Biochemical characterization of *Thermotoga maritima* endoglucanase Cel74 with and without a carbohydrate binding module (CBM)

Swapnil R. Chhabra, Robert M. Kelly*

Department of Chemical Engineering, North Carolina State University, Stinson Drive, P.O. Box 7905, Raleigh, NC 27695-7905, USA

Received 30 July 2002; revised 16 September 2002; accepted 17 September 2002

First published online 7 October 2002

Edited by Stuart Ferguson

Abstract **The genome of the hyperthermophilic bacterium *Thermotoga maritima* (Tm) encodes at least eight glycoside hydrolases with putative signal peptides; the biochemical characteristics of seven of these have been reported previously. The eighth, Tm Cel74, is encoded by an open reading frame of 2124 bp corresponding to a polypeptide of 79 kDa with a signal peptide at the amino-terminus. The gene (lacking the signal peptide) encoding Tm Cel74 was expressed as a 77 kDa monomeric polypeptide in *Escherichia coli* and found to be optimally active at pH 6, 90°C, with a melting temperature of approximately 105°C. The *cel74* gene was previously found to be induced during *T. maritima* growth on a variety of polysaccharides, including barley glucan, carboxymethyl cellulose (CMC), glucomannan, galactomannan and starch. However, while Tm Cel74 was most active towards barley glucan and to a lesser extent CMC, glucomannan and tamarind (xyloglucan), no activity was detected on other glycans, including galactomannan, laminarin and starch. Also, Tm Cel74 did not contain a carbohydrate binding module (CBM), versions of which have been identified in the amino acid sequences of other family 74 enzymes. As such, a CBM associated with a chitinase in another hyperthermophile, *Pyrococcus furiosus*, was used to create a fusion protein that was active on crystalline cellulose; Tm Cel74 lacked activity on this substrate. Based on the cleavage pattern determined for Tm Cel74 on glucan-based substrates, this enzyme likely initiates recruitment of carbohydrate carbon and energy sources by creating oligosaccharides that are transported into the cell for further processing.**

1. Introduction

Endoglucanases (EC 3.2.1.4) catalyze the endo-hydrolysis of β-1,4 linkages between glucose residues in β-glucans that are found in structural biopolymers in plant cell walls. The genome sequence of *Thermotoga maritima*, a hyperthermophilic bacterium [1], indicates the presence of a number of endoglucanases that can be classified according to the glycoside hydrolase families in which they belong [2]: Tm Cel5A (TM1751), Tm Cel5B (TM1752), Tm Cel12A (TM1524), Tm Cel12B (TM1525) and Tm Cel74 (TM0305). The endoglucanases Tm Cel12B and Tm Cel74 are extracellular enzymes as noted by the presence of signal sequences at their respective N-termini, whereas Tm Cel5A, Tm Cel5B and Tm Cel12A are intracellularly located [3]. Biochemical and biophysical properties of the endoglucanases from family 5 (Tm Cel5A) and family 12 (Tm Cel12A, Cel12B) have been reported previously [3–5]. *T. maritima* endo-acting glycoside hydrolase gene expression patterns were found to vary according to polysaccharide growth substrate [3]. *cel5A* and *cel5B* were specifically induced when glucomannan was used as the growth substrate, whereas *cel12A* was induced on carboxymethyl cellulose (CMC). *cel74*, which has not yet been characterized biochemically, was induced by a variety of polysaccharides, including glucomannan, galactomannan, xylan, starch and CMC. Whether the biochemical features of Tm Cel74 endow it with the potential to process such a wide spectrum of substrates is not known. Currently glycoside hydrolase family 74 contains sequences from six other organisms besides *T. maritima*. A notable feature in each of these sequences is the presence of a C-terminal carbohydrate binding module (CBM) which is absent in the Tm Cel74 sequence. CBMs enhance the binding ability of glycoside hydrolases and enable hydrolysis of poorly soluble polysaccharides, such as cellulose [6]. Whether Tm Cel74 has the capacity to hydrolyze insoluble polysaccharides is not known. As such, a family 2 CBM was identified in a thermostable chitinase, Pf Chi18A, from the hyperthermophilic archaeon *Pyrococcus furiosus* [7]. In addition to the biochemical and biophysical characterization of Tm Cel74, a fusion protein, Tm Cel74–Pf CBM2-Chi18A, containing this CBM was investigated for its ability to bind and hydrolyze insoluble cellulose.

