JOURNAL OF BACTERIOLOGY, May 1993, p. 2943–2951
0021-9193/93/102943-09$02.00/0

Vol. 175, No. 10

A Bifunctional Enzyme, with Separate Xylanase and β(1,3-1,4)-Glucanase Domains, Encoded by the *xynD* Gene of *Ruminococcus flavefaciens*

HARRY J. FLINT,* JENNIFER MARTIN, CAROL A. McPHERSON, ANNE S. DANIEL, AND JUN-XIAN ZHANG

Rowett Research Institute, Greenburn Road, Bucksburn, Aberdeen AB2 9SB, United Kingdom

Received 18 December 1992/Accepted 8 March 1993

Adjacent regions of a *Ruminococcus flavefaciens* 17 DNA fragment were found to encode xylanase and β(1,3-1,4)-glucanase activities. Sequencing of this fragment showed that both activities are encoded by a single 2,406-bp open reading frame corresponding to the *xynD* gene. The predicted product has a characteristic signal sequence that is followed by an amino-terminal domain related to family G xylanases, while the carboxy-terminal domain is related to β(1,3-1,4)-glucanases from several other bacterial species. These two domains are connected by a region of unknown function that consists of 309 amino acids and includes a 30-amino-acid threonine-rich sequence. A polypeptide having a molecular weight of approximately 90,000 and exhibiting xylanase and β(1,3-1,4)-glucanase activities was detected in *Escherichia coli* cells carrying the cloned *xynD* gene. This is one of the first cases in which a microbial polysaccharidase has been shown to carry separate catalytic domains active against different plant cell wall polysaccharides within the same polypeptide. *xynD* is one of a family of related genes in *R. flavefaciens* that encode enzymes having multiple catalytic domains, and the amino terminus of XYLD exhibits a high degree of similarity with the corresponding regions of another xylanase, XYLA, which carries two different xylanase catalytic domains.

Microbial degradation of plant xylans requires a range of enzymes, including β(1,4)-xylanases, α-arabinofuranosidases, β(1,4)-xylosidases, α-glucuronidases, and various esterases, reflecting the heterogeneity and variability of the substrate (1, 3, 29, 31). β(1,4)-Xylanases attack the main chain of β(1,4)-linked xylose residues, but vary considerably in their bond preferences and responses to substituents. Multiple xylanases are commonly produced by microorganisms (30). The rumen is the site of rapid degradation of a particularly wide range of plant xylans, and cellulolytic bacteria are among the most efficient ruminal microorganisms at degrading the hemicellulose present in plant cell walls (2, 18). The gram-positive cellulolytic anaerobe *Ruminococcus flavefaciens* 17 is known to possess at least four different genes, designated *xynA*, *xynB*, *xynC*, and *xynD*, which encode β(1,4)-xylanases that lack significant associated cellulase activity (7). *xynA*, *xynB*, *xynC*, and *xynD* correspond to homology groups 4, 2, 3, and 1, respectively, which were originally identified among xylanase clones from a lambda bacteriophage library of *R. flavefaciens* DNA (5). It was shown recently that the *xynA* gene encodes a novel bifunctional enzyme which comprises two distinct xylanase domains joined by an unusual linker region (34). The molecular structure of additional genes is therefore being investigated in order to establish whether this type of organization is present in other xylanases from this species.

Clones carrying the *xynD* gene from *R. flavefaciens* 17 were originally found to differ from other clones by exhibiting associated activity against lichenan [β(1,3-1,4)-glucan] (5). In other species the association of these two activities has been attributed to a single 29-kDa protein in *Clostridium acetobutylicum* (15) and to two closely linked genes in *Bacillus polymyxa* (12). The results reported here show that in *R. flavefaciens* 17 a single gene, *xynD*, encodes a bifunctional enzyme having separate xylanase and β(1,3-1,4)-glucanase domains. The product of the *xynD* gene appears to be one of the first examples of a bifunctional polysaccharidase having two separate catalytic domains within the same polypeptide chain that can act on different polymeric substrates.

MATERIALS AND METHODS

Strains and growth conditions. Isolation of clone L9 from a library of *R. flavefaciens* 17 DNA made in lambda EMBL3 has been described previously (5). L9 phage was propagated in *Escherichia coli* P2392. *E. coli* DH5α and HB101 were used as hosts for isolation and maintenance of subclones and were grown on Luria-Bertani medium containing 50 μg of ampicillin per ml for selection of pUC plasmids.

Molecular biology procedures. For DNA isolation, subcloning, and restriction enzyme analyses we used standard methods (21). DNA sequences were determined for both strands by using the Sanger dideoxy chain termination method (22) and T7 DNA polymerase (Sequenase 2; United States Biochemical Corp., Cleveland, Ohio) in the presence of deaza-dGTP. Sequences were determined from subclones made in pUC13 by using M13 forward and reverse primers or internal oligonucleotide primers (20 mers) synthesized with a Cruachem DNA synthesizer. The sequence shown in Fig. 2 was derived by using DNAs originating from the L956, L9XSH, and L9XRB plasmid subclones (see Fig. 1). The junction between L956 and L9XSH was established from L9XRB sequences. Sequences were analyzed by using the UWGCG and OWL software available through the SERC Seqnet facility (Daresbury, United Kingdom). Multiple alignments were performed by using the ClustalV program, and pairwise comparisons were performed by using the Bestfit program (gap weight, 3; length weight, 0.1).

* Corresponding author.

