

Highly thermostable endo-1,3- β -glucanase (laminarinase) LamA from *Thermotoga neapolitana*: nucleotide sequence of the gene and characterization of the recombinant gene product

Vladimir V. Zverlov,¹† Ilia Y. Volkov,¹ Tatjana V. Velikodvorskaya¹ and Wolfgang H. Schwarz²

Author for correspondence: Wolfgang H. Schwarz. Tel. +49 89 2892 2302. Fax: +49 89 2892 2360. e-mail: wolfgang.schwarz@mikro.biologie.tu-muenchen.de

¹ Institute of Molecular Genetics, Russian Academy of Science, Kurchatov Sq. 46, 123 182 Moscow, Russia

² Institute for Microbiology, Technical University of München, Arcisstr. 21, D-80290 München, Germany

The nucleotide sequence of clone pTT26 (3786 bp), containing the gene for 1,3- β -glucanase LamA (laminarinase) from *Thermotoga neapolitana*, was determined. It contains an ORF encoding a protein of 646 aa (73 328 Da). The central part of the protein is homologous to the complete catalytic domain of bacterial and some eukaryotic endo-1,3- β -D-glucanases and belongs to family 16 of glycosyl hydrolases. This domain is flanked on both sides by one copy on each side of a substrate binding domain homologue (family II). The recombinant laminarinase protein was purified from *Escherichia coli* host cells in two forms, a 73 kDa and a processed 52 kDa protein, both having high specific activity towards laminarin (3100 and 2600 U mg⁻¹, respectively) and K_m values of 2.8 and 2.2 mg ml⁻¹, respectively. Limited activity on 1,3-1,4- β -glucan (lichenan) was detected (90 U mg⁻¹). Laminarin was degraded in an endoglucanase modus, yielding glucose, laminaribiose and -triose as end products. Thus LamA classifies as an endo-1,3(4)- β -glucanase (EC 3.2.1.6). The optimum temperature of the enzymes was 95 °C (73 kDa) and 85 °C (52 kDa) at an optimum pH of 6.2. The superior thermostability of the 73 kDa enzyme is demonstrated by incubation without substrate at 100 °C, where 57% of the initial activity remained after 30 min (82% at 95 °C). Thus, LamA is the most thermostable 1,3- β -glucanase described to date.

Keywords: *Thermotoga neapolitana*, laminarinase, 1,3- β -glucanase, thermostability

INTRODUCTION

Thermotoga neapolitana, a close relative to *T. maritima*, was isolated from marine environments with volcanic activity. It is a saccharolytic, extremely thermophilic, Gram-negative eubacterium which grows up to a maximum temperature of 90 °C (Huber & Stetter, 1992). It readily catabolizes mono- and polysaccharides.

Thermostability of its hydrolytic exoenzymes promises considerable biotechnological application potential. *Thermotogales* also comprise the deepest branch of both the archaeal and eubacterial lineages. Their genes are therefore ideally suited to investigate the evolutionary aspects of the origin and development of enzyme families.

1,3- β -Glucan is a main cell wall component in yeasts and filamentous fungi, and a major structural and storage polysaccharide (laminarin) of the marine macro-alga *Laminaria saccharina*. It also is produced as insoluble exopolysaccharide by some bacteria, e.g. *Alcaligenes faecalis* var. *myxogenes* (curdlan). The β -glucanases degrading 1,3- β -glucans can be divided into two hydrolytic types: exoglucanases releasing exclusively laminari-

† Present address: Institute for Microbiology, Technical University of München, Arcisstr. 21, D-80290 München, Germany.

Abbreviation: Im/HCl, imidazole/HCl.

The GenBank accession number for the nucleotide sequence reported in this paper is Z47974.

biose, produced by germinating cereal grains (e.g. malt), various yeasts and fungi and by some bacteria (e.g. *Bacillus* YK9, Kanzawa *et al.*, 1994); and endoglucanases produced by mesophilic bacteria, e.g. by *Bacillus circulans* strains, *Oerskovia xanthineolytica* (*Cellulomonas cellulans*) (Ferrer *et al.*, 1996) and by Streptomycetes, but also by thermophilic bacteria *Rhodothermus marinus* (Spilliaert *et al.*, 1994), *Clostridium thermocellum* (Schwarz *et al.*, 1988) and *Thermotoga maritima* (Bronnenmeier *et al.*, 1995). Most enzymes have not been characterized in detail. Amino acid sequences of these two enzyme classes are not homologous to each other and seem to have developed independently. They can also be distinguished by their degradation product pattern: whereas the exoglucanases hydrolyse the substrate sequentially and yield only laminaribiose (even on short incubation), the endoglucanases attack 1,3- β -glucan at random to a bulk of different laminari-oligosaccharides with decreasing average lengths on longer incubation. The oligosaccharide pattern in a complete hydrolysis is characteristic for a given enzyme.

1,3- β -Glucanases have a potential for commercial yeast extract production and for the conversion of algal biomass to fermentable sugars. They also have, in combination with chitinases, antimycotic activity for disease protection of plants (Nogi & Horikoshi, 1990). The bacterial endo-1,3- β -D-glucanases are classified in EC 3.2.1.39 (1,3- β -D-glucan glucanohydrolases), but some also have properties of endo-1,3(4)- β -glucanases classified in EC 3.2.1.6 [1,3-(1,3;1,4)- β -D-glucan 3(4)-glucanohydrolases].

A genomic library has previously been constructed from *T. neapolitana* strain Z2706-MC24. It was screened for enzymic activities related to cellulose and hemicellulose degradation. One of the clones (pTT26) was active on the 1,3- β -glucan laminarin and the mixed-linkage glucan lichenan (Dakhova *et al.*, 1993). In another report, a 1,3- β -glucanase activity was isolated from the culture supernatant of *T. maritima*, which is a closely related strain (Bronnenmeier *et al.*, 1995). This β -glucanase was not characterized further.

In this paper we describe the primary structure of the gene and the LamA protein of *T. neapolitana*. We discuss the relationship of the 1,3- β - to the 1,3-1,4- β -glucanases and present a characterization of the purified recombinant laminarinase, which is thermostable and extraordinarily active.

