Identification and Molecular Characterization of the First α-Xylosidase from an Archaeon*

Received for publication, December 27, 1999, and in revised form, April 26, 2000
Published, JBC Papers in Press, May 8, 2000, DOI 10.1074/jbc.M910392199

Marco Moracci‡, Beatrice Cobucci Ponzano‡, Antonio Trincone‡, Stefania Fusco‡,
Mario De Rosa, John van der Oost**, Christoph W. Sensen‡‡, Robert L. Charlebois§§,
and Mose’ Rossi‡¶

From the ‡Institute of Protein Biochemistry and Enzymology, Consiglio Nazionale delle Ricerche (CNR),
80125 Naples, Italy, the ¶Istituto per la Chimica di Molecole di Interesse Biologico, CNR, 80072 Naples, Italy,
the ||Istituto di Farmacologia e Tossicologia, Seconda Universita’ di Napoli, Centro Ricerche Interdipartimentale
Biomateriali (CRIB), 80128 Naples, Italy, the **Laboratory of Microbiology, Department of Biomolecular Sciences,
Wageningen University, NL-6703 CT Wageningen, The Netherlands, the ¶¶Institute for Marine Biosciences,
National Research Council of Canada, Halifax, Nova Scotia B3H 3Z1, Canada, and the §§Department of Biology,
University of Ottawa, Ontario K1N 6N5, Canada, and the ¶¶¶Dipartimento di Chimica Organica e Biologica,
Università di Napoli “Federico II,” 80134 Naples, Italy

We here report the first molecular characterization of an α-xylosidase (XylS) from an Archaeon. Sulfolobus solfataricus is able to grow at temperatures higher than 80 °C on several carbohydrates at acidic pH. The isolated xylS gene encodes a monomeric enzyme homologous to α-gluco- and α-xylosidases, α-amylases, α-galactosidases, glucoamylases and sucrase-isomaltases of the glycosyl hydrolase family 31. xylS belongs to a cluster of four genes in the S. solfataricus genome, including a β-glycosidase, an hypothetical membrane protein homologous to the major facilitator superfamily of transporters, and an open reading frame of unknown function. The α-xylosidase superfamily of transporters, and an open

Celluloses and hemicelluloses are the most abundant polysaccharides in nature. They represent the principal structural component of plant cell walls and are associated with lignin and other polysaccharides. Cellulose is a linear homopolymer of up to 14,000 glucose units linked by β-(1,4) glucosidic bonds forming rigid microfibrils that, in vivo, are assembled to hemicelluloses, mostly xylans and xyloglucans. Xylan is a heteropolymer consisting of a backbone of β-(1,4)-linked D-xylose residues with various branching saccharidic groups (e.g. glucuronic acid, arabinose). Xyloglucan is widely distributed in plants, being the principal hemicellulose component in the primary cell wall (20% of total cell wall) and one of the most abundant storage polysaccharides in seeds (>40% in weight in some species). This polymer is composed of a β-(1,4)-glucan backbone, with α-(1,6)-o-xylose groups linked to about 75% of the glucosyl residues. Thus, the disaccharide isoprimeverose (α-D-xylopyranosyl-(1,6)-D-glucopyranosyl) represents the building block of xyloglucan. Additional ramifications of β-D-galactosyl-(1,2)-o-xylosyl and α-l-fucosyl-(1,2)-β-D-galactosyl-(1,2)-α-o-xylosyl chains, are also α-(1,6)-linked at lesser extent to the main backbone. Xyloglucans from different plant tissues and species greatly vary in molecular mass and chemical composition: storage xyloglucans from seeds are not fucosylated if compared with the same polymer from primary cell walls (1).

The degradation and recycling of cellulose and xylan are well studied processes, involving cellulosytic microorganisms, which play an important role in the biosphere (for reviews see Refs. 2 and 3). Instead, relatively little is known about the mechanism of xyloglucan degradation and the enzymatic systems involved in the metabolism of isoprimeverose and xyloglucan oligosaccharides. In plant seeds, the hydrolysis of xyloglucan occurs after germination, during the mobilization of this storage polysaccharide, and has been extensively studied in nasturtium seeds (Tropaeolum majus L.). In this case, xyloglucan is hydrolyzed by the concerted action of at least four enzymatic activities: an endo-(1,4)-β-glucanase yields xyloglucan fragments that are substrates of a β-galactosidase and of two xyloglucan-oligosaccharide-specific α-xylanase and β-glucosidase enzymes (1). Only few examples have been reported of enzymes able to hydrolyze xyloglucan oligosaccharides from eukaryal and bacterial microorganisms (4–8). Among these, the xylPQ regulon of Lactobacillus pentosus, which encodes for a putative membrane protein transporter and an α-xylosidase, is the only genetic system involved in the metabolism of isoprimeverose described so far (9).

Cellulosytic organisms and enzymes are widespread in Bacteria and Eukarya, whereas β-bond-specific glycosyl hydrolases are extremely rare in Archaea (for reviews see Refs. 10 and 11). In particular, no hyperthermophilic Archaea have been found to grow on cellulose, and only recently it has been reported that the euryarchaeote Pyrococcus furiosus could grow on laminarin.
(β-1,3 linkages) and lichenan (β-1,4 and β-1,3 linkages) and only two endo-glucanases have been identified from this source (12, 13). The utilization of xylglucan in this domain has never been investigated.

