Post-Genomic Characterization of Metabolic Pathways in *Sulfolobus solfataricus*

Jasper Walther

Thesis committee

Thesis supervisors

Prof. dr. J. van der Oost
Personal chair at the laboratory of Microbiology
Wageningen University

Prof. dr. W. M. de Vos
Professor of Microbiology
Wageningen University

Other members

Prof. dr. W.J.H. van Berkel
Wageningen University

Prof. dr. V.A.F. Martins dos Santos
Wageningen University

Dr. T.J.G. Ettema
Uppsala University, Sweden

Dr. S.V. Albers
Max Planck Institute for Terrestrial Microbiology, Marburg, Germany

This research was conducted under the auspices of the Graduate School VLAG

Post-Genomic Characterization of Metabolic Pathways in *Sulfolobus solfataricus*

Jasper Walther

Thesis

Submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. dr. M.J. Kropff,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Monday 23 January 2012
at 11 a.m. in the Aula.

Jasper Walther
Post-Genomic Characterization of Metabolic Pathways in *Sulfolobus solfataricus*,
164 pages.

Thesis, Wageningen University, Wageningen, NL (2012)
With references, with summaries in Dutch and English

ISBN 978-94-6173-203-3

Table of contents

Introduction

Introduction

The discovery that many microorganisms thrive in extreme environments gave scientific grounds to the idea that life can exists not only on earth, but also on other planets both in our solar system and beyond. Analysis of life in these chemically and physically challenging environments revealed where life can exist and therefore where to look for extra-terrestrial life. Moreover, this discovery has also led to a major boost in the biotech industry by the application of different biomolecules that are produced by these extreme organisms. Notable examples include stable enzymes such as proteases (food industry), DNA polymerases (PCR reaction), and xylanases (paper bleaching) from thermophilic and other extremophilic microorganisms (Rothschild and Mancinelli 2001). Moreover, this has led to further stimulation of fundamental research on these extreme organisms as is summarized below.

Archaea: champions of extreme living

Archaea were first discovered by the pioneering work of Carl Woese in the 1970s using the small subunit rRNA sequences to classify organisms (Woese and Fox 1977). The domain Archaea is a second Prokaryotic domain, distinct from the Bacteria and the Eukarya. The cellular data that confirmed the Archaea as a separate branch include the structures of their RNA polymerases, a cell wall without peptidoglycan (Schafer 1996) and plasma membranes containing di and tetra ether-lipids (Reeve 1999). Although their lineage is distinct from the other two lineages, the release of genome information of many organisms showed that Archaea can also be seen as a chimeric of Eukarya and Bacteria. Their core metabolic functions resemble those of Bacteria, while their information processing functions are Eukaryotic (Allers and Mevarech 2005; Ettema, de Vos et al. 2005).

Previously it has been thought that the third domain of life, the Archaea, contained mainly extremophilic species apart from the methanogens that are found in the environmental samples and the gut of many animals. Nowadays it is known that Archaea are abundant in many non-extreme ecosystems like soils and oceans (Chaban, Ng et al. 2006). The domain Archaea does, however, contain some of the most extreme organisms known today. Some of these record holders of extreme living are: (I) *Methanopyrus kandleri*, which is able to grow at 122^{o}C at high pressure (Marteinsson, Birrien et al. 1999; Takai, Nakamura et al. 2008), (II)

Picrophilus torridus, which has an optimal pH of 0.8 and is capable of growth at pH of around zero (Schleper, Puehler et al. 1995), and (III) many Halobacteriales, which are able to grow at saturated NaCl levels (~5.2 M NaCl), some of them having unique morphologies (Boone, Castenholz et al. 2001).

Despite the fact that not all Archaea are extremophiles, most species that make up the two well-established archaeal phyla (Crenarchaeota, Euryarchaeota) (Huber, Hohn et al. 2002) can be found to grow under extreme conditions (Table 1). The Euryarchaeota branch contains organisms which are halophiles, methanogens, meso-, thermo-and barophiles. The Crenarchaeota branch contains hyperthermophilic, acidophilic, psychrophilic and mesophilic species. Due to relatively limited sequence information there is no consensus at present on the phylogenetic position of the (extremophilic) Korarchaeota and Nanoarchaeota. The species forming the Korarchaeota group consists only of species detected with sequence-based techniques applied on environmental samples. Recently an enrichment has been made from which the first complete genome sequence has been determined (Elkins, Podar et al. 2008). Nanoarchaeota appear most closely related to the Euryarchaeota (Brochier-Armanet et al. 2008). The first species of the Nanoarchaeota branch, *Nanoarchaeum equitans*, was reported in 2002 by the group of the extremophile pioneer Karl Stetter as the first symbiotic organism in the Archaeal domain (Huber et al. 2002). Its 0.49 Mb genome sequence reveals that the smallest Archaeal genome to date belongs to *N. equitans.* It is strictly dependent on *Ignicoccus hospitalis*; it cannot be grown in pure culture and no other known organism can support its growth (Waters, Hohn et al. 2003). In addition, comparative genomics has recently revealed the existence of a third archaeal phylum of the Thaumarchaeota (Brochier-Armanet et al. 2008;). Initially classified as ‘mesophilic Crenarchaeota’, the Thaumarchaea form a separate and deep-branching phylum that comprises all the known archaeal ammonia oxidizers (Pester, schleper,et al. 2011).

Table 1: Archaea and their habitats

Environmental Parameter	Type	Definition	Examples	Habitat	Reference
Temperature	Hyperthermophile	Growth >80°C	Sulfolobus	Terrestrial hot spring	(Zillig, Stetter et al. 1980)
			Pyrococcus	Hydrothermal vent	(González, Masuchi et al. 1998)
			strain 121	Hydrothermal vent	(Kashefi and Lovley 2003)
	Thermophile	Growth 60-80 °C	Methanobacterium	Sewage sludge	(Smith, Doucette-Stamm et al. 1997)
	Mesophile	15-60 °C	Methanosphaera	Human gut	(Fricke, Seedorf et al. 2006)
			Methanobrevibacter	Human mouth	(Brusa, Canzi et al. 1993)
	Psychrophile	<15 °C	Methanogenium	Antarctic lake	(Franzmann, Liu et al. 1997)
Salinity	Halophile	2-5 M NaCl	Haloquadratum	Hypersaline pool	(Bolhuis, Poele et al. 2004)
pH	Alkaliphile	pH >9	Natronobacterium	SodaLake	(Xin, Itoh et al. 2001)
	Acidophile	low pH loving (~ pH 2)	Picrophilus	Solfataric spring	(Schleper, Puehler et al. 1995)
Pressure	Barophile	Pressure > 400 atm	Thermococcus	Deep hydrothermal vent	(Marteinsson, Birrien et al. 1999)

Crenarchaeota

The Crenarchaeota, one of the principal kingdoms of the domain Archaea, include *Sulfolobus*, *Pyrobaculum*, and *Pyrodictium* spp. (Figure 1 and 2). The kingdom Crenarchaeota has been defined phylogenetically based on comparative molecular sequence analyses, and its members are therefore primarily defined by sequence similarity. However, like all Archaea, Crenarchaeota are prokaryotic, and possess ether-linked lipid membranes which contain isoprenoid side chains instead of fatty acids. Crenarchaeota cells cover a wide range in shape and size: from small cocci (<1µm in diameter) to long filaments (>100µm in length). The known species display a wide range of cell shapes, including regular cocci clustered in grape-like aggregates (*Staphylothermus*), irregular, lobed cells (*Sulfolobus*; Figure 3), discs (*Thermodiscus*), and almost rectangular rods (*Thermoproteus, Pyrobaculum*). Most species possess flagella (motility) and/or pili (adhesion); an example of the latter is *Pyrodictium* cells that are inter-connected by extensive networks of proteinaceous fibres (Rieger, Rachel, et al. 1995).

Metabolically, Crenarchaeota are quite diverse, ranging from chemo-organotrophs to chemo-lithoautotrophs. They are anaerobes, facultative anaerobes or aerobes, and many utilize sulphur in some way for energy metabolism. Several species are primary producers of organic matter, using carbon dioxide as sole carbon source, and gaining energy by the oxidation of inorganic substances like sulphur and hydrogen, and reduction of sulphur or nitrate. Others grow on organic substrates by aerobic or anaerobic respiration or by fermentation (Chaban, Ng, et al. 2006).

The most spectacular feature of the Crenarchaeota, however, is their tolerance to, and even preference for, extremes of acidity and temperature. While many prefer neutral to slightly acidic pH ranges, members of the Crenarchaeal order Sulfolobales flourish at pH 1-2 and die above pH 7. Optimum growth temperatures range from 75 to 105°C and the maximum temperature of Crenarchaeal growth can be as high as 121°C (*Pyrodictium*-like Strain 121; Kashefi and Lovley 2003). Moreover, *Picrophilus torridus*, has an optimal pH of 0.8 and grows at a pH value of zero (Schleper, Puehler et al. 1995).

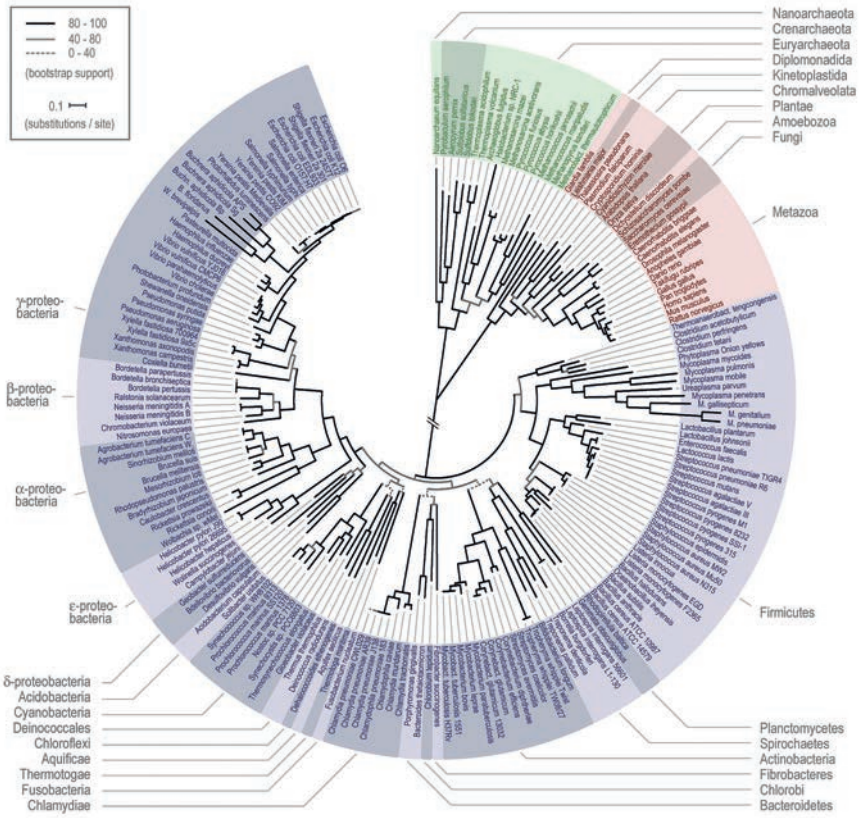

Figure 1: Phylogenetic tree of fully sequenced genomes. The branch separating bacteria from Eukarya and Archaea is shortened for display purposes (Ciccarelli, Doerks et al. 2006).

Sulfolobus species as a model organism

Sulfolobus species are considered model Archaea because of their global abundance (DeLong and Pace 2001) and their relatively easy cultivation on a variety of carbon sources (Grogan 1989). Moreover, they often possess mobile genetic elements, viruses, small plasmids and large conjugative plasmids (Zillig,

Arnold et al. 1998; Lipps 2006), which has resulted in the establishment of genetic tools (Albers, Jonuscheit et al. 2006; Berkner, Grogan et al. 2007).

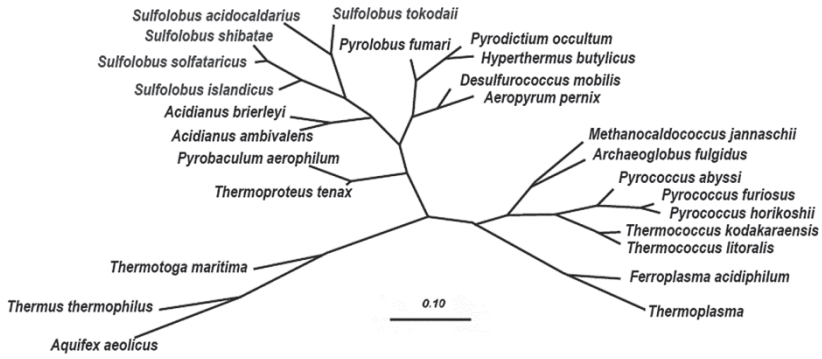

Figure 2: Phylogenetic tree based on Archaeal and Bacterial 16S rRNA sequences (Brouns, Walther et al. 2006). Here the phylogeny of the important *Sulfolobus* species is illustrated (red).

Since its original discovery in Yellowstone (Brock, Brock et al. 1972), species of the genus *Sulfolobus* have been isolated from various solfataric fields, such as *Sulfolobus solfataricus* in Pisciarelli, Italy (Zillig, Stetter et al. 1980), a species provisionally called *Sulfolobus islandicus* in Sogasel, Iceland (Zillig, Kletzin et al. 1994) and *Sulfolobus tokodaii* from Japan's Beppu Hot Springs (Suzuki, Iwasaki et al. 2002). The habitat of the *Sulfolobus* spp. is generally hot acidic mud. Their growth is optimal at temperatures between 80 and 85 °C and at low pH (between 2 and 4). Because of this extreme environment they require protection against the large proton gradient between the interior of the cell (pH 5.5) and the exterior (pH 2-4) (Grogan 2000). To counteract this pH gradient, the membrane potential is inversed, i.e. positive on the inside, in contrast to all other cellular life forms, which have a negative potential on the inside of the membrane (Moll, et al. 1988). Moreover, thermophilic and acidophilic organisms like *S. solfataricus* possess membrane-spanning tetra-ether lipids that form a rigid monolayer membrane, which is nearly impermeable to ions and protons, an important property for maintaining the proton gradient (Slonzewski, Fujisawa, et al. 2009). Their morphology is irregular and lobe-shaped with cell diameters from 0.2 to 2 µm (Figure 3).

Sulfolobus species are generally aerobic, and heterotrophic growth has been reported, during which a range of carbohydrates, yeast extract, and peptide mixtures are oxidized to CO_2 (Grogan 1989; Schönheit and Schäfer 1995). In addition, both autotrophic oxidation, of $S_2O_3^{2-}$, $S_4O_6^{2-}$, S_o and S^{2-} to sulphuric acid, and of H_2 to water, and heterotrophic growth has been described for *S. acidocaldarius* (Shivvers and Brock 1973; Schönheit and Schäfer 1995). It has been suggested that anaerobic respiration (e.g., reduction of NO) by certain Sulfolobales might be possible (She, Singh et al. 2001).

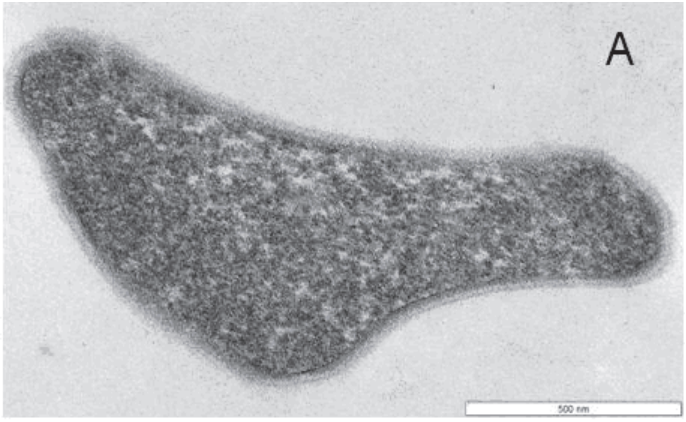

Figure 3: *Sulfolobus solfataricus*. Adjusted from (Ortmann, Brumfield et al. 2008).

Other closely related genera (Figure 2) that have been relatively well-characterized include *Acidianus* (order Sulfolobales, obligatory chemolitho-autotroph, aerobic S_0 oxidation to sulphuric acid, anaerobic S_0 reduction coupled to H_2 oxidation), *Hyperthermus* (order Desulfurococcales, anaerobic, amino acid fermentation) and *Aeropyrum* (order Desulfurococcales, aerobe, heterotroph on starch and peptides). *Aeropyrum pernix* was the first Crenarchaeote for which the complete genome was sequenced (Kawarabayasi, Hino et al. 1999). Its genome size was found to be around 1.7 Mb.

Exploitation of the Archaeal potential

Many industrial processes can become more environmental friendly, sustainable and cost effective by using the large biocatalytic potential of microbial

enzymes. Much research has been done on the applicability of enzymes or whole cells as a cell factory. Enzymes derived from extreme organisms (also called extremozymes) like (hyper)thermophiles are generally very stable with relatively long shelf lives and the ability to work at harsh conditions, high temperature, extreme pH, high pressure and tolerate organic solvents. These conditions have the added benefit that they ensure a reduced risk of contamination and in some cases resulting in an increased solubility of the substrate. A good example of the applicability of heat stable enzymes is in the processing of starch as is detailed below (Crabb and Shetty 1999; Bruins, Janssen et al. 2001).

The conversion of starch to more valuable products such as dextrins, glucose, fructose and trehalose, requires high temperatures. The starch is heated to 95-105^{o}C in order to liquefy the substrate and to make it more accessible to enzymatic attack. To degrade the starch, different enzymes are used, including amylase, glucoamylase, and glucose isomerase. Thermostable enzymes are an obvious choice because they will remain active at these elevated temperatures at which most enzymes from mesophilic organisms will quickly denature. The large scale production of these extreme enzymes is usually done by well-known mesophilic production organisms like *Escherichia coli*, *Bacillus subtilis* and yeast because the knowledge that is already gathered about these organisms ensures a much higher (recombinant) protein yield then in the natural host or another Archaeal expression organism.

Despite the fact that Archaeal fermentation processes often require specialised bioreactor systems and generally result in low biomass concentrations and low productivity, there are examples where they are utilized on a large scale. An example of such a utilization is the bio-oxidation of gold-bearing arsenopyrite-pyrite by *Sulfolobus* cells (Lindström and Gunneriusson 1990). The latter study describes a laboratory scale semi-continuous reactor to enable gold liberation from sulphides by *Sulfolobus* cells. This process is exploited by gold mining companies in South Africa. Some of the largest plants connect several modules of reactors in series, treating approximately 1000 tonnes of gold a day (Norris, Burton et al. 2000) proving the economic feasibility of using whole Archaeal cells as cell-factories.

Post genomic challenges

The first genome sequence of Sulfolobales was the 3.0 Mb genome of *S. solfataricus* (She et al 2001). Following the detailed analysis of this genome, others were established from related *Sulfolobus* species and comparative genomics analyses were performed to predict relevant physiological functions for many genes (She, Singh et al. 2001; Chen, Brugger et al. 2005). As in all studied genomes, many hypothetical genes were found for which a function could not reliably be predicted. Hence, the main challenge of the post-genome era is to integrate classical approaches (physiology, biochemistry, and molecular genetics) with genomics-based, high-throughput approaches (comparative, functional, and structural genomics). In the case of *Sulfolobus* genomes the obvious goals are to (i) identify missing links in central metabolic pathways (degradation and biosynthesis of carbohydrates hexose and pentose, amino acids, nucleotides, and vitamins); (ii) elucidate functions of hypothetical proteins (e.g., those conserved in Archaea and/or Eukarya); and (iii) unravel global regulatory circuits, such as the control of RNA and protein turnover (transcription/exosome, translation/proteasome). Moreover, understanding essential details of the Archaeal cell is not only scientifically very interesting, but this may also contribute to its future application as an industrial “cell factory”.

Aim and outline of this thesis

This thesis presents the results of several integrated approaches to clarify the composition and regulation of the (sugar) metabolism of *Sulfolobus solfataricus*. Using an integrated approach by combining transcriptome, proteome and biochemistry data, we have set the stage for Archaeal systems biology. In this thesis we shed light on the tools needed to obtain a global view of *S. solfataricus* and describe the use of these tools in addressing the unusual metabolism of this acidothermophile. This includes the elucidation of a new pentose degrading pathway, a carotenoid producing pathway and the complete reconstruction of the central metabolic pathways. We have chosen to work with *S. solfataricus* because it is a model Archaeon able to grow at very high temperatures (80^{o}C) and very low pH (1-6). Because of its extremophilic nature, *S.solfataricus* exhibits some unique properties that are very interesting for industrial applications, namely the presence of very stable enzymes able to operate well under harsh conditions. A spin-off of the

here presented work would be the discovery of such enzymes or characteristic metabolic pathways that could be applied in the industry.

Chapter 1 gives a general overview of the characteristics of Archaea with a focus on the model organism *S. solfataricus* and its relatives. It describes the unique properties of Archaea, some of their most extreme habitats and their industrial potential.

Chapter 2 provides an overview of transcriptomic data available for hyperthermophilic Archaea and what new insights it has generated. It also tries to look at the future generation of transcriptomic analysis, its pitfalls and its benefits.

Chapter 3 presents the multi-disciplinary and integrated analysis of the reconstructed central carbohydrate metabolism of *S. solfataricus*. Here Archaeal proteomic and transcriptomic data are for the first time combined in order to get more insight into the regulation of these pathways when two very different carbon sources are provided, namely during growth either on glucose-containing minimal medium or on a peptide-containing rich medium. We could see little fluctuations on protein or transcriptional level, suggesting other ways of regulation for these important pathways.

Chapter 4 presents the great strength of systems approaches that combine a wide range of techniques: (i) transcriptome and proteome analysis, (ii) bioinformatics prediction of gene function and transcription regulator binding site, (iii) heterologous enzyme production and activity analysis. This has resulted in the elucidation of a novel pentose-degrading pathway in prokaryotes. Comparative genomics has been used to reconstruct a scenario for the evolution of this pathway.

Chapter 5 describes the adaptation of *S. solfataricus* to different oxygen levels. *S. solfataricus* is an obligate aerobic organism and its growth rate is therefore highly dependent on the available oxygen. Here the organism is grown under different oxygen levels, and its adaptation to the different oxygen levels is analysed by integrating biochemical and transcriptome data.

Chapter 6 describes the carotenoid biosynthesis pathways in different *Sulfolobus* species. For this a microarray study was done on *S. solfataricus*, *S. shibatae* and a *S. shibatae* carotene-overproducing mutant strain. These organisms mainly produce zeaxanthin (a vital molecule for the continued function of the human eye) and glycosylated zeaxanthin. Genes of different *Sulfolobus* species were cloned in a zeaxanthin-overproducing strain of *E. coli* to show their zeaxanthin modifying capabilities.

Chapter 7 summarizes this thesis, reflects on the obtained results and provides an outlook into some of the future perspectives. Finally, an evaluation is presented of the quickly growing genetic toolkit available for engineering *S. solfataricus*.

References

Albers, S. V., M. Jonuscheit, et al. (2006). Production of recombinant and tagged proteins in the hyperthermophilic archaeon Sulfolobus solfataricus. Appl Environ Microbiol. 72(1): 102-111.

Allers, T. and M. Mevarech (2005). Archaeal genetics - the third way. Nat Rev Genet. 6(1): 58-73.

Berkner, S., D. Grogan, et al. (2007). Small multicopy, non-integrative shuttle vectors based on the plasmid pRN1 for Sulfolobus acidocaldarius and Sulfolobus solfataricus, model organisms of the (cren-)archaea. Nucleic Acids Res. 35(12): e88.

Bolhuis, H., E. M. T. Poele, et al. (2004). Isolation and cultivation of Walsby's square archaeon. Environmental Microbiology. 6(12): 1287-1291.

Boone, D. R., R. W. Castenholz, et al. (2001). Bergey's manual of systematic bacteriology / George M. Garrity, editor-in-chief. New York, Springer.

Brochier-Armanet C, Boussau B, Gribaldo S, Forterre P. (2008) Mesophilic Crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. Nat Rev Microbiol. 6(3):245-252.

Brock, T. D., K. M. Brock, et al. (1972). Sulfolobus: a new genus of sulfur-oxidizing bacteria living at low pH and high temperature. Archiv Für Mikrobiologie. 84(1): 54-68.

Brouns, S. J., J. Walther, et al. (2006). Identification of the missing links in prokaryotic pentose oxidation pathways: evidence for enzyme recruitment. J Biol Chem. 281(37): 27378-27388.

Bruins, M., A. Janssen, et al. (2001). Thermozymes and their applications. Applied Biochemistry and Biotechnology. 90(2): 155-186.

Brusa, T., E. Canzi, et al. (1993). Methanogens in the human intestinal tract and oral cavity. Current Microbiology. 27(5): 261-265.

Chaban, B., S. Y. Ng, et al. (2006). Archaeal habitats--from the extreme to the ordinary. Can J Microbiol. 52(2): 73-116.

Chen, L., K. Brugger, et al. (2005). The genome of Sulfolobus acidocaldarius, a model organism of the Crenarchaeota. J Bacteriol. 187(14): 4992-4999.

Ciccarelli, F. D., T. Doerks, et al. (2006). Toward automatic reconstruction of a highly resolved tree of life. Science (New York, N.Y.) . 311(5765): 1283-1287.

Crabb, W. D. and J. K. Shetty (1999). Commodity scale production of sugars from starches. Curr Opin Microbiol. 2(3): 252-256.

DeLong, E. F. and N. R. Pace (2001). Environmental diversity of bacteria and archaea. Systematic Biology. 50(4): 470-478.

Elkins, J. G., M. Podar, et al. (2008). A korarchaeal genome reveals insights into the evolution of the Archaea. Proc Natl Acad Sci U S A. 105(23): 8102-8107.

Ettema, T. J., W. M. de Vos, et al. (2005). Discovering novel biology by in silico archaeology. Nat Rev Microbiol. 3(11): 859-869.

Franzmann, P. D., Y. Liu, et al. (1997). Methanogenium frigidum sp. nov., a psychrophilic, H2-using methanogen from Ace Lake, Antarctica. International Journal of Systematic Bacteriology. 47(4): 1068-1072.

Fricke, W. F., H. Seedorf, et al. (2006). The genome sequence of Methanosphaera stadtmanae reveals why this human intestinal archaeon is restricted to methanol and H2 for methane formation and ATP synthesis. Journal of Bacteriology. 188(2): 642-658.

González, J. M., Y. Masuchi, et al. (1998). Pyrococcus horikoshii sp. nov., a hyperthermophilic archaeon isolated from a hydrothermal vent at the Okinawa Trough. Extremophiles: Life Under Extreme Conditions. 2(2): 123-130.

Grogan, D. W. (1989). Phenotypic characterization of the archaebacterial genus Sulfolobus: comparison of five wild-type strains. J Bacteriol. 171(12): 6710-6719.

Grogan, D. W. (2000). The question of DNA repair in hyperthermophilic archaea. Trends in Microbiology. 8(4): 180-185.

Huber, H., M. J. Hohn, et al. (2002). A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont. Nature. 417(6884): 63-67.

Kashefi, K. and D. R. Lovley (2003). Extending the Upper Temperature Limit for Life. Science. 301(5635): 934.

Kawarabayasi, Y., Y. Hino, et al. (1999). Complete genome sequence of an aerobic hyper-thermophilic crenarchaeon, Aeropyrum pernix K1. DNA Research: An International Journal for Rapid Publication of Reports on Genes and Genomes. 6(2): 83-101, 145-152.

Lindström, E. B. and L. Gunneriusson (1990). Thermophilic bioleaching of arsenopyrite using Sulfolobus and a semi-continuous laboratory procedure. Journal of Industrial Microbiology and Biotechnology. 5(6): 375-382.

Lipps, G. (2006). Plasmids and viruses of the thermoacidophilic crenarchaeote Sulfolobus. Extremophiles: Life Under Extreme Conditions. 10(1): 17-28.

Marteinsson, V. T., J. L. Birrien, et al. (1999). Thermococcus barophilus sp. nov., a new barophilic and hyperthermophilic archaeon isolated under high hydrostatic pressure from a deep-sea hydrothermal vent. International Journal of Systematic Bacteriology. 49 Pt 2: 351-359.

Matsumi, R., Atomi, H., Driessen, A.J., Van der Oost, J. (2011) Isoprenoid biosynthesis in Archaea - Biochemical and Evolutionary implications. Res. Microbiol. 162, 39-52

Moll, R. and Schäfer, G. (1988). Chemiosmotic H^+ cycling across the plasma membrane of the thermoacidophilic archaebacterium Sulfolobus acidocaldarius. FEBS Letters. 232(2): 359-363

Norris, P. R., N. P. Burton, et al. (2000). Acidophiles in bioreactor mineral processing. Extremophiles. 4(2): 71-76.

Ortmann, A. C., S. K. Brumfield, et al. (2008). Transcriptome analysis of infection of the archaeon Sulfolobus solfataricus with Sulfolobus turreted icosahedral virus. J Virol. 82(10): 4874-4883.

Pester M, Schleper C, Wagner M. (2011). The Thaumarchaeota: an emerging view of their phylogeny and ecophysiology. Curr Opin Microbiol. 14(3):300-306

Reeve, J. N. (1999). Archaebacteria then ... Archaes now (are there really no archaeal pathogens?). J Bacteriol. 181(12): 3613-3617.

Rieger, G., R. Rachel, R. Herman, K. O. Stetter. (1995). Ultrastructure of the Hyperthermophilic Archaeon Pyrodictium abyssi. J. of struct. biol. 115: 78-87.

Rothschild, L. J. and R. L. Mancinelli (2001). Life in extreme environments. Nature. 409(6823): 1092-1101.

Schafer, G. (1996). Bioenergetics of the archaebacterium Sulfolobus. Biochim Biophys Acta. 1277(3): 163-200.

Schelert, J., V. Dixit, V. Hoang, J. Simbahan, M. Drozda, and P. Blum. 2004. Occurrence and characterization of mercury resistance in the hyperthermophilic archaeon Sulfolobus solfataricus by use of gene disruption. J. Bacteriol. **186:** 427-437.

Schleper, C., G. Puehler, et al. (1995). Picrophilus gen. nov., fam. nov.: a novel aerobic, heterotrophic, thermoacidophilic genus and family comprising archaea capable of growth around pH 0. J Bacteriol. 177(24): 7050-7059.

Schönheit, P. and T. Schäfer (1995). Metabolism of hyperthermophiles. World Journal of Microbiology and Biotechnology. 11(1): 26-57.

She, Q., R. K. Singh, et al. (2001). The complete genome of the crenarchaeon Sulfolobus solfataricus P2. Proceedings of the National Academy of Sciences of the United States of America. 98(14): 7835-7840

Shivvers, D. W. and T. D. Brock (1973). Oxidation of elemental sulfur by Sulfolobus acidocaldarius. Journal of Bacteriology. 114(2): 706-710.

Solonczewski, J. L., M. Fujisawa, M. Dopson, T. A. Krulwich (2009). Cytoplasmic pH Measurement and Homeostasis in Bacteria and Archaea. Adv. in Micr. Phys. 58: 1-79.

Smith, D. R., L. A. Doucette-Stamm, et al. (1997). Complete genome sequence of Methanobacterium thermoautotrophicum deltaH: functional analysis and comparative genomics. Journal of Bacteriology. 179(22): 7135-7155.

Suzuki, T., T. Iwasaki, et al. (2002). Sulfolobus tokodaii sp. nov. (f. Sulfolobus sp. strain 7), a new member of the genus Sulfolobus isolated from Beppu Hot Springs, Japan. Extremophiles: Life Under Extreme Conditions. 6(1): 39-44.

Takai, K., K. Nakamura, et al. (2008). Cell proliferation at 122 degrees C and isotopically heavy CH4 production by a hyperthermophilic methanogen under high-pressure cultivation. Proc Natl Acad Sci U S A. 105(31): 10949-10954.

Waters, E., M. J. Hohn, et al. (2003). The genome of Nanoarchaeum equitans: insights into early archaeal evolution and derived parasitism. Proceedings of the National Academy of Sciences of the United States of America. 100(22): 12984-12988.

Woese, C. R. and G. E. Fox (1977). Phylogenetic structure of the prokaryotic domain: the primary kingdoms. Proc Natl Acad Sci U S A. 74(11): 5088-5090.

Xin, H., T. Itoh, et al. (2001). Natronobacterium nitratireducens sp. nov., a aloalkaliphilic archaeon isolated from a soda lake in China. International Journal of Systematic and Evolutionary Microbiology. 51(Pt 5): 1825-1829.

Zillig, W., H. P. Arnold, et al. (1998). Genetic elements in the extremely thermophilic archaeon Sulfolobus. Extremophiles: Life Under Extreme Conditions. 2(3): 131-140.

Zillig, W., A. Kletzin, et al. (1994). Screening for Sulfolobales, their Plasmids and their Viruses in Icelandic Solfataras. Systematic and applied microbiology. 16(4): 609-628.

Zillig, W., K. O. Stetter, et al. (1980). The Sulfolobus-"Caldariella" group: Taxonomy on the basis of the structure of DNA-dependent RNA polymerases. Archives of Microbiology. 125(3): 259-269.

Zhang C, Guo L, Deng L, Wu Y, Liang Y, Huang L, She Q. (2010) Revealing the essentiality of multiple archaeal pcna genes using a mutant propagation assay based on an improved knockout method. Microbiology 156: 3386-3397.

Hot Transcriptomics

Jasper Walther*, Pawel Sierocinski*, John van der Oost.
* authors contributed equally

Archaea 2010; vol. 2010: 14 pages. doi:10.1155/2010/89758

Abstract

DNA microarray technology allows for a quick and easy comparison of complete transcriptomes, resulting in improved molecular insight in fluctuations of gene expression. After emergence of the microarray technology about a decade ago, the technique has now matured and has become routine in many molecular biology laboratories. Numerous studies have been performed that have provided global transcription patterns of many organisms under a wide range of conditions. Initially, implementation of this high-throughput technology has lead to high expectations for ground breaking discoveries. Here an evaluation is performed of the insight that transcriptome analysis has brought about in the field of hyperthermophilic archaea. The examples that will be discussed have been selected on the basis of their impact, either in terms of biological insight or technological progress.

Thermophiles

Forty years ago it was generally accepted that life was not possible at temperatures higher than 60°C. In 1969, however, Thomas Brock and co-workers discovered that the upper temperature limit goes as high as 75°C when micro-organisms were isolated from thermal springs in Yellowstone National Park (Brock and Freeze 1969; Stetter 2006). The pioneering work of Brock set the stage for further exploration of a wide range of thermal ecosystems. Numerous micro-organisms defined as thermophiles have since been found to thrive optimally between 50-80°C, but also many appeared to have their optimal temperature for growth from 80°C to well above 100°C, the hyperthermophiles. Recently it has been shown that some archaea can endure temperatures as high as 122°C and even proliferate in such conditions. Although there are several bacterial representatives in the group as well, most of the known hyperthermophiles belong to the archaea.

Thermophilic organisms can be found in water-containing geothermally heated environments. These volcanic ecosystems are mainly situated along terrestrial and submarine fracture zones where tectonic plates are converging or diverging. The terrestrial biotopes of (hyper)thermophiles are mainly aerobic, sulfur containing solfataric fields with temperature as high as 100°C (depending on the altitude) and the pH in a dual range: either acidic (values from below zero to 4.0 (Angelov and Liebl 2006)) or neutral to slightly alkali (7.0 - 9.0) (Segerer, Burggraf et al. 1993). The marine biotopes for (hyper)thermophiles consist of different hydrothermal systems ranging from shallow to abyssal depths. Temperatures in those anaerobic environments can range up to 400°C and the pH is usually in the range of 5.0 to 8.5.

Progress in culturing thermophilic archaea and in the revolution of DNA sequencing technology has resulted in a rapidly increasing amount of (meta)genomic data on these extreme microorganisms. This has not only led to the discovery of robust biocatalysts, but also to fundamental insight into: (i) physiology - including unique metabolic enzymes, pathways and regulation; (Makarova and Koonin 2003; Ettema, de Vos et al. 2005; Brouns, Walther et al. 2006), (ii) biochemistry - the molecular basis of thermostability of bio-molecules (Cambillau and Claverie 2000; Kumar and Nussinov 2001; Koutsopoulos, van der Oost et al. 2007), and (iii) phylogeny - theories on the evolution of the eukaryotic cell (Rivera and Lake 2004).

The first complete genome analysis of an archaeon, *Methanocaldococcus jannaschii* (Bult, White et al. 1996), was a big step towards confirmation of the monophyletic position of the archaea, with respect to the bacteria and the eukaryotes. In addition, archaea appeared to possess a bacterial-like compact chromosomal organization with clustering of genes as polycistronic units (operons), and with only few interrupted genes (introns). Moreover, the archaeal systems that drive the flow of genetic information (transcription, translation, replication, DNA repair) generally correspond to the core of the eukaryal counterparts. These initial observations of bacterial-like "information storage" and eukaryal-like "information processing" have been confirmed by the analyses of subsequently sequenced hyperthermophilic model archaea: the euryarchaea *Pyrococcus* spp. (*P. furiosus, P. abyssi, P. horikoshii*) as well as the crenarchaea *Sulfolobus* spp. (*S. solfataricus, S. tokodaii, S. acidocaldarius*) (Makarova and Koonin 2003). The comparative analysis of the genome of the hyperthermophilic bacterium *Thermotoga maritima* to *Pyrococcus furiosus* (both isolated from shallow thermal vents at the same beach (Volcano, Italy)), led to the conclusion that horizontal (or lateral) gene transfer substantially contributes to the apparent high degree of genome flexibility (Nelson, Clayton et al. 1999; Koonin, Wolf et al. 2001). In addition, the comparison of closely related species (*P. furiosus, P. abyssi, P. horikoshii*) revealed a high degree of genome plasticity. It was also proposed that the lateral gain, as well as the loss of genes is a modular event (Ettema, van der Oost et al. 2001). Horizontal gene transfer has also been proposed to explain the relatively high degree of homology between genomic loci of the euryarchaeon *Thermoplasma acidophilum* and the crenarchaeon *S. solfataricus*, phylogenetically distant archaea, that inhabit the same environmental niche (65-85°C, pH 2.0). The *Sulfolobus*-like genes in the *T. acidophilum* genome are clustered into at least five discrete regions, again indicating modular recombination of larger DNA fragments (Ruepp, Graml et al. 2000; Frickey and Lupas 2004).

After establishing a genome sequence, comparative genomics analyses is performed to assign potential functions for the identified open reading frames. In the majority of the studied prokaryotic genomes, the fraction of hypothetical and conserved hypothetical genes amounts to 40-60% of the coding regions (Doerks, von Mering et al. 2004). Hence, one of the main challenges of the post-genome era still is to improve the functional annotation of genes by integrating classical approaches (physiology, biochemistry and molecular genetics) with genomics-

based high-throughput approaches (comparative, functional and structural genomics). Obvious targets of comparative and functional analysis of archaeal genomes are the numerous missing links in metabolic pathways as well as the largely unknown regulatory systems with either eukaryal or bacterial characteristics (Makarova and Koonin 2003; Ettema, de Vos et al. 2005).

Archaeal transcriptomics

DNA microarrays have initially been established as high-throughput functional genomics tools to study eukaryotic and bacterial model systems. Initial assumptions suggested that microarray can be used as a general research tool (Ramsay 1998), however after more than a decade of experience it should be concluded that the application of microarray has its pros and cons. The choice of possible microarray approaches ranges from rather simple layouts comparing two states, to relatively complicated multi-state experimental hybridization schemes. The development of appropriate analytical methods has appeared to be a crucial requirement to enable analysis of the more complicated experimental designs, and to allow drawing conclusions from relatively small differences in expression profiles. Consequently, high quality microarray analyses not only require careful experimentation (cultivation, nucleic acid analysis, hybridization) but also state-of-the-art data processing. This has allowed for the high resolution analysis of time course experiments (Lundgren and Bernander 2007) and of multi-condition experiments (Auernik and Kelly 2010). In most recent studies, the majority of DNA microarrays are used either (i) as a pilot experiment that should provide leads for further investigations (Brouns, Walther et al. 2006), (ii) as a refinement tool to confirm previous gene expression studies (Trauger, Kalisak et al. 2008), or (iii) as one of many high throughput methods to be integrated in a systems biology analysis (Albers, Birkeland et al. 2009). Below, selected examples of transcriptome analyses of (hyper)thermophilic archaea are described in more detail. Selection is has been based on technological and/or scientific impact. An overview of archaeal transcriptome studies can be seen in table 1.

Table 1: A list of different archaeal transcriptome publications. This table shows that transcriptome studies are mostly done to elucidate metabolic processes or the behaviour of different archaea in stress situations. The publications are sorted by subject. Per subject the publications are sorted by year of publication. We included some environmental studies because they give a crucial insight in the ecological function of archaeal species. We excluded some of these publications because in our view they focused more on non-archaeal species, which is a subject not related to this article. The studies referring to thermophiles are in bold. The studies more described in this paper in more detail are marked with an asterisk next to the reference.

Organism	Subject Studied	Reference
	Metabolism	
Pyrococcus furiosus	Sulfur metabolism	(Schut, Zhou et al. 2001)*
Halobacterium salinarum NRC-1	Adaptation to phototrophy	(Baliga, Pan et al. 2002)
Haloferax volcanii	Central carbon metabolism	(Zaigler, Schuster et al. 2003)
Pyrococcus furiosus	Central carbon metabolism	(Schut, Brehm et al. 2003)*
Halobacterium salinarum NRC-1	Anaerobic respiration	(Muller and DasSarma 2005)
Methanosarcina mazei	Metabolism of methanogenic substrates	(Hovey, Lentes et al. 2005)
Sulfolobus solfataricus	Central carbon metabolism	(Snijders, Walther et al. 2006)
Sulfolobus solfataricus	Pentose metabolism	(Brouns, Walther et al. 2006)*
Methanosarcina barkeri	Methanogen metabolism/methods	(Culley, Kovacik et al. 2006)
Methanosarcina mazei	Nitrogen metabolism and regulation	(Veit, Ehlers et al. 2006)
Pyrococcus furiosus	Starch metabolism	(Lee, Shockley et al. 2006)
Pyrococcus furiosus	Metabolism of elemental sulfur	(Schut, Bridger et al. 2007)
Halobacterium salinarum R1	Adaptation to phototrophy	(Twellmeyer, Wende et al. 2007)
Methanosarcina acitovorans	Acetate and methanol metabolism	(Li, Li et al. 2007)
Environmental array	Ammonium oxidation	(Rich, Dale et al. 2008)
Metallosphaera sedula	Electron transport chain	(Auernik and Kelly 2008)
Methanosarcina	Methanogenesis	(Ferry and Lessner 2008)
Pyrobaculum aerophilum	Terminal electron acceptor studies	(Cozen, Weirauch et al. 2009)
Thermoproteus tenax	Central carbohydrate metabolism	(Zaparty, Tjaden et al. 2008)
Halobacterium salinarum R1	Phosphate-dependent behaviour	(Wende, Furtwangler et al. 2009)
Halobacterium salinarum NRC-1	Global response to nutrient availability	(Schmid, Reiss et al. 2009)
Haloferax volcanii	D-Xylose metabolism	(Johnsen, Dambeck et al. 2009)
Methanosarcina mazei	Response to nitrogen availability	(Jager, Sharma et al. 2009)
Metallosphaera sedula	Auto- hetero- and mixotrophic growth	(Auernik and Kelly 2010)
Metallosphaera sedula	Bioleaching	(Auernik and Kelly 2010)
	Stress	
Pyrococcus furiosus	Heat shock response	(Shockley, Ward et al. 2003)*
Pyrococcus furiosus	Cold shock response	(Weinberg, Schut et al. 2005)
Halobacterium salinarum NRC-1	UV irradiation	(McCready, Muller et al. 2005)
Methanocaldococcus janaschii	Heat and cold shock	(Boonyaratanakornkit, Simpson et al. 2005)
Methanosarcina barkeri	Heat shock and air exposure	(Zhang, Culley et al. 2006)

Methanocaldococcus janaschii	Pressure stress	(Boonyaratanakornkit, Cordova et al. 2006)
Pyrococcus furiosus	Response to gamma irradiation	(Williams, Lowe et al. 2007)
Methanosarcina mazei	Salt adaptation	(Pfluger, Ehrenreich et al. 2007)
Methanococcus maripaludis	H-limitation and growth rate	(Hendrickson, Haydock et al. 2007)
Halobacterium salinarum NRC-1	Response to change in temperature and salinity	(Coker, Dassarma et al. 2007)
Sulfolobus solfataricus	UV irradiation	(Fröls, Gordon et al. 2007)
Sulfolobus solfataricus; S. acidocaldarius	UV irradiation	(Dorazi, Gotz et al. 2007)
Sulfolobus solfataricus	Heat Shock Response	(Tachdjian and Kelly 2006)*
Halobacterium salinarum NRC-1	UV irradiation	(Baliga, Pan et al. 2002)
Sulfolobus solfataricus	Oxygen stress	(Simon, Walther et al. 2009)
Sulfolobus solfataricus	Oxygen stress	(Kirkpatrick 2010)
Methanococcoides burtonii	Heat stress	(Campanaro, Williams et al. 2010)
Thermococcus kodakaraensis	Heat stress	(Kanai, Takedomi et al. 2010)
Pyrococcus furiosus	Heat stress	(Keese, Schut et al. 2010)
Sulfolobus solfataricus	Heat stress	(Cooper, Daugherty et al. 2009)
Pyrococcus furiosus	Oxidative stress	(Strand, Sun et al. 2010)
Methanohalophilus portucalensis	Hypo- and Hyper-salt stress	(Shih and Lai 2010)
	Replication	
Sulfolobus solfataricus;S. acidocaldarius	Origin of replication	(Lundgren, Andersson et al. 2004)*
Halobacterium salinarum NRC-1	Cell cycle regulation	(Baumann, Lange et al. 2007)
Pyrococcus abyssi	Origin of replication	(Matsunaga, Glatigny et al. 2007)
Sulfolobus acidocaldarius	Cell cycle	(Lundgren and Bernander 2007)*
	Various	
Environmental array	Methanotroph diversity in landfills	(Stralis-Pavese, Sessitsch et al. 2004)
Pyrococci	Genomic DNA hybridization	(Hamilton-Brehm, Schut et al. 2005)
Sulfolobus solfataricus;S. acidocaldarius	RNA decay	(Andersson, Lundgren et al. 2006)
Methanococcus maripaludis	Mutant studies	(Xia, Hendrickson et al. 2006)
Haloferax volcanii	Promoter studies	(Lange, Zaigler et al. 2007)
Thermococcus kodakaraensis	Promotor studies	(Kanai, Akerboom et al. 2007)
Thermococcus kodakaraensis	Archaeal operon prediction	(Santangelo, Cubonova et al. 2008)
Haloferax volcanii	Deletion mutant analysis	(Dambeck and Soppa 2008)
Environmental array	Detection of acidophilic activity	(Garrido, Gonzalez-Toril et al. 2008)
Sulfolobus solfataricus	Viral infection	(Ortmann, Brumfield et al. 2008)*
Sulfolobus	Genomic hybridizations	(Grogan, Ozarzak et al. 2008)
Sulfolobus	Transcription bias near OriC	(Andersson, Pelve et al. 2010)

Sulfolobus solfataricus	Single base resolution map of the genome	(Wurtzel, Sapra et al. 2010)*
Environmental array	Antarctic soil community	(Yergeau, Schoondermark-Stolk et al. 2009)
0Methanosarcina acetivorans	Regulation of genes	(Reichlen, Murakami et al. 2010)
Halobacterium salinarum R1	Control of multiple genes by regulatory proteins	(Schwaiger, Schwarz et al. 2010)
Haloacterium salinarum NRC-1	Physiological readjustments during growth	(Facciotti, Pang et al. 2010)
Environmental array	Methanogens in cattle excreta	(Goberna, Gadermaier et al. 2010)
Environmental array	Gene transfer	(Parnell, Rompato et al. 2010)

Sulfur metabolism

The first microarray analysis reported on either a hyperthermophilic archaeon was a pilot study on *P. furiosus* that focussed on a subset of 271 metabolic genes (Schut, Zhou et al. 2001). This analysis focused on a new sulfur-reducing enzyme complex from *P. furiosus*. The experiment showed at least a two-fold change in signal intensity for about 50 ORFs that were represented on the array. Subsequently, this initial study was followed by the analyses of a complete genome array (Schut, Zhou et al. 2001; Weinberg, Schut et al. 2005) using the same strategy. For most genes the complete ORFs were printed on the array as PCR-amplified fragments. These studies addressed the adaptation of *P. furiosus* cells to the availability of sulfur, different carbon sources, and cold shock.

