University of Groningen

Sugar transport in Sulfolobus solfataricus is mediated by two families of binding protein-dependent ABC transporters

Elferink, M.G.L.; Albers, S.V.; Konings, W.N; Driessen, Arnold

Link to publication in University of Groningen/UMCG research database

Sugar transport in *Sulfolobus solfataricus* is mediated by two families of binding protein-dependent ABC transporters

Marieke G. L. Elferink,[†‡] Sonja-V. Albers,[‡]
Wil N. Konings and Arnold J. M. Driessen*
*Department of Microbiology, Groningen Biomolecular
Sciences and Biotechnology Institute, University of
Groningen, Kerklaan 30, 9751 NN Haren, The
Netherlands.*

Summary

**The extreme thermoacidophilic archaeon *Sulfolobus
solfataricus* grows optimally at 80°C and pH 3 and
uses a variety of sugars as sole carbon and energy
source. Glucose transport in this organism is mediated
by a high-affinity binding protein-dependent ATP-
binding cassette (ABC) transporter. Sugar-binding
studies revealed the presence of four additional
membrane-bound binding proteins for arabinose,
cellobiose, maltose and trehalose. These glycosy-
lated binding proteins are subunits of ABC trans-
porters that fall into two distinct groups: (i)
monosaccharide transporters that are homologous
to the sugar transport family containing a single
ATPase and a periplasmic-binding protein that is
processed at an unusual site at its amino-terminus;
(ii) di- and oligosaccharide transporters, which are
homologous to the family of oligo/dipeptide trans-
porters that contain two different ATPases, and a
binding protein that is synthesized with a typical
bacterial signal sequence. The latter family has not
been implicated in sugar transport before. These data
indicate that binding protein-dependent transport is
the predominant mechanism of transport for sugars
in *S. solfataricus*.**

Introduction

Sulfolobus solfataricus is an obligate aerobe that grows
either lithoautotrophically or chemoheterotrophically in hot
and acid environments. It originates from a solfataric field

with temperatures between 75°C and 90°C and pH values
of 1–3 (Brock *et al.*, 1972). These environments contain
decomposing materials of higher plants, including cellu-
lose and starch that can act as potential carbon sources.
S. solfataricus has been reported to grow on a variety of
reduced organic compounds as sole carbon and energy
source (Grogan, 1989). Although the metabolic pathways
for the degradation of various sugars have been studied in
detail (Schönheit and Schafer, 1995; Kengen *et al.*, 1996),
little is known about how these sugars are transported into
the cell. An exogenous α-amylase has been purified from
S. solfataricus that catalyses the hydrolysis of starch,
dextrin and α-cyclodextrin into maltose and maltodextrins
(Haseltine *et al.*, 1996). These compounds can be
hydrolysed by a soluble cytosolic maltase (α-glucosidase)
to release glucose in the cell (Rolfsmeier and Blum,
1995; Rolfsmeier *et al.*, 1998). A β-glucosidase, initially
described by Cubellis *et al.* (1990), was characterized
by D'Auria *et al.* (1996), but the native substrate for this
enzyme is unknown. Recently, amylolitic activity that
converts starch to trehalose was detected in the culture
supernatant of *S. solfataricus* KM1 cells (Kato *et al.*,
1996). This trehalose-producing activity appears to result
from the co-operative action of two enzymes, a glycosyl-
transferase and an amylase (Kobayashi *et al.*, 1996). The
utilization of sugars as trehalose, maltose and maltotriose
necessitates the presence of uptake systems.

The availability of the partial sequence of the *S.
solfataricus* P2 genome (Sensen *et al.*, 1998) permits
the rapid identification of transport systems in this
organism. *S. solfataricus* contains a number of secondary
transporters that belong to the major facilitator super-
family (MFS) and various binding protein-dependent
transporters belonging to the ATP-binding cassette (ABC)
family of transporters. On the other hand, phospho*enol*
pyruvate-dependent phosphotransferase systems (PTS)
appear to be absent in *S. solfataricus and* in other archaea.
Glucose, galactose and mannose enter *S. solfataricus* cells
via a binding protein-dependent ABC-type transporter
(Albers *et al.*, 1999a). Such systems have also been
found for the uptake of sugars in other archaea as, for
instance, the uptake of maltose and trehalose in the
hyperthermophilic *Thermococcus litoralis* (Xavier *et al.*,
1996). The genes encoding this system have been
cloned and sequenced and shown to be homologous to

the maltose transporter of *Escherichia coli*, MalEFGK (Horlacher *et al.*, 1998). Recently, three ABC transporters of *Haloferax volcanii* have been functionally characterized and shown to be essential for nitrate respiration (Wanner and Soppa, 1999).

In this study, we have undertaken a systematic screening for the presence of sugar-binding proteins in *S. solfataricus*. The data show that *S. solfataricus* uses binding protein-dependent ABC-type transporters for the uptake of a variety of sugars. Surprisingly, some of these systems belong to the family of oligo/dipeptide transporters that have so far only been implicated in peptide transport. The physiological implications of these findings are discussed.

Results

Membrane-bound glycosylated sugar-binding proteins in S. solfataricus

S. solfataricus uses a number of sugars as sole carbon and energy source, including arabinose, glucose, galactose, cellobiose, maltose, sucrose, trehalose, lactose, fructose and melibiose (data not shown). Recently, we reported the identification of a membrane-bound binding protein that functions as a subunit of an ABC-type transport system for glucose, galactose and mannose (Albers *et al.*, 1999a). During the purification of this protein, concanavalin A (ConA) affinity chromatography of detergent-solubilized membranes was used to purify glycosylated membrane proteins selectively (Albers *et al.*, 1999a). As this method allows the purification of a limited number of proteins within the molecular mass range of 30–300 kDa (Fig. 1), we tested this fraction for the presence of other glycosylated substrate-binding proteins. Membranes isolated from cells grown on arabinose were solubilized with Triton X-100 and subjected to ConA column chromatography. The protein fraction eluted with α-methylmannopyranoside (Fig. 1, lane 3) was dialysed and used for substrate-binding assays using radiolabelled substrates. Binding activity could be detected for arabinose, fructose, xylose, glucose, galactose, cellobiose, maltose and trehalose. No binding activity was found for lactose, melibiose, citrate and the amino acids glutamate, leucine and alanine.

