Introduction

The archaeon *Sulfolobus solfataricus* is a thermoacidophile preferring growth at around 80°C and a pH of 2.5 to 3.5. As a thermoacidophile *S. solfataricus* faces two major problems: firstly, the proton permeability of membranes increases with temperature resulting in difficulties to establish proton gradients across the membrane; secondly, the proton concentration in the medium is very high, which increases the pressure on the proton backflow across the membrane even more. It has been shown that *S. solfataricus* antagonizes the increased proton permeability of the membrane by the unique composition of its membrane. Membranes consist of tetraether lipids, which have C₄₀ isoprenoid acyl chains. These lipids span the entire membrane forming a monolayer. Liposomes made of these lipids showed the same proton permeability at high temperature as liposomes from mesophilic organisms measured at their respective growth temperature. The low proton permeability of its membrane enables *S. solfataricus* to establish a pH gradient of 4 pH units (pH 2.5 outside/pH 6.5 inside) resulting in a very high proton motive force, which is reduced by a reversed $\Delta\psi$, inside positive.

Sugar transport in *S. solfataricus*

Because of the large $\Delta$pH, we assumed that *S. solfataricus* would use the proton motive force to fuel uptake of substrates from the medium by secondary transport. These transport systems catalyse the uptake of solutes in symport with protons. Initial research focused on establishing a membrane vesicle system for uptake experiments. Glucose transport studies in whole cells of *S. solfataricus* indicated that this substrate is taken up by a high affinity system with a $K_m$ of 2 $\mu$M (Chapter 2) that utilizes a binding protein. The glucose-binding protein (GlcS) was purified from solubilized membranes of *S. solfataricus*, and found to bind glucose with very high affinity ($K_d$ 480 nM). Interestingly, galactose and mannose are also substrates for GlcS, but not 2-deoxyglucose. GlcS belongs to the group of sugar-binding proteins, but exhibits some unusual features: (i) GlcS is glycosylated, and (ii) seems to remain membrane anchored by means of a transmembrane segment. This in contrast to bacterial binding proteins that either float freely in the periplasm or are anchored to the membrane via an attached lipid. (iii) GlcS possesses a very short, unusual signal sequence (Chapter 5). After N-terminal sequencing of the GlcS protein, the corresponding gene was identified in the *S. solfataricus* P2 database. Upstream of the *glcS* gene, genes encoding an ABC transport operon were found that include two permeases and one ATP-binding protein. In the glycoprotein fraction of solubilized membranes of *S. solfataricus*, additional sugar binding proteins were detected and their substrate specificities were determined (Chapter 4). Binding proteins for arabinose (AraS), cellobiose/cello-oligomers (CbtA), maltose-/maltotriose (MalE) and trehalose (TreS) were analyzed. N-terminal sequencing allowed the identification of the respective ABC transporters in the genome database of *S. solfataricus*. Interestingly, the five described ABC transporters fall into two families: the glucose, arabinose and trehalose transporter are typical sugar ABC transporters, composed of two membrane proteins and one ATP-binding domain, whereas the cellobiose and maltose system are homologous to oligo/dipeptide ABC transporters, consisting of two permeases and
two different ATP-binding proteins. This classification is also found in the structural organization of the binding proteins: AraS, GlcS and TreS are equipped with an unusual short signal sequence (see below), which is followed by a transmembrane segment and a so called linker region consisting of a stretch of 30-60 residues of hydroxylated amino acids. On the other hand, MalE and CbtA contain typical bacterial signal peptides and in addition a carboxyl-terminal transmembrane segment that is proceeded by a serine/threonine (ST)-rich linker region. The ST-rich linker is a putative site for O-linked glycosylation in analogy to other archaeal proteins.

Taken together these data suggest that *S. solfataricus* predominantly uses ABC transporters for sugar uptake. The complete genome indicates the presence of a large number of secondary transporters, but the substrates for these systems have not yet been identified.

This preference for ABC transporters in sugar uptake may be a general feature for (hyper)thermophiles including thermophilic bacteria. The genomes of these organisms also indicate the presence of large numbers of transporters that belong to the sugar and oligo/dipeptide ABC transporters family. It is unlikely that all of the latter systems are involved in peptide transporter, and in analogy with *S. solfataricus*, they may catalyse sugar uptake. *Sulfolobus* and *Thermoplasma* are heterotrophs in contrary to other

![Fig. 1. Model of the regulatory and catalytic cycle of the *E. coli* maltose transport system. At the left site the regulatory cycle is depicted showing MalK, the ATP-binding protein, in the ATP bound state which is able to bind MalT. Upon binding of maltose loaded maltose-binding protein (MBP) ATP is hydrolyzed, which releases MalT and leads via unknown conformational changes to the opening of the translocation pore. The ADP bound to MalK is then readily exchanged by ATP and dependent on the availability of substrate, the complex enters again the catalytic cycle or is arrested in the regulatory cycle by binding of MalT.](image-url)
hyperthermophiles. This means that they are more versatile in using substrates for growth and thus need a greater diversity of transporters.