METHODS

Bacterial strains, plasmids and growth conditions. *Escherichia coli* TG1 {*supE hsdD5 thi* Δ (*lac-proAB*) [*F'* *traD36 proAB⁺ lacI^a lacZ* Δ M15]} was obtained from the Microbial Collection of IMG, RAS, Moscow. Plasmid pTT26, containing the laminarinase gene of *Thermotoga neapolitana* Z2706-MC24, has been described by Dakhova *et al.* (1993). Recombinant *E. coli* cells were grown at 37 °C in Luria-Bertani broth containing 100 μ g ampicillin ml⁻¹.

Cell fractionation. *E. coli*(pTT26) cells were collected by centrifugation and washed in 0.1 vols 25% (w/v) saccharose+0.1 M EDTA, resuspended in the same volume of distilled ice-cold water and shaken for 10 min in the cold. After centrifugation, the supernatant contained the periplasmic proteins. The pellet was resuspended in 0.1 vols 50 mM imidazole/HCl buffer (Im/HCl buffer, pH 6.8) and ultrasonicated (22 kHz, 15–20 min) on ice. The sonicate was centrifuged. The supernatant contained the cytoplasmic fraction; the pellet was resuspended in 50 mM Im/HCl buffer + 1% (v/v) Triton X-100.

Purification of the enzyme. *E. coli*(pTT26) cells (from 2 l culture) were ultrasonicated in 150 ml Im/HCl buffer + 0.1 M NaCl + 0.1% Triton X-100, heated to 85 °C for 40 min and centrifuged. The supernatant was applied to a 1.5 \times 8.0 cm DEAE Toyopearl 650M column, equilibrated with the same buffer and successively washed with Im/HCl buffer + 0.15 M NaCl and Im/HCl buffer + 0.25 M NaCl, and eluted with steps of Im/HCl buffer + 0.4 M NaCl + 0.1% Triton X-100 and Im/HCl buffer + 0.6 M NaCl + 0.1% Triton X-100. Protein fractions were ultrafiltered to wash out the Triton X-100.

DNA sequencing. The DNA sequence was determined from supercoiled double-stranded plasmid DNA by cycle sequencing of both strands (Boehringer Cycle Sequencing Kit) with biotinylated primers. DNA fragments were detected with a GATC 1500 Direct-Blotting Electrophoresis Apparatus using streptavidin-conjugated alkaline phosphatase and NBT-BCIP as chromogenic substrate (Tropix). Sequence data were analysed and compared with the DNASIS/PROSIS for Windows package (Hitachi Software Engineering). Nucleotide and protein sequence databases were screened using the FASTA and BLAST software at the NCBI server (URL address: <http://www.ncbi.nlm.nih.gov>).

Enzymic assays and detection of hydrolysis products. Enzyme aliquots in standard assays were incubated in sodium phosphate/citrate buffer (50 mM, pH 6.2) with 0.7% substrate at 85 °C. Reducing sugars released from polymeric substrates were detected by the 3,5-dinitrosalicylic acid method (Wood & Bhat, 1988). One unit of activity is defined by the release of 1 μ mol glucose equivalent min⁻¹. Digests of laminarin were applied to TLC (Silica-gel plates). Mono- and oligosaccharides were separated in multiple runs in acetonitrile/water (80:20) and detected by spraying with a solution of 1 g diphenylamine and 1 ml aniline in 100 ml acetone. For HPLC analysis (System Gold, Beckman Instruments), 0.5% laminarin was digested in MES buffer (25 mM, pH 6.0) at 90 °C; additional enzyme was added after 10 h. Hydrolysis products were separated at 85 °C in H₂O using an Aminex HPX-42A column (300 \times 7.8 mm; Bio-Rad) fitted with a microguard precolumn (Deash, Bio-Rad) and detected by refractometry.

Oat spelt xylan, CM-cellulose, lichenan and laminarin were obtained from Sigma and Aldrich. Laminaridextrins as size standards for HPLC and TLC were obtained from Seikagaku Co.

RESULTS

Sequencing of clone pTT26 DNA and sequence analysis

The nucleotide sequence (3786 bases) of the previously described clone pTT26 (Dakhova *et al.*, 1993) was determined. An ORF is apparent from base 851 to 2855

ending with an opal stop codon (Fig. 1). A potential ribosome binding site is located at base 908, leading to a coding sequence starting with ATG at base 918 and expressing a protein of 646 aa (73 328 Da). The ORF shows amino acid sequence similarity to bacterial 1,3- β -glucanases and was therefore called *lamA*. It is followed by a downstream 7 GC-pairs hairpin structure of $\Delta G = -22.2$ kcal mol⁻¹ (2864–2883) and a run of 4 T residues, a potential transcription terminator.

The *lamA* gene is preceded on the same DNA strand by an incomplete ORF (*orf1*; base 2–898) with a distance of only 3 bases between the stop codon and the *lamA* Shine–Dalgarno sequence. The 299 determined amino acid residues are 37.5% identical with, for example, the C-terminal region of the β -glucosidase *bglB* sequence of *Cl. thermocellum* (Gräbnitz *et al.*, 1989), a member of family B β -glucosidases. No secondary structure could be predicted in the DNA between *orf1* and *lamA*, and a read-through of at least these two genes seems to be possible. It would end at the palindromic structure downstream of *lamA*.

Within the amino acid sequence of *lamA*, three domains (A1–Cat–A2) can be distinguished in addition to a hydrophobic N terminus with the features of a typical signal peptide (van Heijne, 1986) (Fig. 2). The catalytic domain (Cat) can be assigned to family 16 of glycosyl hydrolases (Henrissat & Bairoch, 1993) by its homology to the complete catalytic domains of bacterial 1,3- β -glucanases and mixed-linkage glucanases of *Bacillus* and other bacteria. Especially high overall homology (42.8% identity in a 257 aa overlap) was found to thermostable *R. marinus* β -glucanase BglA (Spilliaert *et al.*, 1994). The N terminus of the catalytic domain can be exactly defined to amino acid residue Asp-215 by homology to the N termini of mature bacterial 1,3- β -glucanases, the β glII_A isoenzyme of *O. xanthineolytica* LL G109 (Ferrer *et al.*, 1996) and the 28 kDa and 42 kDa 1,3- β -glucanases of *B. circulans* IAM 1165 (Aono *et al.*, 1995). Domains A1 and A2 show homology to each other and to sequences of cellulose-binding domain family II (Coutinho *et al.*, 1992; Tomme *et al.*, 1995). They are not separated from the catalytic domain by obvious PTS-box reminding hinge regions. In contrast to the binding domain of *Cellulomonas fimi* CenC (N1, N2), where two consecutive tandem repeats of family II binding domains are present, in LamA the two copies of the binding domains are separated by the catalytic domain.

