Archaeal homolog of bacterial type IV prepilin signal peptidases with broad substrate specificity
Albers, Sonja-Verena; Szabó, Zalán; Driessen, Arnold J.M.

Published in:
Journal of Bacteriology

DOI:

IMPORTANT NOTE: You are advised to consult the publisher’s version (publisher’s PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher’s PDF, also known as Version of record

Publication date:
2003

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
Archaeal Homolog of Bacterial Type IV Prepilin Signal Peptidases with Broad Substrate Specificity

Sonja-Verena Albers, Zalán Szabó, and Arnold J. M. Driessen*

Department of Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, 9751 NN Haren, The Netherlands

Received 4 November 2002/Accepted 21 April 2003

A large number of secretory proteins in the thermoacidophile Sulfolobus solfataricus are synthesized as a precursor with an unusual leader peptide that resembles bacterial type IV prepilin signal sequences. This set of proteins includes the flagellin subunit but also various solute binding proteins. Here we describe the identification of the S. solfataricus homolog of bacterial type IV prepilin peptidases, termed PibD. PibD is an integral membrane protein that is phylogenetically related to the bacterial enzymes. When heterologously expressed in Escherichia coli, PibD is capable of processing both the flagellin and glucose-binding protein (GlcS) precursors. Site-directed mutagenesis of the GlcS signal peptide shows that the substrate specificity of PibD is consistent with the variations found in proteins with type IV prepilin-like signal sequences of S. solfataricus. We conclude that PibD is responsible for the processing of these secretory proteins in S. solfataricus.

*S. solfataricus is an obligate aerobic thermoacidophilic crenarchaeon that can use a variety of sugars as a sole carbon source for heterotrophic growth (16, 43). Sugar uptake is mediated by high-affinity binding protein-coupled ABC transporters (2). Some of the binding proteins are synthesized as precursors with an unusual amino-terminal signal sequence for membrane targeting that consists of a short positively charged peptide followed by the cleavage site and a hydrophobic, putative membrane-spanning domain. This signal sequence resembles that of bacterial type IV prepilin proteins. In bacteria, type IV pilins are the structural subunits of pili, which are surface-associated structures involved in processes such as surface attachment, twitching motility, or uptake of extracellular DNA (11, 15, 22, 26). The discovery that the precursors of archaeal flagellins contain a signal sequence homologous to bacterial type IV pilin signal sequences has led to the hypothesis that the assembly and structure of archaeal flagella are similar to those of bacterial pili (5, 14). Indeed, the core structures of bacterial pili and the flagellum of the halophilic archaeon Halobacterium salinarum are similar (8). Also, the cleavage of flagellin precursors could be shown by purification and amino-terminal sequencing of flagellins from various archaea (40) and by in vitro cleavage of heterologously expressed prefagellin (10). However, in some archaea, type IV pilin signal sequences are not only confined to flagellin subunits but are also found in precursors of other extracellular proteins (3). This was first demonstrated for the glucose-binding protein from S. solfataricus (2). In addition, two other sugar-binding proteins, i.e., an arabinose- and a trehalose-binding protein, were found to be processed at their amino termini as expected for type IV prepilin-like signal sequences (2,12). A recent analysis of the genome sequence of S. solfataricus suggests that there are 10 proteins with a type IV pilin-like signal sequence (1, 32), 2 of which are putative solute binding proteins. This raises the question of whether these proteins utilize the same signal peptidase as the flagellin subunits (3) or if they are involved in a separate pathway for protein targeting and maturation, as suggested in another study (19).

One of the best-studied type IV prepilin peptidases is PilD from Pseudomonas aeruginosa, a bifunctional enzyme that both cleaves and N-methylates the pilin precursor PilA (36). Whereas methylation is not necessary for pilus function (29), cleavage is an essential activity, as PilD mutants are defective in pilus assembly and accumulate uncleaved PilA precursors in the cytoplasmic membrane (28). Because of the similarity between prepilin and archaeal preflagellin signal sequences, the presence of a gene coding for an enzyme similar to PilD was expected in archaeal genomes. This assumption was recently confirmed by the cloning of FlaK, the preflagellin peptidase from Methanococcus maripaludis (4). However, FlaK and its archaeal homologs show only a very low similarity to bacterial type IV prepilin peptidases, possibly because the bacterial enzyme is bifunctional, while there is no evidence for a modified amino acid at the +1 position of mature archaeal flagellins (12, 40).

