Regulation of Endo-Acting Glycosyl Hydrolases in the Hyperthermophilic Bacterium *Thermotoga maritima* Grown on Glucan- and Mannan-Based Polysaccharides

Swapnil R. Chhabra, Keith R. Shockley, Donald E. Ward, and Robert M. Kelly*

Department of Chemical Engineering, North Carolina State University, Stinson Drive, Box 7905, Raleigh, North Carolina 27695-7905

Received 19 July 2001/Accepted 12 November 2001

The genome sequence of the hyperthermophilic bacterium *Thermotoga maritima* encodes a number of glycosyl hydrolases. Many of these enzymes have been shown in vitro to degrade specific glycosides that presumably serve as carbon and energy sources for the organism. However, because of the broad substrate specificity of many glycosyl hydrolases, it is difficult to determine the physiological substrate preferences for specific enzymes from biochemical information. In this study, *T. maritima* was grown on a range of polysaccharides, including barley β-glucan, carboxymethyl cellulose, carob galactomannan, konjac glucomannan, and potato starch. In all cases, significant growth was observed, and cell densities reached 10⁷ cells/ml. Northern blot analyses revealed different substrate-dependent expression patterns for genes encoding the various endo-acting β-glycosidases; these patterns ranged from strong expression to no expression under the conditions tested. For example, *cel74* (TM0305), a gene encoding a putative β-specific endoglucanase, was strongly expressed on all substrates tested, including starch, while no evidence of expression was observed on any substrate for *lam16* (TM0024), *xyl10A* (TM0061), *xyl10B* (TM0070), and *cel12A* (TM1524), which are genes that encode a laminarase, two xylanases, and an endoglucanase, respectively. The *cel12B* (TM1525) gene, which encodes an endoglucanase, was expressed only on carboxymethyl cellulose. An extracellular mannanase encoded by *man5* (TM1227) was expressed on carob galactomannan and konjac glucomannan and to a lesser extent on carob cellulose. An unexpected result was the finding that the *cel5A* (TM1751) and *cel5B* (TM1752) genes, which encode putative intracellular, β-specific endoglucanases, were induced only when *T. maritima* was grown on konjac glucomannan.

To investigate the biochemical basis of this finding, the recombinant forms of *Man5* (*M₅*, 76,900) and *Cel5A* (*M₅*, 37,400) were expressed in *Escherichia coli* and characterized. *Man5*, a *T. maritima* extracellular enzyme, had a melting temperature of 99°C and an optimum temperature of 90°C, compared to 90 and 80°C, respectively, for the intracellular enzyme *Cel5A*. While *Man5* hydrolyzed both galactomannan and glucomannan, no activity was detected on glucans or xylans. *Cel5A*, however, not only hydrolyzed barley β-glucan, carboxymethyl cellulose, xyloglucan, and lichenin but also had activity comparable to that of *Man5* on galactomannan and higher activity than *Man5* on glucomannan. The biochemical characteristics of *Cel5A*, the fact that *Cel5A* was induced only when *T. maritima* was grown on glucomannan, and the intracellular localization of *Cel5A* suggest that the physiological role of this enzyme includes hydrolysis of glucomannan oligosaccharides that are transported following initial hydrolysis by extracellular glycosidases, such as *Man5*.
49, 55). In some cases, expression of genes associated with individual glycosyl hydrolases has been examined (7, 41). However, how T. maritima orchestrates collective expression of functional subsets of the glycosyl hydrolase inventory in response to the presence of various saccharides in the environment is less clear. Furthermore, because of the broad substrate specificity of many glycosyl hydrolases, it is difficult to determine from genome sequence information which of these enzymes are needed for synergistically breaking down certain complex polysaccharides into their monosaccharide components for subsequent use in central metabolism. To study this, transcription of the genes encoding endo-acting glycosidases in T. maritima was examined by using various complex carbohydrates as primary carbon and energy sources. Among the findings of this study was the fact that while certain genes were expressed on all of the growth substrates examined, other genes were induced only in the presence of specific substrates. Furthermore, the physiological functions of glycosyl hydrolases were determined best by taking into account the results of sequence analyses and biochemical characteristics in conjunction with gene expression information.

**MATERIALS AND METHODS**

**Growth of microorganisms.** T. maritima cells were grown anaerobically at 80°C on artificial seawater supplemented with 0.1% (wt/vol) yeast extract and 0.5% (wt/vol) K2HPO4, 50 mg of KBr, 20 mg of H3BO3, 20 mg of KI, 3 mg of Na2WO4, 2.0 g of MgCl2·6H2O, 0.25 g of NaHCO3, 0.10 g of CaCl2·2H2O, 0.50 g of CaCl2·H2O, 0.25 g of NaH2PO4, 0.10 g of K2HPO4, 50 mg of KBr, 20 mg of H2BO3, 0 mg of K2, 3 mg of Na2WO4·2H2O, and 2 mg of NiCl2·6H2O. K2HPO4 was added after sterilization.


