Journal of Bacteriology, Aug. 1990, p. 4247–4254
0021-9193/90/084247-08$02.00/0
Copyright © 1990, American Society for Microbiology
Vol. 172, No. 8

Cloning, Sequencing, and Expression of a Xylanase Gene from the Anaerobic Ruminal Bacterium *Butyrivibrio fibrisolvens*

BRUNO M. MANNARELLI,* STEVEN EVANS, AND DAVID LEE

Northern Regional Research Center, Agricultural Research Service, Peoria, Illinois 61604

Received 31 July 1989/Accepted 4 May 1990

A gene coding for xylanase activity, *xynA*, from the anaerobic ruminal bacterium *Butyrivibrio fibrisolvens* 49 was cloned into *Escherichia coli* JM83 by using plasmid pUC19. The gene was located on a 2.3-kilobase (kb) DNA insert composed of two adjacent *Eco*RI fragments of 1.65 and 0.65 kb. Expression of xylanase activity required parts of both *Eco*RI segments. In *E. coli*, the cloned xylanase enzyme was not secreted and remained cell associated. The enzyme exhibited no arabinosidase, cellulase, α-glucosidase, or xylosidase activity. The isoelectric point of the cloned protein was approximately 9.8, and optimal xylanase activity was obtained at pH 5.4. The nucleotide sequence of the 1,535-base-pair *Eco*RV-*Eco*RI segment from the *B. fibrisolvens* chromosome that included the *xynA* gene was determined. An open reading frame was found that encoded a 411-amino-acid-residue polypeptide of 46,664 daltons. A putative ribosome-binding site, promoter, and leader sequence were identified. Comparison of the XynA protein sequence with that of the XynA protein from alkalophilic *Bacillus* sp. strain C-125 revealed considerable homology, with 37% identical residues or conservative changes. The presence of the cloned xylanase gene in other strains of *Butyrivibrio* was examined by Southern hybridization. The cloned xylanase gene hybridized strongly to chromosomal sequences in only two of five closely related strains.

Hemicellulose is a major constituent of plant material ingested by browsing or grazing animals and can constitute 30 to 40% of the total carbohydrate (10). Hemicellulose is a heterogeneous mixture of pentose-containing polymers. The predominant polymer is D-xylan, a β(1-4)-linked polymer of xylose with side chains of arabinose and other sugars. In ruminant animals, plant material is degraded by the action of anaerobic bacteria in a large compartment of the intestinal tract called the rumen, located just before the main stomach. Bacteria in the rumen produce xylanases that hydrolyze xylan to utilizable oligosaccharides. In ruminants, xylans represent a significant source of nutritional energy.

The bacterial population of the rumen is a complex ecosystem composed of many different species of bacteria, but it has about 15 to 20 predominant species (17). Members of the species *Butyrivibrio fibrisolvens* are among the most numerous types of bacteria found in the rumen (7, 17) and are one of the most important bacterial species involved in the digestion of xylans (7, 11, 16). Most *B. fibrisolvens* strains are highly xylanolytic and can extensively degrade xylans. *B. fibrisolvens* strain 49 (7) is one of the most xylanolytic strain tested, having both high xylanase activity and the capacity to degrade various types of xylans (16).

In this study we describe the cloning of a gene coding for xylanase activity, *xynA*, from *B. fibrisolvens* 49 into *Escherichia coli*. We examined the expression and location of the cloned xylanase in *E. coli* and present some characteristics of the enzyme. The nucleotide sequence of the *xynA* gene was determined, and the derived encoded polypeptide was compared with the XynA protein from alkalophilic *Bacillus* sp. strain C-125. We also hybridized the gene to homologous chromosomal sequences found in other strains of *Butyrivibrio*.

* Corresponding author.

MATERIALS AND METHODS

Bacterial strains, plasmids, and growth conditions. *B. fibrisolvens* 49 was used as the source of chromosomal DNA. The sources and DNA relatedness of the *B. fibrisolvens* strains used were summarized in a previous study (20). *B. fibrisolvens* strains were grown in RGM medium (20) under strict anaerobic conditions. *E. coli* strains JM83 (38) and HB101 (4) were used as the recipient strains for recombinant plasmids. *E. coli* cells were grown on LB medium (19) or LBA (LB supplemented with ampicillin [75 μg/ml]) for selecting transformants. Bacto-Agar (Difco Laboratories) was included at a final concentration of 2.0% (wt/vol) for agar plates. Plasmids pBR322 and pUC19 were used as the cloning vectors.

Isolation and cloning of DNA. High-molecular-weight chromosomal DNA from *B. fibrisolvens* strains was purified as described previously (20). The DNA was partially digested with *Eco*RI and fractionated on a linear 10 to 40% sucrose density gradient. Fractions containing DNA fragments between 2 and 20 kilobases (kb) long were pooled and ligated into the dephosphorylated *Eco*RI site of pUC19. The ligation mixture was transformed into competent *E. coli* JM83, and transformants were screened on LBA agar plates containing 0.1% 5-bromo-4-chloro-3-indol-β-D-galactopyranoside (X-gal). The resulting white colonies, indicating clones carrying pUC19 with genomic inserts, were picked onto LBA agar plates containing 0.15% Remazol brilliant blue (RBB)–xylan (Sigma Chemical Co.). Colonies producing clear halos indicated putative clones containing hybrid plasmids with genomic inserts coding for xylanase activity. Transformants containing only pUC19 never produced clear halos on RBB-xylan plates.

Characterization of insert DNA. Restriction enzyme cleavage maps were constructed by using single and multiple digests of hybrid plasmids. Restriction enzymes were purchased from Bethesda Research Laboratories and were used as specified by the supplier. Southern hybridization was a modification of the Southern procedure as described by

Thomashow et al. (34), and nick translation of probes was performed with [^{32}P]dATP (NEN Research Products) (19). Arabinosidase, cellulase, α-glucosidase, β-glucosidase, and xylosidase activities of the clones were screened on agar plates with methylumbelliferyl derivatives (Sigma Chemical Co.). LBA agar plates containing either 4-methylumbelliferyl-α-L-arabinoside, 4-methylumbelliferyl-β-D-cellobioside, 4-methylumbelliferyl-α-D-glucoside, 4-methylumbelliferyl-β-D-glucoside, or 4-methylumbelliferyl-β-D-xyloside at 20 μg/ml were prepared, and transformants were streaked directly on these plates. Enzyme activity was detected by the release of 4-methylumbelliferone, which fluoresces when examined under a hand-held UV light (365 nm).