In this study, we have employed an in vitro assay to monitor type IV prepilin-like signal peptidase activity in isolated S. solfataricus membranes. This assay allowed the biochemical demonstration of processing activity toward the precursor forms of flagellin (preFlaB) and glucose-binding protein (preGlcS). We identify a candidate gene in the S. solfataricus genome that encodes a membrane protein that has weak homology to PilD. When expressed in E. coli, processing of both preFlaB and preGlcS could be demonstrated. This finding and a further analysis of the substrate specificity of this enzyme, termed PibD (for peptide involved in biogenesis of prepilin-like proteins), suggest that S. solfataricus employs a single type

* Corresponding author. Mailing address: Department of Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Kerklaan 30, 9751 NN Haren, The Netherlands. Phone: 31-50-3632164. Fax: 31-50-3632154. E-mail: a.j.m.driessen@biol.rug.nl.
TABLE 1. Plasmids used in this study

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Description</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>pET15b</td>
<td>Expression vector with T7 promoter</td>
<td>Novagen</td>
</tr>
<tr>
<td>pSA4</td>
<td>Derivative of pET15b containing the multiple cloning site and C-terminal hexa-His tag of pET324</td>
<td>This study</td>
</tr>
<tr>
<td>pET324</td>
<td>Derivative of pTRC99</td>
<td>This study</td>
</tr>
<tr>
<td>pSA5</td>
<td>Insertion of a C-terminal His tag in the BamHI/XbaI site of pET324</td>
<td>This study</td>
</tr>
<tr>
<td>pET1210</td>
<td>pSA5 carrying glcS as BspHI/BamHI fragment</td>
<td>This study</td>
</tr>
<tr>
<td>pET1215</td>
<td>pET1210 derivative lacking the internal NeoI fragment of glcS</td>
<td>This study</td>
</tr>
<tr>
<td>pET1217</td>
<td>pET1215 derivative with G12A mutation</td>
<td>This study</td>
</tr>
<tr>
<td>pET1212</td>
<td>L13I</td>
<td>This study</td>
</tr>
<tr>
<td>pET1213</td>
<td>G12A/L13I</td>
<td>This study</td>
</tr>
<tr>
<td>pET1214</td>
<td>L13F</td>
<td>This study</td>
</tr>
<tr>
<td>pET1215</td>
<td>G12L</td>
<td>This study</td>
</tr>
<tr>
<td>pET1216</td>
<td>K1A</td>
<td>This study</td>
</tr>
<tr>
<td>pET1217</td>
<td>A10K</td>
<td>This study</td>
</tr>
<tr>
<td>pET1218</td>
<td>K1D</td>
<td>This study</td>
</tr>
<tr>
<td>pET1219</td>
<td>G12R</td>
<td>This study</td>
</tr>
<tr>
<td>pET1280</td>
<td>L13R</td>
<td>This study</td>
</tr>
<tr>
<td>pET1218</td>
<td>L13D</td>
<td>This study</td>
</tr>
<tr>
<td>pZA1</td>
<td>pSA4 carrying flaBΔ1–4 NeoU BamHI fragment</td>
<td>This study</td>
</tr>
<tr>
<td>pZA2</td>
<td>pSA4 carrying flaB as BamHI/BamHI fragment</td>
<td>This study</td>
</tr>
<tr>
<td>pZAS</td>
<td>pSA4 carrying pibD as NcoI/BamHI fragment</td>
<td>This study</td>
</tr>
</tbody>
</table>