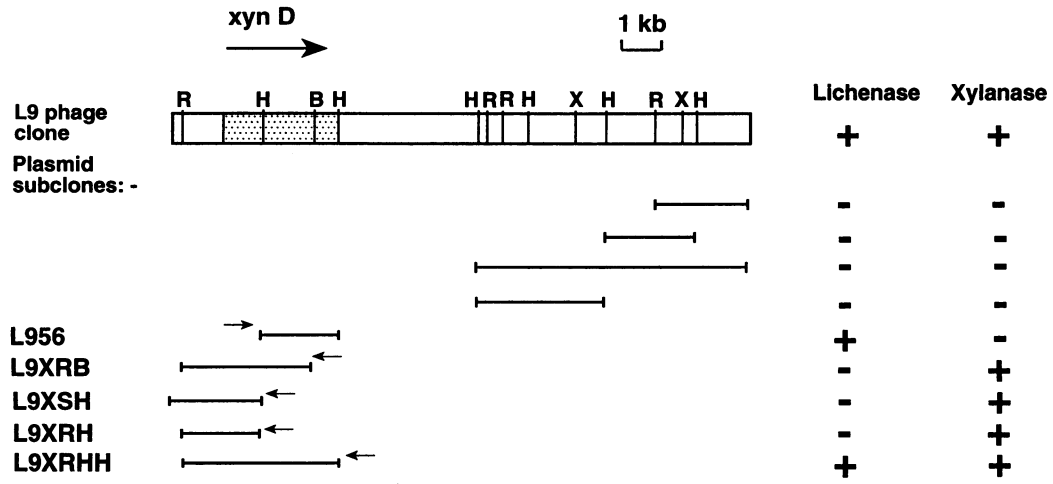

FIG. 1. Identification of regions encoding xylanase and lichenase [β(1,3-1,4)-glucanase] activities in the L9 DNA fragment from *R. flavefaciens* 17. Fragments successfully subcloned in plasmid vector pUC13 are shown beneath the restriction map of the L9 fragment. Terminal fragments were cloned by making use of flanking *Sal*I sites in the vector (sites not shown). The activities shown on the right are based on the ability to produce clearing zones in Luria-Bertani agar containing 0.2% Remazole brilliant blue xylan or 0.1% lichenan (revealed in the latter case by Congo red staining). The *xynD* open reading frame (see Fig. 2) is stippled, and its orientation is indicated by an arrow. Restriction enzyme cleavage sites are shown for *Eco*RI (R), *Hin*dIII (H), *Bam*HI (B), and *Xho*I (X). The smaller arrows indicate the directions of *lacZ* transcription in the active subclones.

Enzyme activities. Plate tests for xylanase, in which Remazole brilliant blue xylan was used, and for lichenase, in which plates carrying lichenan overlays were stained with Congo red (26), were performed as described previously (6). Preparation of *E. coli* cell extracts by sonication and assays for reducing sugar or paranitrophenol release were performed as described previously (6). L9 phage lysate (5) was concentrated 35-fold by ultrafiltration with an Amicon PM10 membrane filter and was dialyzed at 4°C against 25 mM sodium phosphate buffer (pH 6.8) (twice with 100 volumes, once with 200 volumes). Laminarin was used as an assay substrate only after extensive washing and dialysis to remove free reducing sugars. For zymogram analyses we used the methods of Saul et al. (23), as adapted for detection of xylanase activities (34), except that polypeptides were first separated in 12% polyacrylamide gels and activities were detected following renaturation by applying 2% agarose gels containing 0.2% substrate to form a gel sandwich. A 0.2% lichenan solution was prepared by autoclaving a 0.8% stock solution before the lichenan was incorporated into the substrate gel. The *R. flavefaciens* 17 cells used for zymogram analyses were grown for 40 h in defined medium containing 0.2% oat spelt xylan, 0.2% Avicel, 0.2% oat straw, or 0.2% lichenan as an energy source, as described previously (6).

RESULTS

Localization of regions coding for xylanase and lichenase activities. In previous work a lambda bacteriophage clone, L9, carrying an 11.5-kb fragment of *R. flavefaciens* 17 DNA was found to encode xylanase, lichenase [β(1,3-1,4)-glucanase], and β(1,4)-xylosidase activities and a trace of α-L-arabinofuranosidase activity (5). In this study a 1.5-kb fragment encoding only the lichenase activity was isolated from a *Hin*dIII partial digest of phage clone L9 by subcloning in pUC13 (L956) (Fig. 1). In addition, three overlapping fragments (L9XRB, L9XRH, and L9XSH) that specified xylanase activity but not lichenase activity were isolated (Fig. 1). These fragments mapped adjacent to the 1.5-kb lichenase-encoding fragment, confirming earlier evidence that the xylanase activity was encoded close to one end of the L9 insert (5). Neither these subclones nor any of the other subclones of the L9 phage shown in Fig. 1 was found to encode appreciable β-D-xylosidase or α-L-arabinofuranosidase activity. However, we were unable to recover stable plasmid subclones carrying a 2.8-kb *Hin*dIII fragment from the center of the L9 insert (Fig. 1).

Table 1 shows the activities detected in assays of disrupted cells with xylanase-expressing clones L9XSH, L9XRB, and L9XRH and with lichenase-expressing clone L956. L956 also showed activity against the β(1,3)-glucan laminarin (12 nmol/min/mg of protein). Although no plasmid subclones that encoded both xylanase and lichenase activities were obtained, the 1.5-kb lichenase-encoding *Hin*dIII fragment from L956 was inserted in place of the terminal *Hin*dIII-*Bam*HI fragment of xylanase-expressing clone L9XRB in order to restore the arrangement found in the original L9 phage clone (Fig. 1). The resulting construct, L9XRHH, showed both lichenase and xylanase activities

TABLE 1. Activities of pUC13 subclones and constructs carrying partial or complete *xynD* coding sequences[a]

Subclone or construct	Enzyme activities (nmol/min/mg of protein)		
	Xylanase (oat spelt xylan)	β(1,3-1,4)-Glucanase (lichenan)	Carboxymethyl cellulase
L9 phage	34.3	25.6	<1
L956	<1	150	<1
L9XRB	74.9	<1	<1
L9XSH	15.8	<1	<1
L9XRH	7.4	<1	<1
L9XRHH	95.5	83.2	<1

[a] Enzyme assays were performed with sonicated extracts of *E. coli* cells or with dialyzed, concentrated phage lysates by measuring reducing sugar release. The *E. coli* host strains were P2392 for phage L9, HB101 for plasmid subclone L956, and DH5α in all other cases.

(Table 1), and the relative activities were similar to those in the original phage clone L9 when the preparations were assayed under the same conditions (Table 1). Relatively lower xylanase activities were reported previously for L9 phage lysates (5), but these values were from assays in which undialyzed, concentrated lysates were used together with different substrates and assay conditions and are not comparable with those shown in Table 1. Xylanase activity was apparently enhanced in L9XRB and L9XRHH compared with L9XSH and L9XRH (Table 1). No activity was detected in any of the plasmid subclones shown in Table 1 against the *p*-nitrophenyl derivatives *p*-nitrophenyl xyloside, *p*-nitrophenyl cellobioside, *p*-nitrophenyl arabinofuranoside, and *p*-nitrophenyl glucoside.

