APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Apr. 1990, p. 1017–1024
0099-2240/90/041017-08$02.00/0

Vol. 56, No. 4

Cloning, Sequence Analysis, and Expression of Genes Encoding Xylan-Degrading Enzymes from the Thermophile "*Caldocellum saccharolyticum*"

E. LÜTHI, D. R. LOVE, J. MCANULTY, C. WALLACE, P. A. CAUGHEY, D. SAUL, AND P. L. BERGQUIST*

Centre for Gene Technology, Department of Cellular and Molecular Biology, University of Auckland, Auckland 1, New Zealand

Received 10 August 1989/Accepted 18 December 1989

A λ recombinant bacteriophage coding for xylanase and β-xylosidase activity has been isolated from a genomic library of the extremely thermophilic anaerobe "*Caldocellum saccharolyticum*." Partial *Sau*3AI fragments of the λ recombinant DNA were ligated into pBR322. A recombinant plasmid with an insertion of ca. 7 kilobases of thermophilic DNA expressing both enzymatic activities was isolated. The location of the genes has been established by analyzing deletion derivatives, and the DNA sequence of 6.067 kilobases of the insert has been determined. Five open reading frames (ORFs) were found, one of which (ORF1; M_r 40,455) appears to code for a xylanase (XynA) which also acts on *o*-nitrophenyl-β-D-xylopyranoside. Another, ORF5 (M_r 56,365), codes for a β-xylosidase (XynB). The *xynA* gene product shows significant homology with the xylanases from the alkalophilic *Bacillus* sp. strain C125 and *Clostridium thermocellum*.

Xylan is a major component of hemicellulose and is found in large amounts in straw and as a component of hardwood and softwood. The enzymatic hydrolysis of xylan, which is a heteropolymer of the pentose sugar xylose, is accomplished by the action of endo-1,4-β-xylanase (EC 3.2.1.8) and β-D-xylosidase (EC 3.2.1.37). The first enzyme acts on xylan to generate small xylooligosaccharides, and the β-xylosidase hydrolyzes dimers and trimers of xylose to the monomeric sugar (5).

"*Caldocellum saccharolyticum*" is an extremely thermophilic anaerobic microorganism which is capable of growing at 80°C under laboratory conditions. This organism can use xylan as the sole carbon source (4). The temperature optimum and stability of the cellulolytic and xylanolytic enzymes produced by "*C. saccharolyticum*" exceed those reported for other cellulases and xylanases (4, 20, 35, 47). These properties may become important, as xylanases and other hemicellulases may play a role in the enzymatic bleaching of pulp in the manufacture of paper (19, 33). In this report we describe the molecular cloning of a gene encoding a xylanase which also acts on *o*-nitrophenyl-β-D-xylopyranoside (OPNX) and a gene for a β-xylosidase from "*C. saccharolyticum*."

MATERIALS AND METHODS

Bacteria and culture conditions. *Escherichia coli* Q359 (F^- *suII* φ80 P2 [r_k^- m_k^+]) (24) was used to construct the genomic library. *E. coli* TG1 [Δ(*lac pro*) *supE thi hsdS* F′(*traD36 proAB*$^+$ *lacI*^{q} *lacZ*Δ*M15*)] (44) was used to grow up M13mp recombinant bacteriophages and as a host for the different plasmids. Media and culture conditions are described in reference 8.

Nucleic acid procedures and construction of the genomic library. Standard recombinant DNA techniques were used as described previously (26). Preparation of phage and phage DNA is described in reference 24. A genomic library of 14- to 20-kilobase-pair *Sau*3AI fragments of "*C. saccharolyticum*" was prepared in the λ1059 *Bam*HI substitution vector as described previously (24).

Identification of xylanase and β-xylosidase-positive λ1059 clones. Individual plaques were stabbed to duplicate agar plates overlaid with soft agar containing *E. coli* Q359. The plates were incubated overnight at 37°C, and one set of plates was overlaid with 0.8% (wt/vol) soft agar containing 0.8% (wt/vol) xylan coupled to Remazol brilliant blue or oat spelt xylan (Sigma Chemical Co.) plus 1% (wt/vol) OPNX and incubated at 70°C for 2 h. Xylanase activity was detected as a pale blue halo on plates containing xylan coupled to Remazol brilliant blue. β-Xylosidase activity was detected as a diffuse yellow color due to the hydrolysis of ONPX. These plates were subsequently stained with Congo red to detect xylanase activity, which appeared as a yellow halo in a red background (42).

Construction of pNZ1076 and derivatives. DNA of a λ1059 recombinant expressing xylanase and β-xylosidase activity was partially digested with *Sau*3AI, and the fragments were ligated into pBR322 (6) to give pNZ1076, which expressed xylanase and β-xylosidase activity. The approximately 7-kilobase thermophilic DNA fragment inserted in pNZ1076 was deleted with BAL31 exonuclease from the *Sal*I site in the vector portion of pNZ1076, cut with *Eco*RI, and recloned in pCGN566 (41) to give pNZ1400, pNZ1408, pNZ1437, and pNZ1440. Plasmids pNZ1437 and pNZ1440 also have the *Eco*RV fragment deleted. Other deletions were constructed by digesting pNZ1400 with *Hin*cII and religation (pNZ1438) and by directional cloning of restriction enzyme fragments from pNZ1400 in pCGN566, pBS(−), and pUC18 (39, 41, 49). For diagrams of all plasmids, see Fig. 2.

Determination of enzyme activities in *E. coli*. The presence of xylanase activity was tested by using the Congo red staining method (42). Bacteria were grown overnight at 37°C on minimal agar plates, overlaid with 0.8% (wt/vol) soft agar containing 0.8% (wt/vol) oat spelt xylan, and incubated at 70°C for 3 to 5 h prior to staining with Congo red. β-Xylosidase activity and acetyl esterase activity were determined as follows. Cells were grown overnight in L broth at 37°C. A 500-μl portion of cells was spun down and sus-

