



This is a repository copy of *A systems biology approach reveals major metabolic changes in the thermoacidophilic archaeon Sulfolobus solfataricus in response to the carbon source L-fucose versus D-glucose.*

White Rose Research Online URL for this paper:  
<http://eprints.whiterose.ac.uk/106194/>

Version: Accepted Version

---

#### Article:

Wolf, J., Stark, H., Fafenrot, K. et al. (16 more authors) (2016) A systems biology approach reveals major metabolic changes in the thermoacidophilic archaeon *Sulfolobus solfataricus* in response to the carbon source L-fucose versus D-glucose. *Molecular Microbiology*, 102 (5). pp. 882-908. ISSN 0950-382X

<https://doi.org/10.1111/mmi.13498>

---

This is the peer reviewed version of the following article: , Wolf, J. et al (2016), A systems biology approach reveals major metabolic changes in the thermoacidophilic archaeon *Sulfolobus solfataricus* in response to the carbon source L-fucose versus D-glucose. *Molecular Microbiology*, which has been published in final form at <http://dx.doi.org/10.1111/mmi.13498>. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

#### Reuse

Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

#### Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



[eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk)  
<https://eprints.whiterose.ac.uk/>

# **A systems biology approach reveals major metabolic changes in the thermoacidophilic archaeon *Sulfolobus solfataricus* in response to the carbon source L-fucose versus D-glucose**

Jacqueline Wolf<sup>1</sup>, Helge Stark<sup>1</sup>, Katharina Fafenrot<sup>2</sup>, Andreas Albersmeier<sup>3</sup>, Trong K. Pham<sup>4</sup>, Katrin B. Müller<sup>1</sup>, Benjamin Meyer<sup>5</sup>, Lena Hoffmann<sup>5</sup>, Lu Shen<sup>2</sup>, Stefan P. Albaum<sup>3</sup>, Theresa Kouril<sup>2,a</sup>, Kerstin Schmidt-Hohagen<sup>1</sup>, Meina Neumann-Schaal<sup>1</sup>, Christopher Bräsen<sup>2</sup>, Jörn Kalinowski<sup>3</sup>, Phillip C. Wright<sup>4</sup>, Sonja-Verena Albers<sup>5</sup>, Dietmar Schomburg<sup>1,\*</sup> and Bettina Siebers<sup>2</sup>

<sup>1</sup>Department of Bioinformatics and Biochemistry, Technische Universität Braunschweig, 38106 Braunschweig, Germany

<sup>2</sup>Molecular Enzyme Technology and Biochemistry, Biofilm Centre, Universität Duisburg-Essen, 45141 Essen, Germany

<sup>3</sup>Center for Biotechnology - CeBiTec, Universität Bielefeld, 33615 Bielefeld, Germany

<sup>4</sup>Department of Chemical and Biological Engineering, ChELSI Institute, University of Sheffield, S1 3JD, Sheffield, UK

<sup>5</sup>Molecular Biology of Archaea, Institute for Biology II - Microbiology, Universität Freiburg, 79104 Freiburg, Germany

<sup>a</sup>Present address: Department of Biochemistry, Stellenbosch University, Private Bag X1, Matieland 7602, South Africa

**Running title:** Hexose degradation in *S. solfataricus*

**Key words:** carbohydrate metabolism; L-fucose; *Sulfolobus*; nutritional adaptation; systems biology

**\*Corresponding author:**

E-Mail: [d.schomburg@tu-bs.de](mailto:d.schomburg@tu-bs.de)

Phone: +49-531-391 55202

Fax: +49-531-391 55199

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as an 'Accepted Article', doi: 10.1111/mmi.13498

## Summary

Archaea are characterised by a complex metabolism with many unique enzymes that differ from their bacterial and eukaryotic counterparts. The thermoacidophilic archaeon *Sulfolobus solfataricus* is known for its metabolic versatility and is able to utilize a great variety of different carbon sources. However, the underlying degradation pathways and their regulation are often unknown. In this work, we analyse growth on different carbon sources using an integrated systems biology approach. The comparison of growth on L-fucose and D-glucose allows first insights into the genome-wide changes in response to the two carbon sources and revealed a new pathway for L-fucose degradation in *S. solfataricus*. During growth on L-fucose we observed major changes in the central carbon metabolic network, as well as an increased activity of the glyoxylate bypass and the 3-hydroxypropionate/4-hydroxybutyrate cycle. Within the newly discovered pathway for L-fucose degradation the following key reactions were identified: (i) L-fucose oxidation to L-fuconate via a dehydrogenase, (ii) dehydration to 2-keto-3-deoxy-L-fuconate via dehydratase, (iii) 2-keto-3-deoxy-L-fuconate cleavage to pyruvate and L-lactaldehyde via aldolase and (iv) L-lactaldehyde conversion to L-lactate via aldehyde dehydrogenase. This pathway as well as L-fucose transport shows interesting overlaps to the D-arabinose pathway, representing another example for pathway promiscuity in *Sulfolobus* species.

## Introduction

Hostile environments, such as solfatara are characterised by low pH and temperatures above 65 °C and are dominated by members of the third domain of life, the archaea. In the past decades, the thermoacidophilic Crenarchaeon *Sulfolobus solfataricus* has developed into one of the best characterised model organisms with respect to adaptation to various stress factors or different nutrient conditions (Zillig *et al.*, 1980; Nicolaus *et al.*, 1991; Izzo *et al.*, 2005). Furthermore, for *S. solfataricus* a genetic system has been established that allows for the generation of single knock-outs based on the insertion of the *lacS* marker gene (Worthington *et al.*, 2003; Albers and Driessen, 2008). The organism (optimal growth at 80 °C, pH 3.5) is known for its broad sugar degrading capacity, which is reflected in its ability to utilise pentoses, hexoses as well as oligo- and polysaccharides (Grogan, 1989; Ulas *et al.*, 2012).

For the degradation of both, D-glucose and D-galactose, the organism uses a promiscuous branched Entner-Doudoroff pathway, meaning that the same metabolic route is used for degradation of both glucose and galactose. This pathway includes a semi- and a non-phosphorylating branch as a central metabolic route (De Rosa *et al.*, 1984; Lambie *et al.*, 2003; Ahmed *et al.*, 2005; Bräsen *et al.*, 2014). The glucose and galactose metabolism in *S. solfataricus* is well understood and several pathways for pentose degradation have been analysed in significant detail. The degradation of D-arabinose proceeds via formation of D-arabinonate by D-arabinose dehydrogenase and finally leads to 2-oxoglutarate which is fed into the citric acid cycle (Brouns *et al.*, 2006; Brouns *et al.*, 2007). For conversion of D-xylose and L-arabinose, *S. solfataricus* and *S. acidocaldarius* possess a branched degradation pathway, making use of enzymes of the Entner-Doudoroff pathway. The first aldolase-independent route leads to formation of 2-oxoglutarate and is therefore similar to D-arabinose degradation, whereas the second route, also used for L-arabinose degradation, requires aldol cleavage of the 2-keto-3-deoxy-pentanoate intermediate to form glycolaldehyde and pyruvate. Glycolaldehyde is then metabolised to glyoxylate, which enters the citric acid cycle (Nunn *et al.*, 2010).

For pentose formation, the organism possesses a reverse ribulose-monophosphate pathway instead of the classical pentose-phosphate pathway usually found in bacteria and fungi (She *et al.*, 2001; Soderberg, 2005; Orita *et al.*, 2006). Originally this pathway is known as fixation cycle for formaldehyde in methylotrophic bacteria and condenses formaldehyde and ribulose-monophosphate to hexulose-6-phosphate which is then further converted to fructose-6-phosphate (Strom *et al.*, 1974). *S. solfataricus* and most other archaea use this pathway in a reverse manner to synthesize ribulose-5-monophosphate from fructose-6-phosphate. An

isomerase is then finally responsible for formation of ribose-5-phosphate (Soderberg, 2005, Ulas *et al.*, 2012).

The hexose L-fucose (6-deoxygalactose), a stereochemical analogue of the pentose D-arabinose, is naturally present in the cell envelope of numerous pro- and eukaryotes and as constituent of exopolysaccharides (Vanhooren and Vandamme, 1999). In *Escherichia coli* and many other bacterial species, L-fucose is metabolised via a kinase-dependent pathway including the formation of L-fuculose and L-fuculose-1-phosphate, finally yielding dihydroxyacetone-phosphate and lactaldehyde. Dihydroxyacetone-phosphate is then directly used in glycolysis for energy production, whereas lactaldehyde could be either aerobically oxidised to pyruvate (via lactate) or anaerobically converted to 1,2-propanediol (Baldomà and Aguilar, 1988). The latter can be either secreted (Boronat and Aguilar, 1981) or converted to propanol and propionate (Obradors *et al.*, 1988; Petit *et al.*, 2013). Typically L-fucose utilisation operons code for four main fucose processing proteins, a fucose permease, responsible for sugar uptake, an isomerase (EC 5.3.1.25) for conversion to fuculose, a kinase (EC 2.7.1.51) producing the corresponding phosphate and finally an aldolase (EC 4.1.2.17) for conversion to dihydroxyacetone-phosphate and lactaldehyde (Green and Cohen, 1956; Ghalambor and Heath, 1962; Heath and Ghalambor, 1962).

A second non-phosphorylating route for L-fucose degradation was found in *Xanthomonas campestris* (Yew *et al.*, 2006). In this organism, a dehydrogenase (EC 1.1.1.122) first converts L-fucose to the corresponding L-fuconolactone, which is then hydrolysed to L-fuconate. Via formation of 2-keto-3-deoxy-L-fuconate (KDF) this sugar acid is further oxidized to 2,4-diketo-3-deoxy-L-fuconate, which is finally cleaved to pyruvate and lactate. Recent studies revealed the existence of a closely related route for degradation of L-rhamnose in the thermoacidophilic bacterium *Sulfobacillus thermosulfidooxidans* (Bae *et al.*, 2015).

While L-fucose utilisation is well understood in bacteria, very little is known about possible degradation routes in archaea. Previous studies claimed that the thermoacidophilic archaeon *Thermoplasma acidophilum* is able to grow on deoxysugars like, L-fucose or L-rhamnose, however, the underlying pathway of L-fucose degradation is still unknown (Kim *et al.*, 2012). In the genome of *S. solfataricus*, only an  $\alpha$ -L-fucosidase presumably regulated by translational frameshifting was identified so far (Cobucci-Ponzano *et al.*, 2003; Cobucci-Ponzano *et al.*, 2006). In addition no homologs to any of the known enzymes of the *E. coli* L-fucose degradation pathway were identified.

As outlined above the archaeal metabolism is characterized by the presence of new metabolic pathways or often unusual, modified versions of the classical routes. This applies especially to the central carbohydrate metabolism and the modifications often rely on the

utilization of so far unknown enzymes that show no similarity to their respective bacterial and eukaryotic counterparts. Many of these novel enzymes are members of large multifunctional enzyme families and multiple candidates, sometimes promiscuous and/or with broad substrate specificity, were identified in archaeal genomes (e.g. D-glucose dehydrogenase is a member of the medium-chain dehydrogenase/reductase (MDR) superfamily comprising 16 alcohol and sugar dehydrogenases on the *S. solfataricus* genome). Therefore the decipherment of metabolic pathways and the identification of the involved enzymes is challenging in archaea. Here a systems biology approach helps to gain a deeper insight into the utilization and regulation of metabolic pathways in response to different carbon sources. Therefore, to unravel L-fucose degradation and to explore the genome-wide changes during growth on L-fucose compared to D-glucose in *S. solfataricus*, we applied an integrated systems biology approach combining transcriptomics, proteomics, metabolomics, and modelling techniques together with genetic and biochemical studies. During growth on L-fucose only a small set of genes and proteins was differentially expressed, compared to growth on D-glucose. Some of the up-regulated genes/proteins have been previously characterised as enzymes of the D-arabinose degradation pathway (Brouns *et al.*, 2006). In addition to distinct changes within the central carbon metabolic pathways (e.g. gluconeogenesis), the systems biology approach revealed a new promiscuous route for D-arabinose and the naturally occurring L-fucose transport and degradation. In this pathway L-fucose is oxidised to L-fuconate by a promiscuous dehydrogenase and further dehydrated by a dehydratase to KDF. KDF is then cleaved by an aldolase to pyruvate and L-lactaldehyde.

## Results

### *Development of a metabolic model of S. solfataricus P2 and modeling of L-fucose degradation*

On the basis of our genome scale metabolic model of *Sulfolobus solfataricus* P2 (Ulas *et al.*, 2012) that predicts growth for various carbon sources, we performed growth experiments on 36 different substrates. We observed growth of *S. solfataricus* on 16 compounds primarily including carbohydrates of different complexity (Supporting Information Dataset S1). As metabolic fingerprinting revealed some interesting differences in metabolite profiles of cells grown on L-fucose compared to D-glucose, and as L-fucose degradation has never been reported for archaea before, we chose this sugar for an integrated systems biology analysis to explore its metabolic fate in *S. solfataricus*.

To reconstruct possible L-fucose degradation pathways by predicting fluxes through hypothetical catabolic pathways and comparing the results with experimental data, the existing model was first improved using manual and automated annotation techniques as well as experimental data from this work. In the end, the model size increased from formerly 718 reactions and 515 genes to 969 reactions and 604 genes. The updated model is supplied in the SBML (Hucka *et al.*, 2003) and metano (Riemer *et al.*, 2013) format. We adopted the supplied experimental data available from the quantification of major biomass fractions (Table I) and updated the biomass reaction for growth on glucose. In this context, we also included a biomass reaction for growth on L-fucose based on the experimental data. Referring to the biomass fractions, the strongest improvement could be made for the nucleic acids. Especially for RNA the difference between the previously estimated value of 24 % (adopted from *Methanosarcina barkeri* (Feist *et al.*, 2006)) and our current measured experimental data with only 5.4 % RNA for *S. solfataricus*, had a major impact on the model. However, adjustment of the new biomass fractions resulted in a gap of approximately 18 %. To ensure the integrity of the model, this gap was filled with glycogen as it was shown that the glycogen content in members of the genus *Sulfolobus* can increase up to 25 % of total cell dry weight and that glycogen is only synthesised under optimal growth conditions (König *et al.*, 1982; Iglesias and Preiss, 1992), which was the basic assumption for all conducted experiments.

As no L-fucose metabolism in archaea was reported before, we first used our metabolic model to investigate metabolic fluxes and the efficiency of different degradation pathways in *S. solfataricus* (Supporting Information Model S1-S4). For this purpose, three pathway variants were tested (Figure 1). This includes the *E. coli*-like degradation via formation of L-fuculose (Figure 1C), the metabolic route found in *X. campestris* including the formation of 2,4-diketo-3-deoxy-L-fuconate from KDF (Figure 1B) and finally a pathway where KDF is

directly cleaved to lactaldehyde and pyruvate (Figure 1A). For metabolic modeling, we temporarily added reactions for each tested pathway to the model and performed simulations using the L-fucose specific biomass reaction as the objective function. In all scenarios, we fitted the non-growth associated maintenance energy (NGAM) so that the predicted biomass flux would always match the experimentally observed carbon balance (integration of carbon into the biomass). This approach typically yields the highest agreement between predicted and real metabolite fluxes. Finally, since there is experimental evidence that the D-arabinose-1-dehydrogenase (*Sso1300*) accepts L-fucose as substrate (Brouns *et al.*, 2007), this reaction was permanently integrated into the model beforehand.

Both KDF dependent degradation routes (Figure 1A and 1B) would produce 3 NADH per molecule L-fucose and are therefore energetically comparable. In agreement with the lower oxidation state at the C<sub>6</sub> of fucose, the modified Entner-Doudoroff pathway found in *S. solfataricus*, only yields 2 NADH and also no net ATP from one molecule of D-glucose. In *E. coli* L-fucose degradation (Figure 1C) would yield not only 3 NADH but also 1 molecule ATP per molecule fucose, as the formed dihydroxyacetone phosphate is utilized via glyceraldehyde-3-phosphate in the Embden-Meyherhof-Parnas pathway. However, in *S. solfataricus* this pathway would also yield only 3 NADH and no net ATP, as the non phosphorylating glyceraldehyde-3-phosphate dehydrogenase catalyzes the direct oxidation of glyceraldehyde-3-phosphate to 3-phosphoglycerate, omitting any formation of 1,3-bisphosphoglycerate. Therefore, energetics of the three catabolic pathways are similar in *S. solfataricus* and also close to those of glucose degradation, since the major portion of the total energy is gained via the TCA cycle and the respiratory chain. With respect to anabolic reactions, the pathway via fucose-1-phosphate appears to be slightly more efficient, as intermediates of the degradation pathway (dihydroxyacetone phosphate and phosphoenolpyruvate) could also directly be used in gluconeogenesis without additional energy investment. In contrast, on both KDF dependent routes phosphoenolpyruvate has to be synthesized from either pyruvate or oxaloacetate, which requires additional ATP investment. However, fitting of the energy maintenance parameters (GAM & NGAM) to the experimentally observed carbon balance shows that similar values were achieved in all three tested L-fucose degradation routes, thus the differences in the overall energy balance of the three scenarios is very small.

As the experimental results suggests a certain activity of the 3-hydroxypropionate / 4-hydroxybutanoate cycle and the glyoxylate bypass (see below), we further explored the solution space of the tested degradation scenarios using suboptimal flux variance analysis, allowing for a suboptimal biomass flux of 95 %. Compared to growth on glucose, an

increased flux towards the glyoxylate bypass was only feasible in the KDF dependent fucose growth scenarios (Table II). In contrast the possible activity through the complete 3-hydroxypropionate / 4-hydroxybutanoate cycle was comparable under all simulated growth conditions, although parts of the cycle (e.g. sequence from propanoyl-CoA to succinyl-CoA) allow an increased activity in the KDF dependent scenarios (Table II). Finally, flux balance analysis showed that the overall energy demand as represented by GAM and NGAM predicted for growth on L-fucose in the KDF dependent scenario is 42 % higher compared to simulated growth on D-glucose (Table II).

Accepted Article

### ***Comparative phenotypic analysis of *S. solfataricus* P2 grown on L-fucose versus D-glucose***

Growth and substrate uptake of *S. solfataricus* P2 were examined in Brock minimal medium (Brock *et al.*, 1972) either in presence of D-glucose (22 mM) or L-fucose (15 mM). Concentrations of L-fucose above 15 mM had an inhibiting effect on growth of *S. solfataricus*.

The organism showed comparable growth properties and substrate uptake rates on L-fucose and D-glucose (Figure 2A and 2B), with maximal growth rates of  $0.045 \text{ h}^{-1}$  on D-glucose and  $0.040 \text{ h}^{-1}$  on L-fucose, respectively. The molar growth yield per mol of carbon available from the substrate was 10 % lower during growth on L-fucose compared to growth on D-glucose showing that on L-fucose less carbon is integrated into the biomass (Table III).

Under both conditions, the carbon source concentration decreased only slightly during the first 60 hours of cultivation (Figures 2A and 2B), whereas the maximal uptake rate during the exponential growth phase was  $1.21 \text{ mmol}\cdot\text{g}^{-1}\cdot\text{h}^{-1}$  under L-fucose conditions compared to  $1.13 \text{ mmol}\cdot\text{g}^{-1}\cdot\text{h}^{-1}$  on the D-glucose reference (Table III).

It has been shown for *S. solfataricus* that the presence of D-glucose has a negative impact on the expression of the D-arabinose transporter genes (Lubelska *et al.*, 2006), so that we analysed for catabolite repression of L-fucose in presence of D-glucose. For this purpose, the organism was cultivated in the presence of equal molar concentrations of L-fucose and D-glucose (15 mM each) and the culture supernatant was analysed by GC-MS to follow the consumption of external metabolites. *S. solfataricus* P2 consumed both sugars from the media in parallel (Figure 2C). Interestingly, once adapted to the cultivation conditions, *S. solfataricus* seems to prefer L-fucose over D-glucose as the relative amount of L-fucose in the media decreased faster. In comparison to growth on D-glucose or L-fucose alone, the specific growth rate increased to  $0.077 \pm 0.001 \text{ h}^{-1}$  in the presence of both sugars (Figure 2C), whereas the maximum cell dry weight was not affected (Table III).

### *Comparative metabolome analysis*

In order to identify specific metabolic routes used for degradation of L-fucose, a comparative metabolome analysis of *S. solfataricus* P2 grown either on L-fucose or D-glucose was performed using the established protocol for metabolite extraction and analysis by GC-MS (Zaparty *et al.*, 2010) with small modifications. Of 66 detected metabolites (Supporting Information Dataset S2) in cells grown on L-fucose, 38 could be clearly assigned to known compounds by comparison with mass spectra libraries. The remaining 28 compounds belong to unknown but reproducibly detected metabolites of *S. solfataricus*. In contrast, a total number of 67 metabolites were found in cells grown on D-glucose (Supporting Information Dataset S3), of which 44 could be identified as compounds belonging to the organism's primary metabolism. Furthermore, a detailed analysis of the CoA content of *S. solfataricus* using a LC-MS approach was also performed. A total number of 21 CoA-derivatives could be identified under both nutrient conditions (Supporting Information Dataset S2 and S3), whereas malonyl-CoA was found exclusively in L-fucose grown cells.

