Journal of Molecular Catalysis B: Enzymatic 66 (2010) 294–301

Contents lists available at ScienceDirect

Journal of Molecular Catalysis B: Enzymatic

journal homepage: www.elsevier.com/locate/molcatb

A novel thermostable cellulase from *Fervidobacterium nodosum*

Yuguo Wang[a], Xiaonan Wang[a], Rentao Tang[a], Shanshan Yu[a], Baisong Zheng[a], Yan Feng[a,b,*]

[a] Key Laboratory for Molecular Enzymology and Engineering of Ministry of Education, Jilin University, 2519 Jiefang Road, Changchun 130023, China
[b] Key Laboratory of Microbial Metabolism (Ministry of Education) and College of Life Science and Biotechnology, Shanghai Jiao Tong University, 800 Dong-Chuan Road, Shanghai 200240, China

ARTICLE INFO

Article history:
Received 14 March 2010
Received in revised form 7 June 2010
Accepted 16 June 2010
Available online 23 June 2010

Keywords:
Cellulase
Fervidobacterium nodosum
Expression
Characterization
Thermostability

ABSTRACT

A novel cellulase gene encoding a thermostable endoglucanase from the thermophilic eubacterium *Fervidobacterium nodosum* Rt17-B1 was cloned and expressed, which is the first cellulase cloned from the organisms of genus *Fervidobacterium* and designated as *Fn*Cel5A for being a member of glycoside hydrolase family 5, and the enzymatic properties were characterized. The cellulase was overexpressed in *Escherichia coli* with a high protein content and good solubility in water, and could be easily purified. The purified recombinant cellulase shows high hydrolytic activities on carboxylmethyl cellulose, regenerated amorphous cellulose, β-D-glucan from barley and galactomannan, with the optimum temperature of 80–83 °C and the optimum pH of 5.0–5.5. Furthermore, this enzyme is highly thermostable and has a half-life of 48 h at 80 °C. With such a combination of thermostability and high activities, this cellulase is expected to be useful for hydrolysis of cellulosic and hemicellulosic substrates at high temperatures, and for industrial hydrolysis of plant cellulose during long-time processing at the elevated temperatures, particularly in converting biomass into biofuels.

1. Introduction

Converting enormous and renewable biomass into fuels and products is important to the development of a sustainable industrial society [1]. As the most abundant renewable organic compound in biomass and an almost inexhaustible source of raw material, cellulose is a linear homopolymer of β-1,4-linked glucose units and can theoretically be hydrolyzed to pure glucose. Thus, there is great interest in the hydrolysis of cellulose to glucose or other fermentable sugars which could be used for the production of sustainable products and biofuels such as ethanol [2–4]. In the technologies for effectively converting cellulosic biomass to fermentable sugars, cellulases are regarded as one of the key elements. Cellulases comprise three major types of enzymatic activities: endoglucanases, exoglucanases and β-glucosidases. Endoglucanases (EC 3.2.1.4) cut randomly at internal amorphous sites in the cellulose polysaccharide chain. Exoglucanases including cellodextrinases (EC 3.2.1.74) and cellobiohydrolases (EC 3.2.1.91) act in a processive manner on the ends of cellulose polysaccharide chains, liberating either glucose or cellobiose. β-Glucosidases (EC 3.2.1.21) hydrolyze soluble cellodextrins and cellobiose to glucose [5]. Besides cellulose, hemicellulose is another most abundant renewable biomass in plant cell walls [6–8]. Unlike cellulose, hemicelluloses are heterogeneous polymers built up by pentoses (e.g. D-xylose and D-arabinose), hexoses (e.g. D-mannose, D-glucose and D-galactose) and sugar acids. Hemicelluloses (including xylan, galactomannan, glucomannan, mannan, xyloglucan, etc.) are also considered to be the abundant biomass which could be converted into fuels, and the enzymes responsible for the degradation of hemicelluloses would play an important role in the conversion processes [9–11].

Cellulases can be classed in several families according to the comparison of the primary sequences of glycoside hydrolases [12], and the information on cellulases in glycoside hydrolase families is available on the web of CAZy (carbohydrate-active enzymes) [13]. Cellulases comprising three types of activities (endoglucanases, exoglucanases and β-glucosidases) are found in glycoside hydrolase family 1, 3, 5–9, 12, 30, 44, 45, 48, 51, 61 and 74. The traditional IUB Enzyme Nomenclature of glycoside hydrolases is based on their substrate specificity and occasionally on their molecular mechanism, does not reflect the structural features of these enzymes. The classification of glycoside hydrolase families can reflect a direct relationship between sequences and folding similarities, and not take into account substrate specificities which belong to non-related families. The members of glycoside hydrolase family 5 (GH5), for instance, have close evolutionary relationships, similar

Abbreviations: CAZy, carbohydrate-active enzymes; CBM, carbohydrate-binding module; CMC, carboxymethyl cellulose; GH5, glycoside hydrolase family 5; HGT, horizontal gene transfer; *p*NPC, *p*-nitrophenyl β-D-cellobioside; RAC, regenerated amorphous cellulose.

* Corresponding author at: Key Laboratory of Microbial Metabolism (Ministry of Education) and College of Life Science and Biotechnology, Shanghai Jiao Tong University, 800 Dong-Chuan Road, Shanghai 200240, China. Tel.: +86 21 34207189; fax: +86 21 34207189.
E-mail addresses: yfeng2009@mail.sjtu.edu.cn, yfeng@mail.jlu.edu.cn (Y. Feng).

1381-1177/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.molcatb.2010.06.006

structural features and catalytic mechanism, but have several different substrate specificities. Moreover, the number of invariant potentially catalytic amino acid residues in a GH5 protein is so low that the active site residues can be commonly directly predicted [12].

Thermophilic enzymes are usually optimally active between 60 and 80 °C. With their intrinsical stability and activities at high temperatures, thermophilic enzymes have major biotechnological advantages over mesophilic enzymes [14]. As a kind of thermophilic enzymes, thermostable cellulases have some significant advantages as catalysts in the processes for converting cellulose to sugars [15]. Thermophiles have often been proposed as sources of highly active and thermostable enzymes, some thermostable cellulases originating from thermophilic microorganisms were reported [10,16–19]. *Fervidobacterium nodosum*, which was originally isolated from a New Zealand hot spring, is a thermophilic, glycolytic anaerobic eubacterium. It grows at temperatures ranging from 47 to 80 °C (optimum 65–70 °C) and at pH 6.0–8.0 (optimum pH 7.0–7.5) [20]. *F. nodosum* belongs to the order Thermotogales. The genomes of Thermotogales were thought to have frequently undergone horizontal (or lateral) gene transfer (HGT or LGT), and to have complex and incongruent evolutionary histories [21,22]. Recently, the genome sequence data of *F. nodosum* Rt17-B1 has been available, and it seems reasonable to look for the new genes encoding glycoside hydrolases with cellulase or hemicellulase activity in its genome.