Heat shock response

Although hyperthermophiles have a temperature optimum above 80°C, they still can experience heat stress. As in other severe stress conditions, a heat shock will result in retardation or even complete arrest of growth of the organism. This is a consequence of dropping rates of transcription (van de Peppel, Kemmeren et al. 2003); under such conditions protein synthesis appears to be limited to a sub-set of proteins that play a crucial role in dealing with the stress factor to allow survival. When a heat shock is experienced by the cell, two of the biggest threats are the denaturation of proteins and the increased fluidity of the membrane. In order to cope with these problems, hyperthermophilic archaea have developed their own strategies to cope with such conditions. The hyperthermophilic heat shock responses of two distinct hyperthermophilic archaea, *P. furiosus* (Shockley, Ward et al. 2003) and *S. solfataricus* (Tachdjian and Kelly 2006) (Figure 1), were

investigated using transcriptomics. Both organisms seem to react to the same kind of stress differently.

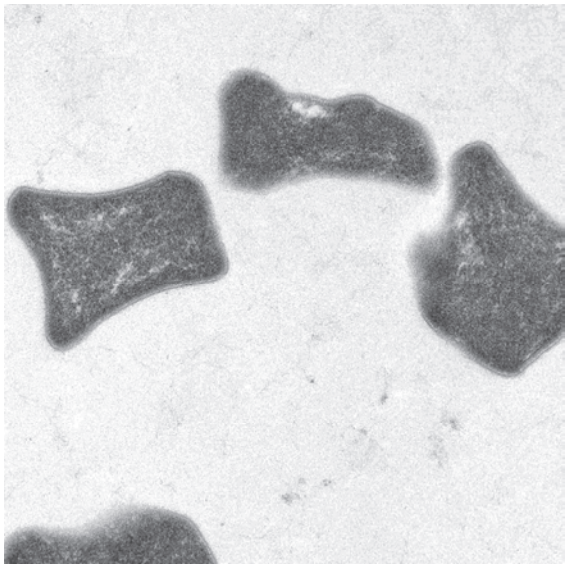

Figure 1: *Sulfolobus solfataricus* cells. (Courtesy of Mark Young)

The heat shock experiment using *P. furiosus* was conducted by growing the cells on a mixture of tryptone and yeast extract at a sub-optimal temperature of 90°C and then shifting the temperature to 105°C (Shockley, Ward et al. 2003). Cells were harvested after 60 minutes and compared to cells grown at 90°C. *P. furiosus* seems to react in several ways: (i) the compatible solutes di-*myo*-inositol-1,1'-phosphate (DIP) and trehalose seem to be produced in order to stabilize its proteins (Santos and da Costa 2002); (ii) proteins were further stabilized by the up regulation of several chaperonin-related genes such as the Hsp60-like thermosome, the Hsp20-like small heat shock protein, and two other proteins (VAT) that are predicted to be involved in both protein unfolding (for proteolyses) and refolding processes; (iii) several genes encoding glycoside hydrolases were up-regulated, either as a general stress response or as a directed adaptation to heat stress that may enhance the production of sugar-based compatible solutes.

The heat shock experiment conducted with *S. solfataricus* was set up differently (Tachdjian and Kelly 2006). The cells were grown at an optimal

temperature of 80°C and then shifted to 90°C. Samples were taken 10 minutes before heat shock, 5, 30 and 60 minutes after heat shock allowing for the elucidation of temporal transcriptome changes. This approach showed that about one-third of the genome (~1000 genes) was differentially regulated in the first 5 minutes. Surprisingly, around 200 of the up-regulated genes were IS elements, showing that almost all of these selfish elements of *S. solfataricus* are activated when the cells encounter (temperature) stress; it may well be that the transposition by itself also contributes to part of the modulated expression of other genes. In contrast to the findings with *P. furiosus,* no evidence was found of induced expression of enzymes involved in compatible solute production. It has been observed that genes that encode different subunits of the RNA polymerase are down-regulated, suggesting that transcription is going down. Furthermore, the gene encoding the DNA polymerase II is down, while several DNA repair related genes have a higher expression. The expression of several transporter genes (eg. Iron, Cobalt, Phosphate, Sulfate, Amino Acids, Arabinose, Glucose, Maltose) went down. Interestingly, also many transcriptional regulators were differentially expressed, namely TetR and the GntR-like repressors. Furthermore the gene encoding the γ-subunit of the thermosome was down-regulated, while the genes encoding the α- and β-subunits were unaffected, which was consistent with the previous findings of a change in composition of the thermosome from 1α:1β:1γ to 2α:1β:0γ (Kagawa, Yaoi et al. 2003). In conclusion, this experiment showed that in *S. solfataricus* the transcriptional response to a heat shock is instantaneous, but apparently not at the level of compatible solutes. Apart from a decrease in growth rate, the overall transcription rate is ceases can be observed as well as a reduced transcription level of the genes encoding the DNA polymerase. Many transcriptional regulators appear to play a role in coping with a heat shock in *S. solfataricus,* and it would be very interesting to establish their specific function, i.e. their target promoters. The difficulty in comparing these two studies is mainly caused by the different sampling approach. In case of *S. solfataricus* the shift has been made from the temperature at which the growth is the fastest, in case of *Pyrococcus* there might be additional variation in the results related to the sub-optimal temperature at the beginning of the experiment.

Viral infections and microorganism interactions

In most environments viral particles significantly outnumber microbial cells, indicating that viral infection is a common threat to the majority of organisms. Hyperthermophiles are not an exception to this rule. Here we discuss two viral infection studies of *S. solfataricus*, both of which have been conducted by using DNA microarrays that contained oligonucleotides corresponding to genes of both *S. solfataricus* as well as genes from selected *S. solfataricus* viruses and plasmids. One study described infection by the lytic virus STIV (*Sulfolobus Turreted Icosahedral Virus)* that usually only kills part of the *S. solfataricus* population in its life cycle (Maaty, Ortmann et al. 2006), whereas comparable analyses have been performed on the well studied lysogenic SSV1 virus (*Sulfolobus shibatae* Virus 1) (Palm, Schleper et al. 1991).

The study of STIV conducted by Ortmann *et al.* (Ortmann, Brumfield et al. 2008) comprises of the isolation of a *S. solfataricus* mutant that is hypersensitive to the studied virus with almost all cells of a culture being killed in the lytic cycle. STIV is a dsDNA virus with a circular genome of 17 kb, containing 37 predicted ORFs. Analysis of the viral transcriptome showed the up-regulation of 47 of the 52 viral microarray probes, which cover the viral genes and some intergenic regions in both directions. Transcription of viral genes was first detected at 8 hpi (hours post infection), whereas at 16 hpi most viral genes are expressed. At 24 hpi a shift takes place from virus replication to preparation for lysis and around this time point most viral genes are expressed; general cell lysis occurs at 32 hpi. Although the expression starts at different time points, no real temporal expression has been observed in this experiment; however, one cannot rule out that this is a resolution issue due to sub-optimal synchronization of the infection cycle. At the early stage of viral gene expression (8 hpi) there are four transcripts and an intergenic region that are being expressed. These genes are most probably responsible for initiation of the early infection process. Expression of most structural viral genes is found at 16 hpi and thereafter. Of the 177 host genes that were differentially regulated (more than 2-fold), of which 124 were up regulated, most are associated with either DNA replication and repair or genes of unknown function, suggesting that STIV uses host proteins to aid the replication of its own DNA. An important up-regulated protein concerns the ESCRTIII homolog, which has recently been reported to be essential for the cell division in *Sulfolobales* (Ettema and Bernander 2009; Samson and Bell 2009); the up-regulation may suggest involvement in the recently

discovered release system for both STIV and SirV that involves unique pyramid-like structures (Figure 2) (Bize, Karlsson et al. 2009; Brumfield, Ortmann et al. 2009). All of the down-regulated host genes were regulated just before cell lysis at 32 hpi, and were associated with metabolism.

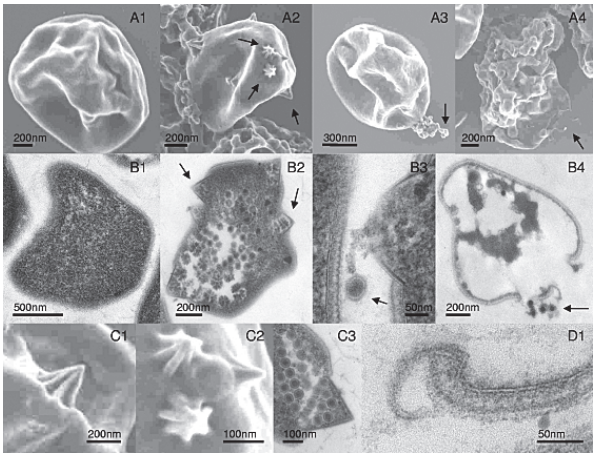

Figure 2: SEM images (row A) and corresponding TEM images (row B) of *S. solfataricus* cells show different stages of infection. (A1 and B1) Noninfected cells. (A2 and B2) Cells infected with STIV displaying membrane protrusions (thin arrows). (A3 and B3) Lysing cells releasing virus (thin arrows) and cell contents. (A4 and B4) Empty cells showing S-layer and broken membrane fragments (thin arrows). Pyramid-like structures from STIV-infected cells observed by SEM (C1 and C2) and TEM (C3) are also shown.(D1) TEM image of broken membrane and S-layer after cell lysis. Scale bars are indicated. (Courtesy of Mark Young)

An infection study of SSV1 with *S. solfataricus* as a host, has been conducted in order to find out more about the transcriptome fluctuations of this lysogenic virus and its host (Frols, Gordon et al. 2007). Initially infection by SSV1 seems not to affect the growth rate of the infected cells; at least partly, the SSV1 genome is integrated at a specific site in the host chromosome (Schleper, Kubo et al. 1992), however, as soon as SSV1 starts to produce and release viral particles, the cell growth is significantly retarded. Viral production can be greatly stimulated after UV induction. The first viral transcripts can already be found at 1 hpi, while most viral genes are active at 8.5 hpi. The viral genes are clustered as 9 operons, comprising both regulatory genes and structural genes. The regulatory genes are the first ones to be transcribed, and the genes coding for the coat protein of the virus are produced at a later stage.

There are more differences between the two studies, and only few similarities. Comparison of the two datasets is not straightforward, mainly because

it compares infection by two distinct types of viruses (lytic vs. lysogenic); in addition there are some methodological differences like the different time points involved, number of time points taken into account, etc. One of the main differences concerns the fact that STIV seems to have a larger impact on the host due to a more profound regulation of host genes (177 instead of 55); this may correlate with its lytic live-cycle. However, to deduce general patterns it will be necessary to compare the transcription profiles during a synchronized infection of additional viruses. A recent study on the infection of the closely related *S. islandicus* with the lytic virus SirV, revealed a dramatic degradation of the host chromosome upon viral assembly and proliferation (Bize, Karlsson et al. 2009); no transcriptome analysis of host genes after infection of this system has yet been reported.

The microarray technique can be used to observe the interactions between two distinct species. One such attempt has been done on a bacteria, *Thermotoga maritima*, which has been grown alone as well as in a co-culture with a archaea, a methanogenic thermophile, *Methanocaldococcus janaschii* (Johnson, Conners et al. 2006). This experiment yielded an interesting view on the importance of the H_2 transfer in hot environment. The experiment focused on a shift from the mid logarithmic growth phase to the early stationary. It has been observed that the growth of *T. maritima* has been boosted 3 to 5-fold due to removal of inhibiting H_2. Also the methane production of *M. jannaschii* has been increased twofold compared with pure culture. The transcriptome analysis of the 2 samples from the early stationary phase showed that in the pure culture of *T. maritima*, 127 genes have been significantly upregulated in comparison with the co-culture. Half of those were associated with the central carbon metabolism. At the same time, in the co-culture of the 113 genes upregulated, the main groups present were ABC transporters and carbohydrate hydrolases. This suggests that the pure culture conditions support the main metabolic pathways while the co-culture conditions seem to boost the scavenging. The scavenging strategy may be boosted by the exopolysaccharide (EPS) produced by the co-culture cells that form aggregates to enhance the hydrogen transfer (Muralidharan, Rinker et al. 1997). Another, less obvious conclusion from the experiment was the confirmation that in this case, a microarray platform designed to analyze one species, can be successfully used to analyze a co-culture condition.

Genome replication and the cell cycle

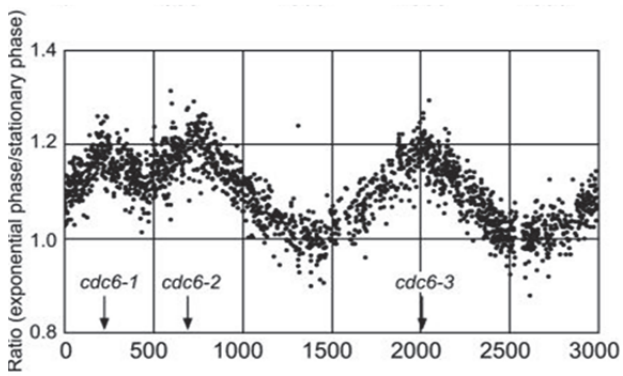

Figure 3: Marker Frequency distributions. Exponential growth vs. stationary phase for S. solfataricus. (Courtesy of Magnus Lundgren) Here DNA from a S. solfataricus cells in exponential phase were compared to DNA from cells in stationary phase. Cells that just have begun growing have more copies of genes at or close to a DNA replication site then DNA further from the replication start site. Therefore genes close to a replication start site will have a higher ratio then genes not close to such a site and this is seen as a peak in both S. acidocaldarius and S. solfataricus were functional (Myllykallio and Forterre 2000; Robinson and Bell 2007).

Up until 2004 it was assumed that genome replication with multiple origins of replication was a typical Eukaryotic-like feature (Myllykallio, Lopez et al. 2000). In 2004, different groups independently discovered that *Sulfolobus* spp. have multiple origins of replication (Contursi, Pisani et al. 2004; Lundgren, Andersson et al. 2004; Robinson, Dionne et al. 2004). Using 2D DNA gels, two origins of replication could be demonstrated in *S. solfataricus*, while a microarray approach (quantification of genomic DNA by hybridizing it with a DNA microarray) was used to prove that *Sulfolobus* spp. has actually three origins of replication (Figure 3). In the latter study *Sulfolobus* cells were treated with acetic acid in order to synchronize the initiation of replication. After removal of the acetic acid inhibition, the cells were harvested at different time points and genomic DNA was extracted and hybridized on a microarray. It was revealed that all three *cdc6*-like *genes* in the figure below. The figure has three clear peaks, showing that S. solfataricus has 3 origins of replication, each peak is located near a predicted cdc6 site.

Although this was a major breakthrough in the field of prokaryotic genome replication, it should be stressed that other archaea (incl. *P. abyssi*) have a single origin of replication (Myllykallio, Lopez et al. 2000). Together with the fact that

none of the known bacterial chromosomes possess multiple origins, this strongly suggests that multiple origins are an archaeal invention, and that the last universal common ancestor (LUCA) most likely possessed a single origin of replication.

The cell cycle of the *Sulfolobus* spp. is relatively well studied and, although some archaeal species show modifications to this model (Maisnier-Patin, Malandrin et al. 2002; Baumann, Lange et al. 2007), it is currently used as archetype of the archaeal cell cycle. An important mechanistic difference, however, concerns the involvement of the ESCRT-III based system in crenarchaea, versus the FtsZ-based, tubulin-directed system in euryarchaea (Makarova, Yutin et al. 2010). *S. solfataricus*, interestingly, possesses both the ESCRT-III encoding genes as well as a gene hypothesized to be a FtsZ paralog (Makarova and Koonin 2003). In 2007, Lundgren and Bernander used a microarray approach to analyze a time series of synchronized cells of *S. acidocaldarius* to show that a cyclic induction of genes that is involved in the cell cycle (Lundgren and Bernander 2007). The cell growth was arrested in the G2 phase by addition of acetic acid (dissipates membrane potential, and inhibits overall metabolic activity at low pH); after resuspending the cells in fresh medium, the synchronized cells started to grow again after 30 minutes. Cells were analyzed at 8 different time points allowing a good overview of global gene expression patterns starting at the G2 phase (0-30 minutes) going all the way through the cycle until the cells are again in the G2 phase (about 200 minutes later). In a parallel study, using a distinct manner of synchronization in which cells are captured at low temperature right after cell division ("the baby machine"), Steve Bell and co-workers presented a cell cycle-dependent transcription of ESCRT-III system components and a Vps4 homolog in *S. acidocaldarius* (Samson, Obita et al. 2008). Interestingly, though not annotated as ESCRT/ Vps4, similar expression profiles of these genes were described in the parallel study mentioned above (Anderson and Dahms). The observed activity of ESCRT-III system in Crenarchaeal cell cycle suggests a common ancestry of cell division mechanisms in archaea and eukarya.

Apart from shedding light on the cell division mechanisms, microarray analysis allowed observing a cyclic expression of different kinases, at least seven transcription factors, as well as the three *cdc6* genes. These findings suggest that the cell cycle is regulated at different levels. Of the three *cdc6* genes, *cd6-1* is the first to be highly expressed, slightly before the G_1/S transition. Shortly after the induction of the first *cdc6* gene, the *cdc6-3* gene is induced, confirming its

secondary role to the *cdc6-1* gene. The gradual induction of the *cdc6-2* gene slightly before the cells approach the G_2 phase suggests a negative regulatory role in chromosome regulation as suggested in earlier studies (Robinson, Dionne et al. 2004). On the other hand, the data from Duggin et al. (Duggin, McCallum et al. 2008) implies that the Cdc6 protein levels during the cell cycle synchronized using the baby machine remain unchanged. The discrepancy between the results is hypothesized to be an effect of two different synchronization methods rather than from the cell cycle itself. Acetate can induce stress in the cells and influence transcription of some stress response related genes. It can also be a result of differential levels of transcript levels and protein, however this possibility is undermined by the fact that other studies showed a correlation between protein and transcript level in case of this gene (Fröls, Gordon et al. 2007; Gotz, Paytubi et al. 2007; Albers, Birkeland et al.).

Pentose metabolism in archaea

Most genomes consist of considerable fractions of hypothetical genes for which a function can not accurately be predicted. These genes are either too distantly related to well established orthologs to be recognized as such; alternatively, they may encode novel types of proteins, either involved in unique processes/bioconversions, or playing a role in a known process but being the result of a non-orthologous gene displacement (Mirkin, Fenner et al. 2003). Microarrays can help elucidating the function of these hypothetical genes, by comparing the transcriptomes in condition where a given process/pathway is expected to be active or not. As such, appropriate transcription profiles could serve as leads for further research.

A good example of a successful microarray-based discovery in archaeal metabolism concerns the elucidation of a pentose-converting pathway in *S. solfataricus*. Unlike many other bacteria and eukaryotes, archaea do not seem to have the classical oxidative pentose phosphate pathway to produce pentose precursors. In addition, until recently the mechanism of the catabolic process of many pentoses in archaea was not understood in great detail (van de Werken, Brouns et al. 2008; Nunn, Johnsen et al. 2010). The analysis of Brouns *et al.* helped to understand how D-arabinose is metabolized by *S. solfataricus*; moreover, insight was gained in the composition of some general pentose oxidation pathways

in both archaea and bacteria (Brouns, Walther et al. 2006). In this study, the microarray technology has been used as an initial step of pathway elucidation and allowed for composing a short list of potential candidate enzymes. Comparison between cells grown on D-arabinose and D-glucose revealed that 16 genes were significantly up-regulated in the first condition. These included the genes encoding the 4 subunits of a previously identified arabinose ABC transporter, a putative sugar permease, and 5 hypothetical enzymes. Comparing the sequences of the intergenic regions revealed the presence of a conserved palindromic motif in promoter regions of 5 of the up-regulated genes: the arabinose ABC transporter operon, and 4 of the hypothetical genes. Production and characterization of the 4 corresponding enzymes has resulted in unraveling the arabinose degrading pathway.

A further *in silico* investigation of the genes resulted in the finding of different but very similar degradation pathways for several C_5 (D- and L-arabinose, D-xylose, hydroxyl-proline) and C_6 (D-glucaric acid, D-galactaric acid) substrates (Brouns, Walther et al. 2006), used by different organisms. Interestingly, all proposed pathways converge at 2,5-dioxopentanoic acid, which is converted to the citric acid cycle intermediate 2-oxoglutaric acid (α-ketoglutarate). This is yet another example of the metabolic tinkering during the evolution of metabolic pathways (Mirkin, Fenner et al. 2003). As biochemical pathways of archaea can be very different from their bacterial/eukaryotic counterparts, DNA microarrays in combination with the currently established gene disruption techniques for *Sulfolobus spp*. (Berkner, Wlodkowski et al. 2010) and *Thermococcus kodakaraensis* (Kanai, Akerboom et al. 2007) may provide a solid basis for subsequent analyses.

Deep sequencing - the high-resolution alternative

The next generation transcriptomics approach is deep sequencing. In deep sequencing protocols, RNA is used to generate complementary DNA (cDNA) that will then be sequenced, generating reads of ~400 nucleotides (454/pyrosequencing (Margulies, Egholm et al. 2005)) and/or reads of ~75 nucleotides (Solexa/SOLiD (Bennett 2004)). A major practical advantage is that this procedure is based on general, species-independent protocols. In addition, it does not need the pre-existing knowledge of the species' genome. Moreover, it allows for comparison of

multiple species in co-culture by simultaneous analysis using the same platform. Because of these features, this technology is frequently used the transcriptomics analysis of environmental samples.

A disadvantage of this approach for analysis of prokaryotic transcriptomes is the overabundance of the rRNA-species, compared to the mRNA-species (only <5% of the total cellular RNA consists of mRNA). This overabundance of non-mRNA species in the sequenced sample results in a high noise factor and also could result in not detecting mRNA that is present in only low amounts. Therefore many protocols rely on the specific removal of rRNA before actual sequencing (Wilhelm and Landry 2009). Most of them are based on techniques that fish out mRNA by using the poly-A tail, which eukarial mRNA posses, but prokaryotes do not. Despite these practical challenges, Sorek and co-workers have successfully analyzed the transcriptome of *S. solfataricus* by deep sequencing, without the removal of the rRNA (Wurtzel, Sapra et al. 2010). They have grown the organism on glucose, cellobiose and cellulose and sequenced the cDNA using the Illumina Genome Analyzer (Solexa). Of the originally proposed set of 3300 genes (She, Singh et al. 2001), the deep-sequencing study managed to correct the annotation of 162 genes, define 80 new ORFs, predict 80 non-coding RNA's, predict a possible hyper-sensitive RNA cleavage site and determine the operon structures of more than 1000 transcriptional units. Moreover, they have found that at least 80 of the *S. solfataricus* operons have overlapping antisense transcripts, a relatively high number (8%) in prokaryotes. These *cis*-encoding transcripts most likely play a role in control of gene expression either at transcriptional or translational level (Waters and Storz 2009).

Standardized procedures

High throughput functional genomics approaches are frequently combined in systems biology approaches aiming at modeling the physiology of microbial cells. A very good example of such a systems approach in mesophilic archaea is a study by Bonneau et al. (Bonneau, Facciotti et al. 2007), in which transcriptome analysis was part of an integrated analysis aiming at the reconstruction of a gene networks in the halophilic archaeon Halobacterium sp. By using different transcription regulators, genetic modification and high throughput methods, a model has been generated that describes the behavior of this network in a range of conditions. Such

a systems approach combined with modeling allows picturing the interactions of an organism and predicting its behavior in the natural environment. The difficulty of such an approach lies in synchronizing a large research project and having a uniform biomaterial to start with.

An example of such a systems biology approach in thermophilic archaea concerns the SulfoSYS project (Albers, Birkeland et al. 2009), which is part of the European SysMO consortium. A major goal of the latter consortium is to establish well-integrated systems biology projects on selected model organisms. A major goal of the SYSMO projects is to perform a multi-disciplinary, functional genomics approach that should be highly reproducible because of the implementation of well described, standard protocols. In the SulfoSYS project the model organism S. solfataricus is cultivated in a very controlled way. The obtained cells are then distributed among the different researchers to perform transcriptomics, proteomics, metabolomics as well as biochemical analyses; eventually the data are included in an integrated metabolic model. The stringency of cultivation and sampling has been important also due to a comparison of cells from different temperature values. As the half-lives of some mRNA particles can be as low as 2 minutes (Bize, Karlsson et al.), a slight difference in sampling may lead to a large difference in the transcript level. The impact of the careful preparation of biological samples in functional genomics analyses, incl. DNA microarray experiments, has not always been appreciated; on the other hand it is generally accepted that this may significantly affect the reproducibility of this approach. The SulfoSYS project puts much weight on careful sample preparation, and on verifying the quality of the obtained cell material before performing actual experiments (Zaparty, Esser et al. 2010); this has resulted in a combined dataset with microarray and deep sequencing data that are in very good agreement (Sierocinski et al., unpublished). The SysMO consortium puts extra weight on giving an unrestricted and easy access to the generated data (Booth 2007). As far as the datasets of respective microarrays are usually freely available, the multitude of standards, methods and platforms severely impedes the possibilities of comparing two data sets with each other. Applying the deposition standards, as Minimum Information About a Microarray Experiment (MIAME) (Brazma, Hingamp et al. 2001), certainly helps to validate the quality of the data; however, a simplified standard for results storage could be proposed to allow quick and efficient analysis of deposited datasets.

Conclusions & Outlook

DNA microarrays have been very successful during the last decade, as a high-throughput research tool that has led to important scientific discoveries, including important findings on cell biological/metabolic features of hyper-thermophilic archaea, as outlined above. The most frequently used DNA micro-arrays (based on oligonucleotides) have restrictions because the probe design is based on previously made assumptions with respect to predicted genes; this implies that small ORFs and non-coding RNAs are generally not included on microarrays. In addition, the commonly used technology only allows for relatively limited numbers of spots can be printed on one slide. The problem of an incomplete set of probes is solved by using tiled DNA microarrays, which are composed of overlapping oligonucleotides. The used probe lengths and the degree of tiling between overlapping probes determine the resolution that can be achieved; typically 2-4 x 10^5 probes are printed per slide, with probe size ranging between 50-75 nucleotides. Tiled arrays cover the two complete strands of the target chromosomes (Mockler, Chan et al. 2005).

New ways of obtaining global transcriptomic data are being investigated. Sequencing cDNA (RNA-seq), although still a developing technique, seems to be very promising (Gilbert, Field et al. 2008). This approach is easier to implement for eukaryotic systems, due to the polyA-based procedure for separating mRNA from the contaminating rRNA. However, despite this practical complication, this technology will also be an important step forward in the transcriptome analysis in prokaryotic systems. In eukaryotes ORF prediction is not as easy as in prokaryotes and this has often led to the development of cDNA libraries for the production of microarrays. RNA-seq, although frequently used in eukaryotic transcriptomics, might become of more importance in future transcriptome studies of bacteria and archaea. Recently some groups have started to gain insight into the expression levels of the complete transcriptome using high-throughput sequencing techniques like 454 deep sequencing (Wilhelm and Landry 2009). Reads of 400 bps can be obtained, at a cost which almost equals the cost for microarray hybridization, with a 97% certainty of prediction the messenger RNA species (Torres, Metta et al. 2008; Wang, Gerstein et al. 2009). This sequencing approach has the advantage that the same platform can be used for different species, resulting in a better interspecies comparison by omitting the cross-platform bias. This opens up the door for environmental transcriptome profiles, allowing for the monitoring of

metagenome-based gene expression in the environment, as opposed to the artificial conditions that are generally imposed on them in a laboratory setting. A further advantage might be that RNA-seq is less prone to signal loss due to mutations that arise during cultivation. Although this technique is not yet readily accessible for most labs, the anticipated reduction of sequencing costs in the near future might make this a very attractive general technique for transcriptome analysis for both eukaryotes and prokaryotes. A decrease in the use of the DNA microarray as a research tool, and an increase of using sequencing-related techniques in this field may be expected (Ledford 2008).

RNA-seq might turn out to be quintessential in examining environmental samples were not all of the components have been known beforehand. For instance they might greatly help to increase our understanding of phage pressure on the potential hosts that takes place in situ by finding more viral transcripts and watching the response of the thermophiles to multiple viruses present in the environment. One can assume that hyperthermophilic environments are a very good target for early attempts of meta-transcriptomic analyses as the ecology of such niches is generally less complex than that of aquatic or soil ecosystems, making it easier to deal with big dataset covering many organisms.

References

1. Brock, T.D. and H. Freeze, Thermus aquaticus gen. n. and sp. n., a nonsporulating extreme thermophile. J Bacteriol, 1969. **98**(1): p. 289-97.
2. Stetter, K.O., Hyperthermophiles in the history of life. Philos Trans R Soc Lond B Biol Sci, 2006. **361**(1474): p. 1837-42; discussion 1842-3.
3. Angelov, A. and W. Liebl, Insights into extreme thermoacidophily based on genome analysis of Picrophilus torridus and other thermoacidophilic archaea. J Biotechnol, 2006. **126**(1): p. 3-10.
4. Segerer, A.H., et al., Life in hot springs and hydrothermal vents. Orig Life Evol Biosph, 1993. **23**(1): p. 77-90.
5. Ettema, T.J., W.M. de Vos, and J. van der Oost, Discovering novel biology by in silico archaeology. Nat Rev Microbiol, 2005. **3**(11): p. 859-69.
6. Makarova, K.S. and E.V. Koonin, Comparative genomics of Archaea: how much have we learned in six years, and what's next? Genome Biol, 2003. **4**(8): p. 115.
7. Brouns, S.J., et al., Identification of the missing links in prokaryotic pentose oxidation pathways: evidence for enzyme recruitment. J Biol Chem, 2006. **281**(37): p. 27378-88.
8. Cambillau, C. and J.M. Claverie, Structural and genomic correlates of hyperthermostability. J Biol Chem, 2000. **275**(42): p. 32383-6.
9. Koutsopoulos, S., J. van der Oost, and W. Norde, Kinetically controlled refolding of a heat-denatured hyperthermostable protein. FEBS J, 2007. **274**(22): p. 5915-23.
10. Kumar, S. and R. Nussinov, How do thermophilic proteins deal with heat? Cell Mol Life Sci, 2001. **58**(9): p. 1216-33.

11. Rivera, M.C. and J.A. Lake, The ring of life provides evidence for a genome fusion origin of eukaryotes. Nature, 2004. **431**(7005): p. 152-5.
12. Bult, C.J., et al., Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii. Science, 1996. **273**(5278): p. 1058-73.
13. Nelson, K.E., et al., Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of Thermotoga maritima. Nature, 1999. **399**(6734): p. 323-9.
14. Koonin, E.V., Y.I. Wolf, and L. Aravind, Prediction of the archaeal exosome and its connections with the proteasome and the translation and transcription machineries by a comparative-genomic approach. Genome Res, 2001. **11**(2): p. 240-52.
15. Ettema, T., J. van der Oost, and M. Huynen, Modularity in the gain and loss of genes: applications for function prediction. Trends Genet, 2001. **17**(9): p. 485-7.
16. Frickey, T. and A.N. Lupas, PhyloGenie: automated phylome generation and analysis. Nucleic Acids Res, 2004. **32**(17): p. 5231-8.
17. Ruepp, A., et al., The genome sequence of the thermoacidophilic scavenger Thermoplasma acidophilum. Nature, 2000. **407**(6803): p. 508-13.
18. Doerks, T., C. von Mering, and P. Bork, Functional clues for hypothetical proteins based on genomic context analysis in prokaryotes. Nucleic Acids Res, 2004. **32**(21): p. 6321-6.
19. Ramsay, G., DNA chips: state-of-the art. Nat Biotechnol, 1998. **16**(1): p. 40-4.
20. Lundgren, M. and R. Bernander, Genome-wide transcription map of an archaeal cell cycle. Proc Natl Acad Sci U S A, 2007. **104**(8): p. 2939-44.
21. Auernik, K.S. and R.M. Kelly, Physiological versatility of the extremely thermoacidophilic archaeon Metallosphaera sedula supported by transcriptomic analysis of heterotrophic, autotrophic, and mixotrophic growth. Appl Environ Microbiol, 2010. **76**(3): p. 931-5.
22. Trauger, S.A., et al., Correlating the transcriptome, proteome, and metabolome in the environmental adaptation of a hyperthermophile. J Proteome Res, 2008. **7**(3): p. 1027-35.
23. Albers, S.V., et al., SulfoSYS (Sulfolobus Systems Biology): towards a silicon cell model for the central carbohydrate metabolism of the archaeon Sulfolobus solfataricus under temperature variation. Biochem Soc Trans, 2009. **37**(Pt 1): p. 58-64.
24. Schut, G.J., J. Zhou, and M.W. Adams, DNA microarray analysis of the hyperthermophilic archaeon Pyrococcus furiosus: evidence for anNew type of sulfur-reducing enzyme complex. J Bacteriol, 2001. **183**(24): p. 7027-36.
25. Weinberg, M.V., et al., Cold shock of a hyperthermophilic archaeon: Pyrococcus furiosus exhibits multiple responses to a suboptimal growth temperature with a key role for membrane-bound glycoproteins. J Bacteriol, 2005. **187**(1): p. 336-48.
26. van de Peppel, J., et al., Monitoring global messenger RNA changes in externally controlled microarray experiments. EMBO Rep, 2003. **4**(4): p. 387-93.
27. Shockley, K.R., et al., Heat shock response by the hyperthermophilic archaeon Pyrococcus furiosus. Appl Environ Microbiol, 2003. **69**(4): p. 2365-71.
28. Tachdjian, S. and R.M. Kelly, Dynamic metabolic adjustments and genome plasticity are implicated in the heat shock response of the extremely thermoacidophilic archaeon Sulfolobus solfataricus. J Bacteriol, 2006. **188**(12): p. 4553-9.
29. Santos, H. and M.S. da Costa, Compatible solutes of organisms that live in hot saline environments. Environ Microbiol, 2002. **4**(9): p. 501-9.
30. Kagawa, H.K., et al., The composition, structure and stability of a group II chaperonin are temperature regulated in a hyperthermophilic archaeon. Mol Microbiol, 2003. **48**(1): p. 143-56.
31. Maaty, W.S., et al., Characterization of the archaeal thermophile Sulfolobus turreted icosahedral virus validates an evolutionary link among double-stranded DNA viruses from all domains of life. J Virol, 2006. **80**(15): p. 7625-35.
32. Palm, P., et al., Complete nucleotide sequence of the virus SSV1 of the archaebacterium Sulfolobus shibatae. Virology, 1991. **185**(1): p. 242-50.

33. Ortmann, A.C., et al., Transcriptome analysis of infection of the archaeon Sulfolobus solfataricus with Sulfolobus turreted icosahedral virus. J Virol, 2008. **82**(10): p. 4874-83.
34. Ettema, T.J. and R. Bernander, Cell division and the ESCRT complex: A surprise from the archaea. Commun Integr Biol, 2009. **2**(2): p. 86-8.
35. Samson, R.Y. and S.D. Bell, Ancient ESCRTs and the evolution of binary fission. Trends Microbiol, 2009. **17**(11): p. 507-13.
36. Brumfield, S.K., et al., Particle assembly and ultrastructural features associated with replication of the lytic archaeal virus sulfolobus turreted icosahedral virus. J Virol, 2009. **83**(12): p. 5964-70.
37. Bize, A., et al., A unique virus release mechanism in the Archaea. Proc Natl Acad Sci U S A, 2009. **106**(27): p. 11306-11.
38. Frols, S., et al., Response of the hyperthermophilic archaeon Sulfolobus solfataricus to UV damage. J Bacteriol, 2007. **189**(23): p. 8708-18.
39. Schleper, C., K. Kubo, and W. Zillig, The particle SSV1 from the extremely thermophilic archaeon Sulfolobus is a virus: demonstration of infectivity and of transfection with viral DNA. Proc Natl Acad Sci U S A, 1992. **89**(16): p. 7645-9.
40. Johnson, M.R., et al., The Thermotoga maritima phenotype is impacted by syntrophic interaction with Methanococcus jannaschii in hyperthermophilic coculture. Appl Environ Microbiol, 2006. **72**(1): p. 811-8.
41. Muralidharan, V., et al., Hydrogen transfer between methanogens and fermentative heterotrophs in hyperthermophilic cocultures. Biotechnol Bioeng, 1997. **56**(3): p. 268-78.
42. Myllykallio, H., et al., Bacterial mode of replication with eukaryotic-like machinery in a hyperthermophilic archaeon. Science, 2000. **288**(5474): p. 2212-5.
43. Lundgren, M., et al., Three replication origins in Sulfolobus species: synchronous initiation of chromosome replication and asynchronous termination. Proc Natl Acad Sci U S A, 2004. **101**(18): p. 7046-51.
44. Robinson, N.P., et al., Identification of two origins of replication in the single chromosome of the archaeon Sulfolobus solfataricus. Cell, 2004. **116**(1): p. 25-38.
45. Contursi, P., et al., Identification and autonomous replication capability of a chromosomal replication origin from the archaeon Sulfolobus solfataricus. Extremophiles, 2004. **8**(5): p. 385-91.
46. Myllykallio, H. and P. Forterre, Mapping of a chromosome replication origin in an archaeon: response. Trends Microbiol, 2000. **8**(12): p. 537-9.
47. Robinson, N.P. and S.D. Bell, Extrachromosomal element capture and the evolution of multiple replication origins in archaeal chromosomes. Proc Natl Acad Sci U S A, 2007. **104**(14): p. 5806-11.
48. Maisnier-Patin, S., et al., Chromosome replication patterns in the hyperthermophilic euryarchaea Archaeoglobus fulgidus and Methanocaldococcus (Methanococcus) jannaschii. Mol Microbiol, 2002. **45**(5): p. 1443-50.
49. Baumann, A., C. Lange, and J. Soppa, Transcriptome changes and cAMP oscillations in an archaeal cell cycle. BMC Cell Biol, 2007. **8**: p. 21.
50. Makarova, K.S., et al., Evolution of diverse cell division and vesicle formation systems in Archaea. Nat Rev Microbiol, 2010. **8**(10): p. 731-41.
51. Samson, R.Y., et al., A role for the ESCRT system in cell division in archaea. Science, 2008. **322**(5908): p. 1710-3.
52. Anderson, R.L. and A.S. Dahms, 2-Keto-3-deoxy-l-arabonate aldolase. Methods Enzymol, 1975. **42**: p. 269-72.
53. Duggin, I.G., S.A. McCallum, and S.D. Bell, Chromosome replication dynamics in the archaeon Sulfolobus acidocaldarius. Proc Natl Acad Sci U S A, 2008. **105**(43): p. 16737-42.
54. Gotz, D., et al., Responses of hyperthermophilic crenarchaea to UV irradiation. Genome Biol, 2007. **8**(10): p. R220.
55. Frols, S., et al., Response of the hyperthermophilic archaeon Sulfolobus solfataricus to UV damage. Journal of Bacteriology, 2007. **189**(23): p. 8708-18 %U http://www.ncbi.nlm.nih.gov/pubmed/17905990.
56. Mirkin, B.G., et al., Algorithms for computing parsimonious evolutionary scenarios for genome evolution, the last universal common ancestor and dominance of horizontal gene transfer in the evolution of prokaryotes. BMC Evol Biol, 2003. **3**: p. 2.

57. Nunn, C.E., et al., Metabolism of pentose sugars in the hyperthermophilic archaea Sulfolobus solfataricus and Sulfolobus acidocaldarius. J Biol Chem, 2010. **285**(44): p. 33701-9.
58. van de Werken, H., S. Brouns, and J. van der Oost, Pentose Metabolism in Archaea. Archaea: new models for prokaryotic biology, 2008: p. 71.
59. Berkner, S., et al., Inducible and constitutive promoters for genetic systems in Sulfolobus acidocaldarius. Extremophiles, 2010. **14**(3): p. 249-59.
60. Kanai, T., et al., A global transcriptional regulator in Thermococcus kodakaraensis controls the expression levels of both glycolytic and gluconeogenic enzyme-encoding genes. J Biol Chem, 2007. **282**(46): p. 33659-70.
61. Margulies, M., et al., Genome sequencing in microfabricated high-density picolitre reactors. Nature, 2005. **437**(7057): p. 376-80.
62. Bennett, S., Solexa Ltd. Pharmacogenomics, 2004. **5**(4): p. 433-8.
63. Wilhelm, B.T. and J.R. Landry, RNA-Seq-quantitative measurement of expression through massively parallel RNA-sequencing. Methods, 2009. **48**(3): p. 249-57.
64. Wurtzel, O., et al., A single-base resolution map of an archaeal transcriptome. Genome Res, 2010. **20**(1): p. 133-41.
65. She, Q., et al., The complete genome of the crenarchaeon Sulfolobus solfataricus P2. Proc Natl Acad Sci U S A, 2001. **98**(14): p. 7835-40.
66. Waters, L.S. and G. Storz, Regulatory RNAs in bacteria. Cell, 2009. **136**(4): p. 615-28.
67. Bonneau, R., et al., A predictive model for transcriptional control of physiology in a free living cell. Cell, 2007. **131**(7): p. 1354-65.
68. Zaparty, M., et al., "Hot standards" for the thermoacidophilic archaeon Sulfolobus solfataricus. Extremophiles, 2010. **14**(1): p. 119-42.
69. Booth, I.R., SysMO: back to the future. Nat Rev Microbiol, 2007. **5**(8): p. 566.
70. Brazma, A., et al., Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. Nat Genet, 2001. **29**(4): p. 365-71.
71. Mockler, T.C., et al., Applications of DNA tiling arrays for whole-genome analysis. Genomics, 2005. **85**(1): p. 1-15.
72. Gilbert, J.A., et al., Detection of large numbers of novel sequences in the metatranscriptomes of complex marine microbial communities. PLoS One, 2008. **3**(8): p. e3042.
73. Torres, T.T., et al., Gene expression profiling by massively parallel sequencing. Genome Res, 2008. **18**(1): p. 172-7.
74. Wang, Z., M. Gerstein, and M. Snyder, RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet, 2009. **10**(1): p. 57-63.
75. Ledford, H., The death of microarrays? Nature, 2008. **455**(7215): p. 847.
76. Baliga, N.S., et al., Coordinate regulation of energy transduction modules in Halobacterium sp. analyzed by a global systems approach. Proc Natl Acad Sci U S A, 2002. **99**(23): p. 14913-149188.
77. Zaigler, A., S.C. Schuster, and J. Soppa, Construction and usage of a onefold-coverage shotgun DNA microarray to characterize the metabolism of the archaeon Haloferax volcanii. Mol Microbiol, 2003. **48**(4): p. 1089-105.
78. Schut, G.J., et al., Whole-genome DNA microarray analysis of a hyperthermophile and an archaeon: Pyrococcus furiosus grown on carbohydrates or peptides. J Bacteriol, 2003. **185**(13): p. 3935-47.
79. Muller, J.A. and S. DasSarma, Genomic analysis of anaerobic respiration in the archaeon Halobacterium sp. strain NRC-1: dimethyl sulfoxide and trimethylamine N-oxide as terminal electron acceptors. J Bacteriol, 2005. **187**(5): p. 1659-67.
80. Hovey, R., et al., DNA microarray analysis of Methanosarcina mazei Go1 reveals adaptation to different methanogenic substrates. Mol Genet Genomics, 2005. **273**(3): p. 225-239.
81. Snijders, A.P., et al., Reconstruction of central carbon metabolism in Sulfolobus solfataricus using a two-dimensional gel electrophoresis map, stable isotope labelling and DNA microarray analysis. Proteomics, 2006. **6**(5): p. 1518-29.

82. Culley, D.E., et al., Optimization of RNA isolation from the archaebacterium Methanosarcina barkeri and validation for oligonucleotide microarray analysis. J Microbiol Methods, 2006. **67**(1): p. 36-43.
83. Veit, K., et al., Global transcriptional analysis of Methanosarcina mazei strain Go1 under different nitrogen availabilities. Mol Genet Genomics, 2006. **276**(1): p. 41-55.
84. Lee, H.S., et al., Transcriptional and biochemical analysis of starch metabolism in the hyperthermophilic archaeon Pyrococcus furiosus. J Bacteriol, 2006. **188**(6): p. 2115-2125.
85. Schut, G.J., S.L. Bridger, and M.W. Adams, Insights into the metabolism of elemental sulfur by the hyperthermophilic archaeon Pyrococcus furiosus: characterization of a coenzyme A- dependent NAD(P)H sulfur oxidoreductase. J Bacteriol, 2007. **189**(12): p. 4431-41.
86. Twellmeyer, J., et al., Microarray analysis in the archaeon Halobacterium salinarum strain R1. PLoS One, 2007. **2**(10): p. e1064.
87. Li, L., et al., Quantitative proteomic and microarray analysis of the archaeon Methanosarcina acetivorans grown with acetate versus methanol. J Proteome Res, 2007. **6**(2): p. 759-771.
88. Rich, J.J., et al., Anaerobic ammonium oxidation (anammox) in Chesapeake Bay sediments. Microb Ecol, 2008. **55**(2): p. 311-20.
89. Auernik, K.S. and R.M. Kelly, Identification of components of electron transport chains in the extremely thermoacidophilic crenarchaeon Metallosphaera sedula through iron and sulfur compound oxidation transcriptomes. Appl Environ Microbiol, 2008. **74**(24): p. 7723-32.
90. Ferry, J.G. and D.J. Lessner, Methanogenesis in marine sediments. Ann N Y Acad Sci, 2008. **1125**: p. 147-57.
91. Cozen, A.E., et al., Transcriptional map of respiratory versatility in the hyperthermophilic crenarchaeon Pyrobaculum aerophilum. J Bacteriol, 2009. **191**(3): p. 782-94.
92. Zaparty, M., et al., The central carbohydrate metabolism of the hyperthermophilic crenarchaeote Thermoproteus tenax: pathways and insights into their regulation. Arch Microbiol, 2008. **190**(3): p. 231-45.
93. Wende, A., K. Furtwangler, and D. Oesterhelt, Phosphate-dependent behavior of the archaeon Halobacterium salinarum strain R1. J Bacteriol, 2009. **191**(12): p. 3852-60.
94. Schmid, A.K., et al., A single transcription factor regulates evolutionarily diverse but functionally linked metabolic pathways in response to nutrient availability. Mol Syst Biol, 2009. **5**: p. 282.
95. Johnsen, U., et al., D-xylose degradation pathway in the halophilic archaeon Haloferax volcanii. J Biol Chem, 2009. **284**(40): p. 27290-303.
96. Jager, D., et al., Deep sequencing analysis of the Methanosarcina mazei Go1 transcriptome in response to nitrogen availability. Proc Natl Acad Sci U S A, 2009. **106**(51): p. 21878-82.
97. Auernik, K.S. and R.M. Kelly, Impact of molecular hydrogen on chalcopyrite bioleaching by the extremely thermoacidophilic archaeon Metallosphaera sedula. Appl Environ Microbiol, 2010. **76**(8): p. 2668-72.
98. McCready, S., et al., UV irradiation induces homologous recombination genes in the model archaeon, Halobacterium sp. NRC-1. Saline Systems, 2005. **1**: p. 3.
99. Boonyaratanakornkit, B.B., et al., Transcriptional profiling of the hyperthermophilic methanarchaeon Methanococcus jannaschii in response to lethal heat and non-lethal cold shock. Environ Microbiol, 2005. **7**(6): p. 789-797.
100. Zhang, W., et al., DNA microarray analysis of anaerobic Methanosarcina barkeri reveals responses to heat shock and air exposure. J Ind Microbiol Biotechnol, 2006. **33**(9): p. 784-90.
101. Boonyaratanakornkit, B., et al., Pressure affects transcription profiles of Methanocaldococcus jannaschii despite the absence of barophilic growth under gas-transfer limitation. Environ Microbiol, 2006. **8**(11): p. 2031-2035.
102. Williams, E., et al., Microarray analysis of the hyperthermophilic archaeon Pyrococcus furiosus exposed to gamma irradiation. Extremophiles, 2007. **11**(1): p. 19-29.
103. Pfluger, K., et al., Identification of genes involved in salt adaptation in the archaeon Methanosarcina mazei Go1 using genome-wide gene expression profiling. FEMS Microbiol Lett, 2007. **277**(1): p. 79-89.
104. Hendrickson, E.L., et al., Functionally distinct genes regulated by hydrogen limitation and growth rate in methanogenic Archaea. Proc Natl Acad Sci U S A, 2007. **104**(21): p. 8930-4.