Preparation of cell extracts and periplasmic proteins. Cell extracts were prepared from 60-ml cultures of *E. coli* in mid-log phase. Cells were suspended in 5.0 ml of 0.05 M sodium phosphate buffer, pH 6.2, and disrupted by sonication at 4°C (10-s bursts for a total of 90 s) with a 4710 series Ultrasonic Homogenizer (Cole Palmer Instrument Co.). The extract was clarified by centrifugation at 10,000 × *g* for 10 min at 4°C. The extracellular, periplasmic, and cellular xylanase activities in *E. coli* strains were determined by the osmotic shock procedure of Cornelis et al. (9). Cyclic phosphodiesterase and β-galactosidase were assayed as markers for periplasmic and cellular proteins, respectively, as described previously (22).

Enzyme assays. Xylanase activity was determined by monitoring the increase in reducing sugar formation from larchwood xylan (Sigma Chemical Co.) with the dinitrosalicylic acid reagent (21). Reaction mixtures contained 60 μl of larchwood xylan (5.0%), 540 μl of 50 mM $NaPO_4$ buffer (pH 6.2), and 400 μl of cell extract. Cell protein was determined by the method of Bradford (5).

Analytical isoelectric focusing and detection of xylanase activity in gels. Vertical polyacrylamide gel isoelectric focusing was performed with 0.75-mm slab gels (7.5% [wt/vol] total acrylamide and 2.7% [wt/vol] *N,N'*-methylene-bisacrylamide) containing 10% (vol/vol) pH 3 to 11 ampholines (Pharmacia) and 10% (vol/vol) glycerol. The anodic buffer was 0.01 M L-aspartate, and the cathodic buffer was 0.01 M β-alanine. Focusing was conducted at 10°C at 5 W (constant power) for 1 h and then at 2,000 V (constant voltage) for 4 h. The pH gradient was estimated by the use of colored isoelectric point markers (pI range, 4.1 to 10.7 at 0 to 2°C; Cal-Biochem). Xylanase activity was determined by the overlay method of Biely et al. (3) with a 2.0% agar overlay containing 0.2% RBB–xylan in 50 mM sodium phosphate buffer, pH 6.2. Concentrated supernatants were obtained from cell-free supernatants of overnight cultures of *B. fibrisolvens* grown in RGM medium. Solid ammonium sulfate was added to 50%, the solution was chilled at 4°C for 30 min and centrifuged at 10,000 × *g* for 10 min, and the precipitate was suspended in 1/10 volume of 50 mM sodium phosphate buffer, pH 6.2.

Nucleotide sequence determination and protein sequence comparison. Restriction fragments from pML110 were cloned into M13mp18 and M13mp19, and single-stranded M13 template DNA was prepared as described previously (1). Large segments of insert DNA from pML110 without known restriction sites were divided into smaller fragments for sequencing by exonuclease III digestion (15). DNA sequencing was performed by the method of Sanger et al. (28) with fluorescent primers. DNA fragments were analyzed on a model 370A automated DNA sequencing system (Applied Biosystems). Amino acid sequences were compared by using the Microgenie alignment program (Beckman Instruments). Homologous amino acids were defined as either identical amino acids or replacements corresponding to conservative amino acid changes. A conservative change is when both amino acids belong to one of the following groups: alanine, serine, and threonine; asparagine and glutamine; aspartic acid and glutamic acid; isoleucine, leucine, methionine, and valine; arginine, histidine, and lysine; or phenylalanine, tryptophan, and tyrosine.

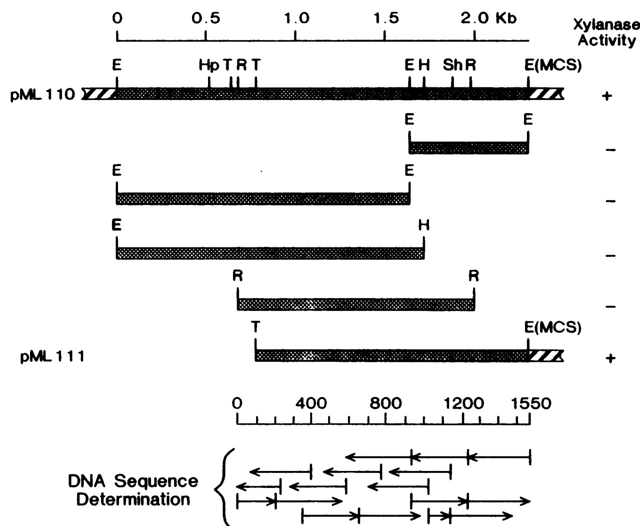

FIG. 1. Restriction endonuclease map of pML110 and subclones and DNA sequencing strategy. pML110 and derivatives were constructed by using pUC19 as the vector. The xylanase activity of the DNA fragments inserted in pUC19 is shown to the right. Striped boxes, pUC19 sequences; cross-hatched boxes, *B. fibrisolvens* DNA sequences. Abbreviations: E, *Eco*RI; R, *Eco*RV; Sh, *Sph*I; H, *Hin*dIII; T, *Taq*I; Hp, *Hpa*I; MCS, multiple cloning site of pUC19. Segments of which nucleotide sequence was determined are indicated by arrows below the restriction map, which indicate the extent and direction of each analysis.

Dendrogram. The dendrogram of *B. fibrisolvens* strain relationships was derived from DNA reassociation values comparing all 13 *B. fibrisolvens* strains examined. DNA reassociation values were determined by the spectrophotometric method and were reported in a previous study (20). The percent relatedness values used to construct the dendrogram were obtained from the DNA reassociation values by using the unweighted pair-group average algorithm (33). Values were calculated with an IBM AT computer and the TAXAN program of David Swartz, University of Maryland, College Park, Md.

RESULTS

Cloning of a xylanase gene from *B. fibrisolvens* in *E. coli*. Approximately 1,500 white colonies resulting from transformed *E. coli* JM83 cells plated onto X-gal plates were picked onto RBB-xylan plates. Two colonies were detected which produced clear halos, indicating the presence of xylanase activity. Both clones were purified by serial streaking onto RBB-xylan plates, and both were stable for xylanase activity. Plasmids isolated from the two clones both contained 2.3-kb DNA inserts with identical restriction maps and oriented in the same direction. The 2.3-kb DNA inserts were composed of two *Eco*RI fragments of 1.65 and 0.65 kb. This plasmid was designated pML110.