Metabolome analysis revealed some significant changes in relative intracellular metabolite concentrations of compounds belonging predominantly either to the glycolytic Entner-Doudoroff-pathway or the gluconeogenic Embden-Meyerhof-Parnas pathway (Figures 3 and 4). In the case of glucose catabolism, gluconate and glycerate showed the strongest difference between the two tested conditions, as these compounds could only be detected in extracts of glucose grown cells (Figure 4). Furthermore, the abundance of pyruvate was reduced in L-fucose extracts by nearly half of the amount detectable in D-glucose grown cells. With respect to metabolites belonging to the upper part of the Embden-Meyerhof-Parnas pathway, we observed the strongest differences in the relative abundances of glucose and glucose-6-phosphate, being reduced by at least 4-fold in L-fucose cell extracts (Table IV). In this context, fructose-6-phosphate was further exclusively detected in D-glucose cell extracts (Figure 4). In agreement with the change in carbon source, noticeable peaks belonging to fucose, fuconolactone and fuconate were found in GC-MS chromatograms of L-fucose grown cells, these were absent in the D-glucose reference samples (Figure 4).

In addition to the small polar metabolites, we also observed some differences in the relative abundance of several CoA intermediates (Figure 3, Table IV). In this context, malonyl-CoA was especially noticeable as this compound was exclusively detected in L-fucose grown cells (Figure 4). Furthermore, also acetoacetyl-CoA showed considerable differences between the two tested conditions, as this compound is normally not stable during solid phase extraction, used for purification of CoA derivatives. Therefore, detection of this compound at significantly higher amounts in L-fucose extracts (Figure 3) indicates high intracellular

Accepted Article

concentrations of this metabolite under L-fucose growth conditions. The metabolic model was used to elucidate the origin and metabolic fate of these two CoA-derivatives in *S. solfataricus*. Acetoacetyl-CoA is linked to lipid biosynthesis (Koga and Morii, 2007) and leucine degradation. However, according to our model predictions, the fluxes through these two pathways do not significantly change between the two nutrient conditions (Supporting Information Figure S1). Additionally, acetoacetyl-CoA is part of the 3-hydroxypropionate/4-hydroxybutyrate cycle, a pathway originally described as carbon-fixation pathway for archaea (Berg *et al.*, 2007; Berg *et al.*, 2010). This pathway also includes the only presence of malonyl-CoA in the metabolic network of *S. solfataricus* P2. In this context, the relative intracellular concentration of other 3-hydroxypropionate/4-hydroxybutyrate-cycle intermediates like crotonyl-CoA, 3-hydroxybutanoyl-CoA, methylmalonyl-CoA and propanoyl-CoA was also found to be increased on L-fucose with fold-changes ranging from 1.27 to 1.76, however not all of these changes were statistically significant (p-value < 0.01; Table IV). Overall, the results suggest an increased activity of the 3-hydroxypropionate/4-hydroxybutyrate cycle under L-fucose conditions.

### ***Global transcriptional and translational response of *S. solfataricus* P2 grown on L-fucose versus D-glucose***

Since metabolome analysis provided evidence for a new L-fucose degradation pathway in *S. solfataricus* P2, we were interested in the carbon-source dependent, global transcriptional and translational response of the organism. From the 3145 genes encoded on the *S. solfataricus* genome, in total 2692 genes were detected to be expressed under both conditions (Supporting Information Dataset S4) using RNAseq analysis. Among these, 32 genes showed a differential transcription between both growth conditions with fold changes of 4-fold or more (Table V). With regard to sugar degradation, we could assign at least six of these to be directly involved in L-fucose metabolism. Under L-fucose conditions, three of the four subunits of the arabinose ABC transporter (*Sso3066* - *Sso3068*) (Elferink *et al.*, 2001) were highly up-regulated (between 8.6 up to 25-fold) as compared to D-glucose conditions. The fourth subunit of this transporter (*Sso3069*) was just below the threshold level (fold change 3.4). Further, the highly up-regulated transcripts in L-fucose samples include the loci *Sso1300*, *Sso3117*, *Sso3118* and *Sso3124*. Interestingly, recent studies could show that all these genes play a central catalytic role in the degradation of the pentose sugar D-arabinose in *S. solfataricus* (Brouns *et al.*, 2006). Additionally, 12 genes with different metabolic functions were up-regulated more than 4-fold at transcript level, including the two subunits of a 2-oxoacid:ferredoxin oxidoreductase (*Sso2128* - *Sso2129*), an acetylornithine deacetylase/succinyl-diaminopimelate desuccinylase (*Sso1007*) (Van de Castele *et al.*, 1990) and a malonyl CoA reductase (*Sso2178*) (Alber *et al.*, 2006; Demmer *et al.*, 2013). The latter is involved in the 3-hydroxypropionate/4-hydroxybutyrate cycle and it should be noted, that all genes of the cycle were found to be expressed under both conditions, with some at relatively high levels (RPKMs ranging from 7 up to 975 (median RPKM of the whole dataset: 69)). Additionally, several enzymes from this pathway showed an increased expression under L-fucose conditions, although fold-changes ranging below the threshold level (fold-changes 2.0-4.0; Supporting Information Dataset S4). Furthermore some additional transporters were found to be induced, including two members of the MFS superfamily (*Sso1305*; *Sso3120*), as well as some proteins with unknown function (Table V).

Among the down-regulated transcripts under L-fucose conditions, a strong decrease was observed for the glucose-1-dehydrogenase (*Sso3003*, fold-change: 0.19). This enzyme is involved in glucose degradation, catalysing the conversion of glucose to glucono-1,5-lactone, the first step within the ED pathway (Giardina *et al.*, 1986; Lamble *et al.*, 2003).

From the remaining 12 down-regulated genes only the function of the alpha-mannosidase (*Sso3006*) has been verified experimentally (Cobucci-Ponzano *et al.*, 2010). EnzymeDetector

and arCOG based annotations, predict the remaining proteins as additional putative sugar dehydrogenases (*Sso3008*, *Sso3011*), a sugar phosphate isomerase (*Sso3050*), mannan endo-1,4-beta- mannosidase (*Sso3007*), glyceraldehyde oxidoreductase (*Sso3009*; large subunit), a predicted sugar phosphate nucleotidyl transferase (*Sso1782*) and a nucleotidyl-sugar reductase (*Sso1783*) (Table V). Furthermore two putative transcriptional regulators (*Sso1536* and *Sso1656*) were repressed in response to L-fucose as carbons source, suggesting them to be involved in regulation of gene expression under D-glucose nutrient conditions.

Accepted Article

The iTRAQ technique was used to investigate the global translational response of *S. solfataricus* P2 grown on the two nutrient conditions. In total, 11397 peptides corresponding to 762 proteins were detected and out of these, 424 proteins could be quantified. 26 of these quantified proteins were significantly regulated under L-fucose compared to D-glucose conditions ( $p$ -value  $< 2.36 \cdot 10^{-5}$ ; Supporting Information Dataset S5). Among these proteins, eight were significantly up-regulated under L-fucose conditions with fold changes of 2-fold or more ( $p$ -values ranging from  $3.3 \cdot 10^{-8}$  to  $1.1 \cdot 10^{-61}$ ; Table V) and given the propensity for iTRAQ to underestimate fold change (Ow *et al.*, 2009), there are likely to be others. The strongest differences in protein production were observed for the proteins Sso3124, Sso3117, Sso1300, Sso3066 and Sso3118 showing fold changes of more than 10-fold as compared to the D-glucose condition (see Table V for details of  $p$ -values). These large fold changes using iTRAQ are often indicative for “off-on” regulation. Four of these proteins (Sso3066, Sso1300, Sso3117 and Sso3124) could be directly assigned to L-fucose metabolism and were also found to be strongly induced on the transcriptome level. Finally, isocitrate lyase (Sso1333; (Uhrigshardt *et al.*, 2002; Nunn *et al.*, 2010)), a dihydroxy-acid dehydratase (Sso3107) (Kim, 2006) and a putative D-xylonate dehydratase (Sso2998) were found in more than 3-fold higher concentrations with  $p$ -values smaller than  $2.6 \cdot 10^{-13}$  under L-fucose conditions (Table V).

From the proteins showing a decreased production level under L-fucose conditions, four were down-regulated more than 2-fold with  $p$ -values ranging from  $2.6 \cdot 10^{-8}$  to  $6.2 \cdot 10^{-13}$  (Table V). These included those proteins already shown to be down-regulated on the transcript level, namely the gene cluster consisting of Sso3003, Sso3006, Sso3008 and Sso3009.

Based on their  $p$ -values, the remaining 14 proteins were found to be significantly regulated, although the corresponding fold-changes were smaller, ranging from 0.60 to 0.69 and 1.24 to 1.82 respectively (Supporting Information Dataset S5), thus suggesting a minor influence on the cell compared to those proteins with larger fold-changes. The predicted function of these proteins indicated, that they were involved in different metabolic processes (e.g. gluconeogenesis, amino acid metabolism) as well as in general cellular processes (e.g. ribosomal proteins, t-RNA synthesis, transcriptional regulation).

### ***Enzymatic activity in cell free extracts and analyses of recombinant proteins of *S. solfataricus* P2***

As transcriptome and proteome analysis showed a strong up-regulation of D-arabinose-degrading enzymes under L-fucose conditions, we were interested in their particular role in L-fucose degradation. With regard to our predicted L-fucose degradation pathway (Figure 5), cell free extracts of *S. solfataricus* P2 grown on either L-fucose or D-glucose were prepared and subjected to different activity assays. Furthermore, candidate genes were cloned and recombinantly expressed and the purified proteins were analysed for their role in L-fucose degradation:

Sugar transport: With respect to our findings from transcriptome and proteome analysis, we analysed the growth behaviour of a mutant of *S. solfataricus* PBL2025 lacking the gene for the ATP binding protein of the arabinose ABC transporter (*Sso3069*). The mutant showed a 20 % decrease in growth on L-fucose and a 60 % decreased growth rate on D-arabinose in comparison to the corresponding wild type. The growth on D-glucose was not affected (Supporting Information Table S1). This indicates that the arabinose ABC transporter is involved in the import of fucose but is not the only transporter responsible. Thus, the two transporters of the MFS superfamily (*Sso1305*; *Sso3120*) may be also involved in L-fucose uptake, although they were not induced that strongly as compared to the arabinose ABC transporter (Table V).

L-Fucose dehydrogenase reaction - The first step in the L-fucose catabolism would require L-fucose dehydrogenase activity, resulting in the formation of L-fuconolactone (Figure 5, reaction 1). Accordingly, extracts from cells grown on L-fucose showed a 23.3-times higher activity for the conversion of L-fucose to L-fuconolactone than extracts from D-glucose grown cells. Furthermore, an increased conversion of D-arabinose was observed, with a similar fold-change of 26.2 (Figure 6 A) suggesting that the L-fucose and D-arabinose oxidation is catalysed by the same enzyme. In fact, the sugar dehydrogenase *Sso1300* is strongly expressed in the presence of L-fucose. The enzyme was recombinantly produced in *E. coli* and the purified enzyme was shown to catalyse the NADP<sup>+</sup> dependent oxidation of L-fucose with higher catalytic efficiency compared to D-arabinose, confirming the previously reported results (Table VI; Supporting Information Figure S2) (Brouns *et al.*, 2006; Brouns *et al.*, 2007).

L-Fuconate dehydratase reaction - In the next step of L-fucose degradation, L-fuconate is metabolised to 2-keto-3-deoxy-L-fuconate (KDF), by action of a dehydratase (Figure 5, reaction 2). In agreement with this, crude extracts of L-fucose grown cells of *S. solfataricus* showed a 24.5-fold increased dehydratase activity using L-fuconate as substrate compared to

the D-glucose reference. Additionally, also an increased activity was observed with D-arabinonate as substrate (fold-change: 33.3; Figure 6B) again indicating that the same enzyme might be involved in D-arabinonate and L-fuconate dehydration. *Sso3124* showed significantly increased expression levels on L-fucose and could be successfully overexpressed and purified as soluble recombinant his tagged protein in *S. acidocaldarius* MW001. Previously described as arabinonate dehydratase (Brouns *et al.*, 2006), *Sso3124* showed significant but lower activities with L-fuconate as substrate compared to D-arabinonate (Table VI).

Aldolase dependent cleavage of KDF - As described for D-glucose, D-xylose and L-arabinose, sugar degradation in *S. solfataricus* proceeds via corresponding 2-keto-3-deoxy sugar acids which are often cleaved in an aldolase-dependent manner (Ahmed *et al.*, 2005; Nunn *et al.*, 2010). An aldolase-dependent cleavage of KDF would result in the formation of L-lactaldehyde and pyruvate (Figure 5, reaction 3). As this reaction is generally reversible, we tested the cell free extracts for the condensation of L-lactaldehyde and pyruvate to KDF using the TBA assay. We found activity in both, cell-free extracts of L-fucose and D-glucose grown cells, with no significant differences (fold-change: 0.83; Figure 6C). This would be consistent, if the 2-keto-3-deoxy-gluconate (KDG) aldolase (*Sso3197*), which was shown to be involved in D-glucose, D-xylose and L-arabinose degradation (Ahmed *et al.*, 2005; Nunn *et al.*, 2010), also catalyse the reversible KDF cleavage. In both, the transcriptomics and proteomics data, the *Sso3197* was detected and showed no significant differential regulation with similar fold changes as observed for cell-free extract measurements. Therefore, the *Sso3197* was heterologously overexpressed in *E. coli* and partially purified via heat precipitation. The recombinant protein catalysed the condensation reaction of L-lactaldehyde and pyruvate with a specific activity of  $1.9 \text{ U mg}^{-1}$  (Table VI; Supporting Information Figure S3). This activity is in the same range as previously observed with glyceraldehyde and pyruvate as substrates (Ahmed *et al.*, 2005).

L-Lactaldehyde dehydrogenase reaction - Aerobic degradation of lactaldehyde results in the formation of lactate and requires dehydrogenase activity (Figure 5, reaction 4). Thus, we tested for lactaldehyde dehydrogenase activity in cell free extracts using either L- or D-lactaldehyde and found that the activity with L-lactaldehyde ( $10 \text{ mU mg}^{-1}$ ) was 5-fold increased in L-fucose adapted cells compared to D-glucose. Only poor conversion of D-lactaldehyde under both conditions could be observed (Figure 6D). The 2,5-dioxopentanoate dehydrogenase (*Sso3117*) was not only strongly expressed under L-fucose conditions (fold-change: 16.4; p-value:  $1.1 \cdot 10^{-61}$ ; Table V), but has also been already shown to possess a broad substrate specificity, converting different aldehydes (Brouns *et al.*, 2006). Consistently,

the dehydrogenase activity in cell-free extracts of L-fucose grown cells towards 2,5-dioxopentanoate ( $140 \text{ mU mg}^{-1}$ ) also increased 2-fold compared to glucose grown cells. The encoding gene Sso3117 was cloned, heterologously overexpressed and purified from *E. coli* cells. The recombinant protein clearly showed dehydrogenase activity towards L-lactaldehyde with  $\text{NADP}^+$  as preferred co-substrate (Table VI). The activity with L-lactaldehyde was however  $> 90\%$  lower compared to 2,5-dioxopentanoate similarly to the activities observed for both substrates in crude extracts.

In addition to the biochemical characterization of the Sso3117 we tested the corresponding knock out mutant for its ability to grow on either L-fucose, D-arabinose or D-glucose. All growth experiments were done in the presence of 0.1 % tryptone as the *S. solfataricus* strain used for mutational analysis required some adaptation time for growth on sugars as sole carbon sources. The mutant was still able to grow in the presence of all tested sugars but compared to the wild type it showed an approximately 20 % and 30 % decreased growth rate on L-fucose and D-arabinose, respectively (Supporting Information Table S2). The growth inhibition caused by the mutation of Sso3117 is therefore comparable to those observed for the transporter mutant, indicating that other aldehyde dehydrogenases of *S. solfataricus* are also able to accept lactaldehyde as substrate.

Lactate dehydrogenase reaction - If lactaldehyde is oxidised to lactate during L-fucose degradation, it has to be converted to pyruvate for further conversion via the tricarboxylic acid (TCA) cycle (Figure 5, reaction 5). To date, there is no experimental evidence for a lactate dehydrogenase in *S. solfataricus*. However, we tested if we could observe this activity in cell-free extracts using pyruvate and NADPH as substrates, because the enzyme normally favours the formation of lactate from pyruvate. We found activity of  $4\text{-}5 \text{ mU mg}^{-1}$  under both nutrient conditions, but without significant differences between L-fucose and D-glucose extracts (fold-change: 1.2, Figure 6E). No lactate oxidizing activity could be observed neither in the soluble fraction after cell disruption nor in the membrane fraction under the described conditions using 2,6-dichloroindophenol (DCPIP) as electron acceptor neither in the presence nor in the absence of phenazine methosulfate (PMS) as an intermediate electron carrier. Also no activity could be observed in the crude extract or membrane fraction using ferrocenium hexafluorophosphate as electron acceptor.

## Discussion

*S. solfataricus* thrives at high temperatures, low pH, often in a nutrient-poor environment and has developed a number of metabolic strategies to adapt best to this hostile habitat. In the present study, we used a systems biology approach to investigate the cellular response to different carbon sources, namely L-fucose vs. D-glucose. Transcriptome and proteome studies revealed that a small set of proteins is differentially expressed under L-fucose compared to D-glucose growth conditions. In combination with metabolic modeling, metabolome analysis, activity assays in cell-free extracts, investigations of the recombinant proteins and mutational analysis, we were able to tackle cellular changes in response to both carbon sources and to identify proteins involved in L-fucose transport and metabolism. Thus, a new L-fucose degradation pathway for *S. solfataricus* P2 was identified (Figure 1A), which is to our knowledge the first description of L-fucose metabolism in archaea.

To explore different strategies for L-fucose degradation in the context of the known metabolic network of *S. solfataricus* we first used our metabolic model to test the feasibility of known L-fucose metabolic pathways. To ensure that our model assumptions resemble conditions that are as close to the *in vivo* ones as possible, we first experimentally determined the major biomass fractions for growth on D-glucose and L-fucose, respectively. The RNA content of *S. solfataricus* was found to be significantly decreased compared to those found in methanogenic archaea. However, it should be considered, that the value primarily adopted from the *M. barkeri* model (Feist *et al.*, 2006) was not experimentally determined but rather calculated based on the RNA content of *E. coli* (Neidhardt *et al.*, 1990). It is known, that the RNA content of microorganisms is strongly dependent on the growth rate and can decrease to about 10 times with decreasing growth rate (Neidhardt *et al.*, 1990). Dauner and Sauer could show, that the RNA content of *Bacillus subtilis* decreases to only 6.5 % during slow growing periods ( $\mu \sim 0.1 \text{ h}^{-1}$ ) (Dauner and Sauer, 2001). Thus, the determined RNA content for *S. solfataricus* fits very well with the strongly decreased growth speed of the organism.

The three investigated catabolic pathway variants for L-fucose degradation were shown to be energetically equally efficient, as all routes would lead to formation of 3 NADH. With respect to anabolic reactions the KDF dependent routes were slightly less efficient than the *E. coli* pathway, as additional energy is required for gluconeogenesis. However, the overall energetic differences between the three tested scenarios were found to be very small, which is reflected by comparable energy maintenance parameters (Table II). It could be suggested, that *S. solfataricus* prefers a KDF dependent pathway for L-fucose degradation, as this route omits the formation of heat labile intermediates like dihydroxyacetone phosphate and

glyceraldehyde-3-phosphate (Kouril *et al.*, 2013), thus being comparable to the non-phosphorylative branch of the Entner-Doudoroff pathway.

Our results suggest the arabinose ABC transporter (Elferink *et al.*, 2001) to be responsible for L-fucose uptake. Surprisingly, no catabolite repression at simultaneous presence of L-fucose and D-glucose was observed, although this effect was previously reported for growth of *S. solfataricus* on a mixture of arabinose and glucose (Lubelska *et al.*, 2006). Additionally, a knock out mutant was still able to utilise both, L-fucose and also D-arabinose, even though the growth rates were decreased compared to the wild type. Thus, the results indicate that at least one other transport system is able to import L-fucose (and D-arabinose) from the media. Two additional putative MFS superfamily transporter proteins were found to be strongly induced on the transcript level (Sso1305 and Sso3120). HHpred (Söding *et al.*, 2005) analysis revealed a remote homology for both transporters (E-values:  $9 \cdot 10^{-27}$  &  $7 \cdot 10^{-29}$ ) to the L-fucose MFS proton symporter from *E. coli* (Dang *et al.*, 2010), suggesting them to be involved in L-fucose uptake. Unfortunately the genetic system in *S. solfataricus* does not allow for the generation of multiple knockouts, so that we could not evaluate the role of these transporters by constructing a corresponding mutant.

