SUMMARY

Challenging metabolic networks at high temperature: The central carbohydrate metabolism of *Sulfolobus solfataricus*

The central carbohydrate metabolism (CCM) of *S. solfataricus* P2 may well reflect much of the complexity and variety of central metabolic pathways in Archaea. Although many unusual enzymes and pathways have been identified in Archaea, the knowledge about their network topology and about their regulation is rather limited. The present study examines whether the CCM of *S. solfataricus* adapts to changes in temperature through network mechanisms. As such, the study is part of the Sulfolobus Systems Biology (“SulfoSYS”) project, which aims to integrate genomic, transcriptomic, proteomic, metabolomic and biochemical data to construct a comprehensive understanding of the network regulation of the CCM of *S. solfataricus* in response to the challenges offered by high and variable temperatures. SulfoSYS also does this by constructing dynamic models of relevant parts of the network (Albers et al., 2009).

An important requirement for network understanding is the determination of the topology of the network, i.e. the elucidation which biochemical reactions and gene products are present and how they connect to form integral pathways. In the course of this project the genome annotations of *S. solfataricus* and *S. acidocaldarius* were up-dated to improve the genome-wide analyses and bioinformatics was used to reconstruct the CCM of *S. solfataricus*. Thereby, in addition to the non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase (GAPN) and 2,5-dioxopentanoate dehydrogenase, three paralogs belonging to the aldehyde dehydrogenase superfamily were identified in the genome of *S. solfataricus*. To address the question if these proteins could indeed fulfill functions in the investigated network, the recombinant proteins were analyzed with respect to their substrate specificity. The studies revealed that *Sso*-1629, *Sso*-1842 represent succinic semialdehyde dehydrogenases, whereas *Sso*-1218 shows activity towards methylmalonate semialdehyde. These three gene products are not part of the CCM network of interest in this study thereby. By contrast I propose a function of an aldehyde oxidoreductase with ferredoxin rather than NAD(P)\(^+\) as electron acceptor, encoded by *Sso*-2636, *Sso*-2637, and *Sso*-2639, converting glyceraldehyde to glycerate in *S. solfataricus*. This proposal leads to a network in *S. solfataricus* with a complete semi-phosphorylative (sp) and non-phosphorylative (np) branch of the Entner-Doudoroff (ED) pathway. Thereby the cells have one route that makes one ATP (from ADP and phosphate) for each glucose molecule producing two molecules of pyruvate, and two routes that make no ATP.

The existence of a continuous network topology on the basis of the genes that encode enzyme activities does not guarantee that the complete topology is active
under all conditions. We have shown however that for the ED pathway all steps are active at the three temperatures we investigated, thus suggesting that the metabolic network could be operational during growth of *S. solfataricus*. We then examined a number of possible mechanisms of thermal adaptation of the network. We could show that the cells do not solve the problem of thermoadaptation in the molecular way, i.e. by having all enzymes temperature independent. Neither is the thermal adaptation based on the simplest network model of thermoadaptation, i.e. that of equal temperature coefficients of all enzymes. The enzymes do have different temperature coefficients. This would lead to variations of steady-state metabolite levels with temperature, unless the mild changes in gene expression with temperature or the $K_M$ dependences on temperature would compensate precisely.

We report some evidence for a combination of metabolic and gene expression regulations as the cell’s response to the temperature-change challenge. This needs to be further examined and confirmed by more detailed modeling and more accurate endometabolomics.

To obtain more insight into the function as well as in the utilization of the two ED branches in *S. solfataricus*, the glycerate kinase (enzyme of the np-ED branch) was investigated further. A deletion strain was constructed which specifically had blocked the sp-ED branch. The glycerate kinase is inhibited at high concentrations of its substrate glycerate. Thereby, glycerate as well as other metabolites of the np-branch accumulated when the sp-ED branch was shut down by the deletion. As shown by growth studies with the deletion strain, this accumulation does not interfere with growth, suggesting that the function of glycerate kinase is to buffer the flux into the lower common shunt of the EMP pathway, and not to keep the concentration of glycerate low.

