MOLECULAR ANALYSIS OF TWO CELLULASE GENES FROM *RUMINOCOCCUS FLAVEFACIENS* FD-1 AND THEIR TRANSCRIPTIONAL REGULATION

WENYEN WANG

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Faculty of Science, University of Cape Tomn.

CAPE TOWN

NOVEMBER 1992

CERTIFICATION OF SUPERVISOR

In terms of paragraph 9 of "General regulations for the degree of Ph.D." I, as supervisor of the candidate, W. Wang, certify that I approve of the incorporation in this thesis material that has already been published or submitted for publication.

Signed by candidate

Profe or J. . Thomson
Head of Department of Microbiology
University of Cape Town

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Professor Jennifer A. Thomson, for her unfailing guidance and encouragement during the course of this study, and for her giving me the chance to see this beautiful country.

I am grateful to Dr. Shez Reid for her helpful discussions and Di James for assisting with the DNA sequences of *celA* and *endA*. I appreciate greatly the many hours spent by G. Wilson in preparation of manuscripts for publication.

Many thanks to A. Jaffray, D. DE Villiers, S. Botman and C. Hendrickse for their help during the course of this study.

I gratefully acknowledge F.-P. Lin and L.-L. Lin for their friendship and helpful dicussions.

A very special thanks to Judy and my parents for always being there for me.

I am grateful for the financial support which I received from the F.R.D., South Africa and the working environment of Department of Microbiology, UCT.

CONTENTS

ABSTRACT

The mesophilic *Ruminococcus flavefaciens* FD-1 (NCDO 2215) is a Gram-positive obligate anaerobic bacterium. The aim of this thesis was to clone, sequence and analyze a cellodextrinase gene (*celA*) and a carboxymethylcellulase gene (*celE*), and study their regulation and induction at the transcriptional level.

The sequence of the *celA* gene from FD-1 was determined and the amino acid sequence of the CelA enzyme (336 amino acid residues) deduced. It showed 40% identity with endoglucanase C of *Clostridium thermocellum* and 27.4% identity with endoglucanase 3 of *Fibrobacter succinogenes*. These three enzymes are grouped into subfamily "A3". The ATG start codon of *celA* is preceded by a GAGG sequence, predicted to be a ribosome binding site. The derived amino acid sequence corresponded to a protein of M_r 38686. SDS-PAGE analysis of *in vitro* and *in vivo* translational products showed that CelA has a molecular mass of ca 39 kDa and was secreted into the *Escherichia coli* periplasmic space. Although CelA has activity on carboxymethylcellulose, further study on the enzyme showed that it degraded cellopentaose and other cellodextrins to predominantly cellobiose. Thus CelA is a cellodextrinase. It also has high activity against *p*-nitrophenyl-ß-D-cellobioside.

A gene, expressing a protein with both carboxymethyl cellulase and xylanase activity, was cloned from *R. flavefaciens* FD-1 using an *E. coli/Bacillus subtilis* shuttle vector, pEB1. The 3.6 kb DNA insert on the plasmid pWF1, which carried the gene, *cel*E, contained an open reading frame of 963 bp encoding 320 amino acid residues with a caculated M_r of 35937. Homology analysis showed 11.6% identity and 55.3% similarity with the N-terminal catalytic region of the cellulase gene of alkalophilic *Bacillus sp.* strain 1139. In order to obtain expression in *E. coli*, the gene had to be transcribed from the lambda P_r promoter.

To determine whether cellulase genes of *R. flavefaciens* FD-1 were regulated at the level of transcription, *cel*A and *cel*E were used as probes against RNA isolated from *R. flavefaciens* FD-1 grown on cellobiose, cellulose or cellotriose. Transcription of both genes was induced when cellulose was added to cells growing in cellobiose. This induction continued after cellulose depletion and after cell division had ceased. Transcription of both genes was also induced by cellotriose although not to the same extent as by cellulose. This suggests that cellotriose and possibly other dextrins may act as key inducers to trigger *cel*A and *cel*E gene expression in *R. flavefaciens* FD-1.

ABBREVIATIONS

A	adenosine
Abs.	absorbance
Ap	ampicillin
ATCC	American Type Culture Collection
ATP	adenosine 5'-triphosphate
bp	base pair
BSA	bovine serum albumin
C-terminal	carboxy terminal
CBH	cellobiohydrolase
Cm	chloramphenicol
CsCl	caesium chloride
CMC	carboxymethylcellulose
CMCase	carboxymethylcellulase
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DNS	dinitrosalicylic acid
DTT	1,4-dithio-L-threitol
EDTA	ethylenediaminetetra-acetic acid
EG	endoglucanase
EtBr	ethidium bromide
h	hour(s)
HCA	hydrophobic cluster analysis
HEC	hydroxyethylcellulose
IPTG	isopropyl-β-D-thiogalactopyranoside
kb	kilobase pairs
Km	kanamycin
l	liter
LB	Luria-Bertani broth
MeUMb	methylumbelliferyl
min	minute(s)
mRNA	messenger RNA
M_r	relative molecular mass
NRF	non-rumen fluid medium for *R. flavefaciens* FD-1
NRF(B)	non-rumen fluid medium with cellobiose
NRF(T)	non-rumen fluid medium with cellotriose
NRF(L)	non-rumen fluid medium with cellulose
N-terminal	amino terminal
OD_{600}	optical density at 600nm
ONPG	*o*-nitrophenyl-β-D-galactopyranoside
ORF	open reading frame

ABBREVIATIONS (cont.)

p	plasmid
PAGE	polyacrylamide gel electrophoresis
PC	phosphate-citrate buffer
*pho*A	gene coding for alkaline phosphatase
pNPG	p-nitrophenyl-β-D-glucopyranoside
pNPC	p-nitrophenyl-β-D-cellobioside
P_R	rightward promoter(phage lambda)
r	resistance (superscript)
RNA	ribonucleic acid
s	second(s)
SDS	sodium dodecyl sulfate
sp(p)	species
TAE	triacetate EDTA buffer
TEMED	*N,N,N',N'*-tetraethylethylenediamine
Tn	transposon
Tris	Tris(hydroxymethyl)aminomethane
U	units of enzyme activity
UV	ultraviolet (light)
v/v	volume/volume
w/v	weight/volume
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
XP	5-bromo-4-chloro-3-indolyl phosphate
::	novel joint (fusion)
()	designates plasmid-carrier state
α	alpha
β	beta
Δ	delta
λ	lambda
μ	micro

CHAPTER 1

GENERAL INTRODUCTION

CHAPTER 1

GENERAL INTRODUCTION

In our search for future sources of food for an expanding population, herbivorous animals are very important. Herbivores, especially ruminants, need not compete with humans for primary plant foods or land suitable for production of these crops. Ruminants can eat fodder and grasses due to the fact that they possess a rumen in which a variety of anaerobic bacterial, fungal and protozoal species degrade the fibre, releasing compounds which the animals can assimilate. Hence the ruminant, equipped with its unique microbial fermentation vat, can serve as a vital link between humans and the most abundant storehouse of organic energy on earth, lignocellulose. However, this degradation process is far from perfect in that approximately 40% of the fibre passes through the animal undigested (Hungate, 1988). In order to utilize this enormous and self-replenishing reservoir of energy, an understanding of the microorganisms involved in lignocellulose breakdown is required.