The experimental results showed that the initial conversion of L-fucose proceeds via formation of L-fuconolactone and L-fuconate catalyzed by enzymes also involved in D-arabinose degradation in *S. solfataricus* (Brouns *et al.*, 2006). Therein, lactone formation is catalyzed via the highly expressed arabinose-1-dehydrogenase (*Sso1300*). Compared to D-arabinose as substrate, with L-fucose we observed higher activities in cell free extracts as well as with the purified recombinant protein, which is in agreement with previous findings (Brouns *et al.*, 2007).

After spontaneous or enzyme-catalysed hydrolysis of L-fuconolactone to L-fuconate, the latter is further converted to KDF by a dehydratase. As we observed high L-fuconate/D-arabinonate dehydratase activity only in cell extracts from L-fucose grown cells only one of the differentially expressed dehydratases is supposed to be responsible for KDF formation. From the four dehydratases which were found to be differentially expressed under L-fucose growth conditions (*Sso3124*, *Sso3118*, *Sso3107* and *Sso2998*), the arabinonate dehydratase (*Sso3124*) showed the highest overall differences in both transcriptome and proteome analyses. In contrast the induction level of the other dehydratases in extracts of L-fucose grown cells is comparatively low, so that we would expect less significant changes in cell extract activity if they play a major role in KDF formation. Accordingly, the recombinant protein Sso3124, heterologously expressed in *S. acidocaldarius*, clearly showed activity with L-fuconate and D-arabinonate as substrates, although the activities observed with the purified

proteins were found to be very low. The purified enzyme preparation seems to contain a major fraction of inactive protein, most probably due to incorrect folding. Preparation of Sso3124 has already been shown to be difficult, as expression in other host systems like *E. coli* lead to a huge amount of protein forming inclusion bodies (Brouns *et al.*, 2006).

According to the current understanding of sugar degradation in *S. solfataricus*, the organism often uses branched pathways including an aldolase dependent and an aldolase independent route (Nunn *et al.*, 2010; Bräsen *et al.*, 2014). With this background, we searched for aldolases that could be involved in the cleavage of KDF to L-lactaldehyde and pyruvate. The transcriptome data revealed four expressed genes (*Sso0286*, *Sso1254*, *Sso1333* and *Sso3197*) with predicted or proven carbon-carbon lyase activity without decarboxylase activity. However, in particular the involvement of *Sso1333* and *Sso1254* in cleavage of KDF seems to be very unlikely as *Sso1333* encodes for the isocitrate lyase (Nunn *et al.*, 2010), a key enzyme of the glyoxylate shunt with a highly different substrate specificity. Further, *Sso1254* shows high structural similarities to citryl-CoA lyases and would therefore be specific for CoA-esters as substrates. From the remaining aldolases, *Sso0286* encodes for bifunctional fructose-bisphosphate aldolase/phosphatase catalysing the condensation of dihydroxyacetone phosphate and D-glyceraldehyde-3-phosphate as part of the gluconeogenesis (Sato *et al.*, 2004; Say and Fuchs, 2010; Kouril *et al.*, 2013). Thus, only the 2-keto-3-deoxy-gluconate aldolase (*Sso3197*) is known to accept dehydrated sugar acids as substrates. Moreover, it is already known for its broad substrate specificity (Lamble *et al.*, 2003; Ahmed *et al.*, 2005; Nunn *et al.*, 2010; Bräsen *et al.*, 2014). We could prove aldolase activity in both, L-fucose and D-glucose cell extracts, using L-lactaldehyde as substrate. Further, also for the purified recombinant enzyme, activity with L-lactaldehyde and pyruvate was demonstrated. With glucose as carbon source, 2-keto-3-deoxy-gluconate aldolase is part of the degradation of D-glucose via the Entner-Doudoroff pathway, which explains the observed expression also in D-glucose extracts. However, the expression of this protein could be expected to be decreased under L-fucose conditions, if not part of L-fucose metabolism (similar to the glucose-1-dehydrogenase; Table V, Supporting Information Figure S5). Thus, no differences, neither in expression level nor in aldolase activity in cell-free extracts, were observed providing additional evidence for its involvement in L-fucose degradation.

L-lactaldehyde is further oxidized to lactate by an aldehyde dehydrogenase which was confirmed by high lactaldehyde dehydrogenase activities found only in cell free extracts of L-fucose grown cells. In addition to the arabinose-1-dehydrogenase, only the 2,5-dioxopentanoate dehydrogenase (*Sso3117*) showed an increased expression under L-fucose conditions (Table V). Accordingly, also the activity with 2,5-dioxopentanoate

appeared 2-fold higher in the cell extracts. Studies with the recombinantly expressed Sso3117 clearly proved that the enzyme also accepts L-lactaldehyde as substrate although with > 90% reduced activity compared to 2,5-dioxopentanoate. However, saturation kinetics with L-lactaldehyde could not be observed up to substrate concentrations of 26 mM (Supporting Information Figure S4). The activity towards L-lactaldehyde has not previously been reported, although the enzyme has been proven to possess activity on various substrates, including the natural substrate 2,5-dioxopentanoate as well as glycol- and D/L-glyceraldehyde (Brouns *et al.*, 2006).

Surprisingly, the knock-out mutant of the aldehyde dehydrogenase Sso3117 is still capable of growing on L-fucose and D-arabinose with 20-30 % reduced growth rates. This indicates the existence of an additional dehydrogenase capable of conversion of lactaldehyde and 2,5-dioxopentanoate. Actually, recent studies showed, that the two succinic semialdehyde dehydrogenase isoenzymes (*Sso1629* and *Sso1842*) also accept 2,5-dioxopentanoate as substrate (Esser *et al.*, 2013). This also explains the rather high activity observed with 2,5-dioxopentanoate in the cell extract of glucose grown cells, as the expression of these two genes was not affected by the carbon source. However, lactaldehyde was not tested as substrate for the two succinic semialdehyde dehydrogenases, but the observed growth phenotype of the Sso3117 mutant suggests for a certain acceptance of this compound. Additionally it may also indicate the existence of a branched pathway for L-fucose degradation, similarly to those observed for D-arabinose (Brouns *et al.*, 2006; Nunn *et al.*, 2010).

As we did not detect any lactate in the culture supernatants under both nutrient conditions this compound has to be further metabolised to pyruvate, which is then directed to the TCA cycle. In L-fucose grown cells, pyruvate is not only channelled into the TCA cycle but also serves as precursor in gluconeogenesis (Figure 7). The latter is supported by higher expression of the phosphoenolpyruvate synthase (Supporting Information Figure S5). In contrast, on D-glucose grown cells, the main flux from pyruvate is primarily directed towards the TCA cycle. Thus, the need for conversion of lactate into pyruvate under L-fucose nutrient conditions becomes obvious. However, the intracellular lactate levels did not differ substantially between the two nutrient conditions (fold change: 1.2). The enzymatic conversion to pyruvate would require activity of a lactate dehydrogenase. To date, there is no experimental evidence for this enzyme in *S. solfataricus*. However, HHpred analysis revealed remote homology of *Sso1004*, a protein with so far uncharacterized function, to the D-lactate dehydrogenase of *E. coli* (E-value:  $2.5 \cdot 10^{-55}$ ; (Dym *et al.*, 2000). The latter was shown to be a peripheral membrane protein, which also accepts L-lactate, although with lower specificity

(Kohn and Kaback, 1973). Furthermore, the operon consisting of *Sso3161*, *Sso3163*, *Sso3165* shares similarities (up to 38 %) to all three glycolate oxidase subunits from *E. coli*, which was shown to be active with D- and L-lactate (Lord, 1972). Both proteins from *S. solfataricus* (*Sso3161*, *Sso3163*, *Sso3165* and *Sso1004*) are supposed to be FAD dependent and soluble enzymes as a FAD binding site and no transmembrane regions are predicted from the sequence. However, we could not observe significant FAD-dependent lactate oxidizing activity under the applied conditions, neither in the soluble nor in the membrane protein fraction, although FAD-dependent lactate oxidation has been reported for the closely related organism *S. tokodaii* (Satomura *et al.*, 2008). Nevertheless, with NAD as final electron acceptor we could observe lactate dehydrogenase activity in the cells extracts, indicating that *S. solfataricus* at least possesses a NAD dependent enzyme.

The results presented in this work clearly demonstrate that L-fucose degradation in *S. solfataricus* occurs via formation of KDF, followed by aldolase dependent cleavage to lactaldehyde (Figure 1A). However, this pathway could not explain the strong increase in expression of the 2-keto-3-deoxy-D-arabinonate dehydratase (*Sso3118*) observed under L-fucose conditions. As D-arabinose degradation in *S. solfataricus* requires the activity of *Sso3118*, its strong expression under L-fucose growth conditions could be a result of the simultaneous expression along with the divergent organized *Sso3117*. A detailed analysis of the genome sequence showed that only 73 base pairs separate these two differentially orientated genes. Even though both genes seem to have their own TATA box located upstream of the transcription start site, only one so-called ARA box could be found within the intergenic region between them. This palindromic motif was found to be present within the promoter regions of all genes important for D-arabinose degradation in *S. solfataricus* and was therefore suggested to act as recognition element for coordinated transcription of these genes (Brouns *et al.*, 2006). Thus, it can be suggested that the induction of *Sso3117* by the presence of L-fucose either causes a simultaneous induction of *Sso3118* or it may give another hint towards the existence of a branched pathway for L-fucose degradation. Branched pathways are often found for sugar degradation in *S. solfataricus* (e.g. L-arabinose, D-xylose (Nunn *et al.*, 2010)), often including an aldolase dependent and an aldolase independent route. Exploring the existence of a potential second pathway in L-fucose degradation is currently the subject of further studies.

In addition to the pathway discussed above, our results also suggest an increased activity of the 3-hydroxypropionate/4-hydroxybutyrate cycle. This was quite surprising, as this pathway is supposed to be active only under autotrophic growth conditions (Berg *et al.*, 2007). However, our model simulations showed that an increased flux through parts of the cycle,

namely formation of succinyl-CoA from propanoyl-CoA via fixation of one molecule of carbon dioxide, is feasible during growth on L-fucose compared to growth on D-glucose (Table II). Furthermore, this part of the pathway becomes essential in the L-fucose growth scenario if we forbid the flux through the glyoxylate bypass in the optimal flux variance analysis (biomass flux: 100 %), otherwise the biomass production would decrease. Considering that the anaplerotic reaction catalyzed by the phosphoenolpyruvate carboxylase (EC 4.1.1.31) is not favourable during growth on L-fucose (as it would withdraw phosphoenolpyruvate required for gluconeogenesis) it could be suggested that the lower part of the 3-hydroxypropionate/4-hydroxybutyrate cycle is used in *S. solfataricus* for replenishment of the TCA cycle via formation of succinyl-CoA. This would be in agreement with the previous finding, that succinyl-CoA and not acetyl-CoA is the key precursor for biosynthesis of central carbon metabolites during autotrophic growth of *Metallosphaera sedula* (Estelmann *et al.*, 2011). However, in the metabolic network of *S. solfataricus* the 3-hydroxypropionate/4-hydroxybutyrate cycle is also the only known pathway involving the production of malonyl-CoA. This intermediate of the first carbon fixation reaction, was detected only in cells grown with L-fucose as carbon source. Moreover, the following enzyme, malonyl-CoA reductase (*Sso2178*) showed an increased expression on the transcript level (RPKM Fuc: 152 vs RPKM Glc: 38). Although, further intermediates from the first part of the 3-hydroxypropionate/4-hydroxybutyrate cycle (e.g. 3-hydroxypropanyl-CoA, acrolyl-CoA) could not be detected in the extracts, all enzymes (except the malonate semialdehyde reductase (*Sso0647*)) involved in the reactions to propanoyl-CoA were found to be expressed at relatively high levels in cells grown on L-fucose (Supporting Information Dataset S4). Considering that during growth on L-fucose less carbon is integrated into the biomass (40 % vs 47 % on D-glucose) and no secretion of compounds was observed it could be suggested that the carbon fixation reactions of the cycle could be used to restore small amounts of carbon dioxide. However, the data presented in this work is not sufficient to unravel the role of the 3-hydroxypropionate/4-hydroxybutyrate cycle for *S. solfataricus* in detail.

The systems biology approach applied in this work revealed only little further changes in the metabolic network of *S. solfataricus* depending on the nutrient conditions. As expected, the strongest differences were observed in such metabolic pathways which depend on the availability of glucose. This is not only reflected by a strong decrease in intracellular glucose levels in cells growing on L-fucose, but also in the down regulation of several sugar converting enzymes, like sugar dehydrogenases (*Sso3008*, *Sso3011*) or the putative glyceraldehyde oxidoreductase (*Sso3009*; large subunit) which is supposed to be involved in the glycolytic Entner-Doudoroff pathway in *S. tokodaii* (Wakagi *et al.*, 2002; Wakagi *et al.*,

2016). Furthermore, a distinct decrease in the intracellular trehalose concentration was observed in L-fucose grown cells (Figure 3). This sugar serves as main compatible solute in *S. solfataricus* and can be synthesized either by cleavage of glycogen via the TreY/TreZ-pathway (Maruta *et al.*, 1996) or directly from glucose and NDP-glucose via the trehalose synthase (TreT) (Qu *et al.*, 2004; Kouril *et al.*, 2008). Under L-fucose conditions trehalose synthesis is much more expensive as gluconeogenesis is necessary to provide the precursor molecules. The induction of several gluconeogenetic enzymes (e.g. PEP synthase, phosphoglycerate kinase) on the protein level reflects this higher need towards glucose synthesis (Supporting Information Dataset S4 & S5 and Figure S5). Additionally the repression of the sugar-nucleotidyl transferase (*Sso1782*) may be connected to trehalose synthesis via the TreT pathway as this protein would provide the precursor NDP-glucose. The decreased availability of trehalose under L-fucose growth conditions may be antagonized by increased putrescine content (Figure 3), as polyamines are likewise known to act as stabilizers of macromolecules (Bachrach, 2005; Kalisiak *et al.*, 2009).

Finally, we observed an increased expression of the isocitrate lyase, a key enzyme of the glyoxylate bypass (Table V). The second enzyme of this cycle, the malate synthase (*Sso1334*) is present under both conditions with no differences in expression. Similar induction patterns of the glyoxylate bypass enzymes have been found in *S. solfataricus* during growth on D-arabinose (Brouns *et al.*, 2006). Activity of this pathway is normally only required if C<sub>2</sub>-providing compounds serve as the main carbon source. However, high activities of both glyoxylate bypass enzymes have been already reported for *S. acidocaldarius* grown on glucose as sole carbon source (Uhrigshardt *et al.*, 2002). Our model simulations allowed relatively high fluxes through the glyoxylate bypass in all tested scenarios with the highest flux allowed in the KDF dependent L-fucose degradation scenario (Table II). Activity of the glyoxylate bypass during growth on L-fucose provides another possibility for replenishment of the TCA cycle. Again, this avoids the anaplerotic reaction catalyzed by phosphoenolpyruvate carboxylase and would be therefore highly favorable under L-fucose growth conditions as phosphoenolpyruvate is required for gluconeogenesis.

## Conclusion

In this work, we describe the response of *S. solfataricus* P2 to different carbon sources and present a new pathway for the uptake and degradation of L-fucose in the thermoacidophilic archaeon. It is also the first detailed description of L-fucose metabolism in hyperthermophilic archaea. Interestingly, L-fucose degradation in *S. solfataricus* is shown to be achieved by promiscuous enzymes, which are partially also involved in the catabolism of D-arabinose, a

sugar not commonly found in nature. To our knowledge, pathway promiscuity for L-fucose and D-arabinose has not been reported before in any organism. The results demonstrate that the organism optimized its metabolism in the course of evolution and is able to use a small set of proteins for utilisation of different carbon sources. This represents an effective metabolic strategy, allowing rapid adaptation to changes in nutrient conditions. Thus, it provides an excellent example of the broad metabolic versatility and the excellent adaptation skills of *S. solfataricus* to a hostile environment.

Accepted Article

## Experimental Procedures

### *Strain and growth conditions*

If not indicated otherwise, *Sulfolobus solfataricus* P2 (Zillig *et al.*, 1980) was grown in defined minimal medium (Brock *et al.*, 1972) containing either 4.4 g·l<sup>-1</sup> D-glucose or 2.5 g·l<sup>-1</sup> L-fucose as sole carbon source. Pre-heated long neck flasks (1000 ml or 500 ml, medium volume 1/5) were inoculated with 200 µl of glycerol stocks (OD<sub>600</sub>: 10), obtained from mid-exponential *S. solfataricus* cultures grown on either D-glucose or L-fucose and prepared as described by Zaparty *et al.* (2010). The cultures (initial OD<sub>600</sub>: 0.01-0.02) were incubated at 80 °C, pH 3.5 and 160 rpm (Thermotron, Infors AG, Switzerland).

The *S. solfataricus* arabinose transporter mutant PBL2025ΔSSO3069 and the 2,5-dioxopentanoate dehydrogenase mutant PBL2025ΔSSO3117 were constructed using homologous recombination as described by Albers and Driessen (Albers and Driessen, 2008) with the primers listed in the Supporting Information (Table S5).

The transporter mutant and the corresponding reference strain PBL2025 (Scheclert *et al.*, 2004) were grown in minimal medium containing carbon equivalent amounts of either, D-glucose (15 mM), L-fucose (15 mM) or D-arabinose (18 mM). Glycerol stocks (OD<sub>600</sub>: 10) obtained from mid-exponential cells grown on 0.2 % tryptone were used for inoculation (initial OD<sub>600</sub>: 0.01-0.02). A minimum of three independent cultures were carried out for each condition.

For growth analysis of PBL2025ΔSSO3117 and the wild type three precultures were used to adapt the cells to the corresponding growth conditions. First, 0.2 % tryptone containing minimal medium were inoculated with tryptone glycerol stocks. This culture was used to inoculate a second preculture, containing 0.1 % tryptone and 15 mM of either D-glucose, L-fucose or D-arabinose. After adaptation to the corresponding sugars a third preculture, containing the same medium were inoculated. The latter was used to inoculate the main culture (100 ml long-neck flask, 50 ml medium) also containing 0.1 % tryptone and 15 mM of one of the appropriate sugars as carbon source (initial OD<sub>600</sub>: 0.01-0.02). Cultures growing with 0.1 % tryptone served as control. All cultures were grown at 76 °C and 120 rpm. Three independent cultures were carried out for each condition.

For analysis of catabolite repression, 3 g·l<sup>-1</sup> D-glucose and 2.5 g·l<sup>-1</sup> L-fucose (both 15 mM) were used as carbon sources and glycerol stocks from tryptone grown cells were used for inoculation (initial OD<sub>600</sub>: 0.01-0.02). Three independent cultures were carried out for each condition.

To determine the stability of the substrates under cultivation conditions additional control cultures were made as described for the catabolite repression experiments but without addition of microorganisms. For GC-MS analysis, 1 ml of supernatant was taken at regular intervals over a whole cultivation period of 125 h.

***Physiological parameters (biomass, biomass composition and extracellular substrate concentrations)***

Cell growth was monitored following the optical density at 600 nm (OD<sub>600</sub>). The specific growth rate was determined by linear regression of time-dependent changes in cell dry weight from at least six data points during the exponential growth phase.

To determine external concentrations of D-glucose and L-fucose, 1 ml samples were taken in regular intervals and centrifuged (20,200 x g, 3 min, RT). The supernatant was analysed for D-glucose concentrations using a commercially available D-glucose UV-assay kit (R-Biopharm, Germany). For quantification of L-fucose 6 µl of the supernatant were mixed with 20 µl ribitol (c = 0.2 g·l<sup>-1</sup>) as internal standard, dried under vacuum and derivatised as described by (Zaparty *et al.*, 2010)). For absolute quantification, a calibration curve with pure L-fucose (Roth, Germany) (2, 4, 6, 8, 10 and 15 mM) was used.

For analysis of catabolite repression, 7 µl of culture supernatant were mixed with 10 µl succinate (c = 1 mM) as internal standard, dried under vacuum and derivatised as described above.

Samples were analysed by GC-MS on a Leco Pegasus 4D GCxGC TOF-MS (Leco, Germany). GC-MS analyses were performed as described below. Chromatograms were processed using the MetaboliteDetector software for targeted analysis (Hiller *et al.*, 2009).