105. Coker, J.A., et al., Transcriptional profiling of the model Archaeon Halobacterium sp. NRC-1: responses to changes in salinity and temperature. Saline Systems, 2007. **3**: p. 6.
106. Dorazi, R., et al., Equal rates of repair of DNA photoproducts in transcribed and non-transcribed strands in Sulfolobus solfataricus. Mol Microbiol, 2007. **63**(2): p. 521-9.
107. Simon, G., et al., Effect of O2 concentrations on Sulfolobus solfataricus P2. FEMS Microbiol Lett, 2009. **299**(2): p. 255-60.
108. Kirkpatrick, R.D., Something old, something new, something borrowed...something blue. Tenn Med, 2010. **103**(8): p. 7-8.
109. Campanaro, S., et al., Temperature-dependent global gene expression in the Antarctic archaeon Methanococcoides burtonii. Environ Microbiol, 2010.
110. Kanai, T., et al., Identification of the Phr-dependent heat shock regulon in the hyperthermophilic archaeon, Thermococcus kodakaraensis. J Biochem, 2010. **147**(3): p. 361-70.
111. Keese, A.M., et al., Genome-wide identification of targets for the archaeal heat shock regulator phr by cell-free transcription of genomic DNA. J Bacteriol, 2010. **192**(5): p. 1292-8.
112. Cooper, C.R., et al., Role of vapBC toxin-antitoxin loci in the thermal stress response of Sulfolobus solfataricus. Biochem Soc Trans, 2009. **37**(Pt 1): p. 123-6.
113. Strand, K.R., et al., Oxidative stress protection and the repair response to hydrogen peroxide in the hyperthermophilic archaeon Pyrococcus furiosus and in related species. Arch Microbiol, 2010. **192**(6): p. 447-59.
114. Shih, C.J. and M.C. Lai, Differentially expressed genes after hyper- and hypo-salt stress in the halophilic archaeon Methanohalophilus portucalensis. Can J Microbiol, 2010. **56**(4): p. 295-307.
115. Matsunaga, F., et al., Genomewide and biochemical analyses of DNA-binding activity of Cdc6/Orc1 and Mcm proteins in Pyrococcus sp. Nucleic Acids Res, 2007. **35**(10): p. 3214-3222.
116. Stralis-Pavese, N., et al., Optimization of diagnostic microarray for application in analysing landfill methanotroph communities under different plant covers. Environ Microbiol, 2004. **6**(4): p. 347-63.
117. Hamilton-Brehm, S.D., G.J. Schut, and M.W. Adams, Metabolic and evolutionary relationships among Pyrococcus Species: genetic exchange within a hydrothermal vent environment. J Bacteriol, 2005. **187**(21): p. 7492-7499.
118. Andersson, A.F., et al., Global analysis of mRNA stability in the archaeon Sulfolobus. Genome Biol, 2006. **7**(10): p. R99.
119. Xia, Q., et al., Quantitative proteomics of the archaeon Methanococcus maripaludis validated by microarray analysis and real time PCR. Mol Cell Proteomics, 2006. **5**(5): p. 868-81.
120. Lange, C., et al., Genome-wide analysis of growth phase-dependent translational and transcriptional regulation in halophilic archaea. BMC Genomics, 2007. **8**: p. 415.
121. Kanai, T., et al., A global transcriptional regulator in Thermococcus kodakaraensis controls the expression levels of both glycolytic and gluconeogenic enzyme-encoding genes. J Biol Chem, 2007. **282**(46): p. 33659-33670.
122. Santangelo, T.J., et al., Polarity in archaeal operon transcription in Thermococcus kodakaraensis. J Bacteriol, 2008. **190**(6): p. 2244-8.
123. Dambeck, M. and J. Soppa, Characterization of a Haloferax volcanii member of the enolase superfamily: deletion mutant construction, expression analysis, and transcriptome comparison. Arch Microbiol, 2008. **190**(3): p. 341-53.
124. Garrido, P., et al., An oligonucleotide prokaryotic acidophile microarray: its validation and its use to monitor seasonal variations in extreme acidic environments with total environmental RNA. Environ Microbiol, 2008. **10**(4): p. 836-50.
125. Grogan, D.W., M.A. Ozarzak, and R. Bernander, Variation in gene content among geographically diverse Sulfolobus isolates. Environ Microbiol, 2008. **10**(1): p. 137-146.
126. Andersson, A.F., et al., Replication-biased genome organisation in the crenarchaeon Sulfolobus. BMC Genomics, 2010. **11**: p. 454.
127. Yergeau, E., et al., Environmental microarray analyses of Antarctic soil microbial communities. ISME J, 2009. **3**(3): p. 340-51.

128. Reichlen, M.J., K.S. Murakami, and J.G. Ferry, Functional analysis of the three TATA binding protein homologs in Methanosarcina acetivorans. J Bacteriol, 2010. **192**(6): p. 1511-7.
129. Schwaiger, R., et al., Transcriptional control by two leucine-responsive regulatory proteins in Halobacterium salinarum R1. BMC Mol Biol, 2010. **11**: p. 40.
130. Facciotti, M.T., et al., Large scale physiological readjustment during growth enables rapid, comprehensive and inexpensive systems analysis. BMC Syst Biol, 2010. **4**: p. 64.
131. Goberna, M., et al., Adaptation of methanogenic communities to the cofermentation of cattle excreta and olive mill wastes at 37 degrees C and 55 degrees C. Appl Environ Microbiol, 2010. **76**(19): p. 6564-71.
132. Parnell, J.J., et al., Functional biogeography as evidence of gene transfer in hypersaline microbial communities. PLoS One, 2010. **5**(9): p. e12919.

Reconstruction of central carbon metabolism in *Sulfolobus solfataricus* using a two-dimensional gel electrophoresis map, stable isotope labelling and DNA microarray analysis

Ambrosius P.L. Snijders*, Jasper Walther*, Stefan Peter, Iris Kinnman, Marjon G.J. de Vos, Harmen J.G. van de Werken, Stan J.J. Brouns, John van der Oost and Phillip C. Wright
* authors contributed equally

Proteomics. 2006 Mar; 6 (5): 1518-29

Abstract

In the last decade, an increasing number of sequenced archaeal genomes have become available, opening up the possibility for functional genomic analyses. Here, we reconstructed the central carbon metabolism in the hyperthermophilic crenarchaeon *Sulfolobus solfataricus* (glycolysis, gluconeogenesis and tricarboxylic acid cycle) on the basis of genomic, proteomic, transcriptomic and biochemical data. A 2-DE reference map of *S. solfataricus* grown on glucose, consisting of 325 unique ORFs in 255 protein spots, was created to facilitate this study. The map was then used for a differential expression study based on ^{15}N metabolic labelling (yeast extract + tryptone grown cells (YT) *vs* glucose grown cells (G)). In addition, the expression ratio of the genes involved in carbon metabolism was studied using DNA microarrays. Surprisingly, only 3 and 14% of the genes and proteins respectively involved in central carbon metabolism showed a greater than two-fold change in expression level. All results are discussed in the light of the current understanding of central carbon metabolism in *S. solfataricus* and will help to obtain a system-wide understanding of this organism.

1 Introduction

Sulfolobus solfataricus is a thermoacidophilic crenarchaeon, which grows between 70 and 90°C and in a pH range of 2-4 (Zillig, Stetter et al. 1980). Its preference for environments hostile to many other organisms makes it an interesting source for novel, thermostable enzymes. *S. solfataricus* has been an attractive crenarchaeal model organism since its isolation in the early 1980s, and the completion of the genomic sequence in 2001 (She, Singh et al. 2001) has further increased its popularity. Currently, 1941 genes (53.11%) in TIGR's comprehensive microbial resource (CMR) database have no known function (Peterson, Umayam et al. 2001). Of the 2977 Open Reading Frames (ORFs) originally identified in the genome of *S. solfataricus*, 40% of the genes are archaea specific, 12% are bacteria specific and 2.3% are shared exclusively with eukaryotes. Currently, genetic tools are under development that will contribute to our understanding of fundamental processes in *Sulfolobus* (Stedman, Schleper et al. 1999; Cannio, Contursi et al. 2001; Contursi, Cannio et al. 2003; Jonuscheit, Martusewitsch et al. 2003; Worthington, Hoang et al. 2003). In order to fully exploit its potential for metabolic engineering, a deeper understanding of the central energy and precursor generating pathways is necessary.

The central metabolic pathways in archaea contain many unique features compared to the classical pathways in bacteria and eukaryotes (Adams, Holden et al. 2001; Verhees, Kengen et al. 2003). In *S. solfataricus*, glucose degradation proceeds via a nonphosphorylated version of the Entner-Doudoroff (ED) pathway (De Rosa, Gambacorta et al. 1984; Schonheit and Schafer 1995; Schafer 1996). In this pathway, glucose is converted into pyruvate through the action of glucose dehydrogenase, gluconate dehydratase, 2-keto-3-deoxy-gluconate (KDG) aldolase, glyceraldehyde dehydrogenase, glycerate kinase, enolase and pyruvate kinase (PK). Recently, experimental evidence has been provided for the operation of the semi-phosphorylated ED pathway in *S. solfataricus* in which KDG is phosphorylated (Ahmed, Ettema et al. 2005). Gluconeogenesis via a reversed ED pathway is unlikely, since the key enzymes in this pathway do not seem to be able to distinguish between glucose and galactose derivatives. In this case, gluconeogenesis via a reversed ED pathway would result in a mixture of glucose and galactose (Lamble, Heyer et al. 2003). Instead, *in silico* analysis of the *Sulfolobus* genomes as well as experimental evidence has revealed the presence of a near complete set of proteins involved in the Embden-Meyerhof-Parnas (EMP) pathway (Verhees,

Kengen et al. 2003), suggested to be active in the gluconeogenic direction rather than in the glycolytic direction (Lamble, Heyer et al. 2003).

In this study, we reconstructed central carbon metabolism and the TriCarboxylic Acid cycle (TCA) cycle on the basis of biochemical, computational, proteomic and DNA microarray data, obtained from cell extracts of *S. solfataricus* grown on sugars and peptides. First of all, a 2-DE map was created to provide a global overview of protein expression under glucose-degrading conditions. This map was then used to investigate the relative abundance of proteins involved in sugar metabolism under minimal or rich media through a ^{15}N metabolic labelling approach. Moreover, DNA microarray analysis was performed to compare mRNA expression under the same conditions. In the last few years, similar transcriptome studies have been conducted with several archaea that utilise different types of glycolysis. These organisms include: *Pyrococcus furiosus* (Schut, Brehm et al. 2003), an obligate anaerobic hyperthermophile with an EMP-like pathway and *Haloferax volcanii* (Zaigler, Schuster et al. 2003) a facultative anaerobic halophile using an ED-like glycolysis. However, there are relatively few studies that combine transcriptomics and proteomics, and none have so far been published for archaea.

Here, we present a study in which both quantitative proteomics and transcriptomics were used to analyse the expression of the genes involved in the central carbon metabolism of *S. solfataricus*.

2 Materials and methods

2.1 Cell growth and harvest

S. solfataricus P2 (DSM1617) was grown aerobically in a rotary shaker at 80 °C in a medium of pH 3.5-4.0 which contained: 2.5 g/L $(NH_4)_2SO_4$, 3.1 g/L KH_2PO_4, 203.3 mg/L $MgCl_2 \cdot 6\ H_2O$, 70.8 mg/L $Ca(NO_3)_2 \cdot 4\ H_2O$, 2 mg/L $FeSO_4 \cdot 7\ H_2O$, 1.8 mg/L $MnCl_2 \cdot 4\ H_2O$, 4.5 mg/L $Na_2B_4O_7 \cdot 2\ H_2O$, 0.22 mg/L $ZnSO_4 \cdot 7\ H_2O$, 0.06 mg/L $CuCl_2 \cdot 2\ H_2O$, 0.03 mg/L $Na_2MoO_4 \cdot 2\ H_2O$, 0.03 mg/L $VOSO_4 \cdot 2\ H_2O$, 0.01 mg/L $CoCl_2 \cdot 6\ H_2O$. The medium was supplemented with Wollin vitamins, and either 0.3% - 0.4% D-glucose (G) or 0.1% Yeast extract and 0.2% Tryptone (YT). The Wollin vitamin stock (100x) contained 2 mg/L D-Biotin, 2 mg/L Folic acid, 10 mg/L Pyridoxine-HCl, 10 mg/L Riboflavin, 5 mg/L Thiamine-HCl, 5 mg/L Nicotinic acid, 5 mg/L DL-Ca-Pantothenate, 0.1 mg/L Vitamin B12, 5

mg/L p-Aminobenzoic acid, 5 mg/L Lipoic acid. Cell growth was monitored by measuring the turbidity at 530 or 600 nm. Cells for the proteome reference map were harvested by centrifugation in the late exponential growth phase at an OD_{530} of 1.0. Cells were washed twice with a 10 mM Tris/HCl Buffer (pH = 7). Subsequently, cells were stored at –20°C until required. During this whole process, considerable care was taken to ensure that culture to culture variation was minimised, and cultures were prepared in at least triplicate. In the case of the ^{15}N labelling experiment, $(^{15}NH_4)_2SO_4$ was used as the nitrogen source. Cells were incubated with ^{15}N ammonium sulphate for at least eight doubling times to allow for full incorporation of the label. After this, the ^{14}N and ^{15}N growth experiments were set up simultaneously. When the optical density reached a value of 0.5, the cultures were mixed. To ensure that equal amounts of biomass were mixed, slight corrections in volume were made in case the OD_{530} was not exactly 0.5. Previously, we have demonstrated that this approach leads to accurate mixing (Snijders, de Vos et al. 2005). Next, cells were pelleted by centrifugation, washed twice with a 10 mM Tris/HCl Buffer (pH = 7) and stored at -20°C. Preparation of cell extracts, 2-DE and protein identification was performed in exactly the same manner for the labelled/unlabelled cells as for the unlabelled cells.

2.2 Preparation of cell extracts

The -20°C frozen cells were thawed and immediately resuspended in 1.5 ml of 10 mM Tris/HCl buffer (pH = 7), and 25 µl of a protease-inhibitor cocktail (Sigma) was added. Cells were disrupted by sonication for 10 minutes on ice ("Soniprep 150", Sanyo). Insoluble cell material was removed by centrifugation at 13,000 rpm for 10 min. The protein concentration of the supernatant was determined using the Bradford Protein Assay (Sigma). The supernatant was subsequently stored at –80°C.

2.3 2-DE

Gels for the reference map were prepared in triplicate. The extract was mixed with a rehydration buffer containing 50 mM DTT (Sigma), 8 M Urea (Sigma), 2% CHAPS (Sigma), 0.2% (w/v) Pharmalyte ampholytes pH 3-10 (Fluka) and Bromophenol Blue (trace) (Sigma). This mixture was designated as the sample mix. Three IPG strips (pH 3-10) (Bio-Rad) were rehydrated with 300 µl (400 µg) of this sample mix. Strips were allowed to rehydrate overnight. IEF was performed

using a 3-step protocol at a temperature of 20°C using a Protean IEF cell (BioRad). In the first step, the voltage was linearly ramped to 250 V over 30 minutes to desalt the strips. Next, the voltage was linearly ramped to 10,000 V over 2.5 half-hour periods. Finally, the voltage was rapidly ramped to 10,000 V for 40,000 V/hours to complete the focussing. At this stage, the strips were stored overnight at –20°C. Focussed strips were first incubated for 15 minutes in a solution containing 6 M Urea, 2% SDS, 0.375 M Tris-HCl (pH 8.8), 20% glycerol, and 2% w/v DTT. After this, the solution was discarded and the strips were incubated in a solution containing 6 M Urea, 2% SDS, 0.375 M Tris-HCl (pH 8.8), 20% Glycerol, and 4% Iodoacetamide. After equilibration, proteins were separated in the second dimension using SDS-PAGE performed using a Protean II Multicell (Bio-Rad) apparatus on 10% T, 2.6% C gels (17 cm x 17 cm x 1 mm). Electrophoresis was carried out with a constant current of 16 mA/gel for 30 min; subsequently the current was increased to 24 mA/gel for another 7 h.

2.4 Protein visualization and image analysis

Gels were stained using Coomassie Brilliant Blue G250 (Sigma). Gels were scanned using a GS-800 densitometer (BioRad) at 100 µm resolution. All spot detection and quantification was performed with PDQUEST 7.1.0 (BioRad). Staining intensity was normalised against the total staining intensity on the gel. Two hundred fifty-five spots were selected for mass spectrometric analysis. For protein quantitation, metabolic labelling was used, and for this gel image was matched to the reference map and protein spots of interest were selected for MS analysis and quantitation.

2.5 Protein isolation and identification by MS

Spots of interest were excised from the stained 2-DE gels by hand, destained with 200 mM ammonium bicarbonate with 40% acetonitrile. The gel pieces were incubated overnight in a 0.4 µg trypsin solution (Sigma) and 50 µl of 40 mM ammonium bicarbonate in 9% acetonitrile. The next day, peptides were extracted in three subsequent extraction steps using 5 µl of 25 mM NH_4HCO_3 (10 minutes, room temperature), 30 µl acetonitrile (15 minutes, 37°C), 50 µl of 5% formic acid (15 minutes, 37°C) and finally with 30 µl acetonitrile (15 minutes, 37°C). All extracts were pooled and dried in a vacuum centrifuge, then stored at –20°C.

The lyophilised peptide mixture was resuspended in 0.1% formic acid in 3% acetonitrile. This mixture was separated on a PepMap C-18 RP capillary column (LC Packings, Amsterdam, The Netherlands) and eluted in a 30-min gradient via an LC Packings Ultimate nanoLC directly onto the mass spectrometer. Peptides were analysed using an Applied Biosystems QStarXL® electrospray ionisation quadrupole time of flight tandem mass spectrometer (ESI qQ-TOF). The data acquisition on the MS was performed in the positive ion mode using Information Dependent Acquisition (IDA). Peptides with charge states 2 and 3 were selected for tandem mass spectrometry. IDA data were submitted to Mascot for database searching in a sequence query type of search (www.matrixscience.com). The peptide tolerance was set to 2.0 Da and the MS/MS tolerance was set to 0.8 Da. A carbamidomethyl modification of cysteine was set as a fixed modification and methionine oxidation was set as a variable modification. Up to one missed cleavage site by trypsin was allowed. The search was performed against the Mass Spectrometry protein sequence DataBase (MSDB; ftp://ftp.ncbi.nih.gov/repository/MSDB/msdb.nam). Molecular Weight Search (MOWSE) (Pappin, Hojrup et al. 1993) scores greater than 50 were regarded as significant.

2.6 Peptide quantitation

In the metabolic labelling experiments, peptide identification of the light (^{14}N) version of the peptide was performed as described in section 2.5. After this, the heavy ^{15}N version of the peptide could be identified by changing the isotope abundance of ^{15}N nitrogen to 100% in the Analyst software data dictionary. Next, the peak area of both version of the same peptide was integrated over time using LC-MS reconstruction tool in the Analyst software. In addition, an extracted ion chromatogram (XIC) was constructed for each peptide. The XIC is an ion chromatogram, which shows the intensity values of a single mass (peptide) over a range of scans. This tool was used to check for chromatographic shifts between heavy and light versions of the same peptide.

2.7 RNA extraction and probe synthesis

Early-log phase cultures (OD_{600} 0.1-0.2) of *S. solfataricus* grown on 0.1% yeast extract and 0.2% tryptone (YT) or 0.3% D-glucose (G) were quickly cooled in ice-water and harvested by centrifugation at 4°C. The RNA extraction was done as described previously (Brinkman, Bell et al. 2002). Preparation of cDNA was done

as follows: to 15 μg of RNA, 5 μg of random hexamers (Qiagen) was added in a total volume of 11.6 μL. This was incubated for 10 min at 72°C after which the mixture was cooled on ice. Next, dATP, dGTP and dCTP (5 μM final concentration) were added, together with 4 μM aminoallyl dUTP (Sigma), 1 μM dTTP, 10 mM dithiotreitol (DTT), 400 U superscript II (Invitrogen) and the corresponding 5x RT buffer in a final volume of 20 μL. The reverse transcriptase reaction was carried out at 42°C for 1 h. To stop the reaction and to degrade the RNA, 2 μL 200 mM EDTA and 3 μL 1 M NaOH were added to the reaction mixture, after which it was incubated at 70°C for 15 min. After neutralisation by the addition of 3 μL 1 M HCl, the cDNA was purified using a Qiagen MinElute kit according to the manufacturer, except that the wash buffer was replaced with 80% v/v ethanol. The cDNA was then labelled using postlabelling reactive CyDye packs (Amersham Biosciences), according to the protocol provided by the company. Differentially labelled cDNA derived from *S. solfataricus* cells grown on either YT or G media was pooled (15 μg labelled cDNA of each sample) and excess label was removed by cDNA purification using the MinElute kit.

2.8 DNA microarray hybridisation, scanning and data analysis

The design and construction of the microarray, as well as the hybridisation was performed as described previously (Lundgren, Andersson et al. 2004; Andersson, Bernander et al. 2005). After hybridisation, the microarrays were scanned at a resolution of 5 μm with a Genepix 4000B scanner (Axon Instruments) using the appropriate laser and filter settings. Spots were analysed with the Genepix pro 5.0 software package (Axon Instruments). Low-quality spots were excluded using criteria that were previously described (Lundgren, Andersson et al. 2004). ^{2}Log-transformed ratios (^{2}log(YT/G)) from the replicate slides were averaged after first averaging the duplicate spots on the array. Statistical significance for the observed ratios was calculated by doing a Significance Analysis of Microarrays (SAM) analysis (Tusher, Tibshirani et al. 2001). Each ^{2}log value represents two hybridisation experiments, performed in duplicate by using cDNA derived from four different cultures of *S. solfataricus*: two grown on YT media and two grown on glucose media. The result of each ORF therefore consisted of eight pairwise comparisons. The ORFs were categorised according to the 20 functional categories of the comprehensive microbial resource (CMR) (Peterson, Umayam et al. 2001).

2.9 Metabolic pathway reconstruction based on biochemical and genomic data

The reconstruction of the main metabolic pathways was performed with BLASTP and PSI-BLAST programmes (Altschul, Madden et al. 1997) on the nonredundant (NR) database of protein sequences (National Center for Biotechnology Information) by using full length or N-terminal protein sequences. All the sequences were derived from verified enzymatic activities of thermophilic or hyperthermophilic archaea unless stated otherwise. The sequences from *S. acidocaldarius* were analysed by BLASTP programme using the complete genome sequence (Chen et al. unpublished; http//dac.molbio.ku.dk). All the assigned enzymatic functions for the proteins of *Sulfolobus solfataricus* P2 were checked with the annotations in public protein databases, such as the BRaunschweig ENzyme DAtabase (BRENDA) (Schomburg, Chang et al. 2004), Clusters of Orthologous Groups of proteins (COG) (Tatusov, Fedorova et al. 2003), InterPro (Mulder, Apweiler et al. 2005) and the fee-based ERGO bioinformatics suite (Overbeek, Larsen et al. 2003). The reconstructed pathways were compared with previous reports (Huynen, Dandekar et al. 1999; Ronimus and Morgan 2003; Verhees, Kengen et al. 2003) and the Kyoto Encyclopaedia of Genes and Genomes (KEGG) (Kanehisa, Goto et al. 2004).

3 Results and discussion

3.1 Generation and application of a 2-DE map

Figure 1 shows an image of the 2-DE reference map for *S. solfataricus*. With Coomassie Brilliant blue G250, approximately 500 spots were visualised. The highest spot count was obtained in the region pI = 5-9, and proteins ranged in size from 15 to 123 kDa (predicted values). In total, 255 spots were selected for Mass Spectrometry (MS) analysis on the basis of their relative high abundance. In addition, faint spots were selected to test the sensitivity of the MS method. In total, 325 unique proteins in 255 spots were identified, with even the faintest spots yielding significant Molecular Weight Search (MOWSE) scores (> 51). All 255 spots were found on the triplicate gels. The complete dataset is presented in the supplementary material. A subset, representing key elements of central energy metabolism and other relevant proteins is discussed more extensively in this paper. The highest MOWSE score, 1362, was achieved for elongation factor 2 (Sso0728,

spot 26). Generally, one peptide (intact mass and tandem mass spectrometry (MS/MS) ion spectrum) was sufficient for confident identification of a *S. solfataricus* protein against the full Mass Spectrometry protein sequence Database (MSDB). In most cases, however, multiple peptides of the same protein were recovered from a spot. On average, the sequence coverage was 30%. The highest sequence coverage (75%) was found for the α-subunit of the proteasome (Sso0738) in spot 213. There was no correlation between the sequence coverage and the protein size. However, larger proteins usually resulted in higher MOWSE scores. This is due to the fact that larger proteins generate a larger number of unique peptides after tryptic digestion. For example, MOWSE scores greater than 800 were only obtained for proteins larger than 48 kDa.

The number of proteins that matched to ORFs that are either hypothetical or conserved hypothetical proteins was 157 (48%). This is similar compared to the expected 53%, on the basis of the genome composition. This was also found in a similar study on the *Methanocaldococcus jannaschii* proteome (Giometti, Reich et al. 2002). Interestingly, there were only two hypothetical proteins amongst the 20 most intense spots, (Sso0029, Sso0099 relating to spots 130 and 224 respectively). The relatively high abundance of those proteins suggests an important function.

Another important observation is that a number of proteins were found in more than one spot. Interestingly, this was true for a large number of proteins involved in the TCA cycle (e.g. 2-oxoacid:ferredoxin oxidoreductase (Sso2815) was found in eight different spots). There are a number of explanations for this: (i) isoforms or posttranslationally modified versions of the protein might be present in the cell (ii) the protein was modified during protein extraction or during 2-DE (e.g. proteolysis, methionine oxidation), (iii) the protein does not resolve well on the gel and therefore “smears” out over a large pH or mass range, or (iv) the denaturating conditions are not strong enough to completely break protein associations. The presence of a protein in multiple spots was also observed in similar proteomic studies (Giometti, Reich et al. 2002). To find posttranslational modifications (PTMs), all mass spectra were searched again but this time with phosphorylation of serine or threonine, and with methylation set as variable modifications. Unfortunately, no consistent results were obtained, and therefore more specific studies targeted to identify PTMs are necessary.

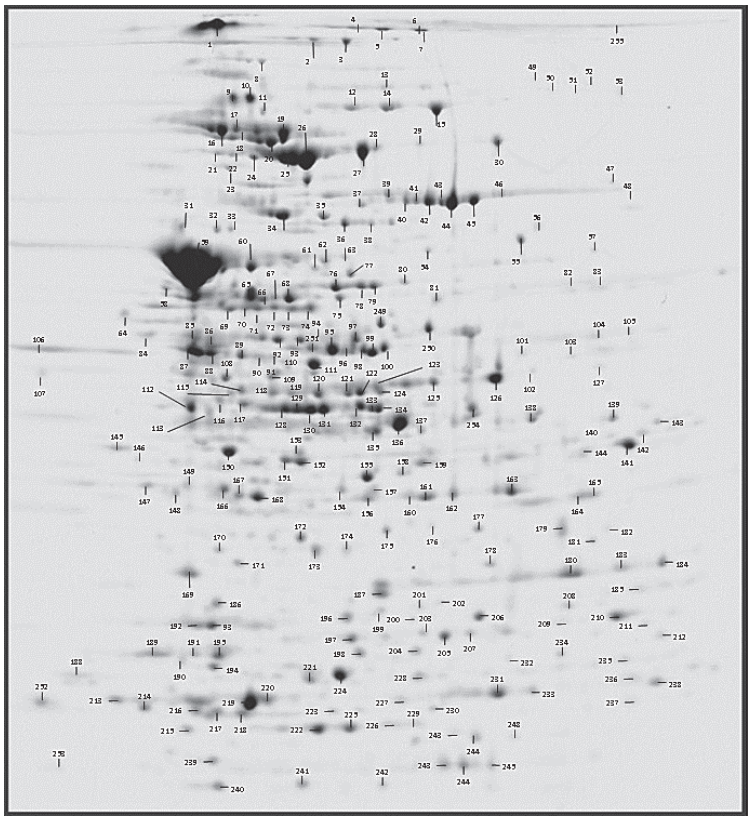

Figure 1. 2-DE reference map for *S. solfataricus* grown on glucose. All numbered spots were subjected to LC-MS-MS analysis. Results are displayed in Table 1 (supplementary).

Another important observation is that a number of proteins were found in more than one spot. Interestingly, this was true for a large number of proteins involved in the TCA cycle (e.g. 2-oxoacid:ferredoxin oxidoreductase (Sso2815) was found in eight different spots). There are a number of explanations for this: (i) isoforms or posttranslationally modified versions of the protein might be present in the cell (ii) the protein was modified during protein extraction or during 2-DE (e.g. proteolysis, methionine oxidation), (iii) the protein does not resolve well on the gel and therefore “smears” out over a large pH or mass range, or (iv) the denaturating conditions are not strong enough to completely break protein associations. The presence of a protein in multiple spots was also observed in similar proteomic studies (Giometti, Reich et al. 2002). To find posttranslational modifications

(PTMs), all mass spectra were searched again but this time with phosphorylation of serine or threonine, and with methylation set as variable modifications. Unfortunately, no consistent results were obtained, and therefore more specific studies targeted to identify PTMs are necessary.

In a number of cases, multiple proteins per spot were found. Often these proteins have similar molecular weights (MW) and iso-electric points (Contursi, Pisani et al.) indicating that the resolution on the gel was insufficient to resolve these proteins into single protein spots. In other cases however, proteins in the same spot differ significantly in MW and pI. These represent biologically interesting cases since these could indicate stable protein associations. An example was found in spot 1, where subunits α, β en γ of aldehyde oxidoreductase (Sso2636, Sso2637, Sso2639) were found.

Protein quantitation was performed on the basis of ^{15}N metabolic labelling as recently described. With this method a number of problems associated with 2-DE (e.g. multiple proteins per spot) can be avoided. Moreover, the reproducibility of gel staining becomes of lesser importance since protein quantitation takes place on the MS (Snijders, de Vos et al. 2005).

Figure 2 shows an example of a TOF-MS spectrum containing both the light and the heavy versions of the peptide IFGSLSSNYVLTK. This peptide is derived from the 2-keto-3-deoxy gluconate aldolase (Sso3197). The light peptide at m/z 714.99 corresponds to the yeast extract + tryptone (YT) grown cells and the heavy peptide at m/z 722.47 corresponds to the glucose (G)-grown cells. The relative abundance of the heavy and light peptide can now be calculated by determining the ratio of the peak areas. Note that the difference in mass between the heavy and light version of the peptide corresponded exactly to the number of nitrogen atoms in the peptide, in this case 15 atoms (m/z = 7.5). Table 1 summarises the differential proteomic data obtained in this way, as well as the corresponding transcriptomic data. In Section 3.3, this data are used for a discussion of the central carbon metabolism in *S. Solfataricus*.

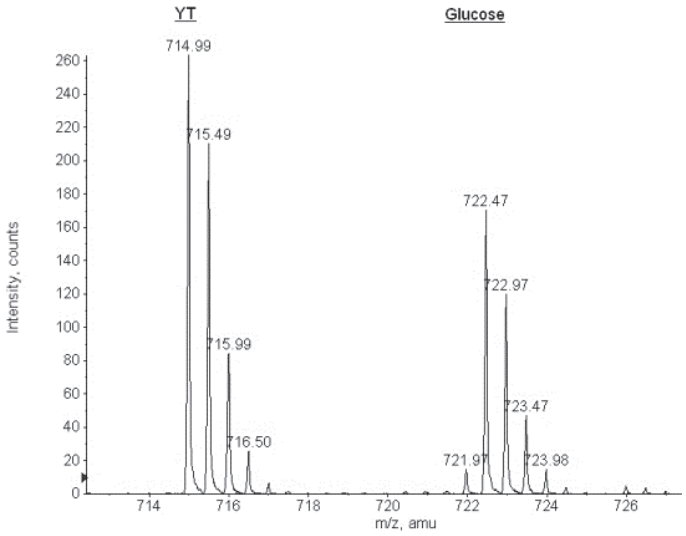

Figure 2. Peptide quantitation.

TOF MS spectrum of a ^{15}N labelled and an unlabeled peptide. The peak on the left at m/z 714.99 represents the unlabeled version of the peptide (protein from cells grown on yeast extract + tryptone (YT)). The peak at the right at m/z 7.22.47 represents the ^{15}N labelled version of the peptide (protein from cells grown on glucose). This peptide was identified as IFGSLSSNYVLTK, corresponding to 2-keto-3-deoxy gluconate aldolase (Sso3197). The ratio between the areas of the heavy and light versions of this peptide was 1.56.2

3.2 Exploration of the transcriptome

In total, 1581 of the 2315 genes printed on the microarray were used in the analysis (selected, according to criteria described in Section 2.8). There were 184 significantly differentially expressed genes ($p<0.05$; p is the statistical certainty that the observed change in ratio is not caused by a biological effect). In total, 135 and 49 genes are up-regulated under glucose and YT conditions respectively. Of these up-regulated genes 23 and 20% were annotated as either hypothetical or conserved hypothetical. Interestingly, these percentages are lower than the expected 53%.

Of the up-regulated genes, 16% and 10% were involved in amino acid metabolism under glucose and YT conditions, respectively. Although knowledge about the amino acid metabolism in S. *solfataricus* is limited, regulation in this functional group was expected since amino acids are expected to be synthesised under glucose conditions and predominantly degraded under YT conditions. This data, therefore, provide an excellent starting point for amino acid metabolism

reconstruction. Future biochemical and proteomic studies are necessary to confirm the exact composition and direction of the responsible pathways.

Interestingly, three genes involved in nitrogen metabolism were regulated: (i) glutamate synthase (Sso0684; 0.15) (ii) glutamine synthase (Sso0366; 0.27) and (iii) glutamate dehydrogenase (Sso2044; 6.29), absolute ratios are given as YT/G. These results show that cells which grow on glucose assimilate nitrogen by the sequential action of glutamine synthase and glutamate synthase. Under YT conditions, glutamate dehydrogenase produces free ammonium by converting glutamate into 2-oxoglutarate. This is necessary because there is an excess of nitrogen bound to carbon when grown in the presence of YT.

Transport and binding proteins are also a major group of up-regulated genes (12 and 8% for glucose and YT respectively). Previously, it was shown that both glucose and YT grown cells have the capacity to transport glucose (Elferink, Albers et al. 2001). This is reflected by the fact that the genes involved in glucose transport were not differentially expressed (Sso2847, Sso2848, Sso2849, Sso2850). In addition, genes involved in dipeptide transport were up-regulated under YT conditions (Sso1282; 2.01 / Sso2615; 1.74 / Sso2616; 1.57). Interestingly, genes involved in maltose transport were slightly up-regulated under glucose conditions (Sso3053; 0.36 / Sso3058; 0.50 / Sso3059; 0.53).

3.3 Metabolic pathway reconstruction

During the last two decades, the main metabolic pathways in *Sulfolobus* spp. have been the subject of extensive experimental research. This has led to a profound understanding of the enzymes and protein complexes that are involved in the glycolysis, the tricarboxylic acid cycle (TCA) and related metabolic conversions (Danson 1988; Verhees, Kengen et al. 2003). The availability of the genome sequences of *S. solfataricus* (She, Singh et al. 2001), *S. tokodaii* (Kawarabayasi, Hino et al. 2001) and *S. acidocaldarius* (Chen et al. unpublished work; http//dac.molbio.ku.dk) has recently allowed for the identification of the genes encoding these proteins by matching full-length or N-terminal protein sequences to the predicted proteomes. A reconstruction of the central carbon metabolic pathways in *S. solfataricus* was performed (Fig. 3). The results should be taken with a degree of caution since significant differences exist in the physiology between the three *Sulfolobus* species (Schafer 1996). Almost all proteins involved in this scheme have been experimentally verified in either *Sulfolobus* spp. or other

hyperthermophilic Archaea, such as *Thermoproteus tenax*, *Archaeoglobus fulgidus*, *Thermoplasma acidophilum*, *P. furiosus*, *Thermococcus kodakaraensis*, *Methanothermus fervidus* and *M. jannaschii*. Moreover, the vast majority of the anticipated proteins in *S. solfataricus* were found on the 2-DE reference map (Fig. 2). On average, the TCA cycle proteins made up approximately 12% of the total staining intensity.

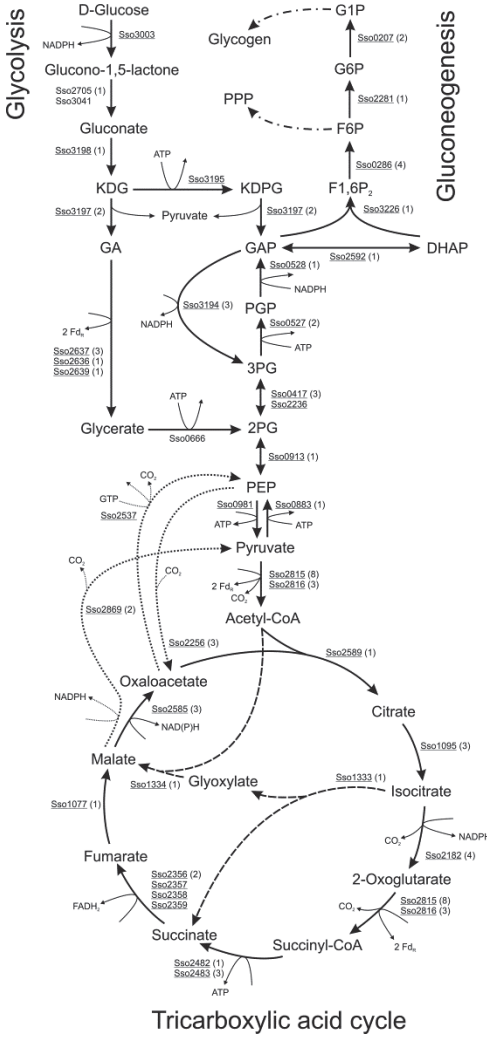

Figure 3. Reconstruction of the central metabolic pathways in *Sulfolobus solfataricus*.
Genes involved in the glycolysis, gluconeogenesis and citric acid cycle were surveyed and are indicated by their locus name. Underlined genes were experimentally verified in *Sulfolobus* or related hyper-thermophilic Archaea (Table 1). The number of spots that were found on the 2-DE reference map is indicated between brackets. The glyoxylate shunt in shown by dashed arrows, while the three to four carbon interconversions are depicted by dotted arrows. Mixed dashed and dotted arrows indicate that the exact pathway to glycogen and pentoses is unknown. The following abbreviations were used: KD(P)G 2-keto-3-deoxy-D-gluconate-(6-phosphate), GA(P) gly-ceraldehyde-(3-phosphate), PGP 1,3-bi-sphosphoglycerate, 3PG 3-phosphoglycerate, 2PG 2-phosphoglycerate, PEP phospho-*enol*pyruvate, DHAP dihydroxyacetone-phosphate, $F1,6P_2$ fructose-1,6-bisphosphate, F6P fructose-6-phosphate, G6P glucose-6-phosphate, G1P glucose-1-phosphate, Fd_R reduced ferredoxin, PPP pentose phosphate pathway. NAD(P)H indicates that both NAD^+ and $NADP^+$ can be used as a cofactor. Arrows represent the presumed physiologically relevant direction of catalysis and are not indicative of enzymatic reversibility.

3.4 Glycolysis and gluconeogenesis

The genus *Sulfolobus* is known to degrade glucose according to a modified version of the Entner-Doudoroff (ED) pathway. While in most cases phosphorylation in the bacterial ED pathway occurs at the level of glucose, gluconate or 2-keto-3-deoxygluconate (KDG), *S. solfataricus* has been reported to utilize a nonphosphorylated version of the ED pathway, which phosphorylates only at the level of glycerate (De Rosa, Gambacorta et al. 1984; Selig, Xavier et al. 1997). Recent experimental findings, however, indicated the presence of a semi-phosphorylated ED pathway, in which KDG is phosphorylated and subsequently cleaved forming pyruvate and glyceraldehyde-3-phosphate (GAP) by the action of the KDG kinase (Sso3195) and the KDG aldolase (Sso3197) respectively. GAP is then oxidised by a nonphosphorylating GAP dehydrogenase (GAPN, Sso3194) forming 3-phosphoglycerate (3PG) (Ahmed, Ettema et al. 2005). The only net difference between the non- and semiphosphorylated pathways is the fact that either reduced ferredoxin (Fd_R) or NADPH is produced, since neither pathway directly yields ATP by substrate level phosphorylation.

The intrinsic irreversibility of several ED enzymes, such as the gluconate dehydratase, the aldehyde oxidoreductase and GAPN, prevents the ED to operate in the gluconeogenic direction, which is, for instance, required to store energy in the form of glycogen (Skorko, Osipiuk et al. 1989). Another important role for the gluconeogenic EMP pathway is the production of fructose-6-phosphate (F6P), which has been proposed to be the main precursor for the Pentose Phosphate Pathway (PPP) (Verhees, Kengen et al. 2003). Except for three kinases (GK glucokinase, PFK phosphofructokinase and PK), the catabolic Embden-Meyerhof-Parnas (EMP) pathway consists of reversible enzymes. Although the genes encoding a GK and PFK were absent, the genes encoding the reversible EMP enzymes were all found in the genome of *Sulfolobus*. Moreover, a gene encoding a fructose-1,6-bisphosphatase (FBPase) was also detected. Because it is known that the catabolic EMP pathway is not operational in *Sulfolobus* (Selig, Xavier et al. 1997), it is likely that these EMP enzymes serve a gluconeogenic role. The simultaneous operation of both the ED and a gluconeogenic EMP pathway, however, requires a strict control of the metabolic flux through the pathway in order to prevent an energetically futile cycle. Allosteric regulation, posttranslational protein modification and regulation at the transcriptional level are common

strategies to modulate the activity and abundance of key enzymes, such as the fructose-1,6-bisphosphatase.

Although glycolysis in *Sulfolobus* is well studied, there are still unconfirmed genes and activities in the pathway. For instance, the transcriptome analysis revealed the expression of one of two putative gluconolactonases (Sso2705) that have generally been omitted in the analysis of the ED pathway, since the reaction from gluconolactone to gluconate also occurs spontaneously (Satory, Furlinger et al. 1997). The expression of the enzyme, however, would suggest a functional role in the metabolism of *Sulfolobus*. Additionally, only one of two phosphoglycerate mutases (Sso0417) that were found in its genome was expressed in both the proteome and transcriptome, while the other type (Sso2236) remained undetected. Expression of the predicted glycerate kinase (Sso0666) was only detected at the mRNA level.

3.5 TCA cycle

Sulfolobus spp. is an obligate aerobe that primarily obtains energy by the oxidation of organic molecules and elemental sulphur (Brock, Brock et al. 1972). This oxidation results in the formation of reduced electron carriers, such as NAD(P)H, Fd_R and $FADH_2$. The majority of these reducing equivalents are generated in the TCA cycle. Per round of the cycle, the succinate-CoA ligase of *Sulfolobus* generates one molecule of ATP, instead of the commonly produced GTP (Danson, Black et al. 1985). Apart from being the main metabolic converter of chemical energy, the TCA cycle intermediates serve an important role as biosynthetic precursors for many cellular components, such as amino acids. Consequently, when too many intermediates are withdrawn from the cycle, they need to be replenished by anaplerotic enzyme reactions. The phospho*enol*pyruvate carboxylase (PEPC), which forms oxaloacetate from phospho*enol*pyruvate, is the only anaplerotic enzyme from *Sulfolobus*, which has been described to date (Sako, Takai et al. 1996; Ettema, Makarova et al. 2004). A gene product with high similarity to known pyruvate carboxylases could not be detected in the predicted proteome of *Sulfolobus*. In the glyoxylate shunt, which is normally only active during growth on acetate, isocitrate and acetyl-CoA are converted into succinate and malate by the action of the isocitrate lyase and the malate synthase. Interestingly, the isocitrate lyase of glucose-grown *S. acidocaldarius* cells co-purified with the aconitase (Uhrigshardt, Walden et al. 2001; Uhrigshardt, Walden

et al. 2002). Not only would this suggest a cytosolic association of the enzymes, but it also suggests that the glyoxylate shunt operates under saccharolytic conditions. This pathway may therefore constitute another way of replenishing four-carbon TCA cycle intermediates.

When there is an excess of TCA intermediates, for instance during growth on proteinaceous substrates, both malate and oxaloacetate can be decarboxylated to pyruvate by the malic enzyme (Bartolucci, Rella et al. 1987). Oxaloacetate can also be converted to phospho*enol*pyruvate by the GTP-dependent carboxykinase (Fukuda, Fukui et al. 2004). These four-to-three carbon conversions then provide the precursors that are required in, for instance, the gluconeogenesis pathway. In contrast to aerobic bacteria and eukaryotes, *Sulfolobus* uses ferredoxin instead of NAD^+ as a cofactor in the formation of acetyl-CoA from pyruvate and succinyl-CoA from 2-oxoglutarate (Kerscher, Nowitzki et al. 1982). The protein complex responsible for both conversions was shown to consist of two subunits; a fused α/γ subunit (Sso2815) and a β subunit (Sso2816) (Zhang, Iwasaki et al. 1996; Fukuda and Wakagi 2002). The genome sequences of the three *Sulfolobus* species, however, revealed several paralogues of ferredoxin-dependent 2-oxoacid oxidoreductases, which might also be involved in these conversions.