*Salopolus solfataricus*, originally isolated from a solfataric field in the area of Naples, Italy (14), is an hyperthermophilic Crenarchaeon able to grow chemoheterotrophically at acidic pH (pH 3–5) and at high temperatures (80–87 °C). *S. solfataricus* was reported to utilize as sole carbon source peptides, several mono- and disaccharides, dextrins, and starch, whereas arabinogalactan, agarose, cellulose, and hemicellulose substrates could not support growth (14, 15). The inspection of the *S. solfataricus* genome sequenced so far revealed two genes with significant homology to the clan GH-C of the glycosyl hydrolase superfamily, consisting of β-1,4-specific xylanases and cellulas (families 11 and 12). In addition, both genes are clustered with a putative β-xylidosidase (16). Although the substrate specificity of these gene products remains to be identified, they could be involved in the degradation and assimilation of plant polysaccharides. The only examples of glycosyl hydrolases involved in the utilization of dextrins, whereas the function in vivo involved in the utilization of dextrins, whereas the function in vivo of the β-glycosidase (LacS), which is under active study in our laboratory, is still obscure (21–23). In an effort to determine the full set of glycosyl hydrolases produced by this hyperthermophilic Archaeon, we have identified a novel α-xylidoside (XylS) with high specificity for isoprimeverose and xyloligosaccharides. We describe here the cloning and heterologous expression of *xylS* and the enzymatic characterization of its gene product. *xylS* maps in a locus of the *Salopolus* chromosome, nearby the β-glycosidase gene (*lacS*), an open reading frame (ORF)3 encoding a putative sugar transport protein (major facilitator superfamily, *msf*) and another open reading frame of unknown function (*orf3*). We show here that XylS and LacS cooperate in the xyloligocarbohydrate hydrolysis in *vivo*. This is the first α-xylidoside described in Archaea: its molecular characterization has implications for the function in *vivo* and the evolution of these enzymes.

**EXPERIMENTAL PROCEDURES**

Archaeal Strain and Cultivation—*S. solfataricus* cells, strain MT4, were grown at 87 °C, pH 3.0 as described previously (14) in a minimal salts medium supplemented either with yeast extract (0.1%) plus sucrose (0.2%), or with tamarind seed xylglucan (0.2%), or with xyloligosaccharides (0.2%), or with isoprimeverose (0.2%) prepared as described below. Growth was monitored spectrophotometrically at 600 nm. In yeast extract plus sucrose medium, the generation time was about 6.5 h.

Substrates—Tamarind seed xylglucan was obtained from Megazyme, Ireland. Xyloligosaccharides were prepared by treatment of polymer with endo-glucanase (Megazyme) in sodium acetate, 25 mM, pH 5.0, for 24 h at 40 °C.

All commercially available substrates and Driselase were purchased from Sigma or Fluka. 4Np-β-isoprimeveroside and isoprimeverose were obtained by enzymatic synthesis as below described.

**Chromatographic Analysis and NMR Spectroscopy**—TLC was performed on 0.25-mm layers of Silica Gel F254 (Merck). Solvents included EtOAc/methanol/water (70:20:10, v/v) for aryl glycosides, acetone/butanol/water (60:20:10, v/v) for disaccharides, and EtOAc/methanol/water (47:40:13, v/v) for malto- and xyloligosaccharides and their cleavage products. Column chromatographic separations were carried out using

1 The abbreviations used are: ORF, open reading frame; COSY, correlation spectroscopy; RT, reverse transcriptase; PCR, polymerase chain reaction; PAGE, polyacrylamide gel electrophoresis; 4Np, 4-nitrophenyl; bp, base pair(s); kb, kilobase(s); nt, nucleotide(s); BRE, TFB-binding element; IPTG, isoprimeveroside; 4Np-β-galactopyranoside.

Siliaca Gel 60 (70–230 mesh; Merck). NMR spectra were recorded on a Bruker AMX 500 (500.13 MHz for 1H and 125.75 MHz for 13C) spectrometer: chemical shifts are given in ppm (δ) scale using solvent signals as internal standard.