2. Materials and methods

2.1. PCR and cloning of cel74

The genomic DNA of *T. maritima* and *P. furiosus* was obtained as described previously [3,8]. PCR amplification of the *cel74* gene was carried out using the following primers with engineered restriction sites for *Nde*I and *Hin*dIII (indicated in bold). Note that the start and stop codons are indicated in bold: 5′-TGGAGCAACT**CATATG**TGGAAATCGGTGG-3′ and 5′-ACGGCGTT**AAGCTT**CAA**TCA**TTCCTCCTTC-3′. The upstream primer was designed such that a start codon was introduced after the region encoding the signal sequence and four additional residues at the N-terminus of the encoded protein. The signal sequence was identified using the program SignalP [1]. The PCR product was purified using a QIAquick PCR purification kit (Qiagen, Valencia, CA, USA) followed by digestion with *Nde*I and *Hin*dIII and separation of the digested products on an agarose gel and purification of the PCR product using the QIAquick gel extraction kit (Qiagen). The double-digested PCR product was

*Corresponding author. Fax: (1)-919-515 3465.
E-mail address: rmkelly@eos.ncsu.edu (R.M. Kelly).

Abbreviations: CBM, carbohydrate binding module; Cel74, endoglucanase (family 74); Chi18A, endochitinase (family 18); CMC, carboxymethyl cellulose; dp, degree of polymerization; Pf, *Pyrococcus furiosus*; Tm, *Thermotoga maritima*

0014-5793/02/$22.00
PII: S0014-5793(02)03493-2

ligated to *Nde*I/*Hin*dIII-digested pET21b(+) (Novagen) and the ligation mixture was transformed into Nova Blue (Novagen) cells. Plasmid DNA was isolated using the Qiagen Miniprep kit (Qiagen) from Nova Blue colonies grown on ampicillin resistance plates. The presence of the *cel74* insert was confirmed by *Nde*I/*Hin*dIII digestion of the isolated plasmid followed by sequencing the *cel74* gene insert at the Sequencing Facility at the University of Georgia (Athens, GA, USA). The deduced amino acid sequence of Cel74 was compared with other sequences in the GenBank database using BLAST [9] available at http://www.ncbi.nlm.nih.gov/. Multiple sequence alignments were carried out using CLUSTALW [10] available at http://www2.ebi.ac.uk/clustalw/.

2.2. Construction of the fusion protein Tm Cel74–Pf CBM2-Chi18A

The gene encoding the fusion protein Tm Cel74–Pf CBM2-Chi18A was constructed as follows. The binding domain in Pf Chi18A (PF1233) was identified using BLAST [9] available at http://www.ncbi.nlm.nih.gov/. PCR amplification of the linker region (257 bp) and CBM region (332 bp) was carried out separately using the following primers: 5′-TTGCTGGATACAATA**CTCGAG**CATATAGCCCAGGGAGG-3′ and 5′-ACCTCTAGAGAACATGAGACTGGGGACAGGGGTAGTTG-3′ for the linker region and 5′-TCCCAGTCTCATGTTCTCTAGAGGT-3′ and 5′-GGGTGTGGGA**GCGGCCGCTCA**GGTTGGACACCATATACCAATTACTTGT-3′ for the CBM region such that there was a 25 bp overlap between the two regions. Restriction sites, *Xho*I and *Not*I (indicated in bold), were engineered into the upstream primer for the linker region and the downstream primer for the CBM region, respectively. The stop codon in the downstream primer for the CBM region is indicated in bold. Primers for the CBM region contained codons for introducing Cys residues at either end of the CBM region. The CBM and linker regions were then extended by PCR to yield a single fragment (589 bp) using the upstream primer for the linker region and downstream primer for the CBM region. PCR amplification of *cel74* was carried out as described above with the following downstream primer containing a *Xho*I restriction site and a mutated stop codon: 5′-AACGGAGTT-**CTCGAG**AAATCCTTCCTCCTTC-3′. This PCR product (2080 bp) and the linker-CBM fragment (589 bp) were then digested with *Xho*I and the products were ligated and separated on a 1% agarose gel. The fragment corresponding to the correct size (2669 bp) was excised from the gel and purified. This ligation product was PCR-amplified using the upstream primer for *cel74* and the downstream product for the CBM region. The amplified ligation product was then double-digested with *Nde*I and *Not*I and ligated to a double-digested pET21b(+) vector followed by transformation into Nova Blue cells as described above. Subsequent steps were similar to those described in the above section. The sequence of the gene encoding the full-length fusion protein was verified at the Sequencing Facility at the University of Georgia.

2.3. Protein expression and purification

The recombinant plasmid containing the *cel74* insert was transformed into BL21(DE3) cells (Novagen). Transformed cells were grown in 1.5 l of LB medium containing carbenicillin (50 μg/ml) at 37°C until the optical density at 600 nm was 0.7. Expression of the *cel74* gene was induced by addition of IPTG (isopropyl-β-D-thiogalactopyranoside) to a final concentration of 1 mM. Cultures were harvested after 5 h by centrifuging at 5000×*g* for 5 min. Cell pellets were re-suspended in Tris–HCl buffer (20 mM, pH 8.0) followed by addition of lysozyme (100 μg/ml, final concentration), sonication and heat treatment of the supernatant at 70°C for 30 min. A second heat treatment step at 90°C for 15 min removed all heat labile *Escherichia coli* proteins and yielded a single homogeneous band as judged by SDS–PAGE. A similar procedure was adopted for the expression and purification of the fusion protein. Protein concentrations were determined by a dye binding method with bovine serum albumin as standard [11].