Nucleotide sequence of the *xynD* gene. It is clear from Fig. 1 that the xylanase and lichenase activities of the L9 region must be encoded separately, either by closely linked genes or as parts of a multifunctional gene product. Nucleotide sequencing of this region revealed a single, long, 2,406-bp open reading frame extending from within the xylanase-encoding fragment almost to the end of the lichenase-encoding fragment (Fig. 2); 7 bp upstream of the proposed ATG start is a sequence (GAGG) complementary to the 3′ ends of *Bacillus subtilis* and *E. coli* 16S rRNAs, which is likely to represent the *R. flavefaciens* ribosome binding site. Identification of the translational start is largely based on the presence of a region encoding a leader peptide immediately following the proposed initiation triplet and subsequent homology with the *xynA* gene product of *R. flavefaciens* (see below) (Fig. 3). Expression of xylanase activity in *E. coli* in the xylanase-positive clones shown in Fig. 1 is assumed to be from a promoter within the *R. flavefaciens* DNA, since transcription from the *lacZ* promoter is in the opposite direction. Sequences that could act as promoters in *E. coli* were present in regions upstream (data not shown), but it is not known whether these sequences correspond to the promoters functioning in *Ruminococcus* sp. Expression of lichenase activity from the L956 clone is assumed to have been due to the *lacZ* promoter since clones in which this fragment was in the opposite orientation exhibited no activity.

Sequence relationships of the *xynD* gene product. The predicted product of the *xynD* gene (XYLD) begins with a sequence in which 4 of the first 8 residues are basic and 15 of the next 18 residues are nonpolar. These are typical features for signal peptides from gram-positive bacteria (28), and a signal peptidase cleavage site is predicted to occur after residue 31 (Fig. 2). A somewhat similar signal sequence is found at the start of the *R. flavefaciens xynA* gene (Fig. 3). The next 213 residues of the *xynD* product, which are designated domain A, exhibit significant sequence similarity with xylanases belonging to family G, as defined by Gilkes et al. (11) (Fig. 3). The closest sequence similarity is with amino-terminal domain A of the *xynA* product of *R. flavefaciens* 17 (Fig. 3). Pairwise (Bestfit) comparisons, in which the signal sequences were omitted, revealed a value of 81.1% identical residues for XYLD domain A and domain A of *R. flavefaciens* XYLA and values of 43.2 and 38.8% identical residues, respectively, between XYLD domain A and the xylanases from *Bacillus pumilus* and *B. subtilis*. The level of similarity at the DNA sequence level was low enough to prevent detectable cross-hybridization of the X4 and L9 phage inserts carrying *xynA* and *xynD*, respectively, in previous work (5), but the initial sequences of these two genes nevertheless encode very similar amino acid sequences.

The carboxy-terminal end (249 residues) of the XYLD polypeptide, which is referred to below as domain C, exhibits substantial sequence similarity to β(1,3-1,4)-glucanases from several other bacterial species (Fig. 4). This domain is known to be responsible for the β(1,3-1,4)-glucanase activity observed in this study for the *xynD* gene product, since the absence of sequences encoding the final 171 amino acids in the L9XRB clone was associated with a loss of lichenase activity (Fig. 1 and Table 1). Pairwise comparisons revealed 53, 51, and 52% identical residues between XYLD domain C and β(1,3-1,4)-glucanases from *B. subtilis*, *Bacillus licheniformis*, and *Clostridium thermocellum*, respectively. Alignments (data not shown) with the β(1,3-1,4)-glucanase from the rumen species *Fibrobacter succinogenes* (25) revealed a much lower number of identical residues (34%), even after we allowed for the rearrangement of subdomains inferred to occur in this case (12). Expression of lichenase activity in the L956 clone, which lacks the amino-terminal coding regions of the *xynD* gene, is assumed to be due to internal starts within the region encoding domains B and C.

Immediately preceding the β(1,3-1,4)-glucanase domain of XYLD (domain C) is a highly threonine-rich sequence in which 26 of the 30 residues are threonines (Fig. 2 and 5). Similar regions are commonly found linking domains of microbial polysaccharidases (11). Preceding this threonine-rich sequence and following the xylanase domain (domain A) in the same polypeptide is a 279-amino-acid region (domain B) whose function and relationships are not known (Fig. 5). This domain does not resemble the linker sequence connecting the two different catalytic domains of the XYLA product, reported previously (34), which is characterized by an unusual repetitive structure with a high content of asparagine and glutamine residues. Searches of the OWL composite protein data base (which includes the SWISSPROT, GenBank, and NBRF data bases) revealed no closely similar sequences except for a 10-residue sequence in domain B (GQWVDLSNSS, positions 362 to 371 in the XYLD translation product) which is identical to a sequence present in a *Bacillus circulans* β-(1,3)-glucanase (33). Significant similarity was limited to this stretch of amino acids, however (data not shown).