Purification of the recombinant 1,3- β -glucanase

The *lamA* gene of clone pTT26 was expressed in *E. coli* host cells in both orientations of the insert DNA, i.e. independently of external *E. coli* promoters. In the *E. coli* host, about 5% of total enzyme activity was present in the periplasmic fraction, 67% in the cytoplasmic fraction and 28% in the insoluble fraction (membrane fraction). Consequently, the enzyme was purified to electrophoretic homogeneity from the *E. coli* cell lysate

(2250 U, 480 mg protein) by heat precipitation (91% yield) followed by DEAE ion exchange chromatography (49% yield), leading to a 470-fold enrichment. Of great advantage for this simple purification scheme was the high negative charge of the protein at pH 6.8, resulting in a strong binding to the DEAE ion exchange material, which allowed for a washing step at high ionic strength before elution (0.25 M NaCl). Two fractions with homogeneous proteins of 52 and 73 kDa, as determined by SDS-PAGE, could be purified by elution from the column with 0.4 and 0.6 M NaCl, respectively. They had a specific activity of 2100 and 2500 U mg⁻¹, respectively. Due to the high hydrophobicity of the enzyme, it had a tendency for aggregation, which made the presence of Triton X-100 unavoidable in all chromatographic steps.

Characterization of the purified LamA

The two purified recombinant protein fractions (52 and 73 kDa) were capable of degrading laminarin (100%) and to a lesser degree lichenan (4.2%); both were inactive towards 1,4- β - and α -glucan substrates (0%) and xylan (0.07%). This characterizes the enzyme as laminarinase. HPLC analysis of the hydrolysis products of both enzyme species (Fig. 3) yielded intermittent laminaridextrins. These were efficiently degraded to glucose, laminaribiose and -triose, which were not further hydrolysed. This degradation pattern is characteristic for endoglucanases. Laminaribiose was not the only end product as is found for exo-1,3- β -glucanases (Kanzawa *et al.*, 1994). An identical pattern was described for the endo-1,3- β -glucanase of *R. marinus* (BglA, Spilliaert *et al.*, 1994) and *B. circulans* KCTC3004 (Lee & Chang, 1995). If laminarin was hydrolysed in the presence of slightly higher salt concentrations (0.1 M sodium phosphate/citrate buffer), intermittently formed laminaritriose was processed preferentially, possibly by transglucosylation, as shown by TLC (data not shown); glucose, laminaribiose and -tetraose were the end products.

Both proteins proved to be stable under extreme temperature conditions in the absence of substrate: 82% of the 73 kDa enzyme activity could still be observed after 30 min at 95 °C, and even 57% at 100 °C, whereas the 52 kDa enzyme was inactivated to 45% at 95 °C. EDTA (10 mM) and a pH above 7.5 or below 4.5 inactivated the 52 kDa enzyme at 95 °C by more than 90%. The 73 kDa enzyme was more stable at low pH.

LamA was even more stable in the presence of substrate: the enzyme had under standard incubation conditions an optimum activity at a temperature of 85 °C (52 kDa) and 95 °C (73 kDa). Optimum enzyme activity was observed around pH 6.1 (52 kDa) and pH 6.3 (73 kDa) (Fig. 4), both having more than two-thirds of their activity between pH 5 and 8.5. The buffer with the highest enzyme stability and activity was 50 mM Im/HCl buffer. A difference in temperature stability with laminarin or lichenan as substrate, as has been shown for *Cl. thermocellum* laminarinase LicA (Schwarz *et al.*, 1988), was not observed.

1 GATCATAAAACCAAACCTCCAGAAAACCTCCTTTCCGAGAAAGGAAATACACAACCTGGCGAAAAGAACGACGTGGCGGTTCATCGTGT 90
 1 I I K P K L P E N F L S E K E I H K L A K K N D V A V I V I 30

91 CAGCAGGATTTCCGAGAAAGGCTATGACAGAAAGCCGGTGAAGGGAGACTTTTACCTTTCTGACGATGAGACTGATCTCATAAAGACTGT 180
 31 S R I S G E G Y D R K P V K G D F Y L S D D E T D L I K T V 60

181 CTCCAGAGAGTCCATGAACAGGGCAAGAAAGTATCGTCTTCTCAACATAGGAAGTCTGTTGAGGTTGTTAGCTGGAGAGATCTGGT 270
 61 S R E F H E Q G K K V I V L L N I G S P V E V V S W R D L L V 90

271 GGATGGGATTTCTCTGTGTGGCAGGCAGGAAACCCGGCAGGATCGTTGCCGATGTTCTACTGGAAGGATCAATCCATCTGGAAA 360
 90 D G I L L V W Q A G Q E T G R I V A D V L T G R I N P S G K 120

361 ACTTCCAACCACCTTTCCGAGAGACTACTCTGATGTACCCTCCTGGACCTTTCTGGAGAGCCAAAGGACAATCCACAGAAGGTGGTCTA 450
 120 L P T T F P R D Y S D V P S W T F P G E P K D N P Q K V V Y 150

451 CGAAGAGGACATCTACTGGGATACAGTACTACGACACCTTCGGTGTGGAGCCGGCTACGAGTTCGGATACGGCCTTTCTTACACGAC 540
 151 E E D I Y V R Y R Y D T F G V E P A Y E F G Y G L S Y T T 180

541 CTTTGAGTACGTTGACCTGAACGTTTCGTTTCGACGTTGAAACCTCAGTACAGAATAGAAAACCGGGCGTTCGTGCAGGAAA 630
 181 F E Y S D L N V S F D G E T L R V Q Y R I E N T G G R A G K 210