**Enzymatic Syntheses of 4Np-β-isoprimeveroside and Isoprimeveroside**—4 µmol of 4Np-α-β-xyloligoside (9.2 mg) were dissolved in 1 ml of sodium acetate buffer, 50 mM, pH 5.0, and added to 67 µmol of 4Np-β-galactopyranoside. The reaction was started at 65 °C by addition of 0.2 mg of XylS. The reaction was complete after 2 h as indicated by complete disappearance of the donor as followed by TLC. The reaction mixture was rotary-evaporated and purified by preparative chromatography (chloroform/acetone, 9:1) of the disaccharide 4Np-β-isoprimeveroside. 1H and 13C NMR spectra in CDCl3/D2O of this compound show the following signals: δ 5.18 (7.0 Hz, H-1 β-Glc), 4.84 (J 3.5 Hz, H-1 α-Xyl); δ 101.2, 99.7, 77.4, 73.6, 74.9, 73.4, 73.2, 71.0, 70.9, 76.2, 62.6. After acetylation (pyridinium/ACO overnight at room temperature), the spectra of acetylated derivative in CDCl3 show the following signals: 1H NMR spectra δ 5.20 (H-1 β-Glc), 5.28 (H-2 β-Glc), 5.34 (H-3 β-Glc), 5.02 (H-4 β-Glc), 3.98 (H-5 β-Glc), 3.81–3.47 (H-6 β-Glc), 5.00 (H-1 α-Xyl), 4.80 (H-2 α-Xyl), 5.50 (H-3 α-Xyl), 4.91 (H-4 α-Xyl), 3.67–3.47 (H-5 α-Xyl). 13C NMR spectra, disaccharide moiety signals: δ 88.4, 71.3, 72.7, 69.1, 73.7, 66.5 (C1-C6 β-Glc); 95.9, 71.3, 69.0, 69.3, 58.7 (C1-C5 α-Xyl); δ 75.7 (c = 26.2, chloroform). Correlation spectroscopy (COSY) and 1H-13C correlation allowed assignments, as above indicated, of all carbon and proton signals in the acetylated derivative and secured about interglycosidic linkage of carbohydrate moieties. The free disaccharide isoprimeveroside was prepared by the action of Driselase on 4Np-β-isoprimeveroside; the product was purified and characterized by NMR spectroscopy after acetylation. Diagnostic signals in the 13C NMR spectra of acetylated disaccharide at 58.43 (C5 α-Xyl) and 66.24 (C-6 of glucose unit) ppm secured about no change in the carbohydrate sequence in the product. Free sugar is re-obtained after methanolyis (sodium carbonate in anhydrous methanol).
run, the gel was stained with Coomassie Blue. The molecular weight markers (Amersham Pharmacia Biotech) were phosphorylase b (94,000), bovine serum albumin (67,000), ovalbumine (43,000), and carbonic anhydrase (30,000). Protein samples are stable in these conditions for several months.

The enrichment of activity was calculated to be 148, 146, and 2 liters of culture yielded about 0.2 mg of pure protein. Protein concentrations were determined by the method of Bradford (24), with bovine serum albumin as standard.

**Enzymatic Assays**—Enzymatic assays were performed on di- oligo- and polysaccharides at the temperatures indicated in sodium acetate or phosphate buffers (50 mM, pH 5.5; measured at room temperature) by incubating the concentrations indicated of substrate and 5–10 μg of XylS in the final volume of 1.0 ml. The enzymatic reaction was linear for up to 30 min, and initial rates of hydrolysis were taken by stopping the reaction in dry ice after 15 min. The amount of glucose produced in the reaction was determined using a glucose oxidase-peroxidase system (GLU kit from Roche Molecular Biochemicals). Enzyme activity is given in micromoles of glycosyl bonds hydrolyzed per minute. The reaction products of malto- and xylotetraclucosacharides hydrolysis were analyzed by TLC.

Unless otherwise indicated, assays with 4-nitrophenyl-glycosides were performed at 65 °C in sodium acetate buffer (50 mM, pH 5.5), 10 mM substrates, and 5–10 μg of XylS. In these conditions the enzymatic reaction was linear for up to 30 min, and the initial rate of hydrolysis was determined by stopping the reaction with 0.8 ml of 1M sodium carbonate.

The amount of 4-nitrophenol released was measured by absorption at 405 nm in a 1-cm cuvette, considering a molar extinction coefficient of 18,300 M⁻¹ cm⁻¹. One unit of enzyme activity was defined as the amount of enzyme catalyzing the hydrolysis of 1 μmol of substrate in 1 min at the conditions described. Assay on 4Np-isoprimeveroside was performed by incubating the substrate (5 mM) in sodium phosphate buffer (50 mM, pH 5.5), in the presence of 2 μg of XylS and 1.2 μg of LacS (molar ratio 1:1). The determination of 4-nitrophenol released and the definition of enzyme units were as described above. Similar results were obtained at XylS/LacS ratios up to 1:6, indicating that the amount used of LacS was not rate-limiting.

Spontaneous hydrolysis of both chromogenic and nonchromogenic substrates resulted in about 1% of the absorbance obtained from the sample containing the enzymes and was subtracted by using blank mixtures without XylS and/or LacS.

Kinetic constants of XylS were measured at the optimal conditions of the different assays at 65 °C, by using substrate concentrations in the range 0.2–100 mM, 2.0–300 mM, and 0.2–7 mM for 4Np-glycosides/isoprimeverose, saccharides, and 4Np-β-isoprimeveroside, respectively. All kinetic data were calculated as the average of at least two experiments and were plotted and refined with the program GraFit (25).

**Enzyme Characterization**—Dependence on temperature was determined by assaying aliquots of homogenous enzyme (4 μg in the temperature range tested) on 100 mM maltose. Thermal stability was tested by incubating pure XylS (0.01 mg ml⁻¹) in sodium acetate buffer (50 mM, pH 5.5) at the indicated temperatures. At intervals, aliquots were withdrawn from the mixture and assayed at 65 °C following maltose (100 mM final concentration) hydrolysis.