Detection of the *xynD* product. The enzyme encoded by clone L9XRHH, which carries the complete *xynD* coding region, was visualized by zymogram techniques involving renaturation of enzyme activity following separation by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (Fig. 6). The results show that the largest polypeptide detected was an approximately 90-kDa polypeptide, and this polypeptide exhibited both xylanase and β(1,3-1,4)-glucanase activities. This probably represents the full-length, bifunctional XYLD product, whose predicted molecular mass is 89.5 kDa. The majority of xylanase and lichenase activity was present in lower-molecular-weight material, however. This is thought to be due to a combination of proteolysis and internal translational starts. For example, the major band of lichenase activity, at approximately 65 kDa, seems likely to be due to an alternative translation start (an ATG at nucleotide 701 [Fig. 2]) which gives a predicted product of 589 amino acids (molecular mass, approximately 66 kDa). This internal ATG is preceded by a region in which seven of eight nucleotides are complementary to the 3′ end of *E. coli* 16S rRNA, giving a predicted ΔG of −10.8 kcal/mol (-45.2×10^3 J/mol) (see reference 27). By comparison, the corresponding value for the ribosome binding region preceding the true initiation codon was −9.4 kcal/mol (-39.3×10^3 J/mol). Another potential ribosome

```
1   CTGCGTTGCAATGATGTATAAATAATACAGTGACCACGGTCACGGAATGAGAGGGCTGAT   60

61  TATGAAAAAAGAGCATATTCAAGCGTTATGCTGCTGCTGTCGGTTTAATGGCTTCAGTTCT  120
      M  K  K  S  I  F  K  R  Y  A  A  A  V  G  L  M  A  S  V  L

121 TATGTTTACCGCAGTACCGACAACGTCAAATGCTGCTGATGACCAGAAAACAGGTAAGGT  180
      M  F  T  A  V  P  T  T  S  N  A  A  D  D  Q  K  T  G  K  V

181 CGGCGGATTTGACTGGGAAATGTGGAACCAGAACTATACCGGTACAGTTTCTATGAATCC  240
      G  G  F  D  W  E  M  W  N  Q  N  Y  T  G  T  V  S  M  N  P

241 GGGCGCAGGTTCATTTACCTGTTCGTGGAGCGGCATTGAGAACTTCCTTGCTCGAATGGG  300
      G  A  G  S  F  T  C  S  W  S  G  I  E  N  F  L  A  R  M  G

301 TAAGAATTATGACGACCAGAAGAAGAACTATAAGGCTTTCGGTGACATCGTGCTTACATA  360
      K  N  Y  D  D  Q  K  K  N  Y  K  A  F  G  D  I  V  L  T  Y

361 TGATGTTGAGTACACTCCCCGCGGAAACTCGTATATGTGTATTTACGGCTGGACAAGAAA  420
      D  V  E  Y  T  P  R  G  N  S  Y  M  C  I  Y  G  W  T  R  N

421 TCCCCTTATGGAATACTATATCGTTGAAGGCTGGGGCGACTGGGAGCCACCCGGAAATGA  480
      P  L  M  E  Y  Y  I  V  E  G  W  G  D  W  E  P  P  G  N  D

481 CGGCGTAGACAATTTCGGTACAACAACTATCGACGGAAAAACATATAAGATACGCAAGTC  540
      G  V  D  N  F  G  T  T  T  I  D  G  K  T  Y  K  I  R  K  S

541 CATGCGTTACAATCAGCCTTCTATCGAGGGAACAAAGACATTCCCGCAGTACTGGAGCGT  600
      M  R  Y  N  Q  P  S  I  E  G  T  K  T  F  P  Q  Y  W  S  V

601 CCGCACTACAAGCGGTTCACGCAATAATACTACAAACTATATGAAGGATCAGGTCTCTGT  660
      R  T  T  S  G  S  R  N  N  T  T  N  Y  M  K  D  Q  V  S  V

661 TACCAAGCATTTTGACGCATGGAGCAAGGCAGGTCTCGACATGAGCGGTACACTTTATGA  720
      T  K  H  F  D  A  W  S  K  A  G  L  D  M  S  G  T  L  Y  E

721 GGTATCACTTAATATCGAAGGCTACAGATCAAACGGTTCTGCCAATGTCAAGAGCATATC  780
      V  S  L  N  I  E  G  Y  R  S  N  G  S  A  N  V  K  S  I  S

781 CTTTGACGGCGGTATTGATATTCCCGATCCCGAGCCGATAAAGCCTGACGAGAACGGCTA  840
      F  D  G  G  I  D  I  P  D  P  E  P  I  K  P  D  E  N  G  Y

841 CTATCTCAAGGAAAATTTTGAGTCCGGCGAAGGAAACTGGAGCGGACGCGGATCTGCAAA  900
      Y  L  K  E  N  F  E  S  G  E  G  N  W  S  G  R  G  S  A  K

901 GGTCAAGTCAAGCTCAGGCTATGACGGAACGAAGGGCATCTTTGTTTCAGGCCGTGAAGA  960
      V  K  S  S  S  G  Y  D  G  T  K  G  I  F  V  S  G  R  E  D
```

binding site (ΔG, -10.2 kcal/mol [-42.7×10^3 J/mol]) was also present before an in-frame ATG codon at nucleotide 200, which could yield a 756-amino-acid product.

