APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Dec. 1992, p. 3864–3867 Vol. 58, No. 12
0099-2240/92/123864-04$02.00/0

The β-Mannanase from "*Caldocellum saccharolyticum*" Is Part of a Multidomain Enzyme

M. D. GIBBS, D. J. SAUL, E. LÜTHI, AND P. L. BERGQUIST*

Centre for Gene Technology, Department of Cellular & Molecular Biology, University of Auckland, Auckland, New Zealand

Received 12 June 1992/Accepted 29 September 1992

The complete sequence of a β-mannanase gene from an anaerobic extreme thermophile was determined, and it shows that the expressed protein consists of two catalytic domains and two binding domains separated by spacer regions rich in proline and threonine residues. The amino-terminal catalytic domain has β-mannanase activity, and the carboxy-terminal domain acts as an endoglucanase. Neither domain shows homology with any other cellulase or hemicellulase sequence at the nucleic acid or protein level.

Hemicellulases are abundant polysaccharides in nature and are thought to have linkages to lignin in wood (8). The major constituents of hemicellulose are the hetero-1,4-β-D-xylans and hetero-1,4-β-D-mannans. The heteroxylans are found mainly in grasses, cereals, and hardwoods, whereas the mannans are more abundant in softwoods (gymnosperms, O-acetylgalactomannans) but are also found in hardwoods (glucomannans). β-Mannan consists of a backbone of mannose residues, but it may contain glucose residues. α-D-galactose and acetyl residues are attached to the backbone in galactoglucomannans (5). Enzymatic hydrolysis of β-mannan is accomplished by β-D-mannanase (1,4-β-D-mannan mannohydrolase, EC 3.2.1.78) and β-mannosidase (β-D-mannoside mannohydrolase, EC 3.2.1.25). β-Mannanases have been isolated from a number of bacterial, fungal, and plant sources (2–4, 9, 12).

A β-mannanase gene from an alkalophilic *Bacillus* sp. has been sequenced, cloned, and expressed in *Escherichia coli* (1), and we reported the cloning, sequencing, and expression of a β-mannanase gene from the obligately anaerobic, extremely thermophilic bacterium "*Caldocellum saccharolyticum*" (10). The mannanase gene was found to be located between the genes for two cellulases on the genome of this organism. Extensive sequence analysis showed that these cellulases are multidomain enzymes (14, 18). Re-examination of DNA sequences in the vicinity of the β-mannanase suggested homology with repetitive domains of the cellulases. In this report, we describe sequence analysis of the β-mannanase and identification of repetitive domains and show that contrary to our earlier report, the mannanase is a catalytic domain of a multidomain protein that also expresses endoglucanase activity.

MATERIALS AND METHODS

Plasmids. The construction of plasmids pNZ1019 and pNZ1609 has been described (10). A detailed restriction map of pNZ1609 was generated, and a number of deletion derivatives were constructed by using restriction enzyme sites in the pBluescript vector (16). The 608-bp *Eco*RV fragment was inserted in each orientation into vector mp10 and sequenced. All plasmids were transformed into strain PB1427 (F^- *thr-1 leuB6 lacY1 supE44*).

Sequence analysis. Sequence analysis was carried out on double- or single-stranded DNA, as appropriate, by using a cycle sequencing protocol provided by Applied Biosystems Ltd. for use with the ABI 373 automated sequencer. All DNA was sequenced on both strands (see Fig. 1).

PCRs. Four primers were designed to allow amplification of domains 1 and 4 by using the polymerase chain reaction (PCR). The primers were designed to incorporate restriction enzyme sites at the 5′ and 3′ ends of the amplified fragment to allow directional cloning of the PCR product in expression vector pJLA602 (15; see Fig. 2B). Primers Man8 and Man11 allowed amplification of the β-mannanase domain, primers Man10 and Man9 allowed amplification of domain 4, and Man8 and Man9 allowed amplification of the entire fragment. The individual domains were amplified in the PCR (2.5 mM $MgCl_2$), cut with restriction enzymes, isolated from a 1% agarose gel using Geneclean (Bio 101), and ligated into pJLA602 cut with *Bam*HI and *Sph*I.