**TABLE 1. Hyperthermophilic glycosyl hydrolases**

<table>
<thead>
<tr>
<th>Type of glycosidase</th>
<th>P. furiosus</th>
<th>P. horikoshi</th>
<th>T. maritima</th>
<th>Aquifex aeolicus</th>
<th>Methanococcus jannaschii</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cellulase Cel1, Cel12</td>
<td>Cell1, Cel15</td>
<td>Cell3, Cel5A, Cel5B, Cel12A, Cel12B, Cel17</td>
<td>Xyl3, Xyl10A, Xyl10B, Xyl31</td>
<td>Amy4A, Amy4B, Amy4C, Amy4D, Amy13A, Amy13B</td>
<td>Amy57, Amy15, Amy57</td>
</tr>
<tr>
<td>Xylanase Man1 Lam16</td>
<td>Man1, Man38</td>
<td>Amy57</td>
<td>Amy57, Amy15A, Amy15B</td>
<td>Amy57</td>
<td>Amy15, Amy57</td>
</tr>
</tbody>
</table>

The following proteins are putative proteins: GH35 in *P. furiosus*; Cel1, Cel5, Man1, Man38, Amy57, Gal57 and amy37 in *P. horikoshi*; Cel1B, Xyl2, Man38A, Man38B, Lam16, Amy4B, Amy4C, Amy4D, Amy13B, Gal42B, Gal53, GH2, GH23A, GH23B, GH28, GH29, GH51, and GH73 in *T. maritima*; Cel8, Amy57, GH23A, GH23B, and GH77 in *A. aeolicus*; and Amy15 in *M. jannaschii*. No hyperthermophilic glycosyl hydrolases have been found in *A. fulgidus*.
starch, *Laminaria digitata* laminarin, and birch wood xylan were obtained from Sigma (St. Louis, Mo.). Guar (Uniguar 150) was obtained from Rhodia (Washington, Pa.). The ratio of D-mannose residues to D-galactose residues in carob galactomannan (3.5:1) was higher than the ratio in guar galactomannan (2:1).

The chromogenic polysaccharides carob azo-galactomannan and azo-carboxymethyl cellulose were obtained from Megazyme. All polysaccharides except carob β-mannan were dissolved in water at a final concentration of 10 mg/ml (1%, wt/vol) as recommended by the manufacturer. β-Mannan was dissolved at a concentration of 10 mg/ml in 10% sodium hydroxide and neutralized with 50% acetic acid. The chromogenic substrates were prepared as described previously (12). Unless indicated otherwise, enzyme assays were done in triplicate at 30°C (Cel5A) or at 90°C (Man5) in 0.5-ml reaction mixtures containing 50 mM sodium phosphate buffer (pH 7.0) or 0.8% (wt/vol) solutions of soluble polysaccharide substrates. The standard deviations for triplicate assays were less than 10%. For other polysaccharides, enzymatic activity was measured according to the nomenclature (12).

Estimation of temperature and pH optima. The temperature dependence of Cel5A was determined by measuring the specific activity of the enzyme with a 0.8% (wt/vol) barley β-glucan solution in 100 mM sodium phosphate buffer (pH 7.0) at various temperatures. Similarly, the temperature dependence of Man5 was determined by measuring the specific activity of the enzyme with 50 mM sodium phosphate buffer (pH 7.0) at various temperatures. The pH dependence of both enzymes was investigated by determining the specific activities of the enzymes with their substrates at pH values between 3.6 and 5.6 with 100 mM sodium acetate buffer, at pH values between 5.4 and 8.4 with 100 mM sodium phosphate buffer, and at pH values between 8.6 and 10.0 with 100 mM glycine–NaOH. Thermostability was determined by incubating the purified enzymes for various lengths of time at 80°C or 90°C in 100 mM sodium phosphate buffer (pH 7.0) and determining the residual activity.

Determination of enzyme kinetics. Kinetic parameters were determined under optimal conditions for both enzymes by using barley β-glucan as the substrate for Cel5A and carob galactomannan for Man5. Reaction rates were determined for substrate concentrations that ranged from approximately 0.4 to 6.0 times the Km. Km and Vmax values were determined from these rates by performing a nonlinear regression analysis with DataFit (Oakdale Engineering, Oakdale, Pa.).