The remaining bands in Fig. 6 are thought to result from the action of *E. coli* proteases on cleavage sites within domain B of XYLD. Thus, cleavage at a site late in domain B should yield a xylanase having a molecular weight of approximately 60,000 and a lichenase having a molecular

```
 961  TACATGGAACGGCGCTTCCATAAACCTCGATGAACTTACATTCAAGGCAGGAGAGACTTA  1020
       T  W  N  G  A  S  I  N  L  D  E  L  T  F  K  A  G  E  T  Y

1021  CAGCTTAGGTACTGCTGTTATGCAGGACTTCGAGTCATCTGTTGACTTCAAGCTTACTCT  1080
       S  L  G  T  A  V  M  Q  D  F  E  S  S  V  D  F  K  L  T  L

1081  TCAGTATACAGATGCTGACGGCAAAGAGAATTATGATGAGGTAAAGACTGTTACTGCTGC  1140
       Q  Y  T  D  A  D  G  K  E  N  Y  D  E  V  K  T  V  T  A  A

1141  AAAGGGACAGTGGGTAGATCTTTCAAATTCTTCATATACGATACCTTCGGGAGCAACAGG  1200
       K  G  Q  W  V  D  L  S  N  S  S  Y  T  I  P  S  G  A  T  G

1201  ACTTGTTCTCTATGTTGAGGTGCCTGAGAGCAAGACAGACTTCTATATGGACGGCGCTTA  1260
       L  V  L  Y  V  E  V  P  E  S  K  T  D  F  Y  M  D  G  A  Y

1261  TGCTGGCGTTAAGGGCACAAAGCCTCTTATCTCCATTTCTTCTCAGTCTGTTGATCCTCC  1320
       A  G  V  K  G  T  K  P  L  I  S  I  S  S  Q  S  V  D  P  P

1321  TGTTACAGAGCCTACAAATCCCACAAATCCTACCGGTCCTTCGGTTACAAAGTGGGGTGA  1380
       V  T  E  P  T  N  P  T  N  P  T  G  P  S  V  T  K  W  G  D

1381  TGCAAACTGTGACGGCGGTGTAGACCTCAGTGATGCTATATTTATCATGCAGTTTCTTGC  1440
       A  N  C  D  G  G  V  D  L  S  D  A  I  F  I  M  Q  F  L  A

1441  AAATCCGAATAAGTATGGTCTTACAGGAACTGAGACAAATCATATGACCAATCAGGGAAA  1500
       N  P  N  K  Y  G  L  T  G  T  E  T  N  H  M  T  N  Q  G  K

1501  GGTCAACGGAGATGTCTGCGAGCATGGTTCGGGACTTACCGGAGACGATGCTGTTTCCAT  1560
       V  N  G  D  V  C  E  H  G  S  G  L  T  G  D  D  A  V  S  I

1561  CCAGAAGTACCTCATCCGTGCCATCAGCGAGCTTCCAGAGTCTTATCTCGAAGGCCATGA  1620
       Q  K  Y  L  I  R  A  I  S  E  L  P  E  S  Y  L  E  G  H  D

1621  TCCTTCAAAGACAACGACCACTACAACAAGGATCACTACTACAACTACAACTACCACAAC  1680
       P  S  K  T  T  T  T  T  T  R  I  T  T  T  T  T  T  T  T  T

1681  GACTACAACATCGAAGACAACGACCACTACTACCACAACTTCTCCCGCAATGCACGGCGG  1740
       T  T  T  S  K  T  T  T  T  T  T  T  T  S  P  A  M  H  G  G

1741  TTACAGAGATCTTGGTACTCCTATGAATACAAGTGCTACAATGATCTCAGATTTCCGTAC  1800
       Y  R  D  L  G  T  P  M  N  T  S  A  T  M  I  S  D  F  R  T

1801  AGGCAAGGCAGGCGACTTCTTTGCATCCGACGGATGGACCAACGGCAAGCCTTTCGACTG  1860
       G  K  A  G  D  F  F  A  S  D  G  W  T  N  G  K  P  F  D  C

1861  CTGGTGGTACAAGCGTAATGCTGTTATCAATGACGGCTGTCTCCAGCTGAGCATTGACCA  1920
       W  W  Y  K  R  N  A  V  I  N  D  G  C  L  Q  L  S  I  D  Q

1921  GAAATGGACAAATGACAAGAATCCCGACTGGGATCCCCGTTATTCCGGCGGTGAGTTCCG  1980
       K  W  T  N  D  K  N  P  D  W  D  P  R  Y  S  G  G  E  F  R
```

```
1981  TACAAATAATTTCTATCACTACGGATACTACGAGTGCTCAATGCAGGCTATGAAGAATGA  2040
       T  N  N  F  Y  H  Y  G  Y  Y  E  C  S  M  Q  A  M  K  N  D

2041  CGGTGTTGTATCATCGTTCTTCACCTATACAGGTCCGTCTGACGATAACCCGTGGGATGA  2100
       G  V  V  S  S  F  F  T  Y  T  G  P  S  D  D  N  P  W  D  E

2101  GATCGATATCGAGATCCTTGGCAAGAACACTACACAGGTTCAGTTCAACTATTACACAAA  2160
       I  D  I  E  I  L  G  K  N  T  T  Q  V  Q  F  N  Y  Y  T  N

2161  CGGTCAGGGAAAGCATGAGAAGCTGTACGACCTCGGCTTTGATTCTTCTGAGGCTTATCA  2220
       G  Q  G  K  H  E  K  L  Y  D  L  G  F  D  S  S  E  A  Y  H

2221  TACATACGGCTTTGACTGGCAGCCGAACTACATCGCATGGTATGTAGACGGCAGGGAAGT  2280
       T  Y  G  F  D  W  Q  P  N  Y  I  A  W  Y  V  D  G  R  E  V

2281  TTATCGTGCTACTCAGGATATTCCCAAGACCCCCGGAAAGATCATGATGAACGCATGGCC  2340
       Y  R  A  T  Q  D  I  P  K  T  P  G  K  I  M  M  N  A  W  P

2341  CGGACTTACAGTTGATGACTGGCTCAAGGCATTCAACGGAAGAACTCCTCTTACAGCTCA  2400
       G  L  T  V  D  D  W  L  K  A  F  N  G  R  T  P  L  T  A  H

2401  CTATCAGTGGGTTACATATAATAAGAACGGTGTTCAGCACAGCAGTCAGGGACAGAACCC  2460
       Y  Q  W  V  T  Y  N  K  N  G  V  Q  H  S  S  Q  G  Q  N  P

2461  CTGGGGTTGATAAAACGCAATAATTGAAGCTT  2492
       W  G  *
```

FIG. 2. Nucleotide sequence and predicted translation product of the *xynD* gene of *R. flavefaciens* 17. The sequence was determined completely for both strands as described in Materials and Methods. The proposed translational start codon and ribosome binding site for the gene are enclosed in boxes. In addition, two of the possible internal start and ribosome binding sites which may be recognized in *E. coli* (see text) are underlined. The putative signal peptide sequence also is underlined.

weight of approximately 30,000, while cleavage at an earlier site could yield a xylanase having a molecular weight of 53,000 and a lichenase having a molecular weight of 37,000. The remaining major bands, a 35-kDa xylanase and a 43-kDa lichenase, could be explained by a protease cutting at two sites within domain B. Only the 35-kDa xylanase band was detected in truncated clone L9XSH, which retains only the first 338 codons of *xynD* (Fig. 6).

When similar approaches were used, up to four major bands of lichenase activity were detected in *R. flavefaciens* cells grown on media containing xylan, lichenan, cellulose, or straw as an energy source. These bands had very approximate molecular masses of 75, 90, 110, and >120 kDa (data not shown). The failure to detect lichenases having molecular weights less than 75,000 therefore supports the view that the smaller lichenase bands in Fig. 6 resulted from expression of *xynD* in *E. coli* and are not major products in *R. flavefaciens*.

DISCUSSION

The evidence presented here shows that the *xynD* gene of *R. flavefaciens* 17 consists of a 2,406-bp open reading frame that encodes a bifunctional polypeptide (XYLD, 802 amino acids) having as its amino-terminal domain (domain A) a family G-related xylanase and as its carboxy-terminal domain (domain C) a β(1,3-1,4)-glucanase. Separating these two domains is a region of unknown function made up of a 279-amino-acid sequence referred to as domain B, followed by a 30-amino-acid threonine-rich sequence (Fig. 5). It has become apparent only recently that certain microbial enzymes concerned with the breakdown of plant cell wall polysaccharides carry separate and distinct catalytic regions within the same polypeptide. Separate catalytic domains have been clearly demonstrated previously in an exocellulase-endocellulase from *Caldocellum saccharolyticum* (23) and in the bifunctional XYLA xylanase from *R. flavefaciens* (34). In the latter case the enzyme carries a family G-related xylanase sequence at its amino terminus and a family F-related xylanase sequence at its carboxy terminus; these two domains are known to differ in their action on xylan, suggesting that the composite XYLA enzyme might gain a broader spectrum of bond cleavage for this substrate (8a). In contrast, the two catalytic domains of the XYLD enzyme attack bonds that are expected to occur in different polysaccharides. Any benefit from the bifunctional organization of XYLD might therefore be expected to come from an enhanced ability to degrade mixtures of xylan and β(1,3-1,4)-glucan chains occurring in plant material, although interaction of the two domains in the degradation of some as-yet-unspecified complex polysaccharide cannot be ruled out.