In models of (hyper)thermophilic systems, the spontaneous i.e. non-enzyme-catalyzed reactions that are due to the thermal instability of intermediates, need to be integrated. Therefore the decay kinetics of a number of CCM intermediates were determined. Especially the triosephosphates, i.e. glyceraldehyde 3-phosphate (GAP; half-life 1.6 min (80°C)), dihydroxyacetone phosphate (DHAP; half-life 5.3 min (80°C)) and 1,3bisphosphoglycerate (1,3BPG; half-life 1.6 min (60°C) (Ahmed *et al*., 2005)) are heat-labile. This adds at least three steps to the network topology defined by the genome network. Accordingly, a small-scale blueprint kinetic model of a part of glycolysis (GAP to pyruvate conversion) was developed by adding the 1,3BPG decay and the additional GAPN reaction of *S. solfataricus* to the model of baker’s yeast (*S. cerevisiae*; (Teusink *et al*., 2000). The resulting model simulates that it should be most advantageous for life at high temperature to redirect the glycolytic flux through the GAPN reaction and to shut down the phosphoglycerate kinase (PGK) during glycolysis. This should prevent an energetically unfavorable futile cycling through PGK, which would make the ATP/pyruvate yield of glucose catabolism negative. We also discuss that
thermolability is not just an energy problem. It may also lead to the accumulation of toxic compounds such as methylglyoxal and glyceraldehyde.

Network function is not only determined by the topology defined by the genome, nor is it by the expression of that topology. The kinetic properties of the enzymes are also a major determinant of whether and where the fluxes flow and which metabolites accumulate. Of the enzymes involved, GAPN, glyceraldehyde 3-phosphate dehydrogenase (GAPDH) and PGK were characterized. This comprised the cloning, expression and detailed analysis at different temperatures (60°C to 88°C) with the aim of parameterizing the blueprint model in terms of the regulatory parameters that determine the flux through these particular reactions. The comparison of kinetic parameters of the enzymes as well as analyses of deletion strains revealed that in *S. solfataricus* GAP only seems to be catabolized by GAPN (due to a high activation by G1P and an increasing GAP oxidation with temperature), while the GAPDH/PGK couple is required for anabolism. The 1,3BPG producing step (PGK) in the anabolic mode slows down at higher temperatures (due to a tight regulation by the energy charge of the cell), while the GAP consuming step (GAPDH) is accelerated at the same time, which might result in minimized amounts of the instable intermediate 1,3 BPG. This proposed flux pattern still has to be confirmed by further modeling studies.

Furthermore, it could be shown that in *S. solfataricus* GAP and DHAP are quickly removed and trapped into fructose 6-phosphate via the bifunctional fructose-1,6-bisphosphate aldolase/phosphatase. The close coupling of these two activities (i.e. phosphatase and aldolase) ensures unidirectional catalysis towards gluconeogenesis and the rapid conversion of heat-labile triosephosphates to thermostable fructose 6-phosphate (Say et al., 2010). The observed regulation patterns might enable *S. solfataricus* to keep the pools of heat-labile compounds small in order to avoid their decay. To understand this regulation mechanism sufficiently is an ongoing challenge. The results obtained here are currently used to construct a precise model of the relevant part of the CCM of *S. solfataricus* (in collaboration with Jacky Snoep and Hans V. Westerhoff). Measurements will be performed to quantify the concentrations of intermediates as well as the carbon and energy loss the organism incurs.

I conclude that *S solfataricus* responds in multifarious and complex ways to the perhaps most pervasive challenge that living organisms are subject to, i.e. that of high and variable temperatures, an instantaneous challenge that cannot be defended against through biological membranes. Therefore, it may be no wonder that the organisms defense is partly immediate (metabolic) and partly and perhaps more strategically indirect, i.e. through gene expression.