Restriction mapping and subcloning. A restriction map of

pML110 and subclones is presented in Fig. 1. The xylanase gene from pML110 was further localized by subcloning fragments from the 2.3-kb insert into pUC19. Neither the 1.65- nor the 0.65-kb *Eco*RI fragment in pUC19 encoded xylanase activity, indicating that the xylanase gene required parts of both *Eco*RI DNA fragments. The internal 1.25-kb *Eco*RV fragment and the 1.7-kb *Eco*RI-*Hin*dIII fragment also did not produce xylanase activity. The 1.45-kb *Eco*RI-*Taq*I fragment did produce xylanase activity. This plasmid was designated pML111 and was constructed by isolating the 1.45-kb *Taq*I fragment from pML110 (in pUC19, a *Taq*I site overlaps the *Eco*RI site), filling in the recessed ends with Klenow fragment, and ligating it to the *Sma*I site of pUC19 by blunt-end ligation. Thus, one end of the xylanase gene is located to the right of the second *Taq*I site of the 1.65-kb *Eco*RI fragment, and the other end is located between the *Eco*RV and *Eco*RI sites of the 0.65-kb *Eco*RI fragment. Twenty other restriction endonucleases, including *Acc*I, *Bam*HI, *Bgl*I, and *Pst*I, did not cut the insert DNA. The *xynA* gene appeared to be transcribed from an endogenous promoter. When either pML110 or pML111 was moved to *E. coli* TG1 [Δ(*lac-pro*) *supE*(F' *traD36 lacI^q lacZΔM15*) (8)], which requires induction for the regulated *lac* promoter, equal clearing of RBB-xylan plates was observed with or without the inducer IPTG (isopropyl-β-D-thiogalactopyranoside).

Expression and localization of xylanase activity in *E. coli*. Similar levels of xylanase activity were expressed by *E. coli* strains JM83 and HB101 containing hybrid plasmids. Strain JM83 containing pML110 or pML111 expressed 14.2 and 18.1 nmol of reduced sugar per min per mg of protein, respectively. Comparable xylanase levels were also observed for *E. coli* strains containing hybrid plasmids composed of the insert DNA from pML110 moved to the *Eco*RI site of pBR322. Strains JM83 and HB101 containing only pUC19 or pBR322 expressed less than 0.1 nmol of reduced sugar per min per mg of protein. The cellular location of xylanase activity in *E. coli* strains containing the hybrid plasmids was also examined. Xylanase activity was found almost exclusively associated with the cells. Most of the xylanase activity (93.2 and 95%) was detected in the cytoplasmic fraction for strains JM83 and HB101, respectively. Low xylanase activity was detected in the culture medium (5 and 5%, respectively) and in the periplasmic fraction (2.8 and 0%, respectively) for strains JM83 and HB101. Control experiments showed that 96% of the β-galactosidase activity (an intracellular enzyme) was located in the cytoplasmic fraction and 85% of the cyclic phosphodiesterase activity (a periplasmic enzyme) was present in the periplasmic fraction.

The xylanase enzyme displayed a broad pH range from 4.5 to 7.5, with optimal activity at pH 5.4. Xylanase activity was not affected by disrupting cells in buffer containing the divalent cation Mg^{2+} or Ca^{2+}, and the enzyme retained appreciable activity when cells were disrupted in distilled water. Specific activity was also not affected when *E. coli* cells harboring hybrid plasmids were grown in minimal medium with various carbohydrates as the carbon source. Cells grown in the presence of arabinose, galactose, glucose, lactose, maltose, mannitol, mannose, sorbitol, sucrose, xylose, or xylan exhibited similar activities. By plate assay, arabinosidase, cellulase, α-glucosidase, or xylosidase activities were not detected in any of the recombinant transformants.

Zymogram of xylanase enzyme. A xylanase activity zymogram of proteins separated by isoelectric focusing is shown in Fig. 2. Two bands of xylanase activity (pI, ca. 5.2 and 8.5) were detected in concentrated cell-free supernatants of *B. fibrisolvens* 49 grown on RGM medium with arabinose as the carbon source. A single band of xylanase activity (pI, ca. 9.8) was detected in cell extracts of *E. coli* JM83(pML111).

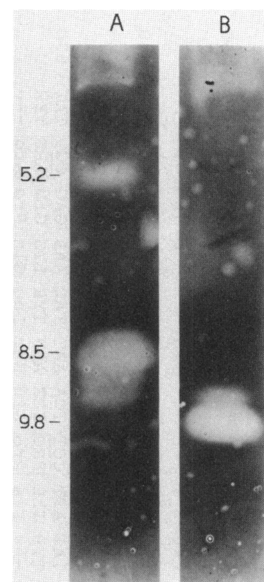

FIG. 2. Zymogram of xylanase activities after isoelectric focusing. (A) 25 μl of 1/10 dilution of concentrated supernatant from an overnight culture of *B. fibrisolvens* 49. (B) 25 μl of cell extract of *E. coli* JM83(pML111). The pIs of the bands were derived from colored pI markers and are indicated.

Nucleotide sequence determination. The nucleotide sequence of both strands of the 1,535-base-pair (bp) *Eco*RV-*Eco*RI fragment from pML110 was determined by using an automated sequencing system from Applied Biosystems (Fig. 3). The sequencing strategy is shown in Fig. 1. Reading right to left, no open reading frame was found that was greater than 200 bp in length. A single large open reading frame was found reading from left to right starting at base 241 and ending at base 1,521. A Shine-Dalgarno sequence, AAAGGAG (30), was found downstream of the start of the open reading frame beginning at base 274. The first start codon was found 8 bases downstream beginning at base 289 and is indicated as the probable start of the prexylanase gene, since it is the only start codon associated with a strong ribosome-binding site. The prexylanase gene, then, would encode a polypeptide of 411 amino acid residues or 46,664 daltons. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis analysis of crude extracts of JM83(pUC19) and JM83(pML110) revealed that pML110 produced a single new protein band slightly larger than 45,000 daltons (data not shown), which is in good agreement with the calculated molecular mass of the prexylanase protein.