The close similarity between the amino-terminal domains encoded by the *R. flavefaciens xynA* and *xynD* genes sug-

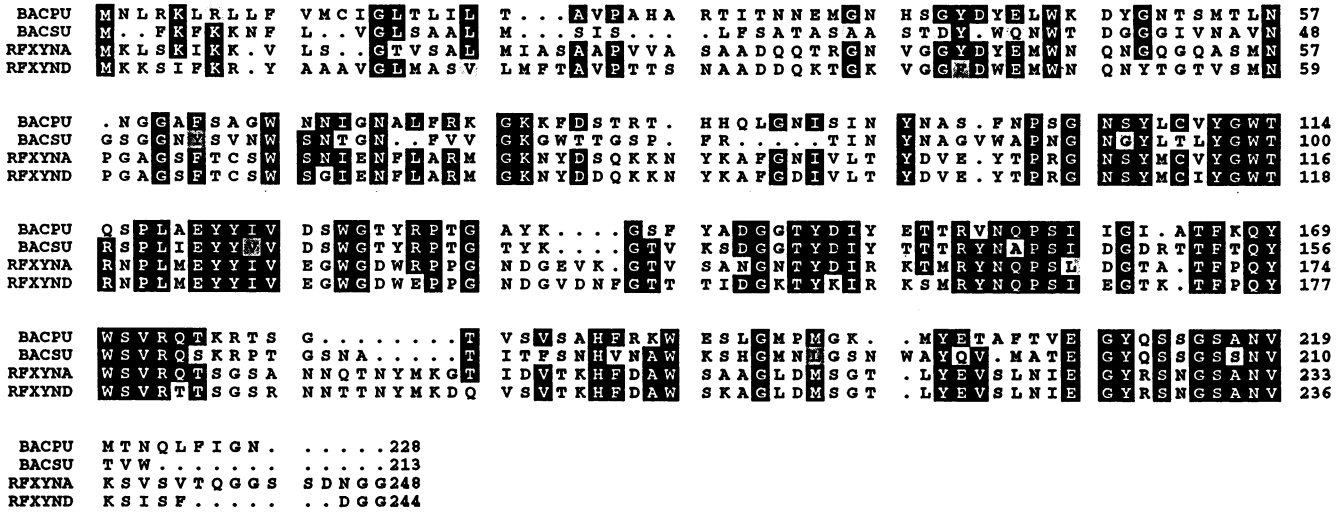

FIG. 3. Relationship of amino-terminal domain A of XYLD with family G xylanases from various bacterial species. RFXYLA and RFXYLD are the amino-terminal xylanase domains of *R. flavefaciens* 17 XYLA (34) and XYLD, respectively. Also shown are the xylanases encoded by the *xynA* genes from *B. subtilis* (BACSU) (20) and *B. pumilus* (BACPU) (8). Conserved residues are indicated by a black background. A grey background indicates substitution of a similar amino acid in one of the four sequences.

gests an evolutionary origin that involved first duplication of an ancestral gene encoding a family G xylanase and then sequence divergence and fusion with genes encoding other catalytic domains. In fact, the amino-terminal catalytic domain of the *xynB* gene product, XYLB, is also related to the amino-terminal catalytic domains of XYLA and XYLD, and additional copies of DNA sequences corresponding to the amino-terminal domains of XYLA and XYLB are known to be present in the *R. flavefaciens* genome (6, 7), suggesting that a still larger family of related xylanase genes may be present.

The *R. flavefaciens* XYLD carboxy-terminal domain belongs to a family of enzymes that includes mixed-linkage β-glucanases from a number of bacterial species (12, 13). These β(1,3-1,4)-glucanases do not hydrolyze β(1,4)-glucan chains but apparently attack β(1,3-1,4)-glucans by cleaving β(1,4) linkages between glucose residues, one of which is bonded at carbon 3 (17). The isolated *R. flavefaciens* XYLD β-glucanase domain exhibited much higher activity against the heteropolymer lichenan than against the homopolymeric β(1,3)-glucan laminarin, but the latter activity indicates some ability to hydrolyze β(1,3)-glycosidic linkages. A β(1,3)-glucanase having a molecular weight of around 25,000 has been purified from *R. flavefaciens* FD1 (4), but this enzyme exhibited no β(1,3-1,4)-glucanase activity and is clearly a different enzyme.

Despite the failure to recover plasmid subclones expressing α-arabinofuranosidase or β-xylosidase, an additional