What is also evident from this reconstruction is that almost all dehydrogenases in the central carbon metabolism of *Sulfolobus* show a clear cofactor preference for $NADP^+$ over NAD^+ (Danson, Black et al. 1985; Bartolucci, Rella et al. 1987; Camacho, Brown et al. 1995; Russo, Rullo et al. 1995; She, Singh et al. 2001; Lamble, Heyer et al. 2003). The only exception to this rule seems to be the malate dehydrogenase, which, at least *in vitro*, uses both electron acceptors equally well (Hartl, Grossebuter et al. 1987). In bacteria and eukaryotes, most NADPH is usually formed in the PPP and used for reductive biosynthesis purposes. In *Sulfolobus*, the apparent enzyme preference for $NADP^+$ would suggest a more general role of its reduced form, in energy conservation by oxidative phosphorylation. Interestingly, as noted by She, *et al.* [2] all genes encoding the NAD(P)H dehydrogenase complex are present in the genome, except the three that encode the subunits which are required for NAD(P)H binding and oxidation. It has been proposed that the reducing equivalents are first transferred to ferredoxin by na NADPH:ferredoxin oxidoreductase, before entering the respiratory chain [2].

3.6 Regulation of the main metabolic pathways

Insight was obtained into the regulation of the genes anticipated in glycolysis, gluconeogenesis and TCA cycle by measuring the relative abundance of their mRNA and protein levels by using a whole-genome DNA microarray and a quantitative proteomics approach respectively (Table 1). In the measurements, 35 out of 41 transcripts ratios were determined, while 29 out of 41 protein ratios were analysed on 2-DE gels. On average the proteomic and transcriptomic data correlate reasonably well. For 26 genes both proteomic data and transcriptomic data are presented. In general, changes at proteomic and transcriptomic level show a similar trend, however, proteomic changes tend to be more pronounced. In only three cases the proteomic data contradict the transcriptome data. This concerns the three subunits for aldehyde dehydrogenase (Sso2639, Sso2636 and Sso2637). However, the fact that these clustered genes show a similar ratio at proteomic or transcriptomic level indicates the consistency of the data. Interestingly, all three subunits were found in the same protein spot on the gel, suggesting that a strong (non-covalent) interaction exists between them. The stability of the protein complex might be affected by stabilising factors such as co-factors that may lead to different degrees of aggregation under different growth conditions. In terms of regulatory effects, the glyceraldehyde-3-phosphate dehydrogenase (non-phosphorylating; GAPN) was up-regulated under glucose conditions, or alternatively, down-regulated during growth in YT media. This is not surprising, since GAP is the crucial intermediate between the ED and gluconeogenic EMP, and too much of the strictly catabolic GAPN would be likely to interfere with gluconeogenesis. The enzymes involved in gluconeogenesis were all slightly up-regulated during growth on YT media, in agreement with expectations. Especially the phospho*enol*pyruvate synthase and the phosphoglycerate kinase, key enzymes of the pathway, appeared to be most differentially expressed.

The expression levels of the TCA-cycle genes were only marginally different under the two conditions. Under glucose conditions, several enzymes of the TCA cycle were slightly induced at proteomic level, including the 2-oxoacid:ferredoxin oxidoreductase, the succinate-CoA ligase, the succinate dehydrogenase and the malate dehydrogenase. This was also true for some enzymes that replenish the four-carbon TCA cycle intermediates, such as the isocitrate lyase and the phospho*enol*pyruvate carboxylase. This ensures that sufficient oxaloacetate is present to serve as biosynthetic precursor and as an acceptor molecule for acetyl-

CoA. The differences may be due to the fact that glucose catabolism mainly results in acetyl-CoA and oxaloacetate formation, whereas peptide degradation probably yields various central intermediates of carbon metabolism, such as pyruvate (Ala, Cys, Trp, Thr, Ser, Gly), acetyl-CoA (Phe, Tyr, Ile, Leu, Lys, Trp, Thr), 2-oxoglutarate (Arg, Gln, His, Pro, Glu), succinyl-CoA (Ile, Met, Val, Thr), fumarate (Phe, Tyr, Asp) and oxaloacetate (Asn, Asp).

4 Concluding remarks

In this study, we have created a proteome reference map for *S. solfataricus* consisting of 325 proteins in 255 spots, and have reconstructed its central carbon metabolic pathways. The expression of the genes in these pathways was analysed by measuring the relative abundance of mRNA and protein under peptide- or sugar-degrading conditions. Surprisingly, most observed differences were small. Despite this, the expression of some key enzymes in glycolysis, gluconeogenesis and TCA cycle were significantly altered. Apart from looking at abundance levels, proteomics studies that focus on the modulation of enzyme activity by protein PTM are now ongoing. These studies will provide additional clues that will reveal the details of regulation of the central carbon metabolism in *S. solfataricus*.

5 References

[1] Zillig, W., Stetter, K.O., Wunderl, S., Schulz, W., Priess, H., et al., Arch. Microbiol., 1980. 125: p. 259-269.
[2] She, Q., Singh, R.K., Confalonieri, F., Zivanovic, Y., Allard, G., et al., Proc Natl Acad Sci U S A, 2001. 98(14): p. 7835-40.
[3] Peterson, J.D., Umayam, L.A., Dickinson, T., Hickey, E.K., White, O., Nucleic Acids Res, 2001. 29(1): p. 123-5.
[4] Cannio, R., Contursi, P., Rossi, M., Bartolucci, S., Extremophiles, 2001. 5(3): p. 153-9.
[5] Stedman, K.M., Schleper, C., Rumpf, E., Zillig, W., Genetics, 1999. 152(4): p. 1397-405.
[6] Worthington, P., Hoang, V., Perez-Pomares, F., Blum, P., J Bacteriol, 2003. 185(2): p. 482-8.
[7] Jonuscheit, M., Martusewitsch, E., Stedman, K.M., Schleper, C., Mol Microbiol, 2003. 48(5): p. 1241-52.
[8] Contursi, P., Cannio, R., Prato, S., Fiorentino, G., Rossi, M., et al., FEMS Microbiol Lett, 2003. 218(1): p. 115-20.
[9] Verhees, C.H., Kengen, S.W., Tuininga, J.E., Schut, G.J., Adams, M.W., et al., Biochem J, 2003. 375(Pt 2): p. 231-46.
[10] Adams, M.W.W., Holden, J.F., Menon, A.L., Schut, G.J., Grunden, A.M., et al., J. Bacteriol., 2001. 183(2): p. 716-724.
[11] Schonheit, P., Schafer, T., World J. Microbiol. Biotechnol., 1995. 11(1): p. 26-57.
[12] Schafer, G., Biochim Biophys Acta, 1996. 1277(3): p. 163-200.
[13] De Rosa, M., Gambacorta, A., Nicolaus, B., Giardina, P., Poerio, E., et al., Biochem J, 1984. 224(2): p. 407-14.
[14] Ahmed, H., Ettema, T.J.G., Tjaden, B., Geerling, A.C.M., van der Oost, J., et al., Biochem J, 2005. in press.
[15] Lamble, H.J., Heyer, N.I., Bull, S.D., Hough, D.W., Danson, M.J., J Biol Chem, 2003.
[16] Lamble, H.J., Heyer, N.I., Bull, S.D., Hough, D.W., Danson, M.J., J Biol Chem, 2003. 278(36): p. 34066-72.

[17] Schut, G.J., Brehm, S.D., Datta, S., Adams, M.W., J Bacteriol, 2003. 185(13): p. 3935-47.
[18] Zaigler, A., Schuster, S.C., Soppa, J., Mol Microbiol, 2003. 48(4): p. 1089-105.
[19] Snijders, A.P., de Vos, M.G., Wright, P.C., J Proteome Res, 2005. 4(2): p. 578-85.
[20] Pappin, D.J., Hojrup, P., Bleasby, A.J., Curr Biol, 1993. 3(6): p. 327-32.
[21] Brinkman, A.B., Bell, S.D., Lebbink, R.J., de Vos, W.M., van der Oost, J., J Biol Chem, 2002. 277(33): p. 29537-49.
[22] Lundgren, M., Andersson, A., Chen, L., Nilsson, P., Bernander, R., Proc Natl Acad Sci U S A, 2004. 101(18): p. 7046-51.
[23] Andersson, A., Bernander, R., Nilsson, P., Bioinformatics, 2005. 21(3): p. 325-32.
[24] Tusher, V.G., Tibshirani, R., Chu, G., Proc Natl Acad Sci U S A, 2001. 98(9): p. 5116-21.
[25] Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., et al., Nucleic Acids Res, 1997. 25(17): p. 3389-402.
[26] Schomburg, I., Chang, A., Ebeling, C., Gremse, M., Heldt, C., et al., Nucleic Acids Res, 2004. 32(Database issue): p. D431-3.
[27] Tatusov, R.L., Fedorova, N.D., Jackson, J.D., Jacobs, A.R., Kiryutin, B., et al., BMC Bioinformatics, 2003. 4(1): p. 41.
[28] Mulder, N.J., Apweiler, R., Attwood, T.K., Bairoch, A., Bateman, A., et al., Nucleic Acids Res, 2005. 33(Database issue): p. D201-5.
[29] Overbeek, R., Larsen, N., Walunas, T., D'Souza, M., Pusch, G., et al., Nucleic Acids Res, 2003. 31(1): p. 164-71.
[30] Huynen, M.A., Dandekar, T., Bork, P., Trends Microbiol, 1999. 7(7): p. 281-91.
[31] Ronimus, R.S., Morgan, H.W., Archaea, 2003. 1(3): p. 199-221.
[32] Kanehisa, M., Goto, S., Kawashima, S., Okuno, Y., Hattori, M., Nucleic Acids Res, 2004. 32(Database issue): p. D277-80.
[33] Giometti, C.S., Reich, C., Tollaksen, S., Babnigg, G., Lim, H., et al., J Chromatogr B Analyt Technol Biomed Life Sci, 2002. 782(1-2): p. 227-43.
[34] Snijders, A.P.L., de Vos, M.G., Koning, B., Wright, P.C., Electrophoresis, 2005. in press.
[35] Elferink, M.G., Albers, S.V., Konings, W.N., Driessen, A.J., Mol Microbiol, 2001. 39(6): p. 1494-503.
[36] Danson, M.J., Adv Microb Physiol, 1988. 29: p. 165-231.
[37] Kawarabayasi, Y., Hino, Y., Horikawa, H., Jin-no, K., Takahashi, M., et al., DNA Res, 2001. 8(4): p. 123-40.
[38] Selig, M., Xavier, K.B., Santos, H., Schonheit, P., Arch Microbiol, 1997. 167(4): p. 217-32.
[39] Skorko, R., Osipiuk, J., Stetter, K.O., J Bacteriol, 1989. 171(9): p. 5162-4.
[40] Satory, M., Furlinger, M., Haltrich, D., Kulbe, K.B., Pittner, F., et al., Biotechnol Lett, 1997. 19(12): p. 1205-08.
[41] Brock, T.D., Brock, K.M., Belly, R.T., Weiss, R.L., Arch Mikrobiol, 1972. 84(1): p. 54-68.
[42] Danson, M.J., Black, S.C., Woodland, D.L., Wood, P.A., FEBS, 1985. 179(1): p. 120-4.
[43] Sako, Y., Takai, K., Uchida, A., Ishida, Y., FEBS Lett, 1996. 392(2): p. 148-52.
[44] Ettema, T.J., Makarova, K.S., Jellema, G.L., Gierman, H.J., Koonin, E.V., et al., J Bacteriol, 2004. 186(22): p. 7754-62.
[45] Uhrigshardt, H., Walden, M., John, H., Anemuller, S., Eur J Biochem, 2001. 268(6): p. 1760-71.
[46] Uhrigshardt, H., Walden, M., John, H., Petersen, A., Anemuller, S., FEBS Lett, 2002. 513(2-3): p. 223-9.
[47] Bartolucci, S., Rella, R., Guagliardi, A., Raia, C.A., Gambacorta, A., et al., J Biol Chem, 1987. 262(16): p. 7725-31.
[48] Fukuda, W., Fukui, T., Atomi, H., Imanaka, T., J Bacteriol, 2004. 186(14): p. 4620-7.
[49] Kerscher, L., Nowitzki, S., Oesterhelt, D., Eur J Biochem, 1982. 128(1): p. 223-30.
[50] Zhang, Q., Iwasaki, T., Wakagi, T., Oshima, T., J Biochem (Tokyo), 1996. 120(3): p. 587-99.
[51] Fukuda, E., Wakagi, T., Biochim Biophys Acta, 2002. 1597(1): p. 74-80.
[52] Russo, A.D., Rullo, R., Masullo, M., Ianniciello, G., Arcari, P., et al., Biochem Mol Biol Int, 1995. 36(1): p. 123-35.
[53] Camacho, M.L., Brown, R.A., Bonete, M.J., Danson, M.J., Hough, D.W., FEMS Microbiol Lett, 1995. 134(1): p. 85-90.
[54] Hartl, T., Grossebuter, W., Gorisch, H., Stezowski, J.J., Biol Chem Hoppe Seyler, 1987. 368(3): p. 259-67.
[55] Kim, S., Lee, S.B., Biochem J, 2005. 387(Pt 1): p. 271-80.
[56] Lamble, H.J., Milburn, C.C., Taylor, G.L., Hough, D.W., Danson, M.J., FEBS Lett, 2004. 576(1-2): p. 133-6.
[57] Buchanan, C.L., Connaris, H., Danson, M.J., Reeve, C.D., Hough, D.W., Biochem J, 1999. 343 Pt 3: p. 563-70.
[58] Brunner, N.A., Brinkmann, H., Siebers, B., Hensel, R., J Biol Chem, 1998. 273(11): p. 6149-56.

[59] Kardinahl, S., Schmidt, C.L., Hansen, T., Anemuller, S., Petersen, A., et al., Eur J Biochem, 1999. 260(2): p. 540-8.
[60] Schramm, A., Siebers, B., Tjaden, B., Brinkmann, H., Hensel, R., J Bacteriol, 2000. 182(7): p. 2001-9.
[61] van der Oost, J., Huynen, M.A., Verhees, C.H., FEMS Microbiol Lett, 2002. 212(1): p. 111-20.
[62] Peak, M.J., Peak, J.G., Stevens, F.J., Blamey, J., Mai, X., et al., Arch Biochem Biophys, 1994. 313(2): p. 280-6.
[63] Hutchins, A.M., Holden, J.F., Adams, M.W., J Bacteriol, 2001. 183(2): p. 709-15.
[64] Hess, D., Kruger, K., Knappik, A., Palm, P., Hensel, R., Eur J Biochem, 1995. 233(1): p. 227-37.
[65] Kohlhoff, M., Dahm, A., Hensel, R., FEBS Lett, 1996. 383(3): p. 245-50.
[66] Siebers, B., Brinkmann, H., Dorr, C., Tjaden, B., Lilie, H., et al., J Biol Chem, 2001. 276(31): p. 28710-8.
[67] Nishimasu, H., Fushinobu, S., Shoun, H., Wakagi, T., Structure (Camb), 2004. 12(6): p. 949-59.
[68] Hansen, T., Wendorff, D., Schonheit, P., J Biol Chem, 2004. 279(3): p. 2262-72.
[69] Solow, B., Bischoff, K.M., Zylka, M.J., Kennelly, P.J., Protein Sci, 1998. 7(1): p. 105-11.
[70] Smith, L.D., Stevenson, K.J., Hough, D.W., Danson, M.J., FEBS Lett, 1987. 225(1,2): p. 277-81.
[71] Lohlein-Werhahn, G., Goepfert, P., Eggerer, H., Biol Chem Hoppe Seyler, 1988. 369(2): p. 109-13.
[72] Janssen, S., Schafer, G., Anemuller, S., Moll, R., J Bacteriol, 1997. 179(17): p. 5560-9.
[73] Puchegger, S., Redl, B., Stoffler, G., J Gen Microbiol, 1990. 136(8): p. 1537-41.
[74] Colombo, S., Grisa, M., Tortora, P., Vanoni, M., FEBS Lett, 1994. 337(1): p. 93-8.

Figures and Tables

Table 1: Relative abundances of mRNA and protein levels of the genes involved in central metabolic pathways of *Sulfolobus solfataricus* grown on yeast extract and tryptone (YT) compared to glucose (G).

	Locus	Enzyme description	EC	COG	Transcriptomics[a]	Proteomics[a]	Reference
			Glycolysis				
	Sso3003	Glucose-1-dehydrogenase	1.1.1.47	1063	NS	NF	(Lamble, Heyer et al. 2003)
	Sso2705	Gluconolactonase	3.1.1.17	3386	1.15 ± 0.07	NF	(Verhees, Kengen et al. 2003)
	Sso3041	Gluconolactonase	3.1.1.17	3386	NF	NF	
	Sso3198	Gluconate dehydratase	4.2.1.39	4948	1.00 ± 0.07	1.42 ± 0.14	(Lamble, Milburn et al. 2004; Kim and Lee 2005)
	Sso3197	2-keto-3-deoxy Gluconate aldolase	4.1.2.-	0329	0.96 ± 0.19	1.55 ± 0.05	(Buchanan, Connaris et al. 1999)
	Sso3195	2-keto-3-deoxy Gluconate kinase	2.7.1.45	0524	1.19 ± 0.15	NF	(Verhees, Kengen et al. 2003)
	Sso3194	Glyceraldehyde-3-phosphate dehydrogenase (non-phosphorylating)	1.2.1.3	1012	0.87 ± 0.10	0.66 ± 0.07	(Brunner, Brinkmann et al. 1998; Ahmed, Ettema et al. 2005)
	Sso2639[c]	Aldehyde oxidoreductase, α-subunit	1.2.7.-	1529	0.65 ± 0.01[b]	4.51 ± 0.78	(Kardinahl, Schmidt et al. 1999)
	Sso2636[c]	Aldehyde oxidoreductase, β-subunit		1319	0.55 ± 0.13[b]	4.89 ± 0.40	
	Sso2637[c]	Aldehyde oxidoreductase, γ-subunit		2080	0.62 ± 0.14[b]	4.22 ± 1.03	
	Sso0666	Glycerate kinase	2.7.1.-	2379	0.70 ± 0.24	NF	(De Rosa, Gambacorta et al. 1984; Verhees, Kengen et al. 2003)
	Sso0981	Pyruvate kinase	2.7.1.40	0469	0.98 ± 0.10	NF	(Schramm, Siebers et al. 2000)
			Glycolysis / Gluconeogenesis				
	Sso0417	Phosphoglycerate mutase	5.4.2.1	3635	1.03 ± 0.13	1.55 ± 0.14	(van der Oost, Huynen et al. 2002)
	Sso2236	Phosphoglycerate mutase	5.4.2.1	0406	NS	NF	

	Sso0913	Enolase	4.2.1.11	0148	1.36 ± 0.35	1.59 ± 0.23	(Peak, Peak et al. 1994)
Gluconeogenesis							
	Sso0883	Phospho*enol*pyruvate synthase	2.7.9.2	0574	1.62 ± 0.08[b]	1.77 ± 0.22	(Hutchins, Holden et al. 2001)
	Sso0527	Phosphoglycerate kinase	2.7.2.3	0126	1.26 ± 0.26	2.30 ± 0.28	(Hess, Kruger et al. 1995)
	Sso0528	Glyceraldehyde-3-phosphate dehydrogenase (phosphorylating)	1.2.1.12	0057	1.07 ± 0.20	1.16 ± 0.02	(Russo, Rullo et al. 1995)
	Sso2592	Triose-phosphate isomerase	5.3.1.1	0149	NF	1.17 ± 0.12	(Kohlhoff, Dahm et al. 1996)
	Sso3226	Fructose-bisphosphate aldolase	4.1.2.13	1830	NS	1.84 ± 0.10	(Siebers, Brinkmann et al. 2001)
	Sso0286	Fructose-bisphosphatase	3.1.3.11	1980	1.24 ± 0.18	1.32 ± 0.05	(Nishimasu, Fushinobu et al. 2004)
	Sso2281	Glucose-6-phosphate isomerase	5.3.1.9	0166	1.01 ± 0.13	1.51 ± 0.10	(Hansen, Wendorff et al. 2004)
	Sso0207	Phosphoglucomutase	5.4.2.2	1109	1.03 ± 0.32	1.55 ± 0.01	(Solow, Bischoff et al. 1998)
Tricarboxylic acid cycle							
	Sso2589	Citrate synthase	2.3.3.1	0372	0.84 ± 0.09	1.02 ± 0.03	(Smith, Stevenson et al. 1987; Lohlein-Werhahn, Goepfert et al. 1988)
	Sso1095	Aconitase	4.2.1.3	1048	1.05 ± 0.14	1.11 ± 0.03	(Uhrigshardt, Walden et al. 2001)
	Sso2182	Isocitrate dehydrogenase	1.1.1.42	0538	1.34 ± 0.65	1.18 ± 0.03	(Camacho, Brown et al. 1995)
	Sso2815[d]	2-oxoacid:ferredoxin oxidoreductase α/γ-subunit	1.2.7.1 1.2.7.3	0674 1014	0.89 ± 0.07	0.56 ± 0.05	(Kerscher, Nowitzki et al. 1982; Zhang, Iwasaki et al. 1996; Fukuda and Wakagi 2002)
	Sso2816[d]	2-oxoacid:ferredoxin oxidoreductase β-subunit		1013	0.85 ± 0.31	0.60 ± 0.02	
	Sso2482	Succinate-CoA ligase, α-subunit	6.2.1.5	0074	0.93 ± 0.25	0.54 ± 0.04	(Danson, Black et al. 1985)

	Sso2483	Succinate-CoA ligase, β-subunit		0045	0.94 ± 0.30	0.51± 0.05	
	Sso2356	Succinate dehydrogenase, subunit A	1.3.99.1	1053	NS	0.58 ± 0.4	(Janssen, Schafer et al. 1997)
	Sso2357	Succinate dehydrogenase, subunit B		0479	0.75 ± 0.28	NF	
	Sso2358	Succinate dehydrogenase, subunit C		2048	0.94 ± 0.27	NF	
	Sso2359	Succinate dehydrogenase, subunit D			0.89 ± 0.16	NF	
	Sso1077	Fumarate hydratase	4.2.1.2	0114	1.08 ± 0.10	1.53 ± 0.09	(Puchegger, Redl et al. 1990; Colombo, Grisa et al. 1994)
	Sso2585	Malate dehydrogenase	1.1.1.37	0039	0.82 ± 0.27	0.69 ± 0.01	(Hartl, Grossebuter et al. 1987)
Glyoxylate shunt							
	Sso1333	Isocitrate lyase	4.1.3.1	2224	0.30 ± 0.07[b]	NF	(Uhrigshardt, Walden et al. 2002)
	Sso1334	Malate synthase	2.3.3.9	2225	1.11 ± 0.47	1.18 ± 0.04	
C3/C4 interconversions							
	Sso2869	Malic enzyme	1.1.1.38	0281	1.05 ± 0.24	1.92 ± 0.15	(Bartolucci, Rella et al. 1987)
	Sso2537	Phosphoenolpyruvate carboxykinase	4.1.1.32	1274	1.42 ± 0.42	NF	(Fukuda, Fukui et al. 2004)
	Sso2256	Phosphoenolpyruvate carboxylase	4.1.1.31	1892	0.83 ± 0.18	0.88 ± 0.17	(Sako, Takai et al. 1996; Ettema, Makarova et al. 2004)

NF: not found, NS: no significant signal.

a relative abundance ratio with standard deviation Yeast extract + Tryptone grown cells / Glucose grown cells (YT/G)

b Probability value (p) smaller than 0.05.

c enzyme complex has broad substrate specificity for aldehydes

d exhibits pyruvate, 2-oxoglutarate and 2-oxobutyrate oxidoreductase activity

Identification of the Missing Links in Prokaryotic Pentose Oxidation Pathways: Evidence for enzyme recruitment

Stan J. J. Brouns, Jasper Walther, Ambrosius P. L. Snijders, Harmen J. G. van de Werken, Hanneke L. D. M. Willemen, Petra Worm, Marjon G. J. de Vos, Anders Andersson, Magnus Lundgren, Hortense F. M. Mazon, Robert H. H. van den Heuvel, Peter Nilsson, Laurent Salmon, Willem M. de Vos, Phillip C. Wright, Rolf Bernander and John van der Oost

J Biol Chem. 2006 Jul; 281 (37): 27378-88

Abstract

The pentose metabolism of Archaea is largely unknown. Here, we have employed an integrated genomics approach including DNA microarray and proteomics analyses to elucidate the catabolic pathway for D-arabinose in *Sulfolobus solfataricus*. During growth on this sugar, a small set of genes appeared to be differentially expressed compared with growth on D-glucose. These genes were heterologously overexpressed in *Escherichia coli*, and the recombinant proteins were purified and biochemically studied. This showed that D-arabinose is oxidized to 2-oxoglutarate by the consecutive action of a number of previously uncharacterized enzymes, including a D-arabinose dehydrogenase, a D-arabinonate dehydratase, a novel 2-keto-3-deoxy-D-arabinonate dehydratase, and a 2,5-dioxopentanoate dehydrogenase. Promoter analysis of these genes revealed a palindromic sequence upstream of the TATA box, which is likely to be involved in their concerted transcriptional control. Integration of the obtained biochemical data with genomic context analysis strongly suggests the occurrence of pentose oxidation pathways in both Archaea and Bacteria, and predicts the involvement of additional enzyme components. Moreover, it revealed striking genetic similarities between the catabolic pathways for pentoses, hexaric acids, and hydroxyproline degradation, which support the theory of metabolic pathway genesis by enzyme recruitment.

Introduction

Pentose sugars are a ubiquitous class of carbohydrates with diverse biological functions. Ribose and deoxyribose are major constituents of nucleic acids, whereas arabinose and xylose are building blocks of several plant cell wall polysaccharides. Many prokaryotes, as well as yeasts and fungi, are able to degrade these polysaccharides, and use the released five-carbon sugars as a sole carbon and energy source. At present, three main catabolic pathways have been described for pentoses. The first is present in Bacteria and uses isomerases, kinases, and epimerases to convert D- and L-arabinose (Ara) and D-xylose (Xyl) into D-xylulose 5-phosphate (Fig. 1A), which is further metabolized by the enzymes of the phosphoketolase or pentose phosphate pathway. The genes encoding the pentose-converting enzymes are often located in gene clusters in bacterial genomes, for example, the araBAD operon for L-Ara (1), the xylAB operon for D-Xyl (Trauger, Kalisak et al.), and the darK-fucPIK gene cluster for D-Ara (3). The second catabolic pathway for pentoses converts D-Xyl into D-xylulose 5-phosphate as well, but the conversions are catalyzed by reductases and dehydrogenases instead of isomerases and epimerases (Fig. 1B). This pathway is commonly found in yeasts, fungi, mammals, and plants, but also in some bacteria (4-6). In a third pathway, pentoses such as L-Ara, D-Xyl, D-ribose, and D-Ara are metabolized non-phosphorylatively to either 2-oxoglutarate (2-OG) or to pyruvate and glycol-aldehyde (Fig. 1C). The conversion to 2-OG, which is a tricarboxylic acid cycle intermediate, proceeds via the subsequent action of a pentose dehydrogenase, a pentonolactonase, a pentonic acid dehydratase, a 2-keto-3-deoxypentonic acid dehydratase, and a 2,5-dioxopentanoate dehydrogenase. This metabolic pathway has been reported in several aerobic bacteria, such as strains of Pseudomonas (7-9), Rhizobium (10, 11), Azospirillum (12), and Herbaspirillum (13). Alternatively, some Pseudomonas and Bradyrhizobium species have been reported to cleave the 2-ke- to-3-deoxypentonic acid with an aldolase to yield pyruvate and glycol-aldehyde (14-16). Despite the fact that these oxidative pathway variants have been known for more than five decades, surprisingly, the majority of the responsible enzymes and genes remain unidentified.

Sulfolobus spp. are obligatory aerobic Crenarchaea that are commonly found in acidic geothermal springs. Among the Archaea, this genus is well known for its broad saccharolytic capacity, which is reflected in their ability to utilize several pentoses and hexoses, as well as oligosaccharides and polysaccharides as a

sole carbon and energy source (17). Although the catabolism of hexoses is well studied (reviewed in Ref. 18), the pathways for pentose degradation have neither been established in *Sulfolobus solfataricus*, nor in any other member of the Archaea (19).

Experimental procedures

All chemicals were of analytical grade and purchased from Sigma, unless stated otherwise. Oligonucleotide primers were obtained from MWG Biotech AG (Ebersberg, Germany).

Growth of Sulfolobus Species

S. solfataricus P2 (DSM1617) was grown in media containing either 3 g/liter D-Ara or D-Glu as previously described (20).

Transcriptomics

Whole genome DNA microarrays containing gene-specific tags representing >90% of the *S. solfataricus* P2 genes (21) were used for global transcript profiling of cultures grown on D-Ara as compared with D-Glu. Total RNA extraction, cDNA synthesis and labeling, hybridization, and scanning were performed as previously described, as were data filtration, normalization, and statistical evaluation (22, 23).

Quantitative Proteomics

The proteome of *S. solfataricus* P2 was studied with a combination of two-dimensional gel electrophoresis, 15N metabolic labeling, and tandem mass spectrometry as previously described (24, 25). Two separate growth experiments were set up: 1) *S. solfataricus* with D-Ara as the carbon source and (14NH4)2SO4 as the nitrogen source; and 2) *S. solfataricus* with D-Glu as the carbon source and (15NH4)2SO4 as the nitrogen source. Next, the 14N and 15N cultures were mixed in equal amounts on the basis of optical density (A530) measurements, proteins were extracted and separated by two-dimensional gel electrophoresis. For the localization of proteins, a previously described two-dimensional gel electrophoresis reference map was used (23). Spots were excised from the gel, and peptides were

quantified on the basis of their relative intensity in the time of flight mass spectrum, according to established methods (23).

Synthesis of Organic Compounds

D-Arabinonate was synthesized from D-Ara as previously described (26). The aldehyde 2,5-dioxopentanoate was synthesized from 1,4-dibromobutane according to reported procedures (27-29).

Gene Cloning and Protein Overexpression

The genes araDH (Sso1300), araD (Sso3124), kdaD (Sso3118), and dopDH (Sso3117) were amplified by PCR from genomic DNA using Pfu TURBO polymerase (Stratagene) and cloned in expression vector pET24d (Novagen) (supplemental materials Table 1). The resulting plasmids were harvested from *Escherichia coli* HB101 transformants by Miniprep (Qiagen), sequenced by Westburg genomics (Leusden, Netherlands), and transformed to *E. coli* expression strain BL21(DE3) containing the tRNA accessory plasmid pRIL (Stratagene).

All proteins were produced according to standard procedures in four 1-liter shaker flasks containing LB medium, but with some exceptions. When the culture A600 reached 0.5, the cultures were cold-shocked by placing them on ice for 30 min to induce host chaperones (20). After that, the expression was started by adding 0.5 mm isopropyl β-D-thiogalactopyranoside, and the cultures were incubated for 12-16 h at 37 °C after which they were spun down (10 min, 5000 × g, 4 °C). At the time of induction, the arabinose dehydrogenase (AraDH) and AraD overexpression cultures were supplemented with 0.25 mm ZnSO4 (30) and 20 mm $MgCl_2$, respectively.

Protein Purification

Pelleted *E. coli* and *S. solfataricus* cells were resuspended in buffer and disrupted by sonication at 0 °C. Afterward, insoluble cell material was spun down (30 min, 26,500 × g, 4 °C) and the *E. coli* supernatants were subjected to heat treatment for 30 min at 75 °C. Denatured proteins were removed by centrifugation (30 min, 26,500 × g, 4 °C) yielding the heat-stable cell-free extract (HSCFE).

AraDH—HSCFE in 20 mm Tris-HCl (pH 7.5) supplemented with 50 mm NaCl was applied to a 20-ml Matrex Red A affinity column (Amicon). After washing the bound protein with 2 column volumes of buffer, the recombinant protein was eluted by a linear gradient of 2 m NaCl.

AraD—HSCFE in 50 mm HEPES-KOH (pH 8.0) supplemented with 50 mm NaCl was applied to a 70-ml Q-Sepharose Fast Flow (Amersham Biosciences) anion exchange column, and eluted in a 2 m NaCl gradient. Fractions containing the recombinant protein were pooled, concentrated with a 30-kDa cut-off filter (Vivaspin), and purified by size exclusion chromatography using a Superdex 200 HR 10/30 column (Amersham Biosciences) and 50 mm HEPES-KOH buffer (pH 8.0) supplemented with 100 mm NaCl as an eluent.

2-Keto-3-deoxy-D-arabinonate Dehydratase (KdaD)—HSCFE in 25 mm NaPi buffer (pH 6.8) was applied to a 70-ml Q-Sepharose Fast Flow (Amersham Biosciences) anion exchange column. Flow-through fractions containing KdaD were collected, loaded onto a 46-ml Bio-Gel hydroxyapatite column (Bio-Rad), and eluted by a linear gradient of 0.5 m NaPi buffer (pH 6.8). Fractions containing the recombinant proteins were pooled, and dialyzed overnight in 50 mm HEPES-KOH (pH 8.0) supplemented with 0.5 mm dithiothreitol (DTT).

2,5-Dioxopentanoate Dehydrogenase (DopDH)—HSCFE in 20 mm HEPES-KOH (pH 8.0) supplemented with 200 mm NaCl and 7.5 mm DTT was purified by affinity chromatography, as described for AraDH. Fractions containing the protein were pooled, concentrated using a 30-kDa cut-off membrane (Vivaspin), and purified by size exclusion chromatography as described for AraD.

Enzyme Assays

Unless stated otherwise, all enzymatic assays were performed in degassed 100 mm HEPES-KOH buffer (pH 7.5) at 70 °C. The optimal pH of catalysis was determined using a 30 mm citratephosphate-glycine buffer system that was adjusted in the range of pH 3-11 at 70 °C. Thermal inactivation assays were performed by incubating 50 µg/ml of enzyme at 70, 80, 85, and 90 °C and drawing aliquots at regular intervals during 2 h followed by a standard activity assay.

Dehydrogenase Assays

Sugar dehydrogenase activity was determined on a Hitachi U-1500 spectrophotometer in a continuous assay using 10 mm D- and L-arabinose, D- and L-xylose, D-ribose, D-lyxose, D- and L-fucose, D- and L-galactose, D-mannose, and D-glucose as a substrate, and 0.4 mm NAD^+ or $NADP^+$ as a cofactor. Aldehyde dehydrogenase reactions were performed using 5 mm 2,5-dioxopentanoate, glycolaldehyde, dL-glyceraldehyde, acetaldehyde, and propionaldehyde in the

presence of 10 mm DTT. Initial enzymatic activity rates were obtained from the increase in absorption at 340 nm (A_{340}), and calculated using a molar extinction coefficient of 6.22 mm^{-1} cm^{-1}.

Dehydratase Assay

Standard reactions were performed using 10 mm potassium D-arabinonate in the presence of 1 mm $MgCl_2$. The formation of 2-keto-3-deoxy-acid reaction products was determined with the thiobarbiturate assay at 549 nm using a molar extinction coefficient of 67.8 mm^{-1} cm^{-1} (31, 32). The effect of different divalent cations on enzymatic activity was investigated by a pre-treatment of the enzyme with 1 mm EDTA for 20 min at 70 °C, followed by a standard assay in the presence of 2 mm divalent metal ions.

Formation of 2-Oxoglutarate and Pyruvate

Enzyme reactions were performed with cell-free extract (CFE) from *S. solfataricus* cultures grown on either D-Ara or D-Glu, which were harvested at mid-exponential phase. The reaction was started by adding 25 µl of 3.5 mg/ml CFE to a mixture containing 10 mm potassium D-arabinonate, 1 mm $MgCl_2$, and either 0.4 mm NAD^+ or $NADP^+$. After an incubation of 2 h at 75°C, the reactions were stopped by placing the tubes on ice. Identical reactions were set up in which the CFE was replaced by the purified enzymes AraD (4.2 µg), KdaD (13.4 µg), and DopDH (3.8 µg). The amount of 2-oxoglutarate in these mixtures was then determined by the reductive amination of 2-oxoglutarate to L-glutamate using purified recombinant Pyrococcus furiosus glutamate dehydrogenase at 60 °C (33). The detection reaction was started by the addition of 5 units of glutamate dehydrogenase to a sample that was supplemented with 10 mm NH4Cl and 0.12 mm NADPH. The formation of pyruvate was determined at 30 °C using 4 units of chicken heart lactate dehydrogenase and 0.1 mm NADH. The conversion of 2-oxoglutarate or pyruvate was continuously monitored on a Hitachi U-1500 spectrophotometer by following the decrease in A340 until substrate depletion occurred. Changes in concentrations of NAD(P)H were calculated as described above.

Determination of the Protein Oligomeric State

The oligomerization state of AraDH, AraD, KdaD, and DopDH was determined by nanoflow electrospray ionization mass spectrometry. For this, the protein was concentrated in the range of 5-15 µm and the buffer was exchanged to 50 or 200 mm ammonium acetate (pH 6.7 or 7.5) by using an Ultrafree 0.5-ml centrifugal filter device with a 5-kDa cut-off (Millipore). Protein samples were introduced into the nanoflow electrospray ionization source of a Micromass LCT mass spectrometer (Waters), which was modified for high mass operation and set in positive ion mode. Mass spectrometry experiments were performed under conditions of increased pressure in the interface region between the sample and extraction cone of 8 mbar by reducing the pumping capacity of the rotary pump (34, 35). Capillary and sample cone voltages were optimized for the different proteins and were in the range of 1.4-1.6 kV and 75-150 V, respectively.

Bioinformatic Analyses

Upstream sequences of the differentially expressed genes were extracted between -200 and +50 nucleotides relative to the open reading frame translation start site. These sequences were analyzed using the Gibbs Recursive Sampler algorithm (36). Possible sequence motifs were checked against all upstream sequences and the complete genome of *S. solfataricus*. A diagram of the sequence motif was created using the WebLogo server.

Protein sequences were retrieved from the National Center for Biotechnology Information (NCBI) and analyzed using PSI-BLAST on the non-redundant data base, and RPS-BLAST on the conserved domain data base. Multiple sequence alignments were built using either ClustalX or TCoffee software. Gene neighborhood analyses were performed using various webserver tools: STRING at the EMBL, Gene Ortholog Neighborhoods at the Integrated Microbial Genomes server of the Joint Genome Institute, and pinned regions at the ERGO bioinformatics suite.

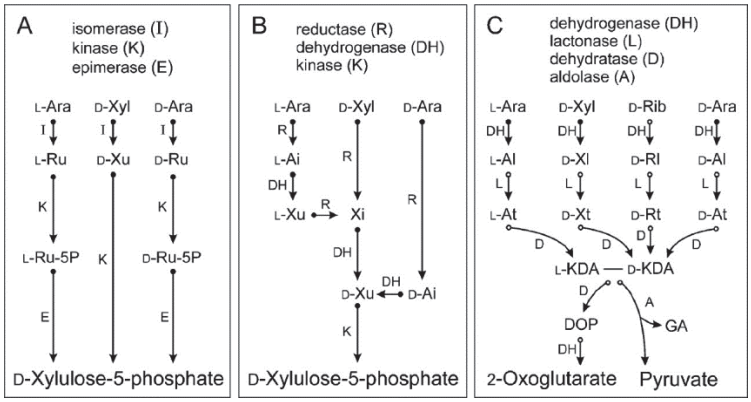

Figure 1: Schematic representation of three types of pentose degrading pathways (A, B, and C).Arrows with an open or closed arrowtail represent enzymatic steps that are performed by unknown proteins or known proteins, respectively. Abbreviations: Ara, arabinose; Xyl, xylose; Rib, ribose; Ru, ribulose; Xu, xylulose; Ai, arabinitol; Xi, xylitol; Al, arabinonolactone; Xl, xylonolactone; Rl, ribonolactone; At, arabinonate; Xt, xylonate; Rt, ribonate; KDA, 2-keto-3-deoxy-arabinonate (also called 2-oxo-4,5-dihydroxypentanoate); DOP, 2,5-dioxopentanoate (also called 2-oxoglutarate semialdehyde); GA, glycolaldehyde.

Transcriptomics—The global transcriptional response of *S. solfataricus* growing exponentially on D-Ara or D-Glu was determined by DNA microarray analysis. The transcriptome comparison between both growth conditions showed that a small set of genes was differentially expressed 3-fold or more (Table 2). The highly expressed genes under D-Ara conditions included all four subunits of the Ara ABC transporter (Sso3066-3069) (39), a putative sugar permease for D-Ara (Sso2718), five of six subunits of the SoxM quinol oxidase complex (Sso2968-2973) (40), and five metabolic genes with general function predictions only (Sso1300, Sso3124, Sso3117, Sso3118, and Sso1303). The differential expression of the gene for the remaining SoxM subunit, i.e. the sulfocyanin subunit SoxE (Sso2972), was just below the threshold level (supplemental materials Table 2). Whereas the expression of the ABC-type transport system genes had been shown to be induced in Ara media previously (39, 41), the differential expression of the SoxM gene cluster was not anticipated.

Results and discussion

S. solfataricus is a model archaeon for studying metabolism and information processing systems, such as transcription, translation, and DNA replication (37, 38). Several halophilic and thermophilic Archaea have been reported to assimilate pentose sugars, but neither the catabolic pathways for these 5-carbon sugars nor the majority of its enzymes are known (17, 19). To close this knowledge gap, we have studied the growth of *S. solfataricus* on the pentose D-Ara using a multidisciplinary genomics approach, and compared the results to growth on the hexose D-Glu. Both culture media supported growth to cell densities of ~2 × 109 cells/ml (A600 2.5) with similar doubling times of around 6 h.

Several enzyme activity assays were performed with CFEs from both cultures to establish a mode of D-Ara degradation (Fig. 1). A 12.3-fold higher $NADP^+$-dependent D-Ara dehydrogenase activity (45.5 milliunits/mg) was detected in D-Ara CFE (Table 1), which indicated the presence of an inducible D-Ara dehydrogenase. D-Ara reductase, D-arabinitol dehydrogenase, and D-Ara isomerase activity were not detected. Activity assays using D-arabinonate indicated that D-Ara CFE contained a 13.9-fold higher D-arabinonate dehydratase activity (7.4 milliunits/mg) than D-Glu CFE (Table 1). Moreover, the multistep conversion of D-arabinonate to 2-OG could readily be demonstrated with D-Ara CFE in the presence of $NADP^+$ (Fig. 2). The formation of pyruvate as one of the products from D-arabinonate was not observed, whereas control reactions with both CFEs and D-gluconate as a substrate did yield pyruvate (data not shown), indicating that the enzymes of the Entner-Doudoroff pathway were operative. In the final step of the pathway, D-Ara CFE contained a 3.6-fold higher activity (255 milliunits/mg) toward the aldehyde 2,5-dioxopentanoate (DOP) using $NADP^+$ as a cofactor. The data suggest that *S. solfataricus* employs an inducible enzyme set that converts D-Ara into the tricarboxylic acid cycle intermediate 2-OG via the pentose oxidation pathway (Fig. 1C).

The genes that were up-regulated under D-Glu conditions encode seven uncharacterized proteins (Sso3073, Sso3089, Sso3104, Sso1312, Sso2884, Sso3085, and Sso3100), the SoxB subunit of the SoxABCD quinol oxidase complex (Sso2657) (42), and a glutamate dehydrogenase (Sso2044) (43) (Table 2). The Glu ABC transporter was not differentially expressed, confirming previous observations (41). The difference in gene expression of subunits SoxA (Sso2658), SoxC (Sso2656), and SoxD (Sso10828) was just below the threshold level (supplemental materials Table 2). Next to the SoxABCD genes, a small gene cluster containing the Rieske iron-sulfur cluster protein SoxL-1 (Sso2660) and Sso2661 to Sso2663 appeared to be expressed with a 2-3-fold difference (supplemental materials Table 2). It thus appears that under D-Glu conditions, the Sox-ABCD quinol oxidase complex is preferentially used, whereas under D-Ara conditions the SoxM-mediated terminal quinol oxidation is favored. Differential use of both oxidase complexes was recently also found in Metallosphaera sedula. Here the SoxABCD genes were expressed at high levels during growth on sulfur, whereas heterotrophic growth on yeast extract induced the production of the SoxM complex (44). Because the aeration and cell density of the D-Ara and D-Glu cultures was similar, the

trigger for the differential expression of the two oxidase complexes in *S. solfataricus* is currently unknown.

Table 1: Properties of the D-Ara degrading enzymes of *S. solfataricus*

	AraDH	AraD	KdaD	DopDH
Description	D-arabinose 1-dehydrogenase	D-arabinonate dehydratase	2-keto-3-deoxy-D-arabinonate dehydratase	2,5-dioxopentanoate dehydrogenase
EC number	1.1.1.117	4.2.1.5	4.2.1.-	1.2.1.26
Locus ID	Sso1300	Sso3124	Sso3118	Sso3117
Uniprot ID	Q97YM2	Q97U96	Q97UA0	Q97UA1
Genbank ID	15898142	15899830	15899826	15899825
COG	1064	4948	3970	1012
PFAM	00107	01188	01557	00171
Sp. activity in *S.s* extracts mU/mg (fold A/G)	45.5 (12.3)	7.4 (13.9)	ND	255 (3.6)
ARA-box present	yes	yes	yes	yes
Subunit size (kDa)	37.3	42.4	33.1	52.3
Oligomerization	Tetramer	Octamer	Tetramer	Tetramer
Substrate specificity, turnover number (s^{-1})	D-arabinose (23.8) L-fucose (26.8) D-ribose (17.7) L-galactose (17.7)	D-arabinonate (1.8)	ND	2,5-dioxopentanoate(8.6) Glycolaldehyde(5.3) DL-glyceraldehyde(4.8)
Cofactor	$NADP^+$, Zn^{2+}	Mg^{2+}	ND	$NADP^+$
App. pH optimum (>50% activity)	8.2 (7.3-9.3)	6.7 (5.2-10.2)	ND	7.8 (6.7-8.2)
Apparent T-opt (°C) (>50% activity)	91 (74->100)	85 (75-92)	ND	ND
Thermal inactivation half-life time	42 min at 85°C 26 min at 90°C	18 min at 85°C <10 min at 90°C	ND	ND

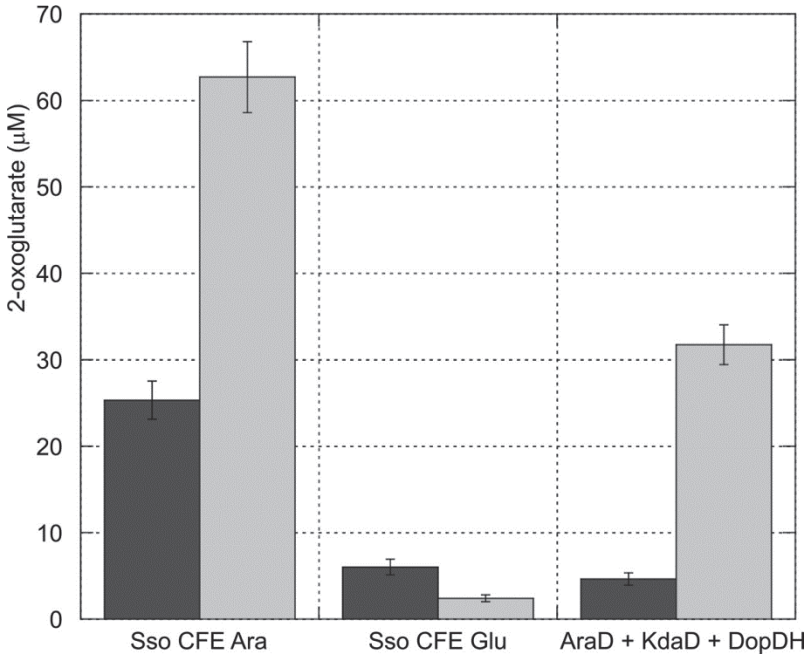

Figure 2: The formation of 2-oxoglutarate from D-arabinonate by extracts of S. solfataricus and by the recombinant enzymes AraD, KdaD and DopDH, in the presence of 0.4 mM NAD^+ (dark gray bars) or $NADP^+$ (light gray bars).