Molecular mass of denatured XylS was determined by SDS-PAGE 7% in both reducing and non-reducing conditions, using the standards described above. Molecular mass of native XylS was determined by gel filtration on a Sephadex column 26/60 HiLoad (Amersham Pharmacia Biotech). Standard molecular weight markers were tyroglobulin (669,000), apoferritin (443,000), β-amylase (200,000), and bovine serum albumin (66,000).

**Reverse Transcriptase-PCR and Northern Blot Analysis**—S. solfataricus cells were grown at midexponential phase (0.6 A₅₅₀ in about 20 h from yeast extract plus sucrose medium. Cells were collected by centrifugation and lysated by three cycles of freeze thawing (2 min at −70 °C; 2 min at 30 °C) in TE buffer, and total RNA was extracted using the RNAeasy Kit (Qiagen, Germany). Contaminating DNA was digested by incubating the mixture and assayed at 65 °C following maltose (100 mM final concentration) hydrolysis.

The new ORF revealed low identity with enzymes from all the three living domains. In particular, the C terminus of XylS (roughly from positions 100 to 700), corresponding to a conserved domain in the ProDom data base (accession number PD001543), produced 35%, 30%, 29%, and 27% identity with the hypothetical protein of Erwinia herbicola (Erwglu), the α-glucosidase from S. solfataricus (MalA), the α-xylidosidas from L. pentosus (XylQ), and Thermotoga maritima (Tmxyl), respectively. Erwglu and Tmxyl enzymes have not been characterized biochemically, whereas MalA has been studied in detail and appears to be a typical α-glucosidase specific for maltose and malto-oligosaccharides (17). XylQ is an α-xylidosidase with high specificity for isoprimeverose (α-D-xylapyranosyl-1,6-α-D-glucopyranose); the xylQ gene is clustered with xylIP, encoding an hypothetical membrane protein transporter, in an operon that is involved in the metabolism of this disaccharide (9). The similarity between XylQ and XylS proteins led us to examine the flanking regions of xylS to find putative membrane protein transporters and to test the substrate specificity of the hypothetical enzyme (see below).

Surprisingly, the analysis of a 15-kilobase (kb) sequenced region, which comprises xylS, revealed that it mapped in the locus of the LacS gene encoding for the β-glycosidase that has been extensively studied in our laboratory (19, 21–23). As reported previously, lacS is preceded, in the same orientation, by a gene encoding a putative membrane protein homologous to the Major Facilitator Superfamily (mfs) genes in E. coli and in the opposite orientation, by a 1785-bp open reading frame (orf3) encoding a 592-amino acid polypeptide of unknown function. xylS maps immediately upstream to orf3 (Fig. 1), and the two genes are transcribed in the same direction. In the xylS promoter region, a consensus archaeal TATA box is present, separated 25 nt from a putative translation start site. A typical

---

2 The structures of the disaccharides used are: cellobiose, β-D-glucopyranosyl(1,4)-D-glucopyranose; isomaltose, α-β-D-glucopyranosyl-(1,6)-β-D-glucopyranose; isomaltose, α-D-maltopyranosyl-(1,6)-D-glucopyranose; maltose, α-D-glucopyranosyl-(1,4)-D-glucopyranose; sucrose, α-D-glucopyranosyl-(1,2)-D-fructofuranose; trehalose, α-D-glucopyranosyl-(1,1)-D-glucopyranose.
been described elsewhere (30).

One matched perfectly with entry P22795, a region of this ORF generated two entries exhibiting significant identity with orf3.

tides downstream, with 15 of 17 nt being identical (Fig. 1).

The transcription preinitiation complex in Archaea contains the TATA box, the BRE, and the translation start site of xylS and lacS genes are boxed. Identical nucleotides are shown in bold. In the consensus of the promoter regions R = A/G, Y = T/C, and W = A/T. Regulatory sequences of msf gene have been described elsewhere (30).

Fig. 1. Physical map of S. solfataricus glycosyl hydrolases locus and alignment of the xylS and lacS promoter regions. Putative regulatory regions, searched by homology with other archaeal genes, are shown as small arrows (promoter) and ball and stick (terminator). The probes used in the Northern blot analysis and the RT-PCR products for xylS and lacS are shown as thick lines. The RT-PCR products for orf3 and the region transversing xylS and orf3 are shown as thin lines. LacS is separated from orf3 by 20 and 2 nt, respectively; the end of xylS and the start of orf3 are separated by 27 nt. The TATA box, the BRE, and the translation start site of xylS and lacS genes are boxed. Identical nucleotides are shown in bold. In the consensus of the promoter regions R = A/G, Y = T/C, and W = A/T. Regulatory sequences of msf gene have been described elsewhere (30).