The substrate uptake rates were calculated using equation (1), where q<sub>S</sub> represents the specific substrate uptake rate and Δc and ΔCDW describe the time-dependent (Δt) changes in substrate concentration and cell dry weight during the exponential growth phase. All measurements were conducted from at least three independent cultivations.

$$q_S = \frac{\Delta c}{\left(\frac{\Delta CDW}{2}\right) \cdot \Delta t} \quad (1)$$

Quantification of total DNA content was done using the *Genomic DNA from tissue Kit* from Macherey&Nagel (Germany). Cells (1 mg dry weight) were resuspended in 180 µl TE buffer (20 mM Tris-HCl, 2 mM EDTA, 1 % Triton X 100, pH 8) containing 20 mg·ml<sup>-1</sup> lysozyme and incubated in an ultrasonic bath at 37 °C for 60 minutes. After lysis 25 µl Proteinase K

were added and incubated at 56 °C with continuous shaking overnight. For depletion of RNA 20 µl preheated (15 min, 95 °C) RNase A (1 mg·ml<sup>-1</sup>) was added and incubated for 5 minutes at room temperature. Further isolation and purification of DNA was performed according to the manufacturer's protocol. Isolated DNA was quantified by measuring absorbance at 260. Absorbance at 280 nm was measured to check for purity of the DNA samples.

For total RNA quantification, cells (0.5 mg dry weight) were resuspended in TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8) containing 2 mg·ml<sup>-1</sup> lysozyme and incubated in an ultrasonic bath at 37 °C for 60 minutes. Isolation and purification of total RNA was performed using the *RNA isolation Kit* from Macherey&Nagel following the manufacturer's instructions. Isolated RNA was quantified by measuring absorbance at 260 nm. Absorbance at 280 nm was measured to check for purity of the RNA samples.

For quantification of total protein fractions, cells (0.5 mg dry weight) were resuspended in lysis puffer (0.5 M NaOH, 2 % SDS) and incubated at 80 °C for 30 minutes in an ultrasonic bath. After centrifugation (10,000 x g, 5 min, 20 °C) the protein concentration of the supernatant was determined using the *Bicinchoninic acid Protein Assay Kit* (Sigma Aldrich, Germany) following the manufacturer's instructions.

### ***Cell harvest and sample processing***

Samples for transcriptomics (10 ml), proteomics (50 ml), intracellular metabolome analysis by GC-MS and HPLC-MS measurements (15 mg and 10 mg cell dry weight, respectively) and enzyme activity measurements (50 ml) were harvested in mid-exponential growth phase at OD<sub>600</sub> = 0.6-0.8, which corresponds to approximately 75 hours of cultivation. To gain homogeneity of the samples, cultures of an overall volume of 2 l were combined prior harvesting. Cell samples for RNA isolation, protein extraction and activity measurements were immediately cooled down in liquid nitrogen for 90 s and harvested by centrifugation (7 min, 6000 x g, 4 °C). After discarding the supernatant, the cell pellets were stored at -80 °C until further processing (as described below).

For metabolome analyses the cell samples were cooled down on ice for 2 min prior centrifugation (7 min, 6000 x g, 20 °C). To extract Coenzyme A (CoA) derivatives for HPLC-MS measurements cells were re-suspended in 1 ml methanol, which contained 0.2 mg·l<sup>-1</sup> biochanin A as internal standard. The suspensions were transferred into 2 ml Precellys tubes (Peqlab, Germany) containing 0.6 g ± 0.03 glassbeads (70 – 110 µm diameter; Kuhmichel Abrasiv GmbH, Germany) and stored at -80 °C until further processing. The extraction of intracellular metabolites was carried out immediately after cell harvest (see below).

## ***Metabolomics***

### ***Extraction of intracellular metabolites for GC-MS analysis***

To quench metabolism and extract intracellular metabolites, the following procedure was applied (modified from Zaparty *et al.*, 2010). After harvesting (see sample processing), the cells were washed twice with 20 ml 0.9 % NaCl (w/v) at room temperature (centrifugation 5 min at 6000 x g, 20 °C). Afterwards cells were quenched in 1.5 ml methanol containing 30 µl ribitol ( $c = 0.2 \text{ g}\cdot\text{l}^{-1}$ ) as an internal standard and lysed in an ultrasonic bath (70 °C, 15 min). Samples were cooled on ice (2 min) and 1.5 ml of deionised water was added and mixed by vortexing for 2 min. Intracellular metabolites were extracted by adding 1 ml chloroform to the samples. After vortexing the samples for 2 min, phase separation was carried out by centrifugation (18,514 x g, 5 min, 4 °C). The upper, polar phase was dried in a vacuum concentrator (SpeedVac, Labconco, Houston, USA) with rotation at 15 °C overnight. The dried extracts were derivatised following the protocol of Zaparty *et al.* (2010) prior to metabolite quantification by GC-MS.

### ***GC-MS analysis of intra- and extracellular metabolites***

Intracellular metabolites, extracellular concentrations of D-glucose and L-fucose during the catabolite repression experiment and L-fucose supernatants for the determination of the L-fucose uptake rate were quantified on a Leco Pegasus 4D GCxGC TOFMS (Leco, Germany) using 1D mode as described by Abu Dawud *et al.* (2012). Gas chromatography was performed on a 30 m, 0.25 mm Zebron ZB-5ms column (Phenomenex, Germany) using 1.2 ml·min<sup>-1</sup> helium flow. Ionisation of analytes was done by positive electron ionization (EI+) at 70 eV. After 0.02 min at 70 °C the temperature was increased to 330 °C with 12 °C·s<sup>-1</sup>, followed by an additional constant temperature period at 330 °C for 5 min. Full scan mass spectra were obtained from 45 to 600 m/z with 8 scans·s<sup>-1</sup> and a solvent delay time of 260 s.

### ***GC-MS data processing and identification***

GC-MS data were processed using the MetaboliteDetector software (version 2.2.N-2013-01-15; (Hiller *et al.*, 2009)) for automated peak detection and deconvolution (Zech *et al.*, 2013). Retention indices of compounds were calculated by comparison to an alkane time standard (van Den Dool and Dec. Kratz, 1963). Identification of compounds was carried out in non-targeted mode by comparison of their specific mass spectra and retention indices to a compound library. This library was created by merging our in-house library with the Golm metabolome database (Kopka *et al.*, 2005). Non-biological and artificial peaks were

eliminated by the aid of blanks. All peak areas were normalised to the internal standard ribitol and peak areas of derivatives were summarised to the corresponding metabolites. Finally, the data was normalised by a central normalisation to the reference condition (D-glucose). Significant changes in metabolite levels were calculated by non-parametric Kruskal-Wallis test (Kruskal and Wallis, 1952);  $p$ -value  $< 0.01$ ) using Benjamini–Yekutieli correction (Benjamini and Yekutieli, 2001) to control the false discovery rate.

#### ***Extraction of intracellular Coenzyme A-derivatives***

Coenzyme A-esters were isolated from frozen cell pellets (see sample processing). Cells were lysed using a Precellys 24 homogeniser (Peqlab, Germany) at  $-10$  °C. The procedure included three cycles of homogenisation (6800 rpm, 30 s with equivalent breaks). The lysate was transferred to 10 ml of ice cold ammonium acetate (25 mM, pH 6) and centrifuged (5 min at 10,000  $\times$  g, 4 °C). CoA-derivatives were extracted on a Strata XL-AW solid phase extraction column (Phenomenex, Germany). The column was equilibrated with 1 ml methanol followed by 1 ml methanol: H<sub>2</sub>O: formic acid (50: 45: 5) and 1 ml H<sub>2</sub>O. Cell lysate was loaded onto the column (800-900 mbar vacuum, 5 min), washed with 1 ml ammonium acetate (25 mM, pH 7.2) and 1 ml methanol, followed by short drying. CoA-esters were eluted with 1 ml methanol containing 5 % (v/v) ammonia. The eluate was dried in a vacuum concentrator (SpeedVac, Labconco, USA) with rotation (15 °C overnight).

#### ***HPLC-MS measurement of Coenzyme A-derivatives***

Buffers, eluents, column and flow rates for HPLC-MS analysis of CoA-derivatives were used as described by Peyraud *et al.* (2009). Dried extracts were resolved in 200  $\mu$ l sample buffer (25 mM ammonium acetate pH 3.5, 2 % methanol) and 50  $\mu$ l were injected on a Dionex ultimate 3000 system (Thermo Scientific Inc., Germany) coupled to a Bruker MicroTOF QII mass spectrometer (Bruker Daltonik GmbH, Germany) equipped with an electrospray ionisation interface. Separation of CoA-intermediates was carried out on a C<sub>18</sub> analytical column (Gemini 150 · 2.0 mm, particle size 3  $\mu$ m; Phenomenex, Germany) at a constant temperature of 35 °C using the following gradient: 1 min 5 % B (methanol): 95 % A (50 mM ammonium formiate, pH 8.1), 18 min gradient to 30 % B: 70 % A, 7 min gradient to 95 % B: 5 % A and a final step at 95 % B: 5 % A for 4 min. MS analysis was done using positive ESI mode with 3 Hz data acquisition and automated MS<sup>2</sup> acquisition with full scan mass spectra from 90 to 1178 m/z. A detailed list of applied MS and MS<sup>2</sup> settings can be found in the Supporting Information (Table S3).

### ***HPLC data processing and peak identification***

Data export to mzXML format and internal mass calibration using the sodium formiate cluster was carried out with DataAnalysis software (version 4.0 SP 5 (Build 283)). Raw data was processed using the XCMS package (Smith *et al.*, 2006; Tautenhahn *et al.*, 2008; Benton *et al.*, 2008) for R. After peak detection the XCMS methods 'group' and 'rector' in two iterations were used for peak alignment and retention time correction and missing values were estimated using the 'fill peaks' method. Detailed parameters of data processing are given in the Supporting Information (Table S4).

Retention times of available CoA standards and the accurate masses of  $M + 2H$  ions were used for peak identification. If no synthetic standard was available for a certain compound, both calculation of molecular mass from  $M + 2H$  and  $M + H$  ions, sum formula prediction from the accurate mass and isotopic pattern were applied for identification, using the DataAnalysis software. Whenever possible,  $MS^2$  fragmentation was used to confirm the presence of the CoA moiety. Data were normalised by a central normalisation to the reference condition (D-glucose) and significant changes in metabolite levels were calculated by non-parametric Kruskal-Wallis test (Kruskal & Wallis, 1952);  $p$ -value  $< 0.01$ ) using Benjamini–Yekutieli correction (Benjamini & Yekutieli, 2001) to control the false discovery rate.

### ***Transcriptomics***

#### ***RNA isolation, sequencing and data analysis***

RNA was isolated from frozen cell pellets (three biological replicates for each condition). Samples were resuspended in 800  $\mu$ l RLT buffer from *RNeasy Mini Kit* (Qiagen, Germany). Cells were disrupted by means of a Ribolyser instrument (Hybaid, Germany) in 30 s at speed-level 6.5. RNA isolation was performed according to the manufacturer's protocol. RNA samples were treated with RNase-free DNase (Qiagen) once in the microtube and a second time on the column followed by purification according to the clean-up protocol.

Ribosomal RNA was depleted using a *RiboZero magnetic kit* for bacteria (Epicentre, USA) with a modified protocol. Only 90  $\mu$ l instead of 225  $\mu$ l magnetic beads were used and for the rRNA removal reaction only 1  $\mu$ g RNA was mixed with 4  $\mu$ l removal solution in a total volume of 20  $\mu$ l. Sequencing libraries for all six samples were prepared with the *TruSeq<sup>®</sup> Stranded mRNA LT kit* (Illumina, USA) starting with the RNA fragmentation step after elution of precipitated RNA in 19  $\mu$ l of the Fragment-Prime-Mix.

Sequencing libraries were quantified with a High-Sensitivity Chip on a Bioanalyzer (Agilent, Germany) and a measurement with a *Quant-iT™ PicoGreen® dsDNA Assay Kit* (Invitrogen, USA) on a Microplate Reader Tecan Infinite 200 (Tecan, Switzerland). Sequencing was performed with a *MiSeq® Reagent Kit v3* (150 cycle) (Illumina) on a *MiSeq®* instrument (Illumina).

Sequencing reads were mapped with Bowtie2 (Langmead and Salzberg, 2012) against the reference genome (*Sulfolobus solfataricus* P2, genome size: 2,992,245 nt, RefSeq ID: NC\_002754.1). Data visualisation and calculation of mapped reads per gene were performed using ReadXplorer (Hilker *et al.*, 2014). Subsequently the *reads per kilobase of gene per million reads* (RPKM) (Mortazavi *et al.*, 2008) were computed for each gene. For comparison, the average values of the RPKMs were calculated for samples grown in medium with D-glucose and for samples grown in medium with L-fucose. A mean log intensity (*A*-value; (Dudoit *et al.*, 2002)) cut off of 2.0 was used to distinguish between background activity and active expression. For determination of regulated genes statistical analysis was performed using DeSeq (Anders and Huber, 2010).

## ***Proteomics***

### ***Protein extraction and iTRAQ labelling***

Proteins were extracted from frozen cell pellets (see sample processing). Cells were washed with cold water before being re-suspended in 1 M tetraethylammonium bromide (TEAB) pH 8.0 containing 0.05 % SDS. Proteins were extracted using glass beads (size 425-600 µm, Sigma Aldrich, UK) in a disruptor (Disruptor Genie, USA) for 8 cycles (alternatively 45 s vortexing and 45 s incubation on ice). Samples were centrifuged at 16,400 x g for 15 min (4 °C) and the supernatants were transferred into new Eppendorf tubes. Total protein concentrations were quantified using the Bradford assay (Bradford, 1976).

100 µg protein of each sample was reduced, alkylated, digested and labelled with iTRAQ 8-plex reagents as described in detail elsewhere (Zaparty *et al.*, 2010). Biological triplicates of each growth condition were used and these samples were labelled with iTRAQ 8-plex reagents as follows (6 reagents used): 113, 114 and 115 were used for cells grown on L-fucose whilst 116, 117 and 118 were used for cells grown on D-glucose (and used as control). Samples were combined before drying in a vacuum concentrator (Concentrator 5301, Eppendorf, Germany).

### ***Peptides separation, mass spectrometry and data analysis***

Dried labelled iTRAQ peptides were fractionated using a HILIC technique on an uHPLC 3000 system (Dionex, UK) operated at a flow rate of  $0.4 \text{ ml}\cdot\text{min}^{-1}$  and a wavelength of 280 nm (Ow *et al.*, 2010). Samples were re-suspended in buffer A (10 mM ammonium formate in 80 % acetonitrile pH 3) before being injected into a  $4.6 \times 200 \text{ mm}$  Poly HYDROXYETHYL-A column (Hichrom Limited, UK). Peptides were separated using a gradient as follows: 5 min of 3 % buffer B (10 mM ammonium formate in 5 % acetonitrile pH 5) followed by a ramp to 25 % buffer B for 10 min, then up to 95 % buffer B for 35 min and finally maintained at 95 % for 10 min before ramping back to 3 % buffer B for 10 min. Fractionated peptides were collected every minute and then dried in a vacuum concentrator (Concentrator 5301, Eppendorf, Germany).

Selected fractions were cleaned using micro spin  $C_{18}$  column (Nest Group, USA). Samples were resuspended in buffer A containing 0.1 % formic acid and 3 % acetonitrile prior to submission to a QStar XL Hybrid ESI Quadrupole time-of-flight tandem mass spectrometer, ESI-qQ-TOF-MS/MS (Applied Biosystems / MDS Sciex (now ABSciex), Canada), coupled with a nano-LC system (LC Packings Ultimate 3000, Dionex, UK). Mass spectrometry parameters were applied as described by Zaparty *et al.* (2010). Briefly, peptides were separated on a PepMap  $C_{18}$  reversed phase capillary column (LC Packings) at a flow rate of  $3 \mu\text{l}\cdot\text{min}^{-1}$  with a gradient generated by increment of buffer B containing 0.1 % formic acid and 97 % acetonitrile. The electrospray ionisation voltage was set at 270 V whilst the mass detector range was set to 350 to 1600 m/z and operated in the positive ion mode. Peptides with +2, +3, and +4 charge states were selected for fragmentation.

Raw mass spectrometry data were directly submitted to Mascot Daemon V2.2.0 with the iTRAQ 8-plex option. The search parameters were set up as follows with MS and MS/MS tolerances: 0.8 and 0.4 Da respectively, ion charges +2, +3 and +4, trypsin used with up to two missed cleavages. Variable modification of methionine and fixed modification of MMTS, iTRAQ 8-plex N-terminal and K were used. Data were searched against the *S. solfataricus* P2 database (2972 entries) downloaded from NCBI (Aug 2014). Furthermore, the false positive rates were also performed by searching the data with a reversed database of *S. solfataricus* P2. Results from the Mascot Daemon were exported into an excel file before submission to our in-house proteomic pipeline (Pham *et al.*, 2010; Bewley *et al.*, 2011) for quantitation and assessment of regulated proteins. The protein identification was obtained from Mascot results while peak areas obtained from Mascot results were used for quantitation. Subsequently a rigorous statistical method was used to determine regulated proteins (Pham *et al.*, 2010).

### ***Activity measurements in cell free extracts***

#### ***Preparation of *S. solfataricus* crude extracts***

Cells (15 mg dry weight; see sample processing) were resuspended in 3 ml 0.1 M TRIS/HCl buffer (pH 7, RT, containing 5 mM DTT and 500  $\mu$ l complete Protease Inhibitor (7 x, Roche). and disrupted by sonication (4 x 2 min pulse/ 1 min cooling). After centrifugation (45 min, 21,114 x g, 4 °C) the supernatant was dialysed (Spectra/Por Dialysis Tubing, MWCO 3.500, Spectrum Laboratories, Inc., Netherlands) overnight at 4 °C against 0.1 M TRIS/HCl buffer (pH 7, RT) with 5 mM DTT. The protein concentration was determined using the Bradford assay (Bradford, 1976).

All enzymatic activity assays were performed with 300  $\mu$ g protein at 70°C in 0.1 M TRIS/HCL buffer (pH 6.5) containing 5 mM MgCl<sub>2</sub> if not stated otherwise. A specord 210 photometer (Analytik Jena, Germany) was used for photometric analysis. Three independent measurements were performed for each assay with control reactions without substrate or crude extract, respectively.

#### ***Dehydrogenase activity assays***

Absorbance measurements were made on Specord 210 (Analytic Jena) photometers.

Dehydrogenase activity was determined in a total volume of 0.5 ml. Initial enzymatic activity rates were obtained by following the formation/depletion of NADH or NADPH, respectively. Three independent measurements were performed for each assay with control reactions without substrate or crude extract, respectively.

L-fucose dehydrogenase assays were performed in presence of 2 mM NADP<sup>+</sup> and started by the addition of 5 mM L-fucose or D-arabinose.

The formation of lactate from pyruvate via lactate dehydrogenase was determined in presence of 0.2 mM NADPH and 5 mM pyruvate was used to start the reaction.

Aldehyde dehydrogenase reactions were performed in 50 mM sodium phosphate buffer (pH 6.0 at 70 °C) containing. 5 mM MgCl<sub>2</sub> and 2 mM NADP<sup>+</sup>. Activity was initiated by the addition of 5 mM L-lactaldehyde, D-lactaldehyde or 2,5-dioxopentanoate, respectively.

NAD-independent lactate dehydrogenase was assayed at 60°C and 600 nm as DCPIP-dependent oxidation of lactate to pyruvate. The assay was performed in 200 mM Tris-HCl buffer (pH 8.5) containing 10 mM L-lactate, 0.1 mM DCPIP and 300  $\mu$ g protein (membrane fraction or crude extract) according to the protocol reported by Satomura *et al.* (2008). For the assays using PMS (1.5 mM) as intermediate electron carrier, the membrane fraction or crude extract were pre-incubated for 5 min at 70°C in 100 mM TRIS-HCL buffer (pH 7.0) in the presence of 1.5 mM PMS and 30 mM L-lactate and the reaction was started by addition of 0.2

mM DCPIP as described previously (Kolaj-Robin *et al.*, 2011). Experiments with ferrocenium hexafluorophosphate ( $\text{Fc}^+\text{PF}_6^-$ ) were conducted at 25 °C in 100 mM potassium phosphate buffer (pH 7.2) containing 300 µg crude extract or membrane fraction, 0.2 mM  $\text{Fc}^+\text{PF}_6^-$  (dissolved in 10 mM HCl) and 10 mM L-lactate according to Lehman *et al.* (1990). The reaction was started by addition of the protein extract and the dehydrogenase activity was followed by the decrease in absorbance at 300 nm derived from the reduction of ferricenium ion.