Proteomics—Protein expression in the soluble proteomes of D-Ara and D-Glu grown *S. solfataricus* cells was compared using a combination of two-dimensional gel electrophoresis, stable isotope labeling, and tandem mass spectrometry. By employing this strategy, five proteins were found with more than a 20-fold difference in expression level (supplemental materials Fig. 1, B-F), including the Ara-binding protein from the Ara ABC transporter (AraS, Sso3066) (39), Sso1300, Sso3124, Sso3118, and Sso3117 (Table 2). Interestingly, the difference in expression level of these genes at the protein level appeared to be more pronounced than at the mRNA level, which ranged from 3.4- to 16-fold. Three other proteins were also produced in higher amounts during growth on D-Ara, albeit only up to a 3-fold difference (Table 2). These were the isocitrate lyase (Sso1333) (45), the phosphoglycerate kinase (Sso0527) (46), and the malic enzyme (Sso2869) (47).

Table 2: Differentially expressed genes (>3-fold different)

Locus name	Microarray [2]log fold(A/G) ± SD (q-value[a])	Proteomics fold(A/G) ± SD	Description	Reference
		High expression on D-Ara		
Sso3066	4.02 ± 0.58 (1.1)	>20	Arabinose ABC transporter, arabinose binding protein (Schwaiger, Schwarz et al.)	Elferink, 2001 #39
Sso3068	3.71 ± 0.97 (1.1)		Arabinose ABC transporter, permease	Elferink, 2001 #39
Sso1300	3.64 ± 0.95 (1.1)	>20	Alcohol dehydrogenase IV[b] D-arabinose 1-dehydrogenase (AraDH)	This study
Sso3067	3.37 ± 0.49 (1.1)		Arabinose ABC transporter, permease	Elferink, 2001 #39
Sso3069	2.97 ± 0.18 (2.5)		Arabinose ABC transporter, ATP binding protein	Elferink, 2001 #39
Sso2968	2.56 ± 1.30 (1.1)		Quinol oxidase subunit (SoxM complex), SoxI	Komorowski, 2002 #40
Sso3124	2.44 ± 1.13 (1.1)	>20	Mandelate racemace / muconate lactonizing enzyme[b] D-arabinonate dehydratase (AraD)	This study
Sso3117	2.39 ± 0.62 (1.1)	>20	Aldehyde dehydrogenase[b] 2,5-dioxopentanoate dehydrogenase (DopDH)	This study
Sso2971	2.24 ± 1.17 (1.1)		Quinol oxidase subunit (SoxM complex), SoxF (Rieske Fe-S protein)	Komorowski, 2002 #40
Sso2973	2.10 ± 1.48 (2.6)		Quinol oxidase subunit (SoxM complex), SoxM (I + III)	Komorowski, 2002 #40
Sso2970	2.09 ± 1.53 (2.6)		Quinol oxidase subunit (SoxM complex), SoxG (cytochrome a_{587})	Komorowski, 2002 #40
Sso3046	1.89 ± 0.87 (1.1)		put. ABC sugar transporter, ATP binding protein	
Sso2969	1.86 ± 1.09 (1.1)		Quinol oxidase (SoxM complex), SoxH subunit (II)	Komorowski, 2002 #40
Sso3118	1.78 ± 0.45 (1.1)	>20	Conserved hypothetical protein[b] 2-keto-3-deoxy-D-arabinonate dehydratase (KdaD)	This study
Sso1303	1.77 ± 0.57 (1.1)		put. pentonic acid dehydratase	
Sso1333	NDE	3.48 ± 0.62	Isocitrate lyase	Uhrigshardt, 2002 #45
Sso0527	NDE	3.45 ± 0.72	Phosphoglycerate kinase	Hess, 1995 #46
Sso2869	NDE	3.06 ± 1.06	Malic enzyme	Bartolucci, 1987 #47

		High expression on D-Glu	
Sso3073	-2.59 ± 0.71 (1.1)	put. Sugar permease	
Sso3089	-2.12 ± 1.12 (1.1)	Hypothetical protein	
Sso3104	-2.04 ± 0.31 (1.1)	Hypothetical protein	
Sso1312	-2.02 ± 0.52 (1.1)	put. Ring oxidation complex / phenylacetic acid degradation rel. protein	
Sso2884	-1.87 ± 0.37 (1.1)	put. 4-carboxymuconolactone decarboxylase	
Sso2657	-1.77 ± 0.87 (1.1)	Quinol oxidase (SoxABCD complex), cytochrome aa3 subunit (SoxB)	Gleissner, 1997 #42
Sso3085	-1.63 ± 0.86 (1.1)	Conserved hypothetical protein	
Sso3100	-1.60 ± 0.88 (1.1)	Hypothetical protein	
Sso2044	-1.60 ± 0.48 (1.1)	L-Glutamate dehydrogenase	

Promoter Motif Analysis—The promoters of the differentially expressed genes were analyzed for the occurrence of DNA sequence motifs that could play a role as cis-acting elements in the coordinated transcriptional control of these genes. The analysis indeed revealed the presence of a palindromic motif (consensus: AACATGTT) in the promoters of Sso3066 (Schwaiger, Schwarz et al.), Sso1300, Sso3124, Sso3118, and Sso3117 genes (Fig. 3). This motif was designated the ARA box and it was always located upstream of the predicted TATA box with a separation of 10 bases. A conserved transcription factor B recognition element appeared to be absent from the interspaced sequence between both boxes. Additional copies of the ARA box were identified further upstream of both Sso3066 and Sso1300. Although primer extension analysis was only performed for the araS gene (41), the promoter architecture suggests that the transcript leader of Sso1300, Sso3124, Sso3118, and Sso3117 will either be very short, or absent. This is in good agreement with the fact that a large proportion of the *S. solfataricus* genes is predicted to be transcribed without a leader (48). The inducibility of the araS promoter has recently been exploited in viral expression vectors that enable recombinant protein production in *S. solfataricus* (49).

Biochemical Characterization of the D-Ara-induced Proteins—The genes that were differentially expressed and contained an Ara box in their promoter were selected and cloned in an *E. coli* expression vector. The resulting proteins were overproduced, purified, and characterized to investigate their role in the metabolism of D-Ara.

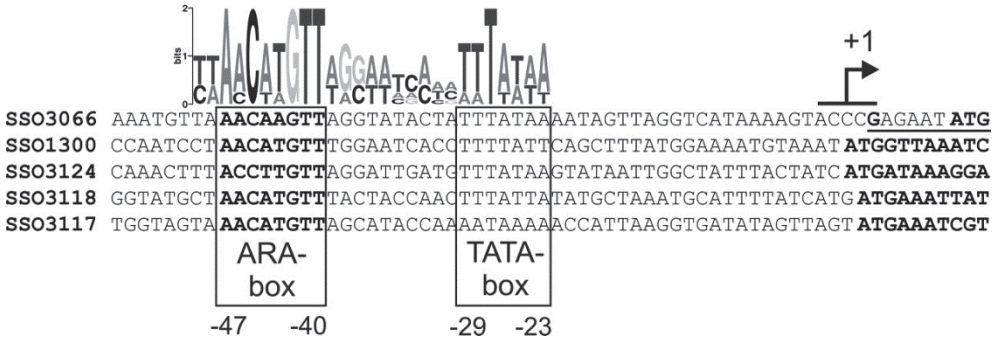

Figure 3: Putative cis-regulatory element (ARA box) and TATA box in upstream sequences of the D-Ara-induced genes. The predicted transcription start site is indicated by +1. Transcripts are underlined (41). Coding sequences are in bold. Additional ARA boxes were found for Sso3066 at -90 to -83 and Sso1300 at -235 to -228 relative to the transcription start sites. Note: a single ARA box is present in the intergenic region between the divergently oriented genes Sso3118 and Sso3117.

AraDH—The putative zinc-containing, medium-chain alcohol dehydrogenase encoded by Sso1300 was efficiently produced and purified using a single step of affinity chromatography (Fig. 4). The enzyme was most active on L-fucose (6-deoxy-L-galactose) (kcat 26.8 s^{-1}), followed by D-Ara (kcat 23.8 s-1), using preferentially $NADP^+$ (Km 0.04 ± 0.01 mm) over NAD^+ (Km 1.25 ± 0.45 mm) as a cofactor. This enzyme was thus likely to account for the elevated D-Ara dehydrogenase activities in *S. solfataricus* CFE. AraDH could also oxidize L-galactose and the D-Ara C2-epimer D-ribose with similar rates (kcat 17.7 s-1) (Table 1). Enzyme activity toward other sugars remained below 7% of the highest activity. Similar substrate specificities and affinities have been found previously for mammalian and bacterial L-fucose or D-Ara dehydrogenases, although these enzymes prefer NAD^+ as a cofactor (50, 51). AraDH was more than 50% active in a relatively narrow pH range from 7.3 to 9.3, with optimal catalysis proceeding at pH 8.2. The thermophilic nature of the enzyme is apparent from its optimal catalytic temperature of 91 °C. The enzyme maintained a half-life of 42 and 26 min at 85 and 90 °C, respectively, indicating that the enzyme is thermostable at physiological growth temperatures of *S. solfataricus*. Native mass spectrometry experiments showed that the intact recombinant AraDH has a molecular mass of 149,700 ± 24 Da. Comparing these data with the expected mass on the basis of the primary sequence (37,291 Da) clearly showed that the protein has a tetrameric structure and contains two zinc atoms per monomer. This is in good agreement with the

tetrameric structure that has been reported for another alcohol dehydrogenase from *S. solfataricus* (Sso2536), which has a 33% identical protein sequence (30). This dehydrogenase, however, prefers aromatic or aliphatic alcohols as a substrate, and NAD^+ over $NADP^+$ as a cofactor. A structural study of AraDH is currently ongoing to explain the observed differences in substrate and cofactor selectivity.

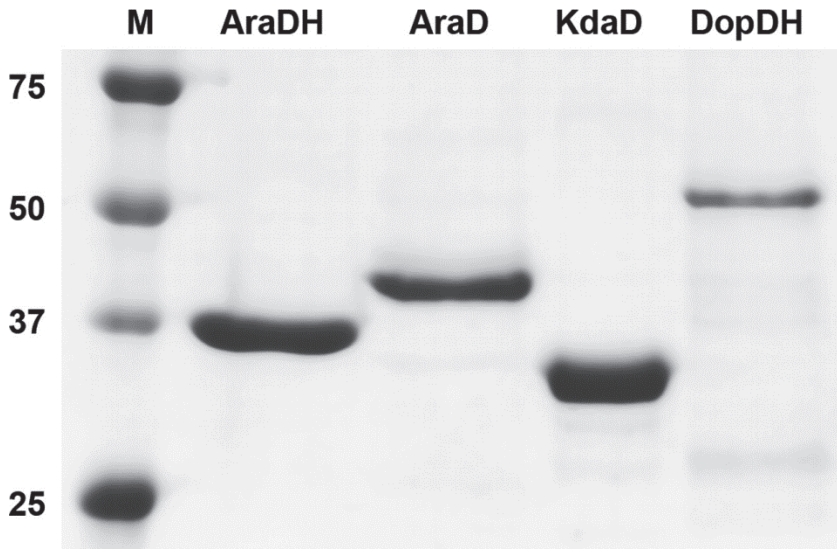

Figure 4: A digital photograph is shown of a Coomassie Blue-stained 8% SDS-PAGE gel that was loaded with purified recombinantly produced enzymes from the D-Ara oxidation pathway of *S. solfataricus*. Marker sizes are indicated in kDa.

Arabinonate Dehydratase (AraD)—The protein encoded by gene Sso3124 was originally annotated as a member of the mandelate racemase and muconate lactonizing enzyme family. This superfamily, which additionally comprises of aldonic acid dehydratases, is mechanistically related by their common ability to abstract α-protons from carboxylic acids (52). Production of the enzyme in *E. coli* yielded app. 10% soluble recombinant protein, which was purified using anion exchange and size exclusion chromatography (Fig. 4). The enzyme was shown to catalyze the strictly Mg^{2+}-dependent dehydration reaction of D-arabinonate to 2-keto-3-deoxy-D-arabinonate (KDA) (supplemental materials Fig. 2A). It is therefore conceivable that this enzyme is largely responsible for the increased levels of D-arabinonate dehydratase activity in *S. solfataricus* extracts. AraD displayed a

maximum turnover rate of 1.8 s-1 at a substrate concentration of 8 mm, whereas higher substrate concentrations imposed severe inhibitory effects on the enzyme (supplemental materials Fig. 2B). No activity was measured with D-gluconate up to 20 mm. More than 50% enzyme activity was observed in a broad pH range of 5.2 to 10.2 with an optimum at pH 6.7 (Table 1). The enzyme was optimally active at 85 °C during which it maintained a half-life time of 18 min. Native mass spectrometry revealed that the protein had a molecular mass of 340,654 ± 63 Da, which corresponds well to an octameric protein assembly (expected monomeric mass is 42,437 Da). The native D-gluconate dehydratase from *S. solfataricus* (GnaD, Sso3198), which has a 23% identical protein sequence, was found to be an octamer as well (32). Interestingly, AraD was only produced as an octamer when the media was supplemented with 20 mm Mg^{2+} during protein overexpression. Without this divalent cation, the recombinant protein was inactive and appeared to be monomeric. Sequence alignment analysis as well as three-dimensional modeling based on a Agrobacterium tumefaciens protein with unknown function (Atu3453, Protein Data Bank code 1RVK) showed that Asp-199, Glu-225, and Glu-251 are likely to be involved in binding the divalent metal ion, which is required to stabilize the enolic reaction intermediate (52).

KdaD—To investigate the possible role of Sso3118, the protein was overproduced in *E. coli*, and subsequently purified (Fig. 4). Surprisingly, although the predicted pI of the enzyme is 5.9, the vast majority of protein did not bind to the anion exchange column at a pH of 8. Moreover, the protein had a tendency to precipitate, which could be reversed and effectively prevented by the addition of 0.5 mm DTT to all buffers. Native mass spectrometry under reducing conditions revealed that the protein had a molecular mass of 132,850 ± 47 Da, which corresponds with a tetrameric quaternary structure (expected monomeric mass of 33,143 Da). The catalytic activity of the protein was investigated by performing indirect enzyme assays using AraD with D-arabinonate as a substrate. A 50% decrease in the yield of KDA was observed when both enzymes were co-incubated in the presence of Mg^{2+}, but this did not result in the formation of either 2-OG or pyruvate. Given the fact that D-arabinonate is converted to 2-OG in D-Ara CFE, this enzyme was anticipated to be responsible for the dehydration of D-KDA to the aldehyde DOP. However, due to the unavailability of D-KDA, it was not possible to show this in a direct enzyme assay. We therefore employed an indirect assay using

AraD, the putative D-KDA dehydratase (KdaD) and the predicted aldehyde dehydrogenase. The results of this assay are described under "DopDH."

According to the Clusters of Orthologous Groups of proteins classification system, the putative KDA dehydratase belongs to COG3970. The catalytic domain of these proteins resembles that of the eukaryal fumarylacetoacetate hydrolase; an enzyme that catalyzes the Mg^{2+}- or Ca^{2+}-dependent hydrolytic cleavage of fumarylacetoacetate to yield fumarate and acetoacetate as the final step of phenylalanine and tyrosine degradation (53). In humans, several mutations in the fumarylacetoacetate hydrolase gene will lead to hereditary tyrosinemia type I, which is mainly characterized by liver defects (54). Members of COG3970 are also homologous to the C-terminal decarboxylase domain of the bifunctional enzyme HpcE from *E. coli*, which in addition consists of an N-terminal isomerase domain (55). This enzyme is active in the homoprotocatechuate pathway of aromatic compounds and is responsible for the Mg^{2+}-dependent decarboxylation of 2-oxo-5-carboxy-hept-3-ene-1,7-dioic acid to 2-hydroxy-hepta-2,4-diene-1,7-dioic acid and its subsequent isomerization to 2-oxo-hept-3-ene-1,7-dioic acid (55). Although the function of these enzyme classes is rather diverse, their structures have revealed similarities in terms of a fully conserved metal-ion binding site and a relatively conserved active site architecture. Multiple sequence alignment analysis of KdaD indicated the presence of a metal binding site consisting of Glu-143, Glu-145, and Asp-164, which may implicate a metal dependent activity as well. Further structural and kinetic studies of KdaD are currently ongoing.

DopDH—The putative aldehyde dehydrogenase encoded by Sso3117 was overproduced in *E. coli*, which resulted in the formation of app. 5% soluble protein. This protein fraction was purified using affinity and size exclusion chromatography (Fig. 4). From native mass spectrometry experiments we could determine a molecular mass of 210,110 Da, which is in reasonable agreement with the expected mass of the tetramer on the basis of the primary sequence (52,290 Da). The measured mass may be somewhat higher due to the binding of small molecules to the protein oligomer. The determined oligomerization state corresponds to that of the closely related aldehyde dehydrogenase ALDH-T from Geobacillus stearothermophilus (56). The aldehyde dehydrogenase was tested for the activity toward different aldehydes and cofactors (Table 1). This indicated that the enzyme preferred $NADP^+$ over NAD^+, and that it oxidized several hydrophilic aldehydes with the highest activity toward DOP followed by glycolaldehyde and dL-

glyceraldehyde. More than 50% enzyme activity was observed in a pH range of 6.7-8.2, with an optimum at pH 7.8. The enzyme was also tested in conjunction with AraD and KdaD for the production of 2-OG or pyruvate. Similar to the activities in D-Ara CFE, these three enzymes were able to form 2-OG and not pyruvate, from D-arabinonate using preferably $NADP^+$ as a cofactor (Fig. 2). Omission of either the cofactor, AraD, KdaD, or DopDH prevented the formation of 2-OG, indicating that all components were essential for the enzymatic conversions, and that KdaD was most likely responsible for the dehydration of D-KDA to DOP.

Extensive kinetic characterization of DopDH proved to be rather complicated, because the enzyme lost nearly all its activity within 1 day after its purification, even in the presence of high concentrations of reducing agents, such as DTT or β-mercaptoethanol. This could be due to the fact that this class of enzymes contains a catalytic cysteine residue (in DopDH Cys-293), which can become irreversibly oxidized, leading to a total loss of enzymatic activity. A rapid inactivation was also observed with ALDH-T from G. stearothermophilus (56).

Central Carbohydrate Metabolism—Some central metabolic routes, such as the glycolysis, gluconeogenesis, and the tricarboxylic acid cycle have been studied extensively in *S. solfataricus*, *Sulfolobus acidocaldarius*, and other Archaea. The availability of their genome sequences (37, 57) as well as the genome sequence of *Sulfolobus tokodaii* (58), has recently allowed a reconstruction of the genes involved in these pathways (23). The effect of the introduction of excess 2-OG resulting from the D-Ara oxidative pathway led to the differential expression of only a few additional genes in these central carbon metabolic routes (Table 2; supplemental materials Fig. 3). The isocitrate lyase, the phosphoglycerate kinase, and the malic enzyme were up-regulated at the protein level under D-Ara conditions. The induction of the malic enzyme might indicate that the main proportion of 2-OG is converted to malate, which is then decarboxylated to pyruvate and acetyl-CoA, respectively, and is then fully oxidized to two molecules of CO2 in one round of the tricarboxylic acid cycle. Although this may seem energetically unfavorable, the net difference in yield between the full degradation of one molecule of D-Glu or D-Ara to CO2 is only one NADPH reduction equivalent in favor of D-Glu, because both degradation schemes lead to 6 reduced ferredoxins, 2 FADH2, 2 ATP, and 6 or 5 NADPH molecules, respectively. It is therefore not surprising that the growth rates under both conditions are similar. The phosphoglycerate kinase may be indicative of increased gluconeogenic activities

that are required under D-Ara conditions. The isocitrate lyase is normally operative in the glyoxylate shunt, but high production levels of the enzyme have also been observed during growth on L-glutamate compared with D-Glu (25). Oxidative deamination of L-glutamate leads to the formation of 2-OG as well, which may inhibit the isocitrate dehydrogenase activity leading to an accumulation of isocitrate. This could trigger the production of the isocitrate lyase, which can bypass this step without the loss of CO2.

Pentose Oxidation Gene Clusters—The comprehensive analysis of conserved gene clustering in multiple genome sequences is becoming an increasingly important tool to predict functionally or physically associated proteins in prokaryotic cells (reviewed in Ref. 59). Genomic context analysis of the genes involved in the D-Ara oxidative pathway of *S. solfataricus* showed that kdaD and dopDH gene orthologs are often located adjacent in prokaryotic genomes. This finding supports the proposed enzymatic functions of an aldehyde producing and an aldehyde oxidative activity. In addition, the analysis uncovered the presence of putative pentose oxidative gene clusters in the genomes of several aerobic proteobacteria, such as members of the genera Burkholderia, Rhizobium, Bradyrhizobium, Agrobacterium, and Pseudomonas. In some cases, the presence of such a gene cluster correlates well with the ability of the organism to assimilate pentoses and with enzymatic activities present in cell extracts (7-11), whereas in other cases biochemical data is not available. Nonetheless, a few of these characteristic gene clusters have been demonstrated genetically to be linked to pentose degradation. Combined with the findings in *S. solfataricus*, this allows the identification of additional enzymatic components in the pentose oxidation pathway and prediction of their enzymatic functions (Fig. 5A).

A putative operon of five genes was found in the genome of the oligotrophic α-proteobacterium Caulobacter crescentus, which was 2.8-11.6-fold up-regulated during growth on D-Xyl as compared with D-Glu (60). Reporter fusion constructs of the CC0823 promoter to the β-galactosidase gene (lacZ) from *E. coli* confirmed that this promoter is highly induced during growth on D-Xyl, and repressed on D-Glu or proteinaceous media (60, 61). Moreover, the disruption of the CC0823 gene prevented the C. crescentus from growth on D-Xyl as a single carbon source (61).

A second pentose degradation gene cluster involved in L-Ara uptake and utilization was found on chromosome II of the pathogenic β-proteobacterium Burkholderia thailandensis. This cluster consisting of nine genes was proposed to

be responsible for the L-Ara degradation to 2-OG (Fig. 5A) (62). Disruption of the araA, araC, araE, and araI genes led to an L-Ara negative phenotype. Reporter gene insertions showed that araC and araE gene expression was repressed during growth in D-Glu media, and induced in L-Ara media. The transfer of the gene cluster to the related bacterium B. pseudomallei enabled this organism to utilize L-Ara as a sole carbon source also (62). Interestingly, an L-Ara dehydrogenase with 80% sequence identity to AraE has recently been characterized from Azospirillum brasiliense (63); an organism that is known to degrade L-Ara to 2-OG (12). The flanking sequences of this gene revealed close homologs of the B. thailandensis araD and araE, which would indicate a similar gene cluster in A. brasiliense (63).

Apart from several bacteria, putative pentose oxidation clusters are also present in some Archaea. In the halophile Haloarcula marismortui, a gene cluster was found on chromosome I that seems to contain all of the necessary components for D-Xyl oxidation, including a gene that has been identified as a D-Xyl dehydrogenase (19) (Fig. 5A).

Components of the Pentose Oxidation Pathway—Careful inspection of the different pentose oxidation gene clusters shows that the gene encoding the final enzymatic step, from DOP to 2-OG, is fully conserved between the different pentose oxidation gene clusters. The remaining analogous enzymatic steps that convert D-Ara, D-Xyl, or L-Ara into DOP are performed by enzymes from distinct COGs (Clusters of Orthologous Groups of proteins) (64) (Fig. 5, A and B, pentose panels). Whereas some of this variation in enzyme use may simply be explained by substrate differences, other variations may be due to the individual adaptation of existing enzymes with similar reaction chemistry, such as the pentose dehydrogenases.

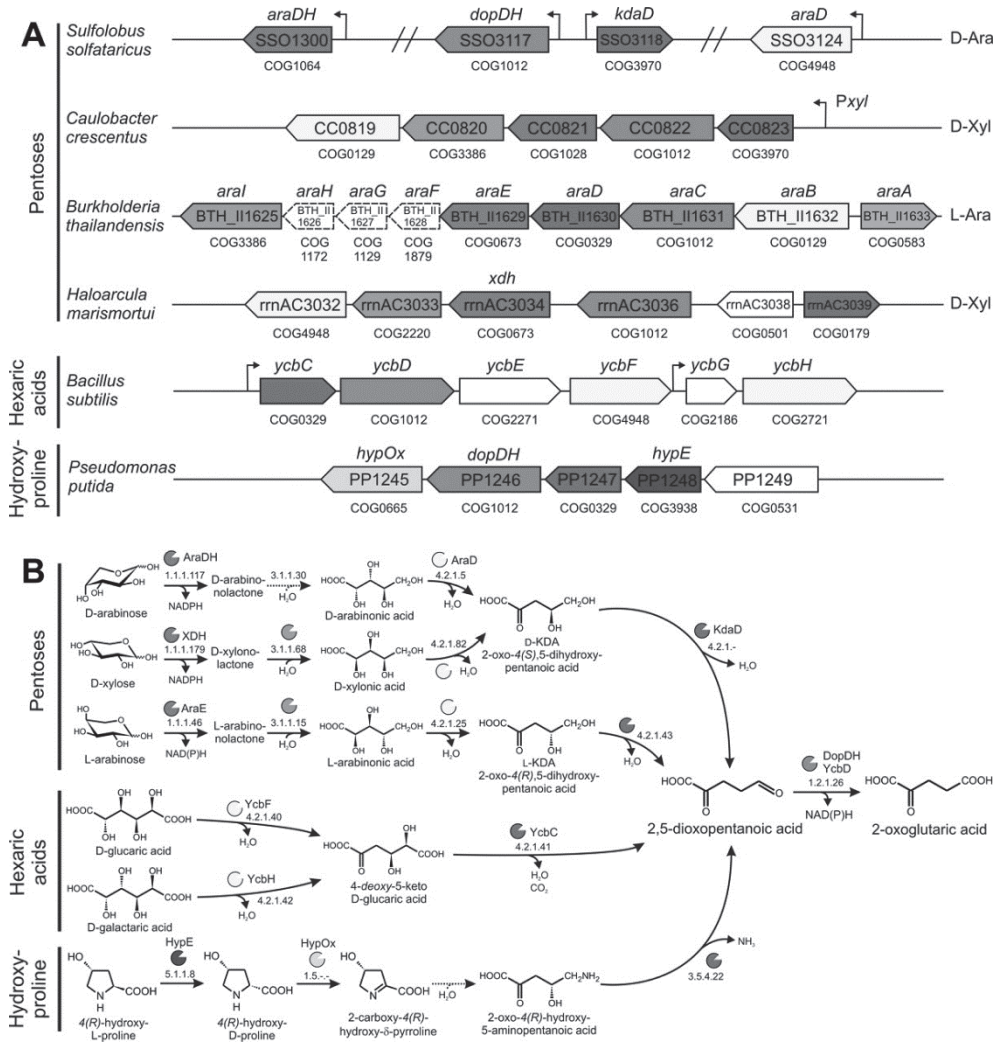

Figure 5: **A**, scheme of the organization of conserved gene clusters involved in the pentose, hexaric acid, and hydroxyproline degradation. Proposed analogous gene functions are indicated in the same color (green, pentose dehydrogenase; orange, pentonolactonase; yellow, aldonic acid dehydratase; red, 2-keto-3-deoxyaldonic acid dehydratase; blue, 2,5-dioxopentanoate dehydrogenase). Dashed genes are displayed smaller than their relative size. Protein family numbers are displayed below each gene according to Clusters of Orthologous Groups of proteins classification system (64). The genes indicated in white or gray encode the following putative functions: araA, transcriptional regulator; araF-araH, L-Ara ABC transporter (periplasmic L-Ara binding protein, ATP-binding protein, permease); rrnAC3038, heat shock protein X; ycbE, glucarate/galactarate permease; ycbG, transcriptional regulator; PP1249, hydroxyproline permease. **B**, schematic representation of the convergence of catabolic pathways for pentoses, hexaric acids (9, 71, 72, 78), and hydroxyproline (73-75) at the level of 2,5-dioxopentanoate. Enzymatic activities are indicated by their EC number. Dashed lines indicate proposed spontaneous reactions.

A striking difference between the set of enzymes responsible for D-Ara degradation in *S. solfataricus* on the one hand, and the predicted sets for D-Xyl degradation in C. crescentus and H. marismortui and L-Ara degradation in B. thailandensis on the other hand, is the apparent absence of an up-regulated lactonase in the hyperthermophile. This enzyme is responsible for the hydrolysis of the lactone, yielding the corresponding linear pentonic acid. Such ring opening reactions are reported to proceed spontaneously at ambient temperatures, albeit at slow rates (65). Overexpressing a lactonase may therefore be advantageous at mesophilic growth temperatures, whereas at 80 °C the spontaneous reaction may well proceed rapidly enough not to be rate-limiting. The pentose oxidation gene clusters seem to be predominated by lactonases of COG3386, which are often annotated as "senescence marker protein-30 family proteins". The genome of *S. solfataricus* contains two of these genes (Sso2705 and Sso3041), but they were not differentially expressed, indicating that they are either not involved or that their basal expression level is sufficient for arabinonolactone hydrolysis. The putative xylonolactonase from H. marismortui, however, is homologous to metal-dependent β-lactamases belonging to COG2220, which catalyze similar lactame-ring opening reactions (66).

Other non-orthologous enzyme components of the pentose oxidation pathway include the pentonic acid dehydratases. Whereas the D-arabinonate dehydratase from *S. solfataricus* belongs to COG4948, the same function seems to be performed by members of COG0129 that are commonly annotated as dihydroxyacid dehydratases (IlvD) or 6-phosphogluconate dehydratases (Edd) (67). A member of this family has recently been characterized from *S. solfataricus* (DHAD, Sso3107), which revealed a broad substrate specificity for aldonic acids (68). However, this gene was not differentially expressed according to the transcriptome or proteome analysis.

The 2-keto-3-deoxy-D-arabinonate dehydratase (COG3970), or a member of the homologous COG0179, appears to be present in D-Ara and D-Xyl degradation gene clusters. Interestingly, in several Burkholderia species, and in A. brasiliense, this gene is replaced by a member of the dihydrodipicinolate synthase family (COG0329, B.th araD). Members of this family catalyze either aldolase or dehydratase reactions via a Schiff base-dependent reaction mechanism by a strictly conserved lysine residue. Interestingly, a detailed study of an L-KDA dehydratase involved in the L-Ara metabolism of P. saccharophila was reported a few decades

ago, but unfortunately, neither the N-terminal sequence of the protein nor the gene sequence was determined (69, 70). The authors found that this enzyme was enantioselective for L-KDA (2-oxo-4(R), 5-dihydroxypentanoate), and that the reaction proceeds via a Schiff-base intermediate. The enzyme activity was not affected by the presence of 1 mm EDTA, which suggests a divalent metal-ion independent reaction. It seems likely that this enzyme is encoded by homologs of the B. thailandensis araD gene, and that the apparent enantioselectivity of this enzyme does not allow a function in the degradation of D-Ara or D-Xyl, which results in a 2-keto-3-deoxypentonic acid with the S-configuration (Fig. 5B).

The aldehyde dehydrogenase from COG1012 is fully conserved in the pentose oxidation gene clusters (Fig. 5A). Strikingly, close homologs of this gene can also be found in hexaric acid degradation gene clusters of Bacillus species (ycbC-ycbI) (71, 72) (Fig. 5A). The same holds for a gene cluster in Pseudomonas putida (PP1245-PP1249) that is likely to be involved in the breakdown of L-hydroxyproline, which is a major constituent of collagen and plant cell wall proteins (73, 74) (Fig. 5B). Apparently, because the degradation of both hexaric acids and L-hydroxyproline is also known to proceed through DOP (9), the genetic information for the conversion of DOP to 2-OG has been shared between multiple metabolic pathways during evolution (Fig. 5, A and B). Apart from the dopDH gene, orthologs of the D-glucarate dehydratase gene (ycbF, COG4948) are observed in the pentose degradation gene clusters of both *S. solfataricus* and H. marismortui, although remarkably, the keto-deoxy-acid dehydratase of COG0329 is found in all three pathways. In the hydroxyproline degradation pathway, this enzyme might function as a deaminase instead (75).

The apparent mosaic of orthologous and non-orthologous proteins involved in the pentose oxidation pathway suggests that some of these enzymatic steps may have evolved by recruitment events between enzymes from the hexaric acid or hydroxyproline degradation pathways, which also make use of DOP as an intermediate and produce 2-OG as the final product (76, 77). The low number of enzymes required, their common cofactor usage, and the large gain of obtaining the hub metabolite 2-OG as the end product of pentose oxidation, may have been the driving force in the creation of this pathway in aerobically respiring Bacteria and Archaea.

Footnotes

The abbreviations used are: 2-OG, 2-oxoglutarate; HSCFE, heat-stable cell-free extract; CFE, cell-free extract; Ara, arabinose; KDA, 2-keto-3-deoxy-arabinonate; DOP, 2,5-dioxopentanoate; AraDH, arabinose dehydrogenase; AraD, arabinonate dehydratase; KdaD, 2-keto-3-deoxy-D-arabinonate dehydratase; DopDH, 2,5-dioxopentanoate dehydrogenase.

References

1 Lee, N., Gielow, W., Martin, R., Hamilton, E., and Fowler, A. (1986) Gene (Amst.) 47, 231-244
2 Rosenfeld, S. A., Stevis, P. E., and Ho, N. W. (1984) Mol. Gen. Genet. 194, 410-415
3 Elsinghorst, E. A., and Mortlock, R. P. (1994) J. Bacteriol. 176, 7223-7232
4 Chiang, C., and Knight, S. G. (1960) Nature 188, 79-81
5 Fossitt, D., Mortlock, R. P., Anderson, R. L., and Wood, W. A. (1964) J. Biol. Chem. 239, 2110-2115
6 Wojtkiewicz, B., Szmidzinski, R., Jezierska, A., and Cocito, C. (1988) Eur. J. Biochem. 172, 197-203
7 Weimberg, R., and Doudoroff, M. (1955) J. Biol. Chem. 217, 607-624
8 Weimberg, R. (1961) J. Biol. Chem. 236, 629-635
9 Dagley, S., and Trudgill, P. W. (1965) Biochem. J. 95, 48-58
10 Duncan, M. J. (1979) J. Gen. Microbiol. 113, 177-179
11 Duncan, M. J., and Fraenkel, D. G. (1979) J. Bacteriol. 137, 415-419
12 Novick, N. J., and Tyler, M. E. (1982) J. Bacteriol. 149, 364-367
13 Mathias, A. L., Rigo, L. U., Funayama, S., and Pedrosa, F. O. (1989) J. Bacteriol. 171, 5206-5209
14 Palleroni, N. J., and Doudoroff, M. (1957) J. Bacteriol. 74, 180-185
15 Dahms, A. S., and Anderson, R. L. (1969) Biochem. Biophys. Res. Commun. 36, 809-814
16 Pedrosa, F. O., and Zancan, G. T. (1974) J. Bacteriol. 119, 336-338
17 Grogan, D. W. (1989) J. Bacteriol. 171, 6710-6719
18 Siebers, B., and Schonheit, P. (2005) Curr. Opin. Microbiol. 8, 695-705
19 Johnsen, U., and Schonheit, P. (2004) J. Bacteriol. 186, 6198-6207
20 Brouns, S. J., Smits, N., Wu, H., Snijders, A. P., Wright, P. C., de Vos, W. M., and van der Oost, J. (2006) J. Bacteriol. 188, 2392-2399
21 Andersson, A., Bernander, R., and Nilsson, P. (2005) Bioinformatics 21, 325-332
22 Lundgren, M., Andersson, A., Chen, L., Nilsson, P., and Bernander, R. (2004) Proc. Natl. Acad. Sci. U. S. A. 101, 7046-7051
23 Snijders, A. P., Walther, J., Peter, S., Kinnman, I., de Vos, M. G., van de Werken, H. J., Brouns, S. J., van der Oost, J., and Wright, P. C. (2006) Proteomics 6, 1518-1529
24 Snijders, A. P., de Vos, M. G., and Wright, P. C. (2005) J. Proteome Res. 4, 578-585
25 Snijders, A. P., de Vos, M. G., de Koning, B., and Wright, P. C. (2005) Electrophoresis 26, 3191-3199
26 Sperber, N., Zaugg, H. E., and Sandstrom, W. M. (1947) J. Am. Chem. Soc. 69, 915-920
27 Kraus, G. A., and Landgrebe, K. (1984) Synthesis 1984, 885
28 Macritchie, J. A., Silcock, A., and Willis, C. L. (1997) Tetrahedron Asymmetry 8, 3895-3902
29 Crestia, D., Guerard, C., Bolte, J., and Demuynck, C. (2001) J. Mol. Catal. B Enzymol. 11, 207-212
30 Esposito, L., Bruno, I., Sica, F., Raia, C. A., Giordano, A., Rossi, M., Mazzarella, L., and Zagari, A. (2003) Biochemistry 42, 14397-14407
31 Warren, L. (1960) Nature 186, 237
32 Kim, S., and Lee, S. B. (2005) Biochem. J. 387, 271-280
33 Lebbink, J. H., Eggen, R. I., Geerling, A. C., Consalvi, V., Chiaraluce, R., Scandurra, R., and de Vos, W. M. (1995) Protein Eng. 8, 1287-1294

34 Schmidt, A., Bahr, U., and Karas, M. (2001) Anal. Chem. 73, 6040-6046
35 Tahallah, N., Pinkse, M., Maier, C. S., and Heck, A. J. (2001) Rapid Commun. Mass Spectrom. 15, 596-601
36 Thompson, W., Rouchka, E. C., and Lawrence, C. E. (2003) Nucleic Acids Res. 31, 3580-3585
37 She, Q., Singh, R. K., Confalonieri, F., Zivanovic, Y., Allard, G., Awayez, M. J., Chan-Weiher, C. C., Clausen, I. G., Curtis, B. A., De Moors, A., Erauso, G., Fletcher, C., Gordon, P. M., Heikamp-de Jong, I., Jeffries, A. C., Kozera, C. J., Medina, N., Peng, X., Thi-Ngoc, H. P., Redder, P., Schenk, M. E., Theriault, C., Tolstrup, N., Charlebois, R. L., Doolittle, W. F., Duguet, M., Gaasterland, T., Garrett, R. A., Ragan, M. A., Sensen, C. W., and van der Oost, J. (2001) Proc. Natl. Acad. Sci. U. S. A. 98, 7835-7840
38 Lundgren, M., and Bernander, R. (2005) Curr. Opin. Microbiol. 8, 662-668.
39 Elferink, M. G., Albers, S. V., Konings, W. N., and Driessen, A. J. (2001) Mol. Microbiol. 39, 1494-1503
40 Komorowski, L., Verheyen, W., and Schafer, G. (2002) Biol. Chem. 383, 1791-1799
41 Lubelska, J. M., Jonuscheit, M., Schleper, C., Albers, S. V., and Driessen, A. J. (2006) Extremophiles, in press
42 Gleissner, M., Kaiser, U., Antonopoulos, E., and Schafer, G. (1997) J. Biol. Chem. 272, 8417-8426
43 Consalvi, V., Chiaraluce, R., Politi, L., Gambacorta, A., De Rosa, M., and Scandurra, R. (1991) Eur. J. Biochem. 196, 459-467
44 Kappler, U., Sly, L. I., and McEwan, A. G. (2005) Microbiology 151, 35-43
45 Uhrigshardt, H., Walden, M., John, H., Petersen, A., and Anemuller, S. (2002) FEBS Lett. 513, 223-229
46 Hess, D., Kruger, K., Knappik, A., Palm, P., and Hensel, R. (1995) Eur. J. Biochem. 233, 227-237
47 Bartolucci, S., Rella, R., Guagliardi, A., Raia, C. A., Gambacorta, A., De Rosa, M., and Rossi, M. (1987) J. Biol. Chem. 262, 7725-7731
48 Londei, P. (2005) FEMS Microbiol. Rev. 29, 185-200
49 Albers, S. V., Jonuscheit, M., Dinkelaker, S., Urich, T., Kletzin, A., Tampe, R., Driessen, A. J., and Schleper, C. (2006) Appl. Environ. Microbiol. 72, 102-111
50 Cline, A. L., and Hu, A. S. (1965) J. Biol. Chem. 240, 4493-4497
51 Mobley, P. W., Metzger, R. P., and Wick, A. N. (1975) Methods Enzymol. 41, 173-177
52 Babbitt, P. C., Mrachko, G. T., Hasson, M. S., Huisman, G. W., Kolter, R., Ringe, D., Petsko, G. A., Kenyon, G. L., and Gerlt, J. A. (1995) Science 267, 1159-1161
53 Bateman, R. L., Bhanumoorthy, P., Witte, J. F., McClard, R. W., Grompe, M., and Timm, D. E. (2001) J. Biol. Chem. 276, 15284-15291
54 St-Louis, M., and Tanguay, R. M. (1997) Hum. Mutat. 9, 291-299
55 Tame, J. R., Namba, K., Dodson, E. J., and Roper, D. I. (2002) Biochemistry 41, 2982-2989
56 Imanaka, T., Ohta, T., Sakoda, H., Widhyastuti, N., and Matsuoka, M. (1993) J. Ferment. Bioeng. 76, 161-167
57 Chen, L., Brugger, K., Skovgaard, M., Redder, P., She, Q., Torarinsson, E., Greve, B., Awayez, M., Zibat, A., Klenk, H. P., and Garrett, R. A. (2005) J. Bacteriol. 187, 4992-4999
58 Kawarabayasi, Y., Hino, Y., Horikawa, H., Jin-no, K., Takahashi, M., Sekine, M., Baba, S., Ankai, A., Kosugi, H., Hosoyama, A., Fukui, S., Nagai, Y., Nishijima, K., Otsuka, R., Nakazawa, H., Takamiya, M., Kato, Y., Yoshizawa, T., Tanaka, T., Kudoh, Y., Yamazaki, J., Kushida, N., Oguchi, A., Aoki, K., Masuda, S., Yanagii, M., Nishimura, M., Yamagishi, A., Oshima, T., and Kikuchi, H. (2001) DNA Res. 8, 123-140
59 Osterman, A., and Overbeek, R. (2003) Curr. Opin. Chem. Biol. 7, 238-251
60 Hottes, A. K., Meewan, M., Yang, D., Arana, N., Romero, P., McAdams, H. H., and Stephens, C. (2004) J. Bacteriol. 186, 1448-1461
61 Meisenzahl, A. C., Shapiro, L., and Jenal, U. (1997) J. Bacteriol. 179, 592-600
62 Moore, R. A., Reckseidler-Zenteno, S., Kim, H., Nierman, W., Yu, Y., Tuanyok, A., Warawa, J., DeShazer, D., and Woods, D. E. (2004) Infect. Immun. 72, 4172-4187
63 Watanabe, S., Kodak, T., and Makino, K. (2006) J. Biol. Chem. 281, 2612-2623
64 Tatusov, R. L., Koonin, E. V., and Lipman, D. J. (1997) Science 278, 631-637
65 Palleroni, N. J., and Doudoroff, M. (1956) J. Biol. Chem. 223, 499-508
66 Deshpande, A. D., Baheti, K. G., and Chatterjee, N. R. (2004) Curr. Sci. 87, 1684-1695

67 Egan, S. E., Fliege, R., Tong, S., Shibata, A., Wolf, R. E., Jr., and Conway, T. (1992) J. Bacteriol. 174, 4638-4646
68 Kim, S., and Lee, S. B. (2006) J. Biochem. (Tokyo) 139, 591-596
69 Stoolmiller, A. C., and Abeles, R. H. (1966) J. Biol. Chem. 241, 5764-5771
70 Portsmouth, D., Stoolmiller, A. C., and Abeles, R. H. (1967) J. Biol. Chem. 242, 2751-2759
71 Sharma, B. S., and Blumenthal, H. J. (1973) J. Bacteriol. 116, 1346-1354
72 Hosoya, S., Yamane, K., Takeuchi, M., and Sato, T. (2002) FEMS Microbiol. Lett. 210, 193-199
73 Yoneya, T., and Adams, E. (1961) J. Biol. Chem. 236, 3272-3279
74 Ramaswamy, S. G. (1984) J. Biol. Chem. 259, 249-254
75 Singh, R. M., and Adams, E. (1965) J. Biol. Chem. 240, 4344-4351
76 Jensen, R. A. (1976) Annu. Rev. Microbiol. 30, 409-425
77 Schmidt, S., Sunyaev, S., Bork, P., and Dandekar, T. (2003) Trends Biochem. Sci. 28, 336-341
78 Jeffcoat, R., Hassall, H., and Dagley, S. (1969) Biochem. J. 115, 977-983

Effect of O_2 concentrations on *Sulfolobus solfataricus* P2

Gwénola Simon*, Jasper Walther*, Nathalie Zabeti, Yannick Combet-Blanc,
Richard Auria, John van der Oost, Laurence Casalot
* authors contributed equally

FEMS microbiology letters 2009 Sept; 299: 255-60

Abstract

Sulfolobus solfataricus P2 was grown aerobically at various O_2 concentrations. Based on growth parameters in microcosms, four types of behavior could be distinguished. At 35% O_2 (v/v; gas phase), the cultures did not grow, indicating a lethal dose of oxygen. For 26–32% O_2, the growth was significantly affected compared with the reference (21%), suggesting a moderate toxicity by O_2. For 16–24% O_2, standard growth was observed. For 1.5–15% O_2, growth was comparable with the reference, but the yield on O_2 indicated a more efficient use of oxygen. These results indicate that *S. solfataricus* P2 grows optimally in the range of 1.5–24% O_2, most likely by adjusting its energy-transducing machinery. To gain some insight into control of the respiratory system, transcriptomes of the strain cultivated at different O_2 concentrations, corresponding to each behavior (1.5%, 21% and 26%), were compared using a DNA microarray approach. It showed differential expression of several genes encoding terminal oxidases, indicating an adaptation of the strain's respiratory system in response to fluctuating oxygen concentrations.

Introduction

It is generally accepted that due to atmospheric photolysis of water, traces of free oxygen were present before the advent of oxygenic photosynthesis (Fenchel & Finlay, 2008). Then, after the appearance of cyanobacteria and photosynthesis, the atmospheric oxygen concentration increased drastically up to >30% before decreasing again to reach the actual concentration (Berner, 1999). At first, microorganisms had, to protect themselves against oxygen before learning how to exploit it for energetic purposes.