631 GGAAGTCTCGCAGGTTTACATCAAGGCACCGAAAGGAAAAATCGACAACCTTCCAGGAACTCAAGCGGTTTACAAAACCGACTTTT 720
 211 E V S Q V Y I K A P K G K I D K P F Q E L K A F H K T R L L 240

721 GAATCTGGAGAGTCTGAAGAAGTGGTCTTGAGATACTGTGAGAGATCTTCAAGTTCACCGTGAAGAATGGGTTGTCGAAGCGGG 810
 241 N P G E S E E V V L E I P V R D L A S F N G E E W V V E A G 270

811 TGAATACGAAGTAAAGGTTGGTGGCTTCGAGGAACATAAACTTAAAGAACGTTTCCGTCGGTGAAGAGAAAGATTCAAACCGTG 900
 271 E Y E V R V G A S R N I K L K G T F S V G E R I E N T G R F K P * 300

901 AAGCAGGAGGGGTATCATGAAAAAAGTGGTCTTGTCTTTTGTATTTCTGTTTTCATCCTGGCTCAGAATATTTTACACAATGGCT 990
 S.D.
 M K K L V L V L L L F P V F I L A Q N I L H N G S 25

991 CCTTCGACGCTCCCATTTCTCATCGCTGGTGTGATATCGAACCACTGCGACGAGCGGTTCTATAACACGCAGAACAACCTGGGTGTTTT 1080
 26 F D A P I L I A G V D I E P P A A D G S I N T Q N N W V F F 55

1081 TCACCAATTCGAACGGTGAAGGCAAGGTTGAAATGAAATAACAACCGGAGGAGACCACCTGGTCTG 1170
 56 T N S E A R V N E I T N G G D H C T W S V 85

1171 TTCAGATAATCCAGTCCCCCATACTGTTGAGAACTCCACAAGTACAGAGTTTTTTTCAAGGCAAGGCCTCCGTTTTCAGAGAAACATAG 1260
 86 Q I I Q S P I R V E K L H K Y R V F F K A K A S V Q R N I G 115

1261 GAGTAAAAATAGGAGAACGGCAGGAAGGGGCTGGCAGCTTACAATCCCGAACGGATGAGTCGGGAGGTATGGTTTTTGAACCTGGAA 1350
 116 V K I G G T A G R G W A A Y N P G T D E S G G M V F E L G T 145

1351 CAGACTGGAAGACGTACGAGTTCGAATTTGTCTGAGACGAGAAACGGATGAAAACGCACGCTTTGAGTTCAGCTGGAAAGTCAACGG 1440
 146 D W K T E F E F V M R Q E T D E N A R F E F Q L G K S T G 175

1441 GGACCGTCTGGATAGACACGGTCTGGATAGACGACGTAGTGTGAAGATGTGGGAACGCTGGAAGTCAAGTGGTGGAGAAAACGAGATCT 1530
 176 T V W I D T V W I D D V V M E D V G T L E V S G E E N E I Y 205

1531 ACACCGAGGAGGATGAGGACAGGTGGAAGACTGGCAACTTGTCTGGAGTCAAGGATTCGATGATGGAGTTATCGATCCAAACGTATGGA 1620
 206 T E E D E D K V E [D W Q L V W S Q E F D D G V I D P N V W N] 235

1621 ACTTTGAGATGAGAAACGGCCATGCAAGGGCATTCTGGCTGGGAAAACGCAGAAGTTCGAGTATTACACCGACAAAACGCATTTCGTTG 1710
 236 F E I G H A K G I P G V L N A E L E Y Y T D K N A F V E 265

1711 AAAACGGCTGCCTCGTATAGAAAGCAAGGAAAGAACAGGTATCCGATGAGTACGGAACCTACGACTACACATCCGCAAGGATAACCAGT 1800
 266 [N G C L V I E A R K E Q V S D E Y G T Y D Y T S A R I T T E] 295

1801 AAGTAAATTCGAATAAAGTACGGAAGATCGAGATAAGGCGAAGCTTCCCAAAGGAAAGGCATCTGGCCGGCTCTCTGGATGCTTG 1890
 296 [G K F E I K Y G K I E I R A K L P K G K G I W P A L W M L G] 325

1891 GAAACAACATCGGTGAAGTTGGATGGCCACCTGTGGTGGATGACATCGGAGATGCTGGGGCAGCAGACAGAAGCCGTTTTCAGAA 1980
 326 [N N I G E W V F D H P F L I M E M L G H D T R T V L R T] 355

1981 CAGCGCACGGTCCAGGTTATCTGGCGTGAAGTATAGGTGGCTTATCATCTCCGAAAGAGTTTCTGACTTCTCAGAAAGACTTCC 2070
 356 [A H G P G Y S G G A S I G V A Y H L P E E V P D F S E D F H] 385

2071 ACGTCTTCTCCATCGAGTGGGACGAAAACGAAGTGGAAATGGTATGTAGATGGCCAGCTGTATCACGTTCTCAGCAAGGACGAACTTGCCG 2160
 386 [V F S I E W D E N E V E W Y V D G Q L Y H V L S K D E L A E] 415

2161 AACTGGACTCGAATGGGTTCTTCGATCACCCGTTCTTCTCATACTGAACGTTGCGATGGGAGGTTACTGGCCGGCTATCCTGACGAAA 2250
 416 [L G L E W V F D H P F L I L N V A M G G Y W P G Y P D E T] 445

2251 CAACGCAGTTTCCACAGAGAAATGTACATCGAATTACATCAGAGTTTACAAGATATGAATCCAGAACTATCACAGGAGAAGTGGACGATT 2340
 446 [T Q F P Q R M Y I D Y I R V Y K D M N P E T I T G E V D D C] 475

2341 GTGAGTACGAACAGTCTCAGCAGCAAACAGGGCCGAAAGTGCCTATGAACAGATAAACAACGGCACATTCGATGAACCGATAGTGAACG 2430
 476 [E Y E Q S Q Q Q T G P E V T Y E Q] I N N G T F D E P I V N D 505

