Journal of Bioenergetics and Biomembranes, Vol. 36, No. 1, February 2004 (© 2004)

Insights Into ABC Transport in Archaea

Sonja-Verena Albers,[1] Sonja M. Koning,[1,2] Wil N. Konings,[1] and Arnold J. M. Driessen[1,3]

In archaea, ATP-binding cassette (ABC) transporters play a crucial role in substrate uptake, export, and osmoregulation. Archaeal substrate-binding-protein-dependent ABC transporters are equipped with a very high affinity for their cognate substrates which provide these organisms with the ability to efficiently scavenge substrates from their environment even when present only at low concentration. Further adaptations to the archaeal way of life are especially found in the domain organization and anchoring of the substrate-binding proteins to the membrane. Examination of the signal peptides of binding proteins of 14 archaeal genomes showed clear differences between euryarchaeotes and crenarchaeotes. Furthermore, a profiling and comparison of ABC transporters in the three sequenced pyrococcal strains was performed.

KEY WORDS: Archaea; ABC transport; substrate-binding protein; signal peptide; *Sulfolobus*; *Pyrococcus*.

INTRODUCTION

ATP-binding cassette (ABC) transporters are an important class of transport proteins, which are widely distributed in all domains of life and are involved in varying processes such as substrate uptake or export, osmosensing and osmoregulation, and antigen processing (Holland and Blight, 1999). In humans, mutations in ABC transporter often result in severe diseases such as cystic fibrosis, adrenoleukodystrophy, while ABC transporters are also associated with multidrug resistance of tumor cells. The domain structure of these transporters comprises two integral membrane proteins and two cytoplasmically located ATPases, which drive the translocation of the substrate by the hydrolysis of ATP. In eukaryotes these four domains are often found in one polypeptide chain whereas in prokaryotes the subunits are often translated as single proteins. In prokaryotes the membrane proteins (permeases) and the ATPases can be present either as homo or heterodimers. Bacterial and archaeal ABC uptake systems comprise a fifth domain, the extracellular substrate-binding protein, which binds the substrate and delivers it to the permease domain (Fig. 1(A)).

In contrast to secondary transporters, which depend on the electrochemical gradient of protons or sodium ions across the cytoplasmic membrane, ABC transporter can accumulate substrates to much higher concentrations inside the cell. Moreover, binding-protein-dependent ABC transporters often have a very high affinity to their substrates, which is mainly achieved by the high binding affinity of the substrate-binding protein. In particular, archaeal ABC sugar transporters appear to be equipped with an exceptional high affinity for the substrate which is in the subnanomolar. This might especially be beneficial for these organisms as they often inhabit carbon source deprived ecological niches.

Here we provide an overview of the progress in the studies of archaeal ABC transporters as well as the comparative analysis of ABC transporters in three pyrococcal genomes. Furthermore, we present new data on the (putative) signal peptides of substrate-binding proteins of ABC transporters of 14 sequenced archaeal genomes.

BINDING-PROTEIN-DEPENDENT ABC TRANSPORTERS

Binding-protein-dependent ABC transporters are the best studied group of transporters in archaea. The

[1] Department of Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Kerklaan 30, 9751 NN Haren, The Netherlands.
[2] Present address: Department of Microbiology, University of Regensburg, Universitätsstr. 31, 93053 Regensburg, Germany.
[3] To whom correspondence should be addressed; e-mail: a.j.m.driessen@biol.rug.nl.

0145-479X/04/0200-0005/0 © 2004 Plenum Publishing Corporation

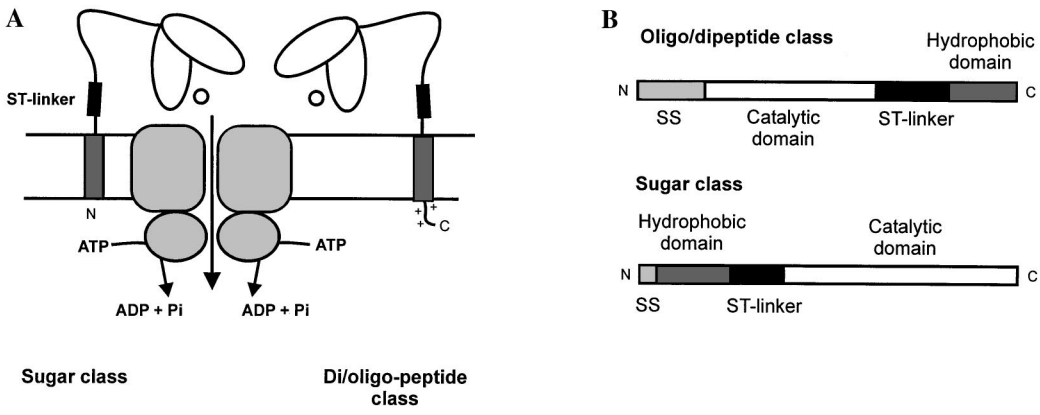

Fig. 1. Domain structure of archaeal ABC transporter. (A) Typical ABC transporter consisting of two permeases, two ATPases, and one membrane anchored substrate-binding protein. The substrate-binding proteins belong to either the sugar or the di/oligopeptide class which differ in membrane topology. (B) Cartoon of the domain organization of precursors of substrate-binding proteins of the sugar and di/oligopeptide class showing the signal peptide.

percentage of ABC transporters in relation to secondary transporters in archaea varies between 39% in *Thermoplasma volcanicum* and 79% in *Pyrobaculum aerophilum* (Paulsen *et al.*, 2000). The subunits of these transporters share a number of typical consensus sequence motifs with their bacterial counterparts. The membrane domains contain the EAAA×3G×9I×LP motif which has been implicated in the interaction with the ATPase subunit (Dassa and Hofnung, 1985). The ATPase subunits exhibit the typical Walker sequences, the Q-loop and the H-region (Higgings, 1992). Table I summarizes the uptake systems in archaea described sofar. Sugar uptake systems have extensively been studied in *Sulfolobus solfataricus*, *Pyrococcus furiosus*, and *Thermococus litoralis* (Albers *et al.*, 1999a,b; Elferink *et al.*, 2001; Horlacher *et al.*, 1998; Koning *et al.*, 2001, 2002; Xavier *et al.*, 1996). These studies concentrated on the identification of substrate, the properties of the substrate-binding proteins, their domain organization, and the nature of the signal peptides (see below).