Differential scanning microcalorimetry. Melting temperatures for Man5 and Cel5A were determined with a NanoDifferential scanning calorimeter (Calorimetry Sciences, Salt Lake City, Utah). Both enzymes were dialyzed against 10 mM sodium phosphate buffer (pH 7.0). A sample added to a cell was maintained at a pressure of 3.0 atm to allow operation at temperatures greater than 100°C. The dialed enzymes were scanned at temperatures ranging from 25 to 125°C by using a scan rate of 1°C/min. Enzyme scans were corrected with a buffer-buffer baseline.

TABLE 2. Oligonucleotides used in this study

<table>
<thead>
<tr>
<th>Locus</th>
<th>Glycosyl hydrolyase gene</th>
<th>Primer sequences</th>
<th>Probe length (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TM0024</td>
<td>lam16</td>
<td>5′-AACGGCTGTCTTGTAATTGAGG-3′, 5′-GAATGGAAAGAACGGATCTG-3′</td>
<td>493</td>
</tr>
<tr>
<td>TM0061</td>
<td>xyl10A</td>
<td>5′-ACGGAGGTCACAGAGAAACGG-3′, 5′-TGGCTCAACTATGATCTCCG-3′</td>
<td>627</td>
</tr>
<tr>
<td>TM0070</td>
<td>xyl10B</td>
<td>5′-GGTACAGAATTGCAAGAAGAA-3′, 5′-ATACCTTCACGAACATGGAG-3′</td>
<td>465</td>
</tr>
<tr>
<td>TM0305</td>
<td>cel74</td>
<td>5′-TGAGCAACATCATATGGGAAACCGGG-3′, 5′-GAAAGGAAGATTTCTGATG-3′</td>
<td>2064</td>
</tr>
<tr>
<td>TM1227</td>
<td>man5</td>
<td>5′-TTCACGTTCCTCACATAGTGGACAGATTGCG-3′, 5′-AGTGAAGGTCATATTCACTATCG-3′</td>
<td>1956</td>
</tr>
<tr>
<td>TM1524</td>
<td>cel12A</td>
<td>5′-CGAACCGTCTCCTGCTTAAGAC-3′, 5′-TTCACACTTCACACCGAACGTG-3′</td>
<td>427</td>
</tr>
<tr>
<td>TM1525</td>
<td>cel12B</td>
<td>5′-AGGTTGTTCCTTACAGGCGGTG-3′, 5′-ITCGCCCGTCTTGTGTTG-3′</td>
<td>724</td>
</tr>
<tr>
<td>TM1751</td>
<td>cel5A</td>
<td>5′-TGATCTTGTATATGCGGATTTATCATGTGTTGATGCC-3′, 5′-AGCCCATCTGATCCGAGTTGACT-3′</td>
<td>951</td>
</tr>
<tr>
<td>TM1752</td>
<td>cel5B</td>
<td>5′-CTCAAGAAGCGTTATCCTACATCG-3′, 5′-CCCTACTATCCACCCATTCCG-3′</td>
<td>324</td>
</tr>
</tbody>
</table>