#### ***Dehydratase activity assay***

Enzyme reactions were performed in presence of either 5 mM L-fuconate or D-arabinonate in 100 mM MES/KOH buffer (pH 6.5 at 70 °C). After 0, 5, 10, 15, 20, 25 and 30 min, 100 µl of the sample were transferred on ice. For the determination of 2-keto-3-deoxy-L-fuconate the reaction was stopped by the addition of 10 µl of 12 % (w/v) trichloroacetic acid. Precipitated protein was removed by centrifugation (16,000 x g, 15 min, 4 °C). Enzymatic activity was determined by using the modified TBA-assay (Buchanan *et al.*, 1999).

#### ***Aldolase activity assay***

Aldolase activity was performed in 50 mM sodium phosphate buffer (pH 6.0) containing 5 mM  $\text{MgCl}_2$ , 5 mM pyruvate and 5 mM L-lactaldehyde or D-lactaldehyde at a total volume of 1 ml. After 0, 5, 10, 15, 20, 25 and 30 min, 100 µl of the sample were withdrawn and stored on ice. The reaction was stopped by the addition of 10 µl of 12 % (w/v) trichloroacetic acid. Precipitated protein was removed by centrifugation (16,000 x g, 15 min, 4 °C) and the amount of formed 2-dehydro-3-deoxy-L-fuconate was analysed by using the modified TBA-assay (Buchanan *et al.*, 1999).

#### ***Preparation of recombinant proteins***

##### ***Construction of expression plasmids***

The coding regions *Sso1300*, *Sso3124*, and *Sso3117* were amplified from genomic DNA of *S. solfataricus* P2 by PCR mutagenesis using Phusion polymerase (Thermo scientific), using the primer sets given in the Supporting Information (Table S6). The amplified DNA fragments were cloned into the respective vectors using the restriction endonucleases also depicted in Table S4. Successful cloning was confirmed by sequencing. Cloning of *Sso3197* has already been described in (Ahmed *et al.*, 2005).

### ***Heterologous expression and protein purification***

Expression of Sso1300 was carried out in *E. coli* BL21 (DE3) RIL transformed with pET15b: Sso1300, and of Sso3117 in *E. coli* Rosetta (DE3) transformed with pSVA3327. Cells were grown at 37°C and 180 rpm in LB medium (pH 7; LB broth, Sigma-Aldrich) supplemented with 100 µg/ml ampicillin and 34 µg/ml chloramphenicol. At an optical density of 0.6-0.8 expression was induced by 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). After continued growth overnight at 22°C and 180 rpm cells were harvested by centrifugation (15 min, 6,000 x g, 4°C) and stored at -80°C.

For expression of Sso3124 *S. acidocaldarius* MW001 was transformed with pSVA1576 as described in Wagner *et al.* (2012). Cells were grown at 76 °C and 180 rpm in Brock minimal medium (pH 3) containing 0.1 % NZA and 0.3 % dextrin. At an optical density of 0.8-1.0 cells were harvested by centrifugation (15 min, 6,000 x g, RT) and stored at -80 °C.

For purification of Sso1300, Sso3124 and Sso3117, cells were resuspended in buffer A (50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, and 7.5 mM DTT, pH 8.0) at a ratio of 3 ml/g cells and disrupted by sonication (3x, 10 min pulse). After centrifugation (45 min, 21,114xg, 4°C) the supernatant was applied to immobilized metal ion affinity chromatography via Ni-TED column (Protino Ni-TED 2000; Macherey-Nagel) pre-equilibrated with with the same buffer; proteins were eluted with buffer A containing 250 mM imidazole. Elution fractions containing the respective proteins were pooled and used for activity assays. The protein content was determined using the Bradford assay using bovine serum albumin as a standard. The KD(P)G aldolase Sso3197 was expressed as described previously (Ahmed *et al.*, 2005) and purified by heat precipitation.

The recombinant proteins were characterized with respect to their role in L-fucose degradation using the assay systems described above.

### ***Metabolic modelling of *S. solfataricus* P2***

#### ***Curation of the genome scale metabolic model of *S. solfataricus****

We used primary literature as well as the databases BRENDA (Chang *et al.*, 2015), MetaCyc (Caspi *et al.*, 2014), KEGG (Kanehisa *et al.*, 2014), BKM-react (Lang *et al.*, 2011) and TransportDB (Ren *et al.*, 2007) as source of information for the manual curation of our metabolic model (Ulas *et al.*, 2012). Semi-automated model curation was carried out using in-house software making use of the integrated genome annotation from the EnzymeDetector database (Quester and Schomburg, 2011). We added new biomass components like calditol (Gambacorta *et al.*, 2002; Yamauchi *et al.*, 2006) and the *Sulfolobus*-specific caldarchaeol

lipids (De Rosa *et al.*, 1980) to the model which will give a better approximation of the carbon and energy demand required for the organism's thermostable cell envelope. Ultimately, the growth associated maintenance energy (GAM) for growth on D-glucose had to be recalculated as previously described (Ulas *et al.*, 2012), resulting in a 10 % increase with a value of 27.31 mmol ATP·g<sub>CDW</sub><sup>-1</sup>. Likewise, the GAM was also calculated for growth on L-fucose using the L-fucose-specific biomass reaction *Biomass-fucose*, resulting in a value of 27.87 mmol ATP·g<sub>CDW</sub><sup>-1</sup>. The non-growth associated maintenance energy (NGAM) for growth on glucose was refitted as well, since we were able to calculate the integration of carbon atoms from the substrate into the biomass using the experimental growth data and the carbon content of the new biomass reaction in the metabolic model. The resulting carbon integration into the biomass of 46.5 % and the new glucose uptake rate of 1.13 mmol·g<sup>-1</sup>·h<sup>-1</sup> were used to fit the NGAM.

### ***Constraint-based modelling***

We used the metano toolbox (<http://metano.tu-bs.de/>) (Riemer *et al.*, 2013) for the curation and computational analysis of the metabolic model. Flux balance analysis (FBA) (Varma and Palsson, 1994; Kauffman *et al.*, 2003; Price *et al.*, 2004) was used to predict flux distributions. Underlying FBA is the assumption of both steady-state conditions and the evolutionary adaptation of the organism or network towards a biological goal (e.g. growth). The computationally efficient implementation of FBA in metano also precludes fluxes in stoichiometrically balanced loops (fluxes that do not contribute to the objective function).

### ***Flux variability analysis (FVA)***

Most metabolic models include more reactions than metabolites, leaving the model under-determined for optimisation methods such as FBA. In order to evaluate the presence of alternate flux distributions, FVA (Mahadevan and Schilling, 2003) can be applied. This method calculates minimal and maximum fluxes for each reaction using the given constraints under the additional prerequisite of optimal or suboptimal (Reed, 2004) optimisation of the objective function.

### ***Split-ratio analysis***

Recently a set of methods was published that facilitates the interpretation of flux distributions in stoichiometric metabolic models in order to identify important metabolic branch points (Riemer *et al.*, 2013). The split-ratio analysis is implemented in metano and

differentiates between consuming and producing fluxes for a designated metabolite. Another tool called AMEBA (Advanced METabolic Branchpoint Analysis; <http://metano.tu-bs.de/ameba>) can be used for the visual representation of these results in form of bipartite graphs.

### ***Simulation parameters***

For the simulation of growth on D-glucose or L-fucose as sole carbon source the uptake of carbonaceous compounds apart from CO<sub>2</sub>, D-glucose or L-fucose was restricted to zero. The uptake rates of D-glucose and L-fucose were restricted to 1.13 mmol·g<sub>CDW</sub><sup>-1</sup>·h<sup>-1</sup> and 1.21 mmol·g<sub>CDW</sub><sup>-1</sup>·h<sup>-1</sup> respectively. The biomass reaction *Biomass* or *Biomass\_fucose* was used as objective function, respectively. For each tested L-fucose scenario, the NGAM was fitted to a carbon balance of 39.9 %, resulting in different NGAM values but identical predictions for growth rates.

### ***Data storage and availability***

The investigation and the complete experimental data sets will be available upon publication on the SEEK homepage at <https://seek.sysmo-db.org>.

## Acknowledgements

We thank Sabine Kaltenhäuser for technical assistance. This work was performed within the e:Bio initiative of the Federal Ministry of Education and Research (BMBF), Germany. The following authors acknowledge the BMBF for financial support: JW and HS (0316188B), TK and KF (0316188A), AA and SPA (0316188D). BM was supported by an ERC starting grant (ARCHAELLUM, 311523) and LH by the BMBF (0316188C) and the DFG (AL1206/4-1). LS was supported by the Mercator Research Center Ruhr (Pr-2013-0010). SVA received intramural funds from the Max Planck Society. MNS was funded by the Federal State of Lower Saxony, Niedersächsisches Vorab (VWZN2889). TKP and PCW acknowledge the EPSRC (EP/1031812/1) and the BBSRC (BB/M018172/1) for financial support.

## Supporting Information

Additional Supporting information can be found in the online version of this article.

## Author Contributions

Conceived and designed the experiments: JW, DS, BS, TK, CB, KSH, MNS, JK, PCW, SVA. Performed the experiments: JW, KF, AA, TKP, KBM, BM, LH, LS. Computational modelling: HS. Analysed the data: JW, HS, KF, AA, TKP, KBM. Wrote the manuscript: JW, HS, KF, AA, TKP with input from DS, BS, SPA, TK, CB, KSH, MNS, JK, PCW and SVA. All authors read and approved the final manuscript.

## Conflict of interest

The authors declare that they have no conflict of interest.

## References

- Abu Dawud, R., Schreiber, K., Schomburg, D., and Adjaye, J. (2012) Human Embryonic Stem Cells and Embryonal Carcinoma Cells Have Overlapping and Distinct Metabolic Signatures. *PLoS ONE* **7**: e39896.
- Ahmed, H., Ettema, T.J.G., Tjaden, B., Geerling, A.C.M., van der Oost, J., and Siebers, B. (2005) The semi-phosphorylative Entner–Doudoroff pathway in hyperthermophilic archaea: a re-evaluation. *Biochem J* **390**: 529.
- Alber, B., Olinger, M., Rieder, A., Kockelkorn, D., Jobst, B., Hugler, M., and Fuchs, G. (2006) Malonyl-Coenzyme A Reductase in the Modified 3-Hydroxypropionate Cycle for Autotrophic Carbon Fixation in Archaeal *Metallosphaera* and *Sulfolobus* spp. *J Bacteriol* **188**: 8551–8559.
- Albers, S.-V., and Driessen, A.J.M. (2008) Conditions for gene disruption by homologous recombination of exogenous DNA into the *Sulfolobus solfataricus* genome. *Archaea* **2**: 145–149.
- Anders, S., and Huber, W. (2010) Differential expression analysis for sequence count data. *Genome Biol* **11**: R106.
- Bachrach, U. (2005) Naturally Occurring Polyamines: Interaction with Macromolecules. *Curr Protein Pept Sci* **6**: 559–566.
- Bae, J., Kim, S.M., and Lee, S.B. (2015) Identification and characterization of 2-keto-3-deoxy-L-rhamnonate dehydrogenase belonging to the MDR superfamily from the thermoacidophilic bacterium *Sulfobacillus thermosulfidooxidans*: implications to L-rhamnose metabolism in archaea. *Extrem Life Extreme Cond* **19**: 469–478.
- Baldomà, L., and Aguilar, J. (1988) Metabolism of L-fucose and L-rhamnose in *Escherichia coli*: aerobic-anaerobic regulation of L-lactaldehyde dissimilation. *J Bacteriol* **170**: 416–421.
- Benjamini, Y., and Yekutieli, D. (2001) The control of the false discovery rate in multiple testing under dependency. *Ann Stat* **29**: 1165–1188.
- Benton, H.P., Wong, D.M., Trauger, S.A., and Siuzdak, G. (2008) XCMS2: processing tandem mass spectrometry data for metabolite identification and structural characterization. *Anal Chem* **80**: 6382–6389.
- Berg, I.A., Kockelkorn, D., Buckel, W., and Fuchs, G. (2007) A 3-Hydroxypropionate/4-Hydroxybutyrate Autotrophic Carbon Dioxide Assimilation Pathway in Archaea. *Science* **318**: 1782–1786.
- Berg, I.A., Ramos-Vera, W.H., Petri, A., Huber, H., and Fuchs, G. (2010) Study of the distribution of autotrophic CO<sub>2</sub> fixation cycles in Crenarchaeota. *Microbiology*. **156**: 256–269.
- Bewley, M.A., Pham, T.K., Marriott, H.M., Noirel, J., Chu, H.-P., Ow, S.Y., *et al.* (2011) Proteomic Evaluation and Validation of Cathepsin D Regulated Proteins in Macrophages Exposed to *Streptococcus pneumoniae*. *Mol Cell Proteomics* **10**: M111.008193–M111.008193.

Boronat, A., and Aguilar, J. (1981) Metabolism of L-fucose and L-rhamnose in *Escherichia coli*: differences in induction of propanediol oxidoreductase. *J Bacteriol* **147**: 181–185.

Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**: 248–254.

Bräsen, C., Esser, D., Rauch, B., and Siebers, B. (2014) Carbohydrate Metabolism in Archaea: Current Insights into Unusual Enzymes and Pathways and Their Regulation. *Microbiol Mol Biol Rev* **78**: 89–175.

Brock, T.D., Brock, K.M., Belly, R.T., and Weiss, R.L. (1972) *Sulfolobus*: a new genus of sulfur-oxidizing bacteria living at low pH and high temperature. *Arch Für Mikrobiol* **84**: 54–68.

Brouns, S.J.J., Turnbull, A.P., Willemsen, H.L.D.M., Akerboom, J., and Oost, J. van der (2007) Crystal Structure and Biochemical Properties of the D-Arabinose Dehydrogenase from *Sulfolobus solfataricus*. *J Mol Biol* **371**: 1249–1260.

Brouns, S.J.J., Walther, J., Snijders, A.P.L., Werken, H.J.G. van de, Willemsen, H.L.D.M., Worm, P., *et al.* (2006) Identification of the Missing Links in Prokaryotic Pentose Oxidation Pathways: EVIDENCE FOR ENZYME RECRUITMENT. *J Biol Chem* **281**: 27378–27388.

Buchanan, C.L., Connaris, H., Danson, M.J., Reeve, C.D., and Hough, D.W. (1999) An extremely thermostable aldolase from *Sulfolobus solfataricus* with specificity for non-phosphorylated substrates. *Biochem J* **343 Pt 3**: 563–570.

Caspi, R., Altman, T., Billington, R., Dreher, K., Foerster, H., Fulcher, C.A., *et al.* (2014) The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. *Nucleic Acids Res* **42**: D459–D471.

Chang, A., Schomburg, I., Placzek, S., Jeske, L., Ulbrich, M., Xiao, M., *et al.* (2015) BRENDA in 2015: exciting developments in its 25th year of existence. *Nucleic Acids Res* **43**: D439–D446.

Cobucci-Ponzano, B., Conte, F., Benelli, D., Londei, P., Flagiello, A., Monti, M., *et al.* (2006) The gene of an archaeal alpha-L-fucosidase is expressed by translational frameshifting. *Nucleic Acids Res* **34**: 4258–4268.

Cobucci-Ponzano, B., Conte, F., Strazzulli, A., Capasso, C., Fiume, I., Pocsfalvi, G., *et al.* (2010) The molecular characterization of a novel GH38  $\alpha$ -mannosidase from the crenarchaeon *Sulfolobus solfataricus* revealed its ability in de-mannosylating glycoproteins. *Biochimie* **92**: 1895–1907.

Cobucci-Ponzano, B., Trincone, A., Giordano, A., Rossi, M., and Moracci, M. (2003) Identification of an archaeal alpha-L-fucosidase encoded by an interrupted gene. Production of a functional enzyme by mutations mimicking programmed -1 frameshifting. *J Biol Chem* **278**: 14622–14631.

Dang, S., Sun, L., Huang, Y., Lu, F., Liu, Y., Gong, H., *et al.* (2010) Structure of a fucose transporter in an outward-open conformation. *Nature* **467**: 734–738.

Dauner, M., and Sauer, U. (2001) Stoichiometric growth model for riboflavin-producing *Bacillus subtilis*. *Biotechnol Bioeng* **76**: 132–143.

- Demmer, U., Warkentin, E., Srivastava, A., Kockelkorn, D., Potter, M., Marx, A., *et al.* (2013) Structural Basis for a Bispecific NADP<sup>+</sup> and CoA Binding Site in an Archaeal Malonyl-Coenzyme A Reductase. *J Biol Chem* **288**: 6363–6370.
- Den Dool, H. van, and Dec. Kratz, P. (1963) A generalization of the retention index system including linear temperature programmed gas–liquid partition chromatography. *J Chromatogr A* **11**: 463–471.
- De Rosa, M., Gambacorta, A., Nicolaus, B., and Bu'Lock, J.D. (1980) Complex lipids of *Caldariella acidophila*, a thermoacidophile archaeobacterium. *Phytochemistry* **19**: 821–825.
- De Rosa, M., Gambacorta, A., Nicolaus, B., Giardina, P., Poerio, E., and Buonocore, V. (1984) Glucose metabolism in the extreme thermoacidophilic archaeobacterium *Sulfolobus solfataricus*. *Biochem J* **224**: 407–414.
- Dudoit, S., Yang, Y.H., Callow, M.J., and Speed, T.P. (2002) Statistical methods for identifying differentially expressed genes in replicated cDNA microarray experiments. *Stat Sin* **12**: 111–140.
- Dym, O., Pratt, E.A., Ho, C., and Eisenberg, D. (2000) The crystal structure of D-lactate dehydrogenase, a peripheral membrane respiratory enzyme. *Proc Natl Acad Sci* **97**: 9413–9418.
- Elferink, M.G.L., Albers, S.-V., Konings, W.N., and Driessen, A.J.M. (2001) Sugar transport in *Sulfolobus solfataricus* is mediated by two families of binding protein-dependent ABC transporters. *Mol Microbiol* **39**: 1494–1503.
- Esser, D., Kouril, T., Talfournier, F., Polkowska, J., Schrader, T., Bräsen, C., and Siebers, B. (2013) Unraveling the function of paralogs of the aldehyde dehydrogenase super family from *Sulfolobus solfataricus*. *Extremophiles* **17**: 205–216.
- Estelmann, S., Hügler, M., Eisenreich, W., Werner, K., Berg, I.A., Ramos-Vera, W.H., *et al.* (2011) Labeling and Enzyme Studies of the Central Carbon Metabolism in *Metallosphaera sedula*. *J Bacteriol* **193**: 1191–1200.
- Feist, A.M., Scholten, J.C.M., Palsson, B.Ø., Brockman, F.J., and Ideker, T. (2006) Modeling methanogenesis with a genome-scale metabolic reconstruction of *Methanosarcina barkeri*. *Mol Syst Biol* **2** <http://msb.embopress.org/cgi/doi/10.1038/msb4100046>. Accessed May 12, 2015.
- Gambacorta, A., Caracciolo, G., Trabasso, D., Izzo, I., Spinella, A., and Sodano, G. (2002) Biosynthesis of calditol, the cyclopentanoid containing moiety of the membrane lipids of the archaeon *Sulfolobus solfataricus*. *Tetrahedron Lett* **43**: 451–453.
- Ghalambor, M.A., and Heath, E.C. (1962) The metabolism of L-fucose. II. The enzymatic cleavage of L-fuculose 1-phosphate. *J Biol Chem* **237**: 2427–2433.
- Giardina, P., Biasi, M.G. de, Rosa, M. de, Gambacorta, A., and Buonocore, V. (1986) Glucose dehydrogenase from the thermoacidophilic archaeobacterium *Sulfolobus solfataricus*. *Biochem J* **239**: 517–522.
- Green, M., and Cohen, S.S. (1956) Enzymatic conversion of L-fucose to L-fuculose. *J Biol Chem* **219**: 557–568.