Oxygen respiration is mediated through various biochemical reactions producing water and variable quantities of partially reduced oxygen intermediaries. Because of their electronic structure, these compounds are highly reactive. They are known as reactive oxygen species (ROS). The generation of the ROS is considered, at least to some extent, to occur in all living organisms in the presence of oxygen. ROS production proceeds via various physiological mechanisms, and they have been reported to affect the immunity, intercellular signaling or regulation of the cellular growth (Cannio et al., 2000b). When their intracellular concentration becomes too high (because of external processes and phenomena), the cell needs to protect itself. ROS can oxidize lipids, proteins, as well as DNA; the latter results in single- and double-strand breaks or in covalent links between DNA and proteins (Cooke et al., 2003). In order to neutralize the toxic effects of an excess of ROS, cells have developed various protection systems that will either degrade the ROS or repair the oxidative damage in proteins and DNA (Apel & Hirt, 2004; Camenisch & Naegeli, 2009).

Respiration is a fundamental cellular process that consists in the step-wise transfer of electrons from an electron donor to a terminal electron acceptor, through a respiratory chain that consists of membrane-embedded protein complexes with their respective redox cofactors. Today, many bacteria and archaea are capable of reducing dioxygen (O2). The terminal electron acceptors of aerobic respiratory chains are dioxygen reductases, which belong to two protein superfamilies: the cytochrome bd (Watanabe et al., 1979) and A-, B- and C-heme copper oxidases (Pereira et al., 2001). Phylogenetic analyses strongly suggest that B-type oxygen reductases are of archaeal origin, and that A-type oxygen reductases were already present before the divergence of bacteria and Archaea (Brochier-Armanet et al., 2009). These membrane-bound enzymes catalyze the reduction of O2 to water using electrons provided by either a quinol derivate or a cytochrome c

(Wikstrom, 1977). They are, therefore, also called quinol oxidases or cytochrome c oxidases according to their electron donor. The domain of archaea includes many hyperthermophile microorganisms. Among them, thermoacidophiles of the order *Sulfolobales* are model organisms that have been studied at the physiological, biochemical and molecular levels. *Sulfolobus solfataricus* P2 was chosen as a model for studying the effects of oxygen (Zillig et al., 1980). Organisms belonging to the genus *Sulfolobus* are generally described as aerobic. However, it has been observed that, in order to isolate *Sulfolobus* strains from a sample taken in Yellowstone Park, it was necessary to add some reductants to the sample (Y. Combet-Blanc, pers. commun.). Indeed, when the samples were kept at room temperature, allowing more oxygen to pass into the liquid phase, no living *Sulfolobales* could subsequently be isolated. The presence of reductants to reduce the amount of dissolved oxygen allowed the isolation of almost only *Sulfolobales*. This observation suggests that oxygen, when present at a high concentration, might be toxic for *Sulfolobus* strains. In this study, we described the first attempts to analyze the effect of increasing oxygen concentrations on *S. solfataricus* P2 growth and gene expression.

Materials and methods

Cultivation

Sulfolobus solfataricus P2 cells (DSM 1617) were grown at 80 °C, 110 r.p.m. and pH=3 on Brock medium (Brock et al., 1972) modified as follows (L−1): 3 g (NH4)2SO4; 95.21 mg MgCl2; 49.2 mg Ca(NO3)2; 3 g KH2PO4; 1.8 mg MnCl2•4H2O; 4.48 mg Na2B4O7•2H2O; 0.22 mg ZnSO4•7H2O; 50 µg CuCl2•2H2O; 30 µg Na2MoO4•2H2O; 30 µg VOSO4•5H2O; 10 µg CoSO4•7H2O; 10 µg NiSO4•6H2O; 2 mg FeSO4•7H2O; 0.5 mL Wolfe's vitamins solution (Freier et al., 1988) and 2 g glucose. The cultures were inoculated with exponential-phase cells, at an initial density at 600 nm of 0.05.

Penicillin bottles (500 mL with 50-mL medium), sealed with Teflon-coated rubber stoppers, were used in microcosm experiments with different oxygen concentrations. Two mass-flows, one for N2 (0−500 mL min−1) and one for O2 (0−200 mL min−1), were connected to the bottle. The total flux was 400 mL min−1. The settings for N2 and O2 flows were calculated in order to reach the

desired O2 concentrations in the headspace (in volume percentage): 1.5%; 3%; 5%; 10%; 15%; 16%; 21%; 24%; 26%; 28%; 32%; and 35%. At 80 °C, these concentrations correspond to 0.87; 1.74; 2.9; 5.79; 8.69; 9.26; 12.16; 13.9; 15.05; 16.21; 18.53; and 20.27 mM O2, respectively (O2 concentration (M)=[total pressure at 80 °C (Pa) × %O2 in the gas phase]/[R×T (K) × 100] with total pressure at 80 °C=1.7 105 Pa and R=8.31 J K−1 mol−1). After 10 min, the system was closed. At least two replicate microcosms were run for each oxygen concentration.

Cell growth analysis

Samples were periodically taken from cultures to measure OD600 nm using a Biowave (WPA, S2100) spectrophotometer. For the determination of the biomass dry weight, culture samples (50 mL) were washed with demineralized water, dried at 80 °C and weighed. Duplicate determinations varied by <1%. Maximum specific growth rates (μ; h^{-1}) were estimated by plotting the total cell concentration vs. time in a log-linear plot. The slope of the curves thus obtained (a straight line during exponential growth) was used as the average specific rate. Molar growth yields were determined as described previously (Malki et al., 1995). The carbon balance was determined by dividing the carbon amount recovered in the biomass (mc-cells) and the CO2 (mc-CO2) by the carbon amount arising from the consumed glucose (mc-glc). Each part was calculated as follows: mc-cells=Δcell×Vcult× 50%; mc-CO2=(total pressure at 80 °C × %CO2×Mc)/(R×T× 100); mc-glc=Δglc consumed×Vcult×nc×Mc with nc=number of carbon in glucose=6 and Mc=12 g mol^{-1} carbon.

Metabolite and gas analyses

Glucose was analyzed by HPLC using an Alliance 2690 HPLC system (Waters, Milford, MA) supplied with an Aminex HPX-87H 300X7, 8 mm column (Bio-Rad, Hercules, CA) and a Spectra System RI 150 (TSP) refractive-index detector. The column was eluted at 35 °C with 0.0005 N H_2SO_4 at a flow rate of 0.6mL·min^{-1}.

O_2 and CO_2 concentrations were measured in 200-μL gas samples by GC (Perichrom PR2100) equipped with a thermal conductivity detector and a concentric column CTR-1 (Alltech). Helium served as a carrier gas at a flow rate of 31 mL min^{-1}. The column temperature was 60 °C, and the injector and detector temperatures were 100 °C. Measurements were performed in duplicate.

RNA extraction, cDNA synthesis and labeling

Samples (50 mL), taken from cultures at $OD_{600\ nm}$=0.3 containing 1.5%, 10%, 21% or 26% O_2 in the gas phase, were centrifuged and snap-frozen in liquid N_2. RNA was extracted as described (Snijders et al., 2006). RNA concentration and quality were determined using a NanoDrop ND-1000 spectrophotometer (Nano-Drop Technologies, Wilmington, DE).

Transcriptomics

The DNA microarrays were custom-designed oligonucleotide arrays containing 3042 S. solfataricus P2 probes and Arabidopsis thaliana oligonucleotides as negative controls (Ocimum Biosolutions); each probe was represented twice on every slide (Ortmann et al., 2008). These microarrays were used for global transcript profiling of cultures grown at different O2 concentrations. Hybridization, scanning, data filtration, normalization and statistical evaluation were performed as described previously (Snijders et al., 2006). Six microarrays per concentration series were hybridized, including dye swaps for all the compared samples.

Microarray data accession number

Raw data, as well as the final log to related changes, have been deposited in the NCBI Gene Expression Omnibus database under series accession number GSE16043.

Results and discussion

Effect of the oxygen concentration on the growth

Sulfolobus solfataricus is defined in the literature as an aerobic microorganism, based on its ability to grow at an atmospheric oxygen concentration. This condition, corresponding to 21% O2 in the gas phase, was chosen as the reference. The growth was monitored following OD600 nm, glucose consumption, O2 consumption and CO2 production in the gas phase (Fig. 1a). The same parameters were followed when different oxygen concentrations were applied to the microcosm. The tested O2 concentrations in the headspace of the microcosm were in the range from 1.5% to 35%. No growth was ever observed for an O2

concentration of 35% (or more). It seems that at this concentration, oxygen has a lethal effect on S. solfataricus. In the same way, no growth was ever observed in the absence of O2. The carbon balance was determined under all the tested conditions when growth was observed (Fig. 2). Three distinct ranges could be identified. From 1.5% to 15% O2 in the gas phase, 42.25 ± 4.75% of carbon from the glucose was recovered in the biomass and the CO2. From 26% to 32%, the recovery of the carbon was even lower (35 ± 7.22%). From 16% to 24%, the range including the reference (21%), the recovery was significantly higher (72.53 ± 8%). The rather low carbon recovery observed for low and high O2 concentrations might suggest the production of secondary metabolites (exopolysaccharides). In terms of the exponential rate (μmax) and the glucose yield (Yglucose: gram biomass per consumed mole glucose), no difference was observed between 1.5% and 24% O2 (Table 1). In contrast, the growth was less efficient for higher concentrations of O2 (>26%) because in this case, the doubling time changed from 10 h to 20 h. This result suggests a toxic effect of oxygen when the concentration reaches 26% in the gas phase. However, when calculating the yield on oxygen (: gram biomass per consumed mole oxygen) or on carbon dioxide (: gram biomass per produced mole CO2), the cells show a different behavior for concentrations ranging from 1.5% to 15% O2 (Table 1). In this case, the yields on both oxygen and carbon dioxide were higher than the ones measured for the reference (21%).

Sulfolobus solfataricus P2 is an obligate aerobic microorganism because O_2 is required for growth. In case of the highest concentrations allowing growth (32%), the same yield on oxygen, as for 21% O_2, was observed. However, at the high concentrations, the growth rate is significantly affected (Table 1), strongly suggesting toxicity of the oxygen. On the other hand, *S. solfataricus* is also able to grow at a very low oxygen concentration. In this case (when the oxygen concentration in the gas phase is <5%), the limiting factor is no longer the carbon and energy source, but the oxygen, because the growth stops when oxygen, but not glucose, is exhausted (Fig. 1b). For concentrations ranging from 1.5% to 15%, in comparison with the 'normal' culture conditions chosen as a reference in this study (21%), the growth is undisturbed as shown by the yield on glucose and the growth rate. The increase in the yields on O_2 and CO_2, when O_2 is provided in a limited amount, most likely suggests a more efficient energy transduction, because for the same amount of glucose used, the same amount of cells is synthesized, but, in the meantime, less oxygen is required. This phenomenon is even stronger for the

lowest O_2 concentration tested (1.5%) because in this case, the yield on oxygen is 1.5 times higher than the one measured for 3% (Fig. 2). These results suggest a change in the energy metabolism of *S. solfataricus* depending on the O_2 concentration. Transcriptomic analyses of these conditions might provide some information on the control of metabolism in response to a fluctuation in oxygen tension.

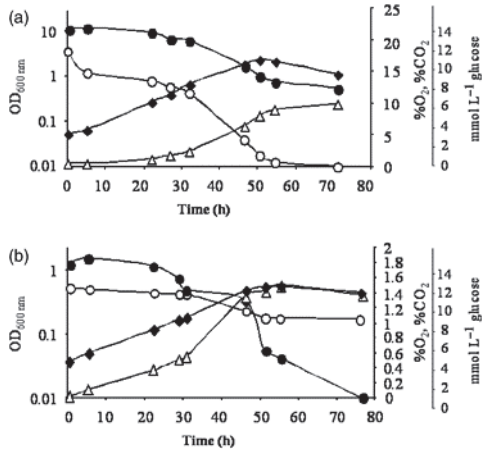

Figure1: Growth, product and substrate kinetics of *Sulfolobus solfataricus* in cultures with 21% O_2 (a) or 1.5% O_2 in the gas phase (b). ◆, $OD_{600\,nm}$; ○, glucose; •, O_2; △, CO_2 concentrations.

Sulfolobus solfataricus P2 is an obligate aerobic microorganism because O_2 is required for growth. In the case of the highest concentrations allowing growth (32%), the same yield on oxygen, as for 21% O_2, was observed. However, at the high concentrations, the growth rate is significantly affected (Table 1), strongly suggesting toxicity of the oxygen. On the other hand, *S. solfataricus* is also able to grow at a very low oxygen concentration. In this case (when the oxygen concentration in the gas phase is <5%), the limiting factor is no longer the carbon and energy source, but the oxygen, because the growth stops when oxygen, but not glucose, is exhausted (Fig. 1b). For concentrations ranging from 1.5% to 15%, in comparison with the 'normal' culture conditions chosen as a reference in this study (21%), the growth is undisturbed as shown by the yield on glucose and the growth rate. The increase in the yields on O_2 and CO_2, when O_2 is provided in a limited amount, most likely suggests a more efficient energy transduction, because for the same amount of glucose used, the same amount of cells is synthesized, but, in the

meantime, less oxygen is required. This phenomenon is even stronger for the lowest O_2 concentration tested (1.5%) because in this case, the yield on oxygen is 1.5 times higher than the one measured for 3% (Fig. 2). These results suggest a change in the energy metabolism of *S. solfataricus* depending on the O_2 concentration. Transcriptomic analyses of these conditions might provide some information on the control of metabolism in response to a fluctuation in oxygen tension.

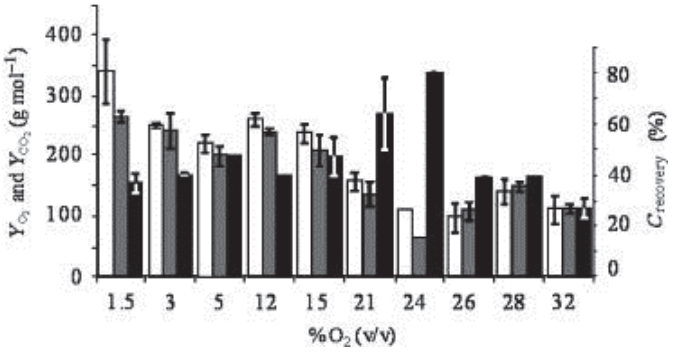

Figure 2: Carbon balance for *S. solfataricus* P2 growth depending on the initial oxygen concentration in the gas phase.

Transcriptomics

The global transcriptional response of S. solfataricus P2 growing exponentially at different oxygen concentrations was determined by DNA microarray analysis. Despite the observed differences of S. solfataricus at the physiological level between 10%, 21% and 26% of O2, no significant changes were detected at the transcriptional level. The apparent absence of regulation at the transcriptional level suggests that the physiological effect of O2 might correspond to a post-transcriptional regulation.

The transcriptome comparison between 1.5% vs. 21% and 1.5% vs. 26% O_2, however, did reveal that some genes were differentially expressed. Even though all the experiments were performed similarly, more pronounced differences in gene expression were observed when comparing the 1.5% sample with the 26% and 21% samples. It seems that this concentration has, on the transcriptomes, more drastic effects that appear in the early exponential phase, almost immediately after exposure to the new O_2 concentration.

Table1: Determination of the growth parameters depending on the initial oxygen concentration in the gas phase

Oxygen Concentration(%)	1.5-15	16-24	26-32
$\mu_{max}(h^{-1})$	0.069 ± 0.011	0.067 ± 0.014	0.035 ± 0.012
$Y_{glucose}(g \cdot mol^{-1})$	63.5 ± 13.5	66.5 ± 11.5	40 ± 18
Y_{O2} $(g \cdot mol^{-1})$	325 ± 75	137.5 ± 37.5	117.5 ± 42.5
Y_{CO2} $(g \cdot mol^{-1})$	237.5 ± 37.5	110 ± 50	125 ± 35

In total, 202 genes and 156 genes were differentially regulated between 1.5% vs. 21% and 1.5% vs. 26%, respectively (Supporting Information, Tables S1 and S2). Only the genes potentially related to oxygen metabolism are listed in Table 2. In this table, the genes upregulated under oxygen-rich conditions encode a superoxide dismutase (Sso0316) (De Vendittis et al., 2001) and six subunits of the SoxM quinol oxidase complex (Sso2968–2973) (Komorowski et al., 2002). Superoxide dismutase plays a role in cell defense against the lethal effect of oxidative stress (Pedone et al., 2004). Therefore, the differential expression of sso0316 was not surprising. Moreover, a higher production of the superoxide dismutase in high-oxygenated cultures of S. solfataricus has been shown previously by Cannio et al. (2000a). The overexpression of the six genes encoding the different subunits of the SoxM supercomplex (sso2968, sso2969, sso2970, sso2971, sso2972, sso2973) suggests its importance under oxygen-rich conditions.

Table 2: Differentially expressed genes. [a] *q* value <5 indicates statistically significant differential expression.

Locus name	Microarray $\log_2$ fold (*q* value)[a]		Description
	1.5% versus 21%	1.5% versus 26%	
Sso2656	3.57 ± 1.95 (>5)	3.18 ± 1.61 (>5)	Quinol oxidase (SoxABC), cytochrome *b* subunit (SoxC)
Sso2657	3.67 ± 1.57 (0)	3.38 ± 0.61 (0)	Quinol oxidase (SoxABC), cytochrome aa_3 subunit (SoxB)
Sso2658	3.11 ± 2.19 (>5)	3.09 ± 1.39 (0)	Quinol oxidase (SoxABC), subunit II (SoxA)
Sso2660	2.22 ± 2.4 (>5)	2.94 ± 1.51 (>5)	Rieske iron-sulfur protein-1 (SoxL-1)
Sso2661	3.06 ± 1.87 (>5)	2.92 ± 1.38 (0)	Hypothetical protein
Sso2662	3.03 ± 1.39 (>5)	3.11 ± 0.66 (0)	Hypothetical protein
Sso10828	2.57 ± 1.43 (2.6)	2.73 ± 1.03 (0)	Quinol oxidase (SoxABC), cytochrome *b* subunit(SoxC)
Sso0316	-1.83 ± 0.4 (0)	-1.49 ± 0.58 (0)	Superoxide dismutase [Fe] (Sod)
Sso2973	-2.31 ± 0.97 (0)	-1.66 ± 1.03 (0.59)	Quinol oxidase-2, subunit I/III, cytochrome aa_3 (SoxM)
Sso2969	-2.54 ± 1.42 (>5)	-1.75 ± 1.21 (0)	Quinol oxidase-2, subunit II (SoxH)
Sso2968	-2.77 ± 1.42 (>5)	-1.74 ± 1.14 (0)	Quinol oxidase-2, putative subunit (SoxI)
Sso2970	-2.53 ± 1.5 (>5)	-1.54 ± 1.78 (>5)	Quinol oxidase-2, cytochrome *b* (SoxG)
Sso2971	-2.15 ± 1.95 (>5)	-1.34 ± 1.54 (>5)	Quinol oxidase-2, Rieske iron-sulfur protein-2 (SoxF)
Sso2972	-2.77 ± 0.69 (>5)	-1.69 ± 0.96 (>5)	Quinol oxidase-2, sulfocyanin (SoxE)

The highly expressed genes under low-oxygen conditions included genes encoding subunits of a putative quinol oxidase complex (Sso2658, Sso2657 and Sso10828) and two uncharacterized proteins (Sso2661 and Sso2662). The putative quinol oxidase complex in *S. solfataricus* (Sso2658, Sso2657, Sso2656, Sso10828) was identified by its strong homology with the SoxABCD complex studied in *S. acidocaldarius* (Lubben et al., 1994). This complex was described as a proton pump showing its important role in the energy metabolism of the archaea. Sso2656 was

annotated as a putative SoxC. Like the other subunits in the complex, SoxC appears to be overexpressed under low-O_2 conditions. The overexpression of the genes associated in the *soxABCD* cluster strongly suggests an important role of the corresponding proteins under low-oxygen conditions. Next to the *soxABCD* genes, a small gene cluster was overexpressed under the same conditions; the cluster encodes a putative Rieske iron–sulfur cluster protein (SoxL-1; Sso2660) as well as two uncharacterized proteins Sso2661 and Sso2662. The fact that *sso10828, sso2656, sso2657, sso2658, sso2660, sso2661* and *sso2662* are coregulated was previously observed when comparing the growth of *S. solfataricus* on glucose and arabinose (Brouns et al., 2006). In addition, upregulation of *soxABCD* was demonstrated at a low iron concentration in a related thermoacidophile, *Metallosphaera sedula* (Auernik & Kelly, 2008). The common behavior of the *soxABCD* regulation in these different cases strongly suggests that it is an operon. During growth at reduced O_2 concentrations, the cell will need oxidases that are more efficient or that have a higher affinity for O_2. SoxM (A-type oxidase) has been characterized in the more distantly related crenarchaeon *Aeropyrum pernix* as an oxidase with a relatively low affinity for O_2, whereas SoxABC (B-type oxidase) has a relatively high oxygen affinity (Ishikawa et al., 2002). Our observation that in *S. solfataricus* P2, under low-oxygen conditions the expression of the SoxABCD quinol oxidase complex is induced, whereas under oxygen-rich conditions the SoxM-mediated terminal quinol oxidation is upregulated, is in perfect agreement with this observation. The same observation was made for the bacterium *Thermus thermophilus*, which contains a B-type oxidase expressed under microaerobic conditions, in addition to an A-type oxidase expressed at high oxygen levels (Mather et al., 1993; Keightley et al., 1995).

Conclusion

This is the first time that a more detailed analysis was conducted on the behavior of *S. solfataricus* P2 at a wide range of O_2 concentrations. According to the results obtained in this study, *S. solfataricus* P2 is able to grow at O_2 concentrations ranging from 1.5% to at least 32%. It seems, however, that the best conditions for growth are at an O_2 concentration of 1.5–24%. Transcriptome analyses showed that several genes were differentially expressed depending on the O_2 concentration. However, the physiological behavior of the strain shows that

significant regulation also occurs at different levels. The microcosm experiments used in this work did not allow to control the oxygen concentration during cultivation. Thus, microcosms are transitory systems where the O_2 concentration is initially defined and decreases during growth. Fermentor cultivations with an oxygen sensor working at high temperatures will be performed in future experiments. This automated system maintains a constant O_2 concentration during the growth. With such a system, we should also be able to analyze the adaptation of the strain to drastic changes in O_2 concentration.

References

Apel K & Hirt H (2004) Reactive oxygen species: metabolism, oxidative stress, and signal transduction. Annu Rev Plant Biol 55: 373–399.

Auernik KS & Kelly RM (2008) Identification of components of electron transport chains in the extremely thermoacidophilic crenarchaeon *Metallosphaera sedula* through iron and sulfur compound oxidation transcriptomes. Appl Environ Microb 74: 7723–7732.

Berner RA (1999) Atmospheric oxygen over Phanerozoic time. P Natl Acad Sci USA 96: 10955–10957.

Brochier-Armanet C, Talla E & Gribaldo S (2009) The multiple evolutionary histories of dioxygen reductases: implications for the origin and evolution of aerobic respiration. Mol Biol Evol 26: 285–297.

Brock TD, Brock KM, Belly RT & Weiss RL (1972) Sulfolobus: a new genus of sulfur-oxidizing bacteria living at low pH and high temperature. Arch Mikrobiol 84: 54–68.

Brouns SJ, Walther J, Snijders AP et al. (2006) Identification of the missing links in prokaryotic pentose oxidation pathways: evidence for enzyme recruitment. J Biol Chem 281: 27378–27388.

Camenisch U & Naegeli H (2009) Role of DNA repair in the protection against genotoxic stress. Exs 99: 111–150.

Cannio R, D'Angelo A, Rossi M & Bartolucci S (2000a) A superoxide dismutase from the archaeon Sulfolobus solfataricus is an extracellular enzyme and prevents the deactivation by superoxide of cell-bound proteins. Eur J Biochem 267: 235–243.

Cannio R, Fiorentino G, Morana A, Rossi M & Bartolucci S (2000b) Oxygen: friend or foe? Archaeal superoxide dismutases in the protection of intra- and extracellular oxidative stress. Front Biosci 5: 768–779.

Cooke MS, Evans MD, Dizdaroglu M & Lunec J (2003) Oxidative DNA damage: mechanisms, mutation, and disease. FASEB J 17: 1195–1214.

De Vendittis E, Ursby T, Rullo R, Gogliettino MA, Masullo M & Bocchini V (2001) Phenylmethanesulfonyl fluoride inactivates an archaeal superoxide dismutase by chemical modification of a specific tyrosine residue. Cloning, sequencing and expression of the gene coding for *Sulfolobus solfataricus* superoxide dismutase. Eur J Biochem 268: 1794–1801.

Fenchel T & Finlay B (2008) Oxygen and the spatial structure of microbial communities. Biol Rev 83: 553–569.

Freier D, Mothershed CP & Wiegel J (1988) Characterization of *Clostridium thermocellum* JW20. Appl Environ Microb 54: 204–211.

Ishikawa R, Ishido Y, Tachikawa A, Kawasaki H, Matsuzawa H & Wakagi T (2002) *Aeropyrum pernix* K1, a strictly aerobic and hyperthermophilic archaeon, has two terminal oxidases, cytochrome ba_3 and cytochrome aa_3. Arch Microbiol 179: 42–49.

Keightley JA, Zimmermann BH, Mather MW, Springer P, Pastuszyn A, Lawrence DM & Fee JA (1995) Molecular genetic and protein chemical characterization of the cytochrome ba_3 from *Thermus thermophilus* HB8. J Biol Chem 270: 20345–20358.

Komorowski L, Verheyen W & Schafer G (2002) The archaeal respiratory supercomplex SoxM from *Sulfolobus acidocaldarius* combines features of quinole and cytochrome c oxidases. Biol Chem 383: 1791–1799.

Lubben M, Warne A, Albracht SP & Saraste M (1994) The purified SoxABCD quinol oxidase complex of *Sulfolobus acidocaldarius* contains a novel haem. Mol Microbiol 13: 327–335.

Malki S, Saimmaime I, De Luca G, Rousset M, Dermoun Z & Belaich JP (1995) Characterization of an operon encoding an NADP-reducing hydrogenase in *Desulfovibrio fructosovorans*. J Bacteriol 177: 2628–2636.

Mather MW, Springer P, Hensel S, Buse G & Fee JA (1993) Cytochrome oxidase genes from *Thermus thermophilus*. Nucleotide sequence of the fused gene and analysis of the deduced primary structures for subunits I and III of cytochrome caa_3. J Biol Chem 268: 5395–5408.

Ortmann AC, Brumfield SK, Walther J, McInnerney K, Brouns SJ, Van De Werken HJ, Bothner B, Douglas T, Van De Oost J & Young MJ (2008) Transcriptome analysis of infection of the archaeon *Sulfolobus solfataricus* with Sulfolobus turreted icosahedral virus. J Virol 82: 4874–4883.

Pedone E, Bartolucci S & Fiorentino G (2004) Sensing and adapting to environmental stress: the archaeal tactic. Front Biosci 9: 2909–2926.

Pereira MM, Santana M & Teixeira M (2001) A novel scenario for the evolution of haem–copper oxygen reductases. Biochim Biophys Acta 1505: 185–208.

Snijders AP, Walther J, Peter S, Kinnman I, De Vos MG, Van De Werken HJ, Brouns SJ, Van Der Oost J & Wright PC (2006) Reconstruction of central carbon metabolism in *Sulfolobus solfataricus* using a two-dimensional gel electrophoresis map, stable isotope labelling and DNA microarray analysis. Proteomics 6: 1518–1529.

Watanabe H, Kamita Y, Nakamura T, Takimoto A & Yamanaka T (1979) The terminal oxidase of *Photobacterium phosphoreum*. A novel cytochrome. Biochim Biophys Acta 547: 70–78.

Wikstrom MK (1977) Proton pump coupled to cytochrome c oxidase in mitochondria. Nature 266: 271–273.

Zillig W, Stetter KO, Wunderl S, Schulz W, Priess H & Scholz I (1980) The Sulfolobus-'Caldariella' group: taxonomy on the basis of the structure of DNA-dependent RNA polymerases. Arch Microbiol 125: 259–269.

Carotenoid production in *Sulfolobus*

Jasper Walther, Mark Scaife, Phillip C. Wright, Willem M. de Vos,
John van der Oost

Abstract

Carotenoids have an important function in nature as a colorant or as a protective agent against UV radiation and reactive oxygen species. In humans carotenoids and their derivatives have a vital function in the eye and are therefore considered healthy. For these reasons carotenoids are often used as food additives (colorant) or as a health additive (ß-carotene being a precursor for vitamin A). The production of carotenoid-like molecules in Archaea is not very well understood. Here we have studied the carotenoid pathways in several *Sulfolobus* species. Besides ß-carotene they produce zeaxanthin, a carotenoid crucial for preventing age-related macular degeneration and cataracts. Furthermore they also generate glycosylated zeaxanthin, which is a water-soluble carotenoid. We have used transcriptomics, bioinformatics and heterologous expression to elucidate the carotenoid pathways in different *Sulfolobus* species, and to prove that the previously unknown gene in the carotenoid cluster encodes a zeaxanthin glycosidase.

Introduction

Carotenoids are isoprenoid-based pigments that have been proposed to be invented by the Archaea (Rohmer, Bouvier et al. 1979). Their original function has been proposed to be the reinforcement of the cell membranes. These early carotenoids diversified into a large group of distinct pigments, which are widely distributed in nature, playing a wide variety of different roles (reviewed by Matsumi et al. 2011 and Jackson et al. 2008). Some functions include light harvesting activities or protection against reactive oxygen species. Carotenoids have received considerable attention due to their use as a natural food colorant, healthy food additive and as an additive in cosmetics. Of the many carotenoids described in literature only a few can be produced cost-effectively in useful quantities. A more profound understanding of the carotenoid-producing pathways is expected to contribute to further developing viable production processes. Moreover, this will contribute to our insight in how complex molecules are generated and allow for the comparative analysis of carotenoid production pathways in the domains of the Archaea, Bacteria and Eukarya. Here we will focus on the carotenoid-like compound zeaxanthin because of its relatively wide application.

Carotenoids are generally hydrophobic compounds and therefore their application is limited. Zeaxanthin has, compared to other carotenoids, a relatively high solubility in water. The addition of hydrophilic groups, such as rhamnose or glucose, further improves its solubility and therefore its use as colorant (Fig. 1, Fig. 2). Zeaxanthin is generally regarded as safe and is used as a food-additive (E161h; http://en.wikipedia.org/wiki/E161h). Furthermore, zeaxanthin is one of the two primary xanthophyll carotenoids in the retina of the eye. Various studies have indicated that a zeaxanthin-rich diet can reduce age-related macular degeneration and cataracts (Krishnadev et al. 2010). Hence there is considerable interest in zeaxanthin and its natural sources including food products such as eggs, spinach, broccoli, kiwi and corn (SanGiovanni et al. 2007).

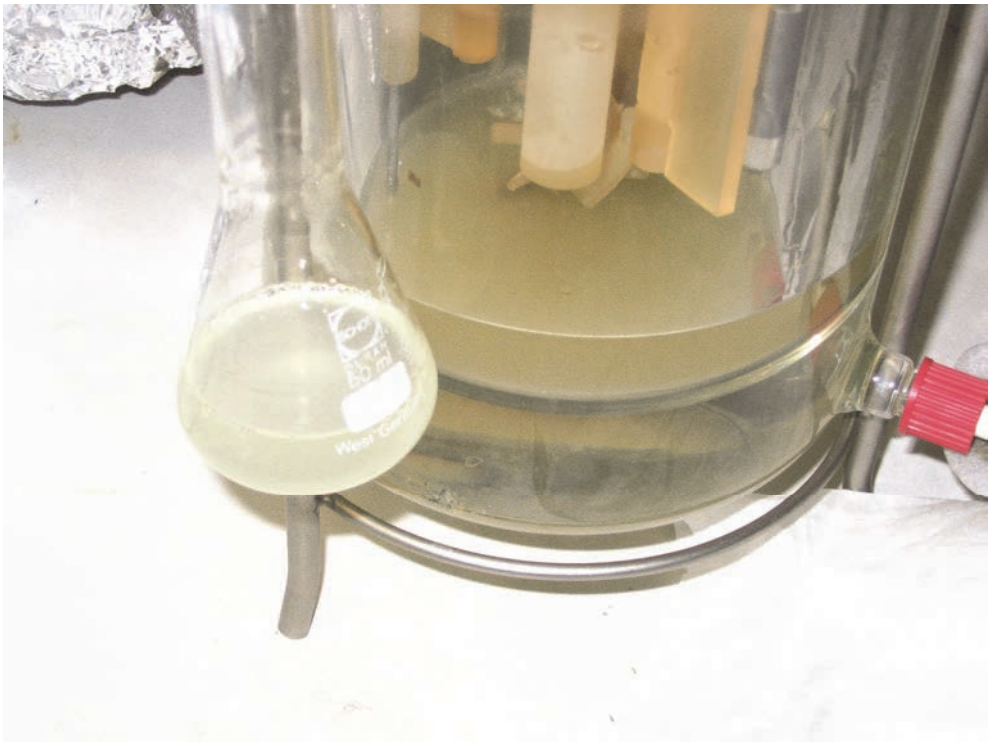

Figure 1: Carotenoid production in *S. solfataricus*. A major difference in colour was observed when comparing *S. solfataricus* grown in normal conditions (flask) and when grown with intense light (bioreactor).

Zeaxanthins (and the glycolsylated forms) are biologically interesting because of their profound protective properties in vivo. One of the main functions of carotenoids in many organisms is the protection from reactive oxygen species (ROS), such as singlet oxygen (1O2) and superoxide (O2-), that are produced by (sun)light. A study by Tatsuzawa et al. (2000) has shown that glycosylated zeaxanthin has very good protective properties against 1O2, about 4.8 times better then ß-carotene, while non-glycosylated zeaxanthin was about 3.1 times better then ß-carotene in protecting against 1O2. The glucosylation of zeaxanthin was first discovered in the Erwinia species. The bacterial CrtX enzyme catalyzes in two additions two glucose molecules to the two alcohol groups of zeaxanthin, resulting in zeaxanthin 3,3'-ß-D-diglucoside. Other known bacteria genera that have a *ctrx* gene are Enterobacter, Cronobacter, Patoea and Pseudomonas.

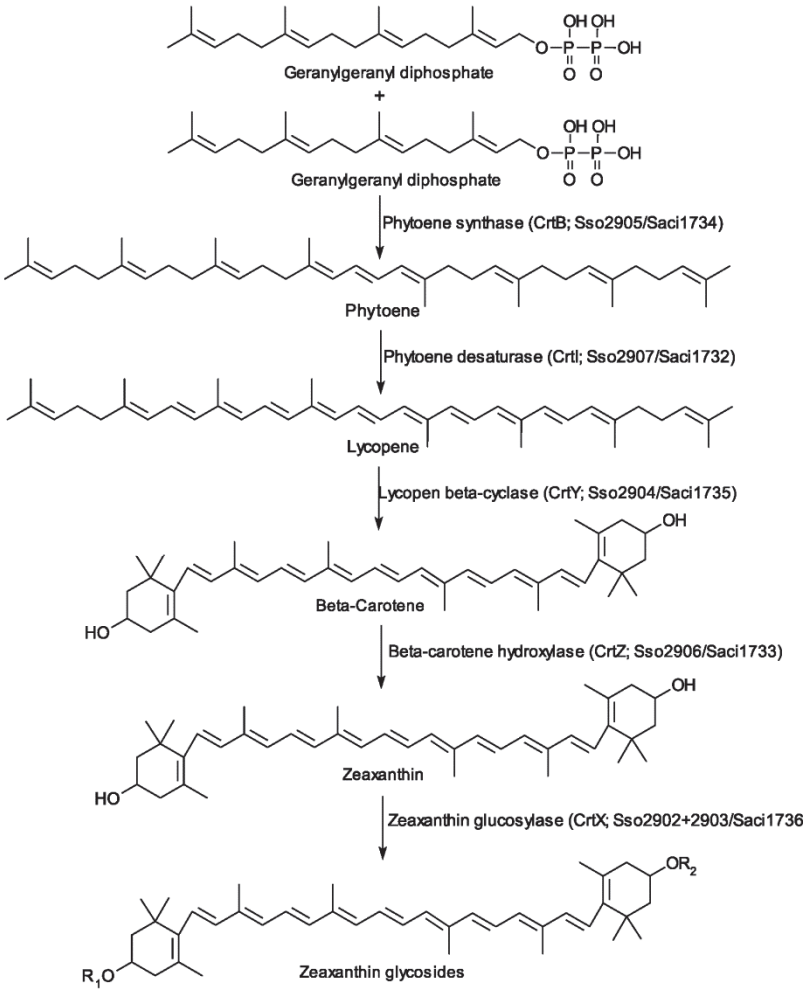

Figure 2: The proposed zeaxanthin pathway for *S. solfataricus* and *S. acidocaldarius*. Bioinformatics, microarray data and the existing literature are used to predict this pathway. The production of different zeaxanthin glycosides is probably not done in one, but in two steps: first, one hydrogen is replaced to form an ester-bond with either glucose or rhamnose or a different sugar derived compound. The second step is the repetition of the first one, but on the other hydroxyl-group at the other side of the zeaxanthin molecule. The name Crtx is choosen because of similarity with the known Crtx enzyme from *Erwinia herbicola* and *E. uredovora* (Sieiro, Poza et al. 2003).

Besides the bacterial genera mentioned above, also several *Sulfolobus* species are also able to produce glycosylated zeaxanthins (Kull and Pfander 1997). The genus *Sulfolobus* belongs to the phylum *Crenarchaeota* and its members are among the best studied Archaea. They are acidothermophiles that grow optimally at temperatures between 80 and 85°C and at a pH between 2 and 5 (Grogan 1989; She et al. 2001). They are obligate aerobes that grow heterotrophically on many different carbon sources. Hence, they have the potential to be developed as production hosts. While the normal carotenoid production levels are relatively low in *Sulfolobus*, compared to other sources such as green algae, a mutant strain of *Sulfolobus shibatae* was described that showed an increased carotenoid production (Grogan 1989). Subsequently, the carotenoids generated by this overproducing strain were characterized biochemically (Kull and Pfander 1997). This has resulted in the discovery of seven different zeaxanthin glycosides, in which either glucose, rhamnose or a combination of both was attached to the zeaxanthin molecule.

Sequence analysis of a gene cluster from *Sulfolobus solfataricus*, a close relative of *Sulfolobus shibatae*, led to the discovery of a gene cluster of four genes that could be involved in the production of carotenoids (Hemmi, Ikejiri et al. 2003). The function of one of the identified genes (*crtY*) predicted to encode a lycopene-β-cyclase, was confirmed by expressing it in *Escherichia coli* and measuring the production of β-carotene. Although the gene-cluster that is involved in the production of β-carotene was correctly predicted in this study, the gene(s) involved in the final step, i.e. the production of glycosylated zeaxanthin, has (have) not been identified (Kull and Pfander 1997). In the present study we have performed a molecular genetics study on two *Sulfolobus* species, *S. solfataricus* and *S. shibatae*, and have elucidated the complete pathway in several *Sulfolobus* species and proposed the presence and activity of a *ctx*-like gene for *Sulfolobus acidocaldarius* and *Sulfolobus islandicus*.

Experimental procedures

Growth conditions

S. solfataricus P2 (DSM1617), *S. shibatae* and a *S. shibatae* mutant as previously described by Grogan (Grogan 1989) were grown in a novel corrosion resistant bioreactor (this thesis, Chapter 7) using medium described previously

(Brouns et al., 2006) with 3g/litre D-glucose as a carbon and energy source. The temperature and dissolved oxygen were actively controlled and kept at 79°C, and at 80% of dissolved oxygen saturation.

Sampling an induction of the carotenoid production

Cells were first grown in complete darkness (wrapped in aluminium foil) until an OD_{600} of 0.5 was reached and the dark sample was taken. Then the light source (Philips Halotone plus 300W; 230V; R&s; 5600 IM) was switched on and after 20 minutes the light sample was taken. The cells were sampled at stirrer level via a sampling tube by allowing an overpressure in the bioreactor. Via the sampling tube the 50 ml samples were quickly poured into pre-chilled Greiner-tubes and cooled in ice-water, after which the cells were harvested by centrifugation at 4°C.

Transcriptomics

The RNA extraction, cDNA synthesis, hybridisation of the microarrays and the analysis of the arrays were done as described previously (Ortmann et al., 2008). Differentially expressed genes have the following criteria: $-1.0 \leq {}^{2}\log \text{value} \geq 1.0$ and the q value should be 5% or lower, using the significance analysis of microarrays (SAM) Stanford Tools Excel plug-in (Tusher et al. 2001)

Cloning

The *crtx*-like genes of *S. shibatae*, *S. solfataricus* (*sso2902* and *ss02903)* and *S. acidocaldarius* (*saci1736)* were cloned into a pBAD24 expression vector (Guzman et al 1995) in an zeaxanthin producing *E. coli* strain. The gene expression was induced with L-arabinose. The *E. coli* strain used was described by Hundle et al. 1995.

HPLC analysis

Modification of zeaxanthin in the transformed *E. coli* strain was tested using HPLC. The HPLC-analysis was done as described by Kull and Pfander in 1997.

Results and Discussion

Fermentation

We studied the carotenoid production in Sulofobus spp, the cells of *Sulfolobus solfataricus (Sso), S. shibatae* wild type (*Ssh*-wt) and the *S. shibatae* mutant (*Ssh*-mutant), a carotene-overproducing mutant of *S. shibatae* described previously (Grogan 1989). For this purpose, these strains were grown separately in a bioreactor that allowed for reproducible growth. To induce carotenoid production, the cultures were grown in the dark until mid-log phase and illuminated by intense light followed by repeated sampling. After 20 minutes of light induction, the light sample was taken. The final density of the cultures varied with a maximum OD_{600} of around 1.5, 1.0 and 2.0 after 24 hours of induction, for the strains *Sso, Ssh*-wt and *Ssh*-mutant, respectively. An orange/brown colour change was clearly visible after light induction in all cultures, especially at the end of the exponential growth phase (Fig. 1).

Table 1: The microarray data concerning the zeaxanthin producing pathway. The data presented here is the ^{2}log ratio of light sample vs. the dark sample. First the organisms were grown in absolute dark (wrapped in aluminium foil) and at an OD_{600} of 0.5 the dark sample was taken. Then the bioreactor was unwrapped and the culture was illuminated with intense light. Twenty minutes after switching on the light, the light sample was taken. The OD_{600} did not change significantly.

Substrate	gene number	S. solfataricus	S. shibataewt	S.shibataemut.	Product
Geranylgeranyl-PP	Sso2905	4.9	2.0	4.9	Phytoene
Phytoene	Sso2907	3.6	-0.2	3.8	Lycopene
Lycopene	Sso2904	2.5	-0.8	1.1	B-carotene
B-carotene	Sso2906	3.6	5.0	3.9	Zeaxanthin
Zeaxanthin	sso2902	no data	no data	no data	Glycosylated Zeaxanthin
Zeaxanthin	sso2903	3.1	0.1	3.5	Glycosylated Zeaxanthin

DNA microarrays

In order to find the genes responsible for the production of (glycosylated) zeaxanthin, a DNA microarray study was performed using RNA isolated from the *Ssh*-wt, *Ssh-mutant* and the *Sso* strains to try to elucidate the last part of the pathway. The used DNA microarray was originally designed for *S. solfataricus*

(Ortman et al., 2008) but was found to be very useful for analysing the transcriptional response of the not yet completely sequenced *S. shibatae*. For both the *Sso* and the *Ssh-mutant* strains, the carotenoid pathway was found to be upregulated: gene cluster Sso2904-2907 (see Figure 2 and 3). For unknown reasons only the two first genes in the two operons are upregulated in the *Ssh-wt* strain. Overall the rest of the transcriptome of this strain reveals very little differential regulation. In this strain very little is differentially regulated. This may suggest that due to the increased carotenoid production, there is no stress in the organism. This is, however, contradicted by the *Ssh-mutant* strain, in which many genes are differentially regulated. The mutations in the *Ssh-mutant* are unknown and therefore we cannot exclude the possibility that the overexpressions of the other genes are due to a single pleiotropic mutation, or to multiple mutations. In table 1 and 2 all differentially expressed genes of *S. solfataricus* are presented, including the relative expression of the homologues of the two *S. shibatae* strains. Surprisingly, only two genes/homologues, which are not part of the carotenoid pathway, are differentially expressed in all 3 strains. The first one is a putative transcriptional regulator (*sso2581*) and therefore a potential candidate for a light-responsive regulator.

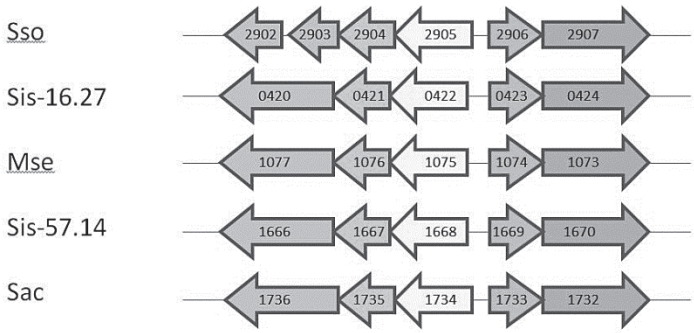

Figure 3**:** Gene cluster for the zeaxanthin pathway for *S. solfataricus*,*S. acidocaldarius*, *Sulfolobus islandicus* Y.G.57.14, *S. islandicus* m16.27 and *Metallosphaera sedula* (Mse). This cluster can be found using a reciprocal BLAST search. Here can be seen that two operons are responsible for the production of ß-carotene, zeaxanthin and glycosylated zeaxanthins.

The second gene (sso2155) is a hypothetical protein with no conserved regions. It can, however, be found in other *Sulfolobus* species and a shorter version can be found in *Acidianus hospitalis* and *Metallosphaera sedula*. It possesses one transmembrane region using TMHMM (Krog, et al 2001), otherwise no known domains can be found.

Carotenoid analysis

HPLC analysis was used to demonstrate that *S. shibatae* and *S. solfataricus* produce carotene and zeaxanthin (not shown). There were a number of unidentified peaks in the HPLC profile that potentially could be glycosylated zeaxanthins; however, since there is no standard available for the different glycosylated zeaxanthins, this could not easily be confirmed.

To show that the a *crtx*-like gene does indeed exist in *Sulfolobus* spp, the *crtx* candidate genes for *S. acidocaldarius* (Sac1736), *S. solfataricus* (Sso2902/ 2903) and *S. shibatae* (see below) were cloned in an *E. coli* overproducing zeaxanthin strain using a pBAD24 plasmid. The carotenoid-like products of the resulting clones were analysed with HPLC (Fig. 4). This *E.coli* host was previously used to express the zeaxanthin cluster of *Erwinia* (Hundle et al. 1995). It appears that *S. solfataricus* has lost the ability to produce glycosylated zeaxanthin, most likely because the *ctrx*-like gene is split, and as such a functional enzyme was not produced in *E. coli*. On the other hand, when the *crtx*-like gene of *S. acidocaldarius* is expressed in *E.coli*, produces different zeaxanthin glycosides than the *S. shibatae* strain. We have not yet been able to determine which groups were added to the zeaxanthin molecule. Because a genome sequence of *S. shibatae* is not available, the sequence of the *ctrx*-like gene is unknown. Therefore, we used the primers from the *S. acidocaldarius*-strain on the *shibatae*-strain. Remarkably, this resulted in a gene-product that could alter the zeaxanthin molecule in the same way as *S. shibatae* strain does.