Downloaded from http://aem.asm.org/ on August 15, 2017 by guest

RESULTS

Analysis of the *T. maritima* genome with respect to glycosyl hydrolases. Examination of the genomic sequences of hyperthermophiles has shown that the glycosyl hydrolases (glycosidases) are widely distributed in these organisms. To date, hyperthermophilic glycosyl hydrolases have been found in at least 34 of the 77 known families (families 1 to 5, 8 to 13, 15, 16, 18, 26, 28, 29, 31, 32, 35, 36, 38, 39, 42 to 44, 48, 51, 53, 57, 65, 67, 74, and 77) (26). Table 1 shows the distribution of these enzymes in various hyperthermophilic microorganisms; glycosyl hydrolases have been classified according to the nomenclature scheme suggested by Henriass et al. (27). Since many hyperthermophilic organisms utilize complex carbohydrates as carbon and energy sources, it is apparent that multienzyme systems are needed to hydrolyze polysaccharides that are too large to be transported across the cell membrane (3, 16). As a result, many endo-acting glycosidases are cell membrane associated or completely secreted (14). Oligosaccharides with various degrees of polymerization produced by endo-acting enzymes are transported into the cell for further processing (14). Once the oligosaccharides are inside the cell, other endo-acting and exo-acting glycosyl hydrolases within the cytoplasm play essential roles in assimilation and catabolism of these compounds to provide saccharides (e.g., glucose and galactose) to metabolic pathways (14, 31). From this perspective, putative polysaccharide-degrading enzyme systems can be identified from genome sequence data on the basis of the type of glycosidic linkages potentially cleaved by specific glycosyl hydrolases. For instance, in the case of *T. maritima*, glycosidases involved in the degradation of starch (α-glucan) could potentially include two endo-acting enzymes, Amy13A (TM1840) and Amy13B (TM1650), along with exo-acting enzymes, such as...
as Amy4A to Amy4D (TM1834, TM1068, TM0752, and TM0434). Table 3 shows putative polysaccharide-degrading enzyme systems in *T. maritima* associated with degradation of glycosides containing the following linkages: \( /H_9251 \)-glucan, \( /H_9252 \)-glucan, \( /H_9252 \)-xylan, and \( /H_9252 \)-mannan. Table 3 also shows sequences of identifiable N-terminal signal peptides that are associated with some of the endo-acting glycosidases in *T. maritima*. The data indicate that this organism is capable of exporting endo-acting glycosyl hydrolases for degradation of starch and pullulan (Amy13A, Pul13), \( /H_9252 \)-glucan (Cel12B, Cel74), laminarin (Lam16), xylan (Xyl10A, Xyl10B), and mannan (Man5). It is interesting that none of the identifiable exo-acting glycosyl hydrolases have signal peptides, suggesting that oligosaccharides must be imported prior to processing to monosaccharides. Also, there are a number of intracellular endo-acting glycosyl hydrolases (Amy13B, Cel12A, Cel5A, and Cel5B) that could be involved in hydrolysis of transported oligosaccharides. However, it is not clear how *T. maritima* uses its glycosyl hydrolase inventory strategically to recruit carbohydrates as carbon and energy sources or the extent to which particular enzymes are utilized. In particular, the functions of Cel5A and Cel5B, which have no apparent extracellular counterparts, are difficult to assess based on genome sequence information.

**Growth of *T. maritima* on complex carbohydrates and induction of endo-acting glycosidase genes.** To investigate the potential roles of various endo-acting glycosyl hydrolases in the utilization of polysaccharides, *T. maritima* was grown on a range of substrates that should contain both carbon and energy sources based on the information in Table 3. These substrates included konjac glucomannan, a linear polysaccharide containing \( /H_9252 \)-1,4-linkedD-mannopyranose and D-glucopyranose units as backbone residues; carob galactomannan, a heteropolysaccharide consisting of a \( /H_9252 \)-linkedD-mannopyranose and D-glucopyranose residues linked by two types of linkages (\( /H_9252 \)-1,4 and \( /H_9252 \)-1,3); carboxymethyl cellulose, a soluble form of cellulose (made by chloroacetic acid treatment of cellulose) and a linear homopolysaccharide made up of \( /H_9252 \)-1,4-linkedD-glucopyranose residues; and starch, a mixture of the linear \( /H_9251 \)-1,4-linkedD-glucanpyranose homopolysaccharide amylose (15 to 25%) and branched amylopectin (75 to 85%) containing \( /H_9251 \)-1,6-glycosidic linkages (at every 17 to 26 glucose residues) in addition to the \( /H_9251 \)-1,4 bonds. The results of batch growth experiments at 80°C are shown in Fig. 1. As expected based on the information in Table 3, *T. maritima* exhibited significant growth on all of the polysaccharides tested, reaching peak cell densities near or greater than 10^9 cells/ml. The growth rates were highest on the linear polysaccharides barley \( /H_9252 \)-glucan (72 min) and konjac glucomannan (74 min), followed by carboxymethyl cellulose (78 min), carob galactomannan (85 min), and potato starch (119 min).

In an effort to elucidate the patterns of expression of the various glycosidases, total RNA was extracted from cells (early to mid-log phase) grown on all polysaccharide substrates, and Northern blot analysis was used to monitor the patterns of expression of several endo-\( \beta \)-glycosidase genes in *T. maritima*, including cel5A, cel5B, cel12A, cel12B, cel74, man5, lam16, xyl10A, and xyl10B. Northern blots showing the expression of these genes are shown in Fig. 2, and the results are summarized in Table 4. It was observed that a putative endoglucanase gene, cel74, was expressed on all substrates tested, including starch. On the other hand, lam16, xyl10A, xyl10B, and cel12A were not expressed on any of the