To assess the number of individual binding proteins, the specificity of the various binding activities found in the ConA fraction was determined by measuring the binding of radioactive-labelled sugars in the absence and presence of a 10-fold excess of unlabelled sugar. Binding of radiolabelled arabinose (Fig. 2A), fructose (Fig. 2B) and xylose (Fig. 2C) was reciprocally inhibited by the unlabelled sugars indicating that these substrates share a common binding protein. The inhibition studies also
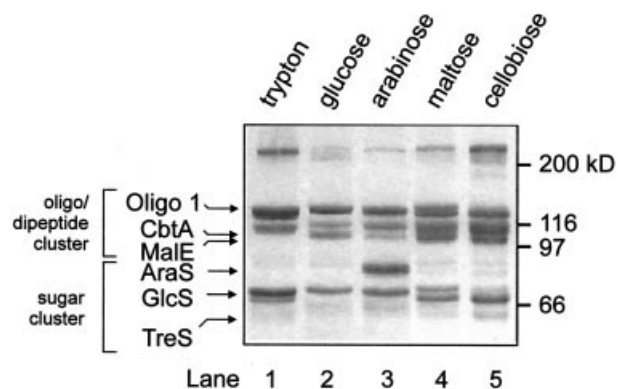

Fig. 1. Coomassie brilliant blue-stained SDS–PAGE of ConA fractions of membrane proteins derived from cells grown on the substrates indicated. The identified binding proteins (BP) are indicated by an arrow: Oligo 1, unknown; CbtA, cellobiose-binding protein; MalE, maltose-binding protein; AraS, arabinose-binding protein; GlcS, glucose-binding protein; and TreS, trehalose-binding protein.

confirmed our earlier report (Albers *et al.*, 1999a) showing that glucose (Fig. 2D) and galactose (Fig. 2E) share a common binding protein. Owing to the higher maltose-binding activity (Fig. 3), the ConA fraction derived from maltose-grown cells was used for the maltose (Fig. 2G) and cellobiose (Fig. 2F) binding inhibition experiment. Inhibition was observed only with the individual unlabelled substrate, suggesting the presence of distinct binding proteins for cellobiose and maltose.

The binding specificity was studied in further detail. D-[^{14}C]-arabinose binding was strongly inhibited by a 10-fold excess of unlabelled D-arabinose (88% inhibition) and L-arabinose (95%), whereas ribose and xylose were without an effect. Apparently, the protein does not discriminate between the stereoisomers of arabinose in analogy with the *E. coli* arabinose-binding protein (Quiocho and Vyas, 1984). We also noted that *S. solfataricus* grows on both D- and L-arabinose, in contrast to what has been reported before (Grogan, 1989). Binding of [^{3}H]-cellobiose was not only strongly inhibited by a 10-fold excess of unlabelled cellobiose (91%), but also by cellotriose (98.5%), cellotetraose (99%) and cellopentaose (99%). The increase in inhibition with polymer length indicates the presence of a cello–oligo-binding protein. Binding of [^{14}C]-maltose was strongly inhibited by a 10-fold excess of unlabelled maltose (76%), weakly by maltotriose (44%) and not at all by trehalose (0%). This binding activity is therefore specific for maltose/maltodextrin, and it differs from the maltose/trehalose-binding protein found in many other organisms (Xavier *et al.*, 1996; Horlacher *et al.*, 1998; Wassenberg *et al.*, 2000). As *S. solfataricus* can grow on trehalose as sole carbon and energy source, and [^{14}C]-trehalose binds to all examined ConA fractions (Fig. 3), an independent trehalose-binding protein must be present in this organism. Taken together, the results suggest the
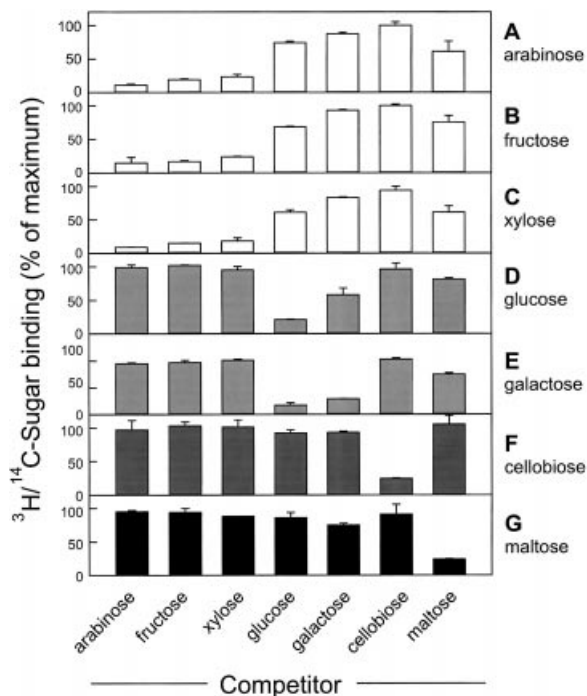
Fig. 2. Substrate specificity of the sugar-binding activity present in the ConA fraction of arabinose (A–E)- and maltose (F and G)-grown cells. Binding of 1 μM of the indicated radiolabelled sugars was determined in the presence and absence of a 10-fold excess of unlabelled competitor substrate. (A) Arabinose; (B) fructose; (C) xylose; (D) glucose; (E) galactose; (F) cellobiose; (G) maltose. The binding in the absence of competitor was set at 100%. The standard deviation shown is for two independent experiments. Different degrees of shading are used to indicate the substrate specificity of the binding activities.

presence of at least five distinct sugar-binding proteins in *S. solfataricus*, i.e. (i) arabinose/fructose/xylose; (ii) galactose/glucose/mannose; (iii) cellobiose and higher derivatives; (iv) maltose/maltodextrin; and (v) trehalose.

Induction and identification of sugar-binding proteins

To correlate the sugar-binding activities to specific proteins present in the ConA fraction, cells were grown on a variety of sugars. The respective ConA fractions were assayed for binding activity of each of the five binding proteins (arabinose, cellobiose, trehalose, glucose and maltose) (Fig. 3). Figure 1 shows the specific polypeptide patterns of the ConA fraction after growth on the indicated sugars and trypton. In cells grown on arabinose, the prominent expression of a protein with a molecular mass of about 80 kDa was evident (Fig. 1, lane 3). The arabinose-binding activity (Fig. 3A) correlated with the relative abundance of the 80 kDa protein in the various fractions (compare Fig. 1 and Fig. 3A), suggesting it to be the arabinose-binding protein. The protein was partially purified from the arabinose ConA fraction

using anion-exchange chromatography and found to bind D-arabinose with a K_D value of 130 nM (data not shown). The 80 kDa protein was blotted on a polyvinylidene difluoride (PVDF) membrane, and the N-terminal amino acid sequence was determined. A BLAST search with the amino acid sequence ISRTAI against the *S. solfataricus* databank yielded a perfect match with an open reading frame (ORF) encoding for a 69 315 Da protein, which we now assign as AraS (see Fig. 6A). AraS shows homology to sugar-binding proteins and, like the glucose-binding protein GlcS (Albers *et al.*, 1999a, b), it lacks a typical signal sequence and appears to be processed in an unusual manner that results in the removal of the first nine amino acids.