IV prepilin-like peptidase to process a broad range of precur-

SULFOLOBUS TYPE IV PREPILIN PEPTIDASE HOMOLOG 3919

MATERIALS AND METHODS

Strains and growth conditions. S. solfataricus P2 (DSM 1617, obtained from the Deutsche Sammlung von Mikroorganismen und Zellkultur GmbH, Brauns-

degew, Germany) was grown aerobically at 80 °C by using the medium described by Brock et al. (6) supplemented with 0.1% sucrose and 0.1% yeast extract. The medium was adjusted to pH 3 with sulfuric acid. Escherichia coli strain DH5α (17) was used for all cloning steps. E. coli strains BL21(DE3) (Novagen, Madison, Wis.) and C43 (DE3) (27) were used for the overproduction of protein. Both strains carry the pACYC-RIL plasmid that encodes additional tRNAs for rare codons (Stratagene, La Jolla, Calif.).

Isolation of membranes. To prepare membranes of E. coli, cells were resus-
pended in 50 mM Tris-Cl (pH 7.5) and 1 mM EDTA containing a small amount of DNAseI. The suspension was subsequently passed through a French pressure

stream from the T7 promoter and removing the N-terminal epitope tag of the vector. To delete unwanted restriction sites, this vector was digested with BamHI and XhoI, the cohesive ends of the vector were filled up with Pwo DNA poly-
mersase, and the product was religated, yielding pSA3. Finally, a 392-bp HindIII fragment from pAMP42 (kindly provided by Antonia Picon) was inte-
grated into the corresponding restriction sites of pSA3, thereby introducing a C-terminal hexa-His tag in frame with a BamHI restriction site. This vector was designated pSA4 and is suitable for cloning and expression of ORFs provided as NcoI/BamHI or BamHI/BamHI fragments.

The gene for gcs was amplified with the primers 5′-CCCCCGGTATGGTGTTATTATTATGTTTTC and reverse, containing BamHI. Again, the native stop codon was deleted. The PCR product was cloned into the vector pSA5, which contains a tac promoter and adds a C-terminal hexa-His tag in frame with the BamHI restriction site, yielding pET2120. pSA5 was constructed by the ligation of the 392-bp BamHI-HindIII fragment from pAMP42 into the expression vector pET324 (42), thereby allowing us to add a C-terminal hexa-His tag in frame with the BamHI restriction site, yielding pET2120. pSA5 was constructed in pET2152 by the Quickchange (Stratagene) method of mutagenesis, which utilizes inverse PCR primed by divergent overlapping primers containing the desired complementary nucleotide ex-

changes. Expression of recombinant genes in E. coli. E. coli strain BL21(DE3)-RIL codon plus (Stratagene) carrying plasmid pZA1, pZA2, pET2120, or pET2125 was grown in JYT medium (16 g of trypton per liter, 10 g of yeast extract per liter, and 5 g of NaCl per liter) supplemented with antibiotics at 37°C until an optical density at 600 nm of 0.5 to 0.8 was reached. Expression was induced by the addition of 0.5 mM isopropyl-beta-D-thiogalactopyranoside (IPTG, Roche Diagnostics). Cells were grown for an additional 4 h and harvested by centrif-

185, 2003
In vitro processing assay. The in vitro preFlaB and preGlcS processing assay was based on the method developed for the Methanococcus voltae preflagellin peptidease (10). Isolated E. coli membranes (about 5 μg of total protein) either containing preFlaB or preGlcS and S. solfataricus membranes (about 10 μg of total protein) were each preincubated for 10 min in 10 μl of assay buffer (25 mM MES [pH 6.5], 0.5% [vol/vol] Triton X-100, 150 mM KCl, and 1 mM EDTA) at room temperature. Both solubilized samples were mixed and incubated for 30 min (preFlaB) or 1 h (preGlcS) at 55°C, unless stated otherwise. The reaction was stopped by the addition of 5 μl of sodium dodecyl sulfate (SDS) sample buffer and analyzed by SDS–12% polyacrylamide gel electrophoresis (PAGE) and subsequent transferring to polyvinylidene difluoride membranes (Roche Diagnostics). Protein was detected by Western immunoblotting by using monoclonal anti-His antibodies (Dianova, Hamburg, Germany). The activity of cloned PbD was tested as described above except that E. coli inner membranes (about 5 μg of total protein) containing overexpressed PbD were used instead of the S. solfataricus membranes.