2431 ATCAGGCCAACACCCGGATGAGTGGTTCATATGGCAGGCCGGTACTATGGAATCAGCGGTGCCAGAGTTTCTGACTACGGTGTCCACAG 2520
 506 Q A N N P D E W F I W Q A G D Y G I S G A R V S D V T D 535

2521 ATGGCTACGCTACATCACGATAGAGGACTCCGGAACAGATACCTGGCACATACAGTTCAACAGTGGATAGGCTCTACAAGGAAAAGA 2610
 536 G Y A Y I T I E D S G T D T W H I Q F N Q W I G L Y K G K T 565

2611 CCTACACCATTTCTTTCAGGGCTAAAGCAGATACACCAAGACCGATAAACGTTGAAGATCCTTCAGAATCACGATCCATGGATCAACTACT 2700
 566 Y T I S F R A K A D T P R P I N V K I L Q N H D P W I N Y F 595

2701 TTGCGCAGACAGTCAATCTCACAACAGAAATGGCAGACGTTTACGTTCACTACACACATCCAGACGATGGGACGAGGTTGTTTCAGATCA 2790
 596 A Q T V N L T T E W Q T F T F T Y T H P D D A D E V V Q I S 625

2791 GTTTCGAACTTGAAAAGAGGCTCAACCCAGATTTACTTCGACGATGTTTCCGGTGGCCCTCAATGATTTCTCCCGGGGTAATGCGCG 2880
 626 F E L G K E A P T T I Y F D D V S V S P Q * 647

2881 GGGTTTTCAAACCTGAACCGAAGAGGGGAGATCGAGAGTGAACATAAGAGCAAGGATGTCTTCATAATGATCGTAGTTCTTTTGTCTTTT 2970

Fig. 1. For legend see facing page.

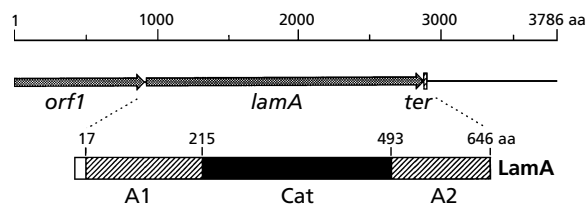


Fig. 2. Structure of the insert DNA and of the LamA protein. *ter* indicates the palindromic sequence; open box, signal peptide; hatched boxes, substrate binding domains A1 and A2; black box, catalytic domain (Cat).

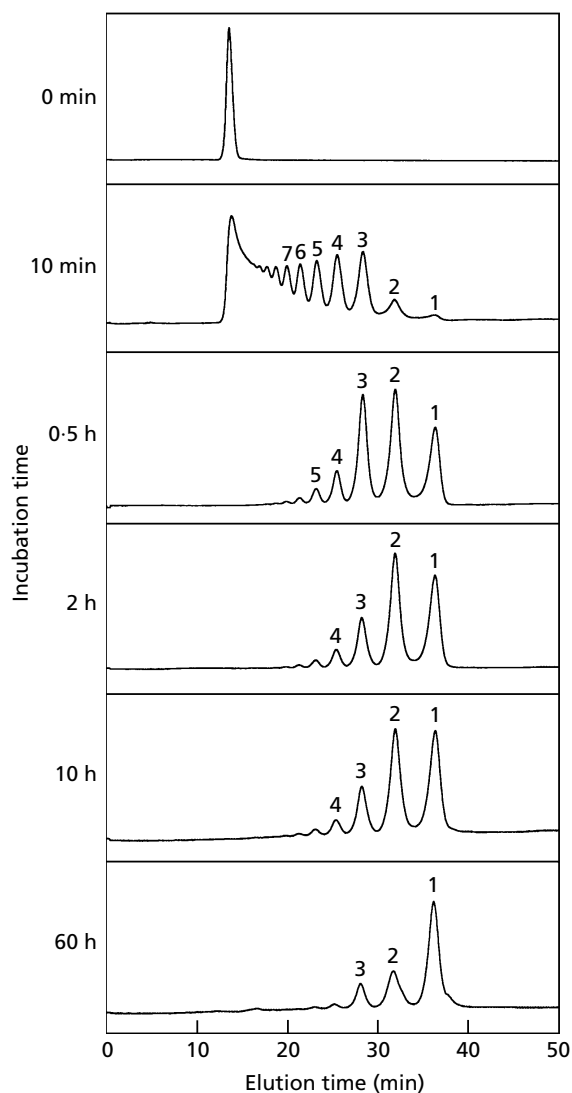


Fig. 3. HPLC analysis of the hydrolysis products after digestion of laminarin with LamA (73 kDa) protein: 1, 2, 3, etc. designate glucose, laminaribiose, -triose, etc.

Fig. 1. Sequence of clone pTT26 (bases 1–2970). Reading frames *orf1* and *lamA* (1,3- β -glucanase) are translated. The postulated Shine–Dalgarno (S.D.) sequence of *lamA* is in bold underlined type, a possible leader peptide of *lamA* is indicated by bold italic type and the catalytic domain (Cat) is boxed. The palindromic sequence downstream of *lamA* is shown by arrows.