For converting cellulosic biomass into fuels, the hydrolytic activities of existing industrial cellulases are not sufficiently high, and cellulase preparations for industrial purposes are too expensive to be widely used. There is considerable interest in looking for economic cellulases with high activities and outstanding stability. We aimed at discovering new suitable candidate cellulases from thermophilic microorganisms. In this paper, we report on the cloning, overexpression, purification and characterization of the novel cellulase encoded by gene Fnod_1560 from *F. nodosum* Rt17-B1. It is the first cellulase from genus *Fervidobacterium* that was cloned and expressed. At high temperature, this cellulase shows high activities and good thermostability. These characteristics might make this cellulase having many advantages and potential applications in the catalytic reaction processes.

2. Materials and methods

2.1. Isolation of the gene encoding the cellulase

The type strain of *F. nodosum* Rt17-B1 (DSMZ 5306; ATCC 35602) was obtained from the Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ) and used as the source of genomic DNA. The organism was grown under strictly anaerobic conditions at 70 °C in thermoanaerobium medium (Medium 144 of DSMZ) [20]. Chromosomal DNA of *F. nodosum* was isolated using the bacteria genomic DNA purification kit (TIANGEN) according to manufacturer's instructions.

The DNA fragment of the Fnod_1560 gene, encoding the mature cellulase protein, was amplified by PCR from *F. nodosum* genomic DNA using *Pfu* DNA polymerase (Fermentas). The primers, 5′-CAGCGC**CATATG**GACCAAAGTGTGAGTAA-3′ and 5′-GACGTC**GGATCC**TTATTTTCCAAGTGCAG-3′, incorporated *Nde*I and *Bam*HI restriction sites (underlined letters). The amplified gene was digested with *Nde*I and *Bam*HI, and was inserted into the similarly restricted vector pET-11a (Novagen) to yield the expression plasmid p*Fn*Cel5A. The inserted fragment was sequenced to verify identity with the anticipated sequence.

The nucleotide sequence encoding the mature cellulase protein has been deposited in the GenBank database under accession number GU944824.

2.2. Domain structure and phylogenetic analysis of the new cellulase

We assembled the protein sequences which were used in this work from the sequence information that was classified, annotated and linked in CAZy database [13]. At the CAZy web page of glycoside hydrolase family 5 (http://www.cazy.org/fam/GH5.html) in October 2009, we chose those proteins which were annotated with appropriate EC (Enzyme Commission) numbers and not identified from unclassified microorganism, then collected the sequence information of those proteins by means of the CAZy web page links.

We used our collected protein sequences to request the SMART (simple modular architecture research tool) service [23,24], and processed these sequences according to the annotations of domain and region sequences in SMART. We only reserved enzymatic catalytic module (catalytic domain) sequences and the undefined regions closest to two terminuses of catalytic modules, for the next phylogenetic analysis. In other words, the fragment sequences of signal peptides, carbohydrate-binding modules (CBMs), transmembrane segments, linkers and other regions unattached to catalytic modules, which are the portions of original full-length protein sequences, were excluded. These edited protein sequences were used for further analysis on the alignment and phylogeny.

ClustalX 2.0.12 [25] was used for alignments with default parameters, construction of phylogenetic trees with the neighbor-joining method and bootstrap analyses of N-J trees. The phylogenetic trees were displayed by the MEGA4 software [26].

2.3. Expression and purification of the cellulase

The plasmid p*Fn*Cel5A was transformed into the *E. coli* BL21-CodonPlus (DE3)-RIL strain (Stratagene) to express the cellulase. The cells were cultured in 2 × YT medium (1.6% tryptone, 1% yeast extract, and 0.5% NaCl) containing 100 μg/ml ampicillin at 37 °C to mid-exponential phase (A_{600} = 0.8–1). Isopropyl-β-D-thiogalactoside was then added to a final concentration of 1 mM, and the culture was grown at 28 °C for a further 12 h.

The cells were harvested and suspended. After ultrasonic cell disintegration, the cell lysate was heated at 60 °C for 10 min and centrifuged to remove the cell debris and heat-induced aggregated proteins. The supernatant was dialysed against 50 mM Tris–HCl buffer (pH 8.5) and loaded on a Q Sepharose Fast Flow (anion-exchange) column (GE Healthcare). After the column was washed with 50 mM Tris–HCl buffer (pH 8.5), the enzyme protein was eluted by 1 M NaCl in 50 mM Tris–HCl buffer (pH 8.5). The active fractions were pooled, dialyzed against 50 mM Tris–HCl buffer (pH 7.0), concentrated, and stored at 4 or −20 °C. Enzyme purity was determined by SDS-PAGE analysis. Protein concentrations were determined according to the Bradford method, and bovine serum albumin was used as the standard.

2.4. Assay of enzyme activity

The hydrolytic activities of the cellulase were measured after 5 or 30 min of incubation at 80 °C in 50 mM phosphate–citrate buffer (pH 5.5) in the presence of 1% (w/v) polysaccharides. The polysaccharide substrates were CMC (carboxymethyl cellulose sodium salt, medium viscosity; Fluka), RAC (regenerated amorphous cellulose; being prepared as described by Zhang et al.) [27], Avicel (PH-101; Fluka), xylan from birch wood (Fluka), β-D-glucan from barley (β-1,4–1,3-glucan; Sigma), soluble starch, galactomannan (locust bean gum from *Ceratonia siliqua* seeds;

Sigma), mannan from *Saccharomyces cerevisiae* (Sigma). The reducing sugars released from the substrates were determined with the 3,5-dinitrosalicylic acid reagent [28]. If the amount of the reducing sugars was too little to be determined, we appropriately prolonged the incubation time or increased the enzyme concentration. One unit (IU) of enzyme activity was defined as the amount of enzyme that releases 1 μmol reducing sugars per min. Hydrolysis of *p*-nitrophenyl β-D-cellobioside (*p*NPC) was measured at 80 °C in 50 mM phosphate–citrate buffer (pH 5.5), and one unit (IU) of activity was defined as the amount of enzyme liberating 1 μmol of *p*-nitrophenol per min.