Table 2: The differential expressed genes of *S. solfataricus* with and without light exposure. In the last two columns the data for the homologues in the two *S. shibatae* strains are presented. A positive number means upregulation in the light induced sample.

		S. solfataricus	S. shibatae mt	S. shibatae wt
Systematic Name	Possible function	^{2}log average	^{2}log average	^{2}log average
SSO2911	Phosphoadenosinephosphosulfate reductase	-1.3	-0.8	-0.8
SSO2909	Sulfite reductase hemoprotein	-1.2	-0.9	-0.8
SSO2912	Sulfateadenylyltransferase	-1.1	-0.9	0.0
SSO2914	Hypotheticalprotein	-1.1	-0.8	-0.5
SSO2913	Hypotheticalprotein	-1.0	-0.8	0.2
SSO0436	Thiazolebiosyntheticenzyme	1.3	1.1	0.0
SSO1079	Hypotheticalprotein	2.6	2.5	1.3
SSO1080	ABC transporterpermeaseprotein	1.8	1.7	0.8
SSO1227	Benzenemonooxygenaseoxygenasesubunit	1.7	1.8	-0.1
SSO1320	Hypotheticalprotein	1.2	1.2	0.3
SSO1324	ThiaminebiosynthesisproteinthiC	1.7	1.3	0.0
SSO1593	Thiaminetransporter	1.7	1.4	0.1
SSO1815	Hypotheticalprotein	2.9	3.0	0.2
SSO1872	hyptheticalprotein	2.8	2.6	-0.2
SSO1877	Hypotheticalprotein	1.6	0.9	0.7
SSO1941	ThiaminebiosynthesisproteinthiC	2.0	1.7	-0.2
SSO2089	Hypotheticalprotein	1.7	2.1	0.0
SSO2155	Hypotheticalprotein	2,5	2,7	1,9
SSO2178	Aspartate-semialdehyde dehydrogenase	1,8	-0,1	0,0
SSO2581	Put transc regulator asnCfam	1,9	1,8	2,1

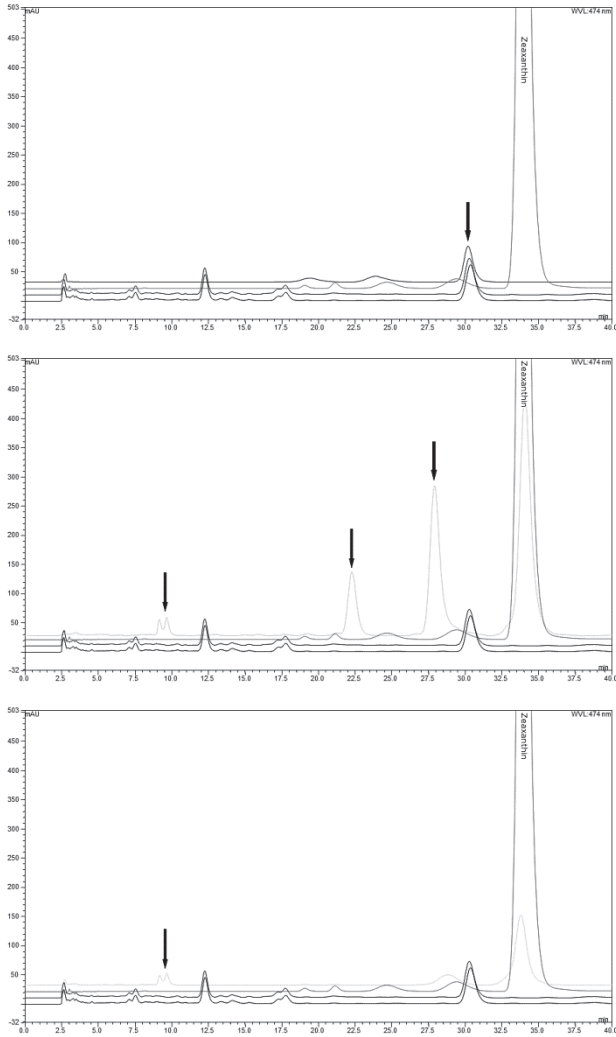

Figure 4: HPLC-graphs of the glycosylated zeaxanthin. The *ctrx*-like genes of *S. solfataricus* (top), *S. acidocaldarius* (middle) and *S. shibatae* (bottom) were cloned into an *E. coli* overproducing zeaxanthin strain. Here can be seen that the *crtx*-like gene of *S. solfataricus* does not work in *E. coli*. The *crtx*-like gene in *S. acidocaldarius* produces more than one product, which is different than the glycosylated zeaxanthins that the *S. shibatae* strain produces.

Conclusion and discussion

Carotenoids are of great use as natural colorants and as a health additive. Not much is known about the archaeal carotenoid pathways. This study has tried to elucidate the carotenoid pathways in 4 *Sulfolobus* strains (*S. solfataricus, S. shibatae, S. acidocaldarius* and *S. islandicus*). Microarray studies have been performed with *S. solfataricus* and the unsequenced *S. shibatae*. A bioinformatics study was used to identify the carotenoid genes in *S. acidocaldarius* and *S. islandicus*. The microarray studies have shed some light on the *crtx*-like gene that is responsible for the glycosylation of the zeaxanthin molecule by these organisms. This work has shown that a very well preserved cluster is operational in these organisms. A Crtx-like enzyme is active in *S. shibatae*, which most likely couples glucose and/or rhamnose to zeaxanthin (Kull et al. 1997). *S. solfataricus* has probably lost the ability of producing glycosylated zeaxanthins, due to the fact the most likely candidate (*sso2902*/*sso2903*) is truncated. *S. acidocaldarius* has been shown to modify zeaxanthin, when the *saci1736* gene is expressed in the zeaxanthin-producing *E.coli* strain, as the exact structure of the product is not known. Further research should be done to elucidate the exact glycosylation pattern of the zeaxanthins produced by these *Sulfolobus* species.

References

Armstrong, G. A. (1997). Genetics of Eubacterial Carotenoid Biosynthesis: a Colorful Tale. Annu. Rev. Microbiol. 51: 629-659.

Brouns, S. J., J. Walther, et al. (2006). Identification of the missing links in prokaryotic pentose oxidation pathways: evidence for enzyme recruitment. J Biol Chem. 281(37): 27378-27388.

Grogan, D. W. (1989). "Phenotypic characterization of the archaebacterial genus Sulfolobus: comparison of five wild-type strains." J Bacteriol 171(12): 6710-6719.

Guzman, L.M., et al., Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. J Bacteriol, 1995. 177(14): p. 4121-30.

Hemmi, H., S. Ikejiri, et al. (2003). "Fusion-type lycopene [beta]-cyclase from a thermoacidophilic archaeon Sulfolobus solfataricus." Biochemical and biophysical research communications 305(3): 586-591.

Hundle, B. S., D. A. O'Brien, M. Alberti, P. Beyer, and J. E. Hearst. (1992). Functional expression of zeaxanthin glucosyltransferase from Erwinia herbicola and a proposed uridine diphosphate binding site. Proc Natl Acad Sci U S A. 89(19): 9321–9325.

Jackson, H., Braun, C. L., en Ernst, H. (2008). The chemistry of novel xanthophyll carotenoids. Am. J. Cardiol 101, 50D-57D.

Krishnadev N., Meleth A. D., Chew E. Y. (2010). "Nutritional supplements for age-related macular degeneration". Current Opinion in Ophthalmology 21(3): 184–189.

Krogh, A., B. Larsson, G. von Heijne, E. L.L Sonnhammer. (2001). Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. J. of Mol. Biol. 305 (3): 567-580.

Kull, D. R. and H. Pfander (1997). "Isolation and structure elucidation of carotenoid glycosides from the thermoacidophilic archaea *Sulfolobus shibatae*." Journal of Natural Products 60 (4): 371-374.

Matsumi, R., Atomi, H., Driessen, A.J., Van der Oost, J. (2011) Isoprenoid biosynthesis in Archaea - Biochemical and Evolutionary implications. Res. Microbiol. 162: 39-52.

Rohmer, M., P. Bouvier, et al. (1979). "Molecular evolution of biomembranes: structural equivalents and phylogenetic precursors of sterols." Proc Natl Acad Sci USA 76: 847-851.

Ortmann AC, Brumfield SK, Walther J, McInnerney K, Brouns SJ, Van De Werken HJ, Bothner B, Douglas T, Van De Oost J&Young MJ *(*2008*)* Transcriptome analysis of infection of the archaeon*Sulfolobussolfataricus* with *Sulfolobus* turreted icosahedral virus. J Virol82*:* 4874–4883.

SanGiovanni J. P., Chew E. Y., Clemons T. E., et al. (2007). "The relationship of dietary carotenoid and vitamin A, E, and C intake with age-related macular degeneration in a case-control study: AREDS Report No. 22". Archives of Ophthalmology 125(9): 1225–1232.

Sieiro, C., M. Poza, et al. (2003). "Genetic basis of microbial carotenogenesis." Int Microbiol 6(1): 11-16 %U http://view.ncbi.nlm.nih.gov/pubmed/12730708.

Tatsuzawa, H., T. Maruyama, N. Misawa, K. Fujimori, M. Nakano. (2000). Quenching of singlet oxygen by carotenoids produced in Escherichia coli - attenuation of singlet oxygen-mediated bacterial killing by carotenoids. FEBS Letters 484: 280-284.

Tusher, V.G., R. Tibshirani, G. Chu. (2001). ed/12730708"crobial carote microarrays applied to the ionizing radiation response."crobial carot. Sci USA 98:5116-5121.

Summary and general discussion

Summary and general conclusion

Archaea are widely spread throughout nature and well-known for their extreme lifestyles. Their industrial potential is high and diverse. Heat-loving Archaea have the advantage that they are safe to work with (no known organism of this group can be directly linked to a fatal disease), and their biomolecules are extremely stable. The best known example of a successful application concerns the use of the thermostable DNA polymerases of *Pyrococcus* and *Thermococcus* species in DNA sequencing and amplification by the polymerase chain reaction (PCR) (Vander Horn, Davis et al. 1997; Biles and Connolly 2004). Other thermostable proteins that can be used in industry are cellulose-degrading enzymes of *Pyrococcus* and *Sulfolobus* species to generate glucose as substrate in a wide range of biotechnology fermentations (Allen 1976), xylan-degrading enzymes of *Pyrodictium* and *Thermococcus* species enzymes in the paper industry (Linko, Honkavaara et al. 1984; Oksanen, Pere et al. 2000), chitin-degrading enzymes of *Pyrococcus* and *Thermoccus* species that can be used to generate building blocks for the chemical industry (Imanaka, Fukui et al. 2001), and proteolytic enzymes of *Pyrococcus* and *Thermoccus* species for the detergent industry (Antranikian, Vorgias et al. 2005). Non-protein thermostable biomolecules of Archaeal origin are also of interest: lipids as potential drug delivery systems (Patel and Sprott 1999; Li, Chen et al. 2010), compatible solutes as stabilizers for biomolecules (Santos and da Costa 2001), and antibiotics produced by a *Sulfolobus* strain as potential leads for anti-microbial agents (O'Connor and Shand 2002). Moreover, chaperons and chaperonins serve as protein stabilizers (Ideno, Furutani et al. 2004; Maruyama, Suzuki et al. 2004). The complete cells of Archaea can also be used, for example in the mining industry where bioleaching at temperatures above 65^{o}C is exclusively done by Archaea. For example *Sulfolobus* cells are used to dissolve sulphur-containing metals to be extracted (Chaban, Ng et al. 2006) as they can leach chalcopyrite at a rate of 11.5 mg Cu $L^{-1} \cdot h^{-1}$, with an acquired tolerance of 27 g Cu L^{-1} (Umrania, 2003). An overview of these and other salient features of Archaea as well as their applications is presented in **Chapter 1**. This Chapter also presents a strong focus on the hyperthermophilic Crenarchaeon *Sulfolobus solfataricus* and its close relatives. Finally, it illustrates why Archaea are considered the most extreme branch on the tree of life and why heat-loving organisms are of interest to industry.

Sulfolobus as Model Organism

Although there are several examples of industrial applications, the full potential of Archaea has not yet been captured. This is mainly due to the lack of knowledge and experience with these organisms as is manifested by the low number of genetic tools, limited insight in the production of (heterologous) proteins and insufficient know-how on biomass formation. Hence, much research is focusing on these fields of interest. Most progress has been obtained for the Sulfolobales and this has led to the development of tools for genetic manipulation of *S. solfataricus*, *S. acidiocaldarius* and *S. islandicus*. This has allowed for the generation of gene-knockouts and heterologous gene expression systems for *S. solfataricus* (Schelert 2004, Zhang 2010, (Albers, Jonuscheit et al. 2006; Wang, Duan et al. 2007; Albers and Driessen 2008). For the genetically more stable S. acidocaldarius shuttle vectors have been recently developed (Berkner and Lipps 2008).

The physiological functions and mode of actions of different biomolecules are of continuous interest and a prerequisite to fully understand and appreciate the potential of Archaea and their molecules. We chose to study *Sulfolobus solfataricus* for its stable (heat-resistant) enzymes and specific metabolic potential, the ease of cultivation of this organism, and the relative large amount of knowledge about this heat-loving acidophilic organism. We selected a systems approach to study the behaviour of this organism trying to make steps forward into the unknown, whenever possible trying to link exploration to exploitation. The cultivation of *S.solfataricus* is an essential element in all systems approaches that link genotype to phenotype. Hence, specific attention is given to the advanced culturing systems for this extremophile that have been used in all experimental studies described here (**Chapters 3-6**).

Cultivation of *Sulfolobus solfataricus*: From Test Tube to Erlenmeyer to Bioreactor

The Crenarchaeal *Sulfolobus solfataricus* is a relevant model organism that grows aerobically at 80 ^{o}C and at a pH range of 2-5. Because of these extreme conditions there are minimal problems with contaminations and therefore it is relatively easy to achieve long term steady state operation in a bioreactor. The fact that it grows aerobically also increases the ease of maintaining long term steady

states. It has a doubling time of six hours, which is slow compared to *E. coli* or yeast, but fast in comparison to many Archaeal organisms.

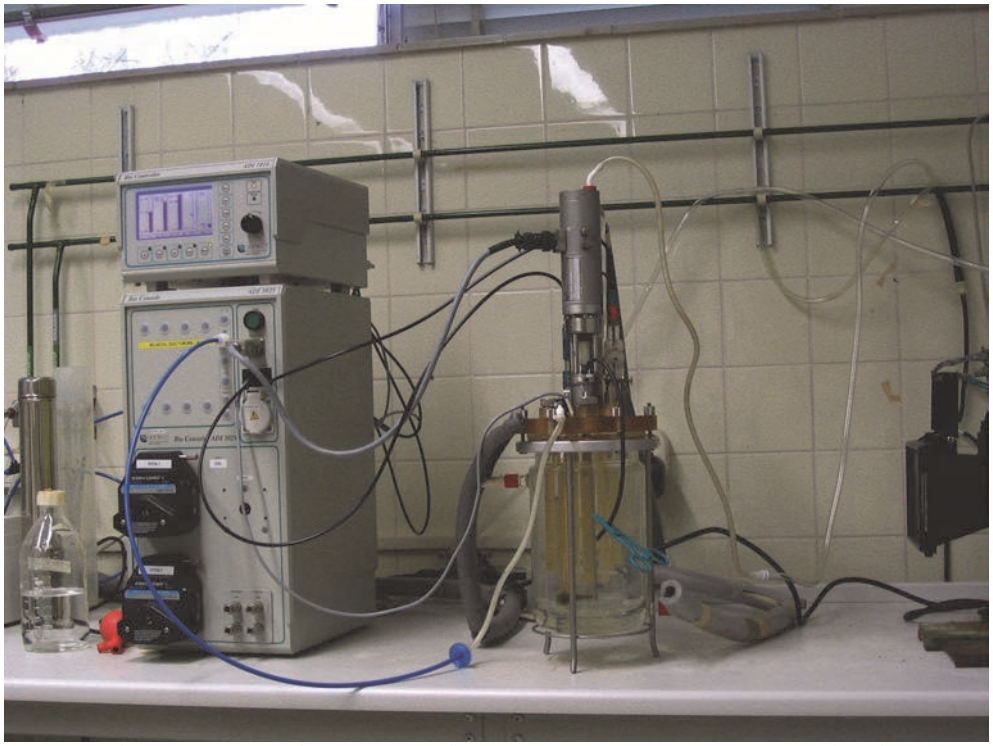

Figure 1: A fermentor suitable for cultivation of acidophilic thermophiles. The outside of the fermentor pot is heated with hot water linked to a water bath. Air can be pumped in via the blue tube and is released via a condenser. The pH, temperature and the oxygen level in the culture are continuously measured. Via the black tubing in the front, samples can be taken very quickly at any desired height in the fermentor.

To allow high temperature cultivation of *S. solfataricus* some adaptations are required in comparison of the growth of mesophilic organisms *E. coli*. For the growth of small cultures (20-50 mL) a water bath was exchanged for an oil bath to avoid rapid evaporation. For very small volumes (up to 5 mL) test tubes with loosely placed metal stoppers can be used. For larger volumes Erlenmeyer flasks can be utilized in which elongated necks ensure a profound reduction of the evaporation of the medium in the Erlenmeyer; the long neck allows for condensation of evaporated water, resulting in a rather stable volume of the culture medium. The loosely placed metal stoppers are used to maintain an aerobic head

space. The volume range for this type of batch growth is 20 mL to 2 L (for the volumes above 50 mL a large 1-3 L Erlenmeyer is used, with baffles to ensure proper aeration of the medium). To allow for better control of relevant parameters, a more sophisticated bioreactor has been constructed. The parameters of interest for increased control are substrate concentration, oxygen concentration, temperature and pH. Control of these parameters ensures reproducible growth and reduced "noise" in the data, which is especially important for sensitive analyses such as transcriptomics (Breitling 2006; van der Veen, Oliveira et al. 2009).

Bioreactors made for mesophilic organisms like *E. coli* are not well suited for extremophilic organisms that grow optimally at high temperatures and low pH because of the quick corrosion of the metal parts when in contact with the medium (Schiraldi, Acone et al. 2001). This includes the stirrer and the headplate; moreover, the acid medium might eventually result in damage of the stirrer-engine if not well protected. Therefore a highly robust bioreactor had to be developed to allow cultivation of *Sulfolobus* species.

At the start of this research project there were few commercial bioreactors available that are suited to the extreme cultivation of *S. solfataricus*. The reactor that was available (Hezayen, Rehm et al. 2000) was expensive, not easy to clean, and prone to damage and corrosion of parts. Therefore we decided to design our own bioreactor for the use in lab-research (Fig. 1). This was based on an Applicon 2 L bioreactor with the same double walled glass reactor vessel ensuring a stable and easy obtainable high temperature, as well as the same engine for the stirrer. However, the stirrer itself and the head plate design was different due to our own needs. The selected material was polysulfone: a rigid, robust and transparent material retaining its properties between -100 ^{o}C and 150 ^{o}C with a glass transition temperature of 185 ^{o}C. It has a very high dimensional stability, the size change when exposed to boiling water is less than 0.1%, but most importantly polysulfone is highly resistant to many chemicals in a pH range of 2-13 and it can be cleaned and sterilized by bleach. The disadvantage of this material is that it obviously cannot be welded and due to its operating environment glue could not be used. We solved this problem by making as many parts as possible from one plate or using tightly fitting pins to ensure a stable construction. The two main probes we used to measure the pH and the oxygen concentration, each had specific difficulties. The heat was the main problem for the probes and the low amount of oxygen implied that accurate oxygen measurements were difficult. Furthermore, the metal oxygen

probe appeared not to be fully resistant to the operating environment. The pH probe was made of glass and thus resistant to the operating conditions; however, evaporation of the buffer solution inside the probe caused problems. With regular maintenance and check-ups the probes functioned well, although the half-life of these probes at high temperatures is relatively short. Despite these problems we were able to design and operate a bioreactor that was easy to use with more easily replaceable parts and most importantly it resulted in reproducible growth (**see Chapters 3-6**).

Towards systems analysis in archaea – starting with transcriptomics

Systems analysis includes the integration of all available omics data and is increasingly used in the analysis of Archaea (**Chapters 3 and 4**). However, most attention has been given to archaeal transcriptome analysis and hence the most important literature on heat-loving Archaea is summarized (**Chapter 2**). DNA microarray systems are an important tool to monitor the expression of the complete set of RNAs in biological systems. Most microarray systems allow for a holistic analysis of gene expression, which makes it easier to study transcription regulation processes, and as such to establish regulons; the latter may result in the discovery of physiological links of genes and their products. An alternative tool to study mRNA expression concerns RT-PCR. Although the data obtained by RT-PCR generally has a higher resolution then the data obtained by microarrays, this method is only suited for the analysis of a limited set of genes. The presented review (**Chapter 2**) includes the study of different Archaea, both anaerobic and aerobic. It briefly describes the first successful attempt to study the behaviour of a heat loving Archaeon (*Pyrococcus furiosus*), which included only 271 genes. After this pioneering study, the field moved rapidly to complete transcriptome analyses in different thermophilic Archaea. An overview is provided in which different studies are grouped on the basis of the research topic: heat shock, cold shock, viral infections, cell cycle studies and metabolic studies. Moreover, attention is given to the increasing quality of transcriptome data due to technical improvements as well as optimization of related statistical analyses. Finally, new approaches are discussed that are based on next generation sequencing of cDNA derived from the complete RNA populations of the species.

Towards systems analysis in archaea – the metabolism of *Sulfolbus*

In the experimental chapters (**Chapters 3-6**) various systems approaches are applied to gain understanding of metabolic pathways in *Sulfolobus*. **Chapter 3** describes the study of the central carbon pathways, consisting of the (non-) phosphorilated Entner-Douderoff (ED) pathway and the citric acid cycle. Different functional genomic approaches were applied on the model organism *Sulfolobus solfataricus* to study the response of growth on different carbon sources, D-Glucose vs. Tryptone and Yeast Extract. The complete transcriptome was studied using PCR-based microarrays. In addition the proteome was studied using 2D-electrophoresis map in combination with ^{13}N- labelling technique to determine protein fluctuations. Despite the large difference in medium, very few significant differences on protein or RNA level were observed for the two conditions. Therefore regulation of these pathways does in all probability not occur through changes in protein abundance but presumably rather by direct changes in enzyme activity. This is unlike two thermophilic Euryarchaea: *Thermococcus kodaaraensis* (Kanai, Akerboom et al. 2007) and *Pyrococcus furiosus* (Schut, Brehm et al. 2003) where extensive regulation of glycolytic genes was observed in a similar situation.

Chapter 4 describes the study of the degradation of D-arabinose through a similar approach as was described in chapter 3. *S. solfataricus* was grown on either D-arabinose or D-glucose and a comprehensive transcriptome and proteome study was carried out. The result of these studies was not only elucidation of the D-arabinose degradation route, but also a general prokaryotic pentose, hexaric acids and hydroxyproline degradation route, which supports the theory of metabolic pathway genesis by enzyme recruitment. Also this study predicted a cis-regulatory element to induce the arabinose degrading pathway when needed. The enzymes involved in the proposed pathway were cloned, expressed and their function was biochemically measured. This showed that using these enzymes, D-arabinose can be degraded to 2-oxogluterate, one of the metabolites that are part of the citric acid cycle.

Chapter 5 reports on the effects of different oxygen concentrations on the behaviour of *Sulfolobus solfataricus*. The oxygen amount can be controlled relatively easily in a bioreactor, which is crucial for rapid and reproducible growth. Based on growth experiments in microcosms, different types of behaviour could be seen. At 35% (v/v gas phase) the cultures did not grow, indicating that *S. solfataricus* experiences a lethal dose of oxygen. At 26-32% growth was impaired,

suggesting a moderate toxicity compared to the reference (21%). In the ranges 16-24% of oxygen, standard growth was observed, suggesting that *S. solfataricus* is comfortable in these oxygen ranges. For the lower amounts of oxygen (1.5-15%), the growth was comparable to the reference, but the respiratoryefficiency was increased. To get some more insight into this behaviour, we looked at the transcriptome. It showed differential expression of several genes, including genes encoding terminal oxidases, indicating that the organism adapts to lower oxygen concentrations by adapting its respiratory machinery.

Chapter 6 describes the zeaxanthin pathway in the *Sulfolobus* species. Zeaxanthin is a colorant and of vital importance for the function of the human eye. In this chapter the genes responsible for zeaxanthin production are presented. For this, DNA microarrays, bioinformatics as well as molecular genetics techniques were used. A *crtx*-like gene is operational in most of the known *Sulfolobus* species that is able to attach sugar-like molecules to zeaxanthin, which improves its solubility in water, which is very important in many food uses. We have cloned this *crtx*-like gene of *S. solfataricus*, *S. shibatae*, and *S. acidocaldarius* in a zeaxanthin overproducing *E. coli* strain. It has been demonstrated that the gene products of *S. shibatae* and *S. acidocaldarius* were responsible for attaching sugar-like molecules to zeaxanthin. The *ctrx*-like gene of *S. solfataricus* was not operating in *E. coli*. This is probably due to the fact that the gene is truncated. This chapter has further improved the understanding of archaeal carotenoid pathways and it has shown that the *Sulfolobus* species are able to modify zeaxanthin, although each species produces different zeaxanthin modifications.

Systems biology of *Sulfolobus*

During this research we have tried to integrate different holistic data sets (genomic, transcriptomic and proteomic) to elucidate metabolic processes (**chapter 3 and 4**). Showing that these kind of approaches are synergistic. A step further in the comprehension of the cellular processes in *S. solfataricus* would be to include data on the metabolome and its use in genome-based modelling.

The main focus of systems biology is to elucidate the complex cellular processes in order to predict cellular phenotypes via a mathematical model. In order to do this, systems biology uses a holistic approach towards understanding the interconnected processes inside the cell and the regulation thereof. It uses

different (holistic) approaches: genomic sequencing; transcriptomics; proteomics; metabolomics and in rare cases systems microscopy (Lock and Stromblad 2010), but also specific biochemical data. For unicellular models yeast is the organism of choice for system biologists, and the field of Archaea is unfortunately largely forgotten. Therefore a collaborative project, of different European Universities, has been set up to model the central carbon pathway in *Sulfolobus solfataricus* (Albers, Birkeland et al. 2009), one of the first Archaeal systems biology projects. Here genomic, transcriptomic, proteomic, metabolomic, kinetic and biochemical information is integrated. It is expected that this system biology project will not only help understanding the mechanism that control the central metabolism of *S. solfataricus*, but also help give a more profound understanding of gene regulation in (Cren)archaea.

References

Adams, M. W. W., J. F. Holden, et al. (2001). "Key role for sulfur in peptide metabolism and in regulation of three hydrogenases in the hyperthermophilic archaeon Pyrococcus furiosus." Journal of Bacteriology **183**(2): 716-724.

Ahmed, H., T. J. G. Ettema, et al. (2005). "The Semi-Phosphorylative Entner-Doudoroff Pathway in Hyperthermophilic Archaea - a Re-Evaluation." Biochem J.

Albers, S. V., N. K. Birkeland, et al. (2009). "SulfoSYS (Sulfolobus Systems Biology): towards a silicon cell model for the central carbohydrate metabolism of the archaeon Sulfolobus solfataricus under temperature variation." Biochem Soc Trans **37**(Pt 1): 58-64.

Albers, S. V. and A. J. Driessen (2008). "Conditions for gene disruption by homologous recombination of exogenous DNA into the Sulfolobus solfataricus genome." Archaea **2**(3): 145-149.

Albers, S. V., M. Jonuscheit, et al. (2006). "Production of recombinant and tagged proteins in the hyperthermophilic archaeon Sulfolobus solfataricus." Appl Environ Microbiol **72**(1): 102-111.

Allen, W. G. (1976). "Potential applications for cellulase enzymes." Biotechnol Bioeng Symp(6): 303-305.

Allers, T. and M. Mevarech (2005). "Archaeal genetics - the third way." Nat Rev Genet **6**(1): 58-73.

Altschul, S. F., T. L. Madden, et al. (1997). "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." Nucleic Acids Res **25**(17): 3389-3402.

Anderson, R. L. and A. S. Dahms (1975). "2-Keto-3-deoxy-l-arabonate aldolase." Methods Enzymol **42**: 269-272.

Andersson, A., R. Bernander, et al. (2005). "Dual-genome primer design for construction of DNA microarrays." Bioinformatics **21**(3): 325-332.

Andersson, A. F., M. Lundgren, et al. (2006). "Global analysis of mRNA stability in the archaeon Sulfolobus." Genome Biol **7**(10): R99.

Andersson, A. F., E. A. Pelve, et al. (2010). "Replication-biased genome organisation in the crenarchaeon Sulfolobus." BMC Genomics **11**: 454.

Angelov, A. and W. Liebl (2006). "Insights into extreme thermoacidophily based on genome analysis of Picrophilus torridus and other thermoacidophilic archaea." J Biotechnol **126**(1): 3-10.

Antranikian, G., C. E. Vorgias, et al. (2005). "Extreme environments as a resource for microorganisms and novel biocatalysts." Adv Biochem Eng Biotechnol **96**: 219-262.

Auernik, K. S. and R. M. Kelly (2008). "Identification of components of electron transport chains in the extremely thermoacidophilic crenarchaeon Metallosphaera sedula through iron and sulfur compound oxidation transcriptomes." Appl Environ Microbiol **74**(24): 7723-7732.

Auernik, K. S. and R. M. Kelly (2010). "Impact of molecular hydrogen on chalcopyrite bioleaching by the extremely thermoacidophilic archaeon Metallosphaera sedula." Appl Environ Microbiol **76**(8): 2668-2672.

Auernik, K. S. and R. M. Kelly (2010). "Physiological versatility of the extremely thermoacidophilic archaeon Metallosphaera sedula supported by transcriptomic analysis of heterotrophic, autotrophic, and mixotrophic growth." Appl Environ Microbiol **76**(3): 931-935.

Baliga, N. S., M. Pan, et al. (2002). "Coordinate regulation of energy transduction modules in Halobacterium sp. analyzed by a global systems approach." Proc Natl Acad Sci U S A **99**(23): 14913-149188.

Bartolucci, S., R. Rella, et al. (1987). "Malic enzyme from archaebacterium Sulfolobus solfataricus. Purification, structure, and kinetic properties." J Biol Chem **262**(16): 7725-7731.

Baumann, A., C. Lange, et al. (2007). "Transcriptome changes and cAMP oscillations in an archaeal cell cycle." BMC Cell Biol **8**: 21.

Bennett, S. (2004). "Solexa Ltd." Pharmacogenomics **5**(4): 433-438.

Berkner, S., D. Grogan, et al. (2007). "Small multicopy, non-integrative shuttle vectors based on the plasmid pRN1 for Sulfolobus acidocaldarius and Sulfolobus solfataricus, model organisms of the (cren-)archaea." Nucleic Acids Res **35**(12): e88.

Berkner, S., A. Wlodkowski, et al. (2010). "Inducible and constitutive promoters for genetic systems in Sulfolobus acidocaldarius." Extremophiles **14**(3): 249-259.

Biles, B. D. and B. A. Connolly (2004). "Low-fidelity Pyrococcus furiosus DNA polymerase mutants useful in error-prone PCR." Nucleic Acids Res **32**(22): e176.

Bize, A., E. A. Karlsson, et al. (2009). "A unique virus release mechanism in the Archaea." Proc Natl Acad Sci U S A **106**(27): 11306-11311.

Bolhuis, H., E. M. T. Poele, et al. (2004). "Isolation and cultivation of Walsby's square archaeon." Environmental Microbiology **6**(12): 1287-1291.

Bonneau, R., M. T. Facciotti, et al. (2007). "A predictive model for transcriptional control of physiology in a free living cell." Cell **131**(7): 1354-1365.

Boone, D. R., R. W. Castenholz, et al. (2001). Bergey's manual of systematic bacteriology / George M. Garrity, editor-in-chief. New York, Springer.

Boonyaratanakornkit, B., J. Cordova, et al. (2006). "Pressure affects transcription profiles of Methanocaldococcus jannaschii despite the absence of barophilic growth under gas-transfer limitation." Environ Microbiol **8**(11): 2031-2035.

Boonyaratanakornkit, B. B., A. J. Simpson, et al. (2005). "Transcriptional profiling of the hyperthermophilic methanarchaeon Methanococcus jannaschii in response to lethal heat and non-lethal cold shock." Environ Microbiol **7**(6): 789-797.

Booth, I. R. (2007). "SysMO: back to the future." Nat Rev Microbiol **5**(8): 566.

Brazma, A., P. Hingamp, et al. (2001). "Minimum information about a microarray experiment (MIAME)-toward standards for microarray data." Nat Genet **29**(4): 365-371.

Breitling, R. (2006). "Biological microarray interpretation: the rules of engagement." Biochim Biophys Acta **1759**(7): 319-327.

Brinkman, A. B., S. D. Bell, et al. (2002). "The Sulfolobus solfataricus Lrp-like protein LysM regulates lysine biosynthesis in response to lysine availability." J Biol Chem **277**(33): 29537-29549.

Brock, T. D., K. M. Brock, et al. (1972). "Sulfolobus: a new genus of sulfur-oxidizing bacteria living at low pH and high temperature." Arch Mikrobiol **84**(1): 54-68.

Brock, T. D., K. M. Brock, et al. (1972). "Sulfolobus: a new genus of sulfur-oxidizing bacteria living at low pH and high temperature." Archiv Für Mikrobiologie **84**(1): 54-68.

Brock, T. D. and H. Freeze (1969). "Thermus aquaticus gen. n. and sp. n., a nonsporulating extreme thermophile." J Bacteriol **98**(1): 289-297.

Brouns, S. J., J. Walther, et al. (2006). "Identification of the missing links in prokaryotic pentose oxidation pathways: evidence for enzyme recruitment." J Biol Chem **281**(37): 27378-27388.

Bruins, M., A. Janssen, et al. (2001). "Thermozymes and their applications." Applied Biochemistry and Biotechnology **90**(2): 155-186.

Brumfield, S. K., A. C. Ortmann, et al. (2009). "Particle assembly and ultrastructural features associated with replication of the lytic archaeal virus sulfolobus turreted icosahedral virus." J Virol **83**(12): 5964-5970.

Brunner, N. A., H. Brinkmann, et al. (1998). "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." J Biol Chem **273**(11): 6149-6156.

Brusa, T., E. Canzi, et al. (1993). "Methanogens in the human intestinal tract and oral cavity." Current Microbiology **27**(5): 261-265.

Buchanan, C. L., H. Connaris, et al. (1999). "An extremely thermostable aldolase from Sulfolobus solfataricus with specificity for non-phosphorylated substrates." Biochem J **343 Pt 3**: 563-570.

Bult, C. J., O. White, et al. (1996). "Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii." Science **273**(5278): 1058-1073.

Camacho, M. L., R. A. Brown, et al. (1995). "Isocitrate dehydrogenases from Haloferax volcanii and Sulfolobus solfataricus: enzyme purification, characterisation and N-terminal sequence." FEMS Microbiol Lett **134**(1): 85-90.

Cambillau, C. and J. M. Claverie (2000). "Structural and genomic correlates of hyperthermostability." J Biol Chem **275**(42): 32383-32386.

Campanaro, S., T. J. Williams, et al. (2010). "Temperature-dependent global gene expression in the Antarctic archaeon Methanococcoides burtonii." Environ Microbiol.

Cannio, R., P. Contursi, et al. (2001). "Thermoadaptation of a mesophilic hygromycin B phosphotransferase by directed evolution in hyperthermophilic Archaea: selection of a stable genetic marker for DNA transfer into Sulfolobus solfataricus." Extremophiles **5**(3): 153-159.

Chaban, B., S. Y. Ng, et al. (2006). "Archaeal habitats--from the extreme to the ordinary." Can J Microbiol **52**(2): 73-116.

Chen, L., K. Brugger, et al. (2005). "The genome of Sulfolobus acidocaldarius, a model organism of the Crenarchaeota." J Bacteriol **187**(14): 4992-4999.

Ciccarelli, F. D., T. Doerks, et al. (2006). "Toward automatic reconstruction of a highly resolved tree of life." Science (New York, N.Y.) **311**(5765): 1283-1287.

Coker, J. A., P. Dassarma, et al. (2007). "Transcriptional profiling of the model Archaeon Halobacterium sp. NRC-1: responses to changes in salinity and temperature." Saline Systems **3**: 6.

Colombo, S., M. Grisa, et al. (1994). "Molecular cloning, nucleotide sequence and expression of a Sulfolobus solfataricus gene encoding a class II fumarase." FEBS Lett **337**(1): 93-98.

Contursi, P., R. Cannio, et al. (2003). "Development of a genetic system for hyperthermophilic Archaea: expression of a moderate thermophilic bacterial alcohol dehydrogenase gene in Sulfolobus solfataricus." FEMS Microbiol Lett **218**(1): 115-120.

Contursi, P., F. M. Pisani, et al. (2004). "Identification and autonomous replication capability of a chromosomal replication origin from the archaeon Sulfolobus solfataricus." Extremophiles **8**(5): 385-391.

Cooper, C. R., A. J. Daugherty, et al. (2009). "Role of vapBC toxin-antitoxin loci in the thermal stress response of Sulfolobus solfataricus." Biochem Soc Trans **37**(Pt 1): 123-126.

Cozen, A. E., M. T. Weirauch, et al. (2009). "Transcriptional map of respiratory versatility in the hyperthermophilic crenarchaeon Pyrobaculum aerophilum." J Bacteriol **191**(3): 782-794.

Crabb, W. D. and J. K. Shetty (1999). "Commodity scale production of sugars from starches." Curr Opin Microbiol **2**(3): 252-256.

Culley, D. E., W. P. Kovacik, Jr., et al. (2006). "Optimization of RNA isolation from the archaebacterium Methanosarcina barkeri and validation for oligonucleotide microarray analysis." J Microbiol Methods **67**(1): 36-43.

Dambeck, M. and J. Soppa (2008). "Characterization of a Haloferax volcanii member of the enolase superfamily: deletion mutant construction, expression analysis, and transcriptome comparison." Arch Microbiol **190**(3): 341-353.

Danson, M. J. (1988). "Archaebacteria: the comparative enzymology of their central metabolic pathways." Adv Microb Physiol **29**: 165-231.

Danson, M. J., S. C. Black, et al. (1985). "Citric acid cycle enzymes of the archaebacteria: citrate synthase and succinate thiokinase." FEBS **179**(1): 120-124.

De Rosa, M., A. Gambacorta, et al. (1984). "Glucose metabolism in the extreme thermoacidophilic archaebacterium Sulfolobus solfataricus." Biochem J **224**(2): 407-414.

DeLong, E. F. and N. R. Pace (2001). "Environmental diversity of bacteria and archaea." Systematic Biology **50**(4): 470-478.

Doerks, T., C. von Mering, et al. (2004). "Functional clues for hypothetical proteins based on genomic context analysis in prokaryotes." Nucleic Acids Res **32**(21): 6321-6326.

Dorazi, R., D. Gotz, et al. (2007). "Equal rates of repair of DNA photoproducts in transcribed and non-transcribed strands in Sulfolobus solfataricus." Mol Microbiol **63**(2): 521-529.

Duggin, I. G., S. A. McCallum, et al. (2008). "Chromosome replication dynamics in the archaeon Sulfolobus acidocaldarius." Proc Natl Acad Sci U S A **105**(43): 16737-16742.

Elferink, M. G., S. V. Albers, et al. (2001). "Sugar transport in Sulfolobus solfataricus is mediated by two families of binding protein-dependent ABC transporters." Mol Microbiol **39**(6): 1494-1503.

Elkins, J. G., M. Podar, et al. (2008). "A korarchaeal genome reveals insights into the evolution of the Archaea." Proc Natl Acad Sci U S A **105**(23): 8102-8107.

Ettema, T., J. van der Oost, et al. (2001). "Modularity in the gain and loss of genes: applications for function prediction." Trends Genet **17**(9): 485-487.

Ettema, T. J. and R. Bernander (2009). "Cell division and the ESCRT complex: A surprise from the archaea." Commun Integr Biol **2**(2): 86-88.

Ettema, T. J., W. M. de Vos, et al. (2005). "Discovering novel biology by in silico archaeology." Nat Rev Microbiol **3**(11): 859-869.

Ettema, T. J., K. S. Makarova, et al. (2004). "Identification and functional verification of archaeal-type phosphoenolpyruvate carboxylase, a missing link in archaeal central carbohydrate metabolism." J Bacteriol **186**(22): 7754-7762.

Facciotti, M. T., W. L. Pang, et al. (2010). "Large scale physiological readjustment during growth enables rapid, comprehensive and inexpensive systems analysis." BMC Syst Biol **4**: 64.

Ferry, J. G. and D. J. Lessner (2008). "Methanogenesis in marine sediments." Ann N Y Acad Sci **1125**: 147-157.

Franzmann, P. D., Y. Liu, et al. (1997). "Methanogenium frigidum sp. nov., a psychrophilic, H2-using methanogen from Ace Lake, Antarctica." International Journal of Systematic Bacteriology **47**(4): 1068-1072.

Fricke, W. F., H. Seedorf, et al. (2006). "The genome sequence of Methanosphaera stadtmanae reveals why this human intestinal archaeon is restricted to methanol and H2 for methane formation and ATP synthesis." Journal of Bacteriology **188**(2): 642-658.

Frickey, T. and A. N. Lupas (2004). "PhyloGenie: automated phylome generation and analysis." Nucleic Acids Res **32**(17): 5231-5238.

Fröls, S., P. M. Gordon, et al. (2007). "Response of the hyperthermophilic archaeon Sulfolobus solfataricus to UV damage." J Bacteriol **189**(23): 8708-8718.

Fukuda, E. and T. Wakagi (2002). "Substrate recognition by 2-oxoacid:ferredoxin oxidoreductase from Sulfolobus sp. strain 7." Biochim Biophys Acta **1597**(1): 74-80.

Fukuda, W., T. Fukui, et al. (2004). "First characterization of an archaeal GTP-dependent phosphoenolpyruvate carboxykinase from the hyperthermophilic archaeon Thermococcus kodakaraensis KOD1." J Bacteriol **186**(14): 4620-4627.

Garrido, P., E. Gonzalez-Toril, et al. (2008). "An oligonucleotide prokaryotic acidophile microarray: its validation and its use to monitor seasonal variations in extreme acidic environments with total environmental RNA." Environ Microbiol **10**(4): 836-850.

Gilbert, J. A., D. Field, et al. (2008). "Detection of large numbers of novel sequences in the metatranscriptomes of complex marine microbial communities." PLoS One **3**(8): e3042.

Giometti, C. S., C. Reich, et al. (2002). "Global analysis of a "simple" proteome: Methanococcus jannaschii." J Chromatogr B Analyt Technol Biomed Life Sci **782**(1-2): 227-243.

Goberna, M., M. Gadermaier, et al. (2010). "Adaptation of methanogenic communities to the cofermentation of cattle excreta and olive mill wastes at 37 degrees C and 55 degrees C." Appl Environ Microbiol **76**(19): 6564-6571.

González, J. M., Y. Masuchi, et al. (1998). "Pyrococcus horikoshii sp. nov., a hyperthermophilic archaeon isolated from a hydrothermal vent at the Okinawa Trough." Extremophiles: Life Under Extreme Conditions **2**(2): 123-130.

Gotz, D., S. Paytubi, et al. (2007). "Responses of hyperthermophilic crenarchaea to UV irradiation." Genome Biol **8**(10): R220.

Grogan, D. W. (1989). "Phenotypic characterization of the archaebacterial genus Sulfolobus: comparison of five wild-type strains." J Bacteriol **171**(12): 6710-6719.

Grogan, D. W. (2000). "The question of DNA repair in hyperthermophilic archaea." Trends in Microbiology **8**(4): 180-185.

Grogan, D. W., M. A. Ozarzak, et al. (2008). "Variation in gene content among geographically diverse Sulfolobus isolates." Environ Microbiol **10**(1): 137-146.

Hamilton-Brehm, S. D., G. J. Schut, et al. (2005). "Metabolic and evolutionary relationships among Pyrococcus Species: genetic exchange within a hydrothermal vent environment." J Bacteriol **187**(21): 7492-7499.

Hansen, T., D. Wendorff, et al. (2004). "Bifunctional phosphoglucose/phosphomannose isomerases from the Archaea Aeropyrum pernix and Thermoplasma acidophilum constitute a novel enzyme family within the phosphoglucose isomerase superfamily." J Biol Chem **279**(3): 2262-2272.

Hartl, T., W. Grossebuter, et al. (1987). "Crystalline NAD/NADP-dependent malate dehydrogenase; the enzyme from the thermoacidophilic archaebacterium Sulfolobus acidocaldarius." Biol Chem Hoppe Seyler **368**(3): 259-267.

Hemmi, H., S. Ikejiri, et al. (2003). "Fusion-type lycopene [beta]-cyclase from a thermoacidophilic archaeon Sulfolobus solfataricus." Biochemical and biophysical research communications **305**(3): 586-591.

Hendrickson, E. L., A. K. Haydock, et al. (2007). "Functionally distinct genes regulated by hydrogen limitation and growth rate in methanogenic Archaea." Proc Natl Acad Sci U S A **104**(21): 8930-8934.

Hess, D., K. Kruger, et al. (1995). "Dimeric 3-phosphoglycerate kinases from hyperthermophilic Archaea. Cloning, sequencing and expression of the 3-phosphoglycerate kinase gene of Pyrococcus woesei in Escherichia coli and characterization of the protein. Structural and functional comparison with the 3-phosphoglycerate kinase of Methanothermus fervidus." Eur J Biochem **233**(1): 227-237.

Hezayen, F. F., B. H. Rehm, et al. (2000). "Polymer production by two newly isolated extremely halophilic archaea: application of a novel corrosion-resistant bioreactor." Appl Microbiol Biotechnol **54**(3): 319-325.

Hovey, R., S. Lentes, et al. (2005). "DNA microarray analysis of Methanosarcina mazei Go1 reveals adaptation to different methanogenic substrates." Mol Genet Genomics **273**(3): 225-239.

Huber, H., M. J. Hohn, et al. (2002). "A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont." Nature **417**(6884): 63-67.

Hutchins, A. M., J. F. Holden, et al. (2001). "Phosphoenolpyruvate synthetase from the hyperthermophilic archaeon Pyrococcus furiosus." J Bacteriol **183**(2): 709-715.

Huynen, M. A., T. Dandekar, et al. (1999). "Variation and evolution of the citric-acid cycle: a genomic perspective." Trends Microbiol **7**(7): 281-291.

Ideno, A., M. Furutani, et al. (2004). "Expression of foreign proteins in Escherichia coli by fusing with an archaeal FK506 binding protein." Appl Microbiol Biotechnol **64**(1): 99-105.

Imanaka, T., T. Fukui, et al. (2001). "Chitinase from Thermococcus kodakaraensis KOD1." Methods Enzymol **330**: 319-329.

Jager, D., C. M. Sharma, et al. (2009). "Deep sequencing analysis of the Methanosarcina mazei Go1 transcriptome in response to nitrogen availability." Proc Natl Acad Sci U S A **106**(51): 21878-21882.