The glucose-binding activity is expressed constitutively (Figs 1 and 3), but is at its highest levels in cells grown on trypton. The abundance of the 65 kDa protein (Fig. 1) correlates with the glucose-binding activities (Fig. 3D) in the various fractions. This protein is identical to GlcS that was purified to homogeneity (Albers *et al.*, 1999a). Like GlcS, the trehalose-binding protein is also expressed under all growth conditions (Fig. 3C). The cellobiose- (Fig. 3B) and, to a lesser extent, maltose (Fig. 2E)-binding activities were increased when cells were grown on the respective substrates. Transport studies with glucose and maltose confirmed the results obtained with the binding assays. Both glucose- and maltose-grown cells showed a substantial glucose uptake rate, but the

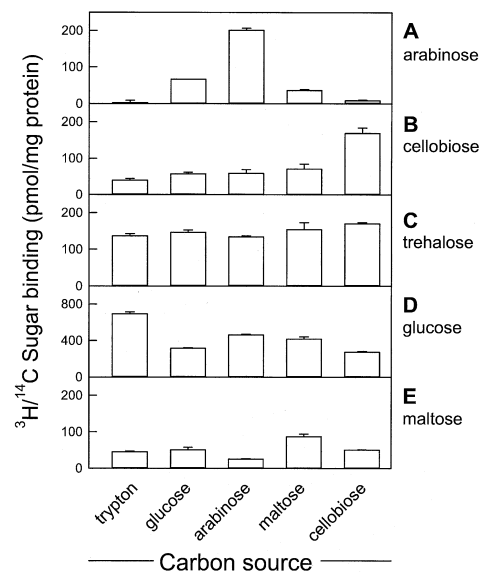**Fig. 3.** Binding of sugars by the ConA fraction derived from cells grown on various substrates. ConA fractions were isolated from cells grown on trypton, glucose, arabinose, maltose or cellobiose. Binding of radiolabelled arabinose (A), cellobiose (B), trehalose (C), glucose (D) and maltose (E) was determined using a substrate concentration of 1 μM. The standard deviation shown is for two independent experiments.
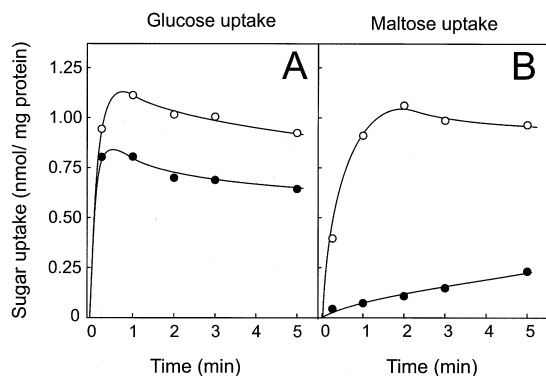
Fig. 4. Uptake of glucose (A) and maltose (B) by *S. solfataricus* cells grown on glucose (closed circles) and maltose (open circles). Radiolabelled glucose and maltose were added to a final concentration of 1 μM.

activity was lower in the glucose-grown cells (Fig. 4A). On the other hand, maltose uptake rates were high in maltose-grown cells and low in glucose-grown cells (Fig. 4).

To identify the other binding proteins, anion-exchange chromatography was used to separate the maltose- and trehalose-binding activities in the ConA fraction of maltose-grown cells (Fig. 5). Trehalose-binding activity correlated with a protein of about 60 kDa (Fig. 1). The N-terminal amino acid sequence LSTTTIIGIVV matched an ORF encoding a protein with a molecular mass of 53 264 Da that is homologous to sugar-binding proteins. The mature trehalose-binding protein, TreS, lacks the first four amino acids and is processed at the same site as GlcS and AraS (see Fig. 6A). The maltose-binding activity correlated with a protein of about 100 kDa (Fig. 1). Repeated N-terminal sequence analysis yielded ambiguous results. The best match with the obtained sequence [QY][SA][PFLN][QP]

was QSPQ, which corresponds to the amino-terminus of a mature protein (79 982 Da) synthesized with a typical bacterial signal sequence. This protein, tentatively named MalE, is identical to the gene product of ORF1 that is localized next to the maltase (*malA*) gene on the *S. solfataricus* chromosome (Rolfsmeier *et al.*, 1998). MalE shows homology to oligo/dipeptide-binding proteins. Partial purification of the ConA fraction of cellobiose-grown cells revealed that the cellobiose-binding activity correlated with a protein of about 97 kDa. The N-terminal sequence ASSPFP[SP]TLYL matches with the amino-terminus of a 76 665 kDa protein that, like MalE, is synthesized with a typical signal sequence belonging to the family of oligo/dipeptide-binding proteins (Fig. 6C). This protein is referred to as CbtA.

Taken together, the identified sugar-binding proteins can be divided into two clusters: (i) proteins with a molecular mass ranging from 53 to 70 kDa that are homologous to sugar-binding proteins and are synthesized with an unusual amino-terminal signal; and (ii) proteins with a molecular mass in the range of 76–100 kDa that are homologous to oligo/dipeptide-binding proteins and are synthesized with a typical signal sequence.

Further database searches identified another five binding proteins for which the substrate is unknown. Two are homologous to sugar-binding proteins referred to as sugar 1 and 2, whereas the other three belong to the oligo/dipeptide cluster referred to as oligos 1, 2 and 3. Oligo 1 could be identified as the upper band of the oligo/dipeptide cluster in the arabinose Con A fraction (Fig. 1). N-terminal sequencing yielded the sequence QSTSVQPE, which matches an ORF of 79 110 kDa. The N-termini of the binding proteins of the first cluster are aligned in Fig. 6A and those of the second cluster in Fig. 6B and C.
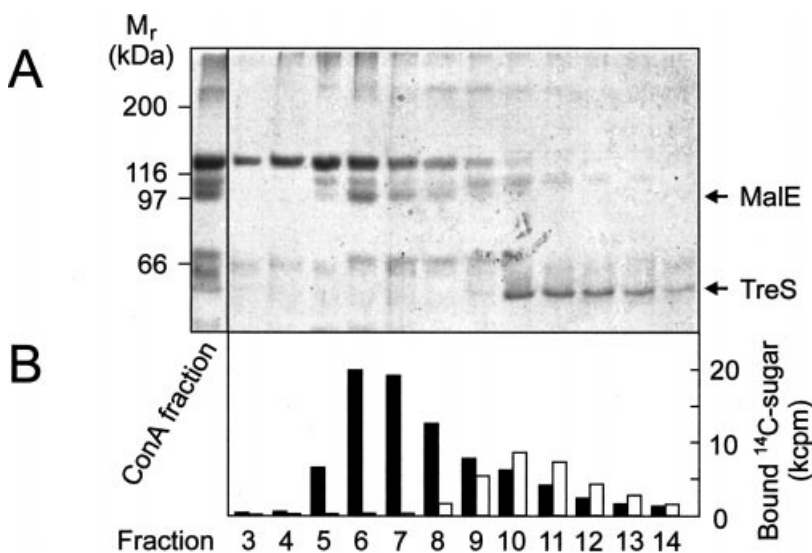