2.5. Effects of pH and temperature on the activity

With CMC being used as a substrate, the pH and temperature optima of the cellulase were determined by measuring the release of reducing sugars from CMC. The pH optimum was determined by measuring the enzyme activity at 80 °C in various pH buffers. The temperature optimum was determined in 50 mM phosphate–citrate buffer (pH 5.5) at different temperatures.

2.6. Temperature stability

Thermal stability studies were carried out by incubating about 2 mg/ml purified enzyme solutions for various lengths of time at different batches of temperatures in 50 mM Tris–HCl buffer (pH 7.0). The residual activity on CMC was determined at 80 °C in 50 mM phosphate–citrate buffer (pH 5.5).

3. Results

3.1. Sequence analysis, expression and purification of the cellulase from F. nodosum

The gene Fnod_1560 of the completely sequenced genome of *F. nodosum* Rt17-B1 (Reference Sequence accession: NC_009718.1; region: 1698601..1699632) has been predicted to encode a cellulase (glycoside hydrolase family 5) by the Pfam web services (http://pfam.janelia.org) [29]. And sequence analysis does not predict the presence of carbohydrate-binding module (CBM) in the putative cellulase protein, which module is considered to have a targeting and disruptive function on crystal cellulose [30]. SignalP 3.0 server (http://www.cbs.dtu.dk/services/SignalP) [31] predicts the presence of a signal peptide for the first 24 N-terminal amino acids. The presence of a signal peptide indicates that the natural enzyme may be a secreted protein, which enables *F. nodosum* to use carbohydrates as carbon source. To express the mature protein in *E. coli*, this putative signal peptide was deleted by the cloning strategy, substituted by a Met residue encoded by an AUG start codon. The mature protein of 320 amino acids which was expressed using *E. coli* cells, exhibited its abilities to hydrolyze cellulosic substrates (Table 1). So we suggest that this protein could be designated as *Fn*Cel5A according to the nomenclature for the enzymes hydrolyzing polysaccharides in plant cell walls [32], the gene encoding this protein was designated as *fncel5a*.

The gene encoding the cellulase contains 19 rare codons (AGA, AGG and AUA), and this bias may represent a mechanism of adaptation to protein thermostability [33]. Therefore the *E. coli* BL21-CodonPlus (DE3)-RIL cells, which contain extra copies of the *argU* (AGA, AGG), *ileY* (AUA), and *leuW* (CUA) tRNA genes, were used for expression. The cellulase was solubly overexpressed in *E. coli* cells harboring plasmid p*Fn*Cel5A and purified by anion-exchange chromatography. The SDS-PAGE analysis showed the contents, purities, and molecular mass of the recombinant cellulase protein (Fig. 1). The theoretical molecular mass of the cellulase protein is 37.6 kDa. Although the SDS-PAGE result of the cellulase was smaller than the theoretical mass, the mass spectrogram of the cellulase (data not shown) indicated the molecular mass is in good agreement with the theoretical mass. So we are sure that the enzyme protein was a completely expressed product of the gene. We speculate that the anomalous behaviour of this protein on SDS-PAGE may be correlated with the stabilization mechanisms which make enzymes from thermophilic extremophiles more compact. For instance, the enzyme in this work contains a large proportion of charged amino acids (8.4% Lys, 3.8% Arg, 5.3% Asp, 8.8% Glu), which can participate in a large network of ion pairs with an increase of protein compactness [34]. Based on the result that the cellulase in this work still had measurable residual activities in 0.25–0.5% (w/v) SDS solutions (Table 2), we think that the cellulase protein would tend to retain as compact structures as possible in order to reduce the loss of enzymatic activity in low concentration of SDS. So the cellulase protein tending towards compactness might migrate faster than the completely denatured polypeptides acting like long rods in the SDS-PAGE buffers containing 0.1% SDS.

Table 1
Activities of the cellulase from *F. nodosum* on various substrates.

Substrate	Specific activity (IU/mg of protein)	Relative activity (%)
CMC	440	100
β-D-Glucan	1360	309
Galactomannan	895	203
RAC	402	91
Avicel	ND[a]	–
Mannan	ND	–
Xylan	ND	–
Soluble starch	ND	–
*p*NPC[b]	0.15	0.034

[a] ND, no detectable activity.
[b] *p*NPC, *p*-nitrophenyl β-D-cellobioside.

3.2. Enzymatic properties and substrate specificity of the cellulase from F. nodosum

The effects of temperature and pH on the activity of the cellulase were measured using CMC as the substrate. We found that

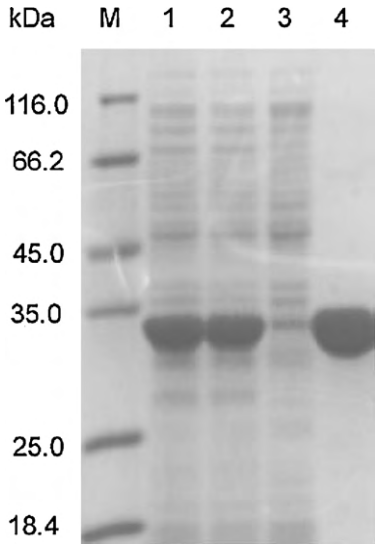

Fig. 1. Analysis of the recombinant protein on SDS-PAGE. Lane M, protein molecular weight marker (116.0, 66.2, 45.0, 35.0, 25.0, 18.4 kDa); lane 1, the cell lysate supernatant of the *E. coli* cells expressing the cellulase from *F. nodosum*; lane 2, supernatant of the cell lysate sample after incubating in 60 °C for 10 min; lane 3, the cell debris and heat-induced aggregated proteins from expression cells; lane 4, enzyme sample purified by a Q Sepharose Fast Flow (anion-exchange) column.

Table 2
Effects of various cations, chelating agent and surfactants on the enzyme activity.

Effectors	Relative activity (%)	
Control	100	100
Cations	5 mM	10 mM
$MnCl_2$	105	112
$CaCl_2$	103	105
NH_4Cl	100	104
$MgCl_2$	95.8	99.5
$BaCl_2$	93.3	107
NaCl	93.3	93.1
$CoCl_2$	79.1	63.1
$NiCl_2$	72.7	62.9
$ZnCl_2$	3.10	2.92
Chelating agent	5 mM	10 mM
EDTA	75.6	92.2
Surfactants	0.25% (w/v)	0.5% (w/v)
Triton X-100	98.8	89.8
SDS	2.4	0.2

the optimum temperature for activity is 80–83 °C (Fig. 2A), and the optimum pH is 5.0–5.5 (Fig. 2B).