Expression and Characterization of XylS—The xylS gene was amplified from S. solfataricus (strain MT4) genomic DNA and was cloned in the vector pT7-SCII. The pT7-SCII-derived plasmid pXyl was used to express the protein in E. coli BL21(DE3) in the presence of 1 mM IPTG. Purification of recombinant XylS a

The thermal activity of XylS is reported in Fig. 3A. The activity on maltose increased sharply up to the optimal temperature of 95 °C, whereas a slight decrease was observed at 95 °C, the highest temperature tested. This behavior led to a discontinuity in the Arrhenius plot at 85 °C; for this reason, the activation energy (E_a) for this substrate (89.5 ± 3 kJ mol⁻¹) was calculated from the slope obtained in the temperature range 55–85 °C. XylS activity showed a sharp dependence on pH when assayed at 70 °C on maltose with a maximum at pH 5.5 in 50 mM sodium acetate or phosphate buffers (not shown). All following characterizations were performed at pH 5.5 in

The thermal activity of XylS is reported in Fig. 3A. The activity on maltose increased sharply up to the optimal temperature of 95 °C, whereas a slight decrease was observed at 95 °C, the highest temperature tested. This behavior led to a discontinuity in the Arrhenius plot at 85 °C; for this reason, the activation energy (E_a) for this substrate (89.5 ± 3 kJ mol⁻¹) was calculated from the slope obtained in the temperature range 55–85 °C. XylS activity showed a sharp dependence on pH when assayed at 70 °C on maltose with a maximum at pH 5.5 in 50 mM sodium acetate or phosphate buffers (not shown). All following characterizations were performed at pH 5.5 in

Table I

| Purification step | Protein Activity | Specific activity | Purification | Yield |
|-------------------|-----------------|------------------|--------------|-------|
| Cell extract      | 520             | units/mg         | 0.18         | 1.0   | 100 |
| 55 °C             | 244             | 0.18             | 1.0          | 100   |
| 65 °C             | 134             | 0.32             | 1.8          | 98    |
| 75 °C             | 98              | 0.40             | 2.2          | 89    |
| Mono Q            | 16              | 1.79             | 9.9          | 66    |
| Sephadex          | 12              | 2.43             | 13.5         | 65    |

a From 16 g of bacterial pellet.
b Assays were performed in sodium acetate, 50 mM buffer, pH 5.5, at 75 °C, by using maltose (100 mM) as substrate.
c Host contaminating activity prevented measurements of XylS activity in cell-free extracts.
Fig. 3. A, thermal activity and derived Arrhenius plot (inset) of XyI S on maltose. B, thermal stability of XyI S at 90 °C (open circles), 92 °C (closed circles), 95 °C (open squares), 97 °C (closed squares), and 100 °C (open triangles). The derived Arrhenius plot is reported in the inset.

both buffer systems.

The residual activity of XyIS on maltose at 65 °C after preincubation at temperatures between 90 °C and 100 °C was followed for up to 2 h (Fig. 3B). Thermal inactivation obeyed first order kinetics at all the temperatures tested: as expected for proteins from hyperthermophiles, the enzyme displayed high stability with a half-life of 38 h at 95 °C, inactivation occurred with coincident aggregation, even at the low protein concentration used (0.01 mg ml⁻¹), resulting in shorter half-lives at 97 °C and 100 °C (6 and 2 min, respectively). The $E_	ext{act}$ for XyIS inactivation, calculated from the Arrhenius plot (inset in Fig. 3B) was 813 kJ mol⁻¹, almost 10-fold higher than the activation energy obtained for the catalyzed reaction on maltose.

**XyIS Promotes Transglycosylation Reactions**—The above general characterization indicated that XyIS efficiently hydro-

lyzes 4Np-α-xyloside. We used this specificity as an advantage to use the enzyme in transglycosylation mode for the synthesis of the disaccharide isoprimeverose (α-D-xylopyranosyl-(1,6)-D-glucopyranose). As described above, family 31 α-glycosyl hydrolases follow a retaining reaction mechanism. This mechanism follows two steps: in the first one the enzyme catalyzes the departure of the aglycon group from the substrate (donor) and the consequent formation of a glycosyl ester intermediate. In the second step, the enzyme is deglycosylated by a nucleophile (acceptor), which attacks the anomic carbon of the donor and cleaves the covalent intermediate leading to the overall retention of the anomeric configuration of the substrate. When a nucleophile different from water intercepts the glycosyl enzyme intermediate, transglycosylation occurs producing glycosylated products.

XyIS was incubated at 65 °C with 4Np-α-xyloside and 4Np-β-glucoside as donor and acceptor, respectively. After 2 h, a disaccharide of 4Np was obtained as the main product. The purification and the characterization of this product by NMR allowed its identification as 4Np-β-isoprimeveroside. By incubation with Driselase, the β-bond of the disaccharide was cleaved and the free disaccharide isoprimeverose was obtained.

These experiments represented a new method for the preparation of isoprimeverose demonstrating that XyIS can be used in transglycosylation mode. Furthermore, the α-anomeric configuration of the product unequivocally indicated that the enzyme followed a retaining mechanism.