Fig. 5. Separation of the maltose- and trehalose-binding activity present in the maltose ConA fraction by MonoQ chromatography. (A) SDS–PAGE and (B) binding of maltose (filled bars) and trehalose (open bars) by the eluted membrane protein fractions. MalE and TreS are indicated by arrows.

A

```
                           ↓
GlcS   : MKRKYPYSLAKGLTSTQIAVIVAVIVIVIIIGVVAGFVLTKGP---------STTAVTTT : 51
AraS   : ---MSRRRLYKAISRTAIIIIVVVIIIAAIAGGLAAYYSSSKPPATSTSLTSTSSSLSVT : 57
sugar1 : MGRKGKKIDYKAISKTLVAVIIVVVIVIAIGGVYAFLSSQHSPAAPS----STTTSFTST : 56
TreS   : --------MRRGLSTTTIIGIVVAIVIIVIGAVAAVTLLSHKP-----------SQVVST : 53
sugar2 : ----------MKALSTLAMAVIIIVVIAVVAAAAYLITSSSHHPSIS-----TTTTPIIAT : 46

GlcS   : VTSTFTTTTTIPSTT--------TSTPSNTVVFYTWWGGGDGG--EALSQIIP-AV : 96
AraS   : TSSTTSTLSSITTTT--------STASSYVVDFINPWGAEDP---VGLKWIGGNFS : 102
sugar1 : TSSTTTSSVLNTSNPQALMQLVGISTPPSTPVTITVWNSYSTSENQAFNETLAQFE : 112
TreS   : TSPSTSQSATSTS-P---------SQVITITYFDDLSPSEAN---ITQKIIIPQFE : 96
sugar2 : NTTAPITLTVVTFSG---------QSANFIQYAGNLFHQLHPN-VQVEVIQYPFSE : 92
```

B

```
                                ↓
oligo1 : ------MKRYKIISTIITVLMVISIGIFAMPILSQSTSVQPEGS : 39
MalE   : MNKKIKNVIGLTALILMALSAFMPFIISSRVVNSQSPQLNPAAS : 45
CtbA   : ----MRKELVLEVGVIFSISVMLFSISGIMIANSASSPFPSTLY : 41
oligo2 : ------MSSLKGLALLSIMLIGIILPSLFLLQTSAQTSLTISPP : 39
oligo3 : --MYSVLSIKDKKIISLLILVATAISPIFAIAQSASSSPASTAI : 43
```

C

```
oligo1 : LALHLVNQSVPEPWWYYTSQVPSSWYTSNDPFVYQTTTTTTTTSTTTTTSTTTTSTTTSTVTTT: 656
MalE   : MSNAQEVAYTQAAWQDLLAFIYGKANTATPPSLLAMLQPST-S-----TVTQTSTSTSTVTTT: 672
CtbA   : ----------AL---------IDYWYAPLYESTIITTTTTSIS------TTITTTTS---AT: 648
oligo2 : ---------------------YYYNLMYYQPYTITISTSTS------TTSSTTTTT-LQTT: 672
oligo3 : ---PGSYLLSVSAYSNTSILINPVTYTLVVLPSITTTTTSTTTS------TTTTTSTTT-TTSV: 855

oligo1 : SVSTTTSVSTSVSTTTATVTTTVTS---SSNTTLYAIIAVVVIIIVIIGVILGLR-R------: 709
MalE   : SASTSVTTTTSMSTT--SVTSTVST---SSGLSTGVIAGIIVVVIIVIIAVVAYVVVR-R----: 720
CtbA   : TATTSVTTTTSVTTTS-ISTTTVTV---TSTSTIPIIIAIVIVIIVIIAAVAILMRR------: 692
oligo2 : TTSIVFGGATNITTSVTSMTTTTSS---SSSTLIYAVIGIVIVIIIIVVAVVLLRGRGRGGPG: 725
oligo3 : STSLSTTTITSVSTVISTVVSTIVSTVVSSASNIGYIAVIVVLIIIIILALAVLLFRR------: 905
```

Fig. 6. Alignment of the amino acid sequences of binding proteins.
A. N-terminal region of binding proteins with homology to sugar-binding proteins.
B. N-terminal region of binding proteins with homology to oligo/dipeptide-binding proteins.
C. C-terminal region of binding proteins with homology to oligo/dipeptide-binding proteins. Experimentally determined N-terminal sequences are underlined. Processing sites of the binding proteins are indicated by an arrow, putative transmembrane segments are highlighted in black, and the long hydrophilic stretch rich in hydroxylated amino acids is indicated in grey.