Cellulose, the major structural component of higher plants, is the most abundant organic compound and renewable resource in the world. Of the yearly 30 billion tons of carbon transformed into organic compounds by higher plants, about one-third is converted into cellulose. Fifty percent of the

wood we use to build our houses is cellulose; a similar percentage of paper is cellulose; more than 90% of the cotton in our clothes is cellulose; and perhaps one-fourth of the diet of our livestock is cellulose (Albersheim, 1965b). Therefore the synthesis, biodegradation and utilization of this polymer has attracted considerable attention. In particular, microbial cellulases are being studied, as they are responsible for most of the biodegradation of cellulose to soluble sugars. Microbial degradation of cellulose is of prime importance in balancing the world's carbon cycle, producing 85 billion metric tons of carbon dioxide per year (Van Soest, 1973; Coughlan, 1985). Due to the growing need for energy, food, chemicals, and the possibility of enzymatically converting the cellulose in forest, agricultural and municipal wastes into useful products to offset oil shortages, the role of cellulolytic microorganisms has attracted a great deal of attention in recent years (Ljungdahl and Eriksson, 1985).

1.1 Structure of lignocellulose.

Lignocellulose, the major carbohydrate constituent of forage, is a collective term for the three major components: cellulose, hemicellulose, and lignin (Van Soest, 1973; Hrmova et al., 1991). Lignin, surrounding the cellulose fibrils, is a highly branched and constitutionally undefined aromatic polymer composed of phenylpropane subunits. It is very recalcitrant to hydrolysis and probably impedes the degradation of lignocellulose (Albersheim, 1965b; Van Soest, 1973; Coughlan, 1985). It was reported that the ability of fungi to degrade lignin depends on both the fungal species and the type of wood. The degradation of lignin by white-rot fungi, such as *Polyporus versicolor*, and soft-rot fungi, such as *Aspergillus fumigatus* is dependent on the presence of oxygen (Amer and Drew, 1980). In addition, the biochemistry and enzymology of the lignin degradative pathway in white-rot fungi such as *Phanerochaete chrysosporium* has been well described (Kirk et al., 1986). The heme-containing ligninases catalyze the cleavage of carbon bonds in the presence of hydrogen peroxide (Kirk et al., 1986). It was concluded that lignin degradation is essentially an aerobic process involving oxidative enzymes or oxidising reagents and is unlikely to occur to any significant extent under the reducing conditions found in the rumen (Chesson and Forsberg, 1988).

Hemicelluloses are a heterogeneous mixture of water-insoluble structural polysaccharides in plant cell walls. The composition of hemicelluloses from several different sources is shown in Table 1.1 (Dehority, 1973). Hemicelluloses can be extracted from plant tissue after removal of lipid and lignin (Whistler and Richards, 1970). The major class of hemicellulose is xylan, which is found in large quantities in annual plants (Gong et al., 1981). It has been shown that β-1,4-xylans are heterogeneous polysaccharides found in the cell walls of all land plants and in almost all plant parts, and that the backbone of β-1,4-xylans consists of β-1,4-linked D-xylosyl residues (Wong et al., 1988). Isolated β-1,4-xylans are generally polydispersed (i.e., varying degree of polymerization) and highly branched heteropolymers with more than one type of substituent. The common substituents found on the backbone of β-1,4-linked D-xylopyranosyl residues are acetyl, arabinosyl and glucuronosyl residues. The frequency and composition of branches in isolated xylans are dependent on their source and the method of their isolation (Wong et al., 1988). The major hemicellulosic constituents of lignocellulosic waste materials are the hetero-1,4-β-D-xylans (arabino-4-O-methylglucuronoxylans and 4-O-methylglucuronoxylans) and the hetero-1,4-β-D-mannans (galactoglucomannans and glucomannans). The heteroxylans constitute the major hemicellulosic component of the Gramineae (grasses and cereals) and Angiosperms (hardwoods),

while the β-mannans are more abundant in Gymnosperms (softwoods) (Dekker, 1983).

Table 1.1 Composition of hemicelluloses from several different sources (Dehority, 1973).

Component	Source of hemicellulose[a]					
	Afalfa	Corn hull	Flax	Fescue	Oat hull	Bromegrass
Arabinose	10.4	35.0	-	20.2	14.0	12.0
Xylose	58.5	48.0	84.7	61.7	66.0	59.3
Glucose	6.9	-	-	9.8	12.0	20.9
Galactose	6.9	7.0	-	8.2	5.0	7.8
Rhamnose	3.9	-	1.5	-	-	
Glucuronic acid	13.5	10.0	13.8		+	

[a] Data presented as % of total hemicellulose; blank spaces indicate that the component was not determined; - indicates the component was not present; + indicates that the component was present but only in trace amounts.

Cellulose is an unbranched polymer of several thousand D-glucose units joined by β-1,4-glycosidic linkages, and is insoluble in water. Most of the world's celluloses are unused and present in plants that either live and die untouched or are used as timber or fuel (Wilkie, 1983). Electron microscopic examination indicates that the crystalline microfibril formed by cellulose chains is the basic structural unit of the cell wall. Cellulose chains tightly crosslink adjacent chains within a microfibril by intermolecular hydrogen bonds. Such microfibrils are the only nonamorphous component of the wall and thus the only microscopically discrete component within the wall (Albersheim, 1965a; 1965b; Chesson and Forsberg, 1988).

Bundles of the above fibrils aggregate to form the inert, insoluble fibres of great strength characteristic of the primary and secondary cell walls of higher plants. Cellobiose is a disaccharide produced upon partial hydrolysis of cellulose and contains the β-1,4-bond, while complete hydrolysis results in release of glucose (Albersheim, 1965a; 1965b; Van Soest, 1973; Ljungdahl and Eriksson, 1985; Coughlan, 1985). In addition to its usual native crystalline form (cellulose I), cellulose can exist in a variety of alternative crystalline forms (allomorphs) which differ in their unit cell dimensions, chain packing schemes, and hydrogen bonding relationships (Weimer et al., 1990).

The proportion of cellulose in plant cell walls varies widely among different species. Broad-leafed plants, including the legumes, tend to be highly lignified while containing a low proportion of hemicellulose. Grasses have a distinctly higher proportion of hemicellulose and lower lignin-to-cellulose ratios. Various forage species are shown in Table 1.2 (Albersheim, 1965a; 1965b; Van Soest, 1973).

Table 1.2 Proportion of hemicellulose, cellulose and lignin in the cell wall of perennial forages (Van Soest, 1973).

Forage			Number forages	Cell wall	Hemi-cellulose (H)	Cellulose (C)	Lignin	Ratio H/C
Genus	Species	Common name			% of cell wall			
	Legumes							
Medica	*gosativa*	alfalfa	14	51%	22	56	21.7	0.39
		leaf	22	20%	30	51	18.2	0.58
		stem		54%	17	64	19.2	0.26
Trifolium	*pratense*	red clover	1	66%	26	59	15.1	0.44
	repens	ladino clover	1	28%	14	68	17.6	0.21
	alexandrinum	persian clover	1	53%	16	65	18.9	0.25
	Grasses Temperate							
Bromus	*inermis*	brome	7	64%	42	50	7.5	0.84
Dactylis	*glomerata*	orchard	3	55%	45	45	7.5	1.00
Festuca	K-31	fescue	4	54%	43	46	7.2	0.93
Phalaris	*arundinacea*	Reed canary	2	62%	40	51	7.5	0.78
Phalaris	*arundinacea*	Reed canary	8	59%	47	49	6.9	0.96
Phleum	*pratense*	timothy	2	57%	45	49	4.9	0.92
	Subtropical and Tropical Grasses							
Cenchrus		Klein	2	66%	53	37	7.7	1.43
		leaf	1	65%	60	34	5.8	1.75
		stem	1	80%	42	45	12.8	0.94
Cynodon	*dactylon*	Bermuda	4	76%	51	38	7.9	1.34
	plectostachyus	star	10	68%	44	45	10.0	0.98
		leaf	1	66%	50	40	7.5	1.24
		stem	1	73%	40	49	11.1	0.81
Digitari	*decumbens*	pangola	11	68%	38	48	11.1	0.79
		leaf	7	68%	38	50	12.9	0.76
		stem	7	70%	37	49	17.2	0.75
Panicum	*maximum*	guinea	2	66%	37	50	8.0	0.74
Pennisetum	*purpureum*	napier	4	63%	34	51	10.4	0.67

1.2 Cellulose degradation by rumen bacteria.

Cellulose utilization by ruminants depends on the synergistic action of rumen microorganisms (Leatherwood, 1973). Structural features such as the degree of crystallinity and available surface area are known to affect the digestion kinetics of cellulose by ruminal microorganisms (Weimer et al., 1990).