**Substrate Specificity of XyIS**—The activity of XyIS on several substrates is reported in Table II. The enzyme revealed clear selectivity for xylose-containing substrates; in particular, the highest activity was found on the disaccharide isoprimeverose. Similar substrate specificity was found at both 65 °C and 85 °C suggesting that at these temperatures no major conformational changes occur in the XyIS active site. The activity on 4Np-β-isoprimeveroside and maltose (α-D-xylopyranosyl-(1,6)-D-glucopyranose) was lower and completely absent on isomaltose (α-D-glucopyranosyl-(1,6)-D-glucopyranose), trehalose (α-D-glucopyranosyl-(1,1)-D-glucopyranose), and sucrose (α-D-glucopyranosyl-(1,2)-D-fructofuranose) (not shown). These findings were confirmed with 4Np-α-glycosides substrates: the activity of XyIS at 65 °C on 4Np-α-D-glucoside is only 1% of that found on 4Np-α-D-xyloside, whereas 4Np-α-D-galactoside, α-D-arabinoside, α-L-rhamnose, α-D-mannoside, and α-L-fucoside were not substrates of the enzyme (not shown).
To test whether XylS could hydrolyze malto-oligosaccharides and to define the mode of action of the enzyme (endo- versus exo-acting), we followed the hydrolysis of maltodextrins ranging from two up to five glucose residues. Although the rates of hydrolysis decrease with the length of the substrates tested (Table II), XylS efficiently hydrolyzed maltotriose and 4Np-α-maltotriose. The analysis of the reaction mixtures by TLC revealed, upon the partial hydrolysis of maltotriose, -tetraose, and -pentaose substrates, the formation of glucose and maltooligosaccharides shortened of one unit. Moreover, when the 4Np-α-maltotriose was used as substrate, its hydrolysis produced glucose and 4Np-α-glucose, whereas 4-nitrophenol was released only after several minutes from the start of the reaction. These results demonstrated that XylS is an exo-acting enzyme that attacks the non-reducing end of the substrate.

To evaluate the activity of XylS on polysaccharides, the enzyme was incubated at 65 °C with starch, amylose, glycogen, pullulan, and xyloglucan from tamarind seeds. No glucose release was observed by glucose oxidase-peroxidase analysis, and no monosaccharides were detected by TLC. This was not surprising: all the α-xylanoligosides identified so far are inactive on xyloglucans and on polysaccharides containing mainly α-bonds. However, when XylS was incubated at 65 °C in the presence of xyloglucan oligosaccharides from tamarind seeds xyloglucan, no glucose could be detected, but the formation of xylose was observed by TLC (Table II). Presumably, XylS, as observed for the α-xylanoligosidase from L. pentosus (9), recognized isoprimeverose units at the non-reducing end of xyloglucan fragments and promoted the release of the xylose residues (see below).

We determined the kinetic parameters of XylS for the best substrates (Tables III; the highest catalytic efficiency values were obtained with 4Np-β-isoprimeveroside and isoprimeverose, confirming the high specificity of the enzyme for these substrates. It is interesting that maltotriose is hydrolyzed at rates comparable to those of maltose, but with higher efficiency, because the enzyme showed increased affinity for the former. Moreover, despite the high $K_m$ value, 4Np-α-xyloligosaccharides are by far the best 4Np-α-glucoside hydrolyzed by the enzyme.

These results strongly indicate that XylS is a true α-xylanoligosidase, highly specific for isoprimeverose and exo acting on xyloglucan oligosaccharide substrates.

### TABLE III

**Kinetic constants of XylS at 65 °C**

| Substrates     | $K_m$ (mM) | $k_{cat}$ (s⁻¹) | $k_{cat}/K_m$ (s⁻¹ mM⁻¹) |
|---------------|------------|-----------------|-------------------------|
| Maltose       | 17.0 ± 3.2 | 1.51 ± 0.07     | 0.09                    |
| Maltotriose   | 3.45 ± 0.97| 0.92 ± 0.04     | 0.27                    |
| Isoprimeverose| 28.9 ± 3.5 | 31.0 ± 1.5      | 1.07                    |
| 4Np-β-isoprimeveroside | 1.72 ± 0.46 | 16.0 ± 1.6     | 9.30                    |
| 4Np-α-xylotriose | 17.0 ± 2.1 | 4.89 ± 0.27     | 0.28                    |
| 4Np-α-glucose  | 2.05 ± 0.44| 0.05 ± 0.00     | 0.02                    |

---

α-Xylosidase from *S. solfataricus*

**Fig. 4. Model of hydrolysis of xyloglucan oligosaccharides by XylS and LacS.** A, typical mixture of xyloglucan oligosaccharides according to Vincken et al. (35) is reported following the nomenclature of Fry et al. (36). To each β-(1,4)-linked α-glucosyl residue in the backbone is given a one-letter code according to its substituents: $G$, glucose residue; $X$, isoprimeverose residue (α-D-xylopyranosyl-(1,6)-β-glucopyranose); $L$, galactosyl-substituted isoprimeverose residue (β-D-galactosyl-(1,2)-α-D-xylopyranosyl-(1,6)-β-D-glucopyranose). Sequences are always written from the non-reducing end of the molecule. Thus, in the sequence XXXG, three isoprimeverose residues form β-1,4-glucosidic bonds among each other and with a glucose that has its reducing end free. **B**, products of action of XylS. The hydrolysis of the isoprimeverose residues at the non-reducing end releases xylose. Xyloglucan oligosaccharides, with glucose moiety at the non-reducing end, become the substrate of LacS. **C**, products of action of LacS. The hydrolysis of the β-(1,4)-linked β-glucosyl residue releases glucose and xyloglucan oligosaccharides shortened of one unit, allowing future attack by XylS.