* Corresponding author.

```
1    GGATCCCCGCAAAGCCTAAAATAAGTACATTTAGAATGATGGCAGAAAATGGTTATATTACCCTTGAATTTACGTTAAGTAAAAATGCTGTGGTGCTTTT  100
101  TGAGGTAAGCAAGGTTGTAGATGAGTCAGATACTTATATAGGACTTGACGATAGTAAAATACCAGGTTATTAGTTGCTTTATAAAATAAAAGGAATGAGG  200
                                                                                                  M  R
201  TGTTTAATTGTGTGCGAAAATTTAGAGATGCTAAACTTATCATTAGCAAAAACATACAAAGATTACTTTAAAATAGGTGCTGCAGTAACTGCGAAAGATT  300
     C  L  I  V  C  E  N  L  E  M  L  N  L  S  L  A  K  T  Y  K  D  Y  F  K  I  G  A  A  V  T  A  K  D  L
301  TAGAAGGAGTTCATAGGGATATTCTTTTGAAGCATTTTAATAGCCTCACACCAGAAAATGCCATGAAGTTTGAAAATATTCATCCAGAAGAGCAGAGATA  400
       E  G  V  H  R  D  I  L  L  K  H  F  N  S  L  T  P  E  N  A  M  K  F  E  N  I  H  P  E  E  Q  R  Y
401  TAATTTTGAAGAGGTTGCCAGGATAAAAGAGTTTGCAATTAAAAATGACATGAAGTTAAGAGGACATACATTTGTTTGGCATAATCAAACTCCGGGGTGG  500
       N  F  E  E  V  A  R  I  K  E  F  A  I  K  N  D  M  K  L  R  G  H  T  F  V  W  H  N  Q  T  P  G  W
501  GTGTTTTTAGATAAGAATGGGGAAGAAGCCTCAAAAGAGTTAGTTATTGAAAGGTTAAGAGAGCATATAAAAACTTTGTGTGAGAGATACAAGGATGTAG  600
       V  F  L  D  K  N  G  E  E  A  S  K  E  L  V  I  E  R  L  R  E  H  I  K  T  L  C  E  R  Y  K  D  V  V
601  TATATGCGTGGGATGTGGTGAACGAAGCAGTAGAAGATAAAACAGAAAAGCTTTGCGAGAATCAAACTGGAGAAAAATTATTGGAGATGATTATATTAA  700
       Y  A  W  D  V  V  N  E  A  V  E  D  K  T  E  K  L  L  R  E  S  N  W  R  K  I  I  G  D  D  Y  I  K
701  AATTGCTTTTGAGATAGCAAGAGAATATGCAGGAGATGCAAAGTTATTTTATAACGATTATAACAATGAAATGCCTTATAAATTAGAAAAAACCTACAAA  800
       I  A  F  E  I  A  R  E  Y  A  G  D  A  K  L  F  Y  N  D  Y  N  N  E  M  P  Y  K  L  E  K  T  Y  K
801  GTTCTAAAAGAGCTTTTAGAAAGAGGTACTCCAATAGATGGAATTGGTATACAAGCACACTGGAATATATGGGATAAAAATCTTGTTAGTAATTTAAAAA  900
       V  L  K  E  L  L  E  R  G  T  P  I  D  G  I  G  I  Q  A  H  W  N  I  W  D  K  N  L  V  S  N  L  K  K
901  AGGCTATAGAAGTATATGCTTCCTTAGGTTTAGAAATTCATATTACAGAACTTGACATTTCAGTATTTGAGTTTGAAGATAAGAGGACTGACTTGTTTGA  1000
       A  I  E  V  Y  A  S  L  G  L  E  I  H  I  T  E  L  D  I  S  V  F  E  F  E  D  K  R  T  D  L  F  E
1001 ACCAACCCCGGAAATGCTTGAACTACAAGCAAAAGTATATGAAGATGTATTTGCAGTTTTTCGAGAATATAAAGATGTAATAACTTCTGTTACATTATGG  1100
       P  T  P  E  M  L  E  L  Q  A  K  V  Y  E  D  V  F  A  V  F  R  E  Y  K  D  V  I  T  S  V  T  L  W
1101 GGTATTAGCGACAGACACACATGGAAAGATAACTTCCCTGTAAAGGGTCGAAAAGATTGGCCTCTCTTATTCGACGTAAATGGAAAACCAAAAGAAGCCT  1200
       G  I  S  D  R  H  T  W  K  D  N  F  P  V  K  G  R  K  D  W  P  L  L  F  D  V  N  G  K  P  K  E  A  L
1201 TGTACAGGATATTAAGATTTTAAAGATTTTTTAACGAAGAAAGGGGTTCTTTTAATATGGCTATCATGCAAATCAACTTTTATTCAAAGATGTTGAAAAA  1300
       Y  R  I  L  R  F  *                                    M  A  I  M  Q  I  N  F  Y  S  K  M  L  K  K
1301 GAACACAACAATTTTGGCCATTTTACCCGTAGATAAACCAGATAAGAAATTCCAGAAAGATGTTGATAGTGAAAATTTGAAAACCTTATATCTTTTGCAT  1400
       N  T  T  I  L  A  I  L  P  V  D  K  P  D  K  K  F  Q  K  D  V  D  S  E  N  L  K  T  L  Y  L  L  H
1401 GGTTATGCTGGTAACTACATGGATTGGTTGTGTGGAGCCCGAATTGTTGAATTATCAATGCGATATAATGTTGCTGTGTTTCTGCCATCAGGTGAAAATA  1500
       G  Y  A  G  N  Y  M  D  W  L  C  G  A  R  I  V  E  L  S  M  R  Y  N  V  A  V  F  L  P  S  G  E  N  S
1501 GTTTTTATTTAGATGATGAAGAAAAGGAAGAATATTTTGGTGAATTTGTGGGAAATGAAATTATAGAATTTACAAGAAGCGTTTTTCCTATTCCTCAAAA  1600
       F  Y  L  D  D  E  E  K  E  E  Y  F  G  E  F  V  G  N  E  I  I  E  F  T  R  S  V  F  P  I  P  Q  K
1601 AAGGGAAAAAACTTTTATTGGCGGTTTATCAATGGGAGGTTACGGTGCTCTTAGAAATGGGCTTAAATATAACAAGAATTTTGTAGGTATAATAGCTTTA  1700
       R  E  K  T  F  I  G  G  L  S  M  G  G  Y  G  A  L  R  N  G  L  K  Y  N  K  N  F  V  G  I  I  A  L
1701 TCATCAGCACTAATAATTCATAAGATTGCAGGTATTCCTAAGGATTATAGGAATGCTTATGCAAGTTATAACTATTATAGACGAGTGTTTGGAGACCTAA  1800
       S  S  A  L  I  I  H  K  I  A  G  I  P  K  D  Y  R  N  A  Y  A  S  Y  N  Y  Y  R  R  V  F  G  D  L  N
1801 ACTCTTTAATAGGTAGCGATAAAGACATAAATGCCTTAGTTACTAAGCTAAAACAAGAAAAAGGTAGTATTCCAAAAATATACATGGCATGCGGCAGAGA  1900
       S  L  I  G  S  D  K  D  I  N  A  L  V  T  K  L  K  Q  E  K  G  S  I  P  K  I  Y  M  A  C  G  R  D
1901 TGACTTTTAGTTCAAGAAAACAGAGATTTATTTAATTTTTTGAAAAATGAAGGTATAGACGTGGTTATGAGGAAGACGAAGGTGGACATGACTGGGAT  2000
       D  F  L  V  Q  E  N  R  D  L  F  N  F  L  K  N  E  G  I  D  V  V  Y  E  E  D  E  G  G  H  D  W  D
2001 TTTTGGAACAAATATATTGCAAATGCTTTTGAGTGGATGAGTAAGGTTTCTGATTAAGTCTTCACGTACCCTGTTTTAAGTTTTACAAATAGATTTGTGG  2100
       F  W  N  K  Y  I  A  N  A  F  E  W  M  S  K  V  S  D  *
2101 GGTGAATAGGTTTTTTTAACACTATTTTATTAAGGAAGAGGATGAAAAATAAAAAAAGTGGACAAATTTCTTGTTAATTGTAATTACATGCATTGCAATG  2200
                                                                                                  M
2201 GTTTTCTTTTTTACATCGTGTACTATTCAGTCTGCTATAGAGCAGAAGAAAACTGTTGAGGAAATCTTGGGAAAAATAGGTGAGAGTGAGGACAAAACAA  2300