Differential fermentation of cellulose allomorphs by ruminal cellulolytic bacteria showed a variety of degradation rates (Weimer, et al., 1990; Weimer, et al., 1991). However various rumen bacteria found in clarified rumen fluid from cows and sheep fed different diets had comparable growth rates (Van Gylswyk, et al., 1992). It has been observed that readily digestible parenchyma cell walls are heavily colonised by bacteria and rapidily digested, whereas the thick, recalcitrant, walls of the vascular and sclerenchymal tissue of the same plant are sparsely colonised (Chesson and Forsberg, 1988). Based on relative numbers in the rumen and relative abilities as pure cultures to attack various forms of cellulose, the four major cellulolytic species are *Ruminococcus flavefaciens*, *R. albus*, *Fibrobacter succinogenes* (formerly *Bacteroides succinogenes*) and *Butyrivibrio fibrisolvens* (Bryant, 1973; Chesson and Forsberg, 1988). All four species have the ability to grow on cellulose as their sole energy and carbon source (Bryant,

1973). Electron microscopic investigations have also demonstrated that although many different morphological types of rumen bacteria adhere to forage fragments, coccoid bacteria are the most common (Akin, 1980). *R. flavefaciens* strains are often found to be more active against crystalline cellulose than strains of *R. albus* (Bryant, 1973; Cunningham et al., 1991). A few other species including *Cillobacterium cellulosolvens* and members of the genus *Clostridium* have occasionally been found, but they appear to be much less important (Bryant, 1973). It has been indicated that the microflora of the rumen of high-arctic Svalbard reindeer are highly effective in fiber digestion and nitrogen metabolism for different food quality and availability during the summer and winter, allowing the animals to survive under the austere nutritional conditions in this habitat (Orpin et al., 1985).

1.3 Adhesion of rumen microorganism to substrates.

It is difficult for an aquatic microorganism found in the rumen to ensure its food supply and to avoid being passively removed by the normal flow of the liquid medium. The ability of cell wall degraders to adhere to plant material is of primary importance and appears an essential first stage in the digestive process (Chesson and Forsberg, 1988). Rumen fungi are proteolytic, unlike the rumen cellulolytic bacteria, and their proteolytic action probably facilitates

the penetration of the proteinaceous layer which prevents cellulolytic bacteria from gaining access to the secondary cell wall (Orpin and Joblin, 1988). Early reports showed that *Cellvibrio fulvus* and *Sporocytophaga myxococcoides* attack cellulose by growing in close contact with the substrate (Berg et al., 1972). However, whether anaerobic fungi penetrate plant cell walls by mechanical force, enzymic action or a combination of the two is unknown (Orpin and Joblin, 1988). It has been demostrated that fungal zoospores attach and produce hyphae that grow and penetrate plant tissues in the rumen of cattle and sheep (Bauchop, 1979). Autolysis of hyphae may be expected to release cellulolytic enzymes which cause erosion at a considerable distance from the growing cells (Berg and Von Hofsten, 1976). These findings strongly suggest that fungi may be reponsible for the primary hydrolysis of lignocellulose-containing plant fibres in some ruminants.

The mechanism of attachment of different rumen bacteria to cellulose may involve specific binding by cell surface-associated enzymes, adhesins (molecules on the microbial cell surface that bind to receptors on the plant material), or possibly non-specific ionic interaction (Chesson and Forsberg, 1988). The ultrastructure and adhesion properties of *R. albus* suggest that adhesion is essential for cellulose hydrolysis (Patterson, et al., 1975; Morris and Cole, 1987; Morris, 1988). *R. flavefaciens* has also been shown to

adhere to cotton cellulose by the cell coat outside the cell wall (Latham, et al., 1978; Bryant, 1986). In *F. succinogenes* adhesion is enhanced by increasing the ionic strength, and markedly reduced by treatment with heat, glutaraldehyde, trypsin, and pronase. However, the degree of adhesion can not be decreased by treatment with dextrinase, the modification of either the amino or carboxyl groups of proteins with formalin or other chemical agents, or the inclusion of either BSA (1%) or Tween 80 (0.5%) (Gong and Forsberg, 1989). These results suggest that adhesion of the bacteria to cellulose requires a specific protein(s) with a unique conformation, but neither the amino nor the carboxyl group of the protein plays an important role in adhesion.

A number of studies have been carried out to determine the effect of a substrate which is resistant to biological degradation in anaerobic environments, such as methylcellulose (MC), on adherence to and hydrolysis of cellulose. Kudo et al. (1987) found that it prevented adhesion of *F. succinogenes* to cellulose, and prevented cellulose digestion when added at a low concentration, although it had little inhibitory effect on cellulolytic activity. However, earlier reports showed that this organism's endoglucanase activity was inhibited by MC (Groleau and Forsberg, 1983) and MC acts as a competitive inhibitor of endo- and exocellulase activity in *R.*

flavefaciens FD-1 (Rasmussen et al., 1988). In addition, electron microscopic analysis showed that MC caused detachment of various cellulolytic ruminal bacterial from cellulose fibers (Kudo et al, 1987; Chesson and Forsberg, 1988). In addition to MC, the adherence of *R. flavefaciens* FD-1 to cellulose is also inhibited by carboxymethyl cellulose (CMC) (White et al, 1988). Thus it might be speculated that MC and other methylated carbohydrates could affect cellulose hydrolysis by ruminal bacteria by two mechanisms. First, the compounds could act as inhibitors of enzyme activity. Second, they could act as surfactants that alter protein structure or block the receptors responsible for adherence (White et al, 1988). Methylated cellulose derivatives may therefore be useful substrates for the determination of bacterial adherence and enzyme mechanisms involved in the hydrolysis of cellulose (White et al, 1988).

1.4 Cellulase systems.

The cellulase systems produced by cellulolytic bacteria are very complicated. The classification of cellulases is based on the activities detected on various substrates. "Complete" cellulases can be detected by their hydrolysis of cotton fibre, filter paper, Solka floc, dyed cellulose, cellulose-agar and Avicel (microcrystalline α-cellulose), which are highly hydrogen-bonded structures. It is generally acknowledged that "complete" cellulases are

composed minimally of three main types of enzymes (Coughlan, 1985; Wood and Kellogg, 1988; Beguin, 1990; Gilkes et al., 1991). Endocellulases (EC. 3.2.1.4) degrade insoluble cellulose polymers by random internal cleavage at β-1,4 bonds along the interior of the cellulose chain. They produce reducing and non-reducing ends while also creating short water soluble chains of up to 7 glucose moieties, called cellodextrins. These enzymes are also referred to as 1,4-β-D-glucan glucanohydrolases, endo-β-1,4-glucanases, CMCases or C_x cellulases. Substrates used for their detection include CMC, hydroxyethyl cellulose and CMC-agar.