RESULTS

In vitro demonstration of type IV prepilin-like signal peptidease activity. To identify the S. solfataricus enzyme responsible for the processing of precursor proteins with a type IV prepilin-like signal sequence, an in vitro assay was established. For this purpose, two proteins with type IV prepilin-like signal sequences were chosen as substrates, flagellin and glucose-binding protein (preGlcS) (2). The flagellin gene was identified in the S. solfataricus genome by using the published amino-terminal sequence of the closely related Sulfolobus shibatae flagellin (13). We discovered one protein encoded by SSO2323, annotated as putative flagellin, and named it according to previously characterized archaeal flagellins FlaB. The genes coding for preGlcS and preFlaB were cloned by PCR into an expression vector suitable for heterologous expression in E. coli, adding a C-terminal hexa-His epitope tag for detection and purification. The annotated sequence of SSO2323 contains two amino-terminal methionine residues at positions 1 (as annotated) and 5, and it is possible that in vivo the second ATG sequence is used as a start codon for translation of flaB. Therefore, both the possible genes were cloned. Unless stated otherwise, the shorter version of preFlaB was used in the following experiments. Overexpressed preFlaB migrated in an SDS-polyacrylamide gel at an apparent molecular mass of about 29 kDa (Fig. 1), which deviates from the predicted mass of 34 kDa, probably due to tight folding of the hydrophobic flagellin. PreGlcS was also expressed but it was partially degraded during overexpression in E. coli (data not shown). However, removal of an internal NcoI fragment of the glcS gene resulted in a truncated version of preGlcS, termed preGlcS*, that could be stably expressed.

Since the precursor proteins expressed cofractionated with cytoplasmic membranes (data not shown), inner membrane vesicles isolated from E. coli cells expressing either preFlaB or preGlcS* were solubilized with Triton X-100 and mixed with S. solfataricus membranes. After incubation at 55°C, cleavage of the amino terminus of preFlaB and preGlcS* was detected by Western immunoblot analysis (Fig. 1B). Processing of preFlaB and preGlcS* resulted in the removal of peptides with lengths of 13 and 12 amino acids, respectively, while the carboxy-terminal His tag remains intact and can be detected with a specific antibody (Fig. 1A). The precursors remained unprocessed when the incubation was performed at 4°C or when incubated in the absence of S. solfataricus membranes. For both precursors, the presence of detergent was required for cleavage to occur (data not shown). In the case of preFlaB, a second, weaker band appeared that was also present in the absence of S. solfataricus membranes. The intensity of this band varied with different preparations of E. coli membranes containing preflagellin, but the activity appeared to be due to degradation in E. coli. PreFlaB lacking the first four amino acids of the annotated sequence was processed more efficiently than the longer version (data not shown). Therefore, we decided to use the shorter protein in our experiments. These data demonstrate that membranes of S. solfataricus harbor a type IV prepilin-like signal peptidease activity.