In *B. fibrisolvens*, the xylanase enzyme is secreted, and the corresponding gene would be expected to contain a leader sequence. The N-terminal amino acid sequence of the prexylanase revealed a 33-residue segment that strongly resembles the general structure of leader sequences (13). The first 12 amino acids contained four lysine and two arginine residues and could form the positively charged "n region." The next 14 amino acids were composed of nonpolar and uncharged polar residues and could form the hydrophobic core "h region." The last seven amino acids conformed to the pattern of residues near signal sequence cleavage sites and form the "c region." By the rules of von

```
            8                   7                                                                    6
GAT CAA AAT ATT ATG TTA ATT ATG ATG AGT ACG AGA CCT TTG TTA GAA ACG CAG GAG GAA TAC TTA ATC CCT TAA ATT TGG TTA ATG CAT   90
                5                                                                               4
ATT GCT CGA ATA TTT TAA TTA TAA AAA AGC GAT ACA ACT TTT AGG TAT TCA AAA GCA ATT TTT GAT AAG AGT AAC TTT TTG GGG GTT GTT  180
     3                      2           1
AGA ATA TAA TTG CAC GAA CAA TAT AAT AAT GCG TAA TAT CGG ACA TCT GCG GCG AAA TAG TCC GAA TAT GCA TTA AGT ATG GCA ACA ATT  270
          -35                            -10
CAC AAA AGG AGA AGA TGT ATG AAA AAA TTC AAA ATC AGA AAA CTG ATG GCA AGA GTA CTG GCA CTT GCC TTA GTT TTC AGC ACA TTT TTT  360
      RBS               MET LYS LYS PHE LYS ILE ARG LYS LEU MET ALA ARG VAL LEU ALA LEU ALA LEU VAL PHE SER THR PHE PHE
                        Met Ile Thr Leu --- Phe Arg Lys Pro Phe Val Ala Gly Leu Ala Ile Ser Leu Leu Val Gly Gly --- ---

                                    ↓
ATG GTA TCT AAG GTA GAT GCC AAC GCT GCC AGC TAC AAT TTA ATG GAG ACC TAT GGA GCA AAA TAT GGG TAC TCT GGA AAC TGT GTG CAT  450
MET VAL SER LYS VAL ASP ALA ASN ALA ALA SER TYR ASN LEU MET GLU THR TYR GLY ALA LYS TYR GLY TYR SER GLY ASN CYS VAL HIS
Gly Ile Gly Asn Val --- --- Ala Ala Ala Gln Gly Gly Pro Pro Lys Ser Phe Gly Glu Asn --- Glu Lys Arg Asn Asp Gln Pro Phe

ACA CAC ATG CTT AGG GAT TCA AGA ATT GTA AAT GCA ATC AAA AAG GAT TCC AAT ATA GTA ACA CTT GGT AAT GAA ATG AAG CCT GAT TAC  540
THR HIS MET LEU ARG ASP SER ARG ILE VAL ASN ALA ILE LYS LYS ASP SER ASN ILE VAL THR LEU GLY ASN GLU MET LYS PRO ASP TYR
Ala Trp Gln Val --- Ala Ser --- Leu --- Ser Glu --- Arg Tyr Gln Glu Gln Phe --- Asp Ile Gly Ala Ala Val Glu Pro --- Tyr

CTT TTG GGA AGC AGA CAG GCA ACT CTT ATT TCG GTT GAT GAA GCA AAG CGT CTT GGT TAC TAT ATT CCT TCA AAT TAC AAA GAG CGT TAT  630
LEU LEU GLY SER ARG GLN ALA THR LEU ILE SER VAL ASP GLU ALA LYS ARG LEU GLY TYR TYR ILE PRO SER ASN TYR LYS GLU ARG TYR
Gln Leu Glu Gly Arg Gln Ala Gln Ile Leu Lys His His --- Ser Leu Val Ala Glu Asn --- Lys Pro Glu Ser Leu Gln Pro Arg ---

GTC CCA AAA ATT GAT TTT AGA ACA GTT GAT GAA GCA GTA AAA ATT TGC TAT GAG AAT GGT CTT AAG ATG AGA GGT CAT ACT CTT GTA TGG  720
VAL PRO LYS ILE ASP PHE ARG THR VAL ASP GLU ALA VAL LYS ILE CYS TYR GLU ASN GLY LEU LYS MET ARG GLY HIS THR LEU VAL TRP
Glu Gly Glu Trp Asn Trp Glu Gly Ala Asp Lys Ile Val Glu Phe Ala Arg Lys His Asn Met Glu Leu Arg Phe His Thr Leu Val Trp

CAT TCA CAG ACC CCA ACC TGG CTA TTC AGA GAG AAT TAT TCA GGA AAT GGC AGA TTT GTA AAT ACA GCT ACT ATG GAT GCT CGT CTT GAG  810
HIS SER GLN THR PRO THR TRP LEU PHE ARG GLU ASN TYR SER GLY ASN GLY ARG PHE VAL ASN THR ALA THR MET ASP ALA ARG LEU GLU
His Ser Gln Val Pro Glu Trp Phe Phe Ile Asp Glu --- Asp Gly Asn --- Arg Met Val Asp Glu Thr Asp Pro Asp Lys Arg Glu ---

TTT TAT GTT AAG TCT GTA ATG GGA CAT TTC TAT TCA GGC AAA TAC GGT AGC ACA CTG GTA TAT TGG GAT GTG TGC AAT GAG ACA CTT CAT  900
PHE TYR VAL LYS SER VAL MET GLY HIS PHE TYR SER GLY LYS TYR GLY SER THR LEU VAL TYR TRP ASP VAL CYS ASN GLU THR LEU HIS
Leu Leu Leu Glu Arg Met Glu Asn His Ile Lys Thr Glu Arg Tyr Lys Asp Asp Val Thr Ser Trp Asp Val Val Asn Glu Val Ile Asp

GCT CAG AAT TCT GGT TGG GAA GCT GTT TAC GGA AGC AAC AAG ACA AAC GCA GTA TAT GTA AAG AAA GCT TTT AAT TAC GCA TAT CAG GTA  990
ALA GLN ASN SER GLY TRP GLU ALA VAL TYR GLY SER ASN LYS THR ASN ALA VAL TYR VAL LYS LYS ALA PHE ASN TYR ALA TYR GLN VAL
Asp Gly Gly Gly Leu Arg Glu Ser Glu Trp --- Tyr Gln Ile Thr Gly Thr Asp Tyr Ile Lys Val Ala Phe Glu --- Thr --- Ala ---

CTT GAG CAG TAC AAG CTT ACA AAT TCC GTT AAG CTT TTC TAC AAC GAC TAC AAT ACA TAC ATG GAA GTC AAT GAT GTG ATA AAA TTA GTA 1080
LEU GLU GLN TYR LYS LEU THR ASN SER VAL LYS LEU PHE TYR ASN ASP TYR ASN THR TYR MET GLU VAL ASN ASP VAL ILE LYS LEU VAL
--- Arg Lys Tyr Gly --- Gly Glu Glu Ala Lys Leu Tyr Ile Asn Asp Tyr Asn Thr Glu Val Pro Ser --- Asp Asp Leu --- Leu Val

AAC TAC ATC AAT CAG GGG AAG AAG GTT TGT GCA GGT GTA GGC ATG CAG TCA CAT CTT GGA ACA GGC TTT CCT TCA GTA GAT TAT TAT ACA 1170
ASN TYR ILE ASN GLN GLY LYS LYS VAL CYS ALA GLY VAL GLY MET GLN SER HIS LEU GLY THR GLY PHE PRO SER VAL ASP TYR TYR THR
Lys --- Leu Glu Gln Gly Val Pro Ile --- Asp Gly Val Gly His Gln Ser His Ile Gln Ile Gly Trp Pro Ser Ile Glu Asp Thr Arg

AAT GCA CTT AAC TCA TTC CTT CGT GCA GGA TTT GAA GTG CAG ATC ACA GAG CTT GAT ATC ACA AAC AAA GGC GAT TAT GAT TTG AAC AAT 1260
ASN ALA LEU ASN SER PHE LEU ARG ALA GLY PHE GLU VAL GLN ILE THR GLU LEU ASP ILE THR ASN LYS GLY ASP TYR ASP LEU ASN ASN
Ala Ser Phe Glu Lys Phe Thr Ser Leu Gly Leu Asp Asn Gln Val Thr Glu Leu Asp Met Ser Leu Tyr Gly --- Trp Pro Pro Thr Gly

TAC GCA TAT CGC CTG TTT AAA AAT ATT AAC GCC GCA AAA AAG AAT GGT GGA AAC ATT TCA TGT ATC ACA TGG TGG GGA CCA TCA GAT GCA 1350
TYR ALA TYR ARG LEU PHE LYS ASN ILE ASN ALA ALA LYS LYS ASN GLY GLY ASN ILE SER CYS ILE THR TRP TRP GLY PRO SER ASP ALA
--- Ala Tyr Thr Ser Tyr Asp Asp Ile Pro Ala Glu Leu Leu Gln Ala Gln Ala Asp Arg --- Gln Leu Phe Tyr Glu Gln Leu Ala Ala

GAG ACA TGG ATT AGA AAT GAA AAG CCA CTT ATC TGG TCA AAT ATT GGT GTT GCA AAG CCA GCC TAT GAT GAG GTT GTA AAG GCA TTT ACA 1440
GLU THR TRP ILE ARG ASN GLU LYS PRO LEU ILE TRP SER ASN ILE GLY VAL ALA LYS PRO ALA TYR ASP GLU VAL VAL LYS ALA PHE THR
Asp --- --- Ile --- Ser Ser Val Thr Phe --- Trp Gly --- Ile Ala Asp Asn His --- Thr Trp Asp Gly Arg Ala Arg Glu Tyr Asn

GAG ACA TTT GGT AAT CCT GGC TCA TTT ACA CCA CAA CCA ACG ATT ACA CCA CAG CCA ACC CCT ACA CCA TCT GGA CAA ACG TAA AAT ACG 1530
GLU THR PHE GLY ASN PRO GLY SER PHE THR PRO GLN PRO THR ILE THR PRO GLN PRO THR PRO THR PRO SER GLY GLN THR ***
Asn Gly Val Gly Ile Asp Ala Pro Phe Val Phe Asp His Asn Tyr Arg Val Lys Pro Ala Tyr Trp Arg Ile --- Ile Asp

AAT TC 1535
```