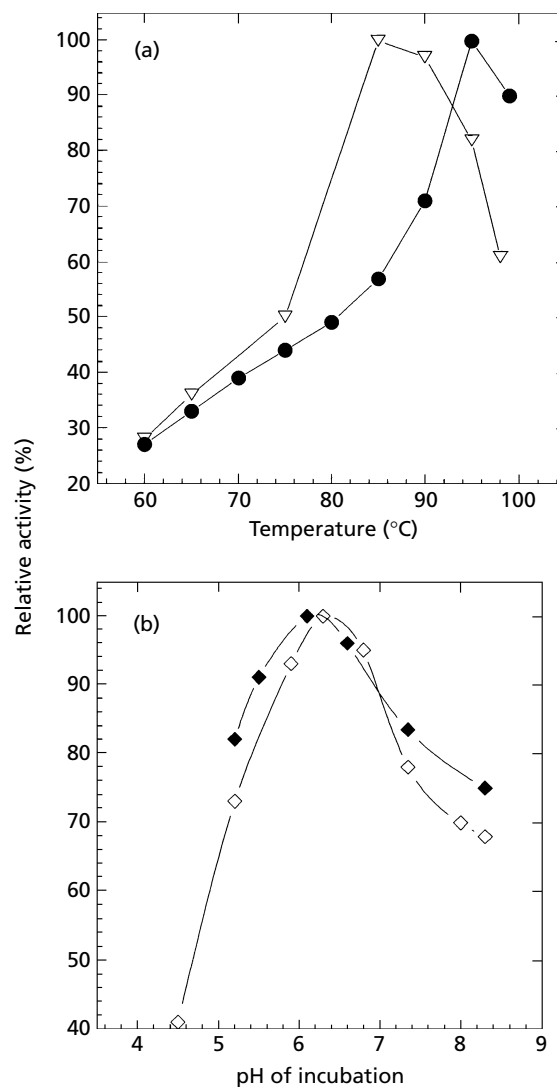


Fig. 4. Biochemical characterization of purified 73 and 52 kDa LamA. (a) Temperature dependence. The 73 kDa protein (●) was incubated for 20 min and the 52 kDa protein (▽) for 15 min with substrate (0.7% laminarin) in 50 mM sodium phosphate/citrate buffer, pH 6.2. (b) pH dependence. The pH of sodium phosphate/citrate buffer (50 mM; pH 5.2–8.3) and sodium acetate buffer (50 mM; pH 4.5) was determined at 20 °C. The 73 kDa protein (◇) was incubated for 12 min at 95 °C, the 52 kDa protein (◆) for 15 min at 88 °C.

The 73 kDa laminarinase was not only stable against thermoinactivation, but also relatively insensitive to inactivation by detergents and solvents: at 1% (w/v) SDS and 10% (v/v) ethanol it was 30 and 40% active, respectively (20 min incubation).

Kinetic parameters were determined for the overall hydrolytic action on laminarin by measuring the initial velocities at each substrate concentration from the slope

of the linear portion of the reaction curve. A V_{max} of 2600 and 3100 U mg^{-1} and a K_m of 2.2 and 2.8 $mg\ ml^{-1}$ for the 52 and 73 kDa proteins, respectively, were determined with sodium phosphate/citrate buffer (70 mM), pH 6.2, at 85 and 95 °C, respectively.

DISCUSSION

The molecular mass of 73328 Da, calculated from the sequence, is in accordance with the estimated molecular mass of the purified 73 kDa recombinant LamA protein from SDS-PAGE. For the 52 kDa β -glucanase, also purified from the recombinant cells, identical enzymic activities and purification profiles, and similar biochemical characteristics were observed, strongly suggesting that the 52 kDa protein resembles a processed form of the 73 kDa gene product. Possible processing sites near the interdomain boundaries outside the catalytic domain (purportedly at the C-terminal side) would yield an enzymically active fragment mass of about 55 kDa and could explain the occurrence of the 52 kDa enzyme. Such a cleavage at interdomain boundaries is common in complex thermostable exoproteins and has been described, for example, for the endo-glucanase CelZ of *Cl. thermocellum* (Jauris *et al.*, 1990).

The high degree of sequence identity with the active site region of *Bacillus* type mixed-linkage glucanase protein sequences (Fig. 5) suggests that the LamA enzyme could also have a retaining stereochemical course of hydrolytic action (Malet *et al.*, 1993). This mechanism requires two functional groups present in the correct spatial setting. According to the molecular model of the hybrid *Bacillus* 1,3-1,4- β -glucanase, the general acid is Glu107 (Glu343 in LamA) and the catalytic nucleophile is Glu103 (Glu338) (Hahn *et al.*, 1995). These residues are, in addition to Asp105 (Asp340) and Gly110 (Gly346), highly conserved throughout the sequences of bacterial lichenases and laminarinases (Fig. 5). Other highly

invariant residues include Ile104 (Ile339) and Ile106 (Ile341).

The lichenases of *B. amyloliquefaciens* or *B. macerans* display an extraordinary substrate specificity on lichenan and hydrolyse only 1,4- β -bonds which are adjacent to 1,3- β -bonds (Anderson & Stone, 1975). The laminarinase of *B. circulans* is exclusively active on laminarin, hydrolysing only 1,3- β -bonds. Other β -glucanases such as LicB of *Cl. thermocellum* (Schimming *et al.*, 1992; Zverlov *et al.*, 1994), which was previously also described as a laminarinase (Zverlov & Velikodvorskaya, 1990), or the laminarinase of *R. marinus* (Spilliaert *et al.*, 1994) are similar to LamA, but much less specific and hydrolyse both substrates to some degree (endo-1,3(4)- β -glucanase). However, in contrast to these enzymes LamA has a much clearer preference for laminarin. This indicates that the enzymes with intermediate substrate specificity are able to hydrolyse both 1,3- β - and 1,4- β -linkages with varying preference for either the 1,3- β - or the 1,4- β -linkage.

All 1,3- β -glucanases contain insertions in homologous positions within the highly conserved region compared to 1,3-1,4- β -glucanases: an inserted Thr-Cys-Gly337 sequence and a Met342 residue in LamA (Fig. 5). These insertions purportedly result in a structural rearrangement of the active site and could influence the shift in substrate specificity. The areas outside β -strand K, containing the active site amino acid residues, are much less conserved, e.g. the region around the substrate-binding residues Asn121, Tyr123 and Glu131 (β -strands L and M and loop regions) in the hybrid *Bacillus* β -glucanase. This possibly reflects the difference in substrate-binding activity towards 1,3- β - or 1,3-1,4- β -glucans and in extension of the hydrolytic activity towards 1,3- β -linkages. Surprisingly, β -sheets N and O, lying within the outside shell of the 'jelly roll' (Hahn *et al.*, 1995), are highly conserved, pointing to a possible crucial role for the structural integrity of the β -glucanase