Janssen, S., G. Schafer, et al. (1997). "A succinate dehydrogenase with novel structure and properties from the hyperthermophilic archaeon Sulfolobus acidocaldarius: genetic and biophysical characterization." J Bacteriol **179**(17): 5560-5569.

Johnsen, U., M. Dambeck, et al. (2009). "D-xylose degradation pathway in the halophilic archaeon Haloferax volcanii." J Biol Chem **284**(40): 27290-27303.

Johnson, M. R., S. B. Conners, et al. (2006). "The Thermotoga maritima phenotype is impacted by syntrophic interaction with Methanococcus jannaschii in hyperthermophilic coculture." Appl Environ Microbiol **72**(1): 811-818.

Jonuscheit, M., E. Martusewitsch, et al. (2003). "A reporter gene system for the hyperthermophilic archaeon Sulfolobus solfataricus based on a selectable and integrative shuttle vector." Mol Microbiol **48**(5): 1241-1252.

Kagawa, H. K., T. Yaoi, et al. (2003). "The composition, structure and stability of a group II chaperonin are temperature regulated in a hyperthermophilic archaeon." Mol Microbiol **48**(1): 143-156.

Kanai, T., J. Akerboom, et al. (2007). "A global transcriptional regulator in Thermococcus kodakaraensis controls the expression levels of both glycolytic and gluconeogenic enzyme-encoding genes." J Biol Chem **282**(46): 33659-33670.

Kanai, T., J. Akerboom, et al. (2007). "A global transcriptional regulator in Thermococcus kodakaraensis controls the expression levels of both glycolytic and gluconeogenic enzyme-encoding genes." J Biol Chem **282**(46): 33659-33670.

Kanai, T., S. Takedomi, et al. (2010). "Identification of the Phr-dependent heat shock regulon in the hyperthermophilic archaeon, Thermococcus kodakaraensis." J Biochem **147**(3): 361-370.

Kanehisa, M., S. Goto, et al. (2004). "The KEGG resource for deciphering the genome." Nucleic Acids Res **32**(Database issue): D277-280.

Kardinahl, S., C. L. Schmidt, et al. (1999). "The strict molybdate-dependence of glucose-degradation by the thermoacidophile Sulfolobus acidocaldarius reveals the first crenarchaeotic molybdenum containing enzyme--an aldehyde oxidoreductase." Eur J Biochem **260**(2): 540-548.

Kashefi, K. and D. R. Lovley (2003). "Extending the Upper Temperature Limit for Life." Science **301**(5635): 934.

Kawarabayasi, Y., Y. Hino, et al. (2001). "Complete genome sequence of an aerobic thermoacidophilic crenarchaeon, Sulfolobus tokodaii strain7." DNA Res **8**(4): 123-140.

Kawarabayasi, Y., Y. Hino, et al. (1999). "Complete genome sequence of an aerobic hyper-thermophilic crenarchaeon, Aeropyrum pernix K1." DNA Research: An International Journal for Rapid Publication of Reports on Genes and Genomes **6**(2): 83-101, 145-152.

Keese, A. M., G. J. Schut, et al. (2010). "Genome-wide identification of targets for the archaeal heat shock regulator phr by cell-free transcription of genomic DNA." J Bacteriol **192**(5): 1292-1298.

Kerscher, L., S. Nowitzki, et al. (1982). "Thermoacidophilic archaebacteria contain bacterial-type ferredoxins acting as electron acceptors of 2-oxoacid:ferredoxin oxidoreductases." Eur J Biochem **128**(1): 223-230.

Kim, S. and S. B. Lee (2005). "Identification and characterization of Sulfolobus solfataricus D-gluconate dehydratase: a key enzyme in the non-phosphorylated Entner-Doudoroff pathway." Biochem J **387**(Pt 1): 271-280.

Kirkpatrick, R. D. (2010). "Something old, something new, something borrowed...something blue." Tenn Med **103**(8): 7-8.

Kohlhoff, M., A. Dahm, et al. (1996). "Tetrameric triosephosphate isomerase from hyperthermophilic Archaea." FEBS Lett **383**(3): 245-250.

Koonin, E. V., Y. I. Wolf, et al. (2001). "Prediction of the archaeal exosome and its connections with the proteasome and the translation and transcription machineries by a comparative-genomic approach." Genome Res **11**(2): 240-252.

Koutsopoulos, S., J. van der Oost, et al. (2007). "Kinetically controlled refolding of a heat-denatured hyperthermostable protein." FEBS J **274**(22): 5915-5923.

Kull, D. R. and H. Pfander (1997). "Isolation and structure elucidation of carotenoid glycosides from the thermoacidophilic archaea Sulfolobus shibatae." Journal of Natural Products **60**(4): 371-374.

Kumar, S. and R. Nussinov (2001). "How do thermophilic proteins deal with heat?" Cell Mol Life Sci **58**(9): 1216-1233.

Lamble, H. J., N. I. Heyer, et al. (2003). "Metabolic pathway promiscuity in the archaeon Sulfolobus solfataricus revealed by studies on glucose dehydrogenase and 2-keto-3-deoxygluconate aldolase." J Biol Chem **278**(36): 34066-34072.

Lamble, H. J., N. I. Heyer, et al. (2003). "Metabolic pathway promiscuity in the Archaeon *Sulfolobus solfataricus* revealed by studies on glucose dehydrogenase and KDG aldolase." J Biol Chem.

Lamble, H. J., C. C. Milburn, et al. (2004). "Gluconate dehydratase from the promiscuous Entner-Doudoroff pathway in Sulfolobus solfataricus." FEBS Lett **576**(1-2): 133-136.

Lange, C., A. Zaigler, et al. (2007). "Genome-wide analysis of growth phase-dependent translational and transcriptional regulation in halophilic archaea." BMC Genomics **8**: 415.

Ledford, H. (2008). "The death of microarrays?" Nature **455**(7215): 847.

Lee, H. S., K. R. Shockley, et al. (2006). "Transcriptional and biochemical analysis of starch metabolism in the hyperthermophilic archaeon Pyrococcus furiosus." J Bacteriol **188**(6): 2115-2125.

Li, L., Q. Li, et al. (2007). "Quantitative proteomic and microarray analysis of the archaeon Methanosarcina acetivorans grown with acetate versus methanol." J Proteome Res **6**(2): 759-771.

Li, Z., J. Chen, et al. (2010). "Investigation of archaeosomes as carriers for oral delivery of peptides." Biochem Biophys Res Commun **394**(2): 412-417.

Lindström, E. B. and L. Gunneriusson (1990). "Thermophilic bioleaching of arsenopyrite usingSulfolobus and a semi-continuous laboratory procedure." Journal of Industrial Microbiology and Biotechnology **5**(6): 375-382.

Linko, K., P. Honkavaara, et al. (1984). "Heated humidification in major abdominal surgery." Eur J Anaesthesiol **1**(3): 285-291.

Lipps, G. (2006). "Plasmids and viruses of the thermoacidophilic crenarchaeote Sulfolobus." Extremophiles: Life Under Extreme Conditions **10**(1): 17-28.

Lock, J. G. and S. Stromblad (2010). "Systems microscopy: an emerging strategy for the life sciences." Exp Cell Res **316**(8): 1438-1444.

Lohlein-Werhahn, G., P. Goepfert, et al. (1988). "Purification and properties of an archaebacterial enzyme: citrate synthase from Sulfolobus solfataricus." Biol Chem Hoppe Seyler **369**(2): 109-113.

Lundgren, M., A. Andersson, et al. (2004). "Three replication origins in Sulfolobus species: synchronous initiation of chromosome replication and asynchronous termination." Proc Natl Acad Sci U S A **101**(18): 7046-7051.

Lundgren, M. and R. Bernander (2007). "Genome-wide transcription map of an archaeal cell cycle." Proc Natl Acad Sci U S A **104**(8): 2939-2944.

Maaty, W. S., A. C. Ortmann, et al. (2006). "Characterization of the archaeal thermophile Sulfolobus turreted icosahedral virus validates an evolutionary link among double-stranded DNA viruses from all domains of life." J Virol **80**(15): 7625-7635.

Maisnier-Patin, S., L. Malandrin, et al. (2002). "Chromosome replication patterns in the hyperthermophilic euryarchaea Archaeoglobus fulgidus and Methanocaldococcus (Methanococcus) jannaschii." Mol Microbiol **45**(5): 1443-1450.

Makarova, K. S. and E. V. Koonin (2003). "Comparative genomics of Archaea: how much have we learned in six years, and what's next?" Genome Biol **4**(8): 115.

Makarova, K. S., N. Yutin, et al. (2010). "Evolution of diverse cell division and vesicle formation systems in Archaea." Nat Rev Microbiol **8**(10): 731-741.

Margulies, M., M. Egholm, et al. (2005). "Genome sequencing in microfabricated high-density picolitre reactors." Nature **437**(7057): 376-380.

Marteinsson, V. T., J. L. Birrien, et al. (1999). "Thermococcus barophilus sp. nov., a new barophilic and hyperthermophilic archaeon isolated under high hydrostatic pressure from a deep-sea hydrothermal vent." International Journal of Systematic Bacteriology **49 Pt 2**: 351-359.

Maruyama, T., R. Suzuki, et al. (2004). "Archaeal peptidyl prolyl cis-trans isomerases (PPIases) update 2004." Front Biosci **9**: 1680-1720.

Matsunaga, F., A. Glatigny, et al. (2007). "Genomewide and biochemical analyses of DNA-binding activity of Cdc6/Orc1 and Mcm proteins in Pyrococcus sp." Nucleic Acids Res **35**(10): 3214-3222.

McCready, S., J. A. Muller, et al. (2005). "UV irradiation induces homologous recombination genes in the model archaeon, Halobacterium sp. NRC-1." Saline Systems **1**: 3.

Mirkin, B. G., T. I. Fenner, et al. (2003). "Algorithms for computing parsimonious evolutionary scenarios for genome evolution, the last universal common ancestor and dominance of horizontal gene transfer in the evolution of prokaryotes." BMC Evol Biol **3**: 2.

Mockler, T. C., S. Chan, et al. (2005). "Applications of DNA tiling arrays for whole-genome analysis." Genomics **85**(1): 1-15.

Mulder, N. J., R. Apweiler, et al. (2005). "InterPro, progress and status in 2005." Nucleic Acids Res **33**(Database issue): D201-205.

Muller, J. A. and S. DasSarma (2005). "Genomic analysis of anaerobic respiration in the archaeon Halobacterium sp. strain NRC-1: dimethyl sulfoxide and trimethylamine N-oxide as terminal electron acceptors." J Bacteriol **187**(5): 1659-1667.

Muralidharan, V., K. D. Rinker, et al. (1997). "Hydrogen transfer between methanogens and fermentative heterotrophs in hyperthermophilic cocultures." Biotechnol Bioeng **56**(3): 268-278.

Myllykallio, H. and P. Forterre (2000). "Mapping of a chromosome replication origin in an archaeon: response." Trends Microbiol **8**(12): 537-539.

Myllykallio, H., P. Lopez, et al. (2000). "Bacterial mode of replication with eukaryotic-like machinery in a hyperthermophilic archaeon." Science **288**(5474): 2212-2215.

Nelson, K. E., R. A. Clayton, et al. (1999). "Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of Thermotoga maritima." Nature **399**(6734): 323-329.

Nishimasu, H., S. Fushinobu, et al. (2004). "The first crystal structure of the novel class of fructose-1,6-bisphosphatase present in thermophilic archaea." Structure (Camb) **12**(6): 949-959.

Norris, P. R., N. P. Burton, et al. (2000). "Acidophiles in bioreactor mineral processing." Extremophiles **4**(2): 71-76.

Nunn, C. E., U. Johnsen, et al. (2010). "Metabolism of pentose sugars in the hyperthermophilic archaea Sulfolobus solfataricus and Sulfolobus acidocaldarius." J Biol Chem **285**(44): 33701-33709.

O'Connor, E. M. and R. F. Shand (2002). "Halocins and sulfolobicins: the emerging story of archaeal protein and peptide antibiotics." J Ind Microbiol Biotechnol **28**(1): 23-31.

Oksanen, T., J. Pere, et al. (2000). "Treatment of recycled kraft pulps with Trichoderma reesei hemicellulases and cellulases." J Biotechnol **78**(1): 39-48.

Ortmann, A. C., S. K. Brumfield, et al. (2008). "Transcriptome analysis of infection of the archaeon Sulfolobus solfataricus with Sulfolobus turreted icosahedral virus." J Virol **82**(10): 4874-4883.

Overbeek, R., N. Larsen, et al. (2003). "The ERGO genome analysis and discovery system." Nucleic Acids Res **31**(1): 164-171.

Palm, P., C. Schleper, et al. (1991). "Complete nucleotide sequence of the virus SSV1 of the archaebacterium Sulfolobus shibatae." Virology **185**(1): 242-250.

Pappin, D. J., P. Hojrup, et al. (1993). "Rapid identification of proteins by peptide-mass fingerprinting." Curr Biol **3**(6): 327-332.

Parnell, J. J., G. Rompato, et al. (2010). "Functional biogeography as evidence of gene transfer in hypersaline microbial communities." PLoS One **5**(9): e12919.

Patel, G. B. and G. D. Sprott (1999). "Archaeobacterial ether lipid liposomes (archaeosomes) as novel vaccine and drug delivery systems." Crit Rev Biotechnol **19**(4): 317-357.

Peak, M. J., J. G. Peak, et al. (1994). "The hyperthermophilic glycolytic enzyme enolase in the archaeon, Pyrococcus furiosus: comparison with mesophilic enolases." Arch Biochem Biophys **313**(2): 280-286.

Peterson, J. D., L. A. Umayam, et al. (2001). "The Comprehensive Microbial Resource." Nucleic Acids Res **29**(1): 123-125.

Pfluger, K., A. Ehrenreich, et al. (2007). "Identification of genes involved in salt adaptation in the archaeon Methanosarcina mazei Go1 using genome-wide gene expression profiling." FEMS Microbiol Lett **277**(1): 79-89.

Puchegger, S., B. Redl, et al. (1990). "Purification and properties of a thermostable fumarate hydratase from the archaeobacterium Sulfolobus solfataricus." J Gen Microbiol **136**(8): 1537-1541.

Ramsay, G. (1998). "DNA chips: state-of-the art." Nat Biotechnol **16**(1): 40-44.

Reeve, J. N. (1999). "Archaebacteria then ... Archaes now (are there really no archaeal pathogens?)." J Bacteriol **181**(12): 3613-3617.

Reichlen, M. J., K. S. Murakami, et al. (2010). "Functional analysis of the three TATA binding protein homologs in Methanosarcina acetivorans." J Bacteriol **192**(6): 1511-1517.

Rich, J. J., O. R. Dale, et al. (2008). "Anaerobic ammonium oxidation (anammox) in Chesapeake Bay sediments." Microb Ecol **55**(2): 311-320.

Rivera, M. C. and J. A. Lake (2004). "The ring of life provides evidence for a genome fusion origin of eukaryotes." Nature **431**(7005): 152-155.

Robinson, N. P. and S. D. Bell (2007). "Extrachromosomal element capture and the evolution of multiple replication origins in archaeal chromosomes." Proc Natl Acad Sci U S A **104**(14): 5806-5811.

Robinson, N. P., I. Dionne, et al. (2004). "Identification of two origins of replication in the single chromosome of the archaeon Sulfolobus solfataricus." Cell **116**(1): 25-38.

Rohmer, M., P. Bouvier, et al. (1979). "Molecular evolution of biomembranes: structural equivalents and phylogenetic precursors of sterols." Proc Natl Acad Sci U S A **76**(2): 847-851.

Ronimus, R. S. and H. W. Morgan (2003). "Distribution and phylogenies of enzymes of the Embden-Meyerhof-Parnas pathway from archaea and hyperthermophilic bacteria support a gluconeogenic origin of metabolism." Archaea **1**(3): 199-221.

Rothschild, L. J. and R. L. Mancinelli (2001). "Life in extreme environments." Nature **409**(6823): 1092-1101.

Ruepp, A., W. Graml, et al. (2000). "The genome sequence of the thermoacidophilic scavenger Thermoplasma acidophilum." Nature **407**(6803): 508-513.

Russo, A. D., R. Rullo, et al. (1995). "Glyceraldehyde-3-phosphate dehydrogenase in the hyperthermophilic archaeon Sulfolobus solfataricus: characterization and significance in glucose metabolism." Biochem Mol Biol Int **36**(1): 123-135.

Sako, Y., K. Takai, et al. (1996). "Purification and characterization of phosphoenolpyruvate carboxylase from the hyperthermophilic archaeon Methanothermus sociabilis." FEBS Lett **392**(2): 148-152.

Samson, R. Y. and S. D. Bell (2009). "Ancient ESCRTs and the evolution of binary fission." Trends Microbiol **17**(11): 507-513.

Samson, R. Y., T. Obita, et al. (2008). "A role for the ESCRT system in cell division in archaea." Science **322**(5908): 1710-1713.

Santangelo, T. J., L. Cubonova, et al. (2008). "Polarity in archaeal operon transcription in Thermococcus kodakaraensis." J Bacteriol **190**(6): 2244-2248.

Santos, H. and M. S. da Costa (2001). "Organic solutes from thermophiles and hyperthermophiles." Methods Enzymol **334**: 302-315.

Santos, H. and M. S. da Costa (2002). "Compatible solutes of organisms that live in hot saline environments." Environ Microbiol **4**(9): 501-509.

Satory, M., M. Furlinger, et al. (1997). "Continuous enzymatic production of lactobionic acid using glucose-fructose oxidoreductase in an ultrafiltration membrane reactor." Biotechnol Lett **19**(12): 1205-1208.

Schafer, G. (1996). "Bioenergetics of the archaebacterium Sulfolobus." Biochim Biophys Acta **1277**(3): 163-200.

Schafer, G. (1996). "Bioenergetics of the archaebacterium Sulfolobus." Biochim Biophys Acta **1277**(3): 163-200.

Schiraldi, C., M. Acone, et al. (2001). "Innovative fermentation strategies for the production of extremophilic enzymes." Extremophiles **5**(3): 193-198.

Schleper, C., K. Kubo, et al. (1992). "The particle SSV1 from the extremely thermophilic archaeon Sulfolobus is a virus: demonstration of infectivity and of transfection with viral DNA." Proc Natl Acad Sci U S A **89**(16): 7645-7649.

Schleper, C., G. Puehler, et al. (1995). "Picrophilus gen. nov., fam. nov.: a novel aerobic, heterotrophic, thermoacidophilic genus and family comprising archaea capable of growth around pH 0." J Bacteriol **177**(24): 7050-7059.

Schmid, A. K., D. J. Reiss, et al. (2009). "A single transcription factor regulates evolutionarily diverse but functionally linked metabolic pathways in response to nutrient availability." Mol Syst Biol **5**: 282.

Schomburg, I., A. Chang, et al. (2004). "BRENDA, the enzyme database: updates and major new developments." Nucleic Acids Res **32**(Database issue): D431-433.

Schonheit, P. and T. Schafer (1995). "Metabolism of Hyperthermophiles." World Journal of Microbiology & Biotechnology **11**(1): 26-57.

Schönheit, P. and T. Schäfer (1995). "Metabolism of hyperthermophiles." World Journal of Microbiology and Biotechnology **11**(1): 26-57.

Schramm, A., B. Siebers, et al. (2000). "Pyruvate kinase of the hyperthermophilic crenarchaeote Thermoproteus tenax: physiological role and phylogenetic aspects." J Bacteriol **182**(7): 2001-2009.

Schut, G. J., S. D. Brehm, et al. (2003). "Whole-genome DNA microarray analysis of a hyperthermophile and an archaeon: Pyrococcus furiosus grown on carbohydrates or peptides." J Bacteriol **185**(13): 3935-3947.

Schut, G. J., S. D. Brehm, et al. (2003). "Whole-genome DNA microarray analysis of a hyperthermophile and an archaeon: Pyrococcus furiosus grown on carbohydrates or peptides." J Bacteriol **185**(13): 3935-3947.

Schut, G. J., S. L. Bridger, et al. (2007). "Insights into the metabolism of elemental sulfur by the hyperthermophilic archaeon Pyrococcus furiosus: characterization of a coenzyme A- dependent NAD(P)H sulfur oxidoreductase." J Bacteriol **189**(12): 4431-4441.

Schut, G. J., J. Zhou, et al. (2001). "DNA microarray analysis of the hyperthermophilic archaeon Pyrococcus furiosus: evidence for anNew type of sulfur-reducing enzyme complex." J Bacteriol **183**(24): 7027-7036.

Schwaiger, R., C. Schwarz, et al. (2010). "Transcriptional control by two leucine-responsive regulatory proteins in Halobacterium salinarum R1." BMC Mol Biol **11**: 40.

Segerer, A. H., S. Burggraf, et al. (1993). "Life in hot springs and hydrothermal vents." Orig Life Evol Biosph **23**(1): 77-90.

Selig, M., K. B. Xavier, et al. (1997). "Comparative analysis of Embden-Meyerhof and Entner-Doudoroff glycolytic pathways in hyperthermophilic archaea and the bacterium Thermotoga." Arch Microbiol **167**(4): 217-232.

She, Q., R. K. Singh, et al. (2001). "The complete genome of the crenarchaeon Sulfolobus solfataricus P2." Proceedings of the National Academy of Sciences of the United States of America **98**(14): 7835-7840

She, Q., R. K. Singh, et al. (2001). "The complete genome of the crenarchaeon Sulfolobus solfataricus P2." Proc Natl Acad Sci U S A **98**(14): 7835-7840.

Shih, C. J. and M. C. Lai (2010). "Differentially expressed genes after hyper- and hypo-salt stress in the halophilic archaeon Methanohalophilus portucalensis." Can J Microbiol **56**(4): 295-307.

Shivvers, D. W. and T. D. Brock (1973). "Oxidation of elemental sulfur by Sulfolobus acidocaldarius." Journal of Bacteriology **114**(2): 706-710.

Shockley, K. R., D. E. Ward, et al. (2003). "Heat shock response by the hyperthermophilic archaeon Pyrococcus furiosus." Appl Environ Microbiol **69**(4): 2365-2371.

Siebers, B., H. Brinkmann, et al. (2001). "Archaeal fructose-1,6-bisphosphate aldolases constitute a new family of archaeal type class I aldolase." J Biol Chem **276**(31): 28710-28718.

Sieiro, C., M. Poza, et al. (2003). "Genetic basis of microbial carotenogenesis." Int Microbiol **6**(1): 11-16 %U http://view.ncbi.nlm.nih.gov/pubmed/12730708.

Simon, G., J. Walther, et al. (2009). "Effect of O2 concentrations on Sulfolobus solfataricus P2." FEMS Microbiol Lett **299**(2): 255-260.

Skorko, R., J. Osipiuk, et al. (1989). "Glycogen-bound polyphosphate kinase from the archaebacterium Sulfolobus acidocaldarius." J Bacteriol **171**(9): 5162-5164.

Smith, D. R., L. A. Doucette-Stamm, et al. (1997). "Complete genome sequence of Methanobacterium thermoautotrophicum deltaH: functional analysis and comparative genomics." Journal of Bacteriology **179**(22): 7135-7155.

Smith, L. D., K. J. Stevenson, et al. (1987). "Citrate synthase from the thermophilic archaebacteria Thermoplasma acidophilum and Sulfolobus acidocaldarius." FEBS Lett **225**(1,2): 277-281.

Snijders, A. P., M. G. de Vos, et al. (2005). "Novel approach for Peptide quantitation and sequencing based on (15)n and (13)c metabolic labeling." J Proteome Res **4**(2): 578-585.

Snijders, A. P., J. Walther, et al. (2006). "Reconstruction of central carbon metabolism in Sulfolobus solfataricus using a two-dimensional gel electrophoresis map, stable isotope labelling and DNA microarray analysis." Proteomics **6**(5): 1518-1529.

Snijders, A. P. L., M. G. de Vos, et al. (2005). "A fast method for quantitative proteomics based on a combination between two-dimensional electrophoresis and 15N metabolic labeling." Electrophoresis **in press**.

Solow, B., K. M. Bischoff, et al. (1998). "Archael phosphoproteins. Identification of a hexosephosphate mutase and the alpha-subunit of succinyl-CoA synthetase in the extreme acidothermophile Sulfolobus solfataricus." Protein Sci **7**(1): 105-111.

Stedman, K. M., C. Schleper, et al. (1999). "Genetic requirements for the function of the archaeal virus SSV1 in Sulfolobus solfataricus: construction and testing of viral shuttle vectors." Genetics **152**(4): 1397-1405.

Stetter, K. O. (2006). "Hyperthermophiles in the history of life." Philos Trans R Soc Lond B Biol Sci **361**(1474): 1837-1842; discussion 1842-1833.

Stralis-Pavese, N., A. Sessitsch, et al. (2004). "Optimization of diagnostic microarray for application in analysing landfill methanotroph communities under different plant covers." Environ Microbiol **6**(4): 347-363.

Strand, K. R., C. Sun, et al. (2010). "Oxidative stress protection and the repair response to hydrogen peroxide in the hyperthermophilic archaeon Pyrococcus furiosus and in related species." Arch Microbiol **192**(6): 447-459.

Suzuki, T., T. Iwasaki, et al. (2002). "Sulfolobus tokodaii sp. nov. (f. Sulfolobus sp. strain 7), a new member of the genus Sulfolobus isolated from Beppu Hot Springs, Japan." Extremophiles: Life Under Extreme Conditions **6**(1): 39-44.

Tachdjian, S. and R. M. Kelly (2006). "Dynamic metabolic adjustments and genome plasticity are implicated in the heat shock response of the extremely thermoacidophilic archaeon Sulfolobus solfataricus." J Bacteriol **188**(12): 4553-4559.

Takai, K., K. Nakamura, et al. (2008). "Cell proliferation at 122 degrees C and isotopically heavy CH4 production by a hyperthermophilic methanogen under high-pressure cultivation." Proc Natl Acad Sci U S A **105**(31): 10949-10954.

Tatusov, R. L., N. D. Fedorova, et al. (2003). "The COG database: an updated version includes eukaryotes." BMC Bioinformatics **4**(1): 41.

Torres, T. T., M. Metta, et al. (2008). "Gene expression profiling by massively parallel sequencing." Genome Res **18**(1): 172-177.

Trauger, S. A., E. Kalisak, et al. (2008). "Correlating the transcriptome, proteome, and metabolome in the environmental adaptation of a hyperthermophile." J Proteome Res **7**(3): 1027-1035.

Tusher, V. G., R. Tibshirani, et al. (2001). "Significance analysis of microarrays applied to the ionizing radiation response." Proc Natl Acad Sci U S A **98**(9): 5116-5121.

Twellmeyer, J., A. Wende, et al. (2007). "Microarray analysis in the archaeon Halobacterium salinarum strain R1." PLoS One **2**(10): e1064.

Uhrigshardt, H., M. Walden, et al. (2001). "Purification and characterization of the first archaeal aconitase from the thermoacidophilic Sulfolobus acidocaldarius." Eur J Biochem **268**(6): 1760-1771.

Uhrigshardt, H., M. Walden, et al. (2002). "Evidence for an operative glyoxylate cycle in the thermoacidophilic crenarchaeon Sulfolobus acidocaldarius." FEBS Lett **513**(2-3): 223-229.

van de Peppel, J., P. Kemmeren, et al. (2003). "Monitoring global messenger RNA changes in externally controlled microarray experiments." EMBO Rep **4**(4): 387-393.

van de Werken, H., S. Brouns, et al. (2008). "Pentose Metabolism in Archaea." Archaea: new models for prokaryotic biology: 71.

van der Oost, J., M. A. Huynen, et al. (2002). "Molecular characterization of phosphoglycerate mutase in archaea." FEMS Microbiol Lett **212**(1): 111-120.

van der Veen, D., J. M. Oliveira, et al. (2009). "Analysis of variance components reveals the contribution of sample processing to transcript variation." Appl Environ Microbiol **75**(8): 2414-2422.

Vander Horn, P. B., M. C. Davis, et al. (1997). "Thermo Sequenase DNA polymerase and T. acidophilum pyrophosphatase: new thermostable enzymes for DNA sequencing." Biotechniques **22**(4): 758-762, 764-755.

Veit, K., C. Ehlers, et al. (2006). "Global transcriptional analysis of Methanosarcina mazei strain Go1 under different nitrogen availabilities." Mol Genet Genomics **276**(1): 41-55.

Verhees, C. H., S. W. Kengen, et al. (2003). "The unique features of glycolytic pathways in Archaea." Biochem J **375**(Pt 2): 231-246.

Wang, Y., Z. Duan, et al. (2007). "A novel Sulfolobus non-conjugative extrachromosomal genetic element capable of integration into the host genome and spreading in the presence of a fusellovirus." Virology **363**(1): 124-133.

Wang, Z., M. Gerstein, et al. (2009). "RNA-Seq: a revolutionary tool for transcriptomics." Nat Rev Genet **10**(1): 57-63.

Waters, E., M. J. Hohn, et al. (2003). "The genome of Nanoarchaeum equitans: insights into early archaeal evolution and derived parasitism." Proceedings of the National Academy of Sciences of the United States of America **100**(22): 12984-12988.

Waters, L. S. and G. Storz (2009). "Regulatory RNAs in bacteria." Cell **136**(4): 615-628.

Weinberg, M. V., G. J. Schut, et al. (2005). "Cold shock of a hyperthermophilic archaeon: Pyrococcus furiosus exhibits multiple responses to a suboptimal growth temperature with a key role for membrane-bound glycoproteins." J Bacteriol **187**(1): 336-348.

Wende, A., K. Furtwangler, et al. (2009). "Phosphate-dependent behavior of the archaeon Halobacterium salinarum strain R1." J Bacteriol **191**(12): 3852-3860.

Wilhelm, B. T. and J. R. Landry (2009). "RNA-Seq-quantitative measurement of expression through massively parallel RNA-sequencing." Methods **48**(3): 249-257.

Williams, E., T. M. Lowe, et al. (2007). "Microarray analysis of the hyperthermophilic archaeon Pyrococcus furiosus exposed to gamma irradiation." Extremophiles **11**(1): 19-29.

Woese, C. R. and G. E. Fox (1977). "Phylogenetic structure of the prokaryotic domain: the primary kingdoms." Proc Natl Acad Sci U S A **74**(11): 5088-5090.

Worthington, P., V. Hoang, et al. (2003). "Targeted disruption of the alpha-amylase gene in the hyperthermophilic archaeon Sulfolobus solfataricus." J Bacteriol **185**(2): 482-488.

Wurtzel, O., R. Sapra, et al. (2010). "A single-base resolution map of an archaeal transcriptome." Genome Res **20**(1): 133-141.

Xia, Q., E. L. Hendrickson, et al. (2006). "Quantitative proteomics of the archaeon Methanococcus maripaludis validated by microarray analysis and real time PCR." Mol Cell Proteomics **5**(5): 868-881.

Xin, H., T. Itoh, et al. (2001). "Natronobacterium nitratireducens sp. nov., a aloalkaliphilic archaeon isolated from a soda lake in China." International Journal of Systematic and Evolutionary Microbiology **51**(Pt 5): 1825-1829.

Yergeau, E., S. A. Schoondermark-Stolk, et al. (2009). "Environmental microarray analyses of Antarctic soil microbial communities." ISME J **3**(3): 340-351.

Zaigler, A., S. C. Schuster, et al. (2003). "Construction and usage of a onefold-coverage shotgun DNA microarray to characterize the metabolism of the archaeon Haloferax volcanii." Mol Microbiol **48**(4): 1089-1105.

Zaigler, A., S. C. Schuster, et al. (2003). "Construction and usage of a onefold-coverage shotgun DNA microarray to characterize the metabolism of the archaeon Haloferax volcanii." Mol Microbiol **48**(4): 1089-1105.

Zaparty, M., D. Esser, et al. (2010). ""Hot standards" for the thermoacidophilic archaeon Sulfolobus solfataricus." Extremophiles **14**(1): 119-142.

Zaparty, M., B. Tjaden, et al. (2008). "The central carbohydrate metabolism of the hyperthermophilic crenarchaeote Thermoproteus tenax: pathways and insights into their regulation." Arch Microbiol **190**(3): 231-245.

Zhang, Q., T. Iwasaki, et al. (1996). "2-oxoacid:ferredoxin oxidoreductase from the thermoacidophilic archaeon, Sulfolobus sp. strain 7." J Biochem (Tokyo) **120**(3): 587-599.

Zhang, W., D. E. Culley, et al. (2006). "DNA microarray analysis of anaerobic Methanosarcina barkeri reveals responses to heat shock and air exposure." J Ind Microbiol Biotechnol **33**(9): 784-790.

Zillig, W., H. P. Arnold, et al. (1998). "Genetic elements in the extremely thermophilic archaeon Sulfolobus." Extremophiles: Life Under Extreme Conditions **2**(3): 131-140.

Zillig, W., A. Kletzin, et al. (1994). "Screening for Sulfolobales, their Plasmids and their Viruses in Icelandic Solfataras." Systematic and applied microbiology **16**(4): 609-628.

Zillig, W., K. O. Stetter, et al. (1980). "The Sulfolobus-"Caldariella" group: Taxonomy on the basis of the structure of DNA-dependent RNA polymerases." Archives of Microbiology **125**(3): 259-269.

Zillig, W., K. O. Stetter, et al. (1980). "The Sulfolobus-Caldariella group: taxonomy on the basis of the structure of DNA-dependent RNA polymerases." Arch. Microbiol. **125**: 259-269.

Nederlandse Samenvatting

Nederlandse Samenvatting

Het onderzoek, zoals beschreven in dit proefschrift, heeft zich gericht op het organisme *Sulfolobus solfataricus*, welke tot de groep organismen genaamd archaea behoort. Aan de hand van de hoofdstukken zal ik een samenvatting van dit proefschrift geven. Hoofdstuk 1 en 2 zijn gebasseerd op een literatuurstudie. In hoofdstuk 3 tot en met 6 staat het onderzoek beschreven dat ik samen met collega's als promovendus in het lab uitgevoerd heb.

In **hoofdstuk 1**, de inleiding, wordt ingegaan op een groep organismen; de archaea. Het leven is ruwweg in drie groepen onder te verdelen: de eukaryoten, de bacteriën en de archaea. De eukaryoten omvatten meercelligen, zoals mensen, dieren en planten, maar ook eencelligen zoals gisten. De bacteriën, alleen eencelligen, zijn bekend geworden door de vele ziektes die ze kunnen veroorzaken, maar zijn zeker ook van belang voor het kunnen overleven van mensen en dieren. De derde groep, archaea, wat oudste betekent, is de meest recent ontdekte groep organismen. Deze eencelligen zijn te vinden in onze darmen, in de grond maar ook in bijvoorbeeld de oceanen. Het opmerkelijke aan deze groep is dat het de meest extreme soorten op aarde bevat. Sommige soorten die tot de archaea behoren kunnen namelijk groeien bij een temperatuur van 122 ^{o}C, bij een pH van rond de 0 (dit is zo zuur dat mensen zouden "oplossen") of in een verzadigde zoutoplossing. Het organisme *Sulfolobus solfataricus*, dat centraal staat in dit proefschrift, groeit optimaal bij een temperatuur van 80 ^{o}C en in een sterk zuur milieu (pH 2-5).

In **hoofdstuk 2** wordt een overzicht gegeven van de studies die wereldwijd gedaan zijn met hitte minnende archaea (archaea die alleen groeien boven 60 ^{o}C). De studies geven inzicht in de expressie van alle genen van het onderzochte organisme (het transcriptoom). Aan het begin van dit onderzoek was de techniek die het mogelijk maakt om naar de expressie van alle genen te kijken revolutionair omdat hiermee de wetenschap veel sneller inzicht kon krijgen in hoe levende organismen op celniveau opereren. Het hoofdstuk begint met een pioniersstudie, waar (slechts) 271 genen van *Pyrococcus furioses* tegelijkertijd werden bekeken, waarna andere studies worden beschreven die naar de expressie van alle genen keken (het gaat hier om duizenden genen). In enkele baanbrekende onderzoeken werd gekeken hoe virussen en de gastheer (hier *S. solfataricus*) met elkaar omgaan op DNA-niveau. Door middel van de microarray techniek zijn er nieuwe

ontdekkingen gedaan op het gebied van de celcyclus en de vermeerdering van het DNA bij de celdeling. Deze publicaties worden in dit hoofdstuk ook besproken.

Hoofdstuk 3 beschrijft het eerste gepubliceerde onderzoek waar data van het transcriptoom en het proteoom (alle eiwitten die geproduceerd worden op één moment) in archaea gecombineerd wordt met een studie naar de citroenzuurcyclus en de glycolyse (de twee belangrijkste stofwisselingsroutes in een cel). Er wordt aangetoond welke genen hierbij een rol spelen. Ook wordt duidelijk dat de regulatie van deze twee routes, in *Sulfolobus solfataricus*, zeer waarschijnlijk niet veroorzaakt wordt door de hoeveelheid enzymen (dit zijn eiwitten die een chemische reactie versnellen) in de cel te veranderen zoals verondersteld werd, maar dat mogelijk de activiteit van die enzymen gereguleerd wordt.

Hoofdstuk 4 beschrijft het onderzoek dat op een moderne manier een metabole route opheldert. In het onderzoek wordt gekeken naar de manier waarop de suiker D-arabinose in *S. solfataricus* omgezet kan worden in grondstoffen voor bouwstenen of energie. D-arabinose wordt in vijf stappen omgezet naar het molecuul 2-oxoglutaraat, welke in de citroenzuurcyclus een rol speelt.

S. solfataricus is een obligaat aeroob organisme, wat betekent dat dit organisme dood gaat als er geen zuurstof aanwezig is. Het komt in de natuur voor in modderpoelen, dus een waterige oplossing, en is dus sterk afhankelijk van de hoeveelheid opgeloste zuurstof in het water waarin het zich bevindt. Het is algemeen bekend dat gassen slecht oplossen in warm water. Een snelle rekensom levert op dat de hoeveelheid zuurstof welke bij 80 ^{o}C opgelost is, ongeveer de helft is van wat er aan zuurstof opgelost kan worden bij 20 ^{o}C. Hier komt nog bij dat een lage pH de maximale oplosbaarheid van zuurstof nog verder weet te verlagen. Niet de beste situatie als je in water leeft met een zeer lage pH (2-5) en een hoge temperatuur (80 ^{o}C). In **hoofdstuk 5** wordt het onderzoek beschreven dat ingaat op de invloed die de zuurstofconcentratie heeft op onder andere de groeisnelheid van *S. solfataricus*. Het blijkt dat te weinig zuurstof voor een lage groeisnelheid zorgt, maar dat een te hoge concentratie dodelijk is voor het organisme. Het is dus heel belangrijk om de juiste hoeveelheid opgeloste zuurstof in het medium te hebben waarin *S. solfataricus* groeit.

Carotenen zijn van vitaal belang voor het opvangen van licht in het oog. Eveneens draagt deze groep moleculen bij aan het voorkomen van staar en macula degeneratie. Carotenen zijn echter voor meer organismen belangrijk. Carotenen zijn onder andere belangrijk bij de bescherming tegen zuurstof en licht en zorgen

voor de kleuring van bloemen. Verschillende soorten die tot de familie *Sulfolobus* horen kunnen carotenen maken. **Hoofdstuk 6** beschrijft welke genen van belang zijn voor de synthese van carotenen in enkele *Sulfolobus* soorten. Ook is aangetoond dat, hoewel de genetische code van de verschillende genen sterk op elkaar lijken, de verschillende soorten *S. solfataricus, S. shibatae* en *S. acidocaldarius* verschillende carotene-achtige moleculen kunnen produceren.

In **hoofdstuk 7**, de algemene samenvatting en discussie, wordt ingegaan op de ontwikkeling van de bioreactor die gebruikt is om *Sulfolobus* te groeien. Deze bioreactor maakt het mogelijk om *Sulfolobus* te groeien op grotere schaal (1 à 1,5 L; het "normale" volume is 0,020-0,050 L). Een ander voordeel is dat het zuurstofgehalte, de temperatuur en de zuurgraad (pH) via de bioreactor gereguleerd kunnen worden. De schaalvergroting was nodig om genoeg cellen te kunnen oogsten voor de onderzoeken. De gecontroleerde groei zorgt voor kwalitatief betere resultaten.

Dankwoord

Dankwoord

Graag wil ik alle mensen bedanken die belangrijk waren bij het tot stand komen van dit proefschrift.

Allereerst wil ik John bedanken als mijn promotor. John, bedankt voor alles. Ik heb veel van onze samenwerking geleerd en wil je graag bedanken dat ik altijd van je uitgebreide netwerk gebruik heb kunnen maken. Willem, onze samenwerking, vooral aan het eind van dit proefschrift, vond ik prettig: snelle antwoorden en positieve mail.

Naast mijn promotoren wil ik mijn paranimfen Marieke en Pawel bedanken. Pawel, you are a great person and a great friend. I have always liked our coöperation. Marieke, ik heb nog vele goede herinneringen aan de thee-uurtjes, die wij samen op Rijnsteeg hadden. Stan en Harmen: ik wil jullie bedanken voor de vele gesprekken die we samen gevoerd hebben en de vele adviezen die jullie mij gegeven hebben. Jaapie, jij was er altijd om een goede sfeer te maken op bacgen. Verder wil ik al mijn collega's bedanken op bacgen: Matthijs, Magnus, Ronnie, Krisztina, Hao, Suzanne (voor de fantastische week in California), John R, Pierpaolo, Ans, Colin, Mark, Marco, Bart en Servé. Wim, bedankt voor het oplossen van alle computerproblemen. Nees, bedankt voor alle bestellingen. Renée voor het onderhouden van de bieb. Francis en Anja, bedankt voor alle administratieve zaken.

Verder wil ik alle mensen bedanken met wie ik heb samengewerkt in de verschillende onderzoeken. Bram voor een het mogelijk maken van mijn eerste publicatie. Alice, your personality has made a great impact on me. Gwenola, a strong woman and a great mother, I very much liked our coöperation.

Verder wil ik mijn familie bedanken voor alle steun door de jaren heen. Papa en Jetty, bedankt dat jullie altijd klaarstonden. Mam, bedankt voor de goede zorgen. Johan en Akke, fijn dat jullie er waren met raad en daad als dit nodig was.

Als laatste wil ik Marije bedanken. Marije, jij bent er altijd. Altijd bezig om mij verder te helpen en mij te motiveren als mijn motivatie moeilijk te vinden is. Bedankt dat jij voor mij drie dagen in de kou, wind en regen gestaan hebt om die ene foto voor mij te maken, die nu op de voorkant van mijn proefschrift te zien is. Bedankt voor jouw organisatorisch talent en de tijd die je erin gestoken hebt om mijn proefschrift te redigeren.

About the author

About the author

Jasper Walther was born on 13 July 1976 in Leidschendam, the Netherlands. In 1996 he finished secondary school (vwo) at Slauerhoff College in Leeuwarden. The same year he started his MSc. bioprocestechnology at the University in Wageningen. During his study he focused on microbiology, metabolism and fermentation. His study included a six months internship at the Uppsala University, Sweden, where he was schooled in using microarrays for the study of *Sulfolobus solfataricus* and helped to manufacture the first arrays for this organism. In 2002 he started his Ph.D. at the Wageningen University where he studied the metabolism of *S. solfataricus*. In 2006 he exchanged the research in the lab for teaching chemistry at secondary school. He obtained his degree in teaching and is currently employed by Sancta Maria Lyceum in Haarlem.

List of publications

Walther* J., Sierocinski* P., van der Oost, J. **Hot Transcriptomics.** Archaea 2010; vol. 2010: 14 pages. doi:10.1155/2010/897585.

Simon* G., Walther* J., Zabeti N., Combet-Blanc Y., Auria R., van der Oost J., Casalot L. **Effect of O2 concentrations on *Sulfolobus solfataricus* P2.** FEMS microbiology letters 2009 Sept; 299: 255-60.

Ortmann A.C., Brumfield S.K., Walther J., McInnerney K., Brouns S.J.J., van de Werken H.J.G., Bothner B., Douglas T., van de Oost J., Young M.J. **Transcriptome analysis of infection of the archaeon Sulfolobus solfataricus with Sulfolobus turreted icosahedral virus.** J. virol. 2008 Mar; 82(10): 4874-83.

Brouns S.J.J., Walther J., Snijders A.P.L., van de Werken H.J.G., Willemen H.L.D.M., Worm P., de Vos M.G.J., Andersson A., Lundgren M., Mazon H.F.M., van den Heuvel R.H.H., Nilsson P., Salmon L., de Vos W. M., Wright P. C., Bernander R., van der Oost J. **Identification of the missing links in prokaryotic pentose oxidation pathways: evidence for enzyme recruitment.** J Biol Chem. 2006 Jul; 281 (37): 27378-88

Snijders* A.P.L., Walther* J., Peter S., Kinnman I., de Vos M.G.J., van de Werken H.J.G., Brouns S.J.J., van der Oost J., Wright P.C.. **Reconstruction of central carbon metabolism in *Sulfolobus solfataricus* using a two-dimensional gel electrophoresis map, stable isotope labelling and DNA microarray analysis.** Proteomics. 2006 Mar; 6 (5): 1518-29.

Van der Oost J., Walther J., Brouns S.J.J., van de Werken H.J.G., Ambrosius P. L. Snijders, Phillip C. Wright, Anders Andersson, Rolf Bernander and Willem M. de Vos. **Functional Genomics of the Thermo-Acidophilic Archaeon *Sulfolobus Solfataricus*.** In: Rainey, F.A., and Oren, A. (eds.) 2006. Extremophiles – Methods in Microbiology Vol. 35. Elsevier/Academic Press, Amsterdam.

Kluskens L.D., van Alebeek G.J.W.M., Walther J., Voragen A.G.J.,de Vos W.M., van der Oost J. **Characterization and mode of action of an exopolygalacturonase from the hyperthermophillic bacterium *Thermotoga maritima*.** FEBS J. 2005 Nov; 272 (21): 5464-73.

Brouns, S.J.J., Ettema, T.J.G., Stedman, K,M., Walther, J., Smidt, H., Snijders, A.P.L.,Young, M., Bernander, R., Wright, P.C., Siebers, B., Van der Oost, J. **The hyperthermophilic archaeon *Sulfolobus* - from exploration to exploitation.** In:Geothermal Biology and Geochemistry in Yellowstone National Park (Eds. M. Young and B. Inskeep). *2005.*

*Contributed equally

Overview of completed training activities

Discipline specific activities

Courses

Affymetrix course, Wageningen University; 2004.
Bioinformation technology, VLAG; 2004
Systems biology, VLAG; 2005
Advanced course on Applied Genomics of Industrial Fermentation, VLAG; 2005
Advanced course on biocatalysis, Delft University; 2006

Meetings

Bioinformatics in microarray research, Centre for Medical systems Biology; 2004
8th international meeting of the microarray gene expression data society; 2005
Extremophiles 2006*
Thermophiles 2007*
Project meetings; Bacterial genetics (weekly)
PhD meetings; Laboratory of Microbiology (bi-weekly)
Annual meetings Platform Moleculaire Genetica* 2003-2007
Microarray meetings; Laboratory of Microbiology

* Poster presentation

General courses

Scientific publishing
VLAG PhD Week

Optionals

PhD trip, California, USA
Preparing PhD research proposal

Cover: Marije Dorenbos & Jasper Walther
Photo was taken at Námaskarð, Iceland

Print: GVO Drukkers & Vormgevers BV, Ede