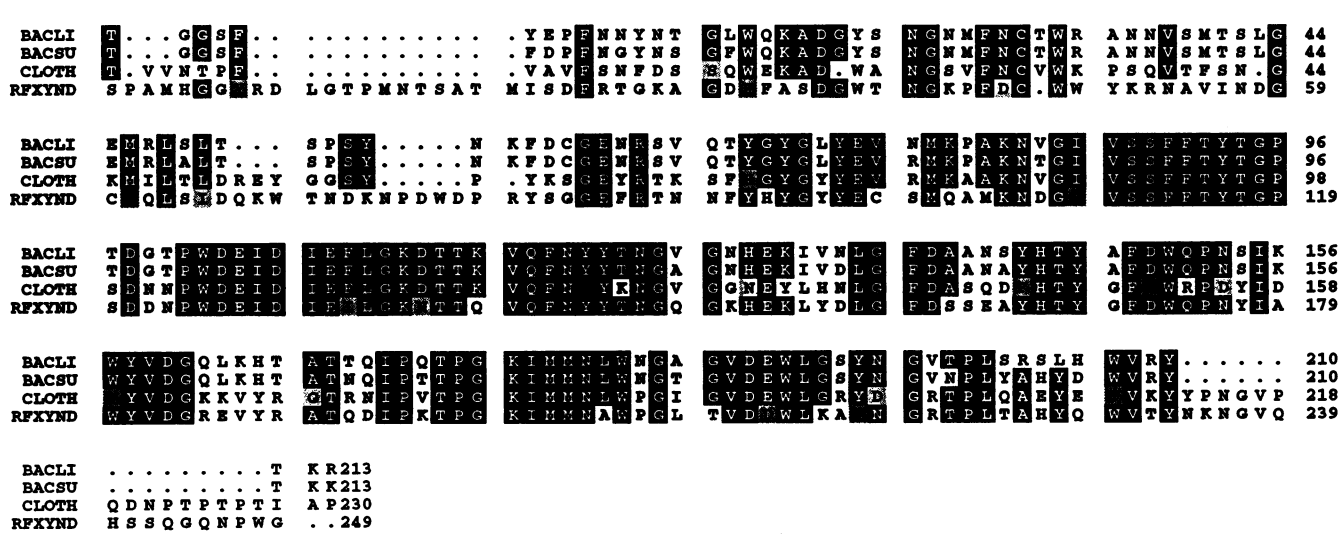

FIG. 4. Relationship of carboxy-terminal domain C of XYLD with β(1,3-1,4)-glucanases from various bacterial species. A multiple alignment with the sequences of the mature enzymes from *B. subtilis* (BACSU) (19), *B. licheniformis* (BACLI) (16), and *Clostridium thermocellum* (CLOTH) (24) is shown. Conserved residues are indicated by a black background. A grey background indicates substitution of a similar amino acid in one of the four sequences.

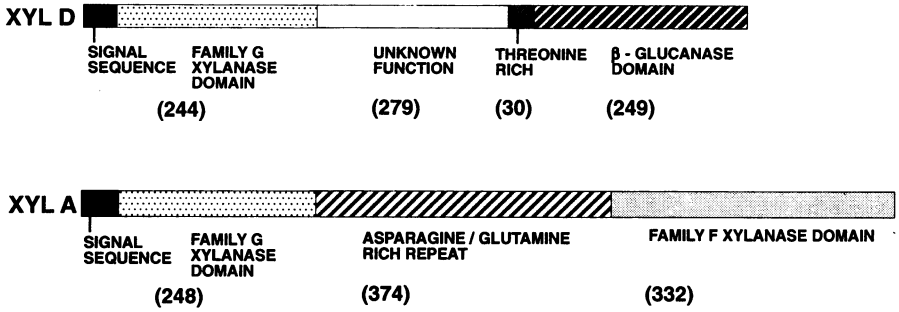

FIG. 5. Comparison of the predicted structures of bifunctional polypeptides XYLD and XYLA encoded by *R. flavefaciens* 17 *xynD* and *xynA* genes, respectively. The structure of the XYLA polypeptide is based on the data of Zhang and Flint (34). Numbers in parentheses refer to amino acid residues.

gene or genes encoding these activities linked to *xynD* must exist, since both activities were detected in the original L9 phage clone (5). The observation that mRNA species hybridizing with the L9 region are induced much more in xylan-grown cells than in cellobiose-grown cells (6) supports the view that this region is important in xylan breakdown in *R. flavefaciens*.

In conclusion, our results provide one of the first examples of a bifunctional enzyme having separate catalytic domains that act on different plant cell wall polysaccharides. It is not yet known whether this organization is the result of particular selection pressures in rumen microorganisms for efficient degradation of a wide variety of plant material under rumen conditions. Multiple catalytic domains have been reported recently in polysaccharidases from a rumen anaerobic fungus (10, 32), but these enzymes exhibit repetition of a single type of catalytic domain and so represent a phenomenon very different from that reported here. It is interesting to note, however, that a gene encoding a bifunctional endoglucanase-mannanase comprising distinct catalytic domains was recently reported from the thermophilic anaerobe *Caldocellum saccharolyticum* (9). Other reports of enzymes having apparently broad substrate specificities and the presence in many larger polysaccharidases of amino acid sequences whose function is not yet known (14) suggest that additional microbial polysaccharidases that have this type of organization may be found.

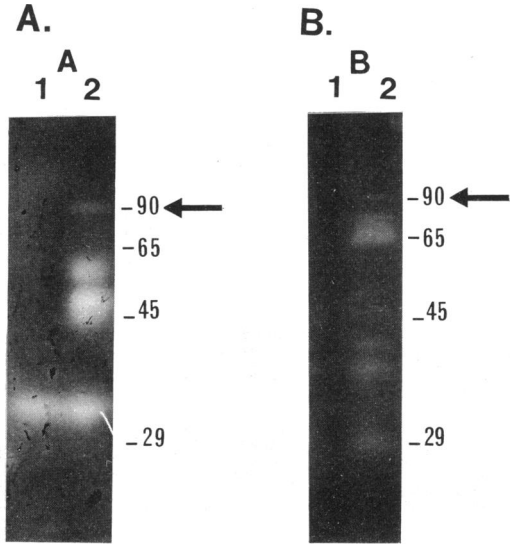

FIG. 6. Detection of the bifunctional *xynD* product synthesized by *E. coli* DH5α/L9XRHH. Polypeptides from sonicated cells were separated on an SDS-polyacrylamide gel. After renaturation the gel was incubated between agarose gels containing xylan (A) or lichenan (B), which were then stained to reveal activity (see Materials and Methods). The position of a polypeptide with a molecular weight of around 90,000 having both xylanase and lichenase activities, produced by L9XRHH (lanes 2), is indicated by the arrow. Smaller active products are thought to be due to internal translational starts or to proteolysis (see text). Lanes 1 contained clone L9XSH. The faint bands apparently present in lane 1 of panel B are in fact due to another sample giving intense bands of lichenase activity which was loaded in the preceding lane of the original gel. The positions of molecular weight markers (jack bean urease, bovine serum albumin, egg albumin, and carbonic anhydrase; molecular sizes shown in kilodaltons) are indicated on the right.

ACKNOWLEDGMENTS

This work was supported by the Scottish Office Agriculture and Fisheries Department.

We are grateful for permission to use the SERC Daresbury Seqnet facility for molecular biology computing.