Exocellulases (EC. 3.2.1.91) cleave crystalline cellulose from the non-reducing end of the chain but show limited activity towards CMC. They are also referred to as exo-1,4-β-D-glucan cellobiohydrolases, exo-β-1,4-glucanases, avicelases, C_1 cellulases or cellobiohydrolases. Substrates used for their detection include amorphous cellulose, cellodextrins and pNPC.

β-glucosidases (EC.3.2.1.21) hydrolyze β-1,4-glucosides such as cellobiose to glucose or glucose plus aromatic residues. They are also referred to as cellobiases or aryl-glucosidases. Substrates used for their detection include cellodextrins, cellobiose and *p*-nitrophenyl-β-D-glucoside (pNPG).

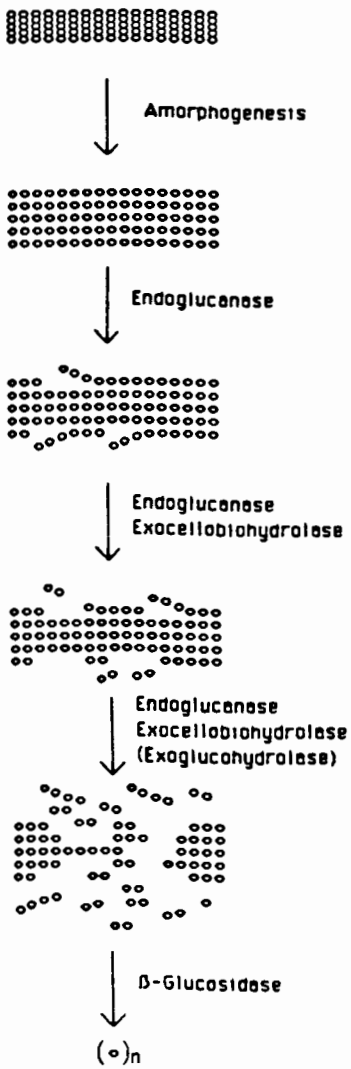
In addition cellodextrinases hydrolyze soluble cellodextrins. They have been found to be cell-associated, and may be located in the periplasm of some bacteria (Huang and Forsberg, 1987; Wang and Thomson, 1990).

Cellulosomes, cell surface-associated complexes of enzymes and proteins responsible for adhesion to and hydrolysis of cellulose, have been found in *Clostridium thermocellum* and some rumen bacteria (Lamed and Bayer, 1988; Cavedon et al., 1990; Morag et al., 1991). During the early stages of growth of *C. thermocellum* on cellulose, the cellulosomes form polycellulosomes which appear as spherical particles on the cell surface and bind the cells to the cellulose. These polycellulosomes have been estimated to have a mass of 50 x 10^6 to 80 x 10^6 Da (Mayer et al., 1987). Later the cells detach from the cellulose and free cellulosomes appear in the culture supernatant. The free cellulosomes may dissociate, at least in part, and bind to and act on crystalline cellulose (Mayer et al., 1987; Lamed and Bayer, 1988). Recently it was reported that the small subunit of the cellulosome of *C. thermocellum* YS acts as a major cellobiohydrolase (Morag et al., 1991). It has also been reported that the cellulase complex of *C. cellulovorans* is a multiprotein complex of large molecular mass, and the enzymes are probably organized on a nonenzymic scaffolding protein whose gene has been cloned and sequenced (Shoseyov and Doi, 1990).

1.5 Mode of action of cellulolytic enzymes.

The different cellulase enzymes with different specificities and modes of actions are thought to act synergistically in the hydrolysis of the insoluble substrate to produce short oligosaccharides, cellobiose and/or glucose (Leatherwood, 1973; Coughlan, 1985; Ljungdahl and Eriksson, 1985). The commonly accepted model for the enzymatic hydrolysis of cellulose is one in which endoglucanases initiate the attack on the substrate. The endwise-acting exocellulases then act on the newly formed non-reducing chain ends. Eventually, the β-glucosidases hydrolyse the cellobiose to glucose (Fig. 1.1A). Competitive adsorption studies led to the postulate that the endoglucanases and cellobiohydrolases adsorb to distinctly different sites on cellulose, corresponding to their sites of hydrolysis (Fig. 1.1B; Coughlan, 1985). Endoglucanases binding to cellulose speeds up the rate of scission by cellobiohydrolase (or endoglucanase) and brings about its desorption (Coughlan, 1985). This also explains the synergistic interaction in cellulose hydrolysis by these enzymes (Coughlan, 1985; Henrissat et al., 1985).

(A)



(B)

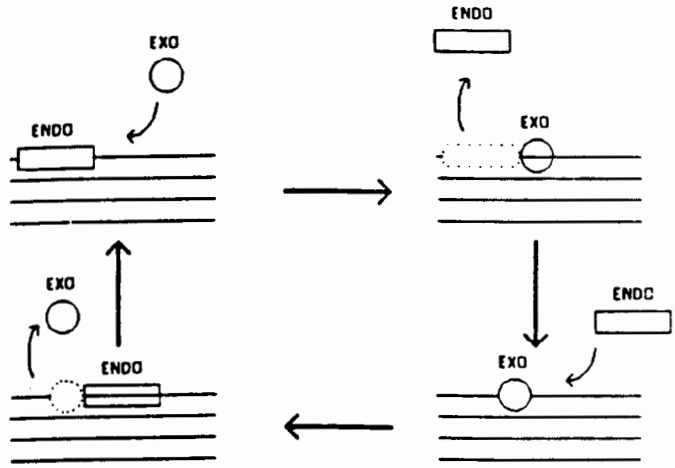
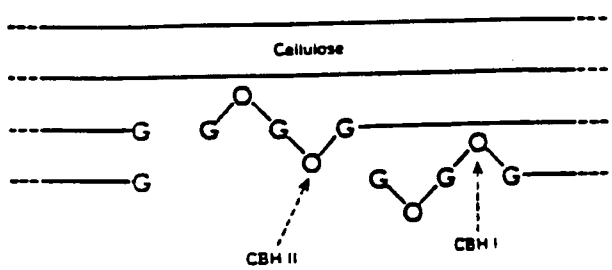
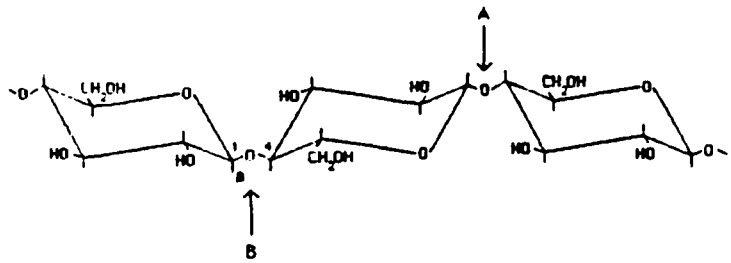


Fig. 1.1
(A) Proposed mechanism for hydrolysis of cellulose at the macromolecular level by complete cellulolytic enzyme systems. (B) Competition between endo- and exoglucanases for adsorption to substrate. (Coughlan, 1985)

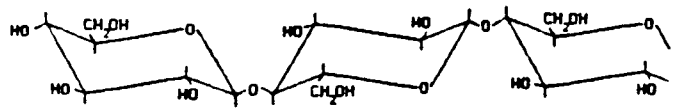
(A)