- Grogan, D.W. (1989) Phenotypic characterization of the archaeobacterial genus *Sulfolobus*: comparison of five wild-type strains. *J Bacteriol* **171**: 6710–6719.
- Heath, E.C., and Ghalambor, M.A. (1962) The metabolism of L-fucose. I. The purification and properties of L-fucose kinase. *J Biol Chem* **237**: 2423–2426.
- Hilker, R., Stadermann, K.B., Doppmeier, D., Kalinowski, J., Stoye, J., Straube, J., *et al.* (2014) ReadXplorer-visualization and analysis of mapped sequences. *Bioinformatics* **30**: 2247–2254.
- Hiller, K., Hangebrauk, J., Jäger, C., Spura, J., Schreiber, K., and Schomburg, D. (2009) MetaboliteDetector: comprehensive analysis tool for targeted and nontargeted GC/MS based metabolome analysis. *Anal Chem* **81**: 3429–3439.
- Hucka, M., Finney, A., Sauro, H.M., Bolouri, H., Doyle, J.C., Kitano, H., *et al.* (2003) The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics* **19**: 524–531.
- Iglesias, A.A., and Preiss, J. (1992) Bacterial glycogen and plant starch biosynthesis. *Biochem Educ* **20**: 196–203.
- Izzo, V., Notomista, E., Picardi, A., Pennacchio, F., and Di Donato, A. (2005) The thermophilic archaeon *Sulfolobus solfataricus* is able to grow on phenol. *Res Microbiol* **156**: 677–689.
- Kalisiak, J., Trauger, S.A., Kalisiak, E., Morita, H., Fokin, V.V., Adams, M.W.W., *et al.* (2009) Identification of a New Endogenous Metabolite and the Characterization of Its Protein Interactions through an Immobilization Approach. *J Am Chem Soc* **131**: 378–386.
- Kanehisa, M., Goto, S., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2014) Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res* **42**: D199–D205.
- Kauffman, K.J., Prakash, P., and Edwards, J.S. (2003) Advances in flux balance analysis. *Curr Opin Biotechnol* **14**: 491–496.
- Kim, S. (2006) Catalytic Promiscuity in Dihydroxy-Acid Dehydratase from the Thermoacidophilic Archaeon *Sulfolobus solfataricus*. *J Biochem (Tokyo)* **139**: 591–596.
- Kim, S.M., Paek, K.H., and Lee, S.B. (2012) Characterization of NADP<sup>+</sup>-specific l-rhamnose dehydrogenase from the thermoacidophilic Archaeon *Thermoplasma acidophilum*. *Extremophiles* **16**: 447–454.
- Koga, Y., and Morii, H. (2007) Biosynthesis of Ether-Type Polar Lipids in Archaea and Evolutionary Considerations. *Microbiol Mol Biol Rev* **71**: 97–120.
- Kohn, L.D., and Kaback, H.R. (1973) Mechanisms of active transport in isolated bacterial membrane vesicles. XV. Purification and properties of the membrane-bound D-lactate dehydrogenase from *Escherichia coli*. *J Biol Chem* **248**: 7012–7017.
- Kolaj-Robin, O., O’Kane, S.R., Nitschke, W., Léger, C., Baymann, F., and Soulimane, T. (2011) Biochemical and biophysical characterization of succinate: Quinone reductase from *Thermus thermophilus*. *Biochim Biophys Acta BBA - Bioenerg* **1807**: 68–79.

- König, H., Skorko, R., Zillig, W., and Reiter, W.-D. (1982) Glycogen in thermoacidophilic archaeobacteria of the genera *Sulfolobus*, *Thermoproteus*, *Desulfurococcus* and *Thermococcus*. *Arch Microbiol* **132**: 297–303.
- Kopka, J., Schauer, N., Krueger, S., Birkemeyer, C., Usadel, B., Bergmüller, E., *et al.* (2005) GMD@CSB.DB: the Golm Metabolome Database. *Bioinformatics* **21**: 1635–1638.
- Kouril, T., Esser, D., Kort, J., Westerhoff, H.V., Siebers, B., and Snoep, J.L. (2013) Intermediate instability at high temperature leads to low pathway efficiency for an *in vitro* reconstituted system of gluconeogenesis in *Sulfolobus solfataricus*. *FEBS J* **280**: 4666–4680.
- Kouril, T., Zaparty, M., Marrero, J., Brinkmann, H., and Siebers, B. (2008) A novel trehalose synthesizing pathway in the hyperthermophilic Crenarchaeon *Thermoproteus tenax*: the unidirectional TreT pathway. *Arch Microbiol* **190**: 355–369.
- Kruskal, W.H., and Wallis, W.A. (1952) Use of Ranks in One-Criterion Variance Analysis. *J Am Stat Assoc* **47**: 583–621.
- Lamble, H.J., Heyer, N.I., Bull, S.D., Hough, D.W., and Danson, M.J. (2003) Metabolic Pathway Promiscuity in the Archaeon *Sulfolobus solfataricus* Revealed by Studies on Glucose Dehydrogenase and 2-Keto-3-deoxygluconate Aldolase. *J Biol Chem* **278**: 34066–34072.
- Langmead, B., and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**: 357–359.
- Lang, M., Stelzer, M., and Schomburg, D. (2011) BKM-react, an integrated biochemical reaction database. *BMC Biochem* **12**: 42.
- Lehman, T.C., Hale, D.E., Bhala, A., and Thorpe, C. (1990) An acyl-coenzyme A dehydrogenase assay utilizing the ferricenium ion. *Anal Biochem* **186**: 280–284.
- Lord, J.M. (1972) Glycolate oxidoreductase in *Escherichia coli*. *Biochim Biophys Acta BBA - Bioenerg* **267**: 227–237.
- Lubelska, J.M., Jonuscheit, M., Schleper, C., Albers, S.-V., and Driessen, A.J.M. (2006) Regulation of expression of the arabinose and glucose transporter genes in the thermophilic archaeon *Sulfolobus solfataricus*. *Extremophiles* **10**: 383–391.
- Mahadevan, R., and Schilling, C.H. (2003) The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. *Metab Eng* **5**: 264–276.
- Maruta, K., Mitsuzumi, H., Nakada, T., Kubota, M., Chaen, H., Fukuda, S., *et al.* (1996) Cloning and sequencing of a cluster of genes encoding novel enzymes of trehalose biosynthesis from thermophilic archaeobacterium *Sulfolobus acidocaldarius*. *Biochim Biophys Acta BBA - Gen Subj* **1291**: 177–181.
- Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., and Wold, B. (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* **5**: 621–628.
- Neidhardt, F.C., Ingraham, J.L., and Schaechter, M. (1990) *Physiology of the bacterial cell: a molecular approach*. Sinauer Associates, Sunderland, Mass.

Nicolaus, B., Trincone, A., Lama, L., Romano, I., Marsiglia, F., and Gambacorta, A. (1991) Adaptation of *Sulfolobus solfataricus* on minimal media. *Biotechnol Lett* **13**: 667–670.

Nunn, C.E.M., Johnsen, U., Schonheit, P., Fuhrer, T., Sauer, U., Hough, D.W., and Danson, M.J. (2010) Metabolism of Pentose Sugars in the Hyperthermophilic Archaea *Sulfolobus solfataricus* and *Sulfolobus acidocaldarius*. *J Biol Chem* **285**: 33701–33709.

Obradors, N., Badía, J., Baldomà, L., and Aguilar, J. (1988) Anaerobic metabolism of the L-rhamnose fermentation product 1,2-propanediol in *Salmonella typhimurium*. *J Bacteriol* **170**: 2159–2162.

Oh, Y.-K., Palsson, B.O., Park, S.M., Schilling, C.H., and Mahadevan, R. (2007) Genome-scale Reconstruction of Metabolic Network in *Bacillus subtilis* Based on High-throughput Phenotyping and Gene Essentiality Data. *J Biol Chem* **282**: 28791–28799.

Orita, I., Sato, T., Yurimoto, H., Kato, N., Atomi, H., Imanaka, T., and Sakai, Y. (2006) The Ribulose Monophosphate Pathway Substitutes for the Missing Pentose Phosphate Pathway in the Archaeon *Thermococcus kodakaraensis*. *J Bacteriol* **188**: 4698–4704.

Ow, S.Y., Noirel, J., Salim, M., Evans, C., Watson, R., and Wright, P.C. (2010) Balancing robust quantification and identification for iTRAQ: Application of UHR-ToF MS. *PROTEOMICS* **10**: 2205–2213.

Ow, S.Y., Salim, M., Noirel, J., Evans, C., Rehman, I., and Wright, P.C. (2009) iTRAQ Underestimation in Simple and Complex Mixtures: “The Good, the Bad and the Ugly.” *Proteome Res* **8**: 5347–5355.

Petit, E., LaTouf, W.G., Coppi, M.V., Warnick, T.A., Currie, D., Romashko, I., *et al.* (2013) Involvement of a bacterial microcompartment in the metabolism of fucose and rhamnose by *Clostridium phytofermentans*. *PloS One* **8**: e54337.

Peyraud, R., Kiefer, P., Christen, P., Massou, S., Portais, J.-C., and Vorholt, J.A. (2009) Demonstration of the ethylmalonyl-CoA pathway by using <sup>13</sup>C metabolomics. *Proc Natl Acad Sci U S A* **106**: 4846–4851.

Pham, T.K., Roy, S., Noirel, J., Douglas, I., Wright, P.C., and Stafford, G.P. (2010) A quantitative proteomic analysis of biofilm adaptation by the periodontal pathogen *Tannerella forsythia*. *PROTEOMICS* **10**: 3130–3141.

Price, N.D., Reed, J.L., and Palsson, B.Ø. (2004) Genome-scale models of microbial cells: evaluating the consequences of constraints. *Nat Rev Microbiol* **2**: 886–897.

Quester, S., and Schomburg, D. (2011) EnzymeDetector: an integrated enzyme function prediction tool and database. *BMC Bioinformatics* **12**: 376.

Qu, Q., Lee, S.-J., and Boos, W. (2004) TreT, a novel trehalose glycosyltransferring synthase of the hyperthermophilic archaeon *Thermococcus litoralis*. *J Biol Chem* **279**: 47890–47897.

Reed, J.L. (2004) Genome-Scale In Silico Models of *E. coli* Have Multiple Equivalent Phenotypic States: Assessment of Correlated Reaction Subsets That Comprise Network States. *Genome Res* **14**: 1797–1805.

Ren, Q., Chen, K., and Paulsen, I.T. (2007) TransportDB: a comprehensive database resource for cytoplasmic membrane transport systems and outer membrane channels. *Nucleic Acids Res* **35**: D274–D279.

Riemer, S., Rex, R., and Schomburg, D. (2013) A metabolite-centric view on flux distributions in genome-scale metabolic models. *BMC Syst Biol* **7**: 33.

Satomura, T., Kawakami, R., Sakuraba, H., and Ohshima, T. (2008) A novel flavin adenine dinucleotide (FAD) containing d-lactate dehydrogenase from the thermoacidophilic crenarchaeota *Sulfolobus tokodaii* strain 7: Purification, characterization and expression in *Escherichia coli*. *J Biosci Bioeng* **106**: 16–21.

Sato, T., Imanaka, H., Rashid, N., Fukui, T., Atomi, H., and Imanaka, T. (2004) Genetic Evidence Identifying the True Gluconeogenic Fructose-1,6-Bisphosphatase in *Thermococcus kodakaraensis* and Other Hyperthermophiles. *J Bacteriol* **186**: 5799–5807.

Say, R.F., and Fuchs, G. (2010) Fructose 1,6-bisphosphate aldolase/phosphatase may be an ancestral gluconeogenic enzyme. *Nature* **464**: 1077–1081.

Schelert, J., Dixit, V., Hoang, V., Simbahan, J., Drozda, M., and Blum, P. (2004) Occurrence and Characterization of Mercury Resistance in the Hyperthermophilic Archaeon *Sulfolobus solfataricus* by Use of Gene Disruption. *J Bacteriol* **186**: 427–437.

She, Q., Singh, R.K., Confalonieri, F., Zivanovic, Y., Allard, G., Awayez, M.J., *et al.* (2001) The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. *Proc Natl Acad Sci* **98**: 7835–7840.

Smith, C.A., Want, E.J., O'Maille, G., Abagyan, R., and Siuzdak, G. (2006) XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Anal Chem* **78**: 779–787.

Soderberg, T. (2005) Biosynthesis of ribose-5-phosphate and erythrose-4-phosphate in archaea: a phylogenetic analysis of archaeal genomes. *Archaea* **1**: 347–352.

Söding, J., Biegert, A., and Lupas, A.N. (2005) The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* **33**: W244–W248.

Strom, T., Ferenci, T., and Quayle, J.R. (1974) The carbon assimilation pathways of *Methylococcus capsulatus*, *Pseudomonas methanica* and *Methylosinus trichosporium* (OB3B) during growth on methane. *Biochem J* **144**: 465–476.

Tautenhahn, R., Bottcher, C., and Neumann, S. (2008) Highly sensitive feature detection for high resolution LC/MS. *BMC Bioinformatics* **9**: 504.

Uhrigshardt, H., Walden, M., John, H., Petersen, A., and Anemüller, S. (2002) Evidence for an operative glyoxylate cycle in the thermoacidophilic crenarchaeon *Sulfolobus acidocaldarius*. *FEBS Lett* **513**: 223–229.

Ulas, T., Riemer, S.A., Zaparty, M., Siebers, B., and Schomburg, D. (2012) Genome-Scale Reconstruction and Analysis of the Metabolic Network in the Hyperthermophilic Archaeon *Sulfolobus solfataricus*. *PLoS ONE* **7**: e43401.

Van de Casteele, M., Demarez, M., Legrain, C., Glansdorff, N., and Pierard, A. (1990) Pathways of arginine biosynthesis in extreme thermophilic archaeo- and eubacteria. *J Gen Microbiol* **136**: 1177–1183.

Vanhooren, P.T., and Vandamme, E.J. (1999) L-Fucose: occurrence, physiological role, chemical, enzymatic and microbial synthesis. *J Chem Technol Biotechnol* **74**: 479–497.

Varma, A., and Palsson, B.O. (1994) Metabolic Flux Balancing: Basic Concepts, Scientific and Practical Use. *Bio/Technology* **12**: 994–998.

Wagner, M., Wolferen, M. van, Wagner, A., Lassak, K., Meyer, B.H., Reimann, J., and Albers, S.-V. (2012) Versatile Genetic Tool Box for the Crenarchaeote *Sulfolobus acidocaldarius*. *Front Microbiol* **3**: 214.

Wakagi, T., Fukuda, E., Ogawa, Y., Kino, H., and Matsuzawa, H. (2002) A novel bifunctional molybdo-enzyme catalyzing both decarboxylation of indolepyruvate and oxidation of indoleacetaldehyde from a thermoacidophilic archaeon, *Sulfolobus* sp. strain 7. *FEBS Lett* **510**: 196–200.

Wakagi, T., Nishimasu, H., Miyake, M., and Fushinobu, S. (2016) Archaeal Mo-Containing Glyceraldehyde Oxidoreductase Isozymes Exhibit Diverse Substrate Specificities through Unique Subunit Assemblies. *PLOS ONE* **11**: e0147333.

Worthington, P., Hoang, V., Perez-Pomares, F., and Blum, P. (2003) Targeted Disruption of the  $\alpha$ -Amylase Gene in the Hyperthermophilic Archaeon *Sulfolobus solfataricus*. *J Bacteriol* **185**: 482–488.

Yamauchi, N., Kamada, N., and Ueoka, H. (2006) The Possibility of Involvement of “Cyclase” Enzyme of the Calditol Carbocycle with Broad Substrate Specificity in *Sulfolobus acidocaldarius*, a Typical Thermophilic Archaea. *Chem Lett* **35**: 1230–1231.

Yew, W.S., Fedorov, A.A., Fedorov, E.V., Rakus, J.F., Pierce, R.W., Almo, S.C., and Gerlt, J.A. (2006) Evolution of enzymatic activities in the enolase superfamily: L-fuconate dehydratase from *Xanthomonas campestris*. *Biochemistry (Mosc)* **45**: 14582–14597.

Zaparty, M., Esser, D., Gertig, S., Haferkamp, P., Kouril, T., Manica, A., *et al.* (2010) “Hot standards” for the thermoacidophilic archaeon *Sulfolobus solfataricus*. *Extrem Life Extreme Cond* **14**: 119–142.

Zech, H., Hensler, M., Koßmehl, S., Drüppel, K., Wöhlbrand, L., Trautwein, K., *et al.* (2013) Adaptation of *Phaeobacter inhibens* DSM 17395 to growth with complex nutrients. *PROTEOMICS* 2851–2868.

Zillig, W., Stetter, K.O., Wunderl, S., Schulz, W., Priess, H., and Scholz, I. (1980) The *Sulfolobus*-“*Caldariella*” group: Taxonomy on the basis of the structure of DNA-dependent RNA polymerases. *Arch Microbiol* **125**: 259–269.

**Table I:** Biomass composition of *S. solfataricus* P2 after growth on D-glucose and L-fucose. Protein, RNA and DNA content were investigated in this study. The values represent the average of at least three independent experiments Errors represent the standard deviation between the experiments.

Carbon source	Component	Cellular content [% (w/w)]	Reference
<b>D-glucose</b>	Proteins	64.3 ± 3.4	This study
	RNA	5.4 ± 0.5	This study
	DNA	1.1 ± 0.1	This study
	Lipids	7.5	(De Rosa <i>et al.</i> , 1980)
	Soluble pool	3.2	(Feist <i>et al.</i> , 2006)
	Glycogen	18.5	(König <i>et al.</i> , 1982; Iglesias and Preiss, 1992)
<b>L-fucose</b>	Proteins	67.9 ± 3.1	This study
	RNA	5.6 ± 0.2	This study
	DNA	1.2 ± 0.0	This study
	Lipids	7.5	(De Rosa <i>et al.</i> , 1980)
	Soluble pool	3.2	(Feist <i>et al.</i> , 2006)
	Glycogen	14.6	(König <i>et al.</i> , 1982; Iglesias and Preiss, 1992)

**Table II:** Maximal allowed fluxes [ $\text{mmol}\cdot\text{g}^{-1}\cdot\text{h}^{-1}$ ] through selected metabolic pathways in *S. solfataricus* for simulated growth on different carbon sources. The glucose scenario was compared to Fucose degradation via the *E. coli* like pathway and Fucose degradation via the two KDF dependent routes. Ratios were calculated as compared to the glucose reference scenario. For simulation of possible fluxes through the 3-hydroxypropionate/4-hydroxybutanoate cycle different scenarios were analyzed, enforcing the flux through different parts of the cycle. Part 1: Reaction 1-6: Acetyl-CoA to Propanoyl-CoA; Part 2: Reactions 7-9: Propanoyl-CoA to Succinyl-CoA; Part 3: Reactions 10-16: Succinyl-CoA to 2 Acetyl-CoA. EC 4.1.1.31 represents the flux through the anaplerotic reaction of the TCA cycle (phosphoenolpyruvate +  $\text{CO}_2 \rightarrow$  oxaloacetate +  $\text{P}_i$ ), whereas EC 4.1.1.32 represents the flux through the anabolic reaction towards gluconeogenesis (oxaloacetate + GTP  $\rightarrow$  phosphoenolpyruvate + GDP +  $\text{CO}_2$ ). For suboptimal flux variance analysis the maximal biomass flux was set to 95 %.

Abbreviations: FBA - flux balance analysis; FVA - flux variance analysis; KDF - 2-keto-3-deoxy-L-fuconate; GAM: growth associated maintenance energy; NGAM - non growth associated maintenance energy; HP/HB cycle - 3-hydroxypropionate/4-hydroxybutanoate cycle.

Suboptimal FVA maximal fluxes [ $\text{mmol}\cdot\text{g}^{-1}\cdot\text{h}^{-1}$ ]	Glucose	Fucose via <i>E. coli</i> like pathway	Fucose via KDF pathways	Ratio <i>E. coli</i>	Ratio KDF
Glyoxylate bypass	0.771	0.719	1.181	0.9	1.5
EC 4.1.1.31	0.771	0.719	0.570	0.9	0.7
EC 4.1.1.32	0.570	0.532	0.99	0.9	1.7
HP/HB cycle part 1	0.285	0.266	0.285	0.9	1.0
HP/HB cycle part 2	0.771	0.719	1.181	0.9	1.5
HP/HB cycle part 3	0.175	0.164	0.177	0.9	1.0
HP/HB cycle part 1 + 2	0.190	0.177	0.190	0.9	1.0
HP/HB cycle complete	0.114	0.106	0.114	0.9	1.0
<b>Energy parameters</b>					
GAM [ $\text{mmol ATP}\cdot\text{g}^{-1}$ ]	27.309	27.865	27.865	1.02	1.02
NGAM [ $\text{mmol ATP}\cdot\text{g}^{-1}\cdot\text{h}^{-1}$ ]	3.281	6.298	5.541	1.92	1.69
ATP hydrolysis due to energy maintenance [ $\text{mmol ATP}\cdot\text{g}^{-1}\cdot\text{h}^{-1}$ ]	5.154	8.049	7.293	1.56	1.42

**Table III:** Specific growth rate  $\mu$  [ $\text{h}^{-1}$ ], maximal substrate uptake rate  $qS_{\text{max}}$  [ $\text{mmol}\cdot\text{g}^{-1}\cdot\text{h}^{-1}$ ], maximum cell dry weight  $\text{CDW}_{\text{max}}$  [ $\text{g}\cdot\text{l}^{-1}$ ] and molar growth yields per substrate carbon  $Y_{\text{SC}}$  [ $\text{g}_{\text{CDW}}/\text{mol}$  of substrate carbon] of *S. solfataricus* P2 grown on L-fucose, D-glucose and on a mixture of both sugars (each 15 mM).

substrate	$\mu$ [ $\text{h}^{-1}$ ] <sup>a</sup>	$qS_{\text{max}}$ [ $\text{mmol}\cdot\text{g}^{-1}\cdot\text{h}^{-1}$ ] <sup>a</sup>	$Y_{\text{SC}}$ [ $\text{g}_{\text{CDW}}/\text{mol}$ of substrate carbon] <sup>a</sup>	
			substrate	$\text{CDW}_{\text{max}}$ [ $\text{g}\cdot\text{l}^{-1}$ ] <sup>a</sup>
L-fucose	$0.040 \pm 0.002$	$1.21 \pm 0.01$	$90.21 \pm 7.14$	$0.52 \pm 0.01$
D-glucose	$0.045 \pm 0.001$	$1.13 \pm 0.11$	$100.6 \pm 12.3$	$0.70 \pm 0.01$
L-fucose/D-glucose	$0.077 \pm 0.001$	ND <sup>b</sup>	ND <sup>b</sup>	$0.67 \pm 0.02$

<sup>a</sup> values represent the average of three independent experiments. Errors represent the standard deviation.