2.4. Estimation of optimum parameters

The temperature dependence of Tm Cel74 was determined by measuring specific activity of the enzyme on barley β-glucan (0.8%, w/v) in sodium acetate buffer (50 mM, pH 5.0) at various temperatures. The pH dependence was determined by measuring the specific activities of the enzyme on barley β-glucan (0.8%, w/v) in a series of pH values between 3.6 and 5.6 with 50 mM sodium acetate buffer, between 5.4 and 8.4 with 50 mM sodium phosphate buffer, and between 8.6 and 10.0 with 50 mM glycine–NaOH. Thermostability was determined by incubating the enzyme for various lengths of times at 90°C in sodium acetate buffer (50 mM, pH 5.0) and then determining the residual activity. Melting temperature for Tm Cel74 was determined with a NanoDifferential scanning calorimeter (Calorimetry Sciences, Salt Lake City, UT, USA) as described previously [3].

2.5. Enzyme activity on soluble and insoluble polysaccharides and oligosaccharides

Enzymatic activity was measured on soluble polysaccharide substrates, β-glucan (barley), CMC, glucomannan (konjac), galactomannan (carob), lichenan (icelandic moss), xyloglucan (tamarind), laminarin (*Laminaria digitata*), starch (potato) and xylan (birchwood) using the dinitrosalicylic acid reducing sugar assay [12], as described previously [3]. Michaelis–Menten kinetic parameters were determined under optimal conditions using barley β-glucan as the substrate. The insoluble polysaccharide Avicel PH102 (microcrystalline cellulose) was obtained from FMC Corporation (Philadelphia, PA, USA). The reaction mixture, consisting of a slurry (1% w/v) of insoluble polysaccharide and enzyme, was incubated at 80°C for 5 h followed by centrifugation at 10 000×*g* for 20 min. Enzymatic activity was determined by measuring the presence of reducing sugars in the supernatant solution using the Somogyi-Nelson assay [12]. One unit of enzyme activity was defined as the amount of enzyme required to release 1 μmol of glucose-equivalent reducing groups per minute. Colomnic acid (poly-2,8-*N*-acetylneuraminic acid) was purchased from Sigma. To check for sialidase activity, enzymatic activity on colomnic acid was monitored at 70°C for 30 min for the liberation of sialic acid residues using the Warren assay.

Cello-oligosaccharides (cellohexaose through cellobiose) were obtained from Seikagaku (Falmouth, MA, USA). Hydrolysis reactions (0.055 ml) on cello-oligosaccharides (10 mM) were performed at 80°C using Cel5A Tm and at 90°C using Tm Cel74 for various time intervals in deionized water and monitored using an Aminex XPX-42A (Bio-Rad, Hercules, CA, USA) high performance liquid chromatography column as described in detail previously [13]. The reaction products were identified from elution times of the cello-oligosaccharide standards.

2.6. Cellulose binding studies

The cellulose binding abilities of Tm Cel74 and Tm Cel74–Pf CBM2-Chi18A were determined by incubating appropriate amounts of enzyme solutions with 0.5 ml Avicel (5% w/v) for 1 h at 35°C followed by centrifugation at 16 000×*g* for 20 min to separate the insoluble polysaccharide. The supernatant was tested for activity on CMC as described in the previous section. The bound fraction was estimated from the difference in activities of the supernatant solutions and control reactions.

Table 1
CBMs present in family 74 glycoside hydrolases

Organism	GenPept#	Sequence length (amino acids)	Identity (%)	Location of CBM	CBM family
T. maritima MSB8	AAD35393.1	707	100	not present	N/A
Caldicellulosiruptor sp. *Tok7B.1*	AAK06388.1	996	39	C-terminus (840–996)	3
S. coelicolor A3(2)	CAA20642.1	890	36	C-terminus (786–886)	2
A. bisporus D649	CAC02964.1	806	34	C-terminus (771–806)	1
A. aculeatus	BAA29031.1	856	33	C-terminus (822–850)	1
A. niger	AAK77227.1	857	33	C-terminus (811–857)	1

Note: N/A not applicable.

