microbiology and Biotechnology, 2003. **62**(2-3): p. 244-255.
141. Sauer, U., et al., *The Soluble and Membrane-bound Transhydrogenases UdhA and PntAB Have Divergent Functions in NADPH Metabolism of Escherichia coli*. Journal of Biological Chemistry, 2004. **279**(8): p. 6613-6619.
142. Agledal, L., M. Niere, and M. Ziegler, *The phosphate makes a difference: Cellular functions of NADP*. Redox Report, 2010. **15**(1): p. 2-10.
143. Churamani, D., et al., *Determination of cellular nicotinic acid-adenine dinucleotide phosphate (NAADP) levels*. Biochem. J., 2004. **380**(2): p. 449-454.
144. Patel, S., et al., *The endo-lysosomal system as an NAADP-sensitive acidic Ca²⁺ store: role for the two-pore channels*. Cell calcium, 2011. **50**(2): p. 157-167.
145. Lee, H.C., *Mechanisms of calcium signaling by cyclic ADP-ribose and NAADP*. Physiological Reviews, 1997. **77**(4): p. 1133-1164.
146. Chini, E.N. and F.G.S. De Toledo, *Nicotinic acid adenine dinucleotide phosphate: a new intracellular second messenger?* American Journal of Physiology - Cell Physiology, 2002. **282**(6): p. C1191-C1198.
147. Vu, C.Q., et al., *2'-phospho-cyclic ADP-ribose, a calcium-mobilizing agent derived from NADP*. Journal of Biological Chemistry, 1996. **271**(9): p. 4747-4754.
148. Ying, W., *NAD⁺/NADH and NADP⁺/NADPH in cellular functions and cell death: regulation and biological consequences*. Antioxidants & redox signaling, 2008. **10**(2): p. 179-206.
149. Davis, W.B., *Identification of a nicotinamide adenine dinucleotide glycohydrolase and an associated inhibitor in isoniazid-susceptible and -resistant Mycobacterium phlei*. Antimicrob Agents Chemother, 1980. **17**(4): p. 663-8.
150. Everse, J., J.B. Griffin, and N.O. Kaplan, *The pyridine nucleosidase from Bacillus subtilis. Kinetic properties and enzyme-inhibitor interactions*. Arch Biochem Biophys, 1975. **169**(2): p. 714-23.
151. Ghosh, J., et al., *Characterization of Streptococcus pyogenes beta-NAD⁺ glycohydrolase: re-evaluation of enzymatic properties associated with pathogenesis*. J Biol Chem, 2010. **285**(8): p. 5683-94.
152. Kawai, S. and K. Murata, *Structure and Function of NAD Kinase and NADP Phosphatase: Key Enzymes That Regulate the Intracellular Balance of NAD(H) and NADP(H)*. Bioscience, Biotechnology, and Biochemistry, 2008. **72**(4): p. 919-930.
153. Mather, L.H. and M. Knight, *A heat-stable nicotinamide-adenine dinucleotide glycohydrolase from Pseudomonas putida KB1. Partial purification and some properties of the enzyme and an inhibitory protein*. Biochem J, 1972. **129**(1): p. 141-52.
154. Penyige, A., et al., *Evidence of a role for NAD⁺-glycohydrolase and ADP-ribosyltransferase in growth and differentiation of Streptomyces griseus NRRL B-2682: inhibition by m-aminophenylboronic acid*. Microbiology, 1996. **142** (Pt 8): p. 1937-44.
155. Pollak, N., C. Dölle, and M. Ziegler, *The power to reduce: pyridine nucleotides – small molecules with a multitude of functions*. Biochem J, 2007. **402**(2): p. 205-218.
156. Ziegler, M., *New functions of a long-known molecule*. European Journal of Biochemistry, 2000. **267**(6): p. 1550-1564.
157. Magni, G., et al., *Enzymology of NAD⁺ homeostasis in man*. Cellular and Molecular Life Sciences, 2004. **61**(1): p. 19-34.
158. Begley, T.P., et al., *The biosynthesis of nicotinamide adenine dinucleotides in bacteria*, in *Vitamins & Hormones*, 2001, Academic Press. p. 103-119.
159. Magni, G., et al., *Enzymology of NAD⁺ synthesis*, in *Advances in Enzymology and Related Areas of Molecular Biology* 1999. p. 135-182.
160. Gazzaniga, F., et al., *Microbial NAD Metabolism: Lessons from Comparative Genomics*. Microbiology and Molecular Biology Reviews, 2009. **73**(3): p. 529-541.
161. Gossmann, T.I., et al., *NAD⁺ biosynthesis and salvage - a phylogenetic perspective*. FEBS J, 2012. **279**(18): p. 3355-3363.
162. Kurnasov, O., et al., *NAD Biosynthesis: Identification of the Tryptophan to Quinolate Pathway in Bacteria*. Chemistry & Biology, 2003. **10**(12): p. 1195-1204.
163. Sakuraba, H., et al., *L-Aspartate oxidase is present in the anaerobic hyperthermophilic archaeon Pyrococcus horikoshii OT-3: characteristics and role in the de novo biosynthesis of nicotinamide adenine dinucleotide proposed by genome sequencing*. Extremophiles, 2002. **6**(4): p. 275-281.
164. Bi, J., H. Wang, and J. Xie, *Comparative genomics of NAD(P) biosynthesis and novel antibiotic drug targets*. Journal of Cellular Physiology, 2011. **226**(2): p. 331-340.
165. 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.
166. Kawai, S., et al., *Molecular characterization of Escherichia coli NAD kinase*. European Journal of Biochemistry, 2001. **268**(15): p. 4359-4365.
167. Kawai, S., et al., *Inorganic Polyphosphate/ATP-NAD Kinase of Micrococcus flavus and Mycobacterium tuberculosis H37Rv*. Biochemical and Biophysical Research Communications, 2000. **276**(1): p. 57-63.
168. Ochiai, A., et al., *Overexpression, purification, and characterization of ATP-NAD kinase of Sphingomonas sp. A1*. Protein Expression and Purification, 2004. **36**(1): p. 124-130.
169. Butler, J.R. and E.T. McGuinness, *Candida utilis NAD⁺⁺ kinase: Purification, properties and affinity gel studies*. International Journal of Biochemistry, 1982. **14**(9): p. 839-844.
170. Tseng, Y.-M., B.G. Harris, and M.K. Jacobson, *Isolation and characterization of yeast nicotinamide adenine dinucleotide kinase*. Biochimica et Biophysica Acta (BBA) - Enzymology, 1979. **568**(1): p. 205-214.
171. Kobayashi, K., et al., *Essential Bacillus subtilis genes*. Proceedings of the National Academy of Sciences, 2003. **100**(8): p. 4678-4683.
172. Sassetti, C.M., D.H. Boyd, and E.J. Rubin, *Genes required for mycobacterial growth defined by high density mutagenesis*. Molecular Microbiology, 2003. **48**(1): p. 77-84.

Appendices

173. Szaszák, M., et al., *Fluorescence Lifetime Imaging Unravels C. trachomatis Metabolism and Its Crosstalk with the Host Cell*. PLoS Pathog, 2011. **7**(7): p. e1002108.
174. Kornberg, A., *Enzymatic synthesis of triphosphopyridine nucleotide*. Journal of Biological Chemistry, 1950. **182**: p. 805-813.
175. Shi, F., Y. Li, and X. Wang, *Molecular properties, functions, and potential applications of NAD kinases*. Acta Biochimica et Biophysica Sinica, 2009. **41**(5): p. 352-361.
176. Harold, F.M., *The Vital Force: A Study of Bioenergetics*, 1986, New York: W. H. Freeman.
177. Amador-Noguez, D., et al., *Metabolome remodeling during the acidogenic-solventogenic transition in Clostridium acetobutylicum*. Appl Environ Microbiol, 2011. **77**(22): p. 7984-97.
178. Andersen, K.B. and K. von Meyenburg, *Charges of nicotinamide adenine nucleotides and adenylate energy charge as regulatory parameters of the metabolism in Escherichia coli*. J Biol Chem, 1977. **252**(12): p. 4151-6.
179. Bennett, B.D., et al., *Absolute metabolite concentrations and implied enzyme active site occupancy in Escherichia coli*. Nat Chem Biol, 2009. **5**(8): p. 593-9.
180. Veech, R.L., L.V. Eggleston, and H.A. Krebs, *The redox state of free nicotinamide-adenine dinucleotide phosphate in the cytoplasm of rat liver*. Biochemical Journal, 1969. **115**(4): p. 609-619.
181. Fuhrer, T. and U. Sauer, *Different Biochemical Mechanisms Ensure Network-Wide Balancing of Reducing Equivalents in Microbial Metabolism*. Journal of Bacteriology, 2009. **191**(7): p. 2112-2121.
182. Zhao, J., et al., *Effect of zwf gene knockout on the metabolism of Escherichia coli grown on glucose or acetate*. Metabolic Engineering, 2004. **6**(2): p. 164-74.
183. Dauner, M., T. Storni, and U. Sauer, *Bacillus subtilis metabolism and energetics in carbon-limited and excess-carbon chemostat culture*. Journal of Bacteriology, 2001. **183**(24): p. 7308-7317.
184. Christiansen, T., B. Christensen, and J. Nielsen, *Metabolic network analysis of Bacillus clausii on minimal and semirich medium using ¹³C-labeled glucose*. Metabolic Engineering, 2002. **4**(2): p. 159-69.
185. Summers, M.L., et al., *Genetic evidence of a major role for glucose-6-phosphate dehydrogenase in nitrogen fixation and dark growth of the cyanobacterium Nostoc sp. strain ATCC 29133*. Journal of Bacteriology, 1995. **177**(21): p. 6184-94.
186. Obanye, A.I.C., et al., *Correlation between carbon flux through the pentose phosphate pathway and production of the antibiotic methylenomycin in Streptomyces coelicolor A3(2)*. Microbiology, 1996. **142**(1): p. 133-137.
187. Bartek, T., et al., *Comparative ¹³C metabolic flux analysis of pyruvate dehydrogenase complex-deficient, L-valine-producing Corynebacterium glutamicum*. Applied and Environmental Microbiology, 2011. **77**(18): p. 6644-52.
188. Celton, M., et al., *A constraint-based model analysis of the metabolic consequences of increased NADPH oxidation in Saccharomyces cerevisiae*. Metabolic Engineering, 2012. **14**(4): p. 366-79.
189. Marx, A., et al., *Response of the central metabolism in Corynebacterium glutamicum to the use of an NADH-dependent glutamate dehydrogenase*. Metabolic Engineering, 1999. **1**(1): p. 35-48.
190. Jonsbu, E., B. Christensen, and J. Nielsen, *Changes of in vivo fluxes through central metabolic pathways during the production of nystatin by Streptomyces noursei in batch culture*. Applied Microbiology and Biotechnology, 2001. **56**(1-2): p. 93-100.
191. Kletzien, R.F., P.K. Harris, and L.A. Foellmi, *Glucose-6-phosphate dehydrogenase: a "housekeeping" enzyme subject to tissue-specific regulation by hormones, nutrients, and oxidant stress*. FASEB journal : official publication of the Federation of American Societies for Experimental Biology, 1994. **8**(2): p. 174-81.
192. Saliola, M., et al., *Intracellular NADPH levels affect the oligomeric state of the glucose 6-phosphate dehydrogenase*. Eukaryotic cell, 2012. **11**(12): p. 1503-11.
193. Moritz, B., et al., *Kinetic properties of the glucose-6-phosphate and 6-phosphogluconate dehydrogenases from Corynebacterium glutamicum and their application for predicting pentose phosphate pathway flux in vivo*. European journal of biochemistry / FEBS, 2000. **267**(12): p. 3442-52.
194. Conway, T., *The Entner-Doudoroff pathway: history, physiology and molecular biology*. FEMS microbiology reviews, 1992. **9**(1): p. 1-27.
195. Fabris, M., et al., *The metabolic blueprint of Phaeodactylum tricornutum reveals a eukaryotic Entner-Doudoroff glycolytic pathway*. The Plant journal : for cell and molecular biology, 2012. **70**(6): p. 1004-14.
196. Fuhrer, T., E. Fischer, and U. Sauer, *Experimental identification and quantification of glucose metabolism in seven bacterial species*. Journal of Bacteriology, 2005. **187**(5): p. 1581-90.
197. Stournaras, C., P. Maurer, and G. Kurz, *6-phospho-D-gluconate dehydrogenase from Pseudomonas fluorescens. Properties and subunit structure*. European journal of biochemistry / FEBS, 1983. **130**(2): p. 391-6.
198. Sung, J.Y. and Y.N. Lee, *Isoforms of glucose 6-phosphate dehydrogenase in Deinococcus radiophilus*. Journal of microbiology (Seoul, Korea), 2007. **45**(4): p. 318-25.
199. Purwantini, E., T.P. Gillis, and L. Daniels, *Presence of F420-dependent glucose-6-phosphate dehydrogenase in Mycobacterium and Nocardia species, but absence from Streptomyces and Corynebacterium species and methanogenic Archaea*. Fems Microbiology Letters, 1997. **146**(1): p. 129-34.
200. Levy, H.R., *Glucose-6-phosphate dehydrogenases*. Advances in enzymology and related areas of molecular biology, 1979. **48**: p. 97-192.
201. Olavarria, K., D. Valdés, and R. Cabrera, *The cofactor preference of glucose-6-phosphate dehydrogenase from Escherichia coli—modeling the physiological production of reduced cofactors*. FEBS Journal, 2012. **279**(13): p. 2296-2309.
202. Butler, M.J., et al., *Engineering of primary carbon metabolism for improved antibiotic production in Streptomyces lividans*. Applied and Environmental Microbiology, 2002. **68**(10): p. 4731-9.
203. Singh, R., et al., *Aluminum-tolerant Pseudomonas fluorescens: ROS toxicity and enhanced NADPH production*. Extremophiles : life under extreme conditions, 2005. **9**(5): p. 367-73.
204. Zamboni, N., et al., *The Bacillus subtilis yajl Gene Encodes the NADP+-Dependent 6-P-Gluconate Dehydrogenase in the Pentose Phosphate Pathway*. Journal of Bacteriology, 2004. **186**(14): p. 4528-4534.