The hydrolytic activities towards various substrates were measured at 80 °C in pH 5.5 buffer, as listed in Table 1. The cellulase showed very high activities on CMC (440 IU/mg), RAC (402 IU/mg), β-D-glucan from barley (1360 IU/mg) and galactomannan (895 IU/mg), and very low activity on *p*NPC. But it has relatively low or no activity on Avicel (microcrystalline cellulose)

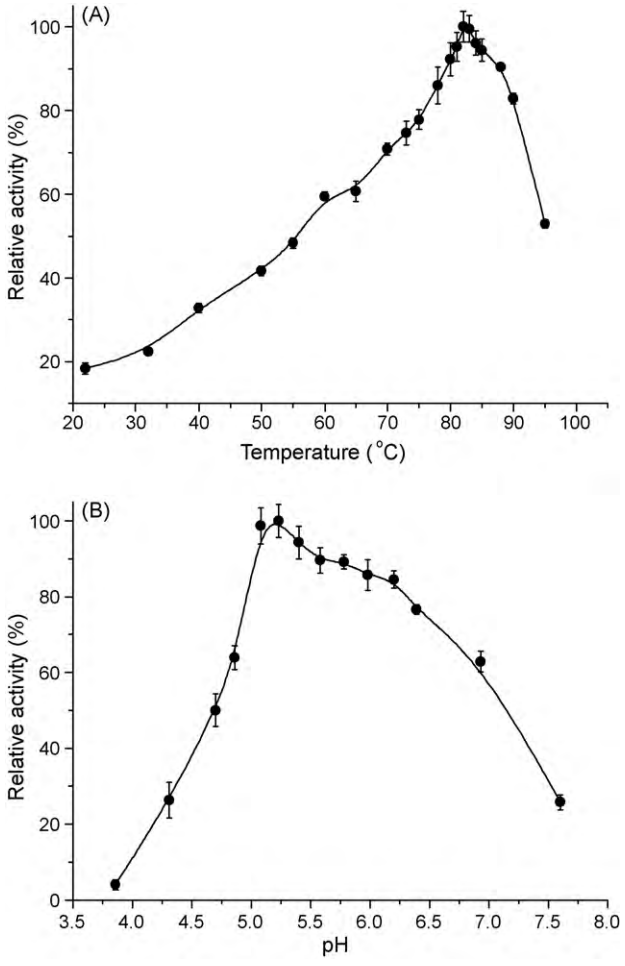

Fig. 2. Effects of temperature and pH on activity of the cellulase from *F. nodosum*. (A) Effect of temperature on the activity of the cellulase on CMC and (B) effect of pH on the activity of the cellulase on CMC.

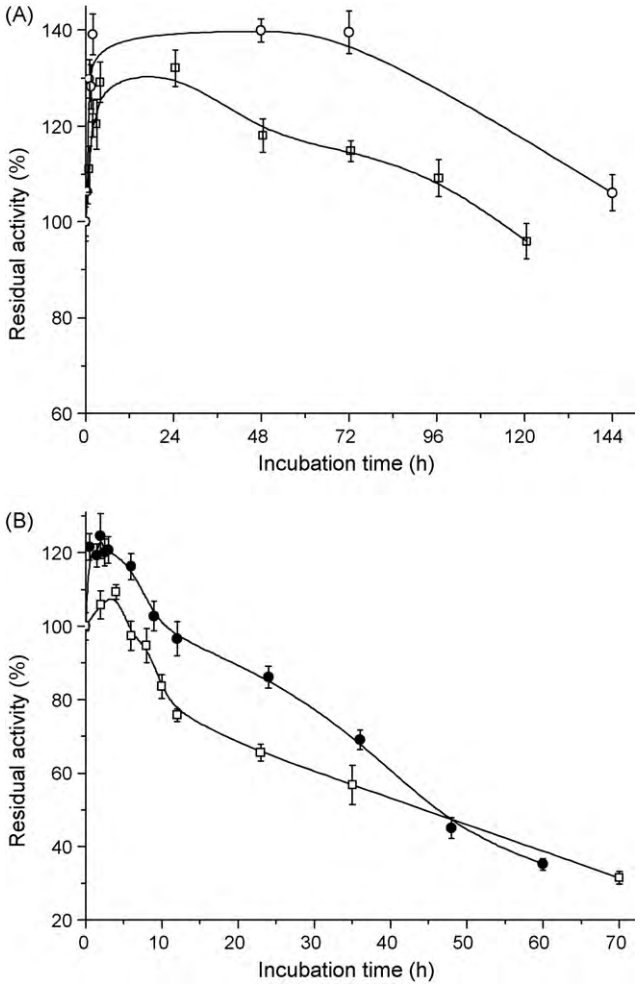

Fig. 3. Thermal stability of the cellulase from *F. nodosum* at different incubation temperatures. The purified cellulase was incubated for various lengths of time at 70 °C (■), 75 °C (○), 80 °C (●) and 83 °C (□). Residual activities on CMC were assayed at 80 °C in 50 mM phosphate–citrate buffer (pH 5.5).

because of lacking any CBM with a disruptive function on the crystalline structure of cellulose [30]. In our later work, the chimeric cellulases, which were the recombinant cellulase proteins linked with some heterologous CBMs, have obvious activities on Avicel (data not shown here).

The effects of various cations at 5 or 10 mM concentration, chelating agent at 5 or 10 mM, and surfactants at 0.25 or 0.5% (w/v) on enzyme activity were examined at 80 °C and pH 5.5 using CMC as the substrate (Table 2). No significant effect was observed in the presence of Mn^{2+}, Ca^{2+}, NH_4^+, Mg^{2+}, Ba^{2+}, Na^+, or surfactant Triton X-100; Co^{2+}, Ni^{2+} and EDTA appreciably inhibited the activity of the enzyme; $ZnCl_2$ and SDS obviously inhibited the enzyme activity.