Table I. Described Solute Transporters in Archaea

ABC transporter	Substrate	K_m for uptake (nM)	K_d for solute binding[a] (nM)	Reference
T. litoralis	Maltose/trehalose	22/17	160	Horlacher *et al.* (1998); Xavier *et al.* (1996)
S. solfataricus	Glucose	2000	480	Albers *et al.* (1999a,b)
	Cellobiose + cellooligomers	—[b]	—	Elferink *et al.* (2001)
	Trehalose	—	—	Elferink *et al.* (2001)
	Maltose/maltotriose	—	—	Elferink *et al.* (2001)
	Arabinose	—	130	Elferink *et al.* (2001)
P. furiosus	Cellobiose + cellooligomers	175	45	Koning *et al.* (2001)
	Maltose/trehalose			Koning et *al.* (2002)
	Maltotrios, maltodextrin	—	270	Evdokimov *et al.* (2001); Koning et *al.* (2002)
Haloferax volcanii	Glucose (anaerobic)	—	—	Wanner *et al.* (1999)
	Molybdate	—	—	Wanner *et al.* (1999)
	Inorganic anions	—	—	Wanner *et al.* (1999)
Methanococcus thermoautrophicum	Phosphate	25	—	Krueger *et al.* (1986)
M. mazei Gö1	Glycine betaine	—	—	Roessler *et al.* (2002)
M. portucalensis	Glycine betaine	23000	—	Lai *et al.* (2000)
M. thermophila TM-1	Glycine betaine	10000	—	Proctor *et al.* (1997)

[a] Solute binding to binding protein.
[b] Not determined.

In bacteria ABC uptake systems are divided into two main classes: the **c**arbohydrate **u**ptake **t**ransporters, the CUT class and the di/oligopeptide class (Schneider, 2001). Interestingly, archaeal sugar ABC transporters are found in both classes. Archaeal sugar transporters, which transport mono- or disaccharides (see Table I), indeed belong functionally and structurally to the CUT class. However, sugar ABC transporters involved in the uptake of disaccharide and oligosaccharides as the cellobiose/β-glucoside transport system of *P. furiosus* (Koning *et al.*, 2001) and the maltose/maltodextrin and the cellobiose/cello-oligomer transporters of *S. solfataricus* (Elferink *et al.*, 2001) show homologies to transporters in the di/oligopeptide class. As in bacterial di/oligopeptide uptake systems (Detmers *et al.*, 2000; Lanfermeijer *et al.*, 2000) the binding proteins can accommodate a broad range of substrates, which vary in length and composition from each other (Koning *et al.*, 2001). Interestingly, the genomes of hyperthermophilic archaea and bacteria contain a great number of transport operons assigned as di/oligopeptide transporter, which are in the vicinity of sugar-degrading enzymes (Nelson *et al.*, 1999). Taken together, it is likely that most of these transporters are not involved in di/oligopeptide transport but instead catalyze the uptake of di- and oligosaccharides.

A set of ABC transporters involved in osmoadaptation have been described for methanogens. These high-affinity transport systems for glycine betaine are activated upon osmotic upshock (Lai *et al.*, 2000; Proctor *et al.*, 1997; Roessler *et al.*, 2002). In *Methanosarcina mazei* Goe1 it was shown that the expression of the *otaA/B/C* gene cluster, which encodes a putative ABC transporter, was highly induced upon growth in medium containing high salt concentrations (Roessler *et al.*, 2002).

The trehalose/maltose ABC transporter of the hyperthermophilic *T. litoralis* has been studied in considerable detail. In early studies the trehalose-maltose-binding protein (TMBP) was identified, and transport was shown to occur with a very high affinity of 20 nM at 80°C (Xavier *et al.*, 1996). Later TMBP and MalK, the ATPase domain of this transporter, were crystallized (Diederichs *et al.*, 2000; Diez *et al.*, 2001; see below). Finally, the entire transporter was heterologously expressed in *E. coli* and purified (Greller *et al.*, 2001). The solubilized complex showed intrinsic ATPase activity, but the ATPase was not further stimulated by the addition of TMBP as observed for the bacterial system (Reich-Slotky *et al.*, 2000). The lack of stimulation might be explained by the difficulties to find detergents which are not precipitating at high temperatures and the absence of archaeal ether-lipids which have been shown to increase the stability of archaeal membrane proteins (Elferink *et al.*, 1992).

Binding Protein

The component of archaeal ABC transporter, which shows the most pronounced difference in comparison with the bacterial counterparts are the substrate-binding proteins. As in the bacterial system the extracellular binding protein captures the substrate and after docking to the permeases, the substrate is released and translocated across the membrane. The structure of substrate-binding proteins contains two lobs which are drawn nearer to each other in a "venus-flytrap-like fashion upon binding of the substrate (Quiocho and Ledvina, 1996). Archaeal substrate-binding proteins have particularly been studied in the hyperthermophiles *P. furiosus*, *T. litoralis*, and *S. solfataricus* (Tabel I). A common feature of all these binding proteins is their glycosylation. Glycosylation was believed to be absent in prokaryotes, but most studied extracellular proteins of archaea have been shown to be glycosylated (Erra-Pujada *et al.*, 1999; Greller *et al.*, 2001; Hettmann *et al.*, 1998; Sumper *et al.*, 1990). Binding proteins isolated from *P. furiosus* contain glucose moieties (Koning *et al.*, 2002), whereas mannose, glucose, galactose, and *N*-acetylglucosamine have been identified in binding proteins of *S. solfataricus* (Elferink *et al.*, 2001). Glycosylation is, however, not required for substrate binding as archaeal binding proteins heterologously expressed in *E.coli* are active in sugar binding (Horlacher *et al.*, 1998; Koning *et al.*, 2001, 2002). Glycosylation may stabilize the extracellular proteins against proteolytic degradation, influence the interaction of the binding proteins with the cell envelope, or affect their thermostability.

As described before the ABC transporters are divided into two main classes: the CUT class and the di/oligopeptide class. This is particularly evident when the domain organization of the binding proteins is examined (Fig. 1(A) and (B)). Binding proteins belonging to the CUT class exhibit an N-terminal signal peptide often followed by a long stretch of hydroxylated amino acids up to 60 residues long, which might act as a flexible linker region between the membrane and the binding domain. A similar site in the S-layer of *Halobacteria* also has been shown to be highly glycosylated (Sumper *et al.*, 1990). Binding proteins of archaea need to be attached to the cytoplasmic membrane as these organisms lack an outer membrane, whereas the S-layer is permeable for small proteins. It is believed that binding proteins of the CUT class are lipidated after translocation across the membrane in analogy to binding proteins of Gram positive bacteria. A conserved cysteine residue directly following the signal sequence cleavage site confirms this idea (see below). Members of the CUT class of *S. solfataricus* constitute a special case: they do not contain

secretory signal peptides, but type IV pili like signal peptides (Albers *et al.*, 1999a,b; Elferink *et al.*, 2001). These signal peptides are normally only found in pilin subunits of bacteria or archaeal flagella. Until now it remains unclear whether these binding proteins assemble to a macromolecular structure like pili or flagella. Nor is it clear what the function of such a structure would be.