Table 2
Properties of extracellular endo-acting glycoside hydrolases from *T. maritima*

Glycoside hydrolase	GenPept#	Activity	Size (kDa)	Signal sequence	CBM family	Optimum conditions		Ref
						T_{opt} (°C)	pH_{opt}	
Tm Cel74	AAD35393.1	1,4-β-glucanase	79	mlrsflilflailgvvfg	no CBM	90	6.0	this work
Tm Cel12B	CAA93274.1	1,4-β-glucanase	32	mrwavllmvvfsallfssevvl	no CBM	85	6.0	[4]
Tm Lam16[a]	AAD35118.1	1,3-β-glucanase	73	mmsrlvfall lfpvfilaqn	4	95	6.2	[28]
Tm Amy13A	CAA72194.1	1,4-α-glucan glucanohydrolase	65	mkvvkkpflllafvflflttscfqts	no CBM	85–90	7.0	[27]
Tm Pul13	CAA04522.1	1,6-α-glucosidase	96	mktklwllvlllsalifs	no CBM	90	6.0	[26]
Tm Xyl10A	CAA86406.1	1,4-β-xylanase	120	mqvrkrrglldvstavlvgilagflgvvla	9/22	90	7.0	[30]
Tm Xyl10B[a]	AAD35164.1	1,4-β-xylanase	41	mkilpsvlil llgcvpvfs	no CBM	90	5.5	[29]
Tm Man5	AAD36302.1	1,4-β-mannanase	77	mrrfmfillvelsfvlfa	27	90	7.0	[25]

[a]Optimum conditions reported for Tn Lam16 and Tn Xyl10B from *T. neapolitana* which are 88% and 83% identical to Tm Lam16 and Tm Xyl10B from *T. maritima*, respectively.

3. Results and discussion

3.1. Sequence analysis of Tm Cel74 and Pf CBM2-Chi18A

The nucleotide sequence of the *cel74* gene (TM0305) corresponded to an open reading frame (ORF) of 2124 bp, encoding a 707 amino acid protein with a predicted molecular mass of 79 kDa [1] and an N-terminal signal peptide region. Elimination of the signal sequence resulted in a protein with a molecular mass of 77 kDa, as confirmed by SDS–PAGE. A BLAST [9] search on the Tm Cel74 sequence reveals sequence similarity to members of family 74 of glycoside hydrolases as shown in Table 1 (see web annex for sequence alignment; http://www.elsevier.com/PII/S0014579302034932). In addition to other putative glucanases, Tm Cel74 exhibits homology (36% sequence identity) to a putative sialidase from *Clostridium acetobutylicum* [14]. Unlike Tm Cel74, the putative glycoside hydrolases in family 74 from *Caldicellulosiruptor* sp., *Streptomyces coelicolor*, *Agaricus bisporus*, *Aspergillus aculeatus* and *Aspergillus niger* contain non-catalytic domains in the form of CBMs at their respective C-termini (Table 1) [15]. CBMs, contiguous amino acid sequences within a carbohydrate-active enzyme with a discrete fold having carbohydrate binding activity [15], are currently classified into 31 families based on amino acid sequence similarity [16]. CBMs present in glycoside hydrolases from family 74 belong to families 1, 2 and 3 and are involved in the binding to insoluble cellulose [16]. Other endo-acting glycoside hydrolases from *T. maritima* (see Table 2) contain CBMs belonging to families 4 (β-1,3-glucan binding), 9, 22 (xylan binding) and 27 (mannan binding) associated with the extracellular glycoside hydrolases Tm Lam16, Tm Xyl10A and Tm Man5, respectively [17].

To determine the significance of the lack of a CBM in Tm Cel74, a fusion protein consisting of Tm Cel74 and a CBM (CBM2) from an endochitinase (PF1233) from *P. furiosus*, Pf Chi18A, was constructed (Gao and Kelly, unpublished data). CBMs from family 2 have been shown to bind crystalline cellulose [18,19] and contain the consensus sequence xCxxxxWxxxxxNxxxWxxxxxxxWxxxxxxxxWNxxxxxGxxxxxxxxxCx, where x can be any amino acid. The sequence alignment of Pf CBM2-Chi18A with two other members of CBM family 2 is shown in Fig. 1. In a recent study on Cf CBM2-Xyn10A from *Cellulomonas fimi* it was shown that effective binding of a family 2 CBM to crystalline cellulose requires an aromatic group at position 71, a Trp residue at position 53 and either a Tyr or Trp at position 16 [20,21]. Although these residues are strictly conserved in Pf CBM2-Chi18A, it lacks the presence of Cys residues present at either end of other CBMs from this family. These residues were introduced in the final sequence of the fusion protein. Pf CBM2-Chi18A was attached to the C-terminus of Tm Cel74 using a 64 amino acid Ser/Thr rich linker region. The fusion protein Tm Cel74–Pf CBM2-Chi18A had a molecular mass of 95 kDa, which was confirmed by SDS–PAGE.