205. Purwantini, E. and L. Daniels, *Purification of a novel coenzyme F420-dependent glucose-6-phosphate dehydrogenase from Mycobacterium smegmatis*. Journal of Bacteriology, 1996. **178**(10): p. 2861-6.
206. Choi, J.-C., H.-D. Shin, and Y.-H. Lee, *Modulation of 3-hydroxyvalerate molar fraction in poly(3-hydroxybutyrate-3-hydroxyvalerate) using Ralstonia eutropha transformant co-amplifying phbC and NADPH generation-related zwf genes*. Enzyme and Microbial Technology, 2003. **32**(1): p. 178-185.
207. Lee, W.-H., et al., *Enhanced production of ϵ -caprolactone by overexpression of NADPH-regenerating glucose 6-phosphate dehydrogenase in recombinant Escherichia coli harboring cyclohexanone monooxygenase gene*. Applied Microbiology and Biotechnology, 2007. **76**(2): p. 329-338.
208. Lim, S.-J., et al., *Amplification of the NADPH-related genes zwf and gnd for the oddball biosynthesis of PHB in an E. coli transformant harboring a cloned phbCAB operon*. Journal of Bioscience and Bioengineering, 2002. **93**(6): p. 543-549.
209. Marx, A., et al., *Metabolic phenotype of phosphoglucose isomerase mutants of Corynebacterium glutamicum*. Journal of Biotechnology, 2003. **104**(1-3): p. 185-197.
210. Chin, J.W. and P.C. Cirino, *Improved NADPH Supply for Xylitol Production by Engineered Escherichia Coli with Glycolytic Mutations*. Biotechnology Progress, 2011. **27**(2): p. 333-341.
211. Wang, Y., K.-Y. San, and G.N. Bennett, *Improvement of NADPH bioavailability in Escherichia coli through the use of phosphofructokinase deficient strains*. Applied Microbiology and Biotechnology, 2013. **97**(15): p. 6883-93.
212. Becker, J., et al., *Amplified Expression of Fructose 1,6-Bisphosphatase in Corynebacterium glutamicum Increases In Vivo Flux through the Pentose Phosphate Pathway and Lysine Production on Different Carbon Sources*. Applied and Environmental Microbiology, 2005. **71**(12): p. 8587-8596.
213. Zhang, J.-D., et al., *Synthesis of optically pure S-sulfoxide by Escherichia coli transformant cells coexpressing the P450 monooxygenase and glucose dehydrogenase genes*. Journal of industrial microbiology & biotechnology, 2011. **38**(5): p. 633-41.
214. Lee, J.-N., H.-D. Shin, and Y.-H. Lee, *Metabolic Engineering of Pentose Phosphate Pathway in Ralstonia eutropha for Enhanced Biosynthesis of Poly- δ -hydroxybutyrate*. Biotechnology Progress, 2003. **19**(5): p. 1444-1449.
215. Smith, K.M., K.-M. Cho, and J.C. Liao, *Engineering Corynebacterium glutamicum for isobutanol production*. Applied Microbiology and Biotechnology, 2010. **87**(3): p. 1045-55.
216. R Poulsen, B., et al., *Increased NADPH concentration obtained by metabolic engineering of the pentose phosphate pathway in Aspergillus niger*. The FEBS journal, 2005. **272**(6): p. 1313-25.
217. Nicolas, C., et al., *Response of the central metabolism of Escherichia coli to modified expression of the gene encoding the glucose-6-phosphate dehydrogenase*. FEBS Letters, 2007. **581**(20): p. 3771-6.
218. Hua, Q., et al., *Responses of the Central Metabolism in Escherichia coli to Phosphoglucose Isomerase and Glucose-6-Phosphate Dehydrogenase Knockouts*. Journal of Bacteriology, 2003. **185**(24): p. 7053-7067.
219. Richhardt, J., S. Bringer, and M. Bott, *Role of the pentose phosphate pathway and the Entner-Doudoroff pathway in glucose metabolism of Gluconobacter oxydans 621H*. Applied Microbiology and Biotechnology, 2013. **97**(10): p. 4315-23.
220. Zhao, J., et al., *Global metabolic response of Escherichia coli to gnd or zwf gene-knockout, based on ¹³C-labeling experiments and the measurement of enzyme activities*. Applied Microbiology and Biotechnology, 2004. **64**(1): p. 91-8.
221. Jiao, Z., et al., *Analysis of metabolic and physiological responses to gnd knockout in Escherichia coli by using C-13 tracer experiment and enzyme activity measurement*. Fems Microbiology Letters, 2003. **220**(2): p. 295-301.
222. Mailloux, R.J., et al., *The Tricarboxylic Acid Cycle, an Ancient Metabolic Network with a Novel Twist*. Plos One, 2007. **2**(8): p. e690.
223. Muro-Pastor, M.I., J.C. Reyes, and F.J. Florencio, *Ammonium assimilation in cyanobacteria*. Photosynthesis research, 2005. **83**(2): p. 135-50.
224. Paul, L., et al., *Integration of regulatory signals through involvement of multiple global regulators: control of the Escherichia coli gltBDF operon by Lrp, IHF, Crp, and ArgR*. BMC Microbiology, 2007. **7**(1): p. 2.
225. Murakami, K., et al., *Role of metal cations in the regulation of NADP-linked isocitrate dehydrogenase from porcine heart*. Biometals : an international journal on the role of metal ions in biology, biochemistry, and medicine, 1997. **10**(3): p. 169-74.
226. Yoon, J.-J., T. Hattori, and M. Shimada, *Purification and characterization of NADP-linked isocitrate dehydrogenase from the copper-tolerant wood-rotting basidiomycete Fomitopsis palustris*. Bioscience, Biotechnology, and Biochemistry, 2003. **67**(1): p. 114-20.
227. Ogawa, T., et al., *Role of phosphoenolpyruvate in the NADP-isocitrate dehydrogenase and isocitrate lyase reaction in Escherichia coli*. Journal of Bacteriology, 2007. **189**(3): p. 1176-8.
228. Cozzzone, A.J., *Regulation of acetate metabolism by protein phosphorylation in enteric bacteria*. Annual review of microbiology, 1998. **52**: p. 127-64.
229. Walsh, K. and D.E. Koshland, *Branch point control by the phosphorylation state of isocitrate dehydrogenase. A quantitative examination of fluxes during a regulatory transition*. Journal of Biological Chemistry, 1985. **260**(14): p. 8430-8437.
230. Beh, M., et al., *Enzymes of the reductive citric acid cycle in the autotrophic eubacterium Aquifex pyrophilus and in the archaeobacterium Thermoproteus neutrophilus*. Archives of Microbiology, 1993. **160**(4): p. 306-311.
231. Kanao, T., et al., *Characterization of isocitrate dehydrogenase from the green sulfur bacterium Chlorobium limicola. A carbon dioxide-fixing enzyme in the reductive tricarboxylic acid cycle*. European journal of biochemistry / FEBS, 2002. **269**(7): p. 1926-31.
232. Shiba, H., et al., *The CO₂ assimilation via the reductive tricarboxylic acid cycle in an obligately autotrophic, aerobic hydrogen-oxidizing bacterium, Hydrogenobacter thermophilus*. Archives of Microbiology, 1985. **141**(3): p. 198-203.
233. Aoshima, M., M. Ishii, and Y. Igarashi, *A novel biotin protein required for reductive carboxylation of 2-oxoglutarate by isocitrate dehydrogenase in Hydrogenobacter thermophilus TK-6*. Molecular Microbiology, 2004. **51**(3): p. 791-8.
234. Chen, R. and S.S. Jeong, *Functional prediction: identification of protein orthologs and paralogs*. Protein science : a publication of the Protein Society, 2000. **9**(12): p. 2344-53.
235. Inoue, H., et al., *Biochemical and molecular characterization of the NAD(+)-dependent isocitrate dehydrogenase from the chemolithotroph Acidithiobacillus thiooxidans*. Fems Microbiology Letters, 2002. **214**(1): p. 127-32.

Appendices

236. Steen, I.H., et al., *Comparison of isocitrate dehydrogenase from three hyperthermophiles reveals differences in thermostability, cofactor specificity, oligomeric state, and phylogenetic affiliation*. The Journal of biological chemistry, 2001. **276**(47): p. 43924-31.
237. Stokke, R., et al., *Biochemical characterization of isocitrate dehydrogenase from Methylococcus capsulatus reveals a unique NAD⁺-dependent homotetrameric enzyme*. Archives of Microbiology, 2007. **187**(5): p. 361-70.
238. Wang, P., et al., *Enzymatic characterization of isocitrate dehydrogenase from an emerging zoonotic pathogen Streptococcus suis*. Biochimie, 2011. **93**(9): p. 1470-5.
239. Wang, P., M. Jin, and G. Zhu, *Biochemical and molecular characterization of NAD⁺-dependent isocitrate dehydrogenase from the ethanologenic bacterium Zymomonas mobilis*. Fems Microbiology Letters, 2012. **327**(2): p. 134-41.
240. Wang, P., et al., *Isocitrate Dehydrogenase from Streptococcus mutans: Biochemical Properties and Evaluation of a Putative Phosphorylation Site at Ser102*. Plos One, 2013. **8**(3): p. e58918.
241. Lloyd, A.J. and D. Weitzman, *Purification and characterization of NAD-linked isocitrate dehydrogenase from Methylophilus methylotrophus*. Biochemical Society Transactions, 1988. **16**(5): p. 871-872.
242. Zhu, G., G.B. Golding, and A.M. Dean, *The Selective Cause of an Ancient Adaptation*. Science, 2005. **307**(5713): p. 1279-1282.
243. Aoshima, M. and Y. Igarashi, *Nondecarboxylating and decarboxylating isocitrate dehydrogenases: oxalosuccinate reductase as an ancestral form of isocitrate dehydrogenase*. Journal of Bacteriology, 2008. **190**(6): p. 2050-5.
244. Steen, I.H., T. Lien, and N.K. Birkeland, *Biochemical and phylogenetic characterization of isocitrate dehydrogenase from a hyperthermophilic archaeon, Archaeoglobus fulgidus*. Archives of Microbiology, 1997. **168**(5): p. 412-20.
245. Banerjee, S., et al., *Comparison of Mycobacterium tuberculosis isocitrate dehydrogenases (ICD-1 and ICD-2) reveals differences in coenzyme affinity, oligomeric state, pH tolerance and phylogenetic affiliation*. BMC Biochemistry, 2005. **6**(1): p. 1-14.
246. Lee, W.-H., et al., *Enhanced production of GDP-L-fucose by overexpression of NADPH regenerator in recombinant Escherichia coli*. Applied Microbiology and Biotechnology, 2011. **91**(4): p. 967-76.
247. Eikmanns, B.J., D. Rittmann, and H. Sahm, *Cloning, sequence analysis, expression, and inactivation of the Corynebacterium glutamicum icd gene encoding isocitrate dehydrogenase and biochemical characterization of the enzyme*. Journal of Bacteriology, 1995. **177**(3): p. 774-82.
248. Becker, J., et al., *Metabolic engineering of the tricarboxylic acid cycle for improved lysine production by Corynebacterium glutamicum*. Applied and Environmental Microbiology, 2009. **75**(24): p. 7866-9.
249. Kabir, M.M. and K. Shimizu, *Metabolic regulation analysis of icd-gene knockout Escherichia coli based on 2D electrophoresis with MALDI-TOF mass spectrometry and enzyme activity measurements*. Applied Microbiology and Biotechnology, 2004. **65**(1): p. 84-96.
250. McDermott, T.R. and M.L. Kahn, *Cloning and mutagenesis of the Rhizobium meliloti isocitrate dehydrogenase gene*. Journal of Bacteriology, 1992. **174**(14): p. 4790-4797.
251. Muro-Pastor, M.I. and F.J. Florencio, *NADP (+)-isocitrate dehydrogenase from the cyanobacterium Anabaena sp. strain PCC 7120: purification and characterization of the enzyme and cloning, sequencing, and disruption of the icd gene*. Journal of Bacteriology, 1994. **176**(9): p. 2718-2726.
252. Lakshmi, T.M. and R.B. Helling, *Selection for citrate synthase deficiency in icd mutants of Escherichia coli*. Journal of Bacteriology, 1976. **127**(1): p. 76-83.
253. Frenkel, R., *Regulation and physiological functions of malic enzymes*. Current topics in cellular regulation, 1975. **9**: p. 157-81.
254. Voegelé, R.T., M.J. Mitsch, and T.M. Finan, *Characterization of two members of a novel malic enzyme class*. Biochimica et biophysica acta, 1999. **1432**(2): p. 275-85.
255. Lerondel, G., et al., *YtsI has the major physiological role of the four paralagous malic enzyme isoforms in Bacillus subtilis*. Journal of Bacteriology, 2006. **188**(13): p. 4727-36.
256. Sauer, U. and B.J. Eikmanns, *The PEP-pyruvate-oxaloacetate node as the switch point for carbon flux distribution in bacteria*. FEMS microbiology reviews, 2005. **29**(4): p. 765-794.
257. Espariz, M., et al., *Identification of malic and soluble oxaloacetate decarboxylase enzymes in Enterococcus faecalis*. The FEBS journal, 2011. **278**(12): p. 2140-51.
258. Kawai, S., et al., *Purification and characterization of a malic enzyme from the ruminal bacterium Streptococcus bovis ATCC 15352 and cloning and sequencing of its gene*. Applied and Environmental Microbiology, 1996. **62**(8): p. 2692-700.
259. Landete, J.M., et al., *Requirement of the Lactobacillus casei MaeKR two-component system for L-malic acid utilization via a malic enzyme pathway*. Applied and Environmental Microbiology, 2010. **76**(1): p. 84-95.
260. London, J., E.Y. Meyer, and S. Kulczyk, *Comparative biochemical and immunological study of malic enzyme from two species of lactic acid bacteria: evolutionary implications*. Journal of Bacteriology, 1971. **106**(1): p. 126-37.
261. Bologna, F.P., C.S. Andreo, and M.F. Drincovich, *Escherichia coli malic enzymes: two isoforms with substantial differences in kinetic properties, metabolic regulation, and structure*. Journal of Bacteriology, 2007. **189**(16): p. 5937-46.
262. Iwakura, M., et al., *Studies on regulatory functions of malic enzymes. V. Comparative studies of malic enzymes in bacteria*. Journal of biochemistry, 1978. **83**(5): p. 1387-94.
263. Morimoto, Y., et al., *Directed evolution of the thermotolerant malic enzyme for improved malate production*. Journal of Bioscience and Bioengineering, 2014. **117**(2): p. 147-52.
264. Chang, G.-G. and L. Tong, *Structure and function of malic enzymes, a new class of oxidative decarboxylases*. Biochemistry, 2003. **42**(44): p. 12721-33.
265. Hsieh, J.-Y. and H.-C. Hung, *Engineering of the cofactor specificities and isoform-specific inhibition of malic enzyme*. The Journal of biological chemistry, 2009. **284**(7): p. 4536-44.
266. Kuo, C.C., et al., *Lysine residues 162 and 340 are involved in the catalysis and coenzyme binding of NADP(+)-dependent malic enzyme from pigeon*. Biochemical and Biophysical Research Communications, 2000. **270**(3): p. 821-5.
267. Yang, Z., et al., *Structural studies of the pigeon cytosolic NADP(+)-dependent malic enzyme*. Protein science : a publication of the Protein Society, 2002. **11**(2): p. 332-41.