The binding proteins of the di/oligopeptide class contain a normal bacterial like signal peptide and, in contrast to the binding proteins of the CUT class, a hydrophobic domain at the extreme C-terminal end, which is preceded by a stretch of hydroxylated amino acids. Therefore, it is assumed that the catalytic domain is membrane anchored by a carboxyl-terminal transmembrane domain. We analyzed the 14 archaeal genomes of the transport database (http://66.93.129.133/transporter/wb/index2.html) (Paulsen *et al.*, 2000) and found that most of the archaeal substrate-binding proteins belong to the CUT class (Table II). *S. solfataricus* shows the highest amount of di/oligopeptide class binding proteins, whereas in methanogens this class is absent.

In addition to the function in uptake of substrates it has recently been shown that some archaeal substrate-binding proteins function as receptors in chemotaxis (Koekova *et al.*, 2002). This is in contrast to bacterial systems, where the primary role is always the function in uptake. For example, the maltose-binding protein of the maltose ABC transporter is mainly involved in maltose uptake but also in sensing the extracellular concentration of maltose in interaction with the chemotaxis transducer Tar (Stock and Surette, 1996). In general, genes encoding substrate-binding proteins are always found in an operon together with the other ABC transporter subunits or in the close vicinity of such genes. In halobacteria, the described substrate-binding proteins BasB (branched amino acids) and CosB (compatible solutes) are found in transcriptional units with their cognate transducer protein (Koekova *et al.*, 2002).

Signal Peptides

Substrate-binding proteins are extracellular proteins that have to be secreted across the cytoplasmic membrane. N-terminal signal peptides target the binding proteins to the membrane for translocation. After translocation, the signal peptide is removed by a signal peptidase. In Gram negative bacteria substrate-binding proteins have a typical secretory signal peptide with a consensus sequence around the cleavage site, which is recognized by the type I signal peptidase. These signal peptides usually have a positively charged N-terminus, a stretch of hydrophobic amino acids, which is followed by the cleavage site which contains small polar amino acids at positions −1 and −3. Substrate-binding proteins of Gram positive bacteria often contain a so-called "lipobox" motif ([I/L/G/A]-[A/G/S]-C) around the cleavage site. This cysteine is lipidated prior to the removal of the signal peptide by a type II signal peptidase. The lipid moiety ensures anchoring of the protein to the membrane.

As described above the archaeal substrate-binding proteins fall into two classes: the sugar class and the

Table II. Distribution of Sugar Class and Di/Oligopeptide Binding Proteins in Various Archaea

Organism	Total number of binding proteins	Sugar class	Oligopeptide class	Ungrouped	Remarks
A. pernix	17	15	2		Sugar class: two type IV
A. fulgidus	13	12	1		
Halobacterium sp. NRC-1	13	12		1 (very short one/two hydrophobic parts)	
M. acetivorans	51	50		1 (very short one/two hydrophobic domains)	
M. kandlerii	3	3	0		
M. jannashii	3	3	0		
M. mazei Goe1	17	16		1 (very short one/two hydrophobic parts)	
P. aerophilum	22	20	2		
P. abyssii	15	11	4		Sugar class: CGI motif at C-terminus
P. horikoshii	13	10	3		Sugar class: CGI motif at C-terminus
S. solfataricus	15	6	8	1	Sugar class: all type IV pili signal peptides
S. tokodaii	6	1 (?)	3	2	
T. volcanicum	5	5	0		Sugar class: 3 type IV-like peptides
T. acidophilum	7	4	2	1	

di/oligopeptide class. These classes differ both in domain organization and the type of signal peptides. As the amino-terminal amino acid sequences of only a few mature secreted archaeal proteins have been determined experimentally, we used the transport class database (http://66.93.129.133/transporter/wb/index2.html) to examine the signal peptides of substrate-binding proteins of 14 archaeal genomes (Table II and Fig. 2). Most of the substrate-binding proteins belong to the sugar class (Table II) implying that these proteins need some means of membrane anchoring upon the cleavage of the signal peptide. Analysis of the halocyanin of the archaeaon *Natronobacter pharaonis* by mass spectroscopy indicated that the N-terminal cysteine of halocyanin is covalently modified by C_{20} diphytanyl diether lipids (Mattar *et al.*, 1994). However, archaeal genomes do not contain homologues of the bacterial lipoprotein signal peptidase. They presumably contain a paralogue with a distinct amino acid sequence. One reason for a new type of enzyme might be the difference of the lipid substrate as the linkage of a C_{20} diphytanlyl diether lipid to a cysteine. This probably involves an entirely distinct enzymatic reaction than the attachment of a fatty acid. Nevertheless, most signal peptides of the examined euryarchaeal substrate-binding proteins contain a clear "lipobox" motif with a cysteine that is preceded by small hydrophobic amino acids (Fig. 2). *Halobacteria* present an interesting case: it has been proposed that probably nearly all extracellular proteins are secreted by the twin arginine (Tat) translocation pathway (Bolhuis, 2002; Rose *et al.*, 2002). Proteins recognized by the Tat translocation pathway contain a double arginine at their extreme N-termini. These proteins are translocated in a folded state (Berks *et al.*, 2000). It is believed that halobacterial proteins already have to fold in the cytoplasm to circumvent precipitation caused by the high internal salt concentration. The halobacterial signal peptides shown in Fig. 2 contain indeed the Tat-motif and in addition after the hydrophobic stretch the lipobox for the attachment of the lipid modification. The cleavage site of the signal peptide of the two *Thermoplasma* strains and *Methanopyrus kandlerii* seem to differ quite substantially from the other euryarchaeota. In *Methanopyrus kandelerii* the hydrophobic part is directly followed by a stretch of three tyrosine residues. This is less pronounced found in *T. acidophilum* and *T. volcanicum*. The signal peptides of the sugar class in the group of analyzed crenarchaeota seem to be even more variable. In *Pyrobaculum aerophilum* the region after the hydrophobic stretch is enriched in glutamine residues, which have been shown in *S. solfataricus* to be at the +1 site of the cleavage site of two binding protein of the di/oligopeptide class (Fig. 2; Albers and Driessen, 2002; Elferink *et al.*, 2001). One of the signal peptides of