     V  F  F  F  T  S  C  T  I  Q  S  A  I  E  Q  K  K  T  V  E  E  I  L  G  K  I  G  E  S  E  D  K  T  N
2301 ATTCAAGGGGGCAACCAGCAACAATGAAAGAGGATGAAGTTGAAGATAATCCTTTAAAAGATGTATATAAAGATTATTTCCTGGTTGGAGCAGCAATTAA  2400
       S  R  G  Q  P  A  T  M  K  E  D  E  V  E  D  N  P  L  K  D  V  Y  K  D  Y  F  L  V  G  A  A  I  N
2401 TGGCTATTCTGTTGAAACTGCTGCTATCAATCATCCTGGTATGGCTGCAATTTTGAAAAAAACTTTAACAGTACAACCCTATCTAATTTGATGAAACAAC  2500
       G  Y  S  V  E  T  A  A  I  N  H  P  G  M  A  A  I  L  K  K  T  L  T  V  Q  P  Y  L  I  *  M  K  Q  Q
2501 AATACCTTTTAGATTATGAAGCTACAAAAGCAAGTAAAAATGGAATGCCAGTGTGTAAATTTGACAGCTGCATTCCTGCTTTACAATTTTGTAAGGAAAA  2600
       Y  L  L  D  Y  E  A  T  K  A  S  K  N  G  M  P  V  C  K  F  D  S  C  I  P  A  L  Q  F  C  K  E  N
2601 TGGCATAAAAATGAGAGGACATGTGTTAGTATGGCATAATCAGACACCAGAATGGTTTTTCCACAAAGACTATGATGTATCGAAACCACTTGTAGATGCT  2700
       G  I  K  M  R  G  H  V  L  V  W  H  N  Q  T  P  E  W  F  F  H  K  D  Y  D  V  S  K  P  L  V  D  A
2701 GCTACTATGGAACGCCGGTTGGAAAGTTATATCAAACAGGTAATTGAATTTTGTCAAAAAAATTATCCCGGTGTAGTCTATTGCTGGGATGTTGTTAACG  2800
       A  T  M  E  R  R  L  E  S  Y  I  K  Q  V  I  E  F  C  Q  K  N  Y  P  G  V  V  Y  C  W  D  V  V  N  E
2801 AAGCTATACTTGATGATGGTTCATGGAGAGAAATCAATAATAATTGGTATACCATTATGAAAGAAAAGTATGTGGAAAAGGCATTTTATTATGCAAGAAA  2900
       A  I  L  D  D  G  S  W  R  E  I  N  N  N  W  Y  T  I  M  K  E  K  Y  V  E  K  A  F  Y  Y  A  R  K
2901 ATATGCCAAAAAAGATGTTGCCCTGTTTTACAATGATTACAATGTTTTTCTCCCTGCAAAGAGAGAAGCAATTTATAATCTTGCTCAGAAACTTAAAGAA  3000
       Y  A  K  K  D  V  A  L  F  Y  N  D  Y  N  V  F  L  P  A  K  R  E  A  I  Y  N  L  A  Q  K  L  K  E
3001 AAAGGATTGATTGACGGGTTGGGTCTTCAACCTACAGTAGGCTTGAATTATCCTGAATTAGATTCTGATGATATAGATTCATTCAAAACGACATTAGAAA  3100
       K  G  L  I  D  G  L  G  L  Q  P  T  V  G  L  N  Y  P  E  L  D  S  D  D  I  D  S  F  K  T  T  L  E  T
3101 CATTTGCAAAACTTGGCTTACAAATTCATATTACTGAGTTAAATTTTGAAATAAAGGGAGATGAGAGCAATCGTACTCCTGAAAATCTCAAAAAACAAGC  3200
       F  A  K  L  G  L  Q  I  H  I  T  E  L  N  F  E  I  K  G  D  E  S  N  R  T  P  E  N  L  K  K  Q  A
3201 AGATAGGTATTACGAAATGATGAAGTTATTATTGAAGGAAGATACTGATAATGGTGGGCCTTGCAACATAACTTGTGTTACTGTTTTTGGTATCTGTGAC  3300
       D  R  Y  Y  E  M  M  K  L  L  L  K  E  D  T  D  N  G  G  P  C  N  I  T  C  V  T  V  F  G  I  C  D
3301 GATTATCCACTATATAAAAATTTTAAGCAGTGCATGTATCTTTGGGATAAAAATTGCAATCCTAAACCATGTTTTTATTCATTTCTCCAAGCAGGTTTAG  3400
       D  Y  P  L  Y  K  N  F  K  Q  C  M  Y  L  W  D  K  N  C  N  P  K  P  C  F  Y  S  F  L  Q  A  G  L  D
3401 ACTGGAAAGCATCTTTATTAAGCAAATAAGAATGAACAACACTTATGGAGAGGAGGAAAATAATGAAAATAACTATTAATTATGGAAAGAGACTTGGGAA  3500
       W  K  A  S  L  L  S  K  *     *              M  E  R  R  K  I  M  K  I  T  I  N  Y  G  K  R  L  G  K
3501 AATAAACAAATTTTGGGCAAAATGTGTTGGAAGCTGTCATGCTACAACTGCGTTAAGAGAAGACTGGCGAAAGCAATTAAAAAAATGTCGTGACGAACTT  3600
       I  N  K  F  W  A  K  C  V  G  S  C  H  A  T  T  A  L  R  E  D  W  R  K  Q  L  K  K  C  R  D  E  L
3601 GGTTTTGAGTATATTCGATTTCATGGTTGGTTGAATGATGATATGAGTGTTTGTTTTAGAAATGATGATGGGCTACTTTCATTCTCATTCTTCAACATAG  3700
       G  F  E  Y  I  R  F  H  G  W  L  N  D  D  M  S  V  C  F  R  N  D  D  G  L  L  S  F  S  F  F  N  I  D
3701 ATTCTATAATTGATTTTCTTTTGGAGATAGGTATGAAACCATTTATTGAACTGAGCTTTATGCCAGAAGCGTTAGCGTCAGGTACAAAGACAGTTTTCCA  3800
       S  I  I  D  F  L  L  E  I  G  M  K  P  F  I  E  L  S  F  M  P  E  A  L  A  S  G  T  K  T  V  F  H
3801 TTACAAAGGAAATATAACACCGCCGAAATCTTATGAAGAATGGGGTCAGCTGATTGAGGAGTTAGCAAGGCATCTTATTAGCAGATATGGGAAAAATGAA  3900
       Y  K  G  N  I  T  P  P  K  S  Y  E  E  W  G  Q  L  I  E  E  L  A  R  H  L  I  S  R  Y  G  K  N  E
3901 GTAAGAGAATGGTTTTTTGAGGTATGGAACGAACCAAATCTAAAGGATTTCTTCTGGGCAGGAACAATGGAAGAATATTTTAAGCTTTACAAATATGCTG  4000
       V  R  E  W  F  F  E  V  W  N  E  P  N  L  K  D  F  F  W  A  G  T  M  E  E  Y  F  K  L  Y  K  Y  A  A
4001 CTTTTGCAATAAAGAAAGTGGACTCTGAACTAAGGGTAGGTGGACCAGCTACTGCAATCGATGCATGGATACCTGAACTAAAAGATTTTTGTACAAAAAA  4100
       F  A  I  K  K  V  D  S  E  L  R  V  G  G  P  A  T  A  I  D  A  W  I  P  E  L  K  D  F  C  T  K  N
4101 TGGTGTTCCAATAGATTTTATTTCAACGCATCAATATCCAACAGATTTAGCATTCAGTACAAGCTCAAATATGGAAGAGGCTATGGCAAAAGCAAAGAGA  4200
       G  V  P  I  D  F  I  S  T  H  Q  Y  P  T  D  L  A  F  S  T  S  S  N  M  E  E  A  M  A  K  A  K  R
4201 GGTGAATTAGCAGAGAGGGTAAAAAAGGCTTTAGAGGAAGCATATCCATTGCCTGTTTACTACACTGAATGGAATAACTCTCCAAGTCCTCGAGACCCAT  4300
       G  E  L  A  E  R  V  K  K  A  L  E  E  A  Y  P  L  P  V  Y  Y  T  E  W  N  N  S  P  S  P  R  D  P  Y
4301 ATCACGACATACCTTACGATGCTGCTTTTATTGTAAAAACAATAATTGACATTATAGATTTACCACTTGGGTGTTATTCTTATTGGACATTTACAGATAT  4400
       H  D  I  P  Y  D  A  A  F  I  V  K  T  I  I  D  I  I  D  L  P  L  G  C  Y  S  Y  W  T  F  T  D  I
```