Structural organization of sugar-binding protein-dependent ABC transporters

After identification of the different binding proteins in the *S. solfataricus* database, adjacent proteins were analysed allowing the identification of the complete operons for the respective transport systems (Fig. 6). By analogy with the binding proteins, the permeases and the ATPases can be divided into the same two categories, the sugar and oligo/dipeptide cluster. The respective operon structures (see Fig. 7A and B) and structural organization resembles that typically found in bacteria (Higgins, 1992), i.e. a binding protein, one or two separate integral membrane subunits and one or two separate peripheral ATPase subunits. The
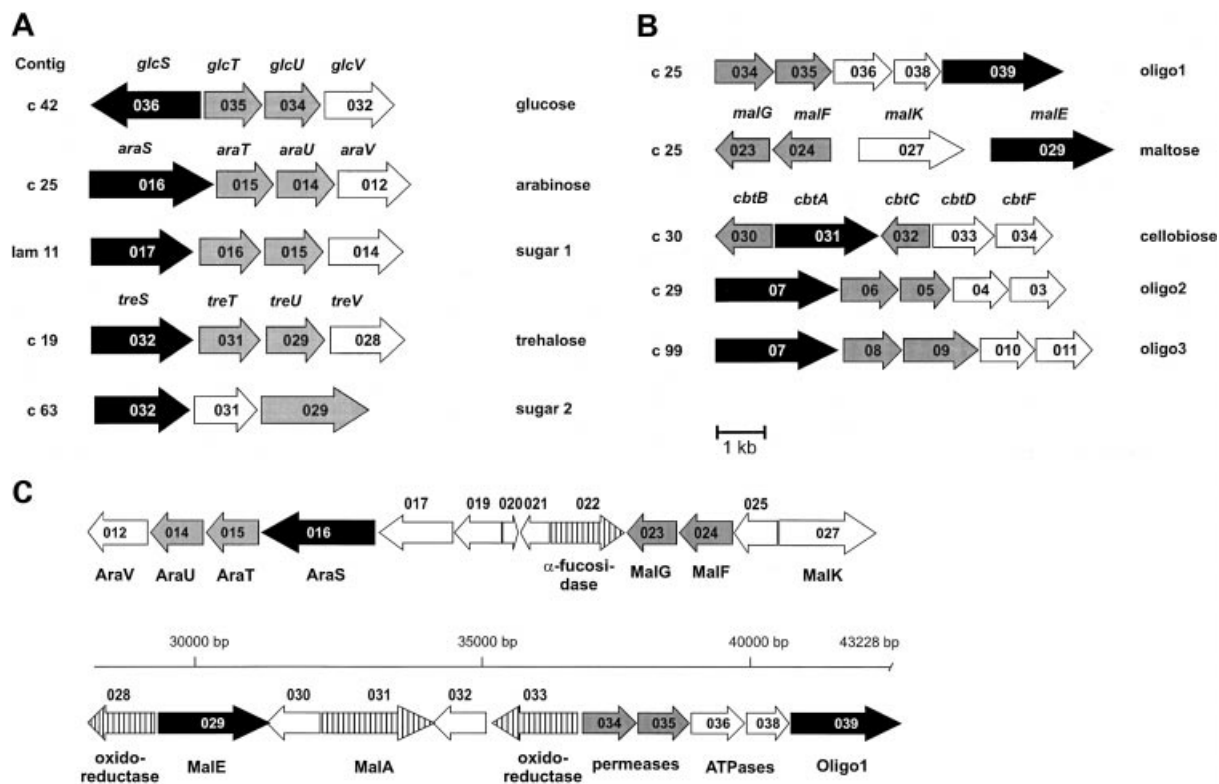
Fig. 7. Schematic representation of the gene organization of the ABC operons derived from the *S. solfataricus* database encoding sugar transporters.
A. Sugar transport cluster.
B. Oligo/dipeptide transport cluster. The length of the arrow indicates relative sizes of the proteins. Binding proteins, permeases and ATP-binding subunits are indicated in black, grey and white respectively. The numbers in the arrows indicate the respective ORFs that belong to the indicated contigs.
C. Clustering of sugar ABC operons and genes related to sugar metabolism on contig 25. Hatched arrows indicate genes encoding a function in sugar metabolism.

integral membrane domains (sugar 2) or peripheral ATPases (MalK) may also occur as fused proteins (Fig. 7).

Sugar-binding proteins are mannose glycosylated

The molecular masses of the identified binding proteins appeared to be smaller than estimated from SDS−PAGE, whereas the protein bands appear diffuse. This is most probably caused by the glycosylation of these proteins as they are selectively retained by ConA chromatography. To reveal the identity of the sugar residues present, binding proteins were analysed with a digoxigenin glycan differentiation kit that contains five different lectins, *Galanthus nivalis* agglutinin (GNA), *Sambucus nigra* agglutinin (SNA), *Maackia amurensis* agglutinin (MAA), peanut agglutinin (PNA) and *Datura stramonium* agglutinin (DSA). These lectins specifically discriminate between five types of glycosylations. The proteins of the arabinose ConA fraction were separated on SDS−PAGE and transferred to nitrocellulose. Blots were treated with the

different lectins to characterize the carbohydrate moieties. Each of the binding proteins could be detected with GNA, whereas no signal was obtained with the other lectins. GNA recognizes $\alpha(1-3)$, $\alpha(1-6)$ and $\alpha(1-2)$ links to terminal mannose, indicating their presence in the carbohydrate moieties of the binding proteins. Total sugar analysis of AraS and the whole arabinose ConA fraction resulted in mannose, glucose, galactose, N-acetyl-glucosamine and a large unknown species (data not shown). The latter may represent 6-sulphoquinovose recently identified in the cytochrome $b_{558/566}$ from *Sulfolobus acidocaldarius* (Zahringer *et al.*, 2000).

Discussion

By means of an integrated proteomic and genomic approach, we have identified a set of membrane-bound sugar-binding proteins in *S. solfataricus*. In addition to the previously described binding protein for glucose, galactose and mannose, GlcS (Albers *et al.*, 1999a), binding proteins could be identified for arabinose, fructose and

xylose (AraS), cellobiose and higher derivatives (CbtA), maltose and maltodextrin (MalE) and trehalose (TreS). Each of these sugars can serve as sole carbon source for the growth of *S. solfataricus*. The gene of each of these binding proteins is contained in an operon that encodes the complete binding protein-dependent ABC transporter. These transporters fall into two groups: the glucose, arabinose and trehalose transporters are homologous to sugar transport systems, whereas the maltose and cellobiose transporters show homology to oligo/dipeptide transporters (Figs 6 and 7). The latter observation is remarkable; such systems have so far been implicated in peptide transport only. MalE and CbtA both bind di- and oligosaccharides, and this recognition of oligomers may be a common feature of the members of ABC transporters of the oligo/dipeptide cluster. Binding of sugar oligomers probably requires a different structure of the substrate-binding pocket compared with regular sugar-binding proteins that predominantly bind monosaccharides. Also, in other archaea, members of the oligo/dipeptide cluster appear to be involved in sugar transport. Recently, we identified a cellobiose transport system in the hyperthermophile *Pyrococcus furiosus.* This system also belongs to the oligo/dipeptide transporter family (S. M. Koning, personal communication). In *Pyrococcus horikoshii*, genes encoding for the subunits of a putative oligo/dipeptide transporter were found to be associated with the sugar-degrading enzyme, β-galactosidase (Kawarabayasi *et al.*, 1998). The complete genome sequence of the thermophilic bacterium *Thermotoga maritima* contains 11 members of the oligo/dipeptide transporter family, nine of which are present in operons that also contain genes involved in sugar metabolism (Nelson *et al.*, 1999). It was postulated that peptide uptake and sugar degradation are co-ordinately regulated. Similar observations were made for *Thermoplasma acidophilum* for which the genome sequence was recently completed (Ruepp *et al.*, 2000). Based on our findings, it seems more likely that most of these systems are involved in sugar uptake and not peptide transport. A clear example of how these sugar transport operons are surrounded by genes encoding proteins involved in sugar metabolism is shown in Fig. 7C, in which the maltose and oligo 1 operons of *S. solfataricus* are located in the vicinity of a gene encoding a maltase (MalA) and α-fucosidase respectively (Fig. 7C). The *malA* gene (contig 025 ORF 031 in the *S. solfataricus* P2 database) and adjacent regions have been characterized by Rolfsmeier *et al.* (1998). Interestingly, a 2.4 kb transcript upstream of *malA*, termed ORF1, was found to be transcribed upon growth on maltose. The gene product of ORF1 is identical to the protein MalE, which we identified as the maltose-binding protein in this study.