---

**Table 3. Polysaccharide-degrading enzyme systems in *T. maritima***

<table>
<thead>
<tr>
<th>Locus</th>
<th>Glycosyl hydrolase</th>
<th>N-terminal sequence</th>
<th>Signal peptide</th>
<th>Mode of action</th>
</tr>
</thead>
<tbody>
<tr>
<td>TM1834</td>
<td>Amy4A</td>
<td>MHSWKGQIS ASAVFVEFVR LVEPLKTF PLG SGTQVTVL QDDRIS IQLAVVQEDQ VSQVYVQEDD</td>
<td>No</td>
<td>Exo</td>
</tr>
<tr>
<td>TM1068</td>
<td>Amy4B</td>
<td>MHTYVPVQG SRTTVTLG GLP NQWAVRL DRV HLYLQK KQHNLK</td>
<td>No</td>
<td>Exo</td>
</tr>
<tr>
<td>TM0752</td>
<td>Amy4C</td>
<td>MIKSSGAVQ NRLATNPQ KQHNLK</td>
<td>No</td>
<td>Exo</td>
</tr>
<tr>
<td>TM0434</td>
<td>Amy4D</td>
<td>MTPUPVQG SVYTVLQG VQHNLK</td>
<td>No</td>
<td>Exo</td>
</tr>
<tr>
<td>TM1840</td>
<td>Amy13A</td>
<td>HKUVKVPPLL LVPYLLTT SQVT</td>
<td>Yes</td>
<td>Endo</td>
</tr>
<tr>
<td>TM1650</td>
<td>Amy13B</td>
<td>MNPRLVY EAAR QVFGKQDDLE LSRLKQG RNVEFNIKTGQPPPTGT</td>
<td>No</td>
<td>Endo</td>
</tr>
<tr>
<td>TM1845</td>
<td>Pef13</td>
<td>MTTKWLWLL TLLASQF STTVFVVR DQKVSQQP</td>
<td>Yes</td>
<td>Endo</td>
</tr>
</tbody>
</table>

---

Note: An arrow indicates the predicted cleavage site of the signal peptide.
substrates tested, and the other family 12 endoglucanase gene, cel12B, was expressed when carboxymethyl cellulose was the primary carbon source (data not shown). Expression of man5, which encodes an extracellular mannanase, was induced on carob galactomannan and konjac glucomannan and to some extent on carboxymethyl cellulose. An unexpected result was that the cel5A and cel5B genes encoding intracellular β-specific endoglucanases were expressed only when T. maritima was grown on konjac glucomannan. The observed transcript sizes for cel5A (1 kb), cel5B (1 kb), man5 (2 kb), cel12B (0.9 kb), and cel74 (2 kb) were 1.5, 1.5, 2.0, 0.9, and 1.5 kb, respectively. Unlike the sizes of the man5 and cel12A transcripts, the observed sizes of the cel5A and cel5B transcripts were larger than the genes.

In case of cel74, the observed transcript size was smaller than the gene. A discrepancy between the sizes of transcripts and the corresponding genes has also been observed for exoglycosidases from Thermotoga neapolitana, in which smaller-than-expected transcript sizes were attributed to selective degradation of unstable transcripts by RNases (41).

Sequence analysis and biochemical properties of recombinant Cel5A and Man5. To further examine the potential biochemical basis for expression of cel5A and cel5B on konjac glucomannan, the recombinant form of Cel5A was produced in Escherichia coli. Since Man5 was also active on konjac glucomannan, it was used as a basis for comparison. The nucleotide sequence of the cel5A gene (TM1751) corresponds to a 951-bp
open reading frame that encodes a 317-amino-acid protein with a predicted molecular mass of 37.3 kDa (39). The Cel5A sequence exhibited the highest level of homology (38% amino acid sequence identity) to the sequence of a family 5 endoglucanase (CelD) from Clostridium cellulolyticum, which has an optimum temperature of 50°C (52). An alignment of the Cel5A sequence with the sequences of other members of glycosyl hydrolase family 5 is shown in Fig. 3, and the data indicate that the following eight amino acid residues in Cel5A that are characteristic of family 5 are conserved: Arg-51, His-95, Asn-135, Glu-137 His-196, Tyr-198, Glu-253, and Trp-286 (positions 375, 420, 459, 460, 520, 522, 577, and 610, respectively). The catalytic residues Glu-137 and Glu-253 act as the proton donor and a nucleophile, respectively. Mutation of any one of these residues resulted in a complete loss of catalytic activity (Chhabra and Kelly, unpublished results). The nucleotide sequence of the man5 gene (TM1227) (39) corresponded to a 2,007-bp open reading frame that encodes a 669-amino-acid protein with a predicted molecular mass of 76.9 kDa (12). The Man5 sequence exhibited the highest level of similarity (46% amino acid sequence identity) to the sequence of a β-mannanase (ManF) from Bacillus stea- rothermophilus, a multidomain family 5 enzyme which has a molecular mass of 76 kDa (19). Conserved residues in Man5 that are characteristic of family 5 include Arg-83, His-164, Asn-209, Glu-211, His-290, Tyr-292, Glu-329, and Trp-362 (12). In Man5, mutation of Glu-329 (nucleophile) results in a complete loss of catalytic activity (Chhabra and Kelly, unpublished results).