Assays for mannanase and cellulase activities. Transformants that carried the correct plasmids were patched onto LB agar plates containing 50 μg of ampicillin per ml, grown overnight at 30°C, and induced at 42°C for 3 to 6 h. Replica plates were overlaid with the appropriate substrate in soft agar and incubated at 70°C for 6 h. Enzymatic activity was determined qualitatively by using the Congo red procedure of Teather and Wood (17) as described previously (14).

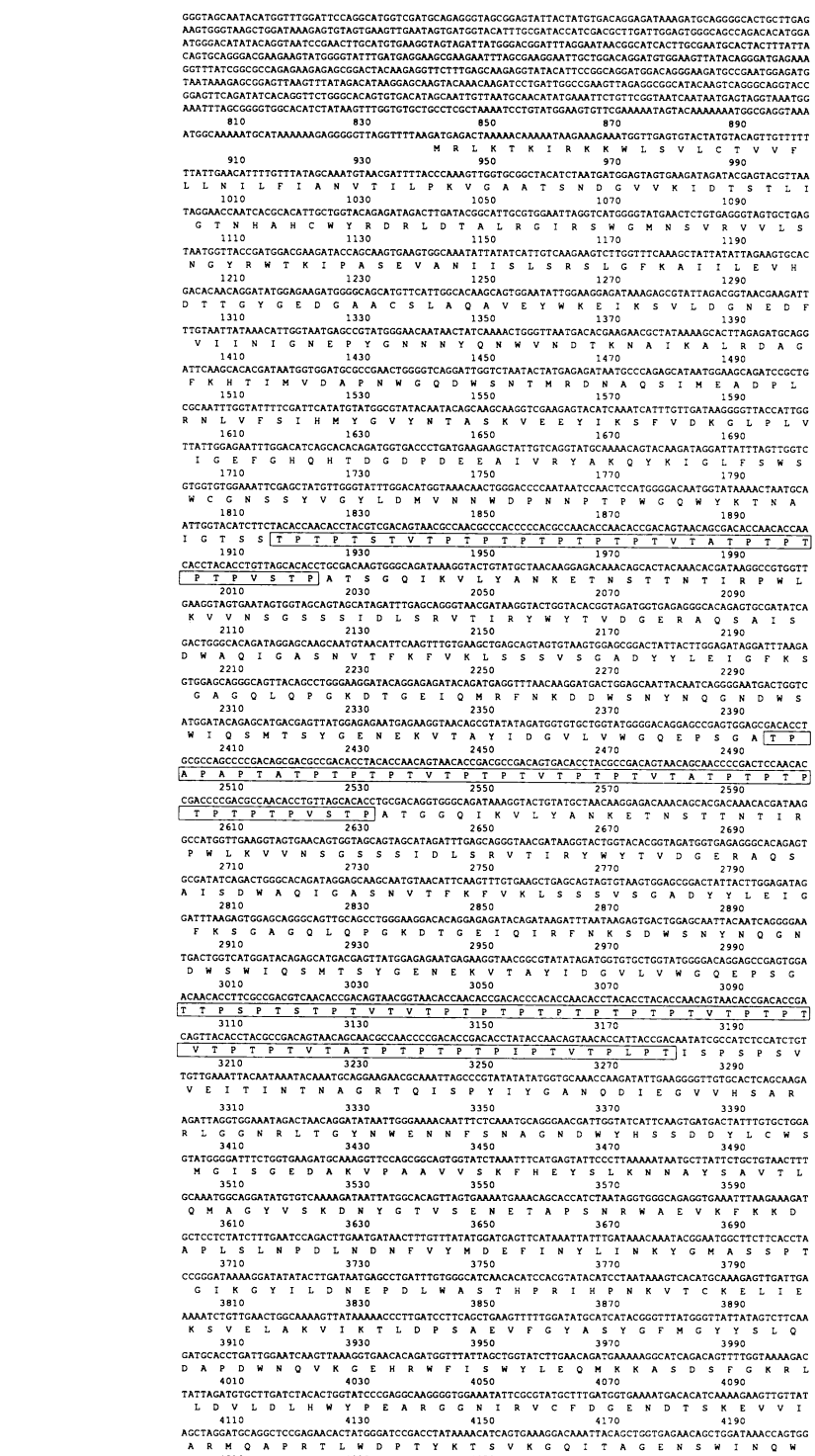
RESULTS AND DISCUSSION

Sequence of the *manA* gene. Examination of the sequence of the *manA* gene (Fig. 1) shows that there is a single open reading frame from positions 841 to 4834 (total, 3,996 bp) which could code for a putative ManA protein of 1,332 amino acids. This statement contrasts with the 1,041-bp (347-amino-acid) open reading frame that we reported previously which has β-mannanase activity (10). This protein is composed of four domains separated by long runs of threonine-proline repeats (PT boxes). This structure is seen in the *celA*-encoded and *celB* cellulases from "*C. saccharolyticum*."

The putative structure of the mannanase gene is shown diagrammatically in Fig. 2, which shows that domains 1 and 4 are catalytic domains and domains 2 and 3 are putative common substrate-binding domains.

Enzyme activity of strains carrying cloned PCR-amplified fragments. Individual transformant colonies of *E. coli* PB1427 carrying plasmids pNZ2005 to pNZ2010 inclusive