Binding proteins of the sugar and oligo/dipeptide cluster differ in their size and type of signal sequence. All binding proteins of the sugar cluster are equipped with an unusual short signal sequence (Fig. 6A) (Albers *et al.*, 1999b). This signal sequence resembles that of the flagellins of methanogenic and halophilic archaea (Kalmokoff and Jarrell, 1991) and of bacterial proteins of the type IV pilin superfamily (Faguy *et al.*, 1994). Although there are some striking similarities, the signal sequences of archaeal preflagellins and binding proteins also differ in some respects (Jarrell *et al.*, 1999). Archaeal preflagellins always harbour a glycine at the −1 position, whereas in the binding protein sequences, the −1 position may contain a glycine or alanine residue. The identity of the archaeal processing peptidase(s) is not known, nor is it clear whether the same cytosolic enzyme processes both preflagellins and binding proteins. Based on the availability of sequenced binding proteins, we now define the archaeal consensus processing sites as $[KR]_{-2}$ $[GA]_{-1}$ $[IL]_{+1}$. For the preflagellins, processing at the N-terminus is necessary for release from the membrane and assembly into a flagellum. With the binding proteins, there is no evidence that they are released from the membrane as well. On the contrary, the hydrophobic stretch of amino acids located directly behind the signal sequence presumably serves to anchor the binding proteins to the membrane, as release of binding activity from the membrane requires detergents. This transmembrane segment is followed by a long stretch of hydroxylated amino acids, i.e. serine and threonine residues. Binding proteins of the oligo/dipeptide cluster are all equipped with a typical bacterial or eukaryotic signal sequence (von Heijne, 1990) (Fig. 6B). They are processed at the outer surface of the membrane, but remain membrane anchored by means of a transmembrane segment present in the carboxyl-terminal region. Strikingly, this transmembrane segment is preceded by a long stretch of hydroxylated amino acids. This region may function as a flexible linker region to couple the catalytic substrate-binding domain to the membrane-anchoring domain. Alternatively, it may represent a site of glycosylation and anchoring to the S-layer. Various extracellular archaeal proteins, such as the pullulanase from *Thermococcus hydrothermalis* (Erra-Pujada *et al.*, 1999) and the S-layer proteins from *Haloferax halobium* and *H. volcanii* (Lechner and Sumper, 1987; Sumper *et al.*, 1990), are equipped with similar types of linker and hydrophobic domains. The linker region of the *H. halobium* S-layer protein was shown to be highly O-glycosylated with galactose and glucose residues (Lechner and Sumper, 1987).

Most of the sugar-binding proteins in *S. solfataricus* are expressed under conditions in which the respective sugar is present in the growth medium (Figs 1, 3). GlcS and TreS appear to be expressed constitutively. As synthesis of α-amylase is repressed by glucose, a catabolite

repression-like system has been proposed for *S. solfataricus* (Haseltine *et al.*, 1996). Carbon regulation of the expression of α-amylase, α-glucosidase and β-glucosidase has been studied in some detail (Haseltine *et al.*, 1999), and a 'global gene regulatory system' was proposed that acts at the level of mRNA abundance to co-ordinate the levels of the three enzymes in response to changes in the nutrient composition of the growth medium. Expression of the binding proteins might be under the control of the same regulatory mechanisms. The glucose uptake system, however, is constitutively expressed, which indicates that it might not only function as a primary uptake system for its substrates from the medium. The system may also be involved in the recruitment of sugars released upon the degradation of extracellular glycoproteins. The sugars used for glycosylation, i.e. galactose and glucose (Lechner and Sumper, 1987) and mannose (this study), are all substrates for the glucose uptake system.

Our data provide biochemical evidence that *S. solfataricus* predominantly uses binding protein-dependent ABC transporters for the uptake of sugars. This is a remarkable observation, as this organism maintains a large ΔpH across its membrane but nevertheless prefers to accumulate sugars at the expense of ATP instead of a proton motive force. The completed genome sequences of other hyperthermophilic organisms, such as *T. maritima* and *P. furiosus*, reveal a large abundance of binding protein-dependent ABC transporters relative to permeases (Nelson *et al.*, 1999; http://wit.mcs.anl.gov). The available genomic information for *S. solfataricus* (about 75% of the genome) also points in this direction. The preference for binding protein-dependent ABC transporters might relate to the high substrate-binding affinity that is usually associated with these systems (K_D values for glucose and arabinose binding are 430 and 130 nM respectively). This may be an essential requirement for these organisms and enable them to scavenge all available nutrients in a substrate-poor environment such as hydrothermal vents in the deep sea or hot sulphuric pools.

Experimental procedures

Culture conditions

Sulfolobus solfataricus P2 was obtained from the Deutsche Sammlung von Mikroorganismen und Zellkultur (Braunschweig, Germany). Cells were grown aerobically at 80°C in the medium described by Brock *et al.* (1972), adjusted to pH 3 with sulphuric acid and supplemented with 0.4% (w/v) of trypton or different sugars as sole carbon and energy source.

Preparation of membranes of S. solfataricus

Cells were suspended in 20 mM Bis-Tris propane (pH 6.5)

and passed twice through a French pressure cell at 800 lb in^{-2}. Unbroken cells were removed by low-spin centrifugation at 3000 *g* for 10 min at 4°C. Membranes were collected by high-spin centrifugation at 100 000 *g* for 45 min at 4°C. The membrane pellet was suspended in 20 mM Bis-Tris propane (pH 6.5) and stored in liquid nitrogen.