Identification and cloning of the S. solfataricus type IV prepilin-like signal peptidease. To identify the type IV prepilin-like signal peptidease, we first searched the S. solfataricus genomic database for a homolog of bacterial enzymes, including those from thermophilic bacteria. However, the BLAST search was unsuccessful. On the other hand, the Cluster of Orthologous Groups (COG) database (http://www.ncbi.nlm.nih.gov/COG [38, 39]) lists several archaeal proteins in the group designated COG1989, signal peptidease, cleaves prepilin like proteins. The COG database relies on a phylogenetic classification of proteins. Therefore, proteins with very little sequence identity are clustered in one group due to phylogenetic relationships (39). Subsequently, protein sequences of archaeal COG1989 members were used to perform BLAST searches in the S. solfataricus genome database. Single significant hits (e < 10−3) were obtained only with MJ1282.1 from Methanocaldococcus jannaschii and AF0936 from Archaeoglobus fulgidus for the same putative membrane protein, SSO0131. The SSO0131 protein shares 19% identical amino acid sequences with AF0936 and 16.1% with MJ1282.1, respectively. However, higher identities are found with proteins that are not included in COG1989: ST2258 from Sulfolobus tokodaii, MA3102 from Methanosarcina acetivorans, and FlaK from M. maripaludis (Fig. 2A). The primary sequence of SSO0131 has a length of 236 amino acids and a predicted topology of five transmembrane segments.
(TMS) with a large cytoplasmic loop between TMS 4 and 5 (Fig. 2B). This topology is similar to most putative archaeal type IV prepilin-like signal peptidases but is unlike that of bacterial enzymes such as *Vibrio cholerae* TcpJ with eight TMS and various large cytoplasmic loops (24, 25). The gene coding for the SSO0131 protein is not contained in an apparent operon structure, but remnants of an insertion element are located upstream of the gene, suggesting a possible rearrangement of the surrounding genome sequence.

To establish whether SSO0131 has cleavage activity in vitro, we PCR amplified and cloned the corresponding gene into an expression vector with a carboxyl-terminal hexa-His tag. The gene was overexpressed in *E. coli*, and the recombinant protein could be detected in membrane preparations by SDS-PAGE. 

![Figure 2](image-url)
and immunoblotting (Fig. 2). Recombinant SSO0131 was tested for in vitro cleavage activity as described above. Both precursor proteins were cleaved in the same way as found with S. solfataricus membranes (Fig. 3, lane 4) but remained unprocessed when incubated with membranes isolated from E. coli cells harboring the empty vector (Fig. 3, lane 5). The reaction is temperature dependent, as the cleavage reaction was less efficient at 20 and 37°C (Fig. 3, lanes 2 and 3). In order to confirm the in vitro cleavage site of FlaB, we made an attempt to purify cleaved FlaB and determine the N-terminal sequence of the protein. However, when cleaved at 55°C, flagellin was insoluble in various buffers, including those containing 2% Triton X-100 or 1% N-lauroylsarcosine in combination with 8 M urea or 6 M guanidine hydrochloride, respectively. When the cleavage reaction was performed at 37°C, the protein could be partially solubilized with 1.5% Triton X-100, 8 M urea, and 20 mM β-mercaptoethanol. The solubilized portion of flagellin was partially purified on a Ni-NTA-Sepharose column, but the eluted protein resisted N-terminal sequencing, although no N-terminal blockage was detected (data not shown).

In conclusion, we demonstrated that the SSO0131 protein is a thermostable type IV prepilin-like signal peptidase of S. solfataricus. We therefore name the corresponding gene pibD (protein involved in biogenesis of prepilin-like proteins). PibD cleaves both the preFlaB and preGlcS and is therefore potentially involved in the secretion of two proteins with completely unrelated functions.

Signal sequence specificity of the S. solfataricus PibD protein. To determine the substrate specificity of PibD and to confirm the exact site of processing, we constructed signal peptide mutants of preGlcS* by site-directed mutagenesis (Fig. 4). All mutants could be overexpressed in E. coli as unprocessed precursor proteins (data not shown). The in vitro cleavage of the preGlcS* signal peptide mutants, either by native S. solfataricus membranes (Fig. 4A) or by E. coli overexpressing recombinant pibD (lower panel). As controls, preGlcS* was incubated in absence or presence of cleavage activity, respectively (first two lanes in each panel).

FIG. 3. Temperature dependent in vitro cleavage of preFlaB (top panel) and preGlcS* (lower panel) by heterologously expressed PibD. Solubilized inner membranes from E. coli containing preFlaB or preGlcS* were incubated with E. coli membranes containing overexpressed PibD protein and incubated at the temperatures indicated. No processing was observed at 20°C or when membranes isolated from cells harboring the empty vector were used.