3.3. Temperature stability of the cellulase from F. nodosum

We incubated the enzyme solutions in 50 mM Tris–HCl buffer (pH 7.0) for various lengths of time at 70, 75, 80 and 83 °C, and then measured the residual activities on CMC. It can be seen from Fig. 3 that the cellulase is very stable at 70–83 °C. The half-lives of the cellulase at 80 and 83 °C were about 48 and 46 h (Fig. 3B), respectively. The residual activities of the cellulase after incubation at 70 or 75 °C were not obviously on the decrease for over 5 days (Fig. 3A). Incubation at 70–83 °C for the first hour even stimulated the activity. Probably, as an enzyme working at high temperatures, the recombinant cellulase protein might not fold into the best functional

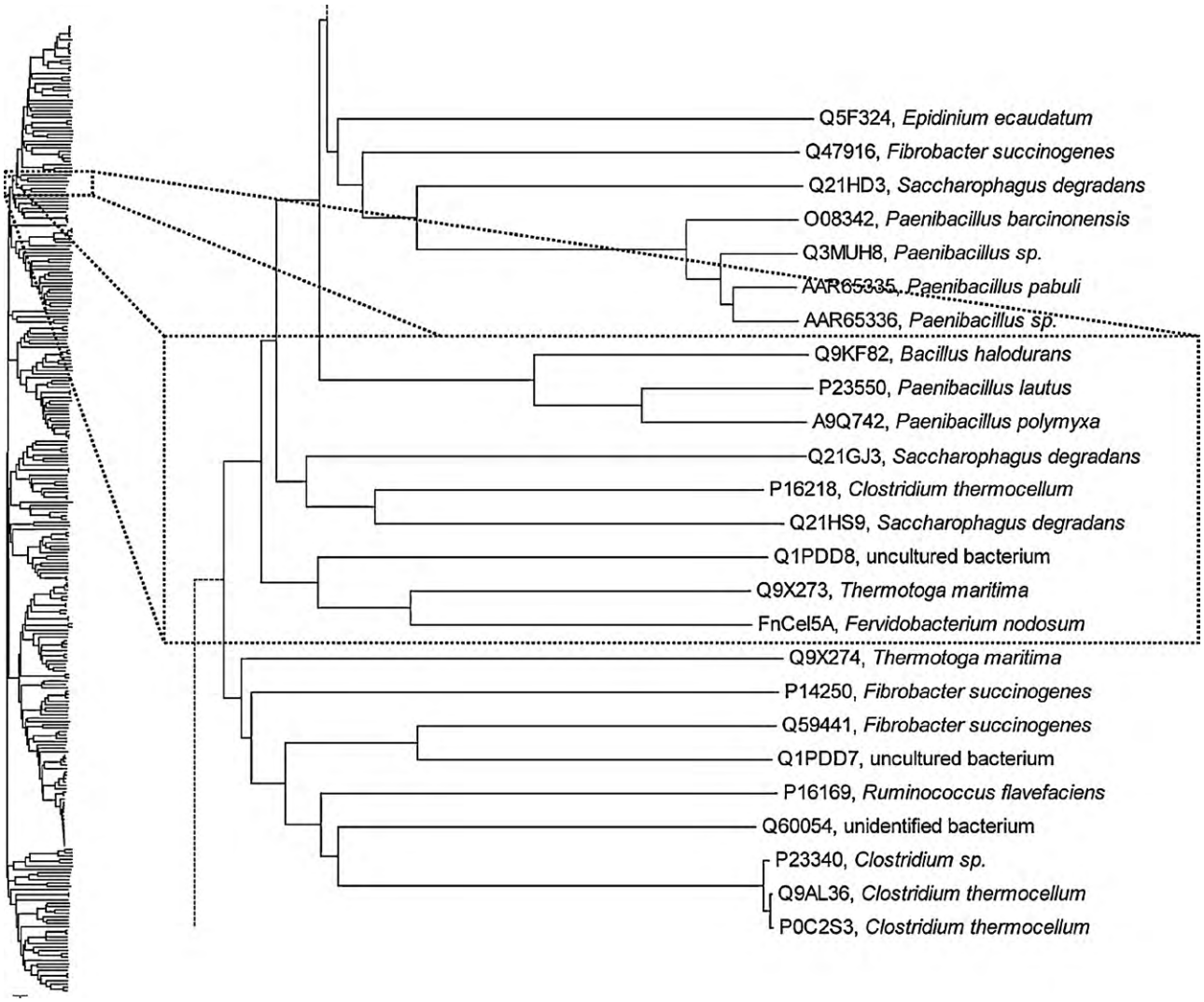

Fig. 4. Unrooted phylogenetic tree for the glycoside hydrolases related to the cellulase from *F. nodosum*. ClustalX 2.0.12 was used for alignment with default parameters, construction of a phylogenetic tree with the neighbor-joining method and bootstrap analysis of N-J trees. The tree was drawn using MEGA 4.0 with computing linearized tree. The GH5 phylogenetic tree is simplified at the left side. And the phylogenetic relationships of some glycoside hydrolases closely related to the cellulase from *F. nodosum* (identified by FnCel5A) in this tree are shown at the right side in detail. These hydrolases are represented by their accession numbers of the databases and the names of their source organisms.

three-dimensional structure when it was expressed and stored at low temperatures, and needs to be activated by heat treatment to refold and form the natural active conformation.

3.4. Analyses of phylogeny and catalytic amino acid residues

We aligned 378 protein sequences of glycoside hydrolase family 5 in our dataset (see Section 2 for details), and constructed an unrooted phylogenetic tree, which was displayed in the simplified form at the left part of Fig. 4 and the detail one of some enzymes closely related to this cellulase at the right part of Fig. 4. Those nine enzymes closely related to the cellulase in the broken lines box (Fig. 4) were aligned once again, and the alignment result is shown in Fig. 5. According to the suggestion of Bortoli-German et al. [35] and Chhabra et al. [17], we presumed that the conserved amino acid residues of the cellulase in GH5 might include Arg59, His103, Asn143, Glu144, His203, Tyr205, Glu260 and Trp293, which are marked with black boxes (Fig. 5), and that His103, Glu144, His203 and Glu260 may be the catalytic amino acid residues, moreover two Glu residues might be the catalytic nucleophile and proton donor, respectively.