FIG. 1. Nucleotide sequence of the "*Caldocellum*" DNA inserted in pNZ1400 (see Fig. 2, shaded bar). Numbers on the side of the nucleotide sequence denote nucleotide positions. The first 7 nucleotides including the *Bam*HI site are part of the multiple cloning site of the vector pCGN566. Six ORFs are present on the cloned fragment coding for a xylanase/β-xylosidase (ORF1, XynA, positions 195 to 1220), an acetyl esterase (ORF2, XynC, positions 1257 to 2054), two ORFs with unknown function (ORF3, positions 2198 to 2488; ORF4, positions 2491 to 3426), and a β-xylosidase (ORF5, XynB, position 3445 or 3463 to 4908; two translational start sites are possible). A sixth ORF (ORF6, position 5439) reads out of the cloned fragment. Sequences homologous to the *E. coli* SD sequence (11) are underlined in front of each ORF. A * shows stop codons. The putative signal peptide for the *xynA* gene product (positions 195 to 293) is indicated by a box around the amino acids. There are three possible cleavage sites (see text).

pended in 100 μl of 50 mM citrate–100 mM phosphate buffer (pH 6.5) for the β-xylosidase assay. For the acetyl esterase assay, cells were suspended in the same buffer at pH 6. A drop of toluene was then added. For the β-xylosidase assay, 10 μl of ONPX (20 mg/ml in dimethylformamide) was added and the cells were incubated at 70°C for 30 min. Cells expressing β-xylosidase activity were detected by the yellow color produced from the degradation of the substrate (21). For the acetyl esterase assay, 10 μl of α-naphthyl acetate (20 mg/ml in dimethylformamide) was added and the cells were incubated at 70°C. After 30 min, 50 μl of coloring reagent (1 M sodium acetate buffer [pH 4.5] containing 0.01% Fast Corinth Salt V and 10% Tween 20 [36]) was added. Cells expressing acetyl esterase activity were detected by the purple color produced. *E. coli* TG1 was used as a negative control.

SDS-polyacrylamide gel electrophoresis. A 15-ml portion of an overnight culture in L broth was spun down, resuspended in 1.5 ml of L broth, and sonicated for 15 s. Cell debris were removed by centrifugation. Portions (50 μl) of the supernatant were boiled for 5 min in the presence of sodium dodecyl sulfate (SDS)-loading buffer. Samples were electrophoresed in an SDS–10% (wt/vol) polyacrylamide gel by the method of Laemmli (22). For activity staining, SDS was removed and the gel was assayed for xylanase and β-xylosidase by transfer to an agar sheet containing 0.5% (wt/vol) Remazol brilliant blue-xylan or 1% (wt/vol) ONPX. The method is described in reference 2. Xylanase activity was detected as blue halo, whereas β-xylosidase activity was seen as yellow band.

DNA sequence analysis. The sequence of the entire "*Caldocellum*" insert in pNZ1400 (6,067 base pairs) has been determined by dideoxy sequencing (37) of a randomly sheared library and a pseudorandom library (by using the enzymes *Hae*III, *Rsa*I, and *Alu*I) constructed in M13 vectors and after directional cloning of restriction fragments into M13 vectors and plasmids pBS(+) and pBS(−) (1, 39). All computer analysis of the sequence data was carried out by using the Sequence Analysis Software Package of the University of Wisconsin Genetics Computer Group on a Micro VaxII (Digital Equipment Corp.).

RESULTS

Sequence analysis and location of enzyme activities. Sequence analysis of the thermophilic DNA fragment present in pNZ1400 (Fig. 1 and 2) showed five open reading frames (ORFs) which could code for proteins of M_r 40,455 (ORF1, positions 195 to 1220), M_r 30,611 (ORF2, positions 1257 to 2054), M_r 10,731 (ORF3, positions 2198 to 2488), M_r 36,493 (ORF4, positions 2491 to 3426), and M_r 56,365 or 55,551 (ORF5, positions 3445 to 4908). There are two possible translational start sites for ORF5 (positions 3445 or 3463). Each is preceded by a Shine-Dalgarno (SD) sequence. Which ATG is used for translation is not known, since the N-terminal sequence of the ORF5 gene product has not been determined. A sixth ORF (ORF6, position 5439) reads out of the cloned fragment.

Deletion analysis of the cloned thermophilic DNA fragment (Fig. 2) shows that ORF1 codes for an enzyme (called XynA) that acts on xylan and on ONPX, a substrate normally used to detect β-xylosidase activity (21). Xylanase/β-xylosidase activity can be found on the 1,317-base-pair *Bam*HI-*Bal*I fragment of pNZ1435 containing only ORF1. Deletion of 20 amino acids (*Hae*III site, pNZ1445) or 131 amino acids (*Rsa*I site, pNZ1446) from the C terminus of ORF1 abolishes xylanase and β-xylosidase activity. ORF5 codes for a β-xylosidase (called XynB), as can be seen by the enzyme activities encoded by pNZ1438, which contains the C-terminal part of ORF4 and ORF5. Deletion into ORF5 (pNZ1408) abolishes β-xylosidase activity.