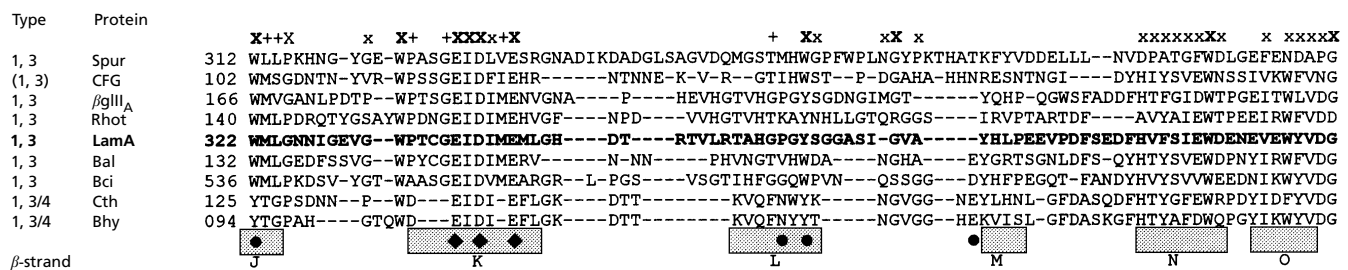


Fig. 5. Sequence alignment around the catalytic amino acid residues. Spur, *Strongylocentrotus purpuratus* 1,3- β -glucanase (g1488257); CFG, coagulation factor G subunit a of Horseshoe Crab (D16622); β gIII_A, *O. xanthineolytica* 1,3- β -glucanase isoenzyme IIA (U56935); Rhot, *R. marinus* 1,3- β -glucanase (P45798); LamA, *T. neapolitana* 1,3- β -glucanase (Z47974); Bal, *B. alkalophilus* AG-430 1,3- β -glucanase (Nogi & Horikoshi, 1990); Bci, *B. circulans* 1,3- β -glucanase A1 (P23903); Cth, *Cl. thermocellum* 1,3-1,4- β -glucanase LicB (P37074); Bhy, *B. amyloliquefaciens*/*B. macerans* hybrid 1,3-1,4- β -glucanase H(A16-M) (Hahn *et al.*, 1995). ◆, Catalytic site amino acid residues; ●, substrate binding site residues in *B. amyloliquefaciens*/*B. macerans* hybrid β -glucanase H(A16-M). Shaded boxes indicate β -strands of β -sheets II (J, K, L, M) and I (N, O) of the hybrid *Bacillus* β -glucanase. Conserved amino acid residues in all (X), in most (x) or in 1,3- β -glucanase sequences (+) are indicated.

molecule (Fig. 5). It would be highly interesting to resolve the three-dimensional structure of this type of enzyme and to clarify the difocused hydrolytic activity continuously shifting within this enzyme family from 1,4- β - to 1,3- β -linkages.

A surprising sequence homology of LamA was found with eukaryotic protein sequences, the horseshoe crab 1,3- β -D-glucan-sensitive coagulation factor G α subunit precursor (Seki *et al.*, 1994) and the 1,3- β -glucanase from sea urchin (*Strongylocentrotus purpuratus*) eggs (Bachman & McClay, 1996) (Fig. 5). So far, no glucanase activity of coagulation factor G has been described. Since sequences are also similar outside the active site in parts of the protein backbone, which preserve the three-dimensional structure of the enzyme, and since the temperature range of the sea urchin enzyme is much closer to that of the bacterial enzymes than to the normal environmental conditions in which sea urchins live, it has to be postulated that 1,3- β -glucanase genes have been horizontally transferred from bacteria to eukaryotes.

No function in substrate hydrolysis could be attributed to the non-catalytic domains A1 and A2. Despite its homology to cellulose binding domains, the complete 73 kDa enzyme did not bind to crystalline cellulose (Avicel; data not shown); this does not exclude their function in binding to other insoluble substrates occurring, for example, in plant cell walls or crab shells. However, the 52 kDa enzyme, obviously missing either A1 or A2, loses thermostability and shows a slightly lower pH optimum. This points to a possible role of these flanking domains at least in stabilizing the overall structure of the enzyme core as was shown for the xylanase of *T. neapolitana* (Zverlov *et al.*, 1996).

Family 16 of glycosyl hydrolases contains a number of thermotolerant β -glucanases. The hitherto most thermostable β -glucanases are *Cl. thermocellum* LicB (Zverlov *et al.*, 1994) and *R. marinus* BglA (Spilliaert *et al.*, 1994) with temperature optima of 80 and 85 °C, respectively. The complete enzyme LamA from *T. neapolitana* is by far superior to all other 1,3- β -glucanases in temperature stability (95 °C) and specific activity (3100 U mg⁻¹).

ACKNOWLEDGEMENTS

This work was supported by a grant from INTAS (94-1328) and by travel grants from the Deutsche Forschungsgemeinschaft DFG to V.V.Z., G.A.V. and W.H.S. A grant of the Volkswagenstiftung to K. Bronnenmeier has enabled a part of the stay of V.V.Z. in München, which was very helpful for completing the project. We are very grateful to W. L. Staudenbauer for critically reading the manuscript.