FIG. 4. In vitro processing of preGlcS* signal peptide mutants. (A) The signal sequence of GlcS is shown at the top; the cleavage site is indicated by a gap. The mutations introduced are listed with the nomenclature used in panel B. (B) GlcS* precursors with altered signal sequences were incubated with membranes from either S. solfataricus membranes (top panel) or E. coli overexpressing recombinant pibD (lower panel). As controls, preGlcS* was incubated in absence or presence of cleavage activity, respectively (first two lanes in each panel).
TABLE 2. Alignment of the N-terminal parts of S. solfataricus secretory proteins with a putative type IV prepilin-like signal sequencea

<table>
<thead>
<tr>
<th>ORF no.</th>
<th>Signal sequence</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSO0118</td>
<td>MKGYKLKRKGGLSILRVLVTLLVGLYASNGEFTSQNAS</td>
<td>Hypo</td>
</tr>
<tr>
<td>SSO0099</td>
<td>MSRSFKSNRRKKRGLSGLTTITGVAIVVGAATLRRGRPSQVST</td>
<td>TreS</td>
</tr>
<tr>
<td>SSO2146</td>
<td>MMASRKKRKGGLSGLTTITGVAIVVGAATLRRGRPSQVST</td>
<td>Hypo</td>
</tr>
<tr>
<td>SSO2152</td>
<td>MLNTVRKGLGSOLTMVLLASVILTVSILFTVQGPSEGSGWY</td>
<td>Hypo</td>
</tr>
<tr>
<td>SSO2323</td>
<td>MLKEYNKVRKGGLAGILDAIIIFAITASVYLAIVINM5LEVFAQKAS</td>
<td>FlaB</td>
</tr>
<tr>
<td>SSO2681</td>
<td>MKRYKLGPLALVLVILVALAIAVSLlYYFFVGLRHSITTG</td>
<td>Hypo</td>
</tr>
<tr>
<td>SSO2847</td>
<td>MKRKYPSICLKLSTQATLAVIVVIIIIGVVGTVLFQGSTTAVT</td>
<td>GlkS</td>
</tr>
<tr>
<td>SSO0037</td>
<td>MKGVIASSILAILIIVVSVVELLILLYLILELARKMSNIAYS</td>
<td>Hypo</td>
</tr>
<tr>
<td>SSO2684</td>
<td>MGISAAITFLVLLVIIIAVYTVYHLNNAVNQYGLY</td>
<td>Hypo</td>
</tr>
<tr>
<td>SSO0489</td>
<td>MKGFSFLAVVIIIIIVIAAVAGIPFVSQGSGNNTTST</td>
<td>PBP</td>
</tr>
<tr>
<td>SSO1017</td>
<td>MMVKLAISSIFSLIVMILTVLSPYTFQTYTNS5IQAN</td>
<td>Hypo</td>
</tr>
<tr>
<td>SSO1171</td>
<td>MRRKSEQKDICYKAIKTLVAVVIIIVVIAIGGYALSSLQH5PAAPST</td>
<td>sugar</td>
</tr>
<tr>
<td>SSO2846</td>
<td>MEQGYKRAISITSTAIADVIIIVILIIVGVVAFQGMSHATSS</td>
<td>Hypo</td>
</tr>
<tr>
<td>SSO3066</td>
<td>MSRRLLYKAISSITAIIIIVVIIIIAAAAGLAAAAYSSSRPATST</td>
<td>AraS</td>
</tr>
<tr>
<td>SSO2712</td>
<td>MKALSTLAMAVVIIITAVVAAAYLITSSHFISPSTTTT</td>
<td>Sugar</td>
</tr>
<tr>
<td>SSO3140</td>
<td>MKSAALSSFLTITLILLLVLLFALLYLPNNS5PFY5QSOQ</td>
<td>Hypo</td>
</tr>
</tbody>
</table>

Consensus: KGLS

a Positively charged residues in the signal peptide are shown in boldface while hydrophobic amino acids in the core of the signal sequence are underlined. The N-terminal sequence of proteins with boldface numbers has been determined experimentally. Abbreviations used: Hypo, hypothetical protein; PBP, putative phosphate-binding protein.