* Corresponding author.



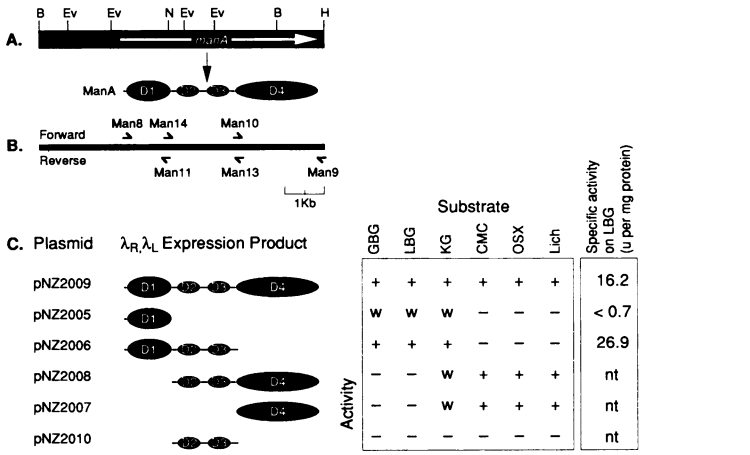


FIG. 2. Subcloning of β-mannanase binding and catalytic domains by PCR. (A) Restriction enzyme sites and structure of the full-length β-mannanase domains. Regions: D1, domain 1 of β-mannanase; D4, domain 4 (carboxymethylcellulase); D2 and D3, putative common substrate-binding domains. DNA restriction site abbreviations: B, *Bam*HI; Ev, *Eco*RV; H, *Hin*dIII; N, *Nco*I. (B) Locations of forward and reverse PCR primers in the amplification of catalytic and putative binding domains of β-mannanase and endoglucanase. Nucleotide sequences of PCR primers: Man8, 5′-AGAAATGCATGCGTGTACTATGTACAG-3′; Man9, 5′-TTTTC CAGATCTTAAGCTTCACATTTG-3′; Man10, 5′-CAGGAAGCAT GCAAATTAGCCCGTATA-3′; Man11, 5′-TGCATTAGATCTTTA CCATTGTCCCCA-3′; Man13, 5′-TTGGTTGGATCCATATATAT ACGGGCT-3′; Man14, 5′-TAATGGCATGCGTACATCTTCTACA CC-3′. (C) Diagrammatic representation of cloned portions of *manA* and their enzymatic activity as measured with plate tests and by release of reducing sugar. Units are expressed as micromoles of reducing sugar released from locust bean gum per min. +, strong clearing in Congo red assay (14); w, weak activity, −, no activity; nt, not tested. Substrate abbreviations: GBG, guar bean gum; LBG, locust bean gum; KG, konjac glucomannan; CMC, carboxymethyl cellulose; OSX, oat spelt xylan; Lich, lichenan.