3.2. Biochemical characterization of Tm Cel74 and the fusion protein Tm Cel74–Pf CBM2-Chi18A

Purified Tm Cel74 had an optimum pH around 6.0 and an optimum temperature of 90°C for the hydrolysis of barley β-glucan. The enzyme exhibited high thermostability at 90°C with a half-life of around 5 h and had a denaturing temperature of 105°C, as determined by differential scanning calorimetry. Table 2 lists some properties of other endo-acting

```
                          10        20        30        40        50        60
                           |         |         |         |         |         |
Cf CBM2-Xyn10A      SGPAGCQVLWGVN-QWNTGFTANVTVKNTSSAPVDGWTLTFSFPSGQQVTQAWSSTVTQS
Sc CBM2-Cel74       TG--ACEVTYTVTNQWPGGFQADVRLTNTGTSAWNGWSLDWSFPGGQEVTRMWNAEHTQA
Pf CBM2-Chi18A      SG----SLEVKVN-DWGSGAEYDVTLNLDG--QYD-WTVKVKLAPGATVGSFWSANKQEG
                    :*    .:   *. :*  *   :* :.  .    : *::  .:. *  *   *.:   :.

                          70        80        90       100       110
                           |         |         |         |         |
Cf CBM2-Xyn10A      GSAVTVRNAPWNGSIPAGGTAQFGFNGSHTGTNAAPTAFSLNGTPCTVG-
Sc CBM2-Cel74       GTSVTARNVGWNAGVAPGASVGFGFTGSRSGTNAEPEGFAVAGRACPTAT
Pf CBM2-Chi18A      NGYVIFTPVSWNKG--PTATFGFIVNGPQ-GDKVEEITLEINGQVIDIWT
                    .  *    . ** .  . .:  * ..*.: * :.    : : *
```

Fig. 1. Sequence alignment of family 2 CBMs. Key: CBMs from the glycoside hydrolases in the organisms as indicated. Cf CBM2-Xyn10A: xylanase from *C. fimi*; Sc CBM2-Cel74: endoglucanase from *S. coelicolor*; Pf CBM2-Chi18A: chitinase from *P. furiosus*. Note: an asterisk indicates identical or conserved residues; a colon indicates conserved substitutions; a period indicates semi-conserved substitutions.

glycoside hydrolases in *T. maritima* that are characterized by the presence of signal peptides at their respective N-termini. Interestingly, most of these extracellular enzymes, including Tm Cel74, have temperature optima around 90°C, 10° higher than the optimum growth temperature of *T. maritima*.

Tm Cel74 was tested for the hydrolysis of several soluble and insoluble polysaccharides. This enzyme showed detectable levels of activity only on soluble glucan-based substrates containing a β-1,4 linkage (Table 3), including β-glucan (705 U/mg), xyloglucan (163 U/mg), CMC (121 U/mg) and glucomannan (107 U/mg). In a recent study on another endoglucanase from family 74, An Cel74 (EglC) from *A. niger*, highest hydrolytic activity was observed on xyloglucan which was around 19-fold higher than that on β-glucan (barley) and CMC [22]. Hydrolysis of barley β-glucan by Tm Cel74 followed Michaelis–Menten kinetics with an apparent k_{cat}/K_m of 336 ml/mg/s. No activity was detected on the soluble polysaccharides laminarin, starch, β-xylan and colomnic acid. Table 5 compares the properties of this enzyme with those of the other endoglucanases present in *T. maritima*, namely Cel5A, Cel12A and Cel12B.

Hydrolysis of cello-oligosaccharides ranging from cellohexaose through cellobiose was also investigated. Cellohexaose hydrolysis proceeded very slowly (Fig. 2A) resulting in the formation of small amounts of cellotetraose and cellobiose. No degradation of oligosaccharides smaller than cellohexaose was observed under the conditions tested. This confirmed the endo mode of action of Tm Cel74 and suggests that this enzyme hydrolyzes β-glucans into oligosaccharides with a degree of polymerization (dp) six units or higher. In contrast, the intracellular endoglucanase Cel5A (TM1751) from *T. maritima* readily hydrolyzed cellohexaose into smaller subunits, cellotetraose, cellotriose and cellobiose (Fig. 2B). Tm Cel5A also degraded the oligosaccharides cellopentaose and cellotetraose to smaller subunits (data not shown). Similarly, the other intracellular endoglucanase from *T. maritima*, Tm Cel12A, was shown to degrade oligosaccharides up to a dp of two [5].

Addition of a cellulose binding domain to Tm Cel74 increased its binding on the insoluble polysaccharide Avicel. Incubation of Tm Cel74–Pf CBM2-Chi18A with a 5% solution of Avicel at room temperature for 1 h resulted in 30% of the hybrid protein being bound to the polysaccharide. No binding was observed for Tm Cel74 without the CBM. Tm Cel74–Pf CBM2-Chi18A released 0.12 μmol glucose equivalents/μmol enzyme per minute from Avicel at 80°C, whereas Tm Cel74 was inactive on this substrate (Table 4). The activity on the soluble polysaccharide CMC was roughly the same for both proteins.