**Expression Levels of xylS and Surrounding Genes in S. solfataricus**—To test the expression levels of xylS and surrounding genes in vivo, total RNA was extracted from *S. solfataricus* cells grown on yeast extract with added sucrose medium in exponential growth and analyzed by RT-PCR and Northern blot. Isoprimeverose and xyloglucan, neither in polymeric nor in oligosaccharidic form, could support growth.

RT-PCR revealed the presence of lacS, orf3, and xylS transcripts (Fig. 5A). Moreover, a longer transcript, bridging orf3 and xylS coding regions, was clearly visible, suggesting that the two genes are cotranscribed. Northern blot analysis (Fig. 5B) confirmed these results indicating that both xylS and orf3 genes were actively transcribed in vivo. The length of the lacS transcript was of the expected dimensions (1.6 kb), whereas the probes relative to both xylS and orf3 genes revealed an identical signal corresponding to a transcript of about 4.6 kb. However, with the xylS probe, a second transcript of a size compatible with the xylS gene alone (2.2 kb) was visible as well. These results strongly indicate the linkage of xylS and orf3 genes in an operon and suggest that xylS-independent expression could be controlled by the regulatory sequences flanking the gene, which are missing in orf3 (Fig. 1). The lacS gene shows two transcription initiation sites mapping −9 and +1 with respect to the A of the translational start codon (30). The high similarity between xylS and lacS promoter regions (Fig. 1) could suggest similar transcription initiation sites also for the xylS gene.
In the framework of our mechanistic studies on thermophilic glycosyl hydrolases (37), the reaction mechanism of XylS was analyzed by testing its ability to function in transglycosylation mode. This approach allowed us to demonstrate experimentally that XylS follows the retaining mechanism catalyzing the formation of the α-(1,6) bond between xylose and glucose and represented a new method for the enzymatic synthesis of isoprimeverose, which was previously obtained only from natural sources (9). The high specificity constant of XylS for 4Np-isoprimeverose, which is itself a substrate of the enzyme, explains the high specificity and the relatively low yields of the synthetic reaction.

Two aspartic acid residues were found to be involved in catalytic activity in family 31 glycosyl hydrolases: residues Asp-505 and Asp-1394 for the sucrase-isomaltase (the enzyme has two homologous active sites) and Asp-518 in the human lysosomal α-glucosidase were identified by affinity labeling with conduritol B epoxide and by site-directed mutagenesis (38, 39). These aspartic acid residues and a highly conserved glutamic acid residue fall in the PROSITE consensus motif (G/F)(L/I/V/M/F)WxDM(NS/A/E), a hallmark of enzymes from family 31, suggesting their involvement in catalysis. The corresponding residues in XylS are Asp-353 and Glu-356. It has been suggested that the aspartic acid residues identified in sucrase-isomaltase and lysosomal α-glucosidase constitute the catalytic nucleophile in this family. However, as already noticed, the assignment of the role played in catalysis by these residues remains equivocal (40). In fact, conduritol epoxide derivatives occasionally have labeled active site residues different from the nucleophile, and the levels of residual enzymatic activity of the human lysosomal α-glucosidase mutant Asp-518 → Asn was incompatible with the essential function proposed (39). For these reasons, the nucleophile of the reaction of family 31 enzymes remains to be identified unequivocally by accurate kinetic studies of site-directed mutants and by using more specific mechanism-based inactivators.

The present study for the first time describes an enzyme from an Archaeon potentially involved in xyloglucan degradation. Screening the completed genome sequences of hyperthermophilic Archaea (Methanococcus jannaschii, Archaeoglobus fulgidus, Aquifex aeolicus, Pyrococcus abyssi, P. horikoshii, and P. furiosus) did not result in any significant homology. In the hyperthermophilic Bacterium T. maritima, however, a putative α-xyllosidase was found (Tmxyl; entry TM0308) that shares homology with XylS and XylQ. It is interesting that this putative α-xyllosidase is part of a cluster of six ORFs, including an endo-glucanase, an α-1-fucosidase, an 1-fucose isomerase, an oligopeptide ABC transporter, and a β-galactosidase (entries TM0305, TM0306, TM0307, TM0309, and TM0310, respectivel (41). The functions assigned to these ORFs suggest their possible involvement in xyloglucan utilization, considering that α-1-fucosyl groups are frequently found in xyloglucans from plant primary cell walls. The characterization of the products of these genes will clarify their function in vivo.

The present identification of xylS gene and the substrate specificity of its gene product strongly suggests involvement of this enzyme in the degradation of di- and oligosaccharides containing α-1,6-linked xylose, which are the building blocks of xyloglucan. Moreover, the cooperation of XylS and LacS in the degradation of xyloglucan oligosaccharides in vitro and the vicinity of the encoding genes on the S. solfataricus chromosome some could suggest that the two enzymatic systems are functionally related also in vivo. However, the complex structure of xyloglucan would require the combined action of several enzymatic activities and protein transporters for its efficient hydrolysis and assimilation. No evidence of such systems was found.