The recombinant versions of Cel5A and Man5 expressed in E. coli were purified by heat treatment in addition to column chromatography. Recombinant Cel5A, purified twofold from the heat-treated E. coli crude extract by using ion-exchange chromatography, had a specific activity of 17 U/mg on azo-carboxymethyl cellulose. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis analysis of the recombinant protein resulted in a band at 37 kDa. Purification of Man5 has been reported elsewhere (45). Cel5A has a temperature optimum of 80°C and a pH optimum of 6. Under these conditions, Cel5A had a half-life of 18 h and a melting temperature of 90°C. The enzyme followed Michaelis-Menten kinetics on barley β-glucan. The apparent $k_{cat}/K_m$ on this substrate was estimated to be $1,112 \text{ ml s}^{-1} \text{ mg}^{-1}$. As reported previously, Man5 had an apparent $k_{cat}/K_m$ of 150 ml s$^{-1} \text{ mg}^{-1}$ on carob galactomannan under optimal conditions (90°C and pH 7) (45). At 90°C, Man5 has a half-life of 3 h. Properties of both enzymes are summarized in Table 5.

The specific activities of both enzymes were determined for a number of polysaccharide substrates (Table 6). These polysaccharides differed in their backbone compositions, types of glycosidic linkages, and side chain residues. Man5 exhibited activity on mannan-based substrates, including the linear polysaccharides β-mannan and konjac glucomannan and the branched galactomannans from carob and guar. However, no activity was detected on glucan-based substrates. The specific activities on all mannan-based substrates were the same order of magnitude, suggesting that the galactose side chain interaction was not important in galactomannan hydrolysis by this enzyme. Cel5A exhibited activity on both mannan-based and glucan-based polysaccharides. This enzyme exhibited approximately threefold-higher activity on guar and carob galactomannan than on linear β-mannan, suggesting that there was a positive interaction with galactose side chains. However, the degree of galactose substitution did not seem to affect the specific activities significantly. Among the glucan-based substrates tested, the highest activity was observed on barley β-glucan, but Cel5A did not have any detectable activity on L. digitata laminarin or potato starch. Interestingly, Cel5A had higher activity than Man5 on konjac glucomannan.

**DISCUSSION**

*T. maritima* utilizes many simple and complex carbohydrates, including glucose, sucrose, maltose, starch, galacto- and glucomannans, carboxymethyl cellulose and xylan, as growth substrates. The recently released genome sequence of *T. maritima* revealed the presence of a large number of glycosidases, and the percentage of predicted coding sequences involved in sugar metabolism in *T. maritima* is more than twice the percentages seen in the other eubacterial and archaeal species sequenced to date (39). Based on comparative genomics, it was suggested that there has been extensive lateral gene transfer between this bacterium and members of the archaeal domain, particularly Pyrococcus horikoshii (39, 56). In a recent study of transport systems encoded in the genomes of 18 prokaryotic organisms, it was found that *T. maritima* possesses a large number of transporters for sugars and for oligopeptides and comparatively few transporters for amino acids (46).
In order to understand the regulatory networks in organisms whose genome sequences are available, it is necessary to couple insights from bioinformatic approaches with the results of physiological and biochemical studies. For *T. maritima*, the propensity to utilize specific carbohydrates as growth substrates and not to utilize others can be understood from this perspective. *T. maritima* has been found to grow on a range of soluble complex polysaccharides. Although the *T. maritima* genome encodes a number of \( \beta \)-glucan-hydrolyzing enzymes, there is no indication that there is a complex cellulose degradation system, such as the cellulosome found in certain *Clostridium* species (21) and in the anaerobic rumen bacterium *Ruminococcus albus* (43). The cellulosome, which is typically composed of between 14 and 26 polypeptides that are presumptively coexpressed, synergistically hydrolyzes crystalline cellulose, whereas its constituent polypeptides, alone or in mixtures, do not. Many of these polypeptides are catalytically active with soluble substrates and can be characterized as endoglucanases, xylanases, and cellodextrinases (5). The *T. maritima* intracellular endoglucanase, Cel5A, exhibits sequence homology to endoglucanases CelD and CelH from *C. cellulolyticum* and *C. thermocellum* (21), respectively, which are components of the cellulosome (5, 21). However, cellulases that can bind to and initiate hydrolysis of insoluble forms of cellulose are apparently not present in *T. maritima*. The process by which fermentative anaerobes such as *C. thermocellum* and *C. cellulolyticum* developed the capacity to utilize insoluble forms of cellulose is unclear.