DISCUSSION

Here we have identified the S. solfataricus type IV prepilin peptidase, PibD, by a combinatorial genomic approach. We used sequences of putative archaeal type IV prepilin peptides listed in the COG database in order to conduct a BLAST search of the S. solfataricus genome. This identified SSO0131 as a possible candidate gene. Importantly, processing activity of native S. solfataricus membranes and recombinant SSO0131 (PibD) overexpressed in E. coli could be demonstrated by using an assay previously used for the identification and characterization of the peptidases from various bacteria (37) and methanogenic archaea (4, 10). This assay also allowed for the detection of the processing of both the precursors of the flagellin and the glucose-binding protein. Although the optimal growth temperature of S. solfataricus is 80°C (43), we could not detect cleavage activity when the assay was performed at temperatures above 60°C (data not shown). This phenomenon is due to the clouding point of the Triton X-100 detergent that was used for the membrane solubilization. At temperatures above 63°C, Triton X-100 loses its solubilizing properties.

Since the recombinant PibD in isolated E. coli inner membranes showed processing activity for both precursor proteins tested, our results indicate that a single enzyme of S. solfataricus is capable of processing these distinct secretory proteins (3). FlaB ultimately assembles into the flagellum structure, whereas GlcS has to function as an extracellular sugar-binding subunit of an ABC transporter.

Cleavage of preGlcS*, an internally truncated version of full-length preGlcS, appeared less efficient than that of pre-FlaB. This could be due to incorrect folding, which possibly interferes with its interaction with PibD. However, we favor the explanation that the difference in cleavage efficiency is due to intrinsic properties of the signal peptides, such as the number of positively charged residues in the signal sequence, which is six and four for preFlaB and preGlcS, respectively. Indeed, a signal peptide mutant of preGlcS* which contained an additional lysine residue in the signal peptide was more efficiently processed than the wild-type protein (Fig. 4). Jarrell and colleagues (19) suggested that the archaeal type IV prepilin-like signal peptidase homologs are in fact preflagellin peptidases, i.e., enzymes with a dedicated role in the processing of flagellin precursors. According to these authors’ hypothesis, other pro
teins with type IV prepilin-like signal sequences (such as the sugar-binding proteins GlcS, TreS, and AraS from \textit{S. solfataricus}) would require another processing enzyme. It was argued that the highly conserved amino acids in the signal peptide of all archaeal flagellins, i.e., glycine at $-1$ and lysine or arginine at positions $-2$ and $-3$, are essential for recognition by the prefagellin peptidase and thus deviation from this consensus sequence would require a second peptidase with altered substrate specificity. Processing of the flagellin precursor by the \textit{M. voltae} peptidase indeed appears very sensitive toward alterations of the signal sequence. Most mutations result in partial or complete loss of cleavage activity (41). However, \textit{S. solfataricus} PibD seems to be equipped with a much broader specificity. It processes preGlcS*, even though the signal sequence lacks a lysine or arginine at position $-3$, whereas an \textit{M. voltae} flagellin mutant with alanine or glutamate at the same position was cleaved only partially (41). Moreover, a lysine-to-alanine mutation of the GlcS* signal sequence at position $-2$ did not affect the cleavage reaction, and glycine at position $-1$ could be changed to alanine (Fig. 4) without loss of cleavage activity. Recently, we identified a number of putative precursor proteins with type IV prepilin-like signal sequences (1). Rescanning the genome with the current consensus sequence revealed a number of new candidate proteins (Table 2). These signal sequences have a relatively conserved core of four amino acids around the cleavage site with the consensus sequence [K/R][G/A][L/I/F][S/T/A] ($-1$ and $+1$ positions relative to the processing site shown in boldface). We have previously shown that the sugar-binding proteins GlcS and TreS are processed between a glycine and a leucine residue, while AraS is processed between alanine and isoleucine (2). Taken together, these data suggest that PibD is capable of processing a wide variety of proteins with a type IV prepilin-like signal sequence, including binding proteins (12) and prefagellin. Intriguingly, the genome sequence of another species from the genus \textit{Sulfolobus}, \textit{S. tokodaii} (21), does not code for homologs of the \textit{S. solfataricus} sugar-binding proteins that contain a type IV prepilin-like signal sequence (unpublished data). Hence, \textit{S. solfataricus} seems to have undergone a specialization concerning the variety of solute binding proteins and also the mechanism of secretion of these proteins.