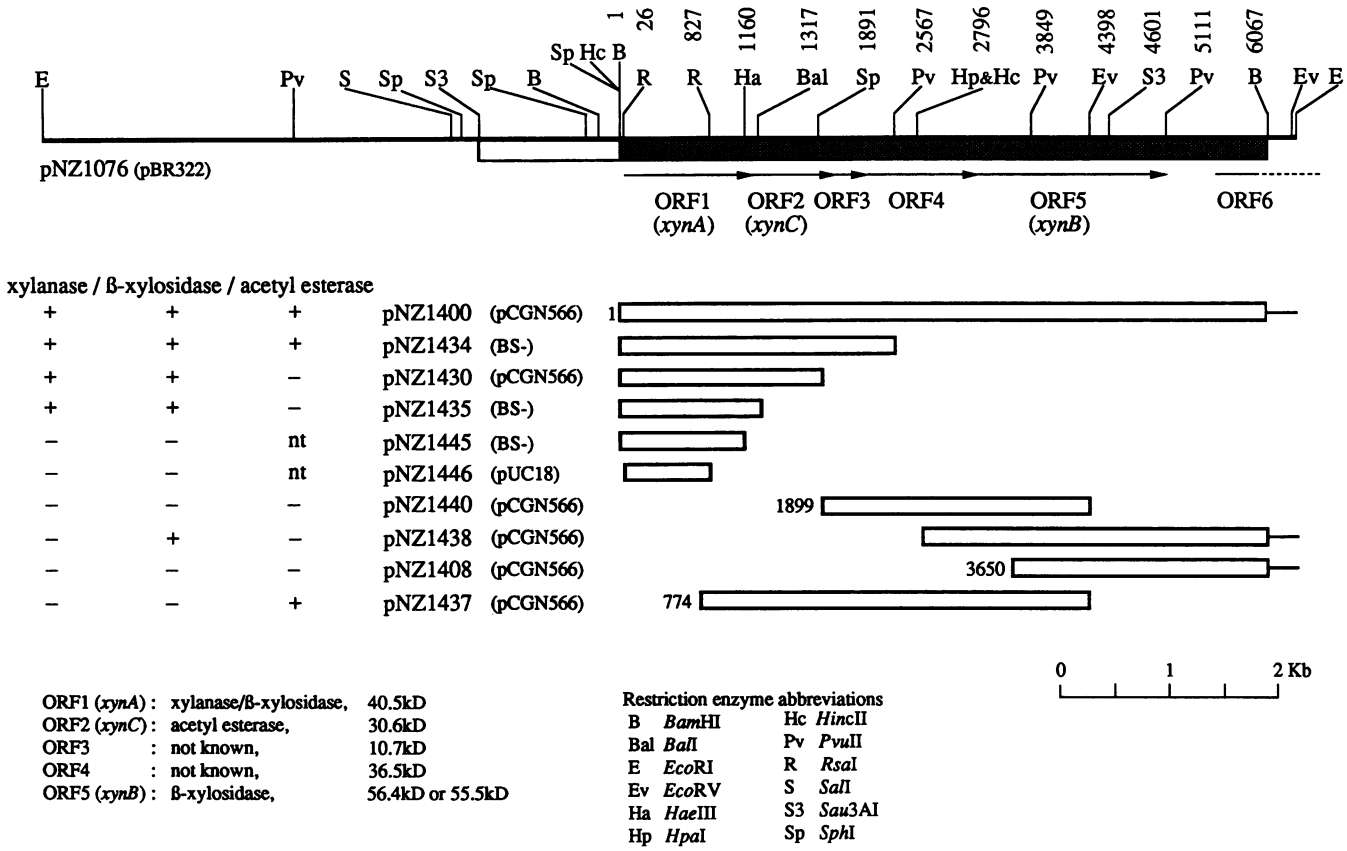

FIG. 2. Restriction map of thermophilic DNA cloned in pNZ1076 (▨). The shaded bar indicates the sequenced part of the cloned fragment and is identical with the insert in pNZ1400. The extent of thermophilic DNA remaining in derivatives of pNZ1076 constructed by Bal31 exonuclease digestion or by subcloning of specific restriction enzyme fragments is shown as enclosed bars. Numbers in front of corresponding bars indicate the end of the BAL31 deletion according to the sequence (Fig. 1). Vectors used for subcloning are shown in parentheses beside plasmid numbers. Numbers above restriction enzyme sites give the distance in base pairs from the *Bam*HI site of the multiple cloning site of pCGN566 which is present in pNZ1400. Restriction enzyme sites clustered around the *Bam*HI site at position 1 are part of the multiple cloning site. Not all *Rsa*I and *Hae*III sites are shown. The *Sau*3AI and *Hin*cII sites in pBR322 are not shown. The arrows show the ORFs that can be identified by computer analysis. Only part of ORF6 is present on the cloned fragment. There are two translational start sites for ORF5, each preceded by a SD sequence; hence, two molecular masses are given (in kilodaltons [kD]). No enzymatic activities could be found for ORF3 and ORF4. Symbols: +, enzyme activity; −, no enzyme activity; nt, not tested. Kb, Kilobases.

Preliminary results suggest that ORF2 could code for an acetyl esterase (called XynC). *E. coli* TG1(pNZ1434) degraded α-naphthyl acetate, a substrate used for acetyl esterases (36), whereas no acetyl esterase activity could be found by using pNZ1430 or pNZ1435. Acetyl esterase activity does not seem to be encoded by ORF3 or ORF4, since no enzyme activity was detected in TG1 (pNZ1440).

Ribosomal-binding sites. Sequences homologous to SD sequences from *E. coli* and *Bacillus subtilis* (11, 27, 29) can be located from 4 to 13 base pairs upstream of the initiation codon for all putative proteins, but the similarities are lowest for ORF3 and ORF4 (Table 1). ORF1, ORF3, and ORF4 do not have the GGAGGA sequence most often found in SD sequences (11). Comparison of the two SD sequences corresponding to the two possible translational start sites of ORF5 (Table 1) shows that SD1 (corresponding to the ATG at position 3463 [Fig. 1]) has good homology to both the *E. coli* and *B. subtilis* SD sequences, whereas SD2 shows only minor homology. This observation suggests that SD1 may be preferentially used as a ribosome-binding site in *E. coli*.

Putative promoter sequences. Several *E. coli* ρ^{70}-like and *B. subtilis* ρ^{55}-like promoter structures (17, 18) can be found upstream of all ORFs. Transcription of the ORFs seems not to be dependent on plasmid promoter sequences, since the genes are cloned in the opposite direction to the tetracycline resistance promoter of pBR322 (6) or to the *lacZ* promoter region of pCGN566 (41). The promoter regions for ORF2 and ORF5 might be positioned within the structural gene of ORF1 or ORF4, respectively, since there are only a few nucleotides between the end of one ORF and the beginning of the other (Fig. 1). These sequences seem to be used for transcription in *E. coli*, as shown by the enzyme activities encoded by pNZ1437 or pNZ1438 (Fig. 2). In these plasmids the ORF upstream of ORF2 or ORF5 has been deleted, suggesting that ORF2 and ORF5 are expressed independently. Experiments are being carried out to study the transcriptional organization of the ORFs.

Signal sequences. Only ORF1 appears to have a signal sequence in broad conformity with the rules given in reference 45. This observation is to be expected, since xylanases are thought to be secreted enzymes (5). A hydropathy plot of ORF1 shows a hydrophobic region within the first 10 to 15 amino acids (data not shown), and there is a charged amino acid within the first 5 amino acids (Arg-2), as well as a 16-amino-acid, mainly hydrophobic sequence, which contains two negatively charged amino acids (Glu-8 and Glu-11)

TABLE 1. Comparison of ribosome-binding site sequences of ORFs found on the sequenced DNA fragment from "*C. saccharolyticum*" with the SD sequences from *B. subtilis* and *E. coli*

Organism[a]	Sequence[b]
E. coli	C T A G T G G A G G A A T
B. subtilis	A G A A A G G A G G T G A T C
"*C. saccharolyticum*"	
ORF1 (*xynA*)	T T A G T T G C T T T A T
ORF2 (*xynC*)	A G A A A G G G G T T C T T
ORF3	T G T T A A T T G T A A T
ORF4	A C A G T A C A A C C C T A T C
ORF5 (*xynB*) SD1	G G A G A G G A G G A A A
ORF5 (*xynB*) SD2	A G C A A A T A A G A A T G
ORF6	G T T T C T G A G G T T T

[a] There are two possible translational start sites for ORF5, each preceded by an SD sequence. ORF1 codes for a xylanase/β-xylosidase (XynA); ORF2 is probably an acetyl esterase (XynC); and ORF5 codes for a β-xylosidase (XynB). No enzyme activity has been found for ORF3 and ORF4. ORF6 reads out of the cloned fragment.

[b] Homology with the *E. coli* SD sequence (11) is shown as *; homology with the *B. subtilis* SD sequence (27, 29) is underlined.

and two polar amino acids (Asn-9 and Asn-14) followed by a stretch of charged and hydrophobic amino acids with small side chains. The cleavage site could be either after the two alanine residues (positions 29 and 30) or after Thr-32 or Ala-33 (Fig. 3).

The presence of charged and polar amino acids in the hydrophobic core and at the C-terminal end of the ORF1 signal peptide is quite atypical and cannot be found in the xylanases from the alkalophilic *Bacillus* sp. strain C125 (15) or *B. pumilus* (9), which are compared in Fig. 3. It is also different from the signal peptide found in the *celB* gene of "*C. saccharolyticum*" (Fig. 3) (38) and from most of the procaryotic signal peptides compiled in reference 45. A similar structure can be found for the signal peptides of penicillinases from *B. cereus* (28) and *B. licheniformis* (Fig. 3) (31). The *B. licheniformis* penicillinase signal peptide is 50 amino acids long, and 14 of the last 22 are charged and polar amino acids. In *B. cereus* the signal peptide is 45 amino acids long, and 7 of the last 14 are charged or polar residues.

Homology comparisons. Comparison of the predicted amino acid sequence of ORF1 (*xynA* gene) with those of other xylanases, β-xylosidases, glucanases, or β-glucosidases compiled in the GenBank data base revealed significant similarities to the following enzymes (Fig. 4): (i) xylanase of alkalophilic *Bacillus* sp. strain C125 (*xynA*) (15); (ii) C-terminal domain containing the xylanase activity of the xylanase (*xynZ*) of *Clostridium thermocellum* (13); (iii) cellobiohydrolase domain of the "*Caldocellum*" bifunctional exocellulase-endocellulase (*celB*) (38); (iv) exoglucanase *cex* from *Cellulomonas fimi* (32); and (v) ORF4 (function unknown) (see above). We also found similarity between ORF1 and the N-terminal sequence of the xylanase from the yeast *Cryptococcus albidus* (7; data not shown). It has been shown by others (46) that this enzyme is similar to the exoglucanase from *Cellulomonas fimi* and to the xylanase from the alkalophilic *Bacillus* sp. strain C125, which are compared in Fig. 4. We found no similarities to the xylanases or β-xylosidases reported from *B. circulans* (48), *B. subtilis* (34), and *B. pumilus* (9, 30).