The acquisition and processing of complex carbohydrates by *T. maritima* likely involve a range of glycosyl hydrolases and transporters, and the latter have not been identified and characterized yet. The results of the present study indicate that glucomannan induces the \( \beta \)-glucan endoglycosidase genes cel5A (TM1751) and cel5B (TM1752) and the putative \( \beta \)-mannan endoglycosidase gene cel74 (TM0305), as well as the \( \beta \)-mannan endoglycosidase gene man5 (TM1227), in *T. maritima*. Man5 and Cel74 are extracellular enzymes (Table 3) that could degrade glucomannan into smaller subunits, lular endoglucanase, Cel5A, exhibits sequence homology to endoglucanases CelD and CelH from *C. cellulolyticum* and *C. thermocellum* (51) and *Clostridium thermocellum* (21), respectively, which are components of the cellulosome (5, 21). However, cellulases that can bind to and initiate hydrolysis of insoluble forms of cellulose are apparently not present in *T. maritima*. The process by which fermentative anaerobes such as *C. thermocellum* and *C. cellulolyticum* developed the capacity to utilize insoluble forms of cellulose is unclear.

The acquisition and processing of complex carbohydrates by *T. maritima* likely involve a range of glycosyl hydrolases and transporters, and the latter have not been identified and characterized yet. The results of the present study indicate that glucomannan induces the \( \beta \)-glucan endoglycosidase genes cel5A (TM1751) and cel5B (TM1752) and the putative \( \beta \)-specific endoglucanase gene cel74 (TM3058), as well as the \( \beta \)-mannan endoglycosidase gene man5 (TM1227), in *T. maritima*. Man5 and Cel74 are extracellular enzymes (Table 3) that could degrade glucomannan into smaller subunits.

![FIG. 3. Multiple sequence alignment for *T. maritima* Cel5A and other proteins. CelD_Cc, endoglucanase CelD from *C. cellulolyticum*; CelH_Ct, endoglucanase CelH from *C. thermocellum*; Cel5A_Tm, endoglucanase Cel5A from *T. maritima*. The highlighted regions are conserved regions in family 5 glycosidases as suggested by Ethier et al. (19). An asterisk indicates identical or conserved residues; a colon indicates semiconserved substitutions; a period indicates semiconserved substitutions.](http://aem.asm.org/)

### Table 5. Properties of recombinant glycosidases from *T. maritima*

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Locus</th>
<th>Glycosyl hydrolase family</th>
<th>Putative catalytic residues</th>
<th>Calculated molecular mass (Da)</th>
<th>Optimum temp (°C)</th>
<th>Optimum pH</th>
<th>Half-life (h)</th>
<th>Melting temp (°C)</th>
<th>Apparent ( k_{cat}/K_m ) (ml s(^{-1}) mg(^{-1}))</th>
<th>Substrate used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cel5A</td>
<td>TM1751</td>
<td>5</td>
<td>Glu-137, Glu-253</td>
<td>37,383</td>
<td>80</td>
<td>6</td>
<td>18</td>
<td>90</td>
<td>1,112</td>
<td>( \beta )-Glucan</td>
</tr>
<tr>
<td>Man5</td>
<td>TM1227</td>
<td>5</td>
<td>Glu-211, Glu-329</td>
<td>76,913</td>
<td>90</td>
<td>7</td>
<td>3</td>
<td>99</td>
<td>150</td>
<td>Galactomannan</td>
</tr>
</tbody>
</table>