REFERENCES

- Anderson, M. A. & Stone, B. A. (1975). A new substrate for investigating the specificity of β -glucan hydrolases. *FEBS Lett* **52**, 202–207.
- Aono, R., Hamamura, M., Yamamoto, M. & Asano, T. (1995). Isolation of extracellular 28- and 42-kilodalton 1,3- β -glucanases and comparison of three 1,3- β -glucanases produced by *Bacillus circulans* IAM1165. *Appl Environ Microbiol* **61**, 122–129.
- Bachman, E. S. & McClay, D. R. (1996). Molecular cloning of the first metazoan 1,3- β glucanase from eggs of the sea urchin *Strongylocentrotus purpuratus*. *Proc Natl Acad Sci USA* **93**, 6808–6813.
- Bronnenmeier, K., Kern, A., Liebl, W. & Staudenbauer, W. L. (1995). Purification of *Thermotoga maritima* enzymes for the degradation of cellulosic materials. *Appl Environ Microbiol* **61**, 1399–1407.
- Coutinho, J. B., Gilkes, N. R., Warren, R. A. J., Kilburn, D. G. & Miller, R. C., Jr (1992). The binding of *Cellulomonas fimi* endoglucanase C (CenC) to cellulose and sephadex is mediated by the N-terminal repeats. *Mol Microbiol* **6**, 1243–1252.
- Dakhova, O. N., Kurepina, N. E., Zverlov, V. V., Svetlichnyi, V. A. & Velikodvorskaya, G. A. (1993). Cloning and expression in *Escherichia coli* of *Thermotoga neapolitana* genes coding for enzymes of carbohydrate substrate degradation. *Biochem Biophys Res Commun* **194**, 1359–1364.
- Ferrer, P., Halkier, T., Hedegaard, L., Savva, D., Diers, I. & Asenjo, J. A. (1996). Nucleotide sequence of a 1,3- β -glucanase isoenzyme II_A gene of *Oerskovia xanthineolytica* LL G109 (*Cellulomonas cellulans*) and initial characterization of the recombinant enzyme expressed in *Bacillus subtilis*. *J Bacteriol* **178**, 4751–4757.
- Gräbnitz, F., Rücknagel, K. P., SeiB, M. & Staudenbauer, W. L. (1989). Nucleotide sequence of the *Clostridium thermocellum* *bglB* gene encoding thermostable β -glucosidase B: homology to fungal β -glucosidases. *Mol Gen Genet* **217**, 70–76.
- Hahn, M., Keitel, T. & Heinemann, U. (1995). Crystal and molecular structure at 0.16-nm resolution of the hybrid *Bacillus* endo-1,3-1,4- β -D-glucan 4-glucanohydrolase H(A16-M). *Eur J Biochem* **232**, 849–858.
- van Heijne, G. (1986). A new method for predicting signal sequence cleavage sites. *Nucleic Acids Res* **14**, 4683–4690.
- Henrissat, B. & Bairoch, A. (1993). New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J* **293**, 781–788.
- Huber, R. & Stetter, K. O. (1992). The order *Thermotogales*. In *The Prokaryotes*, 2nd edn, pp. 3809–3815. Edited by A. Balows, H. G. Trüper, M. Dworkin, W. Harder & K.-H. Schleifer. New York: Springer-Verlag.
- Jauris, S., Rücknagel, K. P., Schwarz, W. H., Kratzsch, P., Bronnenmeier, K. & Staudenbauer, W. L. (1990). Sequence analysis of the *Clostridium stercorarium* *celZ* gene encoding a thermoactive cellulase (Avicelase I): identification of catalytic and cellulose-binding domains. *Mol Gen Genet* **223**, 258–267.
- Kanzawa, Y., Kurasawa, T., Kanegae, Y., Harada, A. & Harada, T. (1994). Purification and properties of a new exo-(1 \rightarrow 3)- β -D-glucanase from *Bacillus circulans* YK9 capable of hydrolysing resistant curdlan with formation of only laminari-biose. *Microbiology* **140**, 637–642.
- Lee, D. S. & Chang, H. G. (1995). Cloning and expression of a 1,3- β -glucanase gene from *Bacillus circulans* KCTC3004 in *Escherichia coli*. *Biotech Lett* **17**, 355–360.
- Malet, C., Jiménez-Barbero, J., Bernabé, M., Brosa, C. & Planas, A. (1993). Stereochemical course and structure of the products of the enzymatic action of endo-1,3-1,4- β -D-glucan 4-glucanohydrolase from *Bacillus licheniformis*. *Biochem J* **296**, 753–758.
- Nogi, Y. & Horikoshi, K. (1990). A thermostable alkaline 1,3- β -glucanase produced by alkalophilic *Bacillus* sp. AG-430. *Appl Microbiol Biotechnol* **32**, 704–707.
- Schimming, S., Schwarz, W. H. & Staudenbauer, W. L. (1992). Structure of the *Clostridium thermocellum* gene *licB* and the encoded 1,3-1,4- β -glucanase. A catalytic region homologous to

Bacillus lichenases joined to the reiterated domain of clostridial cellulases. *Eur J Biochem* **204**, 13–19.

Schwarz, W. H., Schimming, S. & Staudenbauer, W. L. (1988). Isolation of a *Clostridium thermocellum* gene encoding a thermostable 1,3- β -glucanase (laminarinase). *Biotechnol Lett* **10**, 225–230.

Seki, N., Muta, T., Oda, T., Iwaki, D., Kuma, K., Miyata, T. & Iwanaga, S. (1994). Horseshoe crab (1,3)- β -D-glucan-sensitive coagulation factor G. A serine protease zymogen heterodimer with similarities to β -glucan-binding proteins. *J Biol Chem* **269**, 1370–1374.

Spilliaert, R., Hreggvidsson, G. O., Kristjansson, J. K., Eggertsson, G. & Palsdottir, A. (1994). Cloning and sequencing of a *Rhodothermus marinus* gene, *bglA*, coding for a thermostable β -glucanase and its expression in *Escherichia coli*. *Eur J Biochem* **224**, 923–930.

Tomme, P., Warren, R. A. J. & Gilkes, N. R. (1995). Cellulose hydrolysis by bacteria and fungi. *Adv Microb Physiol* **37**, 1–81.

Wood, T. M. & Bhat, K. M. (1988). Methods for measuring cellulase activities. *Methods Enzymol* **160**, 87–112.

Zverlov, V. V. & Velikodvorskaya, G. A. (1990). Cloning the *Clostridium thermocellum* thermostable laminarinase in *Escherichia coli*: the properties of the enzyme produced. *Biotechnol Lett* **12**, 811–816.

Zverlov, V. V., Fuchs, K. P., Schwarz, W. H. & Velikodvorskaya, G. A. (1994). Purification and cellulosomal localization of *Clostridium thermocellum* mixed linkage β -glucanase LicB (1,3-1,4- β -D-glucanase). *Biotechnol Lett* **16**, 29–34.

Zverlov, V., Piotukh, K., Dakhova, O., Velikodvorskaya, G. & Borris, R. (1996). The multidomain xylanase A of the hyperthermophilic bacterium *Thermotoga neapolitana* is extremely thermoresistant. *Appl Microbiol Biotechnol* **45**, 245–247.

Received 28 October 1996; revised 16 December 1996; accepted 19 December 1996.