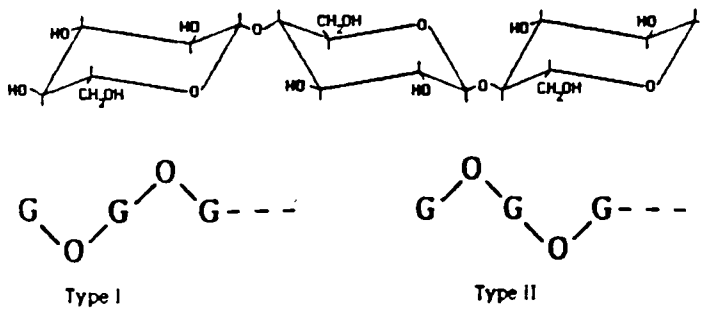
(B)



Cleavage at A gives end group type I



Cleavage at B gives end group type II



(C)

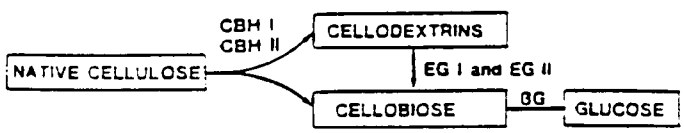


Fig. 1.2 (A) Proposed explanation for exo-exo synergism in the hydrolysis of cellulose (Wood, 1985). (B) Proposed mechanism for the synergistic hydrolysis of cellulose by endo- and exoglucanases and of the need for multiple forms of each enzyme. This in turn means that at least two stereospecific types of exocellobiohydrolase will be required (Coughlan, 1985). (C) Suggested scheme for the enzymatic hydrolysis of cellulose (Enari and Niku-Paavola, 1987).

The second model proposed is for the exo-exo synergism of cellobiohydrolases I and II of *Trichoderma reesei* during cellulose degradation (Fig. 1.2A) and the synergistic action of endo- and exoglucanases (Fig. 1.2B). For the different substrate stereospecificities cellobiohydrolases I and II attack one of the two different types of non-reducing end groups that may be found in the substrate. The cleaved glycosidic linkages of cellulose would re-form rapidly if endoglucanases were to act alone. However this is prevented by the exocellobiohydrolases which remove cellobiose units from the chain ends provided by the endoglucanases. The successive operation of each enzyme on adjacent chains could explain the observed synergism. Cellodextrinases may participate at a later stage by hydrolyzing cellodextrins. β-glucosidases finally cleave the cellobiose into glucose (Fig. 1.2C; Coughlan, 1985; Wood, 1985; Enari and Niku-Paavola, 1987).

The third model proposed (Fig. 1.3) is for phosphate dependent phosphorolysis by cellobiose phosphorylase (EC. 2.4.1.20). Some bacteria grow well on cellobiose, cellodextrins and cellulose, but less well on glucose. It has been suggested that cellotriose and higher cellodextrins are metabolized by transglucosylation to yield cellodextrins and cellobiose (Ayers, 1958; Ayers, 1959; Sasaki et al., 1983; Ljungdahl and Eriksson, 1985). Cleavage of cellobiose is catalyzed by cellobiose phosphorylase to glucose-1-

phosphate and glucose (Heale and Gupta, 1971; Schmiz et al., 1983; Barras et al., 1984). Cellobiose phosphorylase activity was detected in *R. flavefaciens* by both a coupled spectrophotometric assay and a discontinuous assay (Helaszek and White, 1991).

$$\text{Cellodextrin } (G_n) + \text{glucose} \xrightarrow{\text{transglucosylation}} \text{cellodextrin } (G_{n-1}) + \text{cellobiose}$$

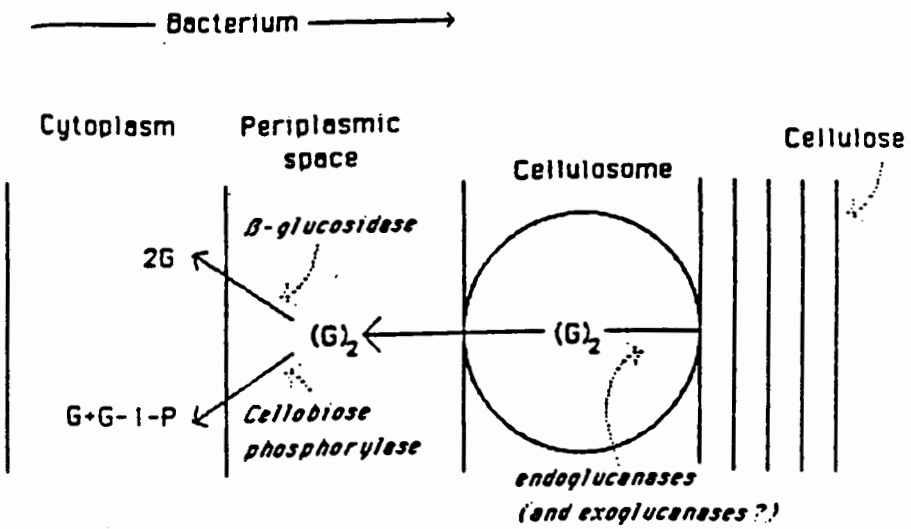
$$\text{Cellobiose} + \text{phosphate} \xrightarrow{\text{cellobiose phosphorylase}} \text{glucose-1-P} + \text{glucose}$$

Fig. 1.3 Proposed mechanism for transglucosylation of cellodextrins and cleavage of cellobiose by cellobiose phosphorylase (Ljungdahl and Eriksson, 1985).

The fourth model proposed is based on the cellulosome which mediates both adherence of the bacterium to the substrate and subsequent degradation (Fig. 1.4A). It has been proposed that the hydrolysis of cellulose by the cellulosome of *C. thermocellum* involves the binding and alignment of the catalytic subunits of the cellulosome at regular intervals along the cellulose fiber, assumed to be mediated by a cellulose-binding factor carrying reiterated anchoring sites for the enzymes. Simultaneous hydrolysis at multiple sites generates small fragments which no longer bind to the fibre lattice (Fig. 1.4B). Electron microscopic studies confirmed that polycellulosomes are located on the cell surface and make contact with cellulose fibers. They decompose and ultimately form free polypeptides in the culture medium

(Mayer et al., 1987; Lamed and Bayer, 1988; Sprey and Bochem, 1991).

(A)



(B)

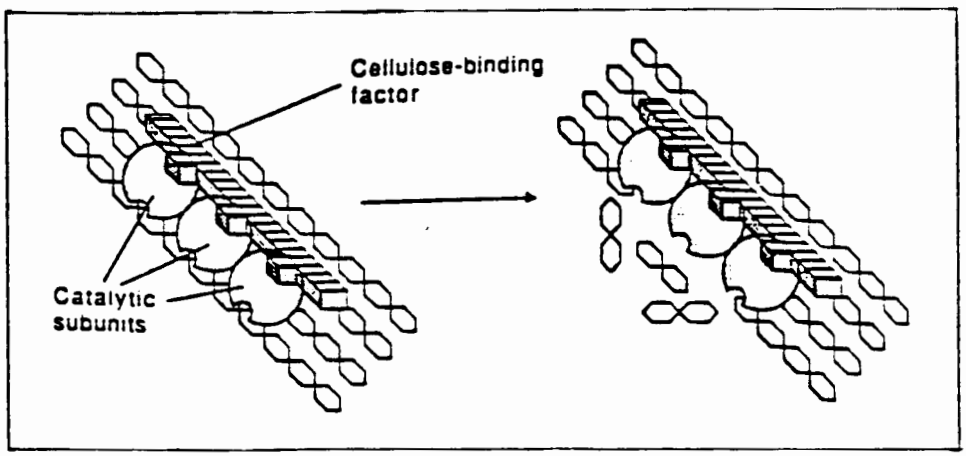


Fig. 1.4 (A) Hydrolysis of cellulose by bacterial cellulolytic enzyme systems (Coughlan, 1985). (B) Hypothetical model for the hydrolysis of cellulose by the cellulosome of *C. thermocellum* (Mayer et al., 1987).