The ATP-binding subunit of the glucose transport system

The ATP-binding subunit of the glucose ABC transporter, GlcV, has been characterized biochemically and structurally (Chapter 3). GlcV shows Mg$^{2+}$-dependent ATP hydrolysis activity and is highly thermostable. The crystal structure revealed that GlcV consists of two domains. The amino-terminal domain constitutes an ATP-binding domain with a fold that is nearly identical to that of *S. typhimurium* HisP and *T. litoralis* MalK. The carboxyl-terminal extension in GlcV has also been found recently in the *T. litoralis* MalK, but not in the *S. typhimurium* HisP. Based on the homology with *E. coli* MalK it can be assumed that this domain is involved in regulation of gene expression. This region may bind and therefore inactivate MalT, the activator of the mal gene expression. In archaea little is known about regulation of sugar transport. There are no *in vivo* studies reported as to yet. The data on *E. coli* MalK indicate that ATP-hydrolysis and binding or release of MalT might be coupled processes. It is assumed that in the ATP bound state MalK does not catalyse transport. This conformation may bind MalT (Fig. 1). Upon contact of the maltose loaded binding protein at the periplasmic site with the translocation domain (MalGF) a conformational change is induced that propagates to MalK via the contact sites (most likely the signature motif) with MalGF. This triggers ATP-hydrolysis, and causes a conformational change of MalK leading to the opening of the transport channel for passage of the substrate, and secondly, release of MalT and subsequently activation of the mal gene expression. After the replacement of ADP by ATP, MalK may return to a conformation that can bind MalT under conditions that maltose is not available. If this model can be extrapolated to the glucose transporter of *T. litoralis*, an transcriptional regulator like MalT should also be present in archaea. However, such protein cannot be found in the genomic database. Further biochemical studies therefore must address the identification of the archaeal MalT-homologue for instance through direct binding studies with GlcV. Information about the conformational change induced in GlcV by the binding and hydrolysis of ATP is needed to understand the molecular mechanism of energy coupling in ABC transporters.

Signal sequence peptides in *S. solfataricus*

Analysis of the *S. solfataricus* database revealed that approximately 4.5% of all proteins contain a putative signal peptide (Chapter 7). Proteins containing secretory signal peptides (including twin arginine signal peptides), lipoprotein signal und type IV pili signal peptides were identified. The major part of these proteins has an unknown function, but some represent substrate-binding proteins, putative extracellular proteases, S-layer proteins and flagellins. The group of proteins containing type IV pili signal peptides is quite large (9 proteins). Most of them are solute binding proteins and one corresponds to the flagellin. From four of these proteins the unusual cleavage site was demonstrated biochemically by amino-terminal sequencing. All archaeal flagellines, which were already purified from a wide variety of organisms, exhibit type IV pilin signal peptides. Unlike all other studied archaea, *S. solfataricus* has only one gene coding for a flagellin, whereas flagella of *S. shibatae* are composed of two different proteins. Next to the flagellin gene in the *S. solfataricus* genome, insertion of a transposon has occurred possibly resulting in the loss of the second flagellin gene.

Bacterial type IV pilins are cleaved by a peptidase, PilD (*Pseudomonas aeruginosa*), which removes the positively charged amino-terminus and concomitantly methylates the new amino-terminus. PilD homologues, responsible for the processing of archaeal flagellins, have
not yet been identified. In *S. solfataricus* not only the flagellin, but also the sugar-binding proteins are processed by a PilD homologue. There is no evident reason, why binding proteins of the sugar cluster are processed by this mechanism. Binding proteins from the di/oligopeptide cluster remain after processing membrane-bound through an amino-terminal transmembrane segment, and, unlike flagellins, these proteins are not immediately incorporated into a larger macromolecule. The question arises, whether the binding proteins from the sugar cluster are translocated via a similar pathway as flagellins and whether they are integrated in a macromolecular structure after processing. Downstream of the flagellin gene, a gene cluster is located that shows homology to bacterial type II secretion systems. In Gram-negative bacteria, such systems are involved in the export of toxines and hydrolases across the outer membrane. These polypeptides form large multimeric structures in the periplasm before further export across the outer membrane. Bacterial type II, III, and IV secretion systems share the basic structure and organization.

The identified type IV pilin signal peptides show that the specificity of the *S. solfataricus* PilD is quite similar to the bacterial homologue. For instance, in bacteria these peptides contain a conserved glycine at the –1 position. This glycine residue can only be substituted for an alanine, suggesting that the bulkiness of this position is critical for recognition. In *S. solfataricus* binding proteins, alanine is found more frequent at –1 position, and this residue is also present in signal peptide of the *T. acidophilum* flagellin. To identify the PilD homologue of *S. solfataricus* an *in vitro* cleavage assay has been established (Chapter 6). The flagellin (FlaB) and GlcS were carboxyl-terminally his-tagged and expressed in *E. coli*. *E. coli* membranes containing these proteins were mixed with *Sulfolobus* membranes and incubated at high temperatures. Under such conditions, specific processing occurred of the preFla and preGlcS proteins. This assay will be used in the future to monitor the purification of the type IV pilin signal peptidase of *S. solfataricus*. 