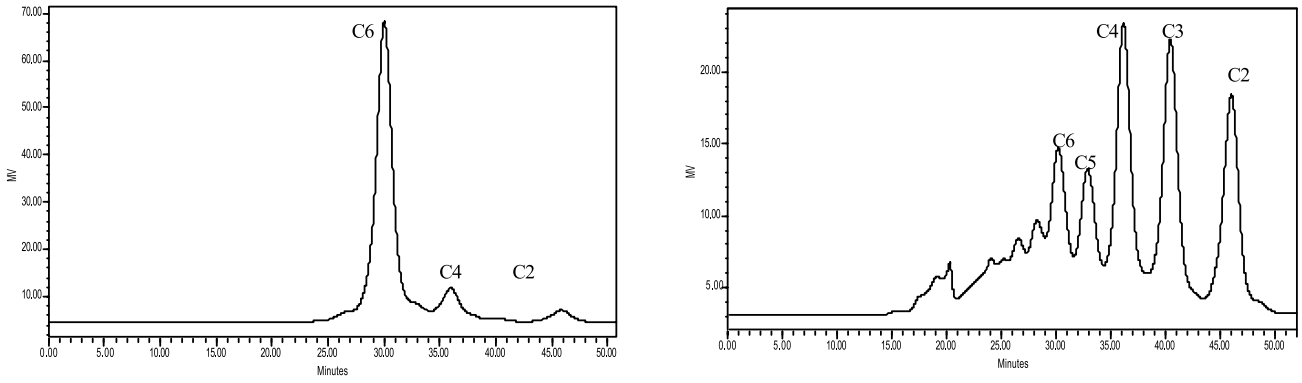
3.3. Summary

The genome sequence of *T. maritima* encodes a number of glycoside hydrolases that belong to families 1–5, 10, 12, 13, 16, 28, 29, 31, 32, 36, 38, 42, 51, 53, 67 and 74 [1,23]. These apparently enable this organism to grow on simple and complex carbohydrates (ranging from glucose [24] to galactomannan [25]). Of these, at least eight are endo-acting glycoside hydrolases [4,25–30] that contain putative signal peptides, suggesting an extracellular role in the acquisition of potential carbon and energy sources by this organism [3]. Although the exact source of polysaccharides in thermophilic environments of *T. maritima* has not yet been established, it can be hypothesized that extracellular polysaccharides generated by other hyperthermophilic bacteria like *Thermococcus litoralis* [31] could be utilized as growth substrates. The ability of *T. maritima* to utilize β-linked glucan substrates can be attributed to the extracellular glucanases, Tm Lam16 (TM0024), Tm Cel12B (TM1525) and Tm Cel74 (TM0305) encoded by its genome (Table 2). Both endo-1,4-β-glucanases, Tm Cel12B and Tm Cel74 Tm, lack the presence of any CBMs whereas the endo-1,3-β-glucanase Tn Lam16 from *T. neapolitana* is flanked by two CBMs, 148 and 161 amino acids long [32].

Table 3
Relative activities of Tm Cel74 on water soluble polysaccharides

Substrate	MW (kDa)	Backbone linkage (backbone sugar)	Relative activity (%)
β-Glucan (barley)	250	β-1,3/4 (glucose)	100
Xyloglucan (tamarind)	N/A	β-1,4 only (glucose)	23
CMC	90	β-1,4 only (glucose)	17
Glucomannan (konjac)	100	β-1,4 only (glucose/mannose)	15
Laminarin (*L. digitata*)	5	β-1,3 only (glucose)	0
Galactomannan (carob)	N/A	β-1,4 only (mannose)	0

N/A: not available.



A. B.

Fig. 2. Cellohexaose hydrolysis. A: Hydrolysis by Tm Cel74 at 90°C for 40 min. B: Hydrolysis by Tm Cel5A at 80°C for 10 min. Note: elution times for oligosaccharides were as follows: cellohexaose (C6): 30 min; cellopentaose (C5): 33 min; cellotetraose (C4): 36 min; cellotriose (C3): 41 min; cellobiose (C2): 46 min.

Although the gene encoding Tm Cel74 is constitutively expressed when *T. maritima* is grown on galactomannan, glucomannan, β-glucan, CMC and starch [3], its biochemical characteristics suggest that the extracellular protein is mainly responsible for producing oligosaccharides from the soluble polysaccharides β-glucan, glucomannan and CMC. These oligosaccharides may then be further transported into the cell for subsequent degradation to smaller units by the intracellular endoglucanases Tm Cel5A, Tm Cel5B and Tm Cel12A. Note the presence of multiple ABC oligopeptide transporter genes (TM0300–TM0304) in the vicinity (16 bp upstream) of *cel74*. The di- and trisaccharides produced by Cel5A and Cel5B may then be converted to glucose either by cellobiose phosphorylase (TM1848) [33] – a family 36 glycosyl transferase [34] – and/or by the β-glucosidase, BglA [35], for use in central metabolism pathways. Although our biochemical analysis suggests that Tm Cel74 acts mainly on β-1,4-linked glucan polymers, the presence of genes encoding the putative enzymes, α-fucosidase (TM0306), fucose isomerase (TM0307) and β-galactosidase (TM0310), downstream of *cel74* suggests that this gene may also be involved in the breakdown of other complex polysaccharides.