1.6 Enzyme inhibition by end product(s).

Cellulase activity can be affected by end products of hydrolysis. In the *Cellulomonas* genus cellobiose, but not glucose, inhibited enzyme activity on both cellulose and carboxymethycellulose (Stewart and Leatherwood, 1976). In *C. thermocellum* YS, the activity of the cellulosome can be activated by calcium ions and thiol reagents but severely inhibited by the end product, cellobiose (Chauvaux et al., 1990; Morag et al., 1991). Activity of the cellodextrinase (CelA) of *R. flavefaciens* FD-1 was also inhibited in the presence of divalent metal cations and cellobiose, the major end product of cellodextrin hydrolysis (Wang and Thomson, 1990; Brown et al., submitted). When *F. succinogenes* subsp. *succinogenes* S85 was grown in microcrystalline cellulose medium combined with either glucose or cellobiose, there was a lag in cellulose digestion until all of the soluble sugar had been utilized (Huang and Forsberg, 1990). One explanation for this is end product inhibition. The other is that glucose and cellobiose are more readily metabolizable than cellulose.

1.7 Structural domains and functions.

The amino acid sequences of cellulases have been compared by amino acid alignment and hydrophobic cluster analysis (HCA; Devereux et al., 1984; Henrissat et al., 1989). The

structure of many enzymes is composed of domains that are more or less conserved in different proteins (Beguin, 1990; Gilkes, et al., 1991). Catalytic domains behave independently of noncatalytic domains as gene deletion or protease truncation can cleave off the latter without loss of catalytic activity. Noncatalytic domains sometimes contain cellulose-binding domains and are commonly connected to the catalytic domains by protease-susceptible linkers (Knowles et al., 1987; Ong et al., 1989).

The enzymes were grouped into eight families based on their conserved catalytic domains (Henrissat et al., 1989; Beguin, 1990; Gilkes, et al., 1991). The multiple cellulase genes of *C. thermocellum* belong to at least four distinct families and most genuinely cellulolytic microbes appear to possess cellulase genes belonging to several different families. The three-dimensional structure of the catalytic core of the endoglucanase (CelD) has a globular structure with an amino-terminal immunoglobulin-like domain tightly packed against a larger catalytic domain which is shaped like an α-barrel of 12 helices connected by loops that form the active site (Juy et al., 1992). The three-dimensional structure of the catalytic core of cellobiohydrolase II (CBHII) from the fungus *T. reesei* is located at the carboxyl-terminal end of a parallel β barrel in an enclosed tunnel. Two aspartic acid residues, located in the center of the tunnel, are

predicted to be the catalytic residues (Rouvinen et al., 1990).

Cellulase enzymes expressing cross-specificities have been reported. Endoglucanase A from *R. albus* SY3 exhibited both endoglucanase and xylanase activities (Poole et al., 1990). Avicelase I of *C. stercorarium* has been characterized as an unusual cellulolytic enzyme combining both endoglucanase and exoglucanase activities (Bronnenmeier et al., 1991). Endoxylanase 2 from *F. succinogens* S85, in addition to its xylan-hydrolyzing activity, showed significant activity on CMC, although it was unable to hydrolyze amorphous (acid-swollen) cellulose (Matte and Forsberg, 1992). Xylanase A of *C. acetobutylicum* ATCC 824 was able to degrade CMC, acid-swollen cellulose, and lichenan but had no activity on laminarin. The activity on cellulose and lichenan reflects the nonspecific nature of this enzyme toward β-1,4 linkages (Lee et al., 1987). A bi-functional cellulase/xylanase clone from *R. flavefaciens* FD-1 has also been reported (Howard and White, 1990).

Cellulose binding domains (CBDs) have a low content of charged amino acids, a high content of hydroxy amino acids and conserved tryptophan, asparagine, and glycine residues (Knowles et al., 1987; Ong et al., 1989; Beguin, 1990; Gilkes, et al., 1991). These domains can enhance the affinity of an enzyme for its substrate and influence its

rate of hydrolysis by as yet unknown mechanisms (Gilkes et al., 1988; Watanabe et al., 1990; Durrant et al., 1991; Meinke et al., 1991; Poole et al., 1991). Deletion of the CBD present in endoglucanase A (EGA) or B (EGB) from *Pseudomonas fluorescens* subsp. *cellulosa* showed no significant effect on the catalytic properties of either protein (Gilbert et al., 1990; Poole et al., 1991). Shoseyov et al. (1992) demonstrated that a large cellulose binding protein (CbpA) of *C. cellulovorans* with no apparent enzyme activity is capable of binding to crystalline cellulose. They suggested that CbpA association with enzyme subunits is necessary for true cellulase activity. It has also been reported that although endoglucanase A (CenA) and B (CenB) from *C. fimi* contain independently functioning CBDs, a sequence of 130 amino acids at the C-terminus of the catalytic domain of CenB can also bind to cellulose and may form a second cellulose binding site on that enzyme (Meinke et al., 1991; Miller et al., 1992).

The linkers (or hinges) between catalytic and non-catalytic domains are variable in length (6 to 59 amino acids) and rich in proline and hydroxyamino acids (Knowles et al., 1987; Ong et al., 1989; Beguin, 1990; Gilkes, et al., 1991). Endoglucanase B from *C. fimi* can be divided into five domains by linker sequences (Meinke et al., 1991), while removal of the linker between the catalytic domain and the CBD of endoglucanase A from *C. fimi* changed the structure

and the properties of both domains, as well as the protease sensitivity of the molecule (Shen et al.,1991). It may be speculated that the function of linkers is to contribute towards enzyme stability and possibly to allow flexibility between the catalytic and cellulose binding domains.

Repeated sequences ranging in length from 20 to 150 amino acids occur in a number of cellulases (Gilkes, et al., 1991). The function(s) of the repeated sequences is unknown, and appears irrelevant to the catalytic activity (Hall et al., 1988; Yague et al., 1990). However loss of the reiterated domains of endoglucanase A from *C. cellulolyticum* H10 enhanced its endo-type characteristics and decreased its catalytic properties on crystalline cellulose. In this case, the presence of the reiterated domains may facilitate the degradation of crystalline cellulose. However it is unclear whether this effect was a property of the domains themselves or an indirect effect inducing a conformational change in the enzyme (Fierobe et al., 1991). It has been shown that domains corresponding to two different types of xylanase are present in the *R. flavefaciens* 17 *xyn*A gene product, and are linked by a long sequence containing repetitive elements rich in asparagine and glutamine residues (Zhang and Flint, 1992).

1.8 Analysis of hybrid genes.

To study the function of different regions of cellulases, hybrid genes have been constructed. Homologous gene segments of (1-3, 1-4)-β-glucanase genes from *B. macerans and B. amyloliquefaciens* were exchanged. By combining the different properties of the enzymes, the *B. macerans* thermostability and the *B. amyloliquefaciens* tolerance towards acidic conditions, the hybrid enzymes exhibited 1.5- to 3- fold higher specific activities than the parental enzymes (Olsen et al., 1991). This method may be a useful technique for obtaining new and improved versions of biologically active proteins. Furthermore, X-ray diffraction and NMR studies of the hybrids can be used to elucidate the structural basis for the different biochemical properties of the two enzymes.