Heijne (35) and Perlman and Halvorson (25), a probable cleavage site can be predicted and is indicated by an arrow in Fig. 3. Thus, the secreted xylanase enzyme would consist of 378 amino acids or 42,906 daltons. Preliminary electrophoretic data for the two xylanases secreted by *B. fibrisolvens* 49 indicate that one of the xylanases has a molecular mass of between 40,000 and 45,000 daltons (unpublished data).

The region upstream of the probable ribosome-binding site is A+T rich, having an A+T content of 68 mol%. Scanning this region for a putative promoter showed that there were at least eight possible Pribnow boxes (−10 regions). These regions are indicated as numbered boxes in Fig. 3. As determined from subcloning experiments, the promoter must be contained in the DNA fragment downstream of the *Taq*I (TCGA) site at base 96. Therefore, only the first four Pribnow boxes can be considered possible −10 regions. The −35 regions to the first four −10 regions all showed relatively weak homology to the consensus sequence of TTGACA in *E. coli* (27). Based on homology, the most likely −35 region is the one associated with the first Pribnow box and is underlined in Fig. 3. This is the first gene sequenced from the genus *Butyrivibrio*. A xylosidase gene (*xylB*) has been cloned from another *Butyrivibrio* strain, GS113 (29). It would be of interest to sequence the control region of this gene for comparison.

The nucleotide sequence of the *xynA* gene from alkalophilic *Bacillus* sp. strain C-125 has been determined (14). The XynA polypeptide from strain C-125 contains 396 amino acid residues and is similar in size to the XynA protein from *B. fibrisolvens*. Comparison of the proteins encoded by the two *xynA* genes revealed considerable homology (Fig. 3). When gaps were introduced into the *xynA* gene from strain C-125 to maximize correspondence, 37% of the residues were identical or corresponded to conservative changes. Particularly strong homology of 50%, whether identical residues or conservative changes, was found in the center regions of these polypeptides within amino acids 137 to 315. Correspondences of over 75% identical residues were found in amino acid regions 138 to 153, 196 to 201, 245 to 253, 276 to 283, and 308 to 313.

Southern hybridization analysis. To confirm the chromosomal origin of the DNA insert in pML110, Southern blots were performed on chromosomal and plasmid DNA digests with ^{32}P-labeled pML110 as the probe and a hybridization temperature of 70°C (Fig. 4). The two original clones of pML110 cut with *Eco*RI yielded three identical bands of equal intensity with sizes of 2.7, 1.65, and 0.65 kb (lanes 1 and 2). Chromosomal DNA from *B. fibrisolvens* 49 cut with *Eco*RI (lane 3) showed two hybridization bands of 1.65 and 0.65 kb, corresponding to the lower-molecular-weight bands of plasmid pML110 cut with *Eco*RI. Strain 49 chromosomal DNA cut with *Pst*I, *Bam*HI, *Acc*I, and *Bgl*I (lanes 4 to 7) produced single hybridization bands of 6.7, 24.0, 12.0, and 7.7 kb, respectively. Detectable hybridization was not observed with *E. coli* DNA digested with *Eco*RI or *Bam*HI (lanes 8 and 9).