In the gram-negative bacterium \textit{P. aeruginosa}, a number of possible alterations at the $+1$ position (phenylalanine) of the prepilin PilA are tolerated in vivo, whereas replacement of the $-1$ glycine for any other amino acid except alanine resulted in complete loss of a functional pilus structure (35). With alanine at position $-1$, PilA was partially processed in vivo. In contrast, the pseudopilin PulG from \textit{Klebsiella oxytoca} remained unpocesssed in vivo when the $-1$ residue (glycine) was changed to alanine, valine, or glutamate (30). Replacement of the glycine at position $-1$ of the \textit{M. voltae} prefagellin signal sequence to alanine resulted in reduced processing activity, and other amino acids were not tolerated at all at this position, resulting in a prefagellin mutant that cannot be cleaved (41). Also, mutants with alterations at position $-2$ (lysine or alanine) or $-3$ (lysine to alanine or glutamate) were not cleaved or partially cleaved, respectively (41). Strikingly, no loss of cleavage activity was observed when these mutations were introduced into the signal sequence of preGlcS*. Hence, the \textit{S. solfataricus} PibD seems to be functionally more similar to PibD and seems to be equipped with a relatively broad substrate specificity.

In \textit{M. jannaschii}, the putative prefagellin peptidase gene (designated \textit{flaK} [4]) is downstream of the flagellin gene cluster and may be cotranscribed with other \textit{fla} genes (4). The genome of \textit{M. jannaschii} (7) codes for a paralog of FlaK, MJ1282.1, which is located elsewhere in the genome. The gene seems to be part of an operon that also codes for two homologs of bacterial type II secretion proteins. Although MJ1282.1 might represent a pseudogene without a function to the cell, it is also possible that it is involved in the processing of other proteins synthesized with a type IV prepilin-like signal sequence. So far the only candidate is the S-layer protein of \textit{M. jannaschii} (3), but experimental evidence is required to confirm this hypothesis.

Although the primary sequence conservation between archaeal and bacterial type IV prepilin-like signal peptides is very low, an evolutionary relationship between these proteins is evident and therefore the catalytic mechanism of signal peptide cleavage reaction may be identical for these enzymes. In the past, cysteines in PibD have been implicated in catalysis (34), although XpsO from \textit{Xanthomonas campestris} can functionally replace PibD despite the fact that it does not contain cysteine residues (18). Also, a recent study on TcpJ of \textit{V. cholerae} suggests that in fact two aspartate residues are essential for processing (24). Unlike the cysteine, the two aspartate residues are highly conserved in all described bacterial type IV prepilin-like signal peptidases. An alignment of the archaeal enzymes reveals the presence of three conserved aspartate residues (Fig. 2). Future studies should be directed at the identification of the catalytic residues by site-directed mutagenesis.

**ACKNOWLEDGMENTS**

This work was supported by a VENI-grant (project code 863.02.001) of The Netherlands Organization for Scientific Research (NWO) to S.-V.A.

We thank Michael Galperin for suggesting the use of the COG database for identification of the \textit{S. solfataricus} type IV prepilin-like signal peptidase. S.-V. Albers and Z. Szabó contributed equally to this work.

**REFERENCES**


27. Miroux, B., and J. E. Walker. 1996. Over-production of proteins in Esche-