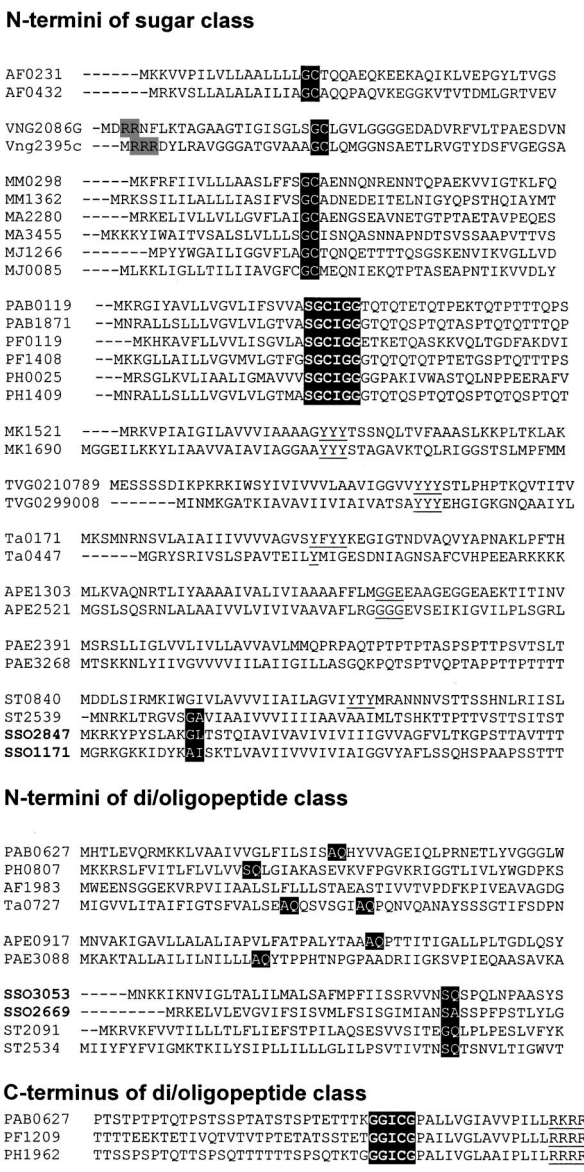

Fig. 2. Different signal peptides of archaeal substrate-binding proteins. The predicted cleavage sites are indicated by white letters on black background. Signal peptides, which have been determined experimentally, are shown in bold. APE, *Aeropyrum pernix*; AF, *Archaeoglobus fulgidus*; VNG, *Halobacterium* sp. NRC-1; MJ, *Methanococcus jannashii*; MA, *Methanosarchina acetovirans*; MM, *Methanosarcina mazei*; MK, *Methanopyrus kandlerii*; PAE, *Pyrobaculum aerophilum*; PAB, *Pyrococcus abyssii*; PF, *Pyrococcus furiosus*; PH, *Pyrococcus horikoshii*; SSO, *Sulfolobus solfataricus*; ST, *Sulfolobus tokodaii*; TA, *Thermoplasma acidophilum*; TVG, *Thermoplasma volcanicum*.

S. tokodaii appears similar to the one found in *M. kandlerii*, whereas all signal peptides of binding proteins from the sugar class of *S. solfataricus* are homologous to type IV prepilin signal peptides (Albers *et al.*, 1999a,b; Elferink

et al., 2001). In vitro assays showed that the binding proteins and the flagellin are indeed processed by a type IV pilin signal peptidase, PibD (Albers *et al.*, 2003). It is unclear why these signal peptides, which otherwise appear only in archaeal flagellins, are so frequently used on extracellular proteins of *S. solfataricus*.

Taken together it seems that euryarchaeal precursors of the sugar-binding proteins belonging to the sugar class are processed by a lipoprotein peptidase whereas in the crenarchaeota it remains uncertain how these proteins are attached to the membrane. Since detergents are required to extract these proteins from the membrane, it appears that they are membrane integral (unpublished results).

As described above the di/oligopeptide class of binding proteins is mainly present in *S. solfataricus* and the *Pyrococcus* strains. As these proteins contain a C-terminal transmembrane domain to anchor them to the membrane they do not need to be lipid modified at their N-termini. So, it is not surprising that the signal peptides do not contain the lipobox motif (Fig. 2), although it is difficult to predict a consensus sequence for the type I signal peptidase. Four extracellular proteins, the S-layer protein and three substrate-binding proteins were N-terminally sequenced (Albers and Driessen, 2002; Elferink *et al.*, 2001). Two of the resulting signal peptide cleavage sites are shown in Fig. 2. The amino acid residues were found to be either alanine or serine at the -1 position and either valine, glutamine, or alanine at the $+1$ position. According to these results cleavage sites are indicated in the N-termini of the other proteins of the di/oligopeptide class (Fig. 2).

Interestingly, the pyrococcal-binding proteins contain a GGICG sequence motif just preceding the C-terminal hydrophobic region (Fig. 2). The presence of lipid modifications close to the C-terminus of glycosylated S-layer proteins has been described for some halophilic archaea (Kikuchi *et al.*, 1999; Konrad and Eichler, 2002). This might suggest that the pyrococcal binding proteins are lipid modified at their C-termini.

The ATP-Binding Protein

The ATP-binding subunit of ABC transporters energizes substrate translocation by the hydrolysis of ATP. Conformational changes in this subunit, presumably caused by the binding of ATP and not its hydrolysis (Verdon *et al.*, 2003), open the translocation pore and allow passage of the substrate. ATP-binding proteins of all domains of life are highly homologous to each other, in particular in the Walker A and B motifs (Walker *et al.*, 1982) and the ABC signature motif (Boos and Shuman, 1998), LSGGQ, which has been shown to have an important function in the dimerization process upon ATP binding (Moody *et al.*, 2002). Several archaeal ATP-binding proteins such as LolD (MJ0769) and LivG (MJ1267) from *Methanococcus jannashii* (Yuann *et al.*, 2001), MalK from the trehalose/maltose transporter of *T. litoralis* (Diederichs *et al.*, 2000) and GlcV from the glucose transporter of *S. solfataricus* (Verdon *et al.*, 2003) have been crystallized. The overall structure of the ATP-binding domain in all four proteins is very similar and comparable to bacterial and eukaryal ATP-binding domains. However, MalK and GlcV exhibit an additional C-terminal domain, showing predominantly β-sheets organized in an OB (oligonucleotide/oligosaccharide binding) fold like structure. This C-terminal extension is found in several bacterial and archaeal ATP-binding proteins. In MalK of *E. coli* this domain has been shown to directly interact with the positive activator of the *mal* operon, MalT (Panagiotidis *et al.*, 1998). In the MalK bound state, MalT cannot induce expression of the *mal* genes and MalK is directly involved in the regulation of this transporter. A MalT homologue has not been found in archaea. The trehalose/maltose operon of *T. litoralis* contains the so-called *trmB* gene that encodes a negative regulator of the maltose/trehalose system (Lee *et al.*, 2003). This would argue against a bacterial-like regulation system of the maltose system and would make the C-terminal extension of MalK obsolete. However, DNA microarray studies in *P. furiosus*, which contains the identical genomic region as *T. litoralis* (DiRuggiero *et al.*, 2000), showed that all the genes are upregulated upon growth on maltose while *trmB* and *malK* stay upregulated even upon growth on peptide sources (Schut *et al.*, 2003). This suggests that MalK also in the archaeal system is involved in the regulation of the expression of the *mal* operon, although this hypothesis awaits further experimental proof.