Partial purification of the binding proteins

Membranes were solubilized by incubation for 30 min at 37°C in the presence of 1% Triton X-100 at a protein concentration of 10 mg ml^{-1} in Bis-Tris propane buffer (pH 6.5). Insoluble material was removed by centrifugation for 15 min at 250 000 *g* in a desktop ultracentrifuge (Beckman Coulter). The supernatant was supplemented with NaCl to a final concentration of 0.5 M and applied to a ConA–Sepharose column (Pharmacia) equilibrated with buffer A (20 mM Bis-Tris propane, pH 6.5, 0.5 M NaCl and 0.05% Triton X-100). The column was washed with five volumes of the same buffer, and bound glycoproteins were eluted by 200 mM α-methylmannopyranoside in buffer A. The protein-containing fractions were pooled and dialysed three times against a 500-fold volume of buffer A without NaCl to remove the salt and α-methylmannopyranoside. The resulting ConA fractions were used for the different binding studies. For subsequent purification, the ConA fraction was applied to an HR 5/5 Mono Q column (Pharmacia) equilibrated with buffer A without NaCl. Proteins were eluted with a linear gradient of 0–500 mM NaCl. Fractions were assayed for the different binding activities, analysed by SDS–PAGE and stained with Coomassie brilliant blue.

Binding assays

Binding of radiolabelled substrates to solubilized membranes or partially purified proteins was measured by a filtration assay. A volume of 50 μl of the ConA fraction was mixed with 50 μl of 0.5 M glycine–HCl buffer (pH 2) and preincubated for 2 min at 60°C. The radiolabelled sugar was added to a final concentration of 1 μM and incubated for 2 min at 60°C. The reaction was stopped by the addition of 2 ml of a chilled 0.1 M LiCl solution, filtered through 0.45 μm pore size nitrocellulose filters (Schleicher and Schüll; BA85) and washed once with 2 ml of the same solution. Filters were dissolved in 2 ml of scintillation fluid (FLO-scint A' Packard Instruments), and the radioactivity was determined. The following radiolabelled compounds were used: D-[1-^{3}H]-arabinose (specific activity 20 Ci mmol^{-1}) was from (American Radiolabeled Chemicals); D-[U-^{14}C]-glucose (291 mCi mmol^{-1}); [U-^{14}C]-maltose (630 mCi mmol^{-1}); D,L-[U-^{14}C]-fructose (340 mCi mmol^{-1}); D-[1-^{14}C]-lactose (57 mCi mmol^{-1}); [U-^{14}C]-xylose (55 mCi mmol^{-1}); [U-^{14}C]-sucrose (540 mCi mmol^{-1}); [^{3}H]-galactose (13700 mCi mmol^{-1}); [^{3}H]-cellobiose (8900 mCi mmol^{-1}); [^{3}H]-melibiose (3.2 Ci mmol^{-1}); [U-^{14}C]-citrate (114 mCi mmol^{-1}); L-[U-^{14}C]-alanine (174 mCi mmol^{-1}); L-[U-^{14}C]-leucine (311 mCi mmol^{-1}); L-[U-^{14}C]-glutamate (251 mCi mmol^{-1}) (all from Amersham); [U-^{14}C]-trehalose (850 mCi mmol^{-1}) was a gift from Dr Boos, University of Konstanz, Germany.

Database searches

The N-terminal sequences were compared with the sequences of proteins described in the *S. solfataricus* P2 database (http://niji.imb.nrc.ca/sulfolobus, Sensen *et al.*, 1998). After identification of the binding proteins, additional BLAST searches were performed to identify the adjacent genes and the organization of the complete operon encoding the ABC transporters.

Direct links to the contigs of the *S. solfataricus* P2 database are: c42 (http://niji.imb.nrc.ca/sulfolobus/sh01e 0442/sh01e0442_home.html); c25 (http://niji.imb.nrc.ca/sul-folobus/sh13a0425/sh13a0425_home.html); lam11 (http://niji. imb.nrc.ca/sulfolobus/l39−2_l11/l39−2_l11_home.html); c19 (http://niji.imb.nrc.ca/sulfolobus/sh22e0419/sh22e0419_home. html); c63 (http://niji.imb.nrc.ca/sulfolobus/sh17g1063/sh17g 1063_home.html); c30 (http://niji.imb.nrc.ca/sulfolobus/sh19h 1230/sh19h1230_home.html); c29 (http://niji.imb.nrc.ca/sulfo-lobus/sh13a0629/sh13a0629_home.html); c99 (http://niji.imb. nrc.ca/sulfolobus/sh01e0999/sh01e0999_home.html).

Sugar analysis

For total sugar analysis, AraS, eluted from SDS–PAGE, and the complete arabinose ConA fraction were subjected to a hot acid hydrolysis treatment [0.1 N trifluoroacetic acid at 80°C for 1 h, or 2 N trifluoroacetic acid at 100°C for 5 h or 4 N HCl at 100°C for 3 h (this sample was re-N-acetylated)]. Released monosaccharides were labelled with 2-aminoacri-done (Molecular Probes), separated by acrylamide gel electrophoresis (Jackson, 1994) and compared with known standards.

The type of glycosylation was determined after blotting of the ConA fraction onto PVDF membranes and incubation with different lectins present in the digoxigenin glycan differentiation kit (Roche Molecular Biochemicals).

Other techniques

Transport studies were performed as described previously (Albers *et al.*, 1999a). *S. solfataricus* cells used were grown in the presence of 0.4% (w/v) glucose or maltose as the sole carbon source. Sequence alignment was performed with CLUSTALW (http://www.ch.embnet.org/software/ClustalW.html). Protein concentrations were determined with the DC kit (Bio-Rad). N-terminal sequencing of the binding proteins blotted on PVDF membranes was performed by the Protein Service Laboratory (University of British Columbia, Vancouver, Canada).

References

Albers, S.V., Elferink, M.G.L., Charlebois, R.L., Sensen, C.W., Driessen, A.J.M., and Konings, W.N. (1999a) Glucose transport in the extremely thermoacidophilic *Sulfolobus solfataricus* involves a high-affinity membrane-integrated binding protein. *J Bacteriol* **181**: 4285–4291.

Albers, S.V., Konings, W.N., and Driessen, A.J. (1999b) A unique short signal sequence in membrane-anchored proteins of Archaea. *Mol Microbiol* **31**: 1595–1596.

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 Microbiol* **84**: 54–68.

Cubellis, M.V., Rozzo, C., Montecucchi, P., and Rossi, M. (1990) Isolation and sequencing of a new β-galactosidase-encoding Archaebacterial gene. *Gene* **94**: 89–94.

D'Auria, S., Morana, A., Febbraio, F., Vaccaro, C., De Rosa, M., and Nucci, R. (1996) Functional and structural proper-ties of the homogeneous beta- glycosidase from the extreme thermoacidophilic archaeon *Sulfolobus solfatar-icus* expressed in *Saccharomyces cerevisiae. Protein Expr Purif* **7**: 299–308.