were patched to L agar-ampicillin plates, grown overnight at 37°C, and replicated to fresh plates which were overlaid with soluble hemicellulose or cellulose substrates.

Figure 2C shows that domain 1 hydrolyzed only mannan, whereas domain 4 hydrolyzed carboxymethyl cellulose, oat spelt xylan, and lichenan and had weak activity on konjac gum. The complete gene had much greater activity than the individual domains, as judged by the extent of the zone of clearing on mannans, suggesting that the binding domains facilitate hydrolysis. We attribute this increased activity to the presence of the binding domain in pNZ2006, as all recombinant plasmids had identical sequence structures between the promoter and the initiation codon, which included the same distance of the expressed proteins from the ribosomal binding site to the ATG. Polyacrylamide gel analysis showed that approximately equal amounts of the proteins were synthesized in *E. coli*. Thus, we conclude that the reduced specific activity of pNZ2005 compared with that of pNZ2006 reflects a reduction of the catalytic activity of the enzyme lacking the binding domain, although we did not

FIG. 1. Sequence of the *Sma*I-*Hin*dIII fragment from the *Caldocellum* genome containing the β-mannanase gene. GenBank accession number, L01257. The proline-threonine-rich areas are boxed.

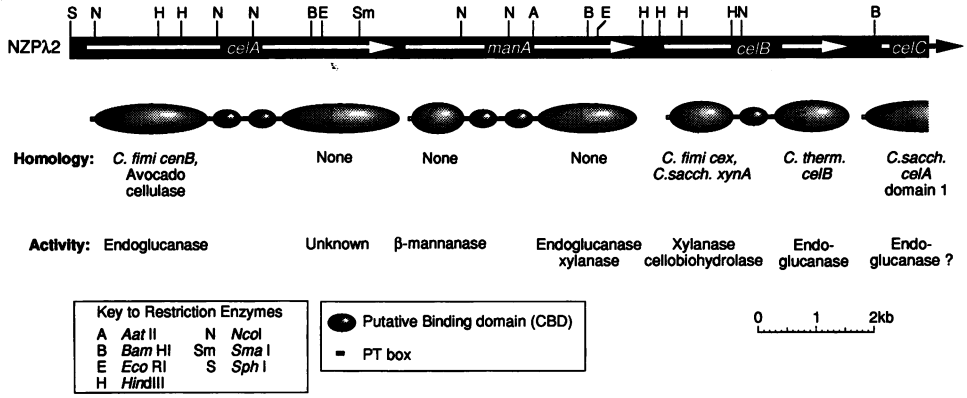


FIG. 3. Diagrammatic representation of multidomain enzymes on a λ recombinant from a genomic library of "*C. saccharolyticum*." All of the DNA indicated was sequenced on both strands.

test the unlikely possibility that these results reflect differences in mRNA stability.

Quantitative assays of the enzymatic activity of partially purified extracts of the PCR-generated recombinants, as measured by release of reducing sugar with locust bean gum as the substrate (10), showed that the β-mannanase had only weak activity without the binding domains which, by themselves or in *trans*, had no activity or effect on the β-mannanase activity of the catalytic domain (Fig. 2C). The specific enzymatic activity of bacteria carrying pNZ2005 was less than previously reported, but this can be accounted for by the different vectors used to construct pNZ2005 and pNZ1611, which carries the truncated *manA* gene without binding domains (10).