Genomic Survey of ABC Transporters in *Pyroccocus* Species

During the last few years the genomes of a large number of prokaryotes have been sequenced, among which three different *Pyrococcus* species (http://www.genoscope.cns.fr/) (Kawarabayasi *et al.*, 1998; Robb *et al.*, 2001). The availability of three *Pyrococcus* genomes, *P. abyssi*, *P. furiosus*, and *P. horikoshii*, provides the unique possibility for species comparison. The three genomes have been compared at the total genome level (Maeder *et al.*, 1999; Zivanovic *et al.*, 2002). However, specific information is lacking about the conservation of certain protein families within the species, in particular the solute transporters.

So far three different ABC transporters have been described in *P. furiosus*, all of which are involved in organic solute uptake, namely a cellobiose/β-glucoside, a trehalose/maltose and a maltodextrin ABC-transporter (Koning *et al.*, 2001, 2002). Genome analysis suggests that there are also systems involved in transport of peptides, inorganic solutes, and drugs. Since in none of the other *Pyrococcus* species, transport studies have been conducted, a genomic survey was performed to provide more insight in the presence of different ABC transporters in the three complete *Pyrococcus* genomes. Although *P. abyssi*, *P. furiosus*, and *P. horikoshii* are closely related, the organisms differ in the total number of putative ABC transporters, which is 18 for *P. abyssi*, 19 for *P. furiosus*, and 17 for *P. horikoshii*. The genes encoding the different components of the putative ABC transporter are usually organized within gene clusters. In these studies, a set of genes is called a putative ABC transporter when the genes encoding the different components are in close vicinity to each other. About half of the putative ABC transporters is predicted to be involved in organic solute uptake. Table III shows an overview of the different ABC transport clusters present in the three completed pyrococcal genomes.

Only *P. furiosus* of the three sequenced species has been reported to be able to grow on a limited number of carbohydrates. However, all three species contain members of the ABC carbohydrate transport family (Table III). It is not clear why *P. abyssi* and *P. horikoshii* are not able to utilize carbohydrates as growth substrates. Both organisms contain orthologs of all genes encoding enzymes involved in the Embden–Meyerhof pathway as found in *P. furiosus*. *P. abyssi* contains a homolog of the cellobiose/β-glucoside transporter of *P. furiosus*. When *P. abyssi* is grown in the presence of cellobiose, this gene cluster is not induced in contrast to *P. furiosus* (Koning *et al.*, 2001). The inability of *P. abyssi* and *P. horikoshii* to utilize carbohydrates may therefore reside in a deficiency to induce the genes encoding the required enzymes.

Several transporters seem to be involved in the transport of ions (Table III). *P. furiosus* contains two gene clusters that belong to the iron/thiamin transport family, while the other two pyrococci harbor only a single member of this family. One of these systems in *P. furiosus* is present on a 16-kb fragment that is flanked by IS elements. On this fragment also the characterized trehalose/maltose transport operon is located (DiRuggiero *et al.*, 2000). This putative iron/thiamin transporter is highly homologous to the system found in *P. abyssi* and *P. horikoshii* and the second system in *P. furiosus*. However, the putative iron/thiamin transporter present on the 16-kb fragment is encoded by four genes that specify a binding protein, two transmembrane domains, and one nucleotide binding domain. The other putative transport cluster, present in all three species, contains a single permease gene. Although the two transport clusters show high sequence homology, it seems unlikely that the transporter present on the 16-kb genomic fragment has evolved from the common transporter by a gene duplication event.

In *P. abyssi* and *P. furiosus* only a single gene cluster seems to be involved in multidrug resistance, while *P. horikoshii* contains two transport clusters that relate to this activity. Only *P. endeavori* and *P. glycovorans* have been tested for sensitivity to antibiotics. Both species are insensitive to a diverse range of antibiotics (Barbier *et al.*, 1999; Pledger and Baross, 1991), and this has been attributed to the high impermeable archaeal cell envelope structure which might prevent entry of these antibiotics into the cell. In contrast, thermophilic bacteria are sensitive toward the same range of antibiotics (Barbier *et al.*, 1999; Pledger and Baross, 1991). The presence of the multidrug transporters, however, now suggests that the high resistance could be caused by the ability to expel these compounds from the cell.

COMPARISONS OF DI/OLIGOPEPTIDE BINDING PROTEIN FAMILY

One of the four *P. furiosus* ORFs encoding a binding protein of the di/oligopeptide binding protein family, PF1209, has been characterized previously as a cellobiose/β-glucoside binding protein (Koning *et al.*, 2001). Although this binding protein is a member of the di/oligopeptide binding protein family, it recognizes several carbohydrates of the β-glucoside class. *P. abyssi* contains a homolog of PF1209, i.e., PAB0627 (72% identity, 79% homology), while a similar gene is missing in *P. horikoshii*. The gene encoding PAB0627 is, however, not induced under conditions where PF1209 is expressed in *P. furiosus* (Koning *et al.*, 2001). The ORF encoding the putative binding protein PF0357 shows the highest amino acid homology to PF1209 (38% identity, 54% homology). As observed for the cellobiose/β-glucoside transporter where a sugar utilization gene (i.e., a β-mannosidase) is present upstream of the transporter gene cluster, a putative β-galactosidase is located upstream of PF0357. Therefore, PF0357 might also encode a carbohydrate binding protein. Likewise, *P. abyssi* and *P. horikoshii* contain a homolog of PF0357 within its vicinity genes encoding sugar metabolizing proteins. The function of these binding proteins is, however, not known.