Chimeric proteins between EngB (CMCase) and EngD (CMCase, cellobiosidase and Avicelase) from *C. cellulovorans* were constructed by exchanging the non-homologous COOH-terminal regions. Chimeric proteins that contained the NH_2-terminus of EngD retained cellobiosidase activity but chimeras with the EngB NH_2-terminus did not. Hydrolysis of crystalline cellulose was observed only with the enzyme containing both the EngD NH_2-terminus and the EngD COOH-terminus. This result suggested that synergism between the EngD NH_2-terminus and COOH-terminus domains is essential for

hydrolysis of crystalline cellulose. As the EngD COOH-terminus has homology to the cellulose binding domain of *C. fimi* cellulase and *P. fluorescens* endoglucanase and xylanase, it may have affinity for cellulose fibers to allow the NH_2-terminal catalytic domain to hydrolyze amorphous regions of crystalline cellulose (Hamamoto et al., 1992).

1.9 Gene regulation.

A generally accepted view on the regulation of synthesis of cellulase enzymes is that small soluble "signal" fragments enter the cell by either active transport or by diffusion, and induce the synthesis of the corresponding enzyme(s) (Sternberg and Mandels, 1979). It has been suggested that organisms secrete low, constitutive levels of polysaccharide hydrolases and that these enzymes allow some hydrolysis of cellulose to release a soluble inducer which can enter the cell and effect induction. Such breakdown product(s) could effect induction in *C. thermocellum*, *T. fusca*, *F. succinogenes* and *T. reesei* (Johnson et al., 1985; Lin and Wilson, 1987; El-Gogary et al., 1989; Huang and Forsberg, 1990).

It has been reported that cellulose induced endoglucanase synthesis in conidia of *T. reesei*, but not in glucose-grown mycelia or protoplasts (Kubicek, 1987). This led to the postulate that close contact between cellulose and cells

could trigger cellulase synthesis (El-Gogary et al., 1989). It has been proposed that the induction of cellulase synthesis by crystalline cellulose in *T. reesei* comprises at least the following two events: (i) initial attack and adsorption of the fungal conidia on cellulose by means of their surface bound basal endoglucanase activities; (ii) conversion of the cellulose breakdown products to the actual intracellular inducer by means of a plasma membrane-bound β-glucosidase (Kubicek, 1987). Early experiments had shown that cellulose induced cellulase in *T. viride* but required time (Mandels et al., 1962), and that an intracellular β-glucosidase, the component responsible for transglycosylation by *T. reesei* (Vaheri et al., 1979), was involved in the control of the inducer, sophorose (β-1,2-glucobiose) (Inglin et al., 1980). Recently, the constitutive β-glucosidase, which is predominantly located in the plasma membrane of *T. reesei* (Umile and Kubicek, 1986), was suggested to be reponsible *in vivo* for formation of sophorose from cellulose hydrolysis products. This suggests that the effect of inhibitors of cellulase induction in *T. reesei* must have been due to inhibition of β-glucosidase (Kubicek, 1987). *T. reesei* cellobiohydrolase and a number of other cellulase genes are also induced by sophorose (Mandels and Reese, 1960; Mandels, et al., 1962; Sternberg and Mandels, 1979; Hrmova, et al., 1986; El-Gogary, et al., 1989; Hrmova, et al., 1991).

Regarding repression of cellulase genes, the expression of cellobiohydrolase I gene of *T. reesei* is repressed by glucose and glycerol (El-Gogary et al., 1989). The *cen*A gene of *C. fimi*, which encodes an endoglucanase, is transcribed at low levels when the cells are grown in glycerol, at high levels in CMC but not at all in glucose. In contrast, the *cex* gene and *cen*C gene, which encode an exoglucanase and an endoglucanase, are transcribed only during growth in CMC. The authors concluded that the *cen*A gene was being expressed at a low constitutive level during growth on glycerol presumably to generate degradation products which can act as inducers (Greenberg et al., 1987a; Moser et al., 1989). The lack of transcription during growth in glucose medium indicates repression at the transcriptional level by this readily assimilated carbon source. It has also been reported that the transcript of an endoglucanase gene could not be detected from cultures of *Thermomonospora fusca* YX grown on glucose or maltose as sole carbon source, but they were readily detected in cultures grown on cellulose, and comparatively lower on cellobiose (Lin and Wilson, 1987; Lin and Wilson, 1988). These results suggest catabolite repression as one mechanism regulating gene expression in this organism.

Sophorose, an efficient inducer of cellulase in the genus *Trichoderma*, was found to be less efficient in *Schizophyllum commune* (Rho et al., 1982) and in a basidiomycete species

(Shewale and Sadana, 1978). In contrast, cellobiose failed to induce cellulases in some *Trichoderma* strains (Royer and Nakas, 1990). It has been reported that polysaccharide-degrading enzyme systems of *Cryptococcus albidus* can be induced by positional isomers of dimers derived from the polysaccharide (Biely and Petrakova, 1984). Using the homodisaccharides cellobiose (or its positional isomers) and xylobiose, it has been shown that the synthesis of cellulases and β-xylanases in *T. reesei* QM9414 (Hrmova et al., 1986) and *Aspergillus terreus* (Hrmova et al., 1989) are under separate regulatory control.

It has been suggested that natural inducers of cellulases may be represented not only by homodisaccharides such as sophorose, cellobiose and xylobiose, but also by various heterodisaccharides (Hrmova et al., 1991). The regulation of the synthesis of cellulose- and xylan-degrading enzyme systems in *A. terreus* was studied using a series of synthetic glucobioses, xylobioses, glucosylxyloses and one xylosylglucose (Hrmova et al., 1991). Results showed that 2-*O*-β-D-glucopyranosyl D-xylose (Glcβ1-2Xyl) was the most powerful inducer of both cellulase and xylanase. However, 2-*O*-β-D-glucopyranosyl D-glucose (Glcβ1-2Glc, sophorose) and 2-*O*-β-D-xylopyranosyl D-xylose (Xylβ1-2Xyl) selectively induced the synthesis of cellulases and β-xylanases, respectively (Hrmova et al., 1991). These results demonstrate that the synthesis of cellulases and β-xylanases

of *A. terreus* are under separate regulatory control, and suggest that heterodisaccharides may occur in nature and play an important role in regulating the synthesis of fibre-degrading enzymes (Hrmova et al., 1991).

1.10 Cellulase gene clusters.

There are some reports of cellulase and hemicellulase genes organized in clusters in *R. flavefaciens* 17 and *P. fluorescens* subspecies *cellulosa* (Gilbert et al., 1988; Flint et al., 1989; Hall et al., 1990). The β-mannanase gene of *C. saccharolyticum* is located on the same genomic fragment as *cel*B which encodes a bifunctional exo- and endocellulase, and both genes were located on the same phage clone (Saul et al., 1989; Saul et al., 1990; Luthi et al., 1991). The *C. saccharolyticum* genes coding for enzymes involved in degradation of xylan also exist as a cluster (Luthi *et al.*, 1990). Multiple enzymatic activities (CMCase, MUCase (4-methylumbelliferyl-β-D-cellobioside) and OBR-HECase (Ostazin brilliant red-hydroxyethyl cellulose) activities) in a single clone from *R. albus* have also been detected (Howard and White, 1988). Genomic clustering of genes coding for hydrolytic enzymes having cellulase, chitinase, amylase, and pectinase activities has been reported in *Cellvibrio mixtus* (Wynne and Pemberton, 1986), and cellulase genes might be clustered on the chomosome of *Streptomyces viridosporus* (Ramachandra et al., 1987). The

physical association between cellulase and hemicellulase genes in *Bacillus polymyxa* has been well-demonstrated (Gosalbes et al., 1991). It was also found that a bacteriophage clone from *R. flavefaciens* 186 encodes CMCase, MUCase and β-glucosidase activities (Huang et al., 1989). The *B. fibrisolvens xyl*B gene, which encodes a bifunctional β-D-xylosidase and α-L-arabinofuranosidase, is located between two incomplete ORFs within 4.2 kb region, and was proposed to be part of a single operon (Utt et al., 1991). Recently, Bagnara-Tardif et al. (1992) reported that two complete endoglucanase genes (*celCCC* and *celCCG*) and two other partial open reading frames (ORF1 and *celCCE*) that probably encode two cellulases or related enzymes were clustered in *C. cellulolyticum*.