The presence of the cloned xylanase gene or homologous genes in other strains of *Butyrivibrio* was examined. *Pst*I digests of chromosomal DNA from 13 different strains of *Butyrivibrio* were probed with the ^{32}P-labeled 0.65-kb *Eco*RI fragment of pML110 (Fig. 5). The two panels in Fig. 5 represent similar gels and blotting conditions except that the hybridizations were performed at different temperatures, 69°C (Fig. 5A) and 64°C (Fig. 5B). At both temperatures, the 0.65-kb *Eco*RI fragment produced dense bands only when hybridized with strains 49 (6.7 kb, lanes 2 and 16), H17c (6.7 kb, lanes 3 and 17), and CF3 (11.0 kb, lanes 7 and 21). A number of faint bands were seen when hybridization was performed under high stringency (Fig. 5A, lanes 2 to 7, 10, and 12), and these bands became more pronounced when the stringency of hybridization was lowered (Fig. 5B, lanes 16 to 21, 24, and 26). In strains 49, H17c, CE-51, CE-52, and 12, the faint bands were the same size, 7.5 kb. Faint bands of 19.0, 4.4, and 3.0 kb were observed in strains CF3, D30g, and D1, respectively. A dendrogram displaying the taxonomic relationships of the *Butyrivibrio* strains in Fig. 5 is shown in Fig. 6. The strains represent isolates from around the world and were obtained from various animals.

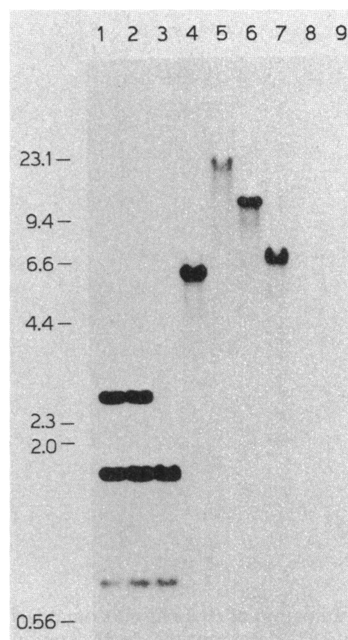

FIG. 4. Hybridization of plasmid and chromosomal DNA digests with pML110. Autoradiogram of DNA digests transferred to nitrocellulose and hybridized with ^{32}P-labeled pML110 at 70°C is shown. Lanes 1 and 2 contain 0.1 μg of the two original clones of pML110 cut with *Eco*RI. These lanes also contain 2.0 μg of sheared salmon sperm DNA. Lanes 3 to 7 contain 2.0 μg of *B. fibrisolvens* 49 chromosomal DNA digested with *Eco*RI, *Pst*I, *Bam*HI, *Acc*I, and *Bgl*I, respectively. Lanes 8 and 9 contain 2.0 μg of *E. coli* chromosomal DNA digested with *Eco*RI and *Bam*HI, respectively. Lambda DNA size markers are indicated (in kilobases).

FIG. 3. Nucleotide sequence of the *xynA* gene from *B. fibrisolvens* 49 and adjacent regions. The DNA strand reading from left to right in the orientation shown in Fig. 1 is depicted. The putative ribosome-binding site (RBS) and −35 promoter region are underlined. The eight possible Pribnow boxes are shown as numbered boxes, with Pribnow box 1 indicated as the putative −10 promoter region. The predicted amino acid sequence for the XynA protein of *B. fibrisolvens* is shown in uppercase letters, with the N-terminal methionine at base 289. The predicted amino acid sequence of the XynA protein of *Bacillus* sp. strain C-125 is shown in lowercase letters. Homologous amino acids between the two protein sequences are boxed. The predicted signal peptide cleavage site is shown with an arrow (↓).

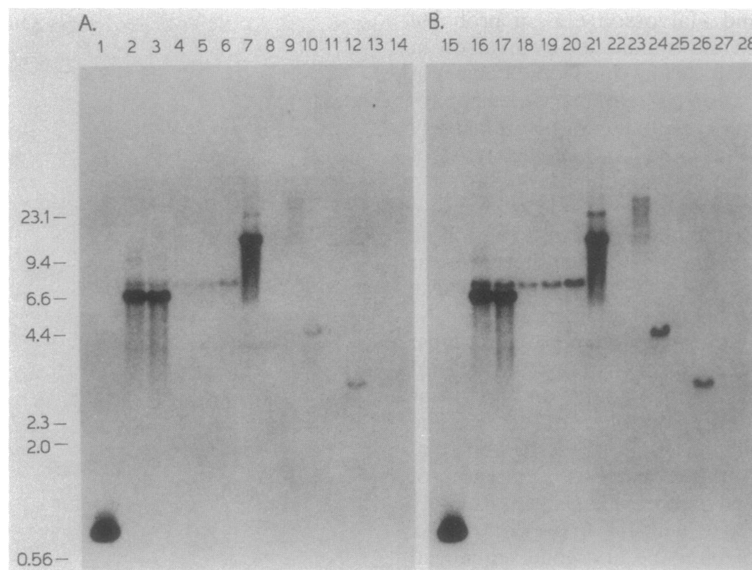

FIG. 5. Autoradiogram of *Pst*I chromosomal DNA digests of *Butyrivibrio* strains probed with ^{32}P-labeled 0.65-kb *Eco*RI fragment. Panels A and B are identical gels except that hybridization was performed at 69°C (A) and 64°C (B). Lanes 1 and 15 contain 0.1 μg of pML110 cut with *Eco*RI. Lanes 2 to 14 and 16 to 28 contain 2.0 μg of *Pst*I chromosomal DNA digests from *B. fibrisolvens* strains 49, H17c, CE-51, CE-52, 12, CF3, NOR-37, PI-7, D30g, C3, D1, E46a, and D16f, respectively. Lambda DNA size markers are indicated (in kilobases).

DISCUSSION

A 2.3-kb *Eco*RI DNA fragment coding for xylanase activity from *B. fibrisolvens* 49 was cloned into *E. coli* by using the vector pUC19. The 2.3-kb fragment was composed of two *Eco*RI segments, 1.65 and 0.65 kb. Since expression of xylanase activity required parts of both *Eco*RI segments, the segments must be adjacent on the *B. fibrisolvens* chromosome. pML110 hybridized to chromosomal DNA digests from *B. fibrisolvens* 49 but not to chromosomal digests of *E. coli*, indicating that the cloned fragment was derived from *B. fibrisolvens*. *Eco*RI-digested chromosomal DNA from strain 49 produced two bands of the sizes expected from the restriction map of pML110.