268. Gourdon, P., et al., *Cloning of the malic enzyme gene from Corynebacterium glutamicum and role of the enzyme in lactate metabolism*. Applied and Environmental Microbiology, 2000. **66**(7): p. 2981-7.
269. Wang, B., et al., *Biochemical properties and physiological roles of NADP-dependent malic enzyme in Escherichia coli*. Journal of microbiology (Seoul, Korea), 2011. **49**(5): p. 797-802.
270. Driscoll, B.T. and T.M. Finan, *Properties of NAD(+)- and NADP(+)-dependent malic enzymes of Rhizobium (Sinorhizobium) meliloti and differential expression of their genes in nitrogen-fixing bacteroids*. Microbiology (Reading, England), 1997. **143** (Pt 2): p. 489-98.
271. Garrido-Pertierra, A., et al., *Properties and function of malate enzyme from Pseudomonas putida*. Biochimie, 1983. **65**(11-12): p. 629-35.
272. Lamed, R. and J.G. Zeikus, *Thermostable, ammonium-activated malic enzyme of Clostridium thermocellum*. Biochimica et biophysica acta, 1981. **660**(2): p. 251-5.
273. Owen, O.E., S.C. Kalhan, and R.W. Hanson, *The Key Role of Anaplerosis and Cataplerosis for Citric Acid Cycle Function*. Journal of Biological Chemistry, 2002. **277**(34): p. 30409-30412.
274. Matula, T.I., I.J. McDonald, and S.M. Martin, *CO₂ fixation by malic enzyme in a species of Micrococcus*. Biochemical and Biophysical Research Communications, 1969. **34**(6): p. 795-802.
275. Fischer, E. and U. Sauer, *A Novel Metabolic Cycle Catalyzes Glucose Oxidation and Anaplerosis in Hungry Escherichia coli*. Journal of Biological Chemistry, 2003. **278**(47): p. 46446-46451.
276. Netzer, R., et al., *Roles of pyruvate kinase and malic enzyme in Corynebacterium glutamicum for growth on carbon sources requiring gluconeogenesis*. Archives of Microbiology, 2004. **182**(5): p. 354-63.
277. Petersen, S., et al., *In Vivo Quantification of Parallel and Bidirectional Fluxes in the Anaplerosis of Corynebacterium glutamicum*. Journal of Biological Chemistry, 2000. **275**(46): p. 35932-35941.
278. Kao, K.C., L.M. Tran, and J.C. Liao, *A global regulatory role of gluconeogenic genes in Escherichia coli revealed by transcriptome network analysis*. The Journal of biological chemistry, 2005. **280**(43): p. 36079-87.
279. Velayudhan, J. and D.J. Kelly, *Analysis of gluconeogenic and anaplerotic enzymes in Campylobacter jejuni: an essential role for phosphoenolpyruvate carboxykinase*. Microbiology (Reading, England), 2002. **148**(Pt 3): p. 685-94.
280. Hansen, E.J. and E. Juni, *Isolation of mutants of Escherichia coli lacking NAD- and NADP-linked malic*. Biochemical and Biophysical Research Communications, 1975. **65**(2): p. 559-66.
281. Oh, M.-K., et al., *Global expression profiling of acetate-grown Escherichia coli*. The Journal of biological chemistry, 2002. **277**(15): p. 13175-83.
282. Meyer, F.M. and J. Stulke, *Malate metabolism in Bacillus subtilis: distinct roles for three classes of malate-oxidizing enzymes*. Fems Microbiology Letters, 2013. **339**(1): p. 17-22.
283. Riedel, C., et al., *Characterization of the phosphoenolpyruvate carboxykinase gene from Corynebacterium glutamicum and significance of the enzyme for growth and amino acid production*. Journal of molecular microbiology and biotechnology, 2001. **3**(4): p. 573-83.
284. Zhou, J., et al., *Atypical Glycolysis in Clostridium thermocellum*. Applied and Environmental Microbiology, 2013. **79**(9): p. 3000-3008.
285. Matsuda, F., et al., *Increased isobutanol production in Saccharomyces cerevisiae by eliminating competing pathways and resolving cofactor imbalance*. Microbial Cell Factories, 2013. **12**: p. 119.
286. Moreira dos Santos, M., et al., *Manipulation of malic enzyme in Saccharomyces cerevisiae for increasing NADPH production capacity aerobically in different cellular compartments*. Metabolic Engineering, 2004. **6**(4): p. 352-63.
287. Hong, S.H. and S.Y. Lee, *Metabolic flux analysis for succinic acid production by recombinant Escherichia coli with amplified malic enzyme activity*. Biotechnology and Bioengineering, 2001. **74**(2): p. 89-95.
288. Shin, J.-A., et al., *5-aminolevulinic acid biosynthesis in Escherichia coli coexpressing NADP-dependent malic enzyme and 5-aminolevulinic synthase*. Journal of microbiology and biotechnology, 2007. **17**(9): p. 1579-84.
289. Stols, L. and M.I. Donnelly, *Production of succinic acid through overexpression of NAD(+)-dependent malic enzyme in an Escherichia coli mutant*. Applied and Environmental Microbiology, 1997. **63**(7): p. 2695-701.
290. Zheng, H., et al., *Production of L-malic acid with fixation of HCO₃(-) by malic enzyme-catalyzed reaction based on regeneration of coenzyme on electrode modified by layer-by-layer self-assembly method*. Journal of Bioscience and Bioengineering, 2009. **107**(1): p. 16-20.
291. Sahn, H., L. Eggeling, and A.A. de Graaf, *Pathway analysis and metabolic engineering in Corynebacterium glutamicum*. Biological chemistry, 2000. **381**(9-10): p. 899-910.
292. Jain, M., R.D. K., and J.G. Zeikus, *High-value organic acids fermentation - emerging processes and products*, in *Bioprocess engineering: the first generation*, T.K.G. (ed.), Editor 1989, Ellis Horwood Ltd.: chichester, England. p. 366-389.
293. Szmant, H.H., *Organic building blocks of the chemical industry*, 1989, New York: John Wiley & Sons.
294. Sophos, N.A., et al., *Aldehyde dehydrogenase gene superfamily: the 2000 update*. Chemico-biological interactions, 2001. **130-132**(1-3): p. 323-37.
295. Sophos, N.A. and V. Vasiliou, *Aldehyde dehydrogenase gene superfamily: the 2002 update*. Chemico-biological interactions, 2003. **143-144**: p. 5-22.
296. Lindahl, R., *Aldehyde dehydrogenases and their role in carcinogenesis*. Critical reviews in biochemistry and molecular biology, 1992. **27**(4-5): p. 283-335.
297. Vasiliou, V., A. Pappa, and D.R. Petersen, *Role of aldehyde dehydrogenases in endogenous and xenobiotic metabolism*. Chemico-biological interactions, 2000. **129**(1-2): p. 1-19.
298. Esser, D., et al., *Unraveling the function of paralogs of the aldehyde dehydrogenase super family from Sulfolobus solfataricus*. Extremophiles : life under extreme conditions, 2013. **17**(2): p. 205-216.
299. Talfournier, F., C. Stines-Chaumeil, and G. Branlant, *Methylmalonate-semialdehyde dehydrogenase from Bacillus subtilis: substrate specificity and coenzyme A binding*. The Journal of biological chemistry, 2011. **286**(25): p. 21971-81.

Appendices

300. Jackson, B., et al., *Update on the aldehyde dehydrogenase gene (ALDH) superfamily*. Human genomics, 2011. **5**(4): p. 283-303.
301. Marchitti, S.A., et al., *Non-P450 aldehyde oxidizing enzymes: the aldehyde dehydrogenase superfamily*. Expert opinion on drug metabolism & toxicology, 2008. **4**(6): p. 697-720.
302. Kelly, G.J. and M. Gibbs, *Nonreversible d-Glyceraldehyde 3-Phosphate Dehydrogenase of Plant Tissues*. Plant Physiology, 1973. **52**(2): p. 111-8.
303. Mateos, M. and A. Serrano, *Occurrence of phosphorylating and non-phosphorylating NADP⁺-dependent glyceraldehyde-3-phosphate dehydrogenases in photosynthetic organisms*. Plant Science, 1992. **84**(2): p. 163-170.
304. Boyd, D.A., D.G. Cvitkovitch, and I.R. Hamilton, *Sequence, expression, and function of the gene for the nonphosphorylating, NADP-dependent glyceraldehyde-3-phosphate dehydrogenase of Streptococcus mutans*. Journal of Bacteriology, 1995. **177**(10): p. 2622-7.
305. Crow, V.L. and C.L. Wittenberger, *Separation and properties of NAD⁺- and NADP⁺-dependent glyceraldehyde-3-phosphate dehydrogenases from Streptococcus mutans*. The Journal of biological chemistry, 1979. **254**(4): p. 1134-42.
306. Brunner, N.A., et al., *NAD⁺-dependent glyceraldehyde-3-phosphate dehydrogenase from Thermoproteus tenax. The first identified archaeal member of the aldehyde dehydrogenase superfamily is a glycolytic enzyme with unusual regulatory properties*. The Journal of biological chemistry, 1998. **273**(11): p. 6149-56.
307. 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 : life under extreme conditions, 2008. **12**(1): p. 75-88.
308. Iddar, A., et al., *Widespread occurrence of non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase among gram-positive bacteria*. International microbiology : the official journal of the Spanish Society for Microbiology, 2005. **8**(4): p. 251-8.
309. Iddar, A., et al., *Expression, purification, and characterization of recombinant nonphosphorylating NADP-dependent glyceraldehyde-3-phosphate dehydrogenase from Clostridium acetobutylicum*. Protein Expression and Purification, 2002. **25**(3): p. 519-26.
310. Iddar, A., et al., *Purification of recombinant non-phosphorylating NADP-dependent glyceraldehyde-3-phosphate dehydrogenase from Streptococcus pyogenes expressed in E. coli*. Molecular and cellular biochemistry, 2003. **247**(1-2): p. 195-203.
311. Ahmed, H., et al., *The semi-phosphorylative Entner-Doudoroff pathway in hyperthermophilic archaea: a re-evaluation*. The Biochemical journal, 2005. **390**(Pt 2): p. 529-40.
312. 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.
313. Brunner, N.A. and R. Hensel, *Nonphosphorylating glyceraldehyde-3-phosphate dehydrogenase from Thermoproteus tenax*. Methods in enzymology, 2001. **331**: p. 117-31.
314. 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. 815-28.
315. Ito, F., et al., *Comparative analysis of two glyceraldehyde-3-phosphate dehydrogenases from a thermoacidophilic archaeon, Sulfolobus tokodaii*. FEBS Letters, 2012. **586**(19): p. 3097-103.
316. Arutyunov, D., et al., *An unusual effect of NADP⁺ on the thermostability of the nonphosphorylating glyceraldehyde-3-phosphate dehydrogenase from Streptococcus mutans*. Biochemistry and cell biology = Biochimie et biologie cellulaire, 2013. **91**(5): p. 295-302.
317. Centeno-Leija, S., et al., *Metabolic and transcriptional response of Escherichia coli with a NADP(+)-dependent glyceraldehyde 3-phosphate dehydrogenase from Streptococcus mutans*. Antonie van Leeuwenhoek, 2013. **104**(6): p. 913-24.
318. Centeno-Leija, S., et al., *Improving poly-3-hydroxybutyrate production in Escherichia coli by combining the increase in the NADPH pool and acetyl-CoA availability*. Antonie van Leeuwenhoek, 2014. **105**(4): p. 687-96.
319. Guo, Z.-P., et al., *Improving ethanol productivity by modification of glycolytic redox factor generation in glycerol-3-phosphate dehydrogenase mutants of an industrial ethanol yeast*. Journal of industrial microbiology & biotechnology, 2011. **38**(8): p. 935-43.
320. Takeno, S., et al., *Engineering of Corynebacterium glutamicum with an NADPH-generating glycolytic pathway for L-lysine production*. Applied and Environmental Microbiology, 2010. **76**(21): p. 7154-60.
321. Brunner, N.A., B. Siebers, and R. Hensel, *Role of two different glyceraldehyde-3-phosphate dehydrogenases in controlling the reversible Embden-Meyerhof-Parnas pathway in Thermoproteus tenax: regulation on protein and transcript level*. Extremophiles : life under extreme conditions, 2001. **5**(2): p. 101-9.
322. Fillinger, S., et al., *Two Glyceraldehyde-3-phosphate Dehydrogenases with Opposite Physiological Roles in a Nonphotosynthetic Bacterium*. Journal of Biological Chemistry, 2000. **275**(19): p. 14031-14037.
323. Koksharova, O., et al., *Genetic and biochemical evidence for distinct key functions of two highly divergent GAPDH genes in catabolic and anabolic carbon flow of the cyanobacterium Synechocystis sp. PCC 6803*. Plant Molecular Biology, 1998. **36**(1): p. 183-194.
324. Schäfer, T. and P. Schönheit, *Gluconeogenesis from pyruvate in the hyperthermophilic archaeon Pyrococcus furiosus: involvement of reactions of the Embden-Meyerhof pathway*. Archives of Microbiology, 1993. **159**(4): p. 354-363.
325. Huang, M., et al., *Multiple Strategies for Metabolic Engineering of Escherichia coli for Efficient Production of Coenzyme Q10*. Chinese Journal of Chemical Engineering, 2011. **19**(2): p. 316-326.
326. Jiang, L.-Y., et al., *Metabolic engineering of Corynebacterium glutamicum for increasing the production of L-ornithine by increasing NADPH availability*. Journal of industrial microbiology & biotechnology, 2013. **40**(10): p. 1143-51.
327. Martinez, I., et al., *Replacing Escherichia coli NAD-dependent glyceraldehyde 3-phosphate dehydrogenase (GAPDH) with a NADP-dependent enzyme from Clostridium acetobutylicum facilitates NADPH dependent pathways*. Metabolic Engineering, 2008. **10**(6): p. 352-9.