4. Discussion

Members of the Thermotogales including *F. nodosum* have been isolated from various high temperature environments, which may produce a number of thermostable enzymes including cellulases [16–18]. The recently released genome sequence of *F. nodosum* Rt17-B1 revealed the presence of a glycoside hydrolase family 5 protein. A PSI-BLAST search [36] in the GenBank database (data not shown) and the phylogenetic analysis (similar to Fig. 4) showed that lots of proteins closely related to this protein were cellulases, and therefore we thought this putative protein might be a cellulase. According to these predictions, we obtained a new thermostable cellulase via genetic engineering technology. In this paper, we describe expression, purification and enzymatic characterization of this novel thermostable cellulase. This cellulase is remarkably active and stable at high temperatures. Specifically, the cellulase shows a high hydrolase activity against CMC at 70–83 °C even for a few days. The intrinsic thermostability of this cellulase, which implies possibilities for prolonged storage and low activity losses during long-time processing at the elevated temperatures, could have some obvious advantages in biochemical experiments, raw

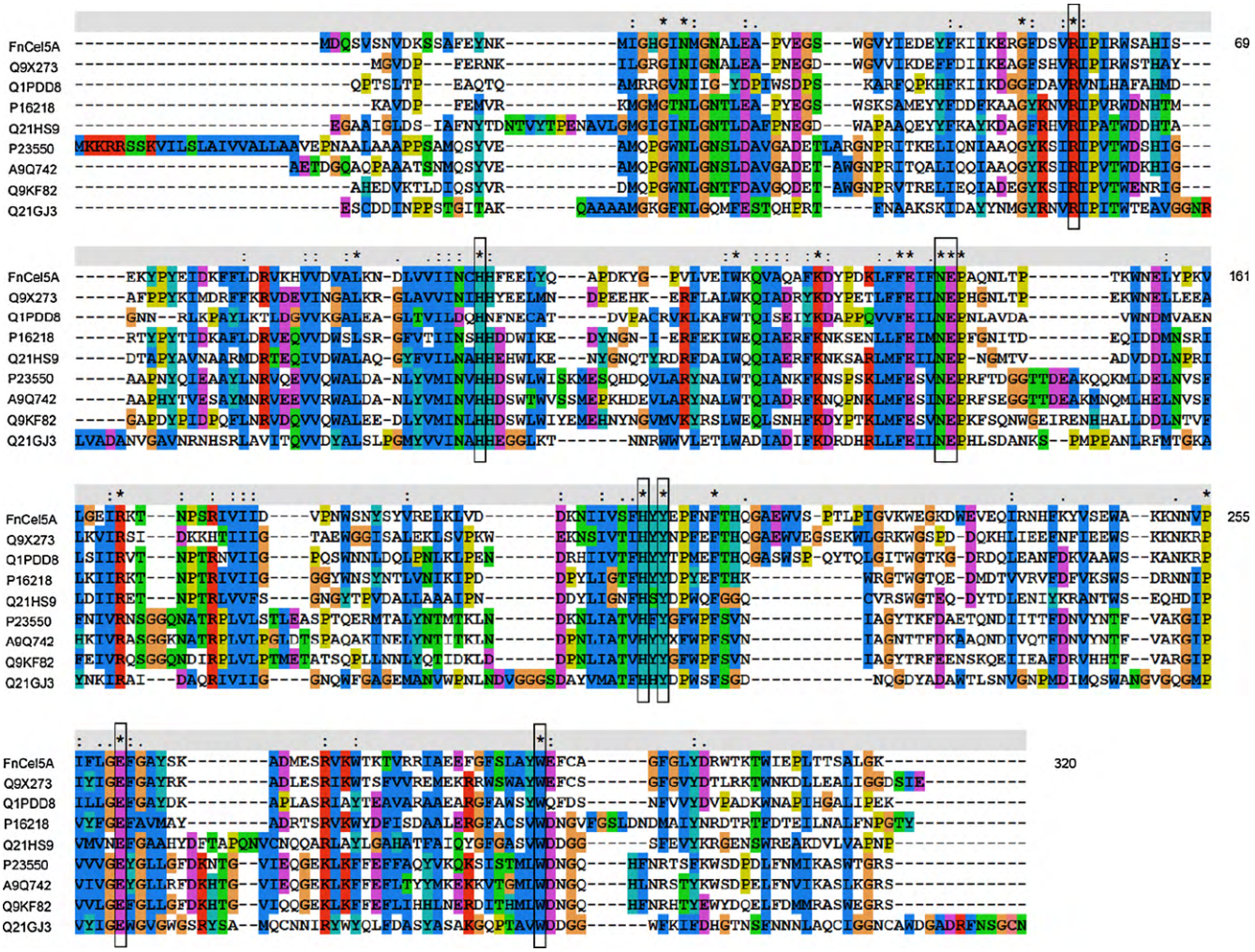

Fig. 5. Multiple sequence alignments for the cellulase from *F. nodosum* and other related proteins. *Fn*Cel5A, the cellulase from *F. nodosum* Rt17-B1; Q9X273, endoglucanase Cel5A from *T. maritima* MSB8; Q1PDD8, endo-1,4-β-D-glucanase from uncultured bacterium; P16218, endo-1,4-glucanase Cel5A from *C. thermocellum* NCIB 10682; Q21HS9, endo-β-1,4-glucanase Cel5B from *S. degradans* 2-40; P23550, endo-1,4-glucanase B from *P. lautus* PL236; A9Q742, cellulase Cel5B from *P. polymyxa* GS01; Q9KF82, β-glucanase CelB from *B. halodurans* C-125; Q21GJ3, endo-β-1,4-glucanase Cel5E from *S. degradans* 2-40. The black box regions are the conserved regions in GH5 enzymes which we presumed. An asterisk indicates identical or conserved residues; a colon indicates conserved substitutions; a period indicates semiconserved substitutions.

cellulosic material pre-treatments and a broad range of industrial processes.