The cloned xylanase did not display arabinosidase, cellulase, α-glucosidase, or xylosidase activity. This is characteristic of true endoxylanases (26). In *B. fibrisolvens*, xylanase activity is constitutive, but higher levels are obtained when the cells are grown in the presence of xylan, and the enzyme is secreted into the medium (16). In *E. coli*, the specific activity of the cloned xylanase gene was not affected by growth in the presence of various sugars, including arabinose, xylose, and xylan. The XynA protein in the *E. coli* host does not cross the cell membrane and is associated with the cells. The location of other bacterial xylanases cloned into *E. coli* seems to depend on whether the donor bacterium is gram-positive or gram-negative. Cloned xylanase genes

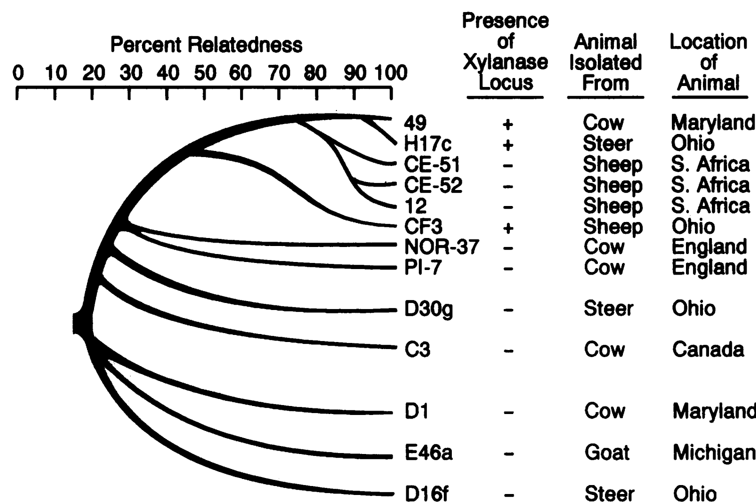

FIG. 6. Dendrogram of taxonomic relationships of *Butyrivibrio* strains. The dendrogram was constructed from the percent relatedness values derived from DNA reassociation values (20) by using an unweighted pair-group average algorithm. The strains were isolated from the gastrointestinal tracts of various animals. The animal source and location of the animal are indicated. Also shown is whether the xylanase locus is present in the strains examined.

from *Bacteroides succinogenes* (32) and *Aeromonas* sp. strain 212 (18), which are gram-negative bacteria, coded for enzymes located in the periplasm. Cloned xylanase genes from *Bacillus subtilis* (2), *Bacillus pumilus* (24), and *Clostridium acetobutylicum* (39), which are gram-positive bacteria, produced enzymes located in the cytoplasm.

As determined from DNA sequence analysis, the *xynA* gene encodes a prexylanase protein of 411 amino acids containing a 33-residue leader sequence. The length of leader sequences is usually between 15 and 30 residues. However, leader sequences between 31 and 44 residues long have been found in the genus *Bacillus* (36). The cloned xylanase appears to be one of two xylanases secreted by *B. fibrisolvens* 49. The size of the XynA protein was calculated at 378 residues or 42,906 daltons, which is similar to the molecular mass observed for one of the secreted xylanases of strain 49. The pI of the prexylanase XynA protein was 9.8. The pI of the processed XynA protein can be calculated from the amino acid sequence. By using the formula of Sillero and Ribeiro (31), the processed xylanase would have a pI of about 8.6, which is close to the actual pI of 8.5 observed for one of the secreted strain 49 xylanases.

Comparison of the XynA protein sequence of strain 49 with that of the protein from strain C-125 showed considerable homology. In the central portion of the enzyme, 50% homology was observed for a stretch of 178 amino acids. Within this long stretch, five smaller regions of amino acids showed correspondences of over 75% identical residues. Sequences of high homology suggest conserved regions that are likely to be important for enzyme function. The central portions of these xylanases may contain the catalytic sites of the enzymes. The clusters of identical amino acids indicate that the genes probably evolved from a common origin. It seems unlikely that a number of stretches of identical residues could result from convergent evolution. Three other procaryotic xylanases from *B. subtilis* (23), *B. pumilus* (12), and *Bacillus circulans* (37) have been sequenced. The proteins encoded by these genes are between 23,000 and 27,000 daltons and are much smaller than the XynA protein. Comparison of these proteins with the XynA protein showed some similarities in the amino acid sequences. However, direct comparison was not practical because of the large gaps that had to be introduced into the smaller proteins to maximize correspondence.

Southern hybridization of *Pst*I digests of chromosomal DNAs from other strains of *Butyrivibrio* probed with the 0.65-kb *Eco*RI fragment, which is composed mostly of xylanase gene sequence, showed that sequences of high homology are present in only a few strains of *Butyrivibrio*. Figure 5 shows strongly hybridizing bands and secondary light bands. The secondary bands disappeared as the temperature was raised from 64°C (Fig. 5B) to 69°C (Fig. 5A) and 70°C (Fig. 4, lane 4). The midpoint of the melting curves of chromosomal DNAs from the *Butyrivibrio* strains used was between 87 and 89°C (data not shown). Hybridization at 69 and 70°C represents rather stringent conditions, while 64°C represents close to an optimal reassociation temperature (6). Incubation temperatures above the optimum allow only highly complementary sequences to reassociate. The light bands must represent sequences with only partial homology to the cloned xylanase gene and could be explained by DNA segments coding for a second xylanase enzyme or other related enzyme. The heavy hybridizing bands, then, represent the xylanase locus.

All the *Butyrivibrio* strains used in this study are xylanolytic. As determined from DNA reassociation data, *B. fibrisolvens* strains 49, H17c, CE-51, CE-52, and 12 represent strains of the same species (group 1 strains). Strain CF3 is closely related to strain 49, while strains NOR-37, PI-7, D30g, C3, D1, E46a, and D16f are moderately related (20). The taxonomic relationships of the strains and the presence of the xylanase locus are summarized in Fig. 6. In group 1 strains, the probe hybridized strongly to DNAs of strains 49 and H17c but not to the DNAs of the other three strains. The probe also hybridized strongly to CF3 DNA and not to chromosomal DNA from any of the other strains. This pattern of hybridization was unexpected. 16S rRNA sequencing data indicate that the *Butyrivibrio* strains used are related (D. Stahl and B. Mannarelli, unpublished data). The strains are similar phenotypically and were originally classified as belonging to the same species. A relatively important enzyme should be present among related strains of bacteria and should at least follow taxonomic relationships. Could the results indicate gene movement among strains? The data may present interesting evolutionary questions.

ACKNOWLEDGMENTS

We thank Larry Tjarks for technical assistance in the sequencing of the *xynA* gene and Terry Whitehead for helping in the mapping of pML110.