328. Wang, Y., K.-Y. San, and G.N. Bennett, *Improvement of NADPH bioavailability in Escherichia coli by replacing NAD(+)-dependent glyceraldehyde-3-phosphate dehydrogenase GapA with NADP (+)-dependent GapB from Bacillus subtilis and addition of NAD kinase*. Journal of industrial microbiology & biotechnology, 2013. **40**(12): p. 1449-60.
329. Bommarreddy, R.R., et al., *A de novo NADPH generation pathway for improving lysine production of Corynebacterium glutamicum by rational design of the coenzyme specificity of glyceraldehyde 3-phosphate dehydrogenase*. Metabolic Engineering, 2014. **25**: p. 30-7.
330. Adachi, O. and M. Ameyama, [26] *d-Glucose dehydrogenase from Gluconobacter su* \square *ydans*, in *Methods in enzymology*, A.W. Willis, Editor 1982, Academic Press. p. 159-163.
331. Lampel, K.A., et al., *Characterization of the developmentally regulated Bacillus subtilis glucose dehydrogenase gene*. Journal of Bacteriology, 1986. **166**(1): p. 238-43.
332. Pauly, H.E. and G. Pfeleiderer, *D-glucose dehydrogenase from Bacillus megaterium M 1286: purification, properties and structure*. Hoppe-Seyler's Zeitschrift fur physiologische Chemie, 1975. **356**(10): p. 1613-23.
333. Eguchi, T., et al., *NADPH regeneration by glucose dehydrogenase from Gluconobacter scleroides for l-leucovorin synthesis*. Bioscience, Biotechnology, and Biochemistry, 1992. **56**(5): p. 701-3.
334. Satoh, Y., et al., *Enzyme-catalyzed poly(3-hydroxybutyrate) synthesis from acetate with CoA recycling and NADPH regeneration in Vitro*. Journal of Bioscience and Bioengineering, 2003. **95**(4): p. 335-41.
335. Ferri, S., K. Kojima, and K. Sode, *Review of glucose oxidases and glucose dehydrogenases: a bird's eye view of glucose sensing enzymes*. Journal of diabetes science and technology, 2011. **5**(5): p. 1068-76.
336. Kataoka, M., et al., *Escherichia coli transformant expressing the glucose dehydrogenase gene from Bacillus megaterium as a cofactor regenerator in a chiral alcohol production system*. Bioscience, Biotechnology, and Biochemistry, 1998. **62**(1): p. 167-9.
337. Kataoka, M., et al., *Stereoselective reduction of ethyl 4-chloro-3-oxobutanoate by Escherichia coli transformant cells coexpressing the aldehyde reductase and glucose dehydrogenase genes*. Applied Microbiology and Biotechnology, 1999. **51**(4): p. 486-90.
338. Kizaki, N., et al., *Synthesis of optically pure ethyl (S)-4-chloro-3-hydroxybutanoate by Escherichia coli transformant cells coexpressing the carbonyl reductase and glucose dehydrogenase genes*. Applied Microbiology and Biotechnology, 2001. **55**(5): p. 590-595.
339. Lu, Y. and L. Mei, *Co-expression of P450 BM3 and glucose dehydrogenase by recombinant Escherichia coli and its application in an NADPH-dependent indigo production system*. Journal of industrial microbiology & biotechnology, 2007. **34**(3): p. 247-53.
340. Park, H.J., et al., *Enantioselective bioconversion using Escherichia coli cells expressing Saccharomyces cerevisiae reductase and Bacillus subtilis glucose dehydrogenase*. Journal of microbiology and biotechnology, 2010. **20**(9): p. 1300-6.
341. Xu, Z., et al., *High-level expression of recombinant glucose dehydrogenase and its application in NADPH regeneration*. Journal of industrial microbiology & biotechnology, 2007. **34**(1): p. 83-90.
342. Yun, H., et al., *Stereospecific synthesis of (R)-2-hydroxy carboxylic acids using recombinant E. coli BL21 overexpressing YiaE from Escherichia coli K12 and glucose dehydrogenase from Bacillus subtilis*. Biotechnology Progress, 2005. **21**(2): p. 366-71.
343. Boonstra, B., et al., *The udhA gene of Escherichia coli encodes a soluble pyridine nucleotide transhydrogenase*. Journal of Bacteriology, 1999. **181**(3): p. 1030-4.
344. Cohen, P.T. and N.O. Kaplan, *Purification and properties of the pyridine nucleotide transhydrogenase from Pseudomonas aeruginosa*. The Journal of biological chemistry, 1970. **245**(11): p. 2825-36.
345. French, C.E., et al., *Cloning, sequence, and properties of the soluble pyridine nucleotide transhydrogenase of Pseudomonas fluorescens*. Journal of Bacteriology, 1997. **179**(8): p. 2761-2765.
346. Voordouw, G., et al., *Structure of Pyridine Nucleotide Transhydrogenase from Azotobacter vinelandii*. European Journal of Biochemistry, 1979. **98**(2): p. 447-454.
347. Bizouarn, T., et al., *The organization of the membrane domain and its interaction with the NADP(H)-binding site in proton-translocating transhydrogenase from E. coli*. Biochimica et Biophysica Acta (BBA) - Bioenergetics, 2002. **1555**(1-3): p. 122-127.
348. Bizouarn, T., et al., *The transmembrane domain and the proton channel in proton-pumping transhydrogenases*. Biochimica et Biophysica Acta (BBA) - Bioenergetics, 2000. **1459**(2-3): p. 284-290.
349. Jackson, J.B., *Proton translocation by transhydrogenase*. FEBS Letters, 2003. **545**(1): p. 18-24.
350. Jackson, J.B., *A review of the binding-change mechanism for proton-translocating transhydrogenase*. Biochimica et biophysica acta, 2012. **1817**(10): p. 1839-46.
351. Jackson, J.B. and U.M. Obiozo, *Proton-Translocating Transhydrogenase in Photosynthetic Bacteria*, in *The Purple Phototrophic Bacteria*, C.N. Hunter, et al., Editors. 2009, Springer Netherlands: Dordrecht. p. 495-508.
352. Pedersen, A., G.B. Karlsson, and J. Rydstrom, *Proton-translocating transhydrogenase: an update of unsolved and controversial issues*. Journal of bioenergetics and biomembranes, 2008. **40**(5): p. 463-73.
353. Hoek, J.B. and J. Rydstrom, *Physiological roles of nicotinamide nucleotide transhydrogenase*. The Biochemical journal, 1988. **254**(1): p. 1-10.
354. Bragg, P.D., P.L. Davies, and C. Hou, *Function of energy-dependent transhydrogenase in Escherichia coli*. Biochemical and Biophysical Research Communications, 1972. **47**(5): p. 1248-55.
355. Jackson, J.B., et al., *Interdomain hydride transfer in proton-translocating transhydrogenase*. Biochimica et biophysica acta, 1998. **1365**(1-2): p. 79-86.
356. Voordouw, G., S.M. van der Vies, and A.P. Themmen, *Why are two different types of pyridine nucleotide transhydrogenase found in living organisms?* European journal of biochemistry / FEBS, 1983. **131**(3): p. 527-33.
357. Canonaco, F., et al., *Metabolic flux response to phosphoglucose isomerase knock-out in Escherichia coli and impact of overexpression of the soluble transhydrogenase UdhA*. Fems Microbiology Letters, 2001. **204**(2): p. 247-52.
358. Zhao, H., et al., *Physiologic roles of soluble pyridine nucleotide transhydrogenase in Escherichia coli as determined by homologous recombination*. Annals of Microbiology, 2008. **58**(2): p. 275-280.

Appendices

359. Angermayr, S.A., M. Paszota, and K.J. Hellingwerf, *Engineering a Cyanobacterial Cell Factory for Production of Lactic Acid*. Applied and Environmental Microbiology, 2012. **78**(19): p. 7098-7106.
360. Boonstra, B., et al., *Cofactor regeneration by a soluble pyridine nucleotide transhydrogenase for biological production of hydromorphone*. Applied and Environmental Microbiology, 2000. **66**(12): p. 5161-6.
361. Qi, H., et al., *Model-driven redox pathway manipulation for improved isobutanol production in Bacillus subtilis complemented with experimental validation and metabolic profiling analysis*. Plos One, 2014. **9**(4): p. e93815.
362. Reddy, G.K., S.N. Lindner, and V.F. Wendisch, *Metabolic Engineering an ATP-neutral EMP pathway in Corynebacterium glutamicum: adaptive point mutation in NADH dehydrogenase restores growth*. Applied and Environmental Microbiology, 2015. **81**: p. 1996-2005.
363. Sanchez, A.M., et al., *Effect of overexpression of a soluble pyridine nucleotide transhydrogenase (UdhA) on the production of poly(3-hydroxybutyrate) in Escherichia coli*. Biotechnology Progress, 2006. **22**(2): p. 420-5.
364. Decorosi, F., et al., *Characterization of a Cr(VI)-sensitive Pseudomonas corrugata 28 mutant impaired in a pyridine nucleotide transhydrogenase gene*. Research in microbiology, 2011. **162**(8): p. 747-55.
365. He, L., et al., *Central metabolic responses to the overproduction of fatty acids in Escherichia coli based on 13C-metabolic flux analysis*. Biotechnology and Bioengineering, 2014. **111**(3): p. 575-585.
366. Jan, J., et al., *Metabolic engineering and transhydrogenase effects on NADPH availability in Escherichia coli*. Biotechnology Progress, 2013. **29**(5): p. 1124-30.
367. Lee, H.C., et al., *High NADPH/NADP+ ratio improves thymidine production by a metabolically engineered Escherichia coli strain*. Journal of Biotechnology, 2010. **149**(1-2): p. 24-32.
368. Chin, J.W., et al., *Analysis of NADPH supply during xylitol production by engineered Escherichia coli*. Biotechnology and Bioengineering, 2009. **102**(1): p. 209-20.
369. Yamauchi, Y., et al., *Enhanced acetic acid and succinic acid production under microaerobic conditions by Corynebacterium glutamicum harboring Escherichia coli transhydrogenase gene pntAB*. The Journal of general and applied microbiology, 2014. **60**(3): p. 112-8.
370. Hickman, J.W., et al., *Link between the membrane-bound pyridine nucleotide transhydrogenase and glutathione-dependent processes in Rhodobacter sphaeroides*. Journal of Bacteriology, 2002. **184**(2): p. 400-9.
371. Imam, S., D.R. Noguera, and T.J. Donohue, *Global insights into energetic and metabolic networks in Rhodobacter sphaeroides*. BMC systems biology, 2013. **7**: p. 89.
372. Cao, Z., et al., *Overexpression and biochemical characterization of soluble pyridine nucleotide transhydrogenase from Escherichia coli*. Fems Microbiology Letters, 2011. **320**(1): p. 9-14.
373. Middleditch, L.E., R.W. Atchison, and A.E. Chung, *Pyridine nucleotide transhydrogenase from Azotobacter vinelandii. Some aspects of its structure*. The Journal of biological chemistry, 1972. **247**(21): p. 6802-9.
374. Whitehead, S.J., et al., *Inhibition of proton-transfer steps in transhydrogenase by transition metal ions*. Biochimica et biophysica acta, 2009. **1787**(10): p. 1276-88.
375. Whitehead, S.J., et al., *Zinc ions selectively inhibit steps associated with binding and release of NAD(P)H during turnover of proton-translocating transhydrogenase*. FEBS Letters, 2005. **579**(13): p. 2863-7.
376. Chang, D.Y., C. Hou, and P.D. Bragg, *Anomalous effect of uncouplers on respiratory chain-linked transhydrogenation in Escherichia coli membranes: evidence for a localized proton pathway?* Archives of Biochemistry and Biophysics, 1992. **293**(2): p. 246-53.
377. Cotton, N.P., et al., *The coupling between protonmotive force and the NAD(P)+ transhydrogenase in chromatophores from photosynthetic bacteria*. European journal of biochemistry / FEBS, 1989. **182**(3): p. 593-603.
378. Lever, T.M., et al., *Purification and properties of the H(+)-nicotinamide nucleotide transhydrogenase from Rhodobacter capsulatus*. European journal of biochemistry / FEBS, 1991. **197**(1): p. 247-55.
379. Hojeberg, B. and J. Rydstrom, *Ca2+-dependent allosteric regulation of nicotinamide nucleotide transhydrogenase from Pseudomonas aeruginosa*. European journal of biochemistry / FEBS, 1977. **77**(2): p. 235-41.
380. Voordouw, G., et al., *Pyridine Nucleotide Transhydrogenase from Azotobacter vinelandii*. European Journal of Biochemistry, 1980. **107**(2): p. 337-344.
381. Bastian, S., et al., *Engineered ketol-acid reductoisomerase and alcohol dehydrogenase enable anaerobic 2-methylpropan-1-ol production at theoretical yield in Escherichia coli*. Metabolic Engineering, 2011. **13**(3): p. 345-52.
382. Rathnasingh, C., et al., *Production of 3-hydroxypropionic acid via malonyl-CoA pathway using recombinant Escherichia coli strains*. Journal of Biotechnology, 2012. **157**(4): p. 633-40.
383. Shi, A., et al., *Activating transhydrogenase and NAD kinase in combination for improving isobutanol production*. Metabolic Engineering, 2013. **16**: p. 1-10.
384. Weckbecker, A. and W. Hummel, *Improved synthesis of chiral alcohols with Escherichia coli cells co-expressing pyridine nucleotide transhydrogenase, NADP+-dependent alcohol dehydrogenase and NAD+-dependent formate dehydrogenase*. Biotechnology letters, 2004. **26**(22): p. 1739-44.
385. Kabus, A., et al., *Expression of the Escherichia coli pntAB genes encoding a membrane-bound transhydrogenase in Corynebacterium glutamicum improves L-lysine formation*. Applied Microbiology and Biotechnology, 2007. **75**(1): p. 47-53.
386. Ceccarelli, E.A., et al., *Functional plasticity and catalytic efficiency in plant and bacterial ferredoxin-NAD(P)H reductases*. Biochimica et biophysica acta, 2004. **1698**(2): p. 155-65.
387. Ma, K. and M.W. Adams, *Ferredoxin:NADP oxidoreductase from Pyrococcus furiosus*. Methods in enzymology, 2001. **334**: p. 40-5.
388. Yan, Z., et al., *Sulfolobus tokodaii ST2133 is characterized as a thioredoxin reductase-like ferredoxin:NADP+ oxidoreductase*. Extremophiles : life under extreme conditions, 2014. **18**(1): p. 99-110.
389. Goss, T. and G. Hanke, *The end of the line: can ferredoxin and ferredoxin NAD(P)H oxidoreductase determine the fate of photosynthetic electrons?* Current protein & peptide science, 2014. **15**(4): p. 385-93.