The binding proteins PF0190 and PF1408 show only weak homology to PF1209 and PF0357, although these are members of the same binding protein family. PF0190

Table III. Overview of and Homologies Between the Predicted ABC Transport Clusters in the Three *Pyrococcus* Species

	P. furiosus					*P. abyssi*					*P. horikoshii*				
Family[a]	Binding protein	Permease(s)		ATPase(s)		Binding protein	Permease(s)		ATPase(s)		Binding protein	Permease(s)		ATPase(s)	
Trk-type K^+	—	PF1856	—	PF1857	—	—	PAB0248	—	PAB0249	—	—	PH1813	—	PH1815	—
Sulphate/molybdate	PF0080	PF0081	—	PF0082	—	PAB0101	PAB0102	—	PAB0103	—	PH0151	PH0154	—	PH0157	—
Phosphate	PF1003	PF1006	PF1007	PF1008	—	PAB2365	PAB0698	PAB0699	PAB0700	—	—	—	—	—	—
Di/oligopeptide	—	—	—	—	—	PAB0091	PAB0092	PAB0093	PAB0094	PAB0095	—	—	—	—	—
Di/oligopeptide[b]	PF1209	PF1210	PF1211	PF1212	PF1213	PAB0627	PAB0628	PAB2363	PAB0630	PAB0631	—	—	—	—	—
Di/oligopeptide	PF0357	PF0358	PF0359	PF0360	PF0361	PAB1343	PAB1344	PAB1345	PAB1346	PAB1347	PH0502	PH0503	PH0504	PH0505	PH0507
Di/oligopeptide	PF1408	PF1409	PF1410	PF1411	PF1412	PAB1871	PAB1872	PAB1873	PAB1874	PAB1875	PH1409	PH1410	PH1411	PH1411	PH1413
Di/oligopeptide	PF0190	PF0191	PF0192	PF0193	PF0194	PAB1193	PAB1194	PAB1195	PAB1196	PAB1197	PH1962	PH1961	PH1960	PH1959	PH1958
Di/oligopeptide	—	—	—	—	—	—	—	—	—	—	PH0807	PH0808	PH0809	PH0810	PH0811
Cobalamine/Fe^{3+}	—	PF0503	—	PF0502	—	—	PAB1535	—	PAB1536	—	—	PH0790	—	PH0791	—
Cobalamine/Fe^{3+}	PF0911	PF0910	—	PF0909	—	PAB0676	PAB0677	—	PAB0678	—	PH1237	PH1236	—	PH1235	—
Cobalt	—	PF0067	—	PF0068	—	—	PAB2261	—	PAB2260	—	—	PH0131	—	PH0132	—
Multidrug	—	PF0582	—	PF0583	—	—	PAB1924	—	PAB1923	—	—	PH0911	—	PH0913	—
Multidrug	—	—	—	—	—	—	—	—	—	—	—	PH0821	—	PH0820	—
Uncharacterized	PF1695	PF1696	PF1697	PF1698	—	PAB0302	PAB0305	PAB0304	PAB0303	—	PH1714	PH1711	PH1712	PH1713	—
Mn^{2+}/Zn^{2+}	PF1774	PF1780	—	PF1779	—	PAB0349	PAB0351	—	PAB0350	—	PH1695	PH1651	—	PH1653	—
Carbohydrate[c]	PF1739	PF1740	PF1741	PF1744	—	—	—	—	—	—	—	—	—	—	—
Carbohydrate	—	—	—	—	—	—	—	—	—	—	PH0753	PH0754	PH0755	PH0756	—
Carbohydrate	PF0119	PF0118	PF0117	PF0116	—	PAB2439	PAB2230	PAB2231	PAB2232	—	PH0206	PH0205	PH0204	PH0203	—
Carbohydrate	PF1967	PF1968	PF1969	PF1970	—	PAB0302	PAB0305	PAB0304	PAB0303	—	PH0025	PH0024	PH0023	PH0022	—
Iron/thiamin	PF1518	PF1520	—	PF1519	—	PAB1835	PAB0543	—	PAB0545	—	PH1349	PH1352	—	PH1350	—
Iron/thiamin	PF1751	PF1748	PF1749	PF1750	—	—	—	—	—	—	—	—	—	—	—
Maltose[d]	PF1938	PF1937	PF1936	PF1933	—	PAB0119	PAB0120	PAB0121	PAB0123	—	—	—	—	—	—

[a] Indicated by the COG database.
[b] β-Glucoside transporter (Koning *et al.*, 2001).
[c] Trehalose/maltose transporter (Koning *et al.*, 2002).
[d] Maltodextrin transporter (Koning *et al.*, 2002).

shows the highest homology with bacterial oligopeptide binding proteins, while PF1408 shows highest homology to dipeptide binding proteins. In both cases, no ORFs encoding carbohydrate hydrolyzing enzymes are located in the vicinity of the transport clusters. This may imply that these ORFs are involved in di- and oligopeptide transport. A homolog of the putative oligopeptide binding protein PF0190 is found in *P. horikoshii* (PH1962). In *P. abyssi*, the N-terminal part of PAB1193 is homologous to the C-terminal part of PF0190, while the C-terminal part of PAB1193 shows homology to the N-terminal part of PF0190. Apparently, gene shuffling has occurred which may have led to a reorientation of the binding protein domains.

Homologs of the putative dipeptide binding protein PF1408 are found both in *P. abyssi* (PAB1871) and *P. horikoshii* (PH1409). The genome of *P. abyssi* contains a second putative dipeptide binding protein, PAB0091, that appears absent in the other *Pyrococcus* species. Also, *P. horiskohii* seems to contain a unique additional putative dipeptide binding protein (PH0807).

COMPARISONS OF THE CARBOHYDRATE-BINDING PROTEIN FAMILY

Five gene clusters in *P. furiosus* seem to encode putative transporters belonging to the carbohydrate ABC transport families. Two of these have been characterized previously as the trehalose/maltose and maltodextrin transporters (Koning *et al.*, 2002). The binding proteins of these transporters, PF1738 and PF1938, show a high homology with the maltose/maltodextrin binding protein, MalE, of *E. coli*. The trehalose/maltose-binding protein, PF1738, is unique for *P. furiosus*, while *P. abyssi* contains a homolog (PAB0123) of the maltodextrin-binding protein encoded by PF1938.

The remaining three ORFs, PF0119, PF1695, and PF1967, seem to belong to the carbohydrate-binding protein family but the identity of the substrate is unknown. Homologs of all three ORFs can be found in both *P. abyssi* (PAB2439, PAB0302, and PAB2333, respectively) and *P. horikoshii* (PH0206, PH1714, and PH0025, respectively). *P. horikoshii* contains an additional unique ORF, PH0753. For all ORFs close homologs can be found in other archaea and bacteria, but in none of these cases, a clear hint for a possible function of the transport clusters is obtained.