Moreover, this cellulase can rapidly hydrolyze not only polysaccharide substrates derived from cellulose, e.g. CMC and RAC, but also some hemicellulosic substrates such as galactomannan (Table 1). This enzyme shows little activity on Avicel which has been used for measuring exoglucanase activity [2], and therefore is not an exoglucanase cellulase. Because CMC is often used for determining endoglucanase activity and RAC is a very good homogeneous substrate for cellulase activity assays [2], we could confirm that this enzyme is a cellulase with the endoglucanase (endo-1,4-β-D-glucanohydrolase) activity when it showed high activities on CMC and RAC. A high activity on β-D-glucan from barley may be due to good solubility of this substrate in water. In addition, as a sort of hemicellulosic substrate [8], galactomannan could be hydrolyzed by this enzyme which shows an activity on hemicellulose. So we concluded that this enzyme is an endoglucanase cellulase (EC 3.2.1.4) with some hydrolytic activities on hemicellulosic substrates. Moreover, the enzymes closely related to this cellulase (Fig. 4) have been reported to be able to hydrolyze several polysaccharide substrates including hemicellulosic substrates, for instance, the endoglucanase Cel5A (Q9X273) from *T. maritima* has activities on mannan, glucomannan and xyloglucan [17], endoglucanase Cel5A from *C. thermocellum* is able to hydrolyze larchwood xylan [37]. Furthermore, the information which had been collected at the web of CAZy in December 2009 showed that GH5 enzymes had 13 known activities of glycoside hydrolases and this number is on the increase. And the GH5 catalytic domains of enzymes have similar $(\beta/\alpha)_8$ barrel protein structures and same retaining catalytic mechanism including one Glu residue as catalytic nucleophile/base and another Glu residue as catalytic proton donor, which may be Glu144 and Glu260 in this cellulase. These facts may be suggest that GH5 enzymes, which have similar structures and catalytic mechanism, could hydrolyze a variety of substrates, just by making some flexible, slight and simple modification. This serves as a good inspiration for us, and it is also an encouragement for people like us that we could find more lots of different types of substrates being hydrolyzed by GH5 enzymes including this cellulase, and may extend the substrate range of this cellulase via the adjustments of the enzyme molecules by artificial molecular design and modification. Actually we had been working with engineering the cellulase molecule to enlarge its application scope based on its high activities and outstanding stability.

This cellulase is the first cellulase cloned from the organisms of genus *Fervidobacterium*. Genus *Fervidobacterium* belongs to the order Thermotogales which is the unique order of phylum Thermo-

togae and includes only one family, namely, the Thermotogaceae. The Thermotogales was placed near the root of the cultured bacterial phylogenetic subtree [38,39]. These organisms were regarded as most slowly evolving lineages in the Eubacteria. On the other hand, the genome plasticities are important features for genetic evolution of the Thermotogales [21,22,40,41], and their genome sequences have been frequently affected by HGT [42,43]. These changes within genes in the order Thermotogales have conveyed a selective advantage to make them increasing their metabolic fitness in their living environments. The preferential transfers are usually related to metabolic genes, for instance cellulase genes, and acquisition of such genes would allow organisms to gain access to new energy and nutrient sources increasing their ability to the environment [43]. HGT events involved in the degradation of carbohydrates from plant cell walls, such as cellulose, indicate an adaptation to the carbohydrate-rich environment [44]. In addition, the phylogenetic tree (Fig. 4) shows that the glycoside hydrolases related to this cellulase are distributed in distantly related organisms. Therefore, according to these features of the order Thermotogales and HGT, and combining with the observational results of the biochemical experiments showing that this enzyme is an excellent cellulase with high activity and stability, we suggest that the gene encoding this cellulase was acquired by its source organism via HGT and had maybe undergone repetitious HGT processes and adaptive innovations among various organisms, and this enzyme is a result of accelerated and relatively complete evolution.

These characteristics and evolutionary processes of this cellulase make it possible for the effective gene transfer and expression in other heterologous organisms, our work of cloning and expressing this gene in *E. coli* is a good example. This cellulase can be overexpressed in *E. coli* with a high protein content, good solubility in water, and a simple process to remove cell debris and heat-induced aggregated impure proteins by heating and centrifugation (Fig. 1). It is the simple and efficient production process coupled with high activity and stability that may make this cellulase becoming a potential industrial enzyme to treat plant cellulose fibers or convert cellulose into fermentable sugars, which may be important raw materials of biofuels. The easy transfer and expression of this gene in other heterologous organisms make it possible to construct a variety of recombinant organisms containing this gene for different applications.

Endoglucanases and exoglucanases can cooperate synergistically to degrade crystalline cellulose, e.g. Avicel (microcrystalline cellulose). Endoglucanases cleave randomly β-1,4-glucosidic bonds only on external cellulose chains made accessible by the peeling or progressive surface erosion action of exoglucanases [27], and generate new cellulose polysaccharide chain ends that can be attacked by exoglucanases [5]. Furthermore, some CBMs with a non-catalytic disruptive function on the crystalline structure of cellulose [30] can destroy the crystal packing of crystalline cellulose and guiding a single polysaccharide chain into the catalytic domains of cellulases. Although CBMs play a key role in hydrolyzing crystalline cellulose and are particularly important for the initiation and processive movement of exoglucanases, the presence of CBMs is not essential for endoglucanase activity [5]. Because of lacking CBMs with a disruptive function, the endoglucanase cellulase in this work has little activity on Avicel. But this cellulase is a good original material for the enzyme engineering and design. Recently, we were successful in introducing several heterologous CBMs to this cellulase, and obtained some chimeric cellulases which have obvious activities on crystal cellulose. Work is in progress to design and select a new type of cellulases which could more efficiently hydrolyze natural plant cellulose and participate in converting enormous and renewable biomass into fuels.

5. Conclusion

A novel cellulase gene encoding a thermostable endoglucanase from the thermophilic eubacterium *F. nodosum* Rt17-B1 was cloned and solubly expressed in *E. coli*, and the recombinant enzyme was purified and characterized. This enzyme is an endoglucanase cellulase with some hydrolytic activities on hemicellulosic substrates. At high temperatures the enzyme shows high hydrolytic activities and thermostability. The enzyme is expected to be useful for hydrolysis of cellulosic and hemicellulosic substrates at high temperatures, particularly for converting biomass into biofuels.

Acknowledgement

This work was supported by the High Technology Project of China (863 Program, 2006AA020202, and 2006AA020203).