390. Shin, M. and D.I. Arnon, *Enzymic mechanisms of pyridine nucleotide reduction in chloroplasts*. The Journal of biological chemistry, 1965. **240**: p. 1405-11.
391. Kramer, D.M. and J.R. Evans, *The importance of energy balance in improving photosynthetic productivity*. Plant Physiology, 2011. **155**(1): p. 70-8.
392. Mulo, P., *Chloroplast-targeted ferredoxin-NADP(+) oxidoreductase (FNR): structure, function and location*. Biochimica et biophysica acta, 2011. **1807**(8): p. 927-34.
393. Munekage, Y., et al., *Cyclic electron flow around photosystem I is essential for photosynthesis*. Nature, 2004. **429**(6991): p. 579-82.
394. Shikanai, T., *Cyclic electron transport around photosystem I: genetic approaches*. Annual review of plant biology, 2007. **58**: p. 199-217.
395. Arakaki, A.K., E.A. Ceccarelli, and N. Carrillo, *Plant-type ferredoxin-NADP+ reductases: a basal structural framework and a multiplicity of functions*. FASEB journal : official publication of the Federation of American Societies for Experimental Biology, 1997. **11**(2): p. 133-40.
396. Balconi, E., et al., *The ferredoxin-NADP+ reductase/ferredoxin electron transfer system of Plasmodium falciparum*. The FEBS journal, 2009. **276**(14): p. 3825-36.
397. Giro, M., N. Carrillo, and A.R. Krapp, *Glucose-6-phosphate dehydrogenase and ferredoxin-NADP(H) reductase contribute to damage repair during the soxRS response of Escherichia coli*. Microbiology (Reading, England), 2006. **152**(Pt 4): p. 1119-28.
398. Krapp, A.R., et al., *The flavoenzyme ferredoxin (flavodoxin)-NADP(H) reductase modulates NADP(H) homeostasis during the soxRS response of Escherichia coli*. Journal of Bacteriology, 2002. **184**(5): p. 1474-80.
399. Lee, Y., et al., *Molecular characterization of fprB (ferredoxin-NADP+ reductase) in Pseudomonas putida KT2440*. Journal of microbiology and biotechnology, 2007. **17**(9): p. 1504-12.
400. Lewis, T.A., et al., *Role for ferredoxin:NAD(P)H oxidoreductase (FprA) in sulfate assimilation and siderophore biosynthesis in Pseudomonas*. Journal of Bacteriology, 2013. **195**(17): p. 3876-87.
401. Seeber, F., A. Aliverti, and G. Zanetti, *The plant-type ferredoxin-NADP+ reductase/ferredoxin redox system as a possible drug target against apicomplexan human parasites*. Current pharmaceutical design, 2005. **11**(24): p. 3159-72.
402. Tang, K.-H., H. Yue, and R.E. Blankenship, *Energy metabolism of Helicobacterium modesticaldum during phototrophic and chemotrophic growth*. BMC Microbiology, 2010. **10**: p. 150.
403. Aliverti, A., et al., *Structural and functional diversity of ferredoxin-NADP+ reductases*. Archives of Biochemistry and Biophysics, 2008. **474**(2): p. 283-291.
404. Carrillo, N. and E.A. Ceccarelli, *Open questions in ferredoxin-NADP+ reductase catalytic mechanism*. European journal of biochemistry / FEBS, 2003. **270**(9): p. 1900-15.
405. Karplus, P.A. and H.R. Faber, *Structural Aspects of Plant Ferredoxin : NADP(+) Oxidoreductases*. Photosynthesis research, 2004. **81**(3): p. 303-15.
406. Musumeci, M.A., E.A. Ceccarelli, and D.L. Catalano-Dupuy, *The Plant-Type Ferredoxin-NADP+ Reductases in Advances in Photosynthesis - Fundamental Aspects*, D.M. Najafpour, Editor 2012, InTech: Rijeka. p. 539-562.
407. Komori, H., et al., *Crystal structure analysis of Bacillus subtilis ferredoxin-NADP(+) oxidoreductase and the structural basis for its substrate selectivity*. Protein science : a publication of the Protein Society, 2010. **19**(12): p. 2279-90.
408. Muraki, N., et al., *Asymmetric dimeric structure of ferredoxin-NAD(P)+ oxidoreductase from the green sulfur bacterium Chlorobaculum tepidum: implications for binding ferredoxin and NADP+*. Journal of Molecular Biology, 2010. **401**(3): p. 403-14.
409. Seo, D., et al., *Purification and characterization of ferredoxin-NADP+ reductase encoded by Bacillus subtilis yumC*. Archives of Microbiology, 2004. **182**(1): p. 80-9.
410. Seo, D. and H. Sakurai, *Purification and characterization of ferredoxin-NAD(P)(+)-reductase from the green sulfur bacterium Chlorobium tepidum*. Biochimica et biophysica acta, 2002. **1597**(1): p. 123-32.
411. Huang, H., et al., *Electron bifurcation involved in the energy metabolism of the acetogenic bacterium Moorella thermoacetica growing on glucose or H2 plus CO2*. Journal of Bacteriology, 2012. **194**(14): p. 3689-99.
412. Rydzak, T., et al., *Insights into electron flux through manipulation of fermentation conditions and assessment of protein expression profiles in Clostridium thermocellum*. Applied Microbiology and Biotechnology, 2014. **98**(14): p. 6497-510.
413. Wang, S., et al., *NADP+ Reduction with Reduced Ferredoxin and NADP+ Reduction with NADH Are Coupled via an Electron-Bifurcating Enzyme Complex in Clostridium kluyveri*. Journal of Bacteriology, 2010. **192**(19): p. 5115-5123.
414. Buckel, W. and R.K. Thauer, *Energy conservation via electron bifurcating ferredoxin reduction and proton/Nad(+) translocating ferredoxin oxidation*. Biochim Biophys Acta, 2013. **1827**(2): p. 94-113.
415. Razquin, P., et al., *Expression of ferredoxin-NADP+ reductase in heterocysts from Anabaena sp.* The Biochemical journal, 1996. **316** (Pt 1): p. 157-60.
416. Ikeda, T., et al., *Ferredoxin-NADP reductase from the thermophilic hydrogen-oxidizing bacterium, Hydrogenobacter thermophilus TK-6*. Fems Microbiology Letters, 2009. **297**(1): p. 124-30.
417. Aliverti, A., et al., *Biochemical and crystallographic characterization of ferredoxin-NADP(+) reductase from nonphotosynthetic tissues*. Biochemistry, 2001. **40**(48): p. 14501-8.
418. Martinez-Julvez, M., M. Medina, and A. Velazquez-Campoy, *Binding thermodynamics of ferredoxin:NADP+ reductase: two different protein substrates and one energetics*. Biophysical journal, 2009. **96**(12): p. 4966-75.
419. Medina, M. and C. Gomez-Moreno, *Interaction of Ferredoxin-NADP(+) Reductase with its Substrates: Optimal Interaction for Efficient Electron Transfer*. Photosynthesis research, 2004. **79**(2): p. 113-31.
420. Onda, Y., et al., *Differential interaction of maize root ferredoxin:NADP(+) oxidoreductase with photosynthetic and non-photosynthetic ferredoxin isoproteins*. Plant Physiology, 2000. **123**(3): p. 1037-45.
421. Yeom, J., et al., *In vitro and in vivo interactions of ferredoxin-NADP+ reductases in Pseudomonas putida*. Journal of biochemistry, 2009. **145**(4): p. 481-91.

Appendices

422. Ewen, K.M., M. Kleser, and R. Bernhardt, *Adrenodoxin: The archetype of vertebrate-type [2Fe–2S] cluster ferredoxins*. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 2011. **1814**(1): p. 111-125.
423. Sancho, J., *Flavodoxins: sequence, folding, binding, function and beyond*. *Cellular and molecular life sciences : CMLS*, 2006. **63**(7-8): p. 855-64.
424. Yamazaki, M. and Y. Ichikawa, *Some properties of the apoenzyme of NADPH-adreno-ferredoxin reductase from bovine adrenocortical mitochondria*. *The International journal of biochemistry*, 1990. **22**(10): p. 1147-52.
425. Sato, Y., et al., *A novel enzymatic system against oxidative stress in the thermophilic hydrogen-oxidizing bacterium Hydrogenobacter thermophilus*. *Plos One*, 2012. **7**(4): p. e34825.
426. Lyon, E.J., et al., *UV-A/blue-light inactivation of the 'metal-free' hydrogenase (Hmd) from methanogenic archaea*. *European journal of biochemistry / FEBS*, 2004. **271**(1): p. 195-204.
427. Peters, J.W., et al., *[FeFe]- and [NiFe]-hydrogenase diversity, mechanism, and maturation*. *Biochimica et biophysica acta*, 2014.
428. Shima, S. and R.K. Thauer, *A third type of hydrogenase catalyzing H₂ activation*. *Chemical record (New York, N Y)*, 2007. **7**(1): p. 37-46.
429. Trchounian, A. and R. Gary Sawers, *Novel insights into the bioenergetics of mixed-acid fermentation: can hydrogen and proton cycles combine to help maintain a proton motive force?* *IUBMB Life*, 2014. **66**(1): p. 1-7.
430. Trchounian, K., et al., *Multiple and reversible hydrogenases for hydrogen production by Escherichia coli: dependence on fermentation substrate, pH and the F(0)F(1)-ATPase*. *Crit Rev Biochem Mol Biol*, 2012. **47**(3): p. 236-49.
431. Vignais, P.M., B. Billoud, and J. Meyer, *Classification and phylogeny of hydrogenases*. *FEMS microbiology reviews*, 2001. **25**(4): p. 455-501.
432. Malki, S., et al., *Characterization of an operon encoding an NADP-reducing hydrogenase in Desulfovibrio fructosovorans*. *Journal of Bacteriology*, 1995. **177**(10): p. 2628-36.
433. Nouailler, M., et al., *Solution structure of HndAc: a thioredoxin-like domain involved in the NADP-reducing hydrogenase complex*. *Protein science : a publication of the Protein Society*, 2006. **15**(6): p. 1369-78.
434. van Haaster, D.J., et al., *Reinvestigation of the steady-state kinetics and physiological function of the soluble NiFe-hydrogenase I of Pyrococcus furiosus*. *Journal of Bacteriology*, 2008. **190**(5): p. 1584-7.
435. Schmitz, O., et al., *HoxE--a subunit specific for the pentameric bidirectional hydrogenase complex (HoxEFUYH) of cyanobacteria*. *Biochimica et biophysica acta*, 2002. **1554**(1-2): p. 66-74.
436. Steuber, J., et al., *A Membrane-Bound NAD(P)(+)-Reducing Hydrogenase Provides Reduced Pyridine Nucleotides during Citrate Fermentation by Klebsiella pneumoniae*. *Journal of Bacteriology*, 1999. **181**(1): p. 241-245.
437. Wells, M.A., et al., *Engineering a non-native hydrogen production pathway into Escherichia coli via a cyanobacterial [NiFe] hydrogenase*. *Metabolic Engineering*, 2011. **13**(4): p. 445-53.
438. Schut, G.J., et al., *Mutational Analyses of the Enzymes Involved in the Metabolism of Hydrogen by the Hyperthermophilic Archaeon Pyrococcus furiosus*. *Frontiers in Microbiology*, 2012. **3**: p. 163.
439. Silva, P.J., et al., *Enzymes of hydrogen metabolism in Pyrococcus furiosus*. *European journal of biochemistry / FEBS*, 2000. **267**(22): p. 6541-51.
440. Casalot, L., et al., *Evidence for a Fourth Hydrogenase in Desulfovibrio fructosovorans*. *Journal of Bacteriology*, 2002. **184**(3): p. 853-856.
441. Lauterbach, L., O. Lenz, and K.A. Vincent, *H(2)-driven cofactor regeneration with NAD(P)(+)-reducing hydrogenases*. *Febs j*, 2013. **280**(13): p. 3058-68.
442. Lipscomb, G.L., et al., *Natural competence in the hyperthermophilic archaeon Pyrococcus furiosus facilitates genetic manipulation: construction of markerless deletions of genes encoding the two cytoplasmic hydrogenases*. *Applied and Environmental Microbiology*, 2011. **77**(7): p. 2232-8.
443. Germer, F., et al., *Overexpression, isolation, and spectroscopic characterization of the bidirectional [NiFe] hydrogenase from Synechocystis sp. PCC 6803*. *The Journal of biological chemistry*, 2009. **284**(52): p. 36462-72.
444. Ma, K., R. Weiss, and M.W.W. Adams, *Characterization of Hydrogenase II from the Hyperthermophilic Archaeon Pyrococcus furiosus and Assessment of Its Role in Sulfur Reduction*. *Journal of Bacteriology*, 2000. **182**(7): p. 1864-1871.
445. Bock, A., et al., *Maturation of hydrogenases*. *Advances in microbial physiology*, 2006. **51**: p. 1-71.
446. Forzi, L. and R.G. Sawers, *Maturation of [NiFe]-hydrogenases in Escherichia coli*. *Biometals : an international journal on the role of metal ions in biology, biochemistry, and medicine*, 2007. **20**(3-4): p. 565-78.
447. Watanabe, S., et al., *Structural basis of [NiFe] hydrogenase maturation by Hyp proteins*. *Biological chemistry*, 2012. **393**(10): p. 1089-100.
448. Mertens, R., et al., *Practical applications of hydrogenase I from Pyrococcus furiosus for NADPH generation and regeneration*. *Journal of Molecular Catalysis B: Enzymatic*, 2003. **24–25**(0): p. 39-52.
449. English, C.M., et al., *Recombinant and in vitro expression systems for hydrogenases: new frontiers in basic and applied studies for biological and synthetic H₂ production*. *Dalton transactions (Cambridge, England : 2003)*, 2009(45): p. 9970-8.
450. Grzeszik, C., et al., *Genes encoding the NAD-reducing hydrogenase of Rhodococcus opacus MR11*. *Microbiology (Reading, England)*, 1997. **143 (Pt 4)**: p. 1271-86.
451. Sun, J., et al., *Heterologous Expression and Maturation of an NADP-Dependent [NiFe]-Hydrogenase: A Key Enzyme in Biofuel Production*. *Plos One*, 2010. **5**(5): p. e10526.
452. Voordouw, G., et al., *Purification and characterization of Desulfovibrio vulgaris (Hildenborough) hydrogenase expressed in Escherichia coli*. *European journal of biochemistry / FEBS*, 1987. **162**(1): p. 31-6.
453. Kawai, S., et al., *Cytosolic NADP phosphatases I and II from Arthrobacter sp. strain KM: implication in regulation of NAD⁺/NADP⁺ balance*. *Journal of basic microbiology*, 2004. **44**(3): p. 185-96.
454. Reidl, J., et al., *NADP and NAD utilization in Haemophilus influenzae*. *Molecular Microbiology*, 2000. **35**(6): p. 1573-81.
455. Fukuda, C., S. Kawai, and K. Murata, *NAD(P)(H) Phosphatase Activities of Archaeal Inositol Monophosphatase and Eubacterial 3'-Phosphoadenosine 5'-Phosphate Phosphatase*. *Applied and Environmental Microbiology*, 2007. **73**(17): p. 5447-5452.