Homologies between binding domains and PT boxes. The genes *celA*, *celB*, and *manA* are closely linked on the genome and were isolated on a single λ clone (NZλP2; Fig. 3). These genes code for multidomain-multifunction enzymes (10, 14, 18). Each shares a common sequence which codes for a domain which is involved in binding of the enzyme to insoluble substrates. Both *manA* and *celA* have two copies of this domain, while *celB* has only one. In addition to the binding domains, these enzymes have two catalytic domains; *celB* has an endoglucanase domain and a domain which, by inference from homologous proteins, is a cellobiohydrolase (exoglucanase). The proteins encoded by *manA* and *celA* have, respectively, mannanase and endoglucanase N-terminal domains and two further C-terminal domains, of which one is an endoglucanase and the other has no recognized enzymatic activity. The domains on each enzyme are separated from each other by PT boxes (long runs of proline-threonine repeats) which are known to act as flexible hinges between functionally distinct portions of proteins.

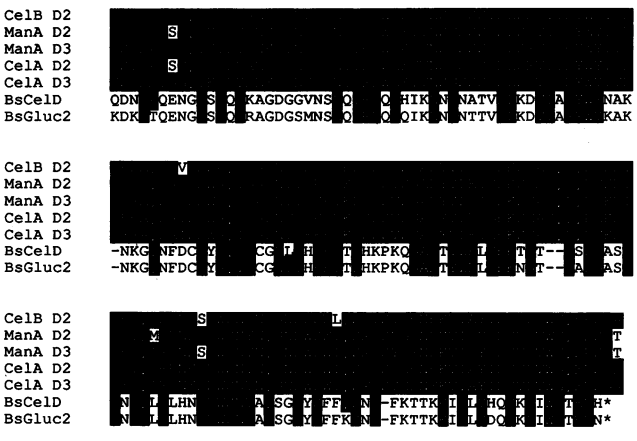


FIG. 4. Alignment of the five common binding domains of CelB, ManA, and CelA with two domains from *B. subtilis* cellulases. Conserved residues are in white letters. The relative positions of the domains CelA D2, ManA D2, ManA D3, CelA D2, and CelA D3 can be seen in Fig. 3. BSCelD, *B. subtilis* CelD (13); BsGluc2, *B. subtilis* βGluc2 (11); *, termination codon.

The five binding domains share a very high level of homology, differing by no more than 3% at either the nucleotide or amino acid level (Fig. 4), and they have moderate homology with two cellulases from *Bacillus subtilis* (11, 13) (Fig. 4).

Din et al. (6) have shown that the binding domain of *Cellulomonas fimi* CenA is involved in nonhydrolytic disruption of cellulose fibers, allowing penetration of the catalytic domain and hydrolysis of the cellulose substrate. We have shown that the activity of ManA is substantially enhanced by the presence of the binding domain in *cis*. The binding domains for these enzymes may be significant in the degradation of solid substrates, but their effect may not be obvious when soluble substrates are used in the laboratory. The overall structure of catalytic domain-PT box-binding domain is common among fungal and bacterial cellulases (7), but the presence of more than one binding domain and catalytic domain is unique to "*C. saccharolyticum*."

ACKNOWLEDGMENTS

This work was funded in part by grants from the University of Auckland Research Grants Committee and Pacific Enzymes Ltd.