LITERATURE CITED

1. **Ausubel, F. M., R. Brent, R. E. Kingston, D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl.** 1988. Current protocols in molecular biology. Greene Publishing Associates & Wiley-Interscience, New York.
2. **Bernier, R., H. Driquez, and M. Desrochers.** 1983. Molecular cloning of a *Bacillus subtilis* xylanase gene in *Escherichia coli*. Gene **26**:59–65.
3. **Biely, P., O. Markovic, and D. Mislovicoca.** 1985. Sensitive detection of endo-1,4-β-glucanases and endo-1,4-β-xylanases in gels. Anal. Biochem. **144**:147–151.
4. **Boyer, H. W., and D. Roulland-Dussoix.** 1969. A complementation analysis of the restriction and modification of DNA in *E. coli*. J. Mol. Biol. **41**:459–472.
5. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72**:248–254.
6. **Brenner, D. J.** 1973. Deoxyribonucleic acid reassociation in the taxonomy of enteric bacteria. Int. J. Syst. Bacteriol. **23**:298–307.
7. **Bryant, M. P., and N. Small.** 1956. The anaerobic monotrichous butyric acid-producing curved rod-shaped bacteria of the rumen. J. Bacteriol. **72**:16–21.
8. **Carter, P., H. Bedouelle, and G. Winter.** 1985. Improved oligonucleotide site-directed mutagenesis using M13 vectors. Nucleic Acids Res. **13**:4431–4443.
9. **Cornelis, P., C. Digneffe, and K. Willemot.** 1982. Cloning and expression of a *Bacillus coagulans* amylase gene in *Escherichia coli*. Mol. Gen. Genet. **186**:507–511.
10. **Dehority, B. A.** 1973. Hemicellulose degradation by rumen bacteria. Fed. Proc. **32**:1819–1825.
11. **Dehority, B. A., and H. W. Scott.** 1967. Extent of cellulose and hemicellulose digestion in various forages by pure cultures of rumen bacteria. J. Dairy Sci. **50**:1136–1141.
12. **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.
13. **Gierasch, L. M.** 1989. Signal sequences. Biochemistry **28**:923–930.
14. **Hamamoto, T., H. Honda, T. Kudo, and K. Horikoshi.** 1987. Nucleotide sequence of the xylanase A gene of alkalophilic *Bacillus* sp. strain C-125. Agric. Biol. Chem. **51**:953–955.
15. **Henikoff, S.** 1987. Unidirectional digestion with exonuclease III in DNA sequence analysis. Methods Enzymol. **155**:156–165.
16. **Hespell, R. B., R. Wolf, and R. J. Bothast.** 1988. Fermentation of

xylans by *Butyrivibrio fibrisolvens* and other ruminal bacterial species. Appl. Environ. Microbiol. **53**:2849–2853.

17. **Hungate, R. E.** 1966. The rumen and its microbes, p. 8–90. Academic Press, Inc., New York.
18. **Kudo, T., A. Atsushi, and K. Horikoshi.** 1985. Molecular cloning and expression of a xylanase gene of alkalophilic *Aeromonas* sp. no. 212 in *Escherichia coli*. J. Gen. Microbiol. **131**:2825–2830.
19. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
20. **Mannarelli, B. M.** 1988. Deoxyribonucleic acid relatedness among strains of the species *Butyrivibrio fibrisolvens*. Int. J. Syst. Bacteriol. **38**:340–347.
21. **Miller, G. L.** 1959. Use of dinitrosalicylic acid reagent for determination of reducing sugars. Anal. Chem. **31**:426–428.
22. **Neu, H. C., and L. A. Heppel.** 1965. The release of enzymes from *Escherichia coli* by osmotic shock and during the formation of spheroplasts. J. Biol. Chem. **240**:3685–3692.
23. **Paice, M. G., R. Bourbonnais, M. Desroches, L. Jurasek, and M. Yagushi.** 1986. A xylanase gene from *Bacillus subtilis*: nucleotide sequence comparison with *Bacillus pumilus* gene. Arch. Microbiol. **144**:201–206.
24. **Panbangred, W., T. Kondo, S. Negoro, A. Shinmyo, and H. Okada.** 1983. Molecular cloning of the genes for xylan degradation of *Bacillus pumilus* and their expression in *Escherichia coli*. Mol. Gen. Genet. **192**:335–341.
25. **Perlman, D., and H. O. Halvorson.** 1983. A putative signal peptidase recognition site and sequence in eukaryotic and prokaryotic signal peptides. J. Mol. Biol. **167**:391–409.
26. **Reilly, P. J.** 1981. Xylanases: structure and function, p. 11–129. *In* A. Hollaender (ed.), Trends in the biology of fermenatations. Plenum Publishing Corp., New York.
27. **Rosenberg, M., and D. Court.** 1979. Regulatory sequences involved in the promotion and termination of RNA transcription. Annu. Rev. Genet. **13**:319–353.
28. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
29. **Sewell, G. W., E. A. Utt, R. B. Hespell, K. F. Mackenzie, and L. O. Ingram.** 1989. Identification of the *Butyrivibrio fibrisolvens* xylosidase gene (*xylB*) coding region and its expression in *Escherichia coli*. Appl. Environ. Microbiol. **55**:306–311.
30. **Shine, J., and L. Dalgarno.** 1975. Determinant of cistron specificity in bacterial ribosomes. Nature (London) **254**:34–38.
31. **Sillero, A., and J. M. Ribeiro.** 1989. Isoelectric points of proteins: theoretical determination. Anal. Biochem. **179**:319–325.
32. **Sipat, A., K. A. Taylor, R. Y. C. Lo, C. W. Forsberg, and P. J. Krell.** 1987. Molecular cloning of a xylanase gene from *Bacteroides succinogenes* and its expression in *Escherichia coli*. Appl. Environ. Microbiol. **53**:477–481.
33. **Sokal, R. R., and C. D. Michener.** 1958. A statistical method for evaluating systemtatic relationships. Univ. Kansas Sci. Bull. **38**:1409–1438.
34. **Thomashow, M. F., V. C. Knauf, and E. W. Nester.** 1981. Relationship between the limited and wide-host-range octopine-type Ti plasmids of *Agrobacterium tumefaciens*. J. Bacteriol. **146**:484–493.
35. **von Heijne, G.** 1983. Patterns of amino acids near signal-sequence cleavage sites. Eur. J. Biochem. **133**:17–21.
36. **Watson, M. E.** 1984. Compilation of published signal sequences. Nucleic Acids Res. **12**:5145–5164.
37. **Yang, R. C., C. R. MacKenzie, and S. A. Narang.** 1988. Nucleotide sequence of a *Bacillus circulans* xylanase gene. Nucleic Acids Res. **16**:7187.
38. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**:103–119.
39. **Zappe, H., D. T. Jones, and D. R. Woods.** 1987. Cloning and expression of a xylanase gene from *Clostridium acetobutylicum* P262 in *Escherichia coli*. Appl. Microbiol. Biotechnol. **27**:57–63.