456. Kawai, S., et al., *MJ0917 in archaeon Methanococcus jannaschii is a novel NADP phosphatase/NAD kinase*. The Journal of biological chemistry, 2005. **280**(47): p. 39200-7.
457. McGuinness, E.T. and J.R. Butler, *NAD⁺ kinase—a review*. The International journal of biochemistry, 1985. **17**(1): p. 1-11.
458. Magni, G., G. Orsomando, and N. Raffaelli, *Structural and functional properties of NAD kinase, a key enzyme in NADP biosynthesis*. Mini reviews in medicinal chemistry, 2006. **6**(7): p. 739-46.
459. Grose, J.H., et al., *Evidence that feedback inhibition of NAD kinase controls responses to oxidative stress*. Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(20): p. 7601-6.
460. Bakke, P., et al., *Evaluation of Three Automated Genome Annotations for Halorhabdus utahensis*. PLoS ONE, 2009. **4**(7): p. e6291.
461. Labesse, G., et al., *Diacylglyceride kinases, sphingosine kinases and NAD kinases: distant relatives of 6-phosphofructokinases*. Trends in biochemical sciences, 2002. **27**(6): p. 273-5.
462. Outten, C.E. and V.C. Culotta, *A novel NADH kinase is the mitochondrial source of NADPH in Saccharomyces cerevisiae*. The EMBO journal, 2003. **22**(9): p. 2015-24.
463. Strand, M.K., et al., *POS5 gene of Saccharomyces cerevisiae encodes a mitochondrial NADH kinase required for stability of mitochondrial DNA*. Eukaryotic cell, 2003. **2**(4): p. 809-20.
464. Turner, W.L., J.C. Waller, and W.A. Snedden, *Identification, molecular cloning and functional characterization of a novel NADH kinase from Arabidopsis thaliana (thale cress)*. The Biochemical journal, 2005. **385**(Pt 1): p. 217-23.
465. Garavaglia, S., A. Galizzi, and M. Rizzi, *Allosteric regulation of Bacillus subtilis NAD kinase by quinolinic acid*. Journal of Bacteriology, 2003. **185**(16): p. 4844-50.
466. Bark, K., et al., *Polyphosphate-dependent enzymes in some coryneform bacteria isolated from sewage sludge*. Fems Microbiology Letters, 1993. **107**(2-3): p. 133-8.
467. Lindner, S.N., et al., *Polyphosphate/ATP-dependent NAD kinase of Corynebacterium glutamicum: biochemical properties and impact of ppnK overexpression on lysine production*. Applied Microbiology and Biotechnology, 2010. **87**(2): p. 583-93.
468. Mori, S., et al., *Molecular conversion of NAD kinase to NADH kinase through single amino acid residue substitution*. The Journal of biological chemistry, 2005. **280**(25): p. 24104-12.
469. Nakamichi, Y., et al., *Conferring the ability to utilize inorganic polyphosphate on ATP-specific NAD kinase*. Scientific Reports, 2013. **3**: p. 2632.
470. Singh, R., et al., *Oxidative Stress Evokes a Metabolic Adaptation That Favors Increased NADPH Synthesis and Decreased NADH Production in Pseudomonas fluorescens*. Journal of Bacteriology, 2007. **189**(18): p. 6665-6675.
471. Gerdes, S.Y., et al., *From genetic footprinting to antimicrobial drug targets: examples in cofactor biosynthetic pathways*. Journal of Bacteriology, 2002. **184**(16): p. 4555-72.
472. Poncet-Montange, G., et al., *NAD kinases use substrate-assisted catalysis for specific recognition of NAD*. The Journal of biological chemistry, 2007. **282**(47): p. 33925-34.
473. Thanassi, J.A., et al., *Identification of 113 conserved essential genes using a high-throughput gene disruption system in Streptococcus pneumoniae*. Nucleic acids research, 2002. **30**(14): p. 3152-62.
474. Zalacain, M., et al., *A global approach to identify novel broad-spectrum antibacterial targets among proteins of unknown function*. Journal of molecular microbiology and biotechnology, 2003. **6**(2): p. 109-26.
475. Lee, H.C., et al., *Thymidine production by overexpressing NAD⁺ kinase in an Escherichia coli recombinant strain*. Biotechnology letters, 2009. **31**(12): p. 1929-36.
476. Li, Z.-J., et al., *Overexpression of NAD kinase in recombinant Escherichia coli harboring the phbCAB operon improves poly(3-hydroxybutyrate) production*. Applied Microbiology and Biotechnology, 2009. **83**(5): p. 939-47.
477. Shi, F., et al., *Overexpression of NAD kinases improves the L-isoleucine biosynthesis in Corynebacterium glutamicum ssp. lactofermentum*. Enzyme and Microbial Technology, 2012. **51**(2): p. 73-80.
478. Rahman, M.M., et al., *Over-expression of NAD kinase in Corynebacterium crenatum and its Impact on L-Arginine Biosynthesis*. Tropical Journal of Pharmaceutical Research, 2012. **11**(6): p. 909-916.
479. Lee, W.-H., et al., *Effects of NADH kinase on NADPH-dependent biotransformation processes in Escherichia coli*. Applied Microbiology and Biotechnology, 2013. **97**(4): p. 1561-9.
480. Klasen, R., S. Bringer-Meyer, and H. Sahm, *Biochemical characterization and sequence analysis of the gluconate:NADP 5-oxidoreductase gene from Gluconobacter oxydans*. Journal of Bacteriology, 1995. **177**(10): p. 2637-43.
481. Wang, S., et al., *NADP-specific electron-bifurcating [FeFe]-hydrogenase in a functional complex with formate dehydrogenase in Clostridium autoethanogenum grown on CO*. Journal of Bacteriology, 2013. **195**(19): p. 4373-86.
482. Berk, H. and R.K. Thauer, *Function of coenzyme F420-dependent NADP reductase in methanogenic archaea containing an NADP-dependent alcohol dehydrogenase*. Arch Microbiol, 1997. **168**(5): p. 396-402.
483. Atomi, H., R. Matsumi, and T. Imanaka, *Reverse gyrase is not a prerequisite for hyperthermophilic life*. Journal of Bacteriology, 2004. **186**(14): p. 4829-4833.
484. Kato, N., H. Yurimoto, and R.K. Thauer, *The physiological role of the ribulose monophosphate pathway in bacteria and archaea*. Biosci Biotechnol Biochem, 2006. **70**(1): p. 10-21.
485. Huber, R., et al., *Thermotoga maritima sp. nov. represents a new genus of unique extremely thermophilic eubacteria growing up to 90°C*. Archives of Microbiology, 1986. **144**(4): p. 324-333.
486. Santangelo, T.J., L. Cubonova, and J.N. Reeve, *Thermococcus kodakarensis Genetics: TK1827-Encoded beta-Glycosidase, New Positive-Selection Protocol, and Targeted and Repetitive Deletion Technology*. Applied and Environmental Microbiology, 2010. **76**(4): p. 1044-1052.
487. Keller, M.A., G. Piedrafita, and M. Ralser, *The widespread role of non-enzymatic reactions in cellular metabolism*. Current Opinion in Biotechnology, 2015. **34**: p. 153-161.
488. Miclet, E., et al., *NMR spectroscopic analysis of the first two steps of the pentose-phosphate pathway elucidates the role of 6-phosphogluconolactonase*. J Biol Chem, 2001. **276**(37): p. 34840-6.

Appendices

489. Rakitzis, E.T. and P. Papandreou, *Reactivity of 6-phosphogluconolactone with hydroxylamine: The possible involvement of glucose-6-phosphate dehydrogenase in endogenous glycation reactions*. *Chemico-biological interactions*, 1998. **113**(3): p. 205-216.
490. Spaans, S.K., J. van der Oost, and S.M. Kengen, *The chromosome copy number of the hyperthermophilic archaeon Thermococcus kodakarensis KOD1*. *Extremophiles*, 2015: p. 1-10.
491. Borges, N., et al., *Thermococcus kodakarensis Mutants Deficient in Di-myo-Inositol Phosphate Use Aspartate To Cope with Heat Stress*. *Journal of Bacteriology*, 2010. **192**(1): p. 191-197.
492. Spitalny, P. and M. Thomm, *A polymerase III-like reinitiation mechanism is operating in regulation of histone expression in archaea*. *Molecular Microbiology*, 2008. **67**(5): p. 958-970.
493. Santangelo, T.J., L. Cubonova, and J.N. Reeve, *Shuttle vector expression in Thermococcus kodakarensis: Contributions of cis elements to protein synthesis in a hyperthermophilic archaeon*. *Applied and Environmental Microbiology*, 2008. **74**(10): p. 3099-3104.
494. Nisar, M.A., et al., *TK1299, a highly thermostable NAD(P)H oxidase from Thermococcus kodakarensis exhibiting higher enzymatic activity with NADPH*. *J Biosci Bioeng*, 2013. **116**(1): p. 39-44.
495. Adams, M.W., et al., *Key role for sulfur in peptide metabolism and in regulation of three hydrogenases in the hyperthermophilic archaeon Pyrococcus furiosus*. *J Bacteriol*, 2001. **183**(2): p. 716-24.
496. Nanninga, H.J. and J.C. Gottschal, *Amino acid fermentation and hydrogen transfer in mixed cultures*. *FEMS Microbiology Letters*, 1985. **31**(5): p. 261-269.
497. Ramsay, I. and P. Pullammanappallil, *Protein degradation during anaerobic wastewater treatment: derivation of stoichiometry*. *Biodegradation*, 2001. **12**(4): p. 247-256.
498. Maru, B.T., et al., *Glycerol fermentation to hydrogen by Thermotoga maritima: Proposed pathway and bioenergetic considerations*. *International Journal of Hydrogen Energy*, 2013. **38**(14): p. 5563-5572.
499. Santangelo, T.J., et al., *Archaeal Intrinsic Transcription Termination In Vivo*. *Journal of Bacteriology*, 2009. **191**(22): p. 7102-7108.
500. Benoist, P. and J. Schwencke, *Native agarose-polyacrylamide gel electrophoresis allowing the detection of aminopeptidase, dehydrogenase, and esterase activities at the nanogram level: enzymatic patterns in some Frankia strains*. *Anal Biochem*, 1990. **187**(2): p. 337-44.
501. Zwietering, M.H., et al., *Modeling of the Bacterial Growth Curve*. *Applied and Environmental Microbiology*, 1990. **56**(6): p. 1875-1881.
502. Vonck, J., et al., *Three-dimensional Structure of A1A0 ATP Synthase from the Hyperthermophilic Archaeon Pyrococcus furiosus by Electron Microscopy*. *Journal of Biological Chemistry*, 2009. **284**(15): p. 10110-10119.
503. Pecoraro, V., et al., *Quantification of ploidy in proteobacteria revealed the existence of monoploid, (mero-)oligoploid and polyploid species*. *Plos One*, 2011. **6**(1): p. e16392.
504. Hildenbrand, C., et al., *Genome Copy Numbers and Gene Conversion in Methanogenic Archaea*. *Journal of Bacteriology*, 2011. **193**(3): p. 734-743.
505. Chant, J., et al., *The protein synthesizing machinery of the archaeobacterium Halobacterium cutirubrum: Molecular characterization*. *Systematic and Applied Microbiology*, 1986. **7**(1): p. 106-114.
506. Breuert, S., et al., *Regulated Polyploidy in Halophilic Archaea*. *Plos One*, 2006. **1**(1): p. e92.
507. Liu, X., et al., *Characterization of the minimal replicon of pHM300 and independent copy number control of major and minor chromosomes of Haloferax mediterranei*. *Fems Microbiology Letters*, 2013. **339**(1): p. 66-74.
508. Majernik, A.I., et al., *DNA content and nucleoid distribution in Methanothermobacter thermautotrophicus*. *Journal of Bacteriology*, 2005. **187**(5): p. 1856-1858.
509. Malandrin, L., H. Huber, and R. Bernander, *Nucleoid structure and partition in Methanococcus jannaschii: an archaeon with multiple copies of the chromosome*. *Genetics*, 1999. **152**(4): p. 1315-1323.
510. Marie, D., D. Vaultot, and F. Partensky, *Application of the novel nucleic acid dyes YOYO-1, YO-PRO-1, and PicoGreen for flow cytometric analysis of marine prokaryotes*. *Applied and Environmental Microbiology*, 1996. **62**(5): p. 1649-55.
511. Lundgren, M., et al., *Cell Cycle Characteristics of Crenarchaeota: Unity among Diversity*. *Journal of Bacteriology*, 2008. **190**(15): p. 5362-5367.
512. Bernander, R. and A. Poplawski, *Cell cycle characteristics of thermophilic archaea*. *Journal of Bacteriology*, 1997. **179**(16): p. 4963-4969.
513. Ventosa, A. and A. Oren, *Halobacterium salinarum nom. corrig., a Name To Replace Halobacterium salinarum (Elazari-Volcani) and To Include Halobacterium halobium and Halobacterium cutirubrum*. *International Journal of Systematic Bacteriology*, 1996. **46**(1): p. 347.
514. Griesse, M., C. Lange, and J. Soppa, *Ploidy in cyanobacteria*. *Fems Microbiology Letters*, 2011. **323**(2): p. 124-131.
515. Maldonado, R., J. Jimenez, and J. Casades, *Changes of ploidy during the Azotobacter vinelandii growth cycle*. *Journal of Bacteriology*, 1994. **176**(13): p. 3911-9.
516. Mendell, J.E., et al., *Extreme polyploidy in a large bacterium*. *Proceedings of the National Academy of Sciences*, 2008. **105**(18): p. 6730-6734.
517. Angert, E.R., K.D. Clements, and N.R. Pace, *The largest bacterium*. *Nature*, 1993. **362**: p. 239-241.
518. Lea-Smith, D.J., et al., *Phycobilisome-Deficient Strains of Synechocystis sp. PCC 6803 Have Reduced Size and Require Carbon-Limiting Conditions to Exhibit Enhanced Productivity*. *Plant Physiology*, 2014. **165**(2): p. 705-714.
519. Cao, Z., C.W. Mueller, and D.A. Julin, *Analysis of the recJ gene and protein from Deinococcus radiodurans*. *DNA repair*, 2010. **9**(1): p. 66-75.
520. Ohtani, N., M. Tomita, and M. Itaya, *An extreme thermophile, Thermus thermophilus, is a polyploid bacterium*. *Journal of Bacteriology*, 2010. **192**(20): p. 5499-505.
521. Soppa, J., *Evolutionary advantages of polyploidy in halophilic archaea*. *Biochemical Society transactions*, 2013. **41**(1): p. 339-343.