REFERENCES

1. **Akino, T., C. Kato, and K. Horikoshi.** 1989. Two *Bacillus* β-mannanases having different COOH termini are produced in *Escherichia coli* carrying pMAH5. Appl. Env. Microbiol. **55:** 3178–3183.
2. **Akino, T., N. Nakamura, and K. Horikoshi.** 1987. Production of β-mannosidase and β-mannanase by an alkalophilic *Bacillus* sp. Appl. Microbiol. Biotechol. **26:**323–327.
3. **Araki, T., and M. Kitamikado.** 1988. Exo-1,4,β-mannanase from *Aeromonas hydrophila*. Methods Enzymol. **160:**583–589.
4. **Araujo, A., and O. P. Ward.** 1990. Hemicellulases of *Bacillus* species: preliminary comparative studies on production and

properties of mannanases and galactanases. J. Appl. Bacteriol. **68**:253–261.
5. **Dekker, R. F. H.** 1985. Biodegradation of hemicelluloses, p. 505–533. *In* T. Higuchi (ed.), Biosynthesis and biodegradation of wood components. Academic Press, Inc., Orlando, Fla.
6. **Din, N., N. R. Gilkes, B. Tekant, R. C. Miller, R. A. J. Warren, and D. G. Kilburn.** 1991. Non-hydrolytic disruption of cellulose fibres by the binding domain of a bacterial cellulase. Bio/Technology **9**:1096–1099.
7. **Gilkes, N. R., B. Henrissat, D. G. Kilburn, R. C. Miller, and R. A. J. Warren.** 1991. Domains in microbial β-1,4-glycanases: sequence conservation, function, and enzyme families. Microbiol. Rev. **55**:303–315.
8. **Kantelinen, A., M. Rättö, J. Sundquist, M. Ranua, L. Viikari, and M. Linko.** 1988. Hemicellulases and their potential role in pulp bleaching, p. 1–4. International Pulp Bleaching Conference, *In* T. J. de Salvo (ed.), Orlando, Fla. Technical Association of the Pulp and Paper Industry, Atlanta.
9. **Kusabe, I., and R. Takahashi.** 1988. β-Mannanase of *Streptomyces*. Methods Enzymol. **160**:611–614.
10. **Lüthi, E., N. B. Jasmat, R. A. Grayling, D. R. Love, and P. L. Bergquist.** 1991. Cloning, sequence analysis, and expression in *Escherichia coli* of a gene coding for a mannanase from the extremely thermophilic bacterium "*Caldocellum saccharolyticum*." Appl. Environ. Microbiol. **57**:694–700.
11. **MacKay, R. M., A. Lo, G. Willide, M. Zuker, S. Baird, M. Dove, F. Moronelli, and V. Seligy.** 1986. Structure of *Bacillus subtilis* endo-β-1,4-glucanase gene. Nucleic Acids Res. **14**:9159–9170.
12. **McCleary, B. V.** 1988. β-D-Mannanases. Methods Enzymol. **160**:596–610.
13. **Nakamura, A., T. Uozumi, and T. Beppu.** 1987. Nucleotide sequence of a cellulase gene of *Bacillus subtilis*. Eur. J. Biochem. **164**:317–320.
14. **Saul, D. J., L. C. Williams, R. A. Grayling, L. W. Chamley, D. R. Love, and P. L. Bergquist.** 1990. *celB*, a gene coding for a bifunctional cellulase from the extreme thermophile "*Caldocellum saccharolyticum*." Appl. Environ. Microbiol. **56**:3117–3124.
15. **Schauder, B., H. Blöcker, R. Frank, and J. E. G. McCarthy.** 1987. Inducible expression vectors incorporating the *Escherichia coli atpE* translational initiation region. Gene **52**:279–283.
16. **Short, J. M., J. M. Fernandez, J. A. Sarge, and W. D. Huse.** 1988. λZAP: a bacteriophage λ expression vector with *in vivo* excision properties. Nucleic Acids Res. **16**:7583–7600.
17. **Teather, R. M., and P. K. Wood.** 1982. Use of Congo red-polysaccharide interactions in enumeration and characterization of cellulolytic bacteria from bovine rumen. Appl. Environ. Microbiol. **43**:777–780.
18. **Te'o, J., D. J. Saul, and P. L. Bergquist.** Unpublished results.