EXPRESSION AND INDUCTION OF BINDING PROTEINS IN *P. furiosus*

To assign a potential function to the different unknown transport clusters in *P. furiosus*, mRNA levels of the binding proteins were determined using Northern hybridization techniques. DIG-labeled DNA probes of the different ORFs were used to hybridize Northern blots containing total RNA isolated from *P. furiosus* cells grown on different substrates, both carbohydrates and peptides.

Of the ORFs which are members of the di/oligopeptide-binding protein family, PF0190 and PF1408 are induced when cells are grown on peptides and pyruvate (Table IV). PF1408, a putative dipeptide binding protein, is also induced after growth on cellobiose and maltose. PF1209, the cellobiose/β-glucoside binding protein (CbtA), is induced not only when cells are grown on cellobiose but also on peptone, tryptone, and casein. The fourth pyrococcal member of this carbohydrate-binding protein family, PF0357, is not induced under any of the tested growth conditions.

The characterized members of the carbohydrate-binding protein family, PF1739 (TMBP) and PF1938 (MDBP), are induced when cells are grown on maltose and starch (Table IV) (Koning *et al.*, 2002). PF1739,

Table IV. Expression of the Predicted *P. furiosus* Organic Solute Binding Proteins After Growth on Different Substrates

Growth substrate	Predicted binding proteins								
	PF0119	PF0190	PF0357	PF1209	PF1408	PF1695	PF1739	PF1938	PF1967
Casein	−	++	−	+	++	−	−	+	++
Cellobiose	−	−	−	++	+	−	−	−	−
Maltose	−	−	−	−	+	−	++	++	++
Peptone	−	++	−	++	++	−	−	+	++
Pyruvate	−	+	−	−	+	−	−	−	+
Starch	−	−	−	−	−	−	++	++	−
Tryptone	−	++	−	++	++	−	−	−	++
Yeast extract	−	+	−	−	−	+	++	−	+

Note. − not induced; + slightly induced; ++ highly induced. Experiments were performed in duplicate.

however, is also expressed when cells are grown on yeast extract, most likely because yeast extract contains trehalose. PF1938, on the other hand, is also induced at low levels when cells are grown on casein and peptone. PF0119 is not induced under any of the tested growth conditions, while PF1695 is induced only at low levels when cells are grown on yeast extract. PF1967 is induced under all growth conditions except when cells are grown on cellobiose and starch.

The results obtained with the Northern hybridization techniques do not provide conclusive evidence about the substrate of the uncharacterized ABC transporters of *P. furiosus*, but the studies suggest that this organism is able to accumulate, and possibly utilize, a greater variety of carbohydrates than previously anticipated.

CONCLUSIONS

Archaea employ binding-protein-dependent ABC transporters for carbon source uptake. The high affinity of the binding proteins may enable them to survive in habitats with low concentration of carbon sources such as sugars and peptides. Interestingly, a number of sugar ABC systems involved in transport of oligosaccharides are structurally and functionally homologous to di/oligopeptide ABC transporter. The examination of putative signal peptides of substrate-binding proteins from archaeal genome sequences showed that euryarchaeal and crenarchaeal signal peptides differ substantially from each other. It will be interesting in the future to prove that the binding proteins of the sugar class are indeed anchored by lipidation and to identify the lipoprotein signal peptidase of archaea.

ACKNOWLEDGMENTS

S.V.A. was supported by a Veni-grant from Dutch Organization for Science Research (NWO). The work of S.K. was supported by the Earth and Life Science Foundation (ALW) which is subsidized by NWO.

REFERENCES

Albers, S. V., and Driessen, A. J. M. (2002). *Arch. Microbiol.* **177**, 209–216.

Albers, S. V., Elferink, M. G., Charlebois, R. L., Sensen, C. W., Driessen, A. J. M., and Konings, W. N. (1999a). *J. Bacteriol.* **181**, 4285–4291.

Albers, S. V., Konings, W. N., and Driessen, A. J. M. (1999b). *Mol. Microbiol.* **31**, 1595–1596.

Albers, S. V., Szabo, Z., and Driessen, A. J. M. (2003). *J. Bacteriol.* **185**, 3918–3925.

Barbier, G., Godfroy, A., Meunier, J. R., Querellou, J., Cambon, M. A., Lesongeur, F., Grimont, P.A., and Raguenes, G. (1999). *Int. J. Syst. Bacteriol.* **49**, 1829–1837.

Berks, B. C., Sargent, F., and Palmer, T. (2000). *Mol. Microbiol.* **35**, 260–274.

Bolhuis, A. (2002). *Microbiology* **148**, 3335–3346.

Boos, W., and Shuman, H. (1998). *Microbiol. Mol. Biol. Rev.* **62**, 204–229.

Dassa, E., and Hofnung, M. (1985). *EMBO J.* **4**, 2287–2293.

Detmers, F. J., Lanfermeijer, F. C., Abele, R., Jack, R. W., Tampe, R., Konings, W. N., and Poolman, B. (2000). *Proc. Natl. Acad. Sci. U.S.A.* **97**, 12487–12492.

Diederichs, K., Diez, J., Greller, G., Muller, C., Breed, J., Schnell, C., Vonrhein, C., Boos, W., and Welte, W. (2000). *EMBO J.* **19**, 5951–5961.

Diez, J., Diederichs, K., Greller, G., Horlacher, R., Boos, W., and Welte, W. (2001). *J. Mol. Biol.* **305**, 905–915.

DiRuggiero, J., Dunn, D., Maeder, D. L., Holley-Shanks, R., Chatard, J., Horlacher, R., Robb, F. T., Boos, W., and Weiss, R. B. (2000). *Mol. Microbiol.* **38**, 684–693.

Eichler, J. (2001). *Eur. J. Biochem.* **268**, 4366–4373.

Elterink, M. G., Albers, S. V., Konings, W. N., and Driessen, A. J. M. (2001). *Mol. Microbiol.* **39**, 1494–1503.

Elferink, M. G. L., De Wit, J. G., Demel, R., Driessen, A. J. M., and Konings, W. N. (1992). *J. Biol. Chem.* **267**, 1375–1381.

Erdokimov, A. G., Anderson, D. E., Routzahn, K. M., and Wough, D. S. (2001). *J. Mol. Biol.* **305**, 891–904.

Erra-Pujada, M., Debeire, P., Duchiron, F., and O'Donohue, M. J. (1999). *J. Bacteriol.* **181**, 3284–3287.

Greller, G., Riek, R., and Boos, W. (2001). *Eur. J. Biochem.* **268**, 4011–4018.

Hettmann, T., Schmidt, C. L., Anemuller, S., Zahringer, U., Moll, H., Petersen, A., and Schafer, G. (1998). *J. Biol. Chem.* **273**, 12032–12040.