522. Zerulla, K., et al., *DNA as a phosphate storage polymer and the alternative advantages of polyploidy for growth or survival*. Plos One, 2014. **9**(4): p. e94819.
523. Zerulla, K. and J. Soppa, *Polyploidy in haloarchaea: advantages for growth and survival*. Frontiers in Microbiology, 2014. **5**: p. 274.
524. Minsky, A., R. Ghirlando, and Z. Reich, *Nucleosomes: a Solution to a Crowded Intracellular Environment?* Journal of Theoretical Biology, 1997. **188**(3): p. 379-385.
525. Sandman, K., S.L. Pereira, and J.N. Reeve, *Diversity of prokaryotic chromosomal proteins and the origin of the nucleosome*. Cellular and molecular life sciences : CMLS, 1998. **54**(12): p. 1350-64.
526. Luijsterburg, M.S., et al., *The Major Architects of Chromatin: Architectural Proteins in Bacteria, Archaea and Eukaryotes*. Critical reviews in biochemistry and molecular biology, 2008. **43**(6): p. 393-418.
527. Maruyama, H., et al., *An alternative beads-on-a-string chromatin architecture in Thermococcus kodakarensis*. EMBO reports, 2013. **14**: p. 711-717.
528. Higashibata, H., et al., *Analysis of DNA Compaction Profile and Intracellular Contents of Archaeal Histones from Pyrococcus kodakarensis KOD1*. Biochemical and Biophysical Research Communications, 1999. **258**(2): p. 416-424.
529. Čuboňová, L., et al., *An Archaeal Histone Is Required for Transformation of Thermococcus kodakarensis*. Journal of Bacteriology, 2012. **194**(24): p. 6864-6874.
530. Maruyama, H., et al., *Histone and TK0471/TrmBL2 form a novel heterogeneous genome architecture in the hyperthermophilic archaeon Thermococcus kodakarensis*. Molecular Biology of the Cell, 2011. **22**(3): p. 386-398.
531. Sandman, K. and J.N. Reeve, *Archaeal histones and the origin of the histone fold*. Current Opinion in Microbiology, 2006. **9**(5): p. 520-525.
532. White, M.F. and S.D. Bell, *Holding it together: chromatin in the Archaea*. Trends in Genetics, 2002. **18**(12): p. 621-626.
533. Odijk, T., *Osmotic compaction of supercoiled DNA into a bacterial nucleoid*. Biophysical Chemistry, 1998. **73**(1-2): p. 23-29.
534. Kubitschek, H., *Cell volume increase in Escherichia coli after shifts to richer media*. Journal of Bacteriology, 1990. **172**(1): p. 94-101.
535. Dame, R.T., *The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin*. Molecular Microbiology, 2005. **56**(4): p. 858-70.
536. Cronan, J.E., *Bacterial Membrane Lipids: Where Do We Stand?* Annual review of microbiology, 2003. **57**(1): p. 203-224.
537. Parsons, J.B. and C.O. Rock, *Bacterial lipids: Metabolism and membrane homeostasis*. Progress in Lipid Research, 2013. **52**(3): p. 249-276.
538. Zhang, Y.-M. and C.O. Rock, *Membrane lipid homeostasis in bacteria*. Nat Rev Micro, 2008. **6**(3): p. 222-233.
539. Zhang, Y.-M. and C.O. Rock, *Transcriptional regulation in bacterial membrane lipid synthesis*. Journal of Lipid Research, 2009. **50**(Supplement): p. S115-S119.
540. Albers, S.-V., et al., *Adaptations of the archaeal cell membrane to heat stress*. Front Biosci, 2000. **5**: p. D813-20.
541. Denich, T.J., et al., *Effect of selected environmental and physico-chemical factors on bacterial cytoplasmic membranes*. Journal of Microbiological Methods, 2003. **52**(2): p. 149-182.
542. Jain, S., A. Caforio, and A.J.M. Driessen, *Biosynthesis of archaeal membrane ether lipids*. Frontiers in Microbiology, 2014. **5**.
543. Koga, Y. and H. Morii, *Biosynthesis of Ether-Type Polar Lipids in Archaea and Evolutionary Considerations*. Microbiology and Molecular Biology Reviews, 2007. **71**(1): p. 97-120.
544. Koga, Y., *Early Evolution of Membrane Lipids: How did the Lipid Divide Occur?* Journal of Molecular Evolution, 2011. **72**(3): p. 274-282.
545. Koga, Y. and H. Morii, *Recent Advances in Structural Research on Ether Lipids from Archaea Including Comparative and Physiological Aspects*. Bioscience, Biotechnology, and Biochemistry, 2005. **69**(11): p. 2019-2034.
546. Matsuno, Y., et al., *Effect of Growth Temperature and Growth Phase on the Lipid Composition of the Archaeal Membrane from Thermococcus kodakarensis*. Bioscience, Biotechnology, and Biochemistry, 2009. **73**(1): p. 104-108.
547. Meador, T.B., et al., *Thermococcus kodakarensis modulates its polar membrane lipids and elemental composition according to growth stage and phosphate availability*. Frontiers in Microbiology, 2014. **5**.
548. Morii, H., et al., *A novel ether core lipid with H-shaped C80-isoprenoid hydrocarbon chain from the hyperthermophilic methanogen Methanothermobacter fervidus*. Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism, 1998. **1390**(3): p. 339-345.
549. Elferink, M.G., et al., *Stability and proton-permeability of liposomes composed of archaeal tetraether lipids*. Biochimica et Biophysica Acta (BBA)-Biomembranes, 1994. **1193**(2): p. 247-254.
550. Lai, D., J. Springstead, and H. Monbouquette, *Effect of growth temperature on ether lipid biochemistry in Archaeoglobus fulgidus*. Extremophiles : life under extreme conditions, 2008. **12**(2): p. 271-278.
551. Sprott, G.D., M. Meloche, and J.C. Richards, *Proportions of diether, macrocyclic diether, and tetraether lipids in Methanococcus jannaschii grown at different temperatures*. Journal of Bacteriology, 1991. **173**(12): p. 3907-3910.
552. Uda, I., et al., *Variation in molecular species of polar lipids from Thermoplasma acidophilum depends on growth temperature*. Lipids, 2001. **36**(1): p. 103-105.
553. Kon, T., et al., *Effects of a squalene epoxidase inhibitor, terbinafine, on ether lipid biosyntheses in a thermoacidophilic archaeon, Thermoplasma acidophilum*. Journal of Bacteriology, 2002. **184**(5): p. 1395-1401.
554. Nemoto, N., et al., *Characterization of the precursor of tetraether lipid biosynthesis in the thermoacidophilic archaeon Thermoplasma acidophilum*. Extremophiles : life under extreme conditions, 2003. **7**(3): p. 235-243.
555. Langworthy, T.A., *Turnover of di-O-phytanylglycerol in Thermoplasma*. Rev. Infect. Dis., 1982. **4**(S266).
556. Poulter, C.D., T. Aoki, and L. Daniels, *Biosynthesis of isoprenoid membranes in the methanogenic archaeobacterium Methanospirillum hungatei*. Journal of the American Chemical Society, 1988. **110**(8): p. 2620-2624.
557. Eguchi, T., Y. Nishimura, and K. Kakinuma, *Importance of the isopropylidene terminal of geranylgeranyl group for the formation of tetraether lipid in methanogenic archaea*. Tetrahedron Letters, 2003. **44**(16): p. 3275-3279.

Appendices

558. Eguchi, T., et al., *Unusual double-bond migration as a plausible key reaction in the biosynthesis of the isoprenoidal membrane lipids of methanogenic archaea*. Chem. Commun., 2000(16): p. 1545-1546.
559. Kramer, J.K.G. and F.D. Sauer, *Changes in the diether-to-tetraether-lipid ratio during cell growth in Methanobacterium thermoautotrophicum*. Fems Microbiology Letters, 1991. **83**(1): p. 45-50.
560. Morii, H. and Y. Koga, *Tetraether type polar lipids increase after logarithmic growth phase of Methanobacterium thermoautotrophicum in compensation for the decrease of diether lipids*. Fems Microbiology Letters, 1993. **109**(2-3): p. 283-287.
561. Lobasso, S., et al., *Coupled TLC and MALDI-TOF/MS analyses of the lipid extract of the hyperthermophilic archaeon Pyrococcus furiosus*. Archaea, 2012. **2012**: p. 957852.
562. Sugai, A., et al., *The Core Lipid Composition of the 17 Strains of Hyperthermophilic Archaea, Thermococcales*. Journal of Oleo Science, 2004. **53**(1): p. 41-44.
563. Kon, T., et al., *Effects of a squalene epoxidase inhibitor, terbinafine, on ether lipid biosyntheses in a thermoacidophilic archaeon, Thermoplasma acidophilum*. J Bacteriol, 2002. **184**(5): p. 1395-401.
564. Sugai, A., et al., *The Core Lipid Composition of the 17 Strains of Hyperthermophilic Archaea, Thermococcales*. Journal of Oleo Science, 2004. **53**(1): p. 41-44.
565. Lobasso, S., et al., *Coupled TLC and MALDI-TOF/MS Analyses of the Lipid Extract of the Hyperthermophilic Archaeon Pyrococcus furiosus*. Archaea, 2012. **2012**: p. 10.
566. Nishihara, M., et al., *Structural analysis by reductive cleavage with LiAlH₄ of an allyl ether choline-phospholipid, archaetidylcholine, from the hyperthermophilic methanoarchaeon Methanopyrus kandleri*. Archaea, 2002. **1**(2): p. 123-131.
567. Sprott, G.D., B.J. Agnew, and G.B. Patel, *Structural features of ether lipids in the archaeobacterial thermophiles Pyrococcus furiosus, Methanopyrus kandleri, Methanothermus fervidus, and Sulfolobus acidocaldarius*. Canadian Journal of Microbiology, 1997. **43**(5): p. 467-476.
568. Tachibana, A., *A novel prenyltransferase, farnesylgeranyl diphosphate synthase, from the haloalkaliphilic archaeon, Natronobacterium pharaonis*. FEBS Letters. **341**(2): p. 291-294.
569. Makarova, K., Y. Wolf, and E. Koonin, *Archaeal Clusters of Orthologous Genes (arCOGs): An Update and Application for Analysis of Shared Features between Thermococcales, Methanococcales, and Methanobacteriales*. Life, 2015. **5**(1): p. 818.
570. Caforio, A., et al., *Formation of the ether lipids archaetidylglycerol and archaetidylethanolamine in Escherichia coli*. Biochemical Journal, 2015. **470**(3): p. 343-355.
571. Guldan, H., R. Sterner, and P. Babinger, *Identification and Characterization of a Bacterial Glycerol-1-phosphate Dehydrogenase: Ni²⁺-Dependent AraM from Bacillus subtilis*. Biochemistry, 2008. **47**(28): p. 7376-7384.
572. Hemmi, H., et al., *(S)-2,3-Di-O-geranylgeranyl glyceryl Phosphate Synthase from the Thermoacidophilic Archaeon Sulfolobus solfataricus: MOLECULAR CLONING AND CHARACTERIZATION OF A MEMBRANE-INTRINSIC PRENYLTRANSFERASE INVOLVED IN THE BIOSYNTHESIS OF ARCHAEAAL ETHER-LINKED MEMBRANE LIPIDS*. Journal of Biological Chemistry, 2004. **279**(48): p. 50197-50203.
573. Isobe, K., et al., *Geranylgeranyl Reductase and Ferredoxin from Methanosarcina acetivorans Are Required for the Synthesis of Fully Reduced Archaeal Membrane Lipid in Escherichia coli Cells*. Journal of Bacteriology, 2014. **196**(2): p. 417-423.
574. Jain, S., et al., *Identification of CDP-Archaeol Synthase, a Missing Link of Ether Lipid Biosynthesis in Archaea*. Chemistry & Biology, 2014. **21**(10): p. 1392-1401.
575. Nemoto, N., T. Oshima, and A. Yamagishi, *Purification and Characterization of Geranylgeranyl glyceryl Phosphate Synthase from a Thermoacidophilic Archaeon, Thermoplasma acidophilum*. The Journal of Biochemistry, 2003. **133**(5): p. 651-657.
576. Soderberg, T., A. Chen, and C.D. Poulter, *Geranylgeranyl glyceryl Phosphate Synthase. Characterization of the Recombinant Enzyme from Methanobacterium thermoautotrophicum*. Biochemistry, 2001. **40**(49): p. 14847-14854.
577. Tachibana, A., et al., *Potassium-Stimulating Mechanism of Geranylgeranyl Diphosphate Synthase of Methanobacterium thermoformicum SF-4*. The Journal of Biochemistry, 1993. **114**(3): p. 389-392.
578. Leferink, N.G.H., et al., *The growing VAO flavoprotein family*. Archives of Biochemistry and Biophysics, 2008. **474**(2): p. 292-301.
579. Liang, P.-H., T.-P. Ko, and A.H.J. Wang, *Structure, mechanism and function of prenyltransferases*. European Journal of Biochemistry, 2002. **269**(14): p. 3339-3354.
580. Ogura, K. and T. Koyama, *Enzymatic Aspects of Isoprenoid Chain Elongation*. Chemical reviews, 1998. **98**(4): p. 1263-1276.
581. Wang, K. and S.-i. Ohnuma, *Chain-length determination mechanism of isoprenyl diphosphate synthases and implications for molecular evolution*. Trends in biochemical sciences, 1999. **24**(11): p. 445-451.
582. Kanai, T., et al., *Identification of the Phr-dependent heat shock regulon in the hyperthermophilic archaeon, Thermococcus kodakaraensis*. Journal of biochemistry, 2010. **147**(3): p. 361-370.
583. Liu, W., et al., *Crystal Structure of the Archaeal Heat Shock Regulator from Pyrococcus furiosus: A Molecular Chimera Representing Eukaryal and Bacterial Features*. Journal of Molecular Biology, 2007. **369**(2): p. 474-488.
584. Vierke, G., et al., *A Novel Archaeal Transcriptional Regulator of Heat Shock Response*. Journal of Biological Chemistry, 2003. **278**(1): p. 18-26.
585. Hemmi, H., et al., *Menaquinone-Specific Prenyl Reductase from the Hyperthermophilic Archaeon Archaeoglobus fulgidus*. Journal of Bacteriology, 2005. **187**(6): p. 1937-1944.
586. Naparstek, S., Z. Guan, and J. Eichler, *A predicted geranylgeranyl reductase reduces the ω-position isoprene of dolichol phosphate in the halophilic archaeon, Haloferax volcanii*. Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids, 2012. **1821**(6): p. 923-933.
587. Huijbers, M.M.E., et al., *Flavin dependent monooxygenases*. Archives of Biochemistry and Biophysics, 2014. **544**: p. 2-17.
588. Nowosielski, M., et al., *Detailed Mechanism of Squalene Epoxidase Inhibition by Terbinafine*. Journal of Chemical Information and Modeling, 2011. **51**(2): p. 455-462.
589. Hopmans, E.C., S. Schouten, and J.S. Sinninghe Damsté, *The effect of improved chromatography on GDGT-based palaeoproxies*. Organic Geochemistry, 2016. **93**: p. 1-6.