1.11 Genetic manipulation of rumen bacteria.

To study the cellulase systems of rumen bacteria, and their regulation and synergism under anaerobic conditions with a view to possibly improving their efficiencies, the ability to genetically manipulate them would be desirable. However due to the lack of suitable vectors and transformation systems most cellulase genes have been cloned and their products analyzed in heterologous hosts. The fact that expression occurs suggests that transcription and translation systems do not provide a major barrier to intergenetic transfer. An additional problem is that apart

from the examples of gene clusters given above, cellulase genes are scattered on the genomes of most microorganisms making it difficult to express multiple cellulase genes in heterologous hosts.

In an attempt to introduce foreign genes into a rumen bacteria, Hazlewood and Teather (1988) reported the transfer of the broad host-range conjugative plasmid RP4 from *E. coli* to *B. fibrisolvens*. However the clones were unstable. Recently high-voltage electroporation, a quick and reproducible technique, was used to introduce vectors derived from lactic acid bacteria plasmids into three freshly isolated *R. albus* strains RC6, RC7 and RC16 and into the *R. albus* type strain ATCC 27210 (Cocconcelli et al., 1992). Southern hybridization was used to demonstrate the presence of the plasmids in the transformants. The efficiency of electrotransformation was 3×10^5 transformants per µg DNA.

A possible factor affecting the viability and expression of DNA in rumen bacteria could be host restriction-modification systems, especially if the capsular and cell wall structures limit DNA entry (Morrison et al., 1992). Partial characterization of a DNA restriction endonuclease from *R. flavefaciens* FD-1 and its inhibition by site-specific adenine methylation has been reported (Morrison, et al.,

1992). Detection and inhibition of such systems may assist the expression of foreign genes in rumen bacteria.

1.12 The importance of *R. flavefaciens* FD-1 in the rumen.

The mesophilic *R. flavefaciens* FD-1 (NCDO 2215) is one of the most important cellulolytic bacterial species in the rumen (Halliwell and Bryant, 1963; Bryant, 1973; Bryant, 1986). It is Gram-positive and non-motile with coccoid cells 0.8-0.9 μm in diameter occurring singly or in pairs and chains (Stewart and Bryant, 1988). Cell walls have a prominent glycoprotein coat containing rhamnose, glucose and galactose (Bryant, 1986). Bacteria adhere strongly to cotton cellulose and to damaged cell walls of forage grass by means of this coat (Bryant, 1986). Almost all strains of *R. flavefaciens* are cellulolytic (Stewart and Bryant, 1988). When incubated with leaves of perennial ryegrass, cells mainly colonise the cut edges of the epidermis, schlerenchyma and phloem cells and digest the epidermis and parenchyma bundle-sheath cells (Latham et al., 1978). However, bacteria do not attach either to the readily degraded mesophyll cells or to the indigestible xylem vessels (Stewart and Bryant, 1988). The G+C content of the DNA is 39 - 44% and the major fermentation product is succinate (Bryant, 1986).

To degrade cellulose with sufficient efficiency to allow for growth *R. flavefaciens* has evolved a typical cellulase system (Pettipher and Latham, 1979a; 1979b; Gardner et al., 1987; Rasmussen et al., 1988; Doerner and White, 1990; Wang and Thomson, 1990). However β-1,4-glucosidase activity has not been detected (Rasmussen et al., 1988). Cellobiose, the major end product of cellulose degadation, is taken up by the cell where cellobiose is cleaved by cellobiose phosphorylase to yield glucose-1-phosphate and glucose for the Embden-Meyerhoff-Parnas (EMP) pathway metabolism (Ayers, 1958; Ayers, 1959; Bryant, 1986; Rasmussen et al., 1988; Helaszek and White, 1991).

In addition to being cellulolytic, most strains of *R. flavefaciens* can also degrade hemicellulose and pectin (Pettipher and Latham, 1979a; 1979b). It has also been reported that several DNA clones from *R. flavefaciens* 17 encode xylanase and lichenanase activities (Flint et al, 1989; 1991).

For these reasons, *R. flavefaciens* FD-1 appears to be a good choice for biochemical and genetic studies leading to an improved understanding of plant fibre digestion in the rumen.

1.13 Cloning of cellulase genes and purification of enzymes from *Ruminococcus* spp.

Cellulase genes that have been cloned from *Ruminococcus* spp. include a cellodextrinase (Wang and Thomson, 1990), a bi-functional cellulase/xylanase (Howard and White, 1990) and two β-glucanase genes from *R. flavefaciens* FD-1 (Doerner et al., 1992). In addition an endoglucanase gene has been cloned from *R. flavefaciens* 17 (Cunningham, et al., 1991) and multiple cellulase genes from *R. flavefaciens* strain 186 (Huang, et al., 1989). Several endoglucanase genes have been cloned from *R. albus* AR67, AR68 (Ware, et al., 1989), *R. albus* 8 (Howard and White, 1988), *R. albus* F-40 (Ohmiya, et al., 1989) and *R. albus* SY3 (Poole, et al., 1990), as well as a β-glucosidase gene from *R. albus* (Ohmiya et al., 1990).

An assessment of the endo-1,4-β-glucanase components of *R. flavefaciens* FD-1 has identified at least six different endoglucanases (Doerner and White, 1990). Although an exoglucanase enzyme has been purified (Gardner, et al., 1987). In addition a (1,3)-β-D-glucanase from *R. flavefaciens* OR18 has been isolated (Erfle and Teather, 1991).

1.14 The aims of this thesis. The study presented here was aimed at gaining further insight at the molecular level into

1.14 The aims of this thesis. The study presented here was aimed at gaining further insight at the molecular level into cellulose degradation by *R. flavefaciens* FD-1. A previously cloned cellulase gene (*celA*) was sequenced and shown to encode a cellodextrinase. A gene bank was established in *E. coli* and an endoglucanase gene (*celE*) cloned and sequenced. Transcriptional regulation of both the *celA* and *celE* genes was shown to be due to induction by cellotriose and possibly other cellodextrins.

CHAPTER 2

NUCLEOTIDE SEQUENCE OF THE *CELA* GENE ENCODING A CELLODEXTRINASE

CHAPTER 2

NUCLEOTIDE SEQUENCE OF THE *CELA* GENE ENCODING A CELLODEXTRINASE

2.0 Summary.

The nucleotide sequence of a 3.6 kb DNA fragment containing a cellodextrinase gene (*cel*A) from *R. flavefaciens* FD-1 was determined. The complete amino acid sequence of the CelA enzyme (336 residues) was deduced and corresponded to a protein of M_r 38686. It showed 40% identity with endoglucanase C of *Clostridium thermocellum* and 27.4% identity with endoglucanase 3 of *Fibrobacter succinogenes*, and was grouped into the same subfamily "A3". The ATG start codon is preceded by a ribosome-binding site sequence, GAGG. Although CelA has activity on carboxymethylcellulose, its degradation of cellopentanose