<sup>b</sup>ND - not determined.

**Table IV:** Fold changes of selected intracellular metabolite levels of *S. solfataricus* P2 grown on L-fucose as compared to cells grown on D-glucose.

Metabolites	Fold change <sup>a</sup>
	Fuc
<b>ED and gluconeogenesis</b>	
glucose <sup>b</sup>	0.01 ± 0.00
glucose-6-phosphate <sup>b</sup>	0.20 ± 0.01
pyruvate <sup>b</sup>	0.55 ± 0.05
<b>CoA-intermediates</b>	
acetyl-CoA <sup>b,c</sup>	1.27 ± 0.05
crotonyl-CoA <sup>c</sup>	1.61 ± 0.25
3-hydroxybutanoyl-CoA <sup>c</sup>	1.27 ± 0.10
methylmalonyl-CoA <sup>c</sup>	1.35 ± 0.14
propanoyl-CoA <sup>b,c</sup>	1.76 ± 0.09
succinyl-CoA <sup>c</sup>	0.63 ± 0.05

<sup>a</sup>Values represent the average of three independent experiments.

<sup>b</sup>statistically significant change (p-value < 0.01)

<sup>c</sup>involved in the 3-hydroxypropionate-4-hydroxybutyrate-cycle (Berg *et al.*, 2007; Berg *et al.*, 2010).

**Table V:** Differentially expressed genes and differentially produced proteins in *S. solfataricus* P2 after growth on L-fucose identified from RNASeq (Transcriptome) and iTRAQ (Proteome) analysis. The fold change represents the relative abundance of every gene or protein averaged from three independent experiments. Functions of proteins were assigned using either the EnzymeDetector (Quester and Schomburg, 2011) or arCOG based annotations.

Locus tag	Function	Fold change <sup>a</sup> (p-value <sup>b</sup> )	
		Transcriptome	Proteome
<i>Sso3124</i>	D-Arabinonate dehydratase	49.9 (4.6 · 10 <sup>-68</sup> )	16.8 (1.9 · 10 <sup>-12</sup> )
<i>Sso3117</i>	2,5-dioxovalerate dehydrogenase	38.4 (8.2 · 10 <sup>-54</sup> )	16.4 (1.1 · 10 <sup>-61</sup> )
<i>Sso1300</i>	D-Arabinose-1-dehydrogenase	28.0 (1.3 · 10 <sup>-151</sup> )	14.5 (9.4 · 10 <sup>-16</sup> )
<i>Sso3066</i>	Arabinose ABC transporter, arabinose-binding protein	25.0 (5.0 · 10 <sup>-103</sup> )	11.7 (3.3 · 10 <sup>-08</sup> )
<i>Sso3118</i>	2-dehydro-3-deoxy-D-arabinonate dehydratase	17.4 (1.1 · 10 <sup>-07</sup> )	10.5 (7.0 · 10 <sup>-38</sup> )
<i>Sso2998</i>	D-xylonate dehydratase*	3.7 (2.1 · 10 <sup>-08</sup> )	3.8 (2.6 · 10 <sup>-13</sup> )
<i>Sso1333</i>	Isocitrate lyase	3.3 (5.1 · 10 <sup>-30</sup> )	3.5 (1.2 · 10 <sup>-34</sup> )
<i>Sso3107</i>	Dihydroxy-acid dehydratase	3.1 (2.1 · 10 <sup>-25</sup> )	3.0 (9.5 · 10 <sup>-36</sup> )
<i>Sso3067</i>	Arabinose ABC transporter, permease	14.2 (1.0 · 10 <sup>-18</sup> )	ND <sup>c</sup>
<i>Sso1004</i>	FAD/FMN-containing dehydrogenase	14.0 (5.7 · 10 <sup>-17</sup> )	ND
<i>Sso1005</i>	Unknown function	12.6 (7.8 · 10 <sup>-13</sup> )	ND
<i>Sso1007</i>	Succinyl-diaminopimelate desuccinylase / Acetylornithine deacetylase *	11.8 (9.3 · 10 <sup>-19</sup> )	ND
<i>Sso3120</i>	MFS family permease*	10.4 (4.1 · 10 <sup>-61</sup> )	ND
<i>Sso3068</i>	Arabinose ABC transporter, permease	8.6 (1.1 · 10 <sup>-59</sup> )	ND
<i>Sso1305</i>	MFS family permease *	8.3 (5.6 · 10 <sup>-21</sup> )	ND
<i>Sso0786</i>	Amino acid specific permease*	7.5 (4.2 · 10 <sup>-06</sup> )	ND
<i>Sso1303</i>	L-alanine-DL-glutamate epimerase or related enzyme of enolase superfamily*	7.3 (2.3 · 10 <sup>-33</sup> )	ND
<i>Sso1009</i>	Amino acid transporter *	5.8 (1.4 · 10 <sup>-09</sup> )	ND
<i>Sso3123</i>	Unknown function	5.4 (9.8 · 10 <sup>-44</sup> )	ND
<i>Sso2128</i>	2-oxoacid:ferredoxin oxidoreductase, gamma subunit*	4.5 (2.7 · 10 <sup>-10</sup> )	ND
<i>Sso2129</i>	2-oxoacid:ferredoxin oxidoreductase, alpha subunit*	4.4 (1.6 · 10 <sup>-10</sup> )	ND
<i>Sso2178</i>	Malonyl-CoA reductase*	4.0 (3.2 · 10 <sup>-31</sup> )	ND
<i>Sso3003</i>	Glucose-1-dehydrogenase	0.19 (5.5 · 10 <sup>-65</sup> )	0.28 (5.6 · 10 <sup>-12</sup> )
<i>Sso3008</i>	Putative sugar dehydrogenase*	0.16 (1.1 · 10 <sup>-39</sup> )	0.33 (2.6 · 10 <sup>-08</sup> )
<i>Sso3009</i>	Glyceraldehyde oxidoreductase, large subunit*	0.23 (9.4 · 10 <sup>-52</sup> )	0.38 (6.2 · 10 <sup>-13</sup> )
<i>Sso3006</i>	Alpha-mannosidase	0.21 (2.1 · 10 <sup>-54</sup> )	0.38 (7.6 · 10 <sup>-09</sup> )
<i>Sso3011</i>	Putative sugar dehydrogenase*	0.20 (5.8 · 10 <sup>-45</sup> )	0.44 (2.4 · 10 <sup>-03</sup> )
<i>Sso1725</i>	Putative CRISPR system CMR subunit Cmr7 2	0.09 (2.1 · 10 <sup>-21</sup> )	ND
<i>Sso3050</i>	Sugar phosphate isomerase/epimerase *	0.11 (4.3 · 10 <sup>-38</sup> )	ND
<i>Sso2967</i>	Unknown function	0.13 (1.4 · 10 <sup>-58</sup> )	ND
<i>Sso3007</i>	Mannan endo-1,4-beta- mannosidase*	0.15 (8.4 · 10 <sup>-54</sup> )	ND

Table V: continued.

Locus tag	Function	Fold change <sup>a</sup> (p-value <sup>b</sup> )	
		Transcriptome	Proteome
<i>Sso1782</i>	Predicted sugar phosphate nucleotidyl transferase*	0.16 ( $5.7 \cdot 10^{-6}$ )	ND
<i>Sso1783</i>	Predicted nucleotidyl-sugar reductase*	0.18 ( $3.1 \cdot 10^{-10}$ )	ND
<i>Sso1536</i>	Transcriptional regulator	0.20 ( $2.1 \cdot 10^{-06}$ )	ND
<i>Sso1656</i>	SWIM zinc finger containing protein	0.20 ( $2.2 \cdot 10^{-10}$ )	ND
<i>Sso3197</i>	Bifunctional 2-keto-3-deoxy-(6-phospho)gluconate/galactonate aldolase	1.1 ( $9.0 \cdot 10^{-01}$ )	0.86 ( $3.6 \cdot 10^{-01}$ )
<i>Sso3161</i>	Fe-S oxidoreductase subunit	1.0 ( $7.1 \cdot 10^{-01}$ )	ND
<i>Sso3163</i>	FAD/FMN-containing dehydrogenase	1.0 ( $6.9 \cdot 10^{-01}$ )	ND
<i>Sso3165</i>	FAD/FMN-containing dehydrogenase	1.1 ( $8.7 \cdot 10^{-01}$ )	ND

<sup>a</sup>Fold change in relative abundance as compared to the D-glucose reference condition.

<sup>b</sup>p-value < 0.01 (transcriptomics) or <  $2.36 \cdot 10^{-5}$  (proteomics) indicates statistically significant differential expression.

<sup>c</sup>ND - not determined.

\*Function predicted.

**Table VI:** Kinetic characterization of recombinant L-fucose degrading enzymes of *S. solfataricus*.

Current annotation	Substrate	$V_{\max}$ (U/mg)	$K_m$ (mM)	$k_{\text{cat}}$ ( $s^{-1}$ )	$k_{\text{cat}}/K_m$ ( $mM^{-1}s^{-1}$ )
<b>D-arabinose dehydrogenase (Sso1300)</b>	L-Fucose (NADP <sup>+</sup> )	26.7 ± 0.89	0.12 ± 0.02	19.4	161.9
	D-Arabinose (NADP <sup>+</sup> )	21.5 ± 0.81	0.22 ± 0.03	12.8	58.2
<b>D-arabinonate dehydratase (Sso3124)</b>	L-Fuconate	0.018 ± 0.002	ND*	0.013	ND
	D-Arabinonate	0.078 ± 0.02	ND	0.055	ND
<b>2-Keto-3-deoxy-gluconate/galactonate aldolase (Sso3197)</b>	L-Lactaldehyde (Pyruvate)	1.9 ± 0.13	8.8 ± 1.15	0.9	0.10
	2-Keto-3-deoxy-D-gluconate	33.8 ± 3.6	10.7 ± 1.8	27.5	2.6
<b>2,5-dioxovalerate dehydrogenase (Sso3117)</b>	L-Lactaldehyde (NADP <sup>+</sup> )	0.17 ± 0.001	ND	0.15	ND

\*ND - not determined.

## Figure Legends

**Figure 1: Overview of three possible routes for L-fucose degradation.** The predicted pathway in *S. solfataricus* (A) is compared to the common ways for L-fucose degradation found in *X. campestris* (B) and *E. coli* (C).

Catalysing enzymes are indicated by their corresponding EC-number. The *S. solfataricus* P2 specific omics data from this work was mapped to the intermediates and the proposed catalysing enzymes. The mapping visualizes the corresponding logarithmic fold changes for every node comparing L-fucose conditions to D-glucose. Yellow nodes indicate higher abundance of protein or metabolite on L-fucose whereas blue nodes represent a decreased state. Note that the metabolites fucose, fuconolactone and fuconate were only found under L-fucose growth conditions, thus the fold-change is assumed to be at least  $> 4.0$ . Catalysing enzymes are divided into two parts (white line) whereas the upper part represents the data from the transcriptome analysis and the proteome data are mapped at the bottom part. Box-shaped, uncoloured nodes represent not detected metabolites and proteins. Dashed lines indicate that in *S. solfataricus* no homologous enzymes are present for this reaction. Note that in *E. coli* (C) a third NADH and one net ATP is produced from utilization of GAP via the classical EMP pathway.

Abbreviations: ABC-TR - ABC transporter; EMP - Embden-Meyerhof-Parnas pathway; GAP - glyceraldehyde-3-phosphate; KDF - 2-keto-3-deoxy-L-fuconate; DKDF - 2,4-diketo-3-deoxy-L-fuconate; L-fuculose-1-P - L-fuculose-1-phosphate; DHAP - dihydroxyacetonephosphate

**Figure 2: Growth and substrate uptake of *S. solfataricus* P2.**

(A) Growth (blue triangles) and substrate uptake (pink squares) on minimal medium containing L-fucose as sole carbon source.

(B) Growth (green triangles) and substrate uptake (red squares) on minimal medium containing D-glucose as sole carbon source.

(C) Growth behaviour (black squares) and time resolved relative abundance of L-fucose (blue bars) and D-glucose (red bars) in the culture supernatant of *S. solfataricus*. Cells were cultivated on minimal medium containing 15 mM of each, L-fucose and D-glucose and supernatants were analysed by GC-MS. For determination of the relative concentration of the two sugars in the media, integrated peak areas of the zero hour sample were set as 100 % and all other values were scaled accordingly.

All values represent the average of three independent cultivations. Error bars represent the standard deviation between the three experiments.

**Figure 3: Comparison of logarithmised normalized peak areas of metabolites found in *S. solfataricus* P2 extracts after cultivation on L-fucose versus D-glucose.** Identified metabolites showing significant changes are labelled (Benjamini-Yekutieli corrected Kruskal-Wallis test, p-value  $< 0.01$ ). Metabolite levels for trehalose and acetoacetyl-CoA are only an approximation of the actual

intracellular amount, as the trehalose signal is always strongly overloaded (especially under glucose conditions) and acetoacetyl-CoA is normally not stable during the extraction process.

Values represent the average of three independent experiments. Error bars represent the standard error between the three experiments.

Abbreviations: G6P – glucose-6-phosphate; MGP – methyl- $\alpha$ -glucopyranoside; 3-OH-2-Me-butyryl-CoA – 3-hydroxy-2-methyl-butyryl-CoA; 4-OH-PheAc – 4-hydroxyphenylacetate.

**Figure 4: Relative metabolite concentrations of compounds solely detected in *S. solfataricus* P2 extracts after cultivation on either L-fucose (A) or D-glucose (B).**

Abbreviations: 2,7-aSH - 2,7-anhydro-sedoheptulose; EA - erythronic acid; F6P - fructose-6-phosphate; fuconolactone - fucono-1,4-lactone; Gal - galactose; GlcA - gluconate; 1,6-aGlc - 1,6-anhydroglucose; Isomalt - isomaltulose.

Values represent the average of three independent experiments. Error bars represent the standard deviation between the three experiments.

**Figure 5: Aldolase-dependent route predicted for L-fucose degradation in *S. solfataricus* P2.**

Enzymes assayed in this work are shown in bold, corresponding reactions are numbered. Lactate dehydrogenase activity could be found in crude extracts although no specific enzyme for this reaction is identified in *S. solfataricus* until now.

**Figure 6: Specific activities of assayed reactions in cell free extracts of L-fucose and D-glucose grown cells of *S. solfataricus* P2.** Blue bars: L-fucose cell free extracts; red bars: D-glucose cell free extracts. All values represent the average of at least three independent measurements. Error bars represent the standard deviation between the measurements

(A) L-Fucose-dehydrogenase assay, using L-fucose or D-arabinose and  $\text{NADP}^+$  as substrates and following NADPH formation by increase in absorbance at 340 nm.

(B) L-fuconate-dehydratase assay, using L-fuconate or D-arabinonate as substrates and following the formation of 2-keto-3-deoxy-fuconate (KDF) formation by thiobarbituric acid (TBA) assay.

(C) KDF aldolase assay. The anabolic direction was analysed by measuring the condensation of pyruvate and L-lactaldehyde to KDF.

(D) Lactaldehyde dehydrogenase assay, using different stereoisomers of lactaldehyde (LA) or 2,5-dioxopentanoate (DOP) and  $\text{NADP}^+$  as substrates.

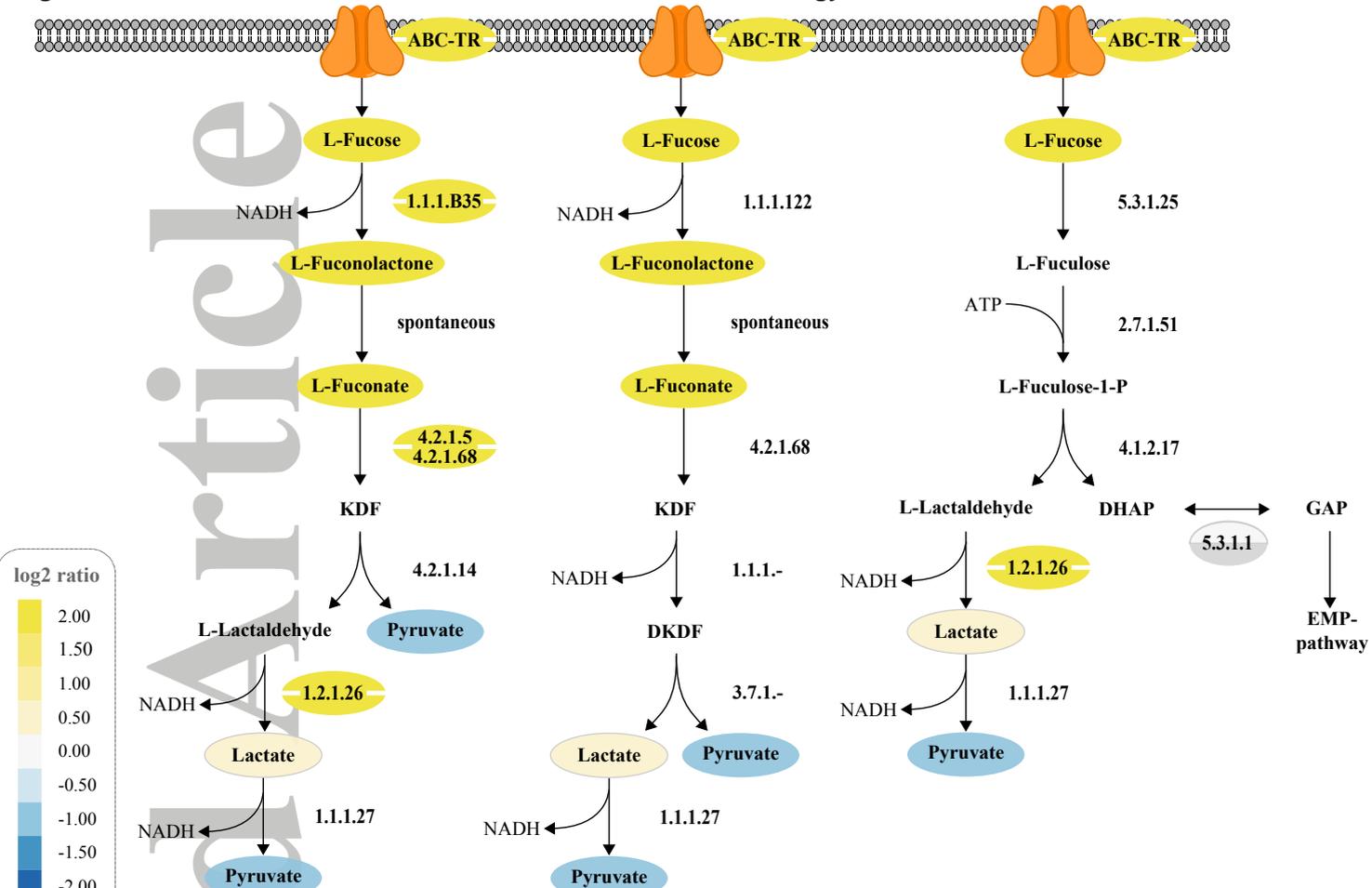
(E) Lactate dehydrogenase assay, using pyruvate and NADPH as substrates.