590. Schouten, S., et al., *Archaeal and Bacterial Glycerol Dialkyl Glycerol Tetraether Lipids in Hot Springs of Yellowstone National Park*. Applied and Environmental Microbiology, 2007. **73**(19): p. 6181-6191.
591. Damsté, J.S.S., et al., *Crenarchaeol: the characteristic core glycerol dibiphytanyl glycerol tetraether membrane lipid of cosmopolitan pelagic crenarchaeota*. Journal of Lipid Research, 2002. **43**(10): p. 1641-1651.
592. Aono, R., et al., *A pentose bisphosphate pathway for nucleoside degradation in Archaea*. Nat Chem Biol, 2015. **11**(5): p. 355-360.
593. Gossen, M. and H. Bujard, *Tight control of gene expression in mammalian cells by tetracycline-responsive promoters*. Proceedings of the National Academy of Sciences of the United States of America, 1992. **89**(12): p. 5547-5551.
594. Lopez de Felipe, F., et al., *Cofactor Engineering: a Novel Approach to Metabolic Engineering in Lactococcus lactis by Controlled Expression of NADH Oxidase*. Journal of Bacteriology, 1998. **180**(15): p. 3804-3808.
595. Lutz, R. and H. Bujard, *Independent and tight regulation of transcriptional units in Escherichia coli via the LacR/O, the TetR/O and AraC/I1-12 regulatory elements*. Nucleic acids research, 1997. **25**(6): p. 1203-1210.
596. Peng, N., et al., *A Synthetic Arabinose-Inducible Promoter Confers High Levels of Recombinant Protein Expression in Hyperthermophilic Archaeon Sulfolobus islandicus*. Applied and Environmental Microbiology, 2012. **78**(16): p. 5630-5637.
597. van de Werken, H.J.G., et al., *Identification of a glycolytic regulon in the archaea Pyrococcus and Thermococcus*. Fems Microbiology Letters, 2006. **260**(1): p. 69-76.
598. Ellis, D.I. and R. Goodacre, *Metabolomics-assisted synthetic biology*. Current Opinion in Biotechnology, 2012. **23**(1): p. 22-28.
599. Toya, Y. and H. Shimizu, *Flux analysis and metabolomics for systematic metabolic engineering of microorganisms*. Biotechnol Adv, 2013. **31**(6): p. 818-826.
600. van Maris, A.A., et al., *Development of Efficient Xylose Fermentation in Saccharomyces cerevisiae: Xylose Isomerase as a Key Component*, in *Biofuels*, L. Olsson, Editor 2007, Springer Berlin Heidelberg, p. 179-204.
601. Albers, S.-V., et al., *Insights into ABC Transport in Archaea*. Journal of bioenergetics and biomembranes, 2004. **36**(1): p. 5-15.
602. Koning, S.M., W.N. Konings, and A.J.M. Driessen, *Biochemical evidence for the presence of two α -glucoside ABC-transport systems in the hyperthermophilic archaeon Pyrococcus furiosus*. Archaea, 2002. **1**(1): p. 19-25.
603. Mardanov, A.V., et al., *Metabolic Versatility and Indigenous Origin of the Archaeon Thermococcus sibiricus, Isolated from a Siberian Oil Reservoir, as Revealed by Genome Analysis*. Applied and Environmental Microbiology, 2009. **75**(13): p. 4580-4588.
604. Xavier, K.B., et al., *High-affinity maltose/trehalose transport system in the hyperthermophilic archaeon Thermococcus litoralis*. Journal of Bacteriology, 1996. **178**(16): p. 4773-7.
605. Farkas, J., et al., *Recombinogenic Properties of Pyrococcus furiosus Strain COM1 Enable Rapid Selection of Targeted Mutants*. Applied and Environmental Microbiology, 2012. **78**(13): p. 4669-4676.
606. Han, D., S. Norris, and Z. Xu, *Construction and transformation of a Thermotoga-E. coli shuttle vector*. BMC Biotechnology, 2012. **12**(1): p. 1-9.
607. Han, D., et al., *Natural transformation of Thermotoga sp. strain RQ7*. BMC Biotechnology, 2014. **14**(1): p. 1-10.
608. Xu, H., D. Han, and Z. Xu, *Expression of Heterologous Cellulases in Thermotoga sp. Strain RQ2*. BioMed Research International, 2015. **2015**: p. 11.
609. Mukund, S. and M.W.W. Adams, *Glyceraldehyde-3-phosphate Ferredoxin Oxidoreductase, a Novel Tungsten-containing Enzyme with a Potential Glycolytic Role in the Hyperthermophilic Archaeon Pyrococcus furiosus*. Journal of Biological Chemistry, 1995. **270**(15): p. 8389-8392.
610. Heinicke, I., et al., *Mutational analysis of genes encoding chromatin proteins in the archaeon Methanococcus voltae indicates their involvement in the regulation of gene expression*. Molecular Genetics and Genomics, 2004. **272**(1): p. 76-87.
611. Weidenbach, K., et al., *Deletion of the archaeal histone in Methanosarcina mazei G61 results in reduced growth and genomic transcription*. Molecular Microbiology, 2008. **67**(3): p. 662-671.
612. Dinger, M.E., G.J. Baillie, and D.R. Musgrave, *Growth phase-dependent expression and degradation of histones in the thermophilic archaeon Thermococcus zilligii*. Molecular Microbiology, 2000. **36**(4): p. 876-885.
613. Sandman, K., et al., *Growth-phase-dependent synthesis of histones in the archaeon Methanothermus fervidus*. Proceedings of the National Academy of Sciences of the United States of America, 1994. **91**(26): p. 12624-12628.
614. Jiang, W., et al., *RNA-guided editing of bacterial genomes using CRISPR-Cas systems*. Nature biotechnology, 2013. **31**(3): p. 233-9.
615. energy, U.S.d.o. *Hydrogen and fuel cell technologies program: fuel cells*. 2010; Available from: http://www1.eere.energy.gov/hydrogenandfuelcells/pdfs/doe_h2_fuelcell_factsheet.pdf.
616. Crabtree, G.W., M.S. Dresselhaus, and M.V. Buchanan, *The hydrogen economy*. Physics Today, 2004. **57**(12): p. 39-44.

Acknowledgements - Dankwoord

As I sit here, in early 2016, I can hardly believe it is over. All those years of hard work finally resulted in the present work: my thesis! Although at times it felt like every day was a new struggle to overcome the disappointment of yet another failing experiment, I am truly grateful for all I have learned. Especially, non-scientifically. Since I realize that this might sound a bit strange in a PhD thesis, let me explain.

As many of you know, doing a PhD in Science can be really challenging. Not only does it require some intelligence and knowledge about the subject but, perhaps even more important, also patience, perseverance and, admittedly, luck. Nevertheless, it is inevitable that sooner or later you have to endure hardships, uncertainties and deprivations, and it was then that I learned the most. It taught me about my strengths and weaknesses, my motivational drivers and energy drains, and let me to belief that you can do almost anything as long as you truly want to and believe in it. Moreover, it not only taught me to endure but to overcome and made me realize that true change happens when you fully embrace the relations and connections around you. And that, in my opinion, is worth more than all the scientific knowledge in the world.

This thesis would not have been possible without the direct and indirect help and support of many different people. It is impossible to list them all here individually. Nevertheless I would like to thank some persons in particular.

First of all, I would like to thank my supervisors **Servé** and **John**. I am very grateful that you gave me the opportunity to do my PhD at the Laboratory of Microbiology, and have to admit that I felt slightly honored when you approached me even before my internship in Berkeley had finished. **Servé**, many thanks for all your support, involvement and patience. You have always made me feel welcome and I really appreciate it that I could always drop by to ask questions or discuss the latest results. You were a great mentor and sparring partner and really taught me a lot. Servé, thanks for everything and I enjoyed working with you.

John, thanks for having me in your group and being my promotor. Your enthusiasm, drive and openness were really inspiring and contributed greatly to the positive atmosphere in your group. I enjoyed your "can do" mentality and endless flow of ideas, and want to especially thank you for proving that a clean desk policy is highly overrated. John, despite your terrible taste in football, I really enjoyed working with you.

I also like to thank all the collaborators and experts without whom the work described in this thesis would not have been possible. In particular, I like to thank **Haruyuki (Harry) Atomi**, **Thomas (Tom) Santangelo**, **Rie Matsumi** and **Melvin Siliakus**. **Harry**, thanks for all your support over the years. I really appreciate that you were always willing to help and that you took the time to think along and answer my questions. I will also never forget the endless supply of burgers you arranged for us when we visited your lab in Kyoto.

Tom, thank you very much for hosting me at the Department of Microbiology of the Ohio State University. From the beginning on you made me feel very welcome and I really enjoyed meeting your family. I would like to thank you for all your input and help over the years and want you to know that this has really helped the project some major leaps forward.

Rie, like Harry and Tom, I would like to thank you for all your help over the years. You were always willing to answer my questions and taught me a lot about *T. kodakarensis*. Moreover, you have been of great help in my final research project, by constructing various disruption strains. It has been a pleasure working with you and I hope our effort soon pays off.

Melvin, the last research chapter of this thesis could not have been realized without your help. Your knowledge on archaeal membranes and my knowledge on hyperthermophiles and the engineering thereof, turned out to be a great combination! Thanks for sharing your knowledge and all your hard work. I really enjoyed working with you and I hope our effort soon pays off.

Moreover, I also would like to thank the group members of the laboratory of biochemical engineering of the Kyoto University for their help over the years, especially **Tamotsu Kanai**, **Takaaki Sato** and **Yuusuke Yokooji**. Similarly, I would like to thank the members of the Department of Microbiology of the Ohio State University and especially

L'ubomíra Čuboňová and **John Reeve**. Furthermore, I would like to thank **Jörg Soppa** of the Goethe University Frankfurt, **Remus Dame** of Leiden University, **Kira Makarova** of the National Center for Biotechnology Information (NCBI), **Jaap Sinninghe Damste** and **Michel Koenen** of the Department of Marine Organic Biogeochemistry of the Royal Netherlands Institute for Sea Research (NIOZ), and **Jolein Gloerich** of the Radboud Center for Proteomics, Glycomics & Metabolomics.

Besides all these collaborators and experts from all around the world, I also want to thank many people closer to home. **To all the people from the Laboratory of Microbiology**, especially from the **Bacterial Genetics group**, thanks for your advice, support, help and all the great times over the years! You are simply with too many to thank you all individually, but I wish you all the best in your (future) careers and hope to see you again!

To my students, **Philana, Fabienne, Wouter, Robbert** and **Frits**. I really enjoyed working with you and even though much of your work did not end up in this thesis, your help was much appreciated! In addition, I would like to thank **Karen Trchounian** for his help with the formate project, which unfortunately also proved to be a dead-end. Similarly, I would like to thank **Bram** and **Sjon** for their help with the *Thermotoga maritima* project.

To the **organizers and participants** of the **PhD trip to China and Japan**, thanks for the great experience! Especially, **Ruud, Sven** and **Peer**. I will never forget our drinking session at a random square in the middle of Beijing, our karaoke session, or Sven being at the wrong train station!

To the **organizers, participants and hosts** of the **PhD trip to the USA and Canada**, thanks for making it a huge success! **Floor, Teunke, Corina, Noora, Javi, Vicente**, it was a lot of fun organizing this awesome trip with you! Teunke, I want to thank you, especially, for all your help with finishing the financial overview. I could not have done it without your help. In addition, I would like to thank **Nico, Daan, Tijn** and **Jasper** for the great time in Canada after the trip. Even though we spend a lot of time crammed up in a car or cottage in the middle of nowhere and we had to endure unexpected snowfall, Dutch neighbors, a bear encounter, and endless hours of Risk, I really enjoyed it!

Rie, Pierpaolo, Raymond, Tessa, Nico, Melvin and Franklin, thanks for being awesome office mates! I really enjoyed all our discussions, chats, running gags, bets, stupid jokes and whatnot.

Daarnaast wil ik natuurlijk ook al mijn vrienden en familie bedanken. Dit boekje is het tastbare bewijs dat ik daadwerkelijk aan iets werkte als ik er voor de zoveelste keer weer eens niet bij kon zijn... **Tinus, Sjur, Wijn, Joer, Freek, Nelis, Sjaap, Cum, Alt, Jerk, Vinnie, Ernie, Selen, Flens, Geert, Midas, Nick, Daan, Jan, Menjo, Karel, Simba, Tiet, Gijs en Nico**, bedankt voor de nodige ontspanning en gezelligheid! **Pap, Mam, Lau** en **alle andere familie**, heel erg bedankt voor jullie interesse, enthousiasme, steun en begrip door de jaren heen. Misschien dat ik het niet altijd even veel heb laten blijken, maar ik heb hier echt enorm veel aan gehad!

Lieve Inge, de laatste woorden zijn voor jou. Om eerlijk te zijn zou ik niet weten waar ik moet beginnen om je te bedanken voor alles wat je voor me betekend hebt. Zonder jouw liefde, steun, flexibiliteit en begrip zou dit boekje er nooit gekomen zijn. Ik weet dat het niet altijd even makkelijk voor je is geweest (en ik misschien niet altijd even gezellig was), maar ik kon desondanks altijd bij je terecht. Lieve Inge, bedankt voor alles, voor wie je bent, en dat we samen zijn.

About the author

Sebastiaan (Bas) Kevin Spaans was born on November 22, 1985, in Leidschendam, the Netherlands. After finishing his pre-university education (VWO) at the Erasmus College in Zoetermeer in 2004, he started his BSc in Biology, with a minor in Management and Marketing, at the Wageningen University. In 2008, he switched to a MSc in Biotechnology at the same university, where he specialized in Cellular and Molecular Biotechnology. His Master thesis was performed in the Molecular Ecology group at the Laboratory of Microbiology of Wageningen University. Under the supervision of Dr. Detmer Sipkema and Prof. Hauke Smidt, he worked on the cultivation and genome analysis of *Halichondria panicea* associated micro-organisms, with the aim to find the producer(s) of natural occurring brominated compounds. In addition, he performed an internship at the Joint BioEnergy Institute based in the San Francisco Bay Area, USA, under the supervision of Dr. Nathan Hillson, Dr. Swapnil Chhabra and John Thorne. Here he was involved in the development of a new molecular tool for the rapid manipulation of chromosomal elements for strain improvement. After obtaining his MSc degree in 2010, he started a PhD project in the Bacterial Genetics group of the Laboratory of Microbiology of Wageningen University, under the supervision of his promotor Prof. John van der Oost and co-promotor Dr. Servé W.M. Kengen. During this PhD project, he worked on the engineering of the hyperthermophilic archaeon *Thermococcus kodakarensis* for improved biohydrogen production, which is presented in this thesis. Moreover, he also briefly worked as a guest employee at the Department of Microbiology of the Ohio State University, USA, where he was supervised by Dr. Thomas Santangelo and received a training in the genetic tools needed for his PhD project. He is currently working as a Science and Business Fellow at Janssen IDV, Leiden, which is part of the pharmaceutical companies of Johnson & Johnson.



overview of completed training activities

Discipline-specific activities

Meetings & conferences

- Archaea meeting (GRC), Lucca (IT), 2013
- Netherlands Biotechnology Congress 14 (NBC), Ede (NL), 2012
- Metabolic Engineering IX conference (ECI), Biarritz (FR), 2012
- Platform Molecular Genetics meeting (NWO/ALW), Lunteren (NL), 2012
- International Thermophiles Conference (FEMS), Big Sky, Montana (USA), 2011

Courses & Training

- Advanced course Microbial Physiology and Fermentation Technology (TU Delft), Delft (NL), 2012
- Training Genetic tools *Thermococcus kodakarensis* (OSU), Ohio (USA), 2011

General courses

- Career perspectives (WGS), Wageningen (NL), 2014
- Scientific Writing (WGS), Wageningen (NL), 2013
- Project & Time Management (WGS), Wageningen (NL), 2012
- Competence assessment (WGS), Wageningen (NL), 2011
- VLAG PhD week (VLAG), Baarlo (NL), 2011

Optionals

- Organization of PhD study trip USA & Canada, 2013
- Microbiology PhD study trip:
 - USA & Canada, 2013
 - Japan & China, 2011
- Preparation of research proposal, 2011
- Bacterial Genetics group meetings
- Microbiology PhD-PostDoc meetings

This work is part of the Open Programme (“Biological hydrogen production beyond current limits” file number 820.02.011), which is financed by the Netherlands Organisation for Scientific Research (NWO).