Erra-Pujada, M., Debeire, P., Duchiron, F., and O'Donohue, M.J. (1999) The type II pullulanase of *Thermococcus hydrothermalis*: molecular characterization of the gene and expression of the catalytic domain. *J Bacteriol* **181**: 3284–3287.

Faguy, D.M., Jarrell, K.F., Kuzio, J., and Kalmokoff, M.L. (1994) Molecular analysis of archaeal flagellins: similarity to the type IV pilin-transport superfamily widespread in bacteria. *Can J Microbiol* **40**: 67–71.

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

Haseltine, C., Rolfsmeier, M., and Blum, P. (1996) The glucose effect and regulation of α-amylase synthesis in the hyperthermophilic archaeon *Sulfolobus solfataricus. J Bacteriol* **178**: 945–950.

Haseltine, C., Montalvo-Rodriguez, R., Bini, E., Carl, A., and Blum, P. (1999) Coordinate transcriptional control in the hyperthermophilic archaeon *Sulfolobus solfataricus. J Bacteriol* **181**: 3920–3927.

von Heijne, G. (1990) The signal peptide. *J Membr Biol* **115**: 195–201.

Higgins, C.F. (1992) ABC transporters: from microorganisms to man. *Annu Rev Cell Biol* **8**: 67–113.

Horlacher, R., Xavier, K.B., Santos, H., DiRuggiero, J., Kossmann, M., and Boos, W. (1998) Archaeal binding protein-dependent ABC transporter: molecular and bio-chemical analysis of the trehalose/maltose transport system of the hyperthermophilic archaeon *Thermococcus litoralis. J Bacteriol* **180**: 680–689.

Jackson, P. (1994) High-resolution polyacrylamide gel electrophoresis of fluorophore-labeled reducing sacchar-ides. *Methods Enzymol* **230**: 250–265.

Jarrell, K.F., Correia, J.D., and Thomas, N.A. (1999) Is the processing and translocation system used by flagellins also used by membrane-anchored secretory proteins in archaea? *Mol Microbiol* **34**: 395–398.

Kalmokoff, M.L., and Jarrell, K.F. (1991) Cloning and

sequencing of a multigene family encoding the flagellins of *Methanococcus voltae*. *J Bacteriol* **173**: 7113–7125.

Kato, M., Miura, Y., Kettoku, M., Shindo, K., Iwamatsu, A., and Kobayashi, K. (1996) Purification and characterization of new trehalose-producing enzymes isolated from the hyperthermophilic archaea, *Sulfolobus solfataricus* KM1. *Biosci Biotechnol Biochem* **60**: 546–550.

Kawarabayasi, Y., Sawada, M., Horikawa, H., Haikawa, Y., Hino, Y., Yamamoto, S., *et al.* (1998) Complete sequence and gene organization of the genome of a hyper-thermophilic archaebacterium, *Pyrococcus horikoshii* OT3. *DNA Res* **5**: 55–76.

Kengen, S.W.M., Stams. A.J.M., and de Vos, W.M. (1996) Sugar metabolism of hyperthermophiles. *FEMS Microbiol Rev* **18**: 119–137.

Kobayashi, K., Kato, M., Miura, Y., Kettoku, M., Komeda, T., and Iwamatsu, A. (1996) Gene cloning and expression of new trehalose-producing enzymes from the hyperthermophilic archaeum *Sulfolobus solfataricus* KM1. *Biosci Biotechnol Biochem* **60**: 1882–1885.

Lechner, J., and Sumper, M. (1987) The primary structure of a procaryotic glycoprotein. Cloning and sequencing of the cell surface glycoprotein gene of halobacteria. *J Biol Chem* **262**: 9724–9729.

Nelson, K.E., Clayton, R.A., Gill, S.R., Gwinn, M.L., Dodson, R.J., Haft, D.H., *et al.* (1999) Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of *Thermotoga maritima*. *Nature* **399**: 323–329.

Quiocho, F.A., and Vyas, N.K. (1984) Novel stereospecificity of the L-arabinose-binding protein. *Nature* **310**: 381–386.

Rolfsmeier, M., and Blum, P. (1995) Purification and characterization of a maltase from the extremely thermophilic crenarchaeote *Sulfolobus solfataricus*. *J Bacteriol* **177**: 482–485.

Rolfsmeier, M., Haseltine, C., Bini, E., Clark, A., and Blum, P.

(1998) Molecular characterization of the alpha-glucosidase gene (malA) from the hyperthermophilic archaeon *Sulfolobus solfataricus*. *J Bacteriol* **180**: 1287–1295.

Ruepp, A., Graml, W., Santos-Martinez, M.L., Koretke, K.K., Volker, C., Mewes, H.W., *et al.* (2000) The genome sequence of the thermoacidophilic scavenger *Thermplasma acidophilum*. *Nature* **407**: 508–513.

Schonheit, P., and Schafer, T. (1995) Metabolism of hyperthermophiles. *World J Microbiol Biotechnol* **11**: 26–57.

Sensen, C.W., Charlebois, R.L., Chow, C., Clausen, I.G., Curtis, B., Doolittle, W.F., *et al.* (1998) Completing the sequence of the *Sulfolobus solfataricus* P2 genome. *Extremophiles* **2**: 305–312.

Sumper, M., Berg, E., Mengele, R., and Strobel, I. (1990) Primary structure and glycosylation of the S-layer protein of *Haloferax volcanii*. *J Bacteriol* **172**: 7111–7118.

Wanner, C., and Soppa, J. (1999) Genetic identification of three ABC transporters as essential elements for nitrate respiration in *Haloferax volcanii*. *Genetics* **152**: 1417–1428.

Wassenberg, D., Liebl, W., and Jaenicke, R. (2000) Maltose-binding protein from the hyperthermophilic bacterium *Thermotoga maritima*: stability and binding properties. *J Mol Biol* **295**: 279–288.

Xavier, K.B., Martins, L.O., Peist, R., Kossmann, M., Boos, W., and Santos, H. (1996) High-affinity maltose/trehalose transport system in the hyperthermophilic Archaeon *Thermococcus litoralis*. *J Bacteriol* **178**: 4773–4777.

Zahringer, U., Moll, H., Hettmann, T., Knirel, Y.A., and Schafer, G. (2000) Cytochrome b558/566 from the archaeon *Sulfolobus acidocaldarius* has a unique Asn-linked highly branched hexasaccharide chain containing 6-sulfoquinovose. *Eur J Biochem* **267**: 4144–4149.