Enzyme characteristics. To estimate the molecular mass of the xylanase/β-xylosidase and β-xylosidase, the proteins produced by pNZ1076 were separated on a denaturing SDS-polyacrylamide gel, the SDS was removed, and the gel was assayed for xylanase and β-xylosidase activity. Activity gels showed ONPX breakdown at positions corresponding to proteins of M_r 42,000 and 76,000 (data not shown), in reasonable agreement with predicted protein sizes from the sequence data. Xylanase activity was also seen at M_r 42,000,

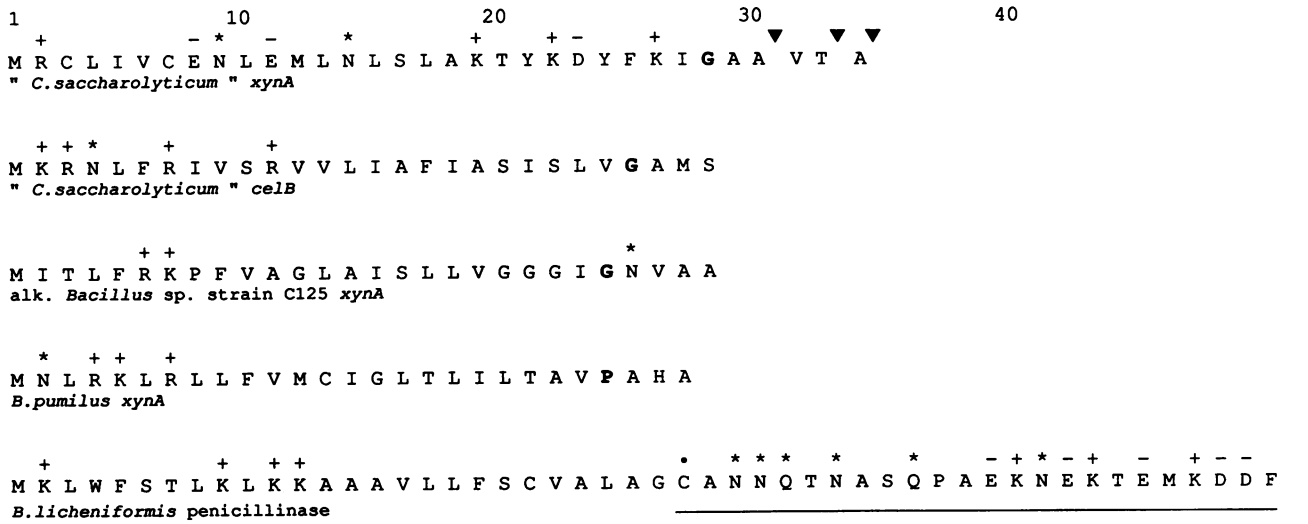

FIG. 3. Comparison of the "*C. saccharolyticum*" xylanase/β-xylosidase (*xynA*) signal peptide with signal peptides of two *Bacillus* xylanases (9, 15) and a penicillinase (31) and the "*C. saccharolyticum*" *celB* gene product (38). Positively charged amino acids are indicated by + above their single-letter code; negatively charged amino acids are indicated by −; polar amino acids are marked by *; and the helix-breaking residues are in bold type. Three possible cleavage sites are shown by the ▼ above the *xynA* signal peptide. A dot marks the cysteine residue used for the glyceride thioether bond in *B. licheniformis* membrane-bound penicillinase. The C-terminal part integrated into the membrane is underlined.

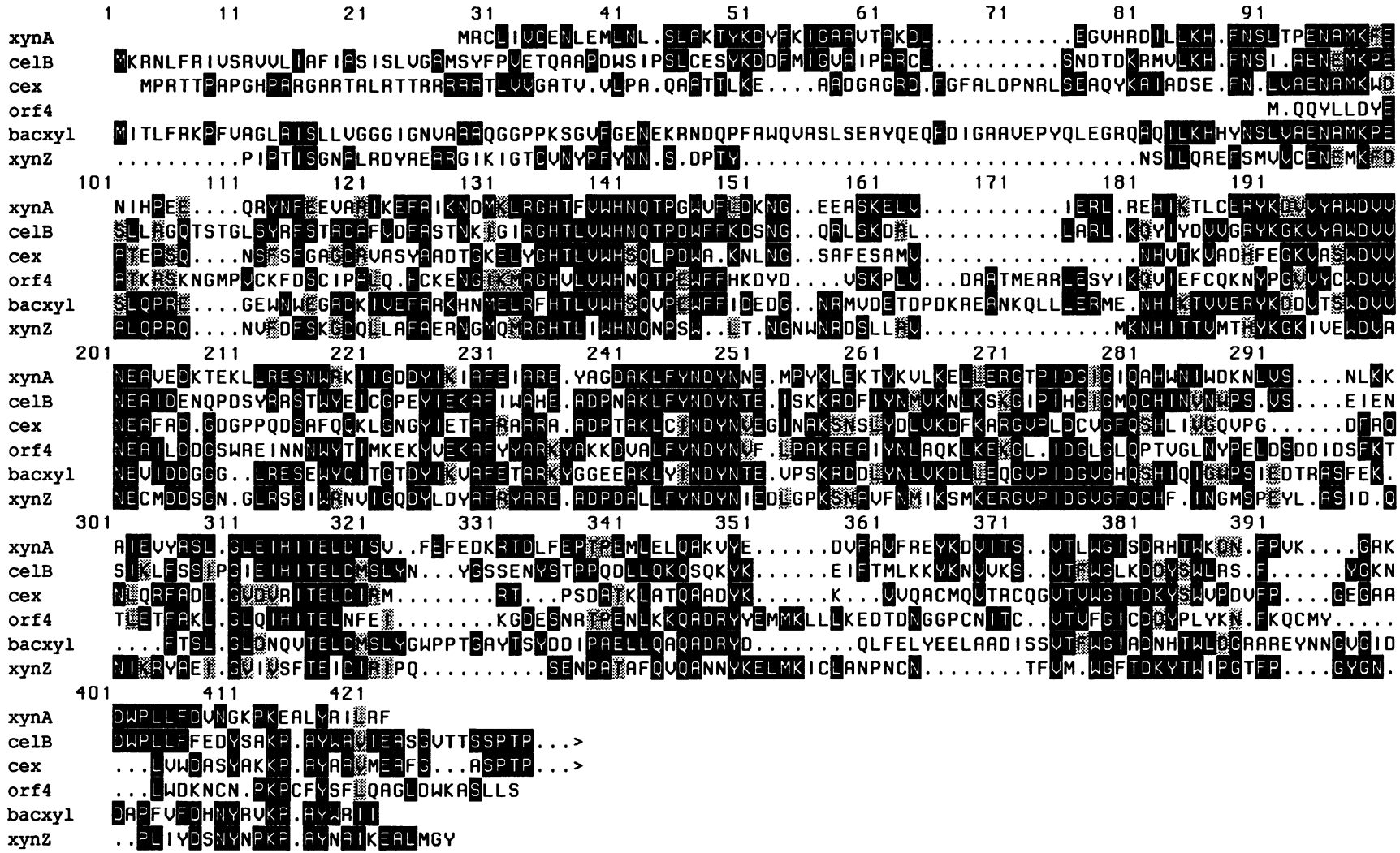

FIG. 4. Alignment of similar regions of primary structure of the "*C. saccharolyticum*" xylanase/β-xylosidase (xynA) and the exoglucanase domains of "*C. saccharolyticum*" cellulase (celB) (38) and *Cellulomonas fimi* (cex) (32), ORF4 (Fig. 1), the alkalophilic *Bacillus* sp. strain C125 xylanase (*xynA*; bacxyl) (15), and the C-terminal domain of the *Clostridium thermocellum xynZ* gene product which contains the xylanolytic activity (from amino acid 509 onwards; xynZ) (13). Identical amino acids are in reversed font. If the amino acids are not present in all of the protein sequences compared, they are shown either in reversed font or in stippled boxes. Dashed lines indicate where it is necessary to introduce gaps for better alignment. Numbers above the lines denote amino acid positions of the proteins.

which is in agreement with the suggestion that the xylanase also acts on ONPX.