The effect of CBMs on cellulase activity has been studied on a number of endoglucanases [6,21,36–38]. Previous studies have suggested that CBMs enhance the enzymatic activity of cellulases by increasing the effective enzyme concentration at the substrate surface or by disrupting the crystalline cellulose structure thereby improving substrate accessibility [6,19,39]. A recent study on family 2 CBMs points to the role of surface aromatics in binding with the staircase-like cellulose surface [20]. Most of the work done to date on hybrid cellulases has been on endoglucanases fused to CBMs from mesophilic organisms. These hybrid proteins have been shown to enhance their activity on crystalline cellulose [19,40,41] in some cases while addition of the binding domain to the catalytic center has had little or no effect in others [42]. Removal of CBMs occurring naturally in certain cellulases has reduced their ability to degrade crystalline cellulose [21,38]. This is the first report of a hyperthermophilic CBM being fused to a hyperthermophilic endoglucanase. The fusion protein Tm Cel74–Pf CBM2-Chi18A had higher activity on microcrystalline cellulose at 80°C than the parent endoglucanase. A recent study on family 6 CBMs found that multiple CBMs are able to bind cooperatively and enhance the binding ability by 20–40-fold relative to individual modules [18]. Indeed, sequence analysis of a number of modular hyperthermophilic glycoside hydrolases reveals the presence of a higher number of binding domains as compared to their mesophilic counterparts, perhaps in order to compensate for the loss of binding affinity at elevated temperatures [18]. Whether the addition of multiple CBMs to Tm Cel74 further enhances its activity on crystalline cellulose is currently under investigation.

Acknowledgements: This work was supported by grants from the U.S. National Science Foundation (Biotechnology Program) and U.S. Department of Energy (Energy Biosciences Program).

References

[1] Nelson, K.E. et al. (1999) Nature 399, 323–329.
[2] Henrissat, B., Teeri, T.T. and Warren, R.A.J. (1998) FEBS Lett. 425, 352–354.
[3] Chhabra, S.R., Shockley, K.R., Ward, D.E. and Kelly, R.M. (2002) Appl. Environ. Microbiol. 68, 545–554.
[4] Liebl, W., Ruile, P., Bronnenmeier, K., Riedel, K., Lottspeich, F. and Greif, I. (1996) Microbiology 142, 2533–2542.
[5] Liebl, W. (2001) Methods Enzymol. 330, 290–300.
[6] Linder, M. and Teeri, T.T. (1997) J. Biotechnol. 57, 15–28.

Table 4
Comparison of Tm Cel74 and Tm Cel74–Pf CBM2-Chi18A

Protein	Size (kDa)	Sequence length (amino acids)	Binding	Enzyme activity (katal/mol)	
				CMC	Avicel
Tm Cel74	79	686	–	282	0
Tm Cel74–Pf CBM2-Chi18A	95	865	+	273	0.12

katal/mol: molecules of product generated/min per molecule of enzyme.

Table 5
Comparison of properties of endoglucanases from *T. maritima*

Endoglucanase	Location	Family	ORF	Size (kDa)	Properties				Specific activity (U/mg)[b]	Ref
					T_{opt} (°C)	$t_{1/2}$[a] (h)	T_{melt} (°C)	pH_{opt}		
Tm Cel5A	Intracellular	5	TM1751	37	80	18	90	6	2345	[3]
Tm Cel12A		12	TM1524	30	90	3	NR[c]	7	1785	[4]
Tm Cel12B[d]	Extracellular	12	TM1525	30	85	9	NR	6	2.91	[4]
Tm Cel74		74	TM0305	79	90	5	105	7	705	this work

[a]Half-life at optimum temperature.
[b]Specific activity measured on barley β-glucan using the dinitrosalicylic acid assay [12]. One unit of enzyme activity is defined as the amount of enzyme required to release 1 μmol of glucose-equivalent reducing groups per minute.
[c]Not reported.
[d]Properties reported for crude enzyme extract.