References

[1] A.J. Ragauskas, C.K. Williams, B.H. Davison, G. Britovsek, J. Cairney, C.A. Eckert, W.J. Frederick Jr., J.P. Hallett, D.J. Leak, C.L. Liotta, J.R. Mielenz, R. Murphy, R. Templer, T. Tschaplinski, Science 311 (2006) 484–489.
[2] Y.H. Percival-Zhang, M.E. Himmel, J.R. Mielenz, Biotechnol. Adv. 24 (2006) 452–481.
[3] M.C. Chang, Curr. Opin. Chem. Biol. 11 (2007) 677–684.
[4] E.M. Rubin, Nature 454 (2008) 841–845.
[5] L.R. Lynd, P.J. Weimer, W.H. van Zyl, I.S. Pretorius, Microbiol. Mol. Biol. Rev. 66 (2002) 506–577.
[6] B.C. Saha, Biotechnol. Adv. 18 (2000) 403–423.
[7] R. Kumar, S. Singh, O.V. Singh, J. Ind. Microbiol. Biotechnol. 35 (2008) 377–391.
[8] M. Pauly, K. Keegstra, Plant J. 54 (2008) 559–568.
[9] N.S. Parachin, S. Siqueira, F.P. de Faria, F.A.G. Torres, L.M.P. de Moraes, J. Mol. Catal. B: Enzym. 59 (2009) 52–57.
[10] M. Zhang, Z. Jiang, L. Li, P. Katrolia, J. Mol. Catal. B: Enzym. 60 (2009) 119–124.
[11] D.Y. Kim, M.K. Han, H.W. Oh, D.S. Park, S.J. Kim, S.G. Lee, D.H. Shin, K.H. Son, K.S. Bae, H.Y. Park, J. Mol. Catal. B: Enzym. 62 (2010) 32–39.
[12] B. Henrissat, Biochem. J. 280 (Pt 2) (1991) 309–316.
[13] B.L. Cantarel, P.M. Coutinho, C. Rancurel, T. Bernard, V. Lombard, B. Henrissat, Nucleic Acids Res. 37 (2009) D233–D238.
[14] C. Vieille, G.J. Zeikus, Microbiol. Mol. Biol. Rev. 65 (2001) 1–43.
[15] P. Turner, G. Mamo, E.N. Karlsson, Microb. Cell Fact. 6 (2007) 9.
[16] J.D. Bok, D.A. Yernool, D.E. Eveleigh, Appl. Environ. Microbiol. 64 (1998) 4774–4781.
[17] S.R. Chhabra, K.R. Shockley, D.E. Ward, R.M. Kelly, Appl. Environ. Microbiol. 68 (2002) 545–554.
[18] S.R. Chhabra, R.M. Kelly, FEBS Lett. 531 (2002) 375–380.
[19] L. Viikari, M. Alapuranen, T. Puranen, J. Vehmaanperä, M. Siika-Aho, Adv. Biochem. Eng. Biotechnol. 108 (2007) 121–145.
[20] B.K.C. Patel, H.W. Morgan, R.M. Daniel, Arch. Microbiol. 141 (1985) 63–69.
[21] K.E. Nelson, R.A. Clayton, S.R. Gill, M.L. Gwinn, R.J. Dodson, D.H. Haft, E.K. Hickey, J.D. Peterson, W.C. Nelson, K.A. Ketchum, L. McDonald, T.R. Utterback, J.A. Malek, K.D. Linher, M.M. Garrett, A.M. Stewart, M.D. Cotton, M.S. Pratt, C.A. Phillips, D. Richardson, J. Heidelberg, G.G. Sutton, R.D. Fleischmann, J.A. Eisen, O. White, S.L. Salzberg, H.O. Smith, J.C. Venter, C.M. Fraser, Nature 399 (1999) 323–329.
[22] O. Zhaxybayeva, K.S. Swithers, P. Lapierre, G.P. Fournier, D.M. Bickhart, R.T. DeBoy, K.E. Nelson, C.L. Nesbø, W.F. Doolittle, J.P. Gogarten, K.M. Noll, Proc. Natl. Acad. Sci. U.S.A. 106 (2009) 5865–5870.
[23] J. Schultz, F. Milpetz, P. Bork, C.P. Ponting, Proc. Natl. Acad. Sci. U.S.A. 95 (1998) 5857–5864.
[24] I. Letunic, T. Doerks, P. Bork, Nucleic Acids Res. 37 (2009) D229–D232.
[25] M.A. Larkin, G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, H. McWilliam, F. Valentin, I.M. Wallace, A. Wilm, R. Lopez, J.D. Thompson, T.J. Gibson, D.G. Higgins, Bioinformatics 23 (2007) 2947–2948.
[26] S. Kumar, M. Nei, J. Dudley, K. Tamura, Brief Bioinform. 9 (2008) 299–306.
[27] Y.H. Zhang, J. Cui, L.R. Lynd, L.R. Kuang, Biomacromolecules 7 (2006) 644–648.
[28] G.L. Miller, Anal. Chem. 31 (1959) 426–428.
[29] R.D. Finn, J. Tate, J. Mistry, P.C. Coggill, S.J. Sammut, H.R. Hotz, G. Ceric, K. Forslund, S.R. Eddy, E.L. Sonnhammer, A. Bateman, Nucleic Acids Res. 36 (2008) D281–D288.
[30] A.B. Boraston, D.N. Bolam, H.J. Gilbert, G.J. Davies, Biochem. J. 382 (2004) 769–781.
[31] J.D. Bendtsen, H. Nielsen, G. von Heijne, S. Brunak, J. Mol. Biol. 340 (2004) 783–795.
[32] B. Henrissat, T.T. Teeri, R.A.J. Warren, FEBS Lett. 425 (1998) 352–354.
[33] M.G. Van der Linden, S.T. de Farias, Extremophiles 10 (2006) 479–481.
[34] C. Cambillau, J.M. Claverie, J. Biol. Chem. 275 (2000) 32383–32386.
[35] I. Bortoli-German, J. Haiech, M. Chippaux, F. Barras, J. Mol. Biol. 246 (1995) 82–94.
[36] S.F. Altschul, T.L. Madden, A.A. Schäffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipman, Nucleic Acids Res. 25 (1997) 3389–3402.

[37] E. Yagüe, P. Béguin, J.P. Aubert, Gene 89 (1990) 61–67.
[38] N.R. Pace, Science 276 (1997) 734–740.
[39] Y.I. Wolf, I.B. Rogozin, N.V. Grishin, E.V. Koonin, Trends Genet. 18 (2002) 472–479.
[40] E.F. Mongodin, I.R. Hance, R.T. Deboy, S.R. Gill, S. Daugherty, R. Huber, C.M. Fraser, K. Stetter, K.E. Nelson, J. Bacteriol. 187 (2005) 4935–4944.
[41] C.L. Nesbø, M. Dlutek, W.F. Doolittle, Genetics 172 (2006) 759–769.
[42] H. Ochman, J.G. Lawrence, E.A. Groisman, Nature 405 (2000) 299–304.
[43] R.G. Beiko, T.J. Harlow, M.A. Ragan, Proc. Natl. Acad. Sci. U.S.A. 102 (2005) 14332–14337.
[44] G. Ricard, N.R. McEwan, B.E. Dutilh, J.P. Jouany, D. Macheboeuf, M. Mitsumori, F.M. McIntosh, T. Michalowski, T. Nagamine, N. Nelson, C.J. Newbold, E. Nsabimana, A. Takenaka, N.A. Thomas, K. Ushida, J.H. Hackstein, M.A. Huynen, BMC Genomics 7 (2006) 22.