REFERENCES

1. **Biely, P.** 1985. Microbial xylanolytic systems. Trends Biotechnol. **3**:286–290.
2. **Dehority, B. A., and H. W. Scott.** 1967. Extent of cellulose and hemicellulose degradation in various forages by pure cultures of rumen bacteria. J. Dairy Sci. **50**:1136–1141.
3. **Dekker, R. F. H.** 1985. Biodegradation of the hemicelluloses, p. 505–533. *In* T. Higuchi (ed.), Biosynthesis and biodegradation of wood components. Academic Press, Inc., Orlando, Fla.
4. **Erfle, J. D., and R. M. Teather.** 1991. Isolation and properties of a (1,3)-β-D-glucanase from *Ruminococcus flavefaciens*. Appl. Environ. Microbiol. **57**:122–129.
5. **Flint, H. J., C. A. McPherson, and J. Bisset.** 1989. Molecular cloning of genes from *Ruminococcus flavefaciens* encoding xylanase and β(1,3-1,4)glucanase activities. Appl. Environ. Microbiol. **55**:1230–1233.
6. **Flint, H. J., C. A. McPherson, and J. Martin.** 1991. Expression of two xylanase genes from the rumen cellulolytic bacterium *Ruminococcus flavefaciens* cloned in pUC13. J. Gen. Microbiol. **137**:123–129.
7. **Flint, H. J., and J.-X. Zhang.** 1992. Molecular biology of xylanases from the rumen cellulolytic anaerobe *Ruminococcus flavefaciens* 17—multiple genes and bifunctional enzymes, p. 301–308. *In* J. Visser, G. Beldman, M. A. Kusters-van Someren, and A. G. J. Voragen (ed.), Xylan and xylanases. Progress in biotechnology, vol. 7. Elsevier, Amsterdam.
8. **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.
8a. **Garcia-Campayo, V., et al.** Submitted for publication.
9. **Gibbs, M. D., D. J. Saul, E. Lüthi, and P. L. Bergquist.** 1992. The β-mannanase from "*Caldocellum saccharolyticum*" is part of a multidomain enzyme. Appl. Environ. Microbiol. **58**:3864–3867.
10. **Gilbert, H. J., G. P. Hazlewood, J. I. Laurie, C. G. Orpin, and G. P. Xue.** 1992. Homologous catalytic domains in a rumen fungal xylanase: evidence for gene duplication and prokaryotic origin. Mol. Microbiol. **6**:2065–2076.
11. **Gilkes, N. R., B. Henrissat, D. G. Kilburn, R. C. Miller, Jr., and R. A. J. Warren.** 1991. Domains in microbial β-1,4-glycanases: sequence conservation, function and enzyme families. Microbiol. Rev. **55**:303–315.
12. **Gosalbes, M. J., J. A. Perez-Gonzalez, R. Gonzales, and A. Navarro.** 1991. Two β-glycanase genes are clustered in *Bacillus polymyxa*: molecular cloning, expression, and sequence analysis of genes encoding a xylanase and an endo-β-(1,3)-(1,4)-glucanase. J. Bacteriol. **173**:7705–7710.
13. **Henrissat, B.** 1991. A classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem. J. **280**:309–316.
14. **Henrissat, B.** 1992. Analysis of hemicellulase sequences. Relationships to other glycanases, p. 97–110. *In* J. Visser, G. Beldman, M. A. Kusters-van Someren, and A. G. J. Voragen (ed.), Xylan and xylanases. Progress in biotechnology, vol. 7. Elsevier, Amsterdam.
15. **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.
16. **Lloberas, J., J. A. Perez-Pons, and E. Querol.** 1991. Molecular cloning, expression and nucleotide sequence of the endo-β-1,3-1,4-D-glucanase gene from *Bacillus licheniformis*. Eur. J. Biochem. **197**:337–343.
17. **Matheson, N. K., and B. V. McCleary.** 1985. Enzymes metabolizing polysaccharides, p. 1–103. *In* G. O. Aspinall (ed.), The polysaccharides, vol. 3. Academic Press, New York.
18. **Morris, E. J., and N. O. Van Gylswyk.** 1980. Comparison of the action of rumen bacteria on cell walls of *Eragrostis tef*. J. Agric. Sci. **95**:313–323.
19. **Murphy, N., D. J. McConnell, and B. A. Cantwell.** 1984. The DNA sequence of the gene and genetic control sites for the excreted *B. subtilis* enzyme β-glucanase. Nucleic Acids Res. **12**:5355–5367.
20. **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.
21. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
22. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
23. **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.
24. **Schimming, S., W. H. Schwarz, and W. L. Staudenbauer.** 1992. Structure of the *Clostridium thermocellum* gene *licB* and the encoded β-1,3-1,4-glucanase. Eur. J. Biochem. **204**:13–19.
25. **Teather, R. M., and J. D. Erfle.** 1990. DNA sequence of a *Fibrobacter succinogenes* mixed-linkage β-glucanase ((1,3-1,4)-β-D-glucan 4-glucanohydrolase) gene. J. Bacteriol. **172**:3837–3841.
26. **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.
27. **Tinoco, I., P. N. Borer, B. Dengler, M. D. Levine, O. C. Uhlenbeck, D. M. Crothers, and J. Gralla.** 1973. Improved estimation of secondary structure in ribonucleic acids. Nature (London) New Biol. **246**:40–41.
28. **Von Heijne, G.** 1988. Transcending the impenetrable: how proteins come to terms with membranes. Biochim. Biophys. Acta **947**:307–333.
29. **Wilkie, K. C. B.** 1983. Hemicellulose. Chemtech **13**:306–319.
30. **Wong, K. K. Y., L. U. L. Tan, and J. N. Saddler.** 1988. Multiplicity of β-1,4-xylanase in microorganisms: functions and applications. Microbiol. Rev. **52**:305–317.
31. **Wood, T. M., A. Castanares, D. C. Smith, S. I. McCrae, and J. A. Brown.** 1992. Xylan-degrading enzyme systems of the fungi *Aspergillus awamari* and *Penicillium pinophilum*: studies on the isolation of over-producing mutant strains and aspects of the enzymology, p. 187–202. *In* J. Visser, G. Beldman, M. A. Kusters-van Someren, and A. G. J. Voragen (ed.), Xylans and xylanases. Progress in biotechnology, vol. 7. Elsevier, Amsterdam.
32. **Xue, G.-P., K. S. Gobius, and C. G. Orpin.** 1992. A novel polysaccharide hydrolase cDNA (*celD*) from *Neocallimastix patriciarum* encoding three multi-functional catalytic domains with high endoglucanase, cellobiohydrolase and xylanase activities. J. Gen. Microbiol. **138**:2397–2403.
33. **Yahata, N., T. Watanabe, Y. Nakamura, Y. Yamanto, S. Kamimiya, and H. Tanaka.** 1990. Structure of the gene encoding β-1,3-glucanase A1 of *Bacillus circulans* WL-12. Gene **86**:113–117.
34. **Zhang, J.-X., and H. J. Flint.** 1992. A bifunctional xylanase encoded by the *xynA* gene of the rumen cellulolytic bacterium *Ruminococcus flavefaciens* 17 comprises two dissimilar domains linked by an asparagine/glutamine-rich sequence. Mol. Microbiol. **6**:1013–1023.