**DISCUSSION**

Among the newly available data from the S. solfataricus genome-sequencing project, a gene encoding for an α-xyllosidase has been identified. The selectivity of the enzyme for xylose containing substrates such as 4Np-α-xylloside, isoprimeverose, and xyloglucan oligosaccharides validates the high similarity found with α-xyllosidase enzymes. The amino acid sequence assigned XylS to family 31 of glycosyl hydrolases, a group of α-glucosidases, glucoamylases, and sucrase-isomaltases from Archaea and Eukarya and that have been sequenced so far all belong to this family, whereas bacterial glucoamylases, α-amylases, and xyloglucan hydrolases are grouped in families 13 and 15. The archaeal α-xyllosidase described here demonstrates that this enzymatic activity is present in all the three living domains. This finding, and the similarity with the enzymes of family 31, suggest that α-xyllosidases and eukaryal isomaltases could have evolved from a common ancestor, whereas bacterial glucoamylases, α-amylases, and α-glucosidases are evolutionarily unrelated to family 31 enzymes. As previously noticed, it is not simple to explain the lack of homology between bacterial α-glucosidases and α-xyllosidases (9) and the presence of hypothetical α-glucosidase enzymes from Alicyclobacillus acidocaldarius, Bacillus thermoamylolyticus, and E. coli, in family 31. Nevertheless, it is not known if these enzymes are indeed α-glucosidases or α-xyllosidases, because they have never been characterized. Data on the substrate specificity of these enzymes would be useful to clarify their evolutionary link to the rest of the family.
in *Sulfolobus* so far. In *S. solfataricus* strain MT4, neither isoprimeverose or oligosaccharidic or polymeric xyloglucans could support growth as minimal carbon sources after 1 week. However, it remains to be established whether this strain can utilize by-products of xyloglucan degradation after regulatory or mutational adaptation as observed on the disaccharides cellobiose and maltose (15).

The analysis of the gene expression revealed that *xylS* is cotranscribed in *vivo* with the orf3 gene, indicating the linkage of these two genes in an operon. Interestingly, a transcript of the dimensions expected for *xylS* gene alone was also observed. This could suggest that *xylS* gene could be expressed also as a single transcriptional unit exhibiting a canonical terminator sequence located 3 nt from the *xylS* stop codon. Unfortunately, the lack of clear homologies found for orf3 made it difficult to assign a functional role to this putative protein. The results of the lack of clear homologies found for orf3 made it difficult to assign a functional role to this putative protein. The results of the lack of clear homologies found for orf3 made it difficult to assign a functional role to this putative protein.

**Acknowledgments**—We thank the staff of Istituto per la Chimica di Molecole di Interesse Biologico-CNR NMR service (S. Zambardino and V. Mirra) for NMR spectra, and G. Imperato and O. Piedimonte for technical assistance. M. Ciaramella, B. Di Lauro, and the other members of M. Moracci’s laboratory are gratefully acknowledged for their comments and support.

**REFERENCES**

1. Crombie, H. J., Chengappa, S., Hellyer A., and Reid, J. S. G. (1998) *Plant J.* 15, 27–38
2. Beguin, P., and Aubert, J. (1994) *FEMS Microbiol. Rev.* 13, 25–58
3. Warren, R. A. J. (1996) *Annu. Rev. Microbiol.* 50, 183–212
4. Matsushita, J., Kato, Y., Pe´rez Lorences, E., Maclachlan, G. A., McNeil, M., Mort, A. J., Reid, J. S. G., Seitz, H. U., Selvendran, R. R., Voragen, A. G. J., and White, A. J. J. (1991) *J. Bacteriol.* 177, 1614–1619
5. Bell, S. D., Koss, P. L., Sigler, P. B., and Jackson, S. P. (1996) *Proc. Nat. Acad. Sci. U. S. A.* 93, 1353–1361
6. Zillig, W., Stetter, K. O., Wunderl, S., Schulz, W., Priess, H., and Schlotz, I. (1980) *Arch. Microbiol.* 125, 259–269
7. York, W. S., Van Halbeek, H., Darvill, A. G., and Albersheim, P. (1990) *Carbohydr. Res.* 200, 9–31
8. Vincken, J., de Keizer, A., Klemm, G., and Voragen, A. G. J. (1995) *Plant Physiol.* 109, 1579–1585
9. Fry, S. C., York, W. S., Albersheim, P., Darvill, A., Hayashi, T., Joselep, J., Kato, Y., Perez Lorenes, E., Maldachlan, G. A., McNeil, M., Mort, A. J., Reid, J. S. G., Seitz, H. U., Selvendran, R. R., Voragen, A. G. J., and White, A. R. (1993) *Plant Physiol.* 89, 1–3
10. Moracci, M., Trincone, A., Perugino, G., Ciaramella, M., and Rossi, M. (1998) *Biochemistry* 37, 17262–17270
11. Quaroni, A., and Semenza, G. (1976) *J. Bacteriol.* 125, 321–328