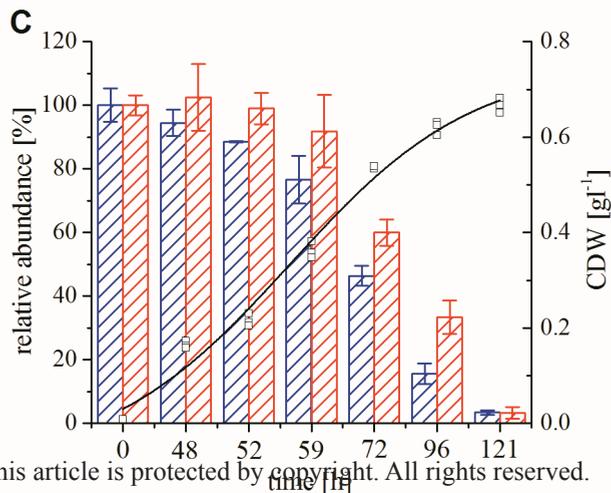
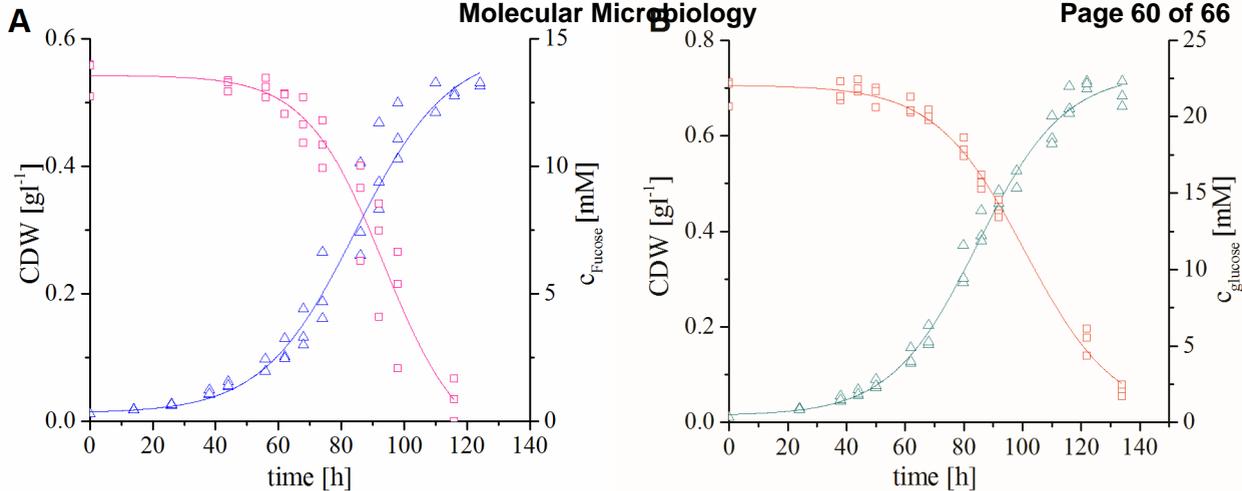
**Figure 7: Predicted flux distributions at the metabolic branch-point pyruvate for growth of *S. solfataricus* P2 on D-glucose and L-fucose minimal media.** Metabolites are represented as blue nodes and enzymes catalyzing individual reactions as yellow nodes. Numbers in metabolite nodes are total fluxes in  $\text{mmol} \cdot \text{g}_{\text{CDW}}^{-1} \cdot \text{h}^{-1}$ . Numbers on the edge labels represent split-ratio (relative producing

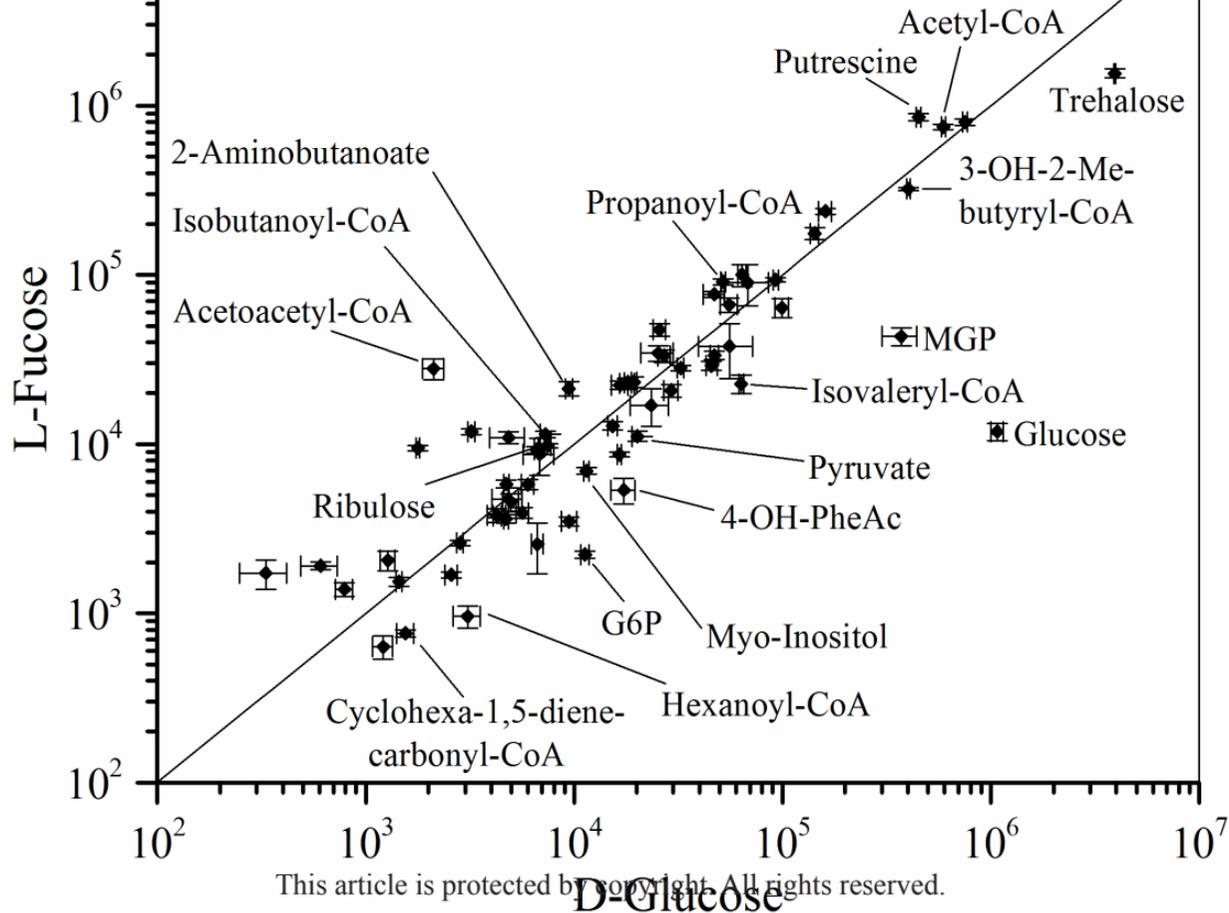
and consuming) fractions of the flux through the adjacent metabolite node. (Top: glucose scenario, bottom: L-fucose scenario) (Riemer *et al.*, 2013). Relative fluxes below 5% are not shown.

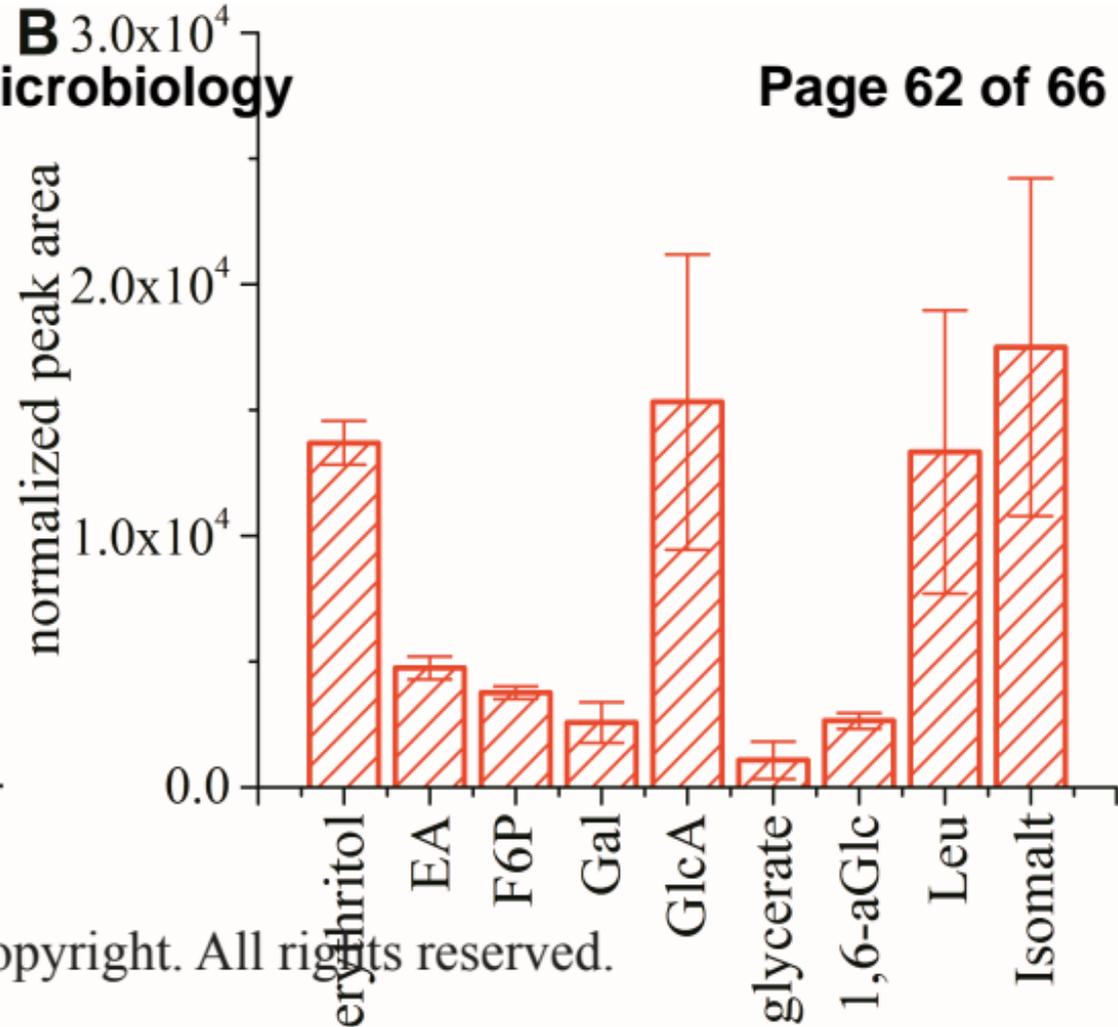
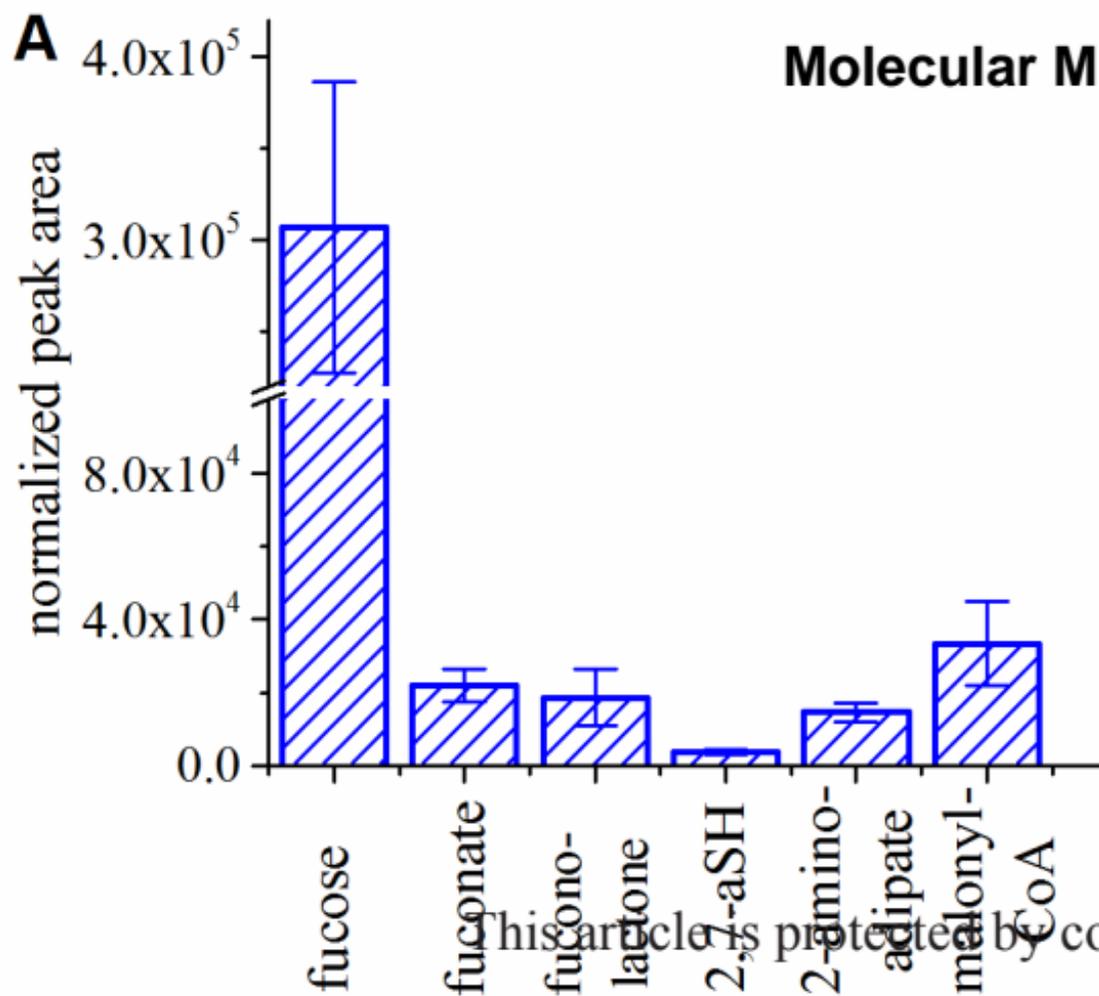
Abbreviations: KDF - 2-keto-3-deoxy-L-fuconate.

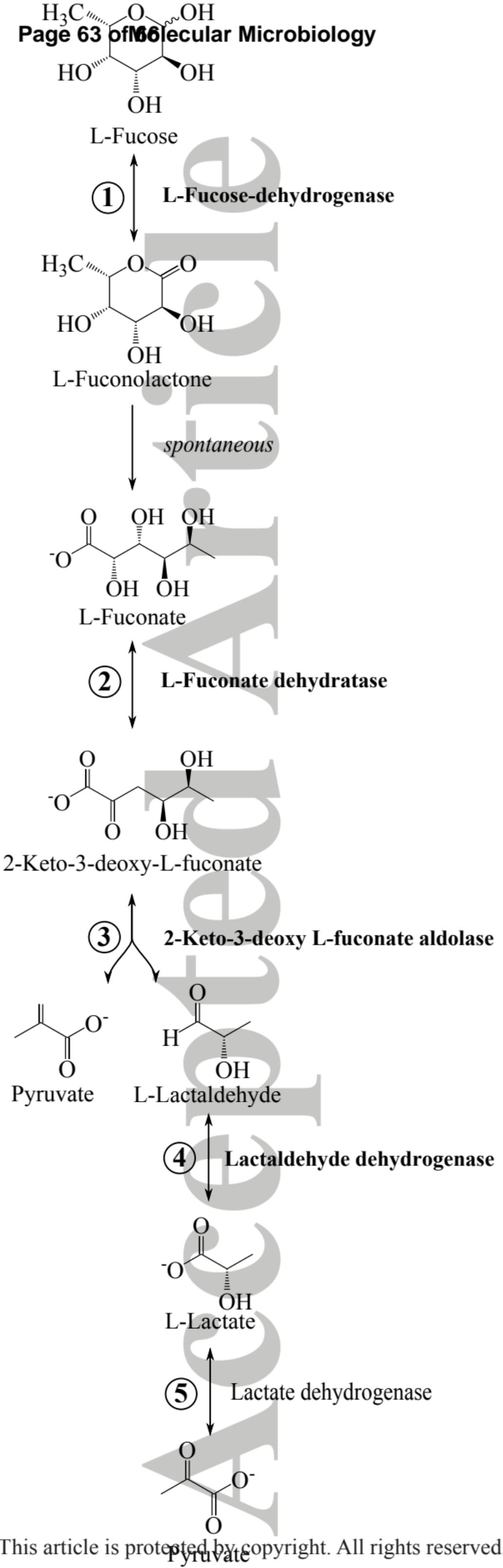
Accepted Article

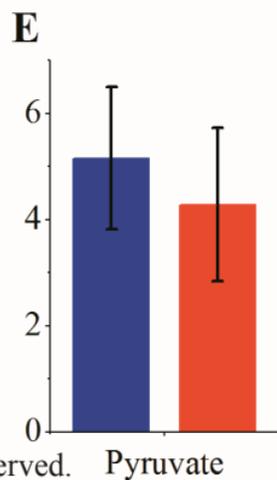
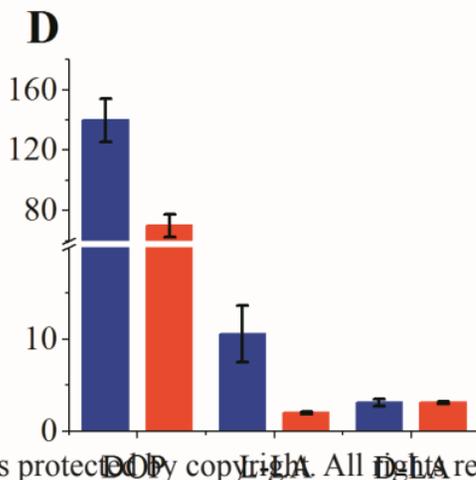
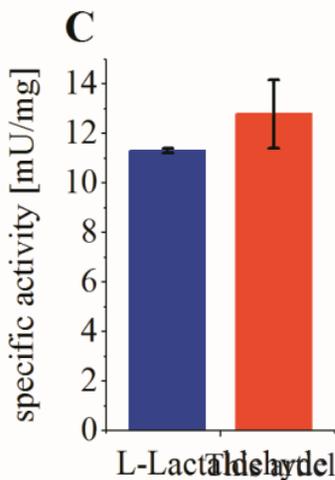
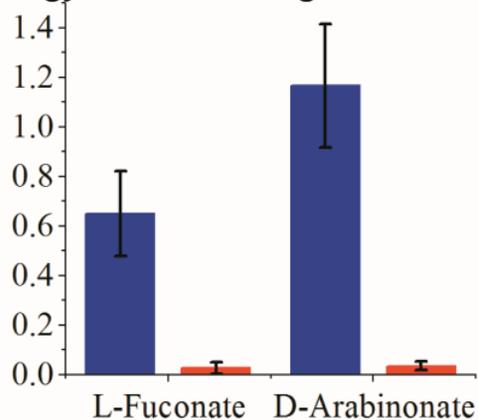
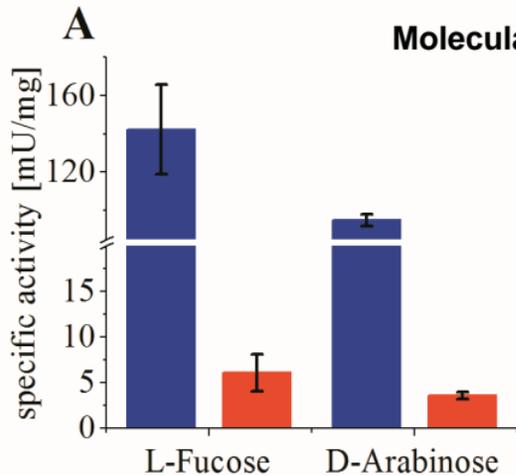












Pyruvate kinase  
EC 2.7.1.40

KDG aldolase  
EC 4.1.2.55

KDF aldolase  
EC 4.1.2.55

L-Lactate dehydrogenase  
EC 1.1.1.27

42.06%  
0.00%

57.78%  
0.52%

0.00%  
49.69%

0.00%  
49.69%

pyruvate  
1.7756  
2.4350

87.44%  
80.15%

5.12%  
1.75%

5.42%  
4.66%

0.00%  
11.69%

Pyruvate synthase  
EC 1.2.7.1

Acetolactate synthase (C6)  
EC 2.2.1.6

Acetolactate synthase (C5)  
EC 2.2.1.6

PEP synthase  
EC 2.7.9.2

Archaeal metabolism often differs from its bacterial and eukaryotic counterparts. Here, we investigate the global metabolic changes in response to growth on L-fucose versus D-glucose in the thermoacidophilic Crenarchaeon *Sulfolobus solfataricus* P2. We present a new pathway for L-fucose degradation in this organism which is also the first description of L-fucose metabolism in hyperthermophilic archaea. Interestingly, L-fucose degradation has shown to be promiscuous in *S. solfataricus*, partially using the enzymes involved in D-arabinose degradation.

Accepted Article

1.1.1.B35

spontaneous

KDF

L-Fuconate

4.2.1.5

4.2.1.68

4.2.1.14

Pyruvate

L-Lactaldehyde

1.2.1.26

Lactate

1.1.1.27

Pyruvate