DISCUSSION

We report here the cloning and sequence structure of a gene (*xynA*) that has xylanase/β-xylosidase activity, of a separate gene that does not show activity against xylan but does have β-xylosidase activity (*xynB*), and of a third gene, which could code for an acetyl esterase (*xynC*). These three genes have little sequence similarity with each other. The *xynA* gene product has sequence similarity with the gene product encoded by the putative ORF4. Consideration of the upstream control sequences of ORF4 suggests that this protein may not be expressed, and it could be a pseudogene by analogy with the *celX* gene described for *Clostridium thermocellum* (14).

Others have noted the homology between the exoglucanase domain of *Cellulomonas fimi*, the alkalophilic *Bacillus* sp. strain C125, and the *Cryptococcus albidus* xylanases and have suggested that the glycosidases from *Cellulomonas*, *Trichoderma*, and *Bacillus* spp. may have evolved from the reshuffling of two catalytic domains and several binding domains (46). The "*Caldocellum*" xylanase/β-xylosidase shares similarities with the catalytic domains described for the exoglucanase of *Cellulomonas fimi*, the cellobiohydrolase domain of "*Caldocellum*" *celB* (38), and the C-terminal domain of the *Clostridium thermocellum xynZ* gene product, which contains the xylanolytic activity (13). It also has similarities with the xylanases of the alkalophilic *Bacillus* sp. strain C125 (15) and *Cryptococcus albidus* (7). It does not show significant similarity to the endoglucanase domains compared in reference 46. Thus, the "*Caldocellum*" xylanase/β-xylosidase enzyme may be the product of reshuffling of catalytic and binding domains, resulting in an enzyme with different substrate specificity.

The similarity of the *xynA* gene product to the exoglucanase domains of *Cellulomonas fimi cex* and "*C. saccharolyticum*" *celB* gene products may be explained by the fact that the enzyme also degrades 4-methylumbelliferyl-β-D-cellobioside (MUC), a substrate used as a test for exoglucanase activity (43). Others also have recorded the action of cloned xylanases on MUC (10, 12, 16) and have presumed that this activity represents nonspecific cleavage of the agluconic bond (12). Purified xylanases have been reported which either hydrolyze xylan only (40) or also show endo- or exoglucanase activity (23). In addition to MUC activity, the *xynA* gene product shows low-level activity on carboxymethyl cellulose, as measured by reducing-sugar assays (our unpublished data).

The xylanase/β-xylosidase does not show the three-domain structure separated by "PT" boxes that characterize the bifunctional enzyme we have identified in "*Caldocellum*" (CelB) (38) and which can also be found in the *Clostridium thermocellum* xylanase (13). Nor does it show the conserved region seen by others for the endoglucanases EGA and EGB of *Clostridium thermocellum*, the endo- and exoglucanases of *Cellulomonas fimi*, and the cellulases of *Trichoderma reesei* (see reference 13 and references therein).

The putative signal peptide for the xylanase/β-xylosidase is quite atypical in comparison with leader sequences found for other xylanases (9, 13, 15, 34) or glucanases (14, 25) because of the presence of charged and polar amino acids in the hydrophobic sequence. It also differs from the "*Caldocellum*" *celB* gene product signal peptide (38). However, the charged C terminus of the *xynA* gene product signal peptide resembles signal peptides found for the penicillinases of *B. cereus* and *B. licheniformis*. The *B. cereus* β-lactamase is almost completely secreted (28), whereas in *B. licheniformis* a substantial proportion is found as hydrophobic membrane-bound form (31). The charged and polar part of the penicillinase is anchored in the membrane via a glyceride thioether bond formed with a cysteine residue. The conserved sequence proposed (31) for the membrane-bound penicillinases and other membrane-bound lipoproteins is -Leu-Ala-Gly-Cys-α-Ser-Asn- (where α designates a neutral or nonpolar residue). This sequence cannot be found in the *xynA* gene product signal peptide, but other signals may be relevant for membrane-bound "*Caldocellum*" proteins.

The temperature optimum of the "*Caldocellum*" *xynA* gene product exceeds those reported for other xylanases, for example, the xylanase of *Clostridium stercorarium*, which has an optimum of 65°C (3). Although the 70°C temperature optimum (our unpublished data) for the *xynA* gene product is not as high as for other enzymes from the same organism (4, 38), this property may be of value in the bleaching of chemical pulp (19, 33).

ACKNOWLEDGMENTS

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

We thank M. D. Gibbs for advice and assistance with computing.

LITERATURE CITED

1. **Bankier, A. T., K. M. Weston, and B. G. Barrell.** 1987. Random cloning and sequencing by the M13/dideoxynucleotide chain termination method. Methods Enzymol. **155**:51–93.
2. **Béguin, P.** 1983. Detection of cellulase activity in polyacrylamide gels using Congo Red-stained agar replicas. Anal. Biochem. **131**:333–336.
3. **Berenger, J. F., C. Frixon, J. Bigliardi, and N. Creuzet.** 1985. Production, purification, and properties of thermostable xylanase from *Clostridium stercorarium*. Can. J. Microbiol. **31**: 635–643.
4. **Bergquist, P. L., D. R. Love, J. E. Croft, M. B. Streiff, R. M. Daniel, and W. H. Morgan.** 1987. Genetics and potential biotechnological applications of thermophilic and extremely thermophilic microorganisms. Biotechnol. Genet. Eng. Rev. **5**: 199–244.
5. **Biely, P.** 1985. Microbial xylanolytic systems. Trends Biotechnol. **3**:286–290.
6. **Bolivar, F., R. L. Rodriguez, P. J. Greene, M. C. Betlach, H. L. Heynecker, H. W. Boyer, J. H. Crosa, and S. Falkow.** 1977. Construction and characterization of new cloning vehicles II. A multipurpose cloning system. Gene **2**:95–113.
7. **Boucher, F., R. Morosoli, and S. Durand.** 1988. Complete nucleotide sequence of the xylanase gene from the yeast *Cryptococcus albidus*. Nucleic Acids Res. **16**:9874.
8. **Croft, J. E., D. R. Love, and P. L. Bergquist.** 1987. Expression of leucine genes from an extremely thermophilic bacterium in *Escherichia coli*. Mol. Gen. Genet. **210**:490–497.
9. **Fukusaki, E., W. Panbangred, A. Shinmyo, and H. Okada.** 1984. The complete nucleotide sequence of the xylanase gene (*xynA*) of *Bacillus pumilus*. FEBS Lett. **171**:197–201.
10. **Gilbert, H. J., D. A. Sullivan, G. Jenkins, L. E. Kellett, N. P. Minton, and J. Hall.** 1988. Molecular cloning of multiple xylanase genes from *Pseudomonas fluorescens* subsp. *cellulosa*. J. Gen. Microbiol. **134**:3239–3247.
11. **Gold, L., and G. D. Stormo.** 1987. Translation initiation, p. 1302–1307. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
12. **Grépinet, O., M.-C. Chebrou, and P. Béguin.** 1988. Purification of *Clostridium thermocellum* xylanase Z expressed in *Escherichia coli* and identification of the corresponding product in the