In order to understand the regulatory networks in organisms whose genome sequences are available, it is necessary to couple insights from bioinformatic approaches with the results of physiological and biochemical studies. For *T. maritima*, the propensity to utilize specific carbohydrates as growth substrates and not to utilize others can be understood from this perspective. *T. maritima* has been found to grow on a range of soluble complex polysaccharides. Although the *T. maritima* genome encodes a number of \( \beta \)-glucan-hydrolyzing enzymes, there is no indication that there is a complex cellulose degradation system, such as the cellulosome found in certain *Clostridium* species (21) and in the anaerobic rumen bacterium *Ruminococcus albus* (43). The cellulosome, which is typically composed of between 14 and 26 polypeptides that are presumptively coexpressed, synergistically hydrolyzes crystalline cellulose, whereas its constituent polypeptides, alone or in mixtures, do not. Many of these polypeptides are catalytically active with soluble substrates and can be characterized as endoglucanases, xylanases, and cellodextrinases (5). The *T. maritima* intracellular endoglucanase, Cel5A, exhibits sequence homology to endoglucanases CelD and CelH from *C. cellulolyticum* and *C. thermocellum* (51) and *Clostridium thermocellum* (21), respectively, which are components of the cellulosome (5, 21). However, cellulases that can bind to and initiate hydrolysis of insoluble forms of cellulose are apparently not present in *T. maritima*. The process by which fermentative anaerobes such as *C. thermocellum* and *C. cellulolyticum* developed the capacity to utilize insoluble forms of cellulose is unclear.

The acquisition and processing of complex carbohydrates by *T. maritima* likely involve a range of glycosyl hydrolases and transporters, and the latter have not been identified and characterized yet. The results of the present study indicate that glucomannan induces the \( \beta \)-glucan endoglycosidase genes cel5A (TM1751) and cel5B (TM1752) and the putative \( \beta \)-specific endoglucanase gene cel74 (TM3058), as well as the \( \beta \)-mannan endoglycosidase gene man5 (TM1227), in *T. maritima*. Man5 and Cel74 are extracellular enzymes (Table 3) that could degrade glucomannan into smaller subunits.
which are transported into the cell for further degradation by intracellular Cel5A and Cel5B. The endoglucanase genes cel12A (TM1751) and cel12B (TM1752) are not induced by glucanomann even though in vitro the glycosidases are active on β-glucans (34). Analysis of the genes in the vicinity of cel5A (TM1753), cel5B (TM1752), and cel74 (TM0305) indicates that multiple oligopeptide ABC transporters (TM1746 to TM1750 and TM0300 to TM0304) are present (39). In the case of man5 (TM1227), in addition to multiple oligopeptide ABC transporter subunits (TM1219 to TM1223) downstream of the gene, there are also a number of sugar ABC transporter subunits (TM1232 to TM1234) upstream of the gene (39). In a recent study of sugar transporters in *Sulfolobus solfataricus*, it was found that the maltose and cellobiose transporters exhibit significant sequence similarity to oligopeptide-dipeptide transporters (18). Oligopeptide transporters have also been found in the vicinity of glycosidase genes in *P. horikoshii* (29) and *Thermoplasma acidophilum* (48). This suggests that oligopeptide transporters in the vicinity of endoglycosidase genes in *T. maritima* are most likely involved in sugar transport.

Growth of *T. maritima* on the polysaccharides carob galactomannan and barley β-glucan resulted in expression of man5 and cel74, respectively. It remains to be seen whether the degradation products of these polysaccharides induce expression of the intracellular exoglycosidases man2 (TM1624) and gal36 (TM1192) for galactomannan oligosaccharide hydrolysis and cel3 (TM0025) for cellobiose saccharide hydrolysis. Recent studies performed with *T. neapolitana* cultures grown on galactomannan indicated that all three activities (β-mannanase [GenBank accession no. AY033477], β-mannosidase [GenBank accession no. AY033395], and α-galactosidase [GenBank accession no. AF011400]) were induced (45). The biochemical and physiological properties of Cel74 are not available yet. Preliminary studies of Cel74 have indicated that this enzyme exhibits activity on the polysaccharides barley β-glucan and carboxymethyl cellulose (Chhabra and Kelly, unpublished results), but expression of Cel74 in the presence of the polysaccharides examined here indicated that it may be active on a broad range of substrates.

The source of polysaccharides in the thermophilic environment of *T. maritima* has not been established yet. However, extracellular polysaccharides generated by hyperthermophiles (25, 33, 47) could be utilized as growth substrates. Transformers for the export of exopolysaccharides (the PST family of transporters) have been identified in a number of other hyperthermophiles (46). For instance, 9% of the total transporters in *P. horikoshii* are involved in macromolecular efflux (46). The inducers of polysaccharide synthesis and export in hyperthermophiles have not been examined yet, nor has the potential relationship of these processes to polysaccharide utilization been examined. To do this, a comprehensive study that involves the use of microarrays to examine regulation patterns involving polysaccharide degradation by glycosidases from *T. maritima* is under way.

**ACKNOWLEDGMENTS**

This work was supported in part by grants from the National Science Foundation and the U.S. Department of Energy, Energy Biosciences Program.

**REFERENCES**

VOL. 68, 2002 REGULATION OF ENDO-ACTING GLYCOSYL HYDROLASES 553