[7] Robb, F.T., Maeder, D.L., Brown, J.R., DiRuggiero, J., Stump, M.D., Yeh, R.K., Weiss, R.B. and Dunn, D.M. (2001) Methods Enzymol. 330, 134–157.
[8] Ward, D.E., Kengen, S.W.M., van der Oost, J. and de Vos, W.M. (2000) J. Bacteriol. 182, 2559–2566.
[9] Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) J. Mol. Biol. 215, 403–410.
[10] Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) Nucleic Acids Res. 22, 4673–4680.
[11] Smith, P.K. et al. (1985) Anal. Biochem. 150, 76–85.
[12] Chaplin, M. (1994) in: Carbohydrate Analysis: A Practical Approach (Chaplin, M.F. and Kennedy, J.F., Eds.), OIRL Press, Oxford.
[13] Bauer, M.W., Driskill, L.E., Callen, W., Snead, M.A., Mathur, E.J. and Kelly, R.M. (1999) J. Bacteriol. 181, 284–290.
[14] Nolling, J. et al. (2001) J. Bacteriol. 183, 4823–4838.
[15] Coutinho, P.M. and Henrissat, B. (1999) in: Genetics, Biochemistry and Ecology of Cellulose Degradation (Ohmiya, K., Hayashi, K., Sakka, Y., Kobayashi, Y., Karita, S. and Kimura, T., Eds.), pp. 15–23.
[16] Coutinho, P.M. and Henrissat, B. (1999) in: Recent Advances in Carbohydrate Bioengineering (Gilbert, H.J., Davies, G., Henrissat, B. and Svensson, B., Eds.), pp. 3–12, The Royal Society of Chemistry, Cambridge.
[17] Sunna, A., Gibbs, M.D. and Bergquist, P.L. (2001) Biochem. J. 356, 791–798.
[18] Boraston, A.B., McLean, B.W., Chen, G., Li, A., Warren, R.A. and Kilburn, D.G. (2002) Mol. Microbiol. 43, 187–194.
[19] Bolam, D.N. et al. (1998) Biochem. J. 331, 775–781.
[20] McLean, B.W., Bray, M.R., Boraston, A.B., Gilkes, N.R., Haynes, C.A. and Kilburn, D.G. (2000) Protein Eng. 13, 801–809.
[21] Hall, J., Black, G.W., Ferreira, L.M., Millward-Sadler, S.J., Ali, B.R., Hazlewood, G.P. and Gilbert, H.J. (1995) Biochem. J. 309, 749–756.
[22] Hasper, A.A., Dekkers, E., van Mil, M., van De Vondervoort, P.J. and de Graaff, L.H. (2002) Appl. Environ. Microbiol. 68, 1556–1560.
[23] Henrissat, B. and Coutinho, P.M. (2001) Methods Enzymol. 330, 183–201.
[24] Huber, R., Langworthy, T.A., Konig, H., Thomm, M., Woese, C.R., Sleytr, U.B. and Stetter, K.O. (1986) Arch. Microbiol. 144, 324–333.
[25] Parker, K.N. et al. (2001) Biotechnol. Bioeng. 75, 322–333.
[26] Bibel, M., Brettl, C., Gosslar, U., Kriegshauser, G. and Liebl, W. (1998) FEMS Microbiol. Lett. 158, 9–15.
[27] Liebl, W., Stemplinger, I. and Ruile, P. (1997) J. Bacteriol. 179, 941–948.
[28] Zverlov, V.V., Volkov, I.Y., Velikodvorskaya, T.V. and Schwarz, W.H. (1997) Microbiology 143, 1701–1708.
[29] Velikodvorskaya, T.V., Volkov, I.Y., Vasilevko, V.T., Zverlov, V.V. and Piruzian, E.S. (1997) Biochemistry 62, 66–70.
[30] Winterhalter, C., Heinrich, P., Candussio, A., Wich, G. and Liebl, W. (1995) Mol. Microbiol. 15, 431–444.
[31] Rinker, K.D. and Kelly, R.M. (2000) Biotechnol. Bioeng. 69, 537–547.
[32] Zverlov, V.V., Volkov, I.Y., Velikodvorskaya, G.A. and Schwarz, W.H. (2001) Microbiology 147, 621–629.
[33] Yernool, D.A., McCarthy, J.K., Eveleigh, D.E. and Bok, J.D. (2000) J. Bacteriol. 182, 5172–5179.
[34] Campbell, J.A., Davies, G.J., Bulone, V.V. and Henrissat, B. (1998) Biochem. J. 329, 719.
[35] Gabelsberger, J., Liebl, W. and Schleifer, K.H. (1993) Appl. Microbiol. Biotechnol. 40, 44–52.
[36] Srisodsuk, M., Lehtio, J., Linder, M., Margolles-Clark, E., Reinikainen, T. and Teeri, T.T. (1997) J. Biotechnol. 57, 49–57.
[37] Belaich, A., Parsiegla, G., Gal, L., Villard, C., Haser, R. and Belaich, J.P. (2002) J. Bacteriol. 184, 1378–1384.
[38] Gilbert, H.J., Hall, J., Hazlewood, G.P. and Ferreira, L.M. (1990) Mol. Microbiol. 4, 759–767.
[39] Shoham, Y., Lamed, R. and Bayer, E.A. (1999) Trends Microbiol. 7, 275–281.
[40] Maglione, G., Matsushita, O., Russell, J.B. and Wilson, D.B. (1992) Appl. Environ. Microbiol. 58, 3593–3597.
[41] Karita, S., Sakka, K. and Ohmiya, K. (1996) J. Ferment. Bioeng. 81, 553–556.
[42] Poole, D.M., Durrant, A.J., Hazlewood, G.P. and Gilbert, H.J. (1991) Biochem. J. 279, 787–792.