Higgings, C. F. (1992). *Annu. Rev. Cell Biol.* **8**, 67–113.

Holland, I. B., and Blight, M. A. (1999). *J. Mol. Biol.* **293**, 381–399.

Horlacher, R., Xavier, K. B., Santos, H., DiRuggiero, J., Kossmann, M., and Boos, W. (1998). *J. Bacteriol.* **180**, 680–689.

Kawarabayasi, Y., Sawada, M., Horikawa, H., Haikawa, Y., Hino, Y., Yamamoto, S., Sekine, M., Baba, S., Kosugi, H., Hosoyama, A., Nagai, Y., Sakai, M., Ogura, K., Otsuka, R., Nakazawa, H., Takamiya, M., Ohfuku, Y., Funahashi, T., Tanaka, T., Kudoh, Y., Yamazaki, J., Kushida, N., Oguchi, A., Aoki, K., and Kikuchi, H. (1998). *DNA Res.* **5**, 55–76.

Kikuchi, A., Sagami, H., and Ogura, K. (1999). *J. Biol. Chem.* **274**, 18011–18016.

Koekova, M. V., Storch, K. F., Klein, C., and Oesterhelt, D. (2002). *EMBO J.* **21**, 2312–2322.

Koning, S. M., Elferink, M. G., Konings, W. N., and Driessen, A. J. M. (2001). *J. Bacteriol.* **183**, 4979–4984.

Koning, S. M., Konings, W. N., and Driessen, A. J. M. (2002) *Archaea* **1**, 19–25.

Konrad, Z., and Eichler, J. (2002). *Biochem. J.* **366**, 959–964.

Krueger, R. D., Harper, S. H., Campbell, J. W., and Fabroeg, D. E. (1986). *J. Bacterial.* **167**, 49–56.

Lai, M. C., Hong, T. Y., and Gunsalus, R. P. (2000). *J. Bacteriol.* **182**, 5020–5024.

Lanfermeijer, F. C., Detmers, F. J., Konings, W. N., and Poolman, B. (2000). *EMBO J.* **19**, 3649–3656.

Lee, S. J., Engelmann, A., Horlacher, R., Qu, Q., Vierke, G., Hebbeln, C., Thomm, M., and Boos, W. (2003). *J. Biol. Chem.* **278**, 983–990.

Maeder, D. L., Weiss, R. B., Dunn, D. M., Cherry, J. L., Ganzalez, J. M., DiRuggiero, J., and Robb, F. T. (1999). *Genetics* **152**, 1299–1305.

Mattar, S., Scharf, B., Kent, S. B., Rodewald, K., Oesterhelt, D., and Engelhard, M. (1994). *J. Biol. Chem.* **269**, 14939–14945.

Moody, J. E., Millen, L., Binns, D., Hunt, J. F., and Thomas, P. J. (2002). *J. Biol. Chem.* **277**, 21111–21114.

Nelson, K. E., Clayton, R. A., Gill, S. R., Gwinn, M. L., Dodson, R. J., Haft, D. H., Hickey, E. K., Peterson, J. D., Nelson, W. C., Ketchum,

K. A., McDonald, L., Utterback, T. R., Malek, J. A., Linher, K. D., Garrett, M. M., Stewart, A. M., Cotton, M. D., Pratt, M. S., Phillips, C. A., Richardson, D., Heidelberg, J., Sutton, G. G., Fleischmann, R. D., Eisen, J. A., and Fraser, C. M. (1999). *Nature* **399**, 323–329.

Panagiotidis, C. H., Boos, W., and Shuman, H. A. (1998). *Mol. Microbiol.* **30**, 535–546.

Paulsen, I. T., Nguyen, L., Sliwinski, M. K., Rabus, R., and Saier, M. H. J. (2000). *J. Mol. Biol.* **301**, 75–100.

Pledger, R. J., and Baross, J. A. (1991). *J. Gen. Microbiol.* **268**, 203–211.

Proctor, L. M., Lai, R., and Gunsalus, R. P. (1997). *Appl. Environ. Microbiol.* **63**, 2252–2257.

Quiocho, F. A., and Ledvina, P. S. (1996). *Mol. Microbiol.* **20**, 17–25.

Reich-Slotky, R., Panagiotidis, C., Reyes, M., and Shuman, H. A. (2000). *J. Bacteriol.* **182**, 993–1000.

Robb, F. T., Maeder, D. L., Brown, J. R., DiRuggiero, J., Stump, M. D., Yeh, R. K., Weiss, R. B., and Dunn, D. M. (2001). *Methods Enzymol.* **330**, 134–157.

Roessler, M., Pflueger, K., Flach, H., Lienard, T., Gottschalk, G., and Mueller, V. (2002). *Appl. Environ. Microbiol.* **68**, 2133–2139.

Rose, R. W., Bruser, T., Kissinger, J. C., and Pohlschroder, M. (2002). *Mol. Microbiol.* **45**, 943–950.

Schneider, E. (2001). *Res. Microbiol.* **152**, 303–310.

Schut, G. J., Brehm, S. D., Datta, S., and Adams, W. W. (2003). *J. Bacteriol.* **185**, 3935–3947.

Stock, J. B., and Surette, M. G. (1996). In *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology: Chemotaxis, Vol. 1* (Neidhardt, R. C. I., Ingraham, J. L., Lin, E. C. C., Low, K. B., Magasanik, B., Reznikoff, W. S., Schaechter, M., and Umbarger, H. E., eds.), American Society for Microbiology, Washington, DC, pp. 551–573.

Sumper, M., Berg, E., Mengele, R., and Strobel, I. (1990). *J. Bacteriol.* **172**, 7111–7118.

Verdon, G., Albers, S. V., Dŷkstra, B. W., Driessen, A. J. M., and Thunnissen, A. M. (2003). *J. Mol. Biol.* **330**, 343–358.

Walker, J. E., Saraste, M., Runswick, M. J., and Gay, N. J. (1982). *EMBO J.* **1**, 945–951.

Wanner, C., and Soppa, J. (1999). *Genetics* **154**, 1417–1428.

Xavier, K. B., Martins, L. O., Peist, R., Kossmann, M., Boos, W., and Santos, H. (1996). *J. Bacteriol.* **178**, 4773–4777.

Yuan, Y. R., Blecker, S., Martsinkeuich, O., Millen, L., Thomas, P. J., and Hunt, J. F. (2001). *J. Biol. Chem.* **276**, 32313–32321.

Zivanovic, Y., Lopez, P., Philippe, H., and Forterre, P. (2002). *Nucleic Acids Res.* **30**, 1902–1910.