culture medium of *C. thermocellum*. J. Bacteriol. **170**:4576–4581.
13. **Grépinet, O., M.-C. Chebrou, and P. Béguin.** 1988. Nucleotide sequence and deletion analysis of the xylanase gene (*xynZ*) of *Clostridium thermocellum*. J. Bacteriol. **170**:4582–4588.
14. **Hall, J., G. P. Hazlewood, P. J. Barker, and H. J. Gilbert.** 1988. Conserved reiterated domains in *Clostridium thermocellum* endoglucanases are not essential for catalytic activity. Gene **69**:30–38.
15. **Hamamoto, T., H.Honda, T. Kudo, and K. Horikoshi.** 1987. Nucleotide sequence of the xylanase gene of alkalophilic *Bacillus* sp. strain C125. Agric. Biol. Chem. **51**:953–955.
16. **Hazlewood, G. P., M. P. M. Romaniec, K. Davidson, O. Grépinet, P. Béguin, J. Millet, O. Raynaud, and J.-P. Aubert.** 1988. A catalogue of *Clostridium thermocellum* endoglucanase, β-glucosidase and xylanase genes cloned in *Escherichia coli*. FEMS Microbiol. Lett. **51**:231–236.
17. **Hoopes, B. C., and W. R. McClure.** 1987. Strategies in regulation of transcription initiation, p. 1231–1240. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
18. **Johnson, W. J., C. P. Moran, Jr., and R. Losick.** 1983. Two RNA polymerase sigma factors from *Bacillus subtilis* discriminate between overlapping promoters for a developmentally regulated gene. Nature (London) **302**:800–804.
19. **Kantelinen, A., M. Rättö, J. Sundquist, M. Ranua, L. Viikari, and M. Linko.** 1988. Hemicellulases and their potential role in bleaching, p. 1–4. *In* T. J. de Salvo (ed.), International Pulp Bleaching Conference, Orlando, Fla. TAPPI Proceedings. Technial Association of the Pulp and Paper Industry, Atlanta.
20. **Knowles, J., P. Lehtovaara, and T. Teeri.** 1987. Cellulase families and their genes. Trends Biotechnol. **5**:255–261.
21. **Lachke, A. H.** 1988. 1,4-β-D-Xylan hydrolase of *Sclerotium rolfsii*. Methods Enzymol. **160**:679–684.
22. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
23. **Lee, S. F., C. W. Forsberg, and J. B. Rattray.** 1987. Purification and characterization of two endoxylanases from *Clostridium acetobutylicum* ATCC 824. Appl. Environ. Microbiol. **53**:644–650.
24. **Love, D. R., and M. Streiff.** 1987. Molecular cloning of a β-glucosidase gene from an extremely thermophilic anaerobe in *E. coli* and *B. subtilis*. Bio/Technology **5**:384–387.
25. **MacKay, R. M., A. Lo, G. Willick, M. Zudier, S. Baird, M. Dove, F. Moranelli, and V. Seligy.** 1986. Structure of a *Bacillus subtilis* endo-β-1,4-glucanase gene. Nucleic Acids Res. **14**:9159–9170.
26. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
27. **McLaughlin, J. R., C. L. Murray, and J. C. Rabinowitz.** 1981. Unique features in the ribosome binding site sequence of the gram-positive *Staphylococcus aureus* β-lactamase gene. J. Biol. Chem. **256**:11283–11291.
28. **Mézes, P. S. F., Y. Q. Yang, M. Hussain, and J. O. Lampen.** 1983. *Bacillus cereus* 569/H β-lactamase. I. Cloning in *Escherichia coli* and signal sequence determination. FEBS Lett. **161**:195–200.
29. **Moran, C. P., Jr., N. Lang, S. F. J. LeGrice, G. Lee, M. Stephens, A. L. Sonnenschein, J. Pero, and R. Losick.** 1982. Nucleotide sequences that signal the initiation of transcription and translation in *Bacillus subtilis*. Mol. Gen. Genet. **186**:339–346.
30. **Moriyama, H., E. Fukusaki, J. Cabrera Crespo, A. Shinmyo, and H. Okada.** 1987. Structure and expression of genes coding for xylan-degrading enzymes of *Bacillus pumilus*. Eur. J. Biochem. **166**:539–545.
31. **Nielsen, B. K., and J. O. Lampen.** 1982. Membrane-bound penicillinases in gram-positive bacteria. J. Biol. Chem. **257**:4490–4495.
32. **O'Neill, G. O., S. H. Goh, R. A. J. Warren, D. G. Kilburn, and R. C. Miller.** 1986. Structure of the gene encoding the exoglucanase of *Cellulomonas fimi*. Gene **44**:325–330.
33. **Paice, M. G., R. Bernier, Jr., and L. Jurasek.** 1988. Viscosity-enhancing bleaching of hardwood kraft pulp with xylanase from a cloned gene. Biotechnol. Bioeng. **32**:235–239.
34. **Paice, M. G., R. Bourbonnais, M. Desrochers, L. Jurasek, and M. Yaguchi.** 1986. A xylanase gene from *Bacillus subtilis*: nucleotide sequence and comparison with *B. pumilus* gene. Arch. Microbiol. **144**:201–206.
35. **Patchet, M. L., T. L. Neal, L. R. Schofield, R. C. Strange, R. M. Daniel, and W. H. Morgan.** 1989. Heat treatment purification of thermostable cellulase and hemicellulase enzymes expressed in *Escherichia coli*. Enzyme Microb. Technol. **11**:113–115.
36. **Poutanen, K., and J. Puls.** 1988. Characteristics of *Trichoderma reesei* β-xylosidase and its use in the hydrolysis of soluble xylans. Appl. Microbiol. Biotechnol. **28**:425–432.
37. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
38. **Saul, D. J., L. C. Williams, D. R. Love, L. W. Chamley, and P. L. Bergquist.** 1989. Nucleotide sequence of a gene from *Caldocellum saccharolyticum* encoding for exocellulase and endocellulase activity. Nucleic Acids Res. **17**:439.
39. **Short, J. M., J. M. Fernandez, J. A. Sarge, and W. D. Huse.** 1988. λZAB: A bacteriophage λ expression vector with *in vivo* excision properties. Nucleic Acids Res. **16**:7583–7600.
40. **Sipat, A., K. A. Taylor, R. Y. C. Lo, C. W. Forsberg, and P. J. Krell.** 1987. Molecular cloning of a xylanase gene from *Bacteriodes succinogenes* and its expression in *Escherichia coli*. Appl. Environ. Microbiol. **53**:477–481.
41. **Stalker, D. M., K. E. McBride, and L. D. Malyj.** 1988. Herbicide resistance in transgenic plants expressing a bacterial detoxification gene. Science **242**:419–423.
42. **Teather, R. M., and P. J. Wood.** 1982. Use of Congo red-polysaccharide interactions in enumeration and characterization of cellulolytic bacteria from the bovine rumen. Appl. Environ. Microbiol. **43**:777–780.
43. **van Tilbeurgh, H., M. Claeyssens, and C. K. de Bruyne.** 1982. The use of 4-methylumbelliferyl and other chromophoric glycosides in the study of cellulolytic enzymes. FEBS Lett. **149**:152–156.
44. **Wain-Hobson, S., P. Sonigo, O. Danos, S. Cole, and M. Alison.** 1985. Nucleotide sequence of the AIDS virus. Cell **40**:9–17.
45. **Watson, M. E. E.** 1984. Compilation of published signal sequences. Nucleic Acids Res. **12**:5145–5164.
46. **West, C. A., A. Elzanowski, L. S. Yeh, and W. C. Barber.** 1989. Homologues of catalytic domains of *Cellulomonas* glucanases found in fungal and *Bacillus* glycosidases. FEMS Microbiol. Lett. **59**:167–172.
47. **Wong, K. K. Y., L. U. L. Tan, and J. N. Saddler.** 1988. Multiplicity of β-1,4-xylanases in microorganisms: functions and applications. Microbiol. Rev. **52**:305–317.
48. **Yang, R. C. A., C. R. MacKenzie, and S. A. Narang.** 1988. Nucleotide sequence of a *Bacillus circulans* xylanase gene. Nucleic Acids Res. **16**:7187.
49. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequence of the M13mp18 and pUC19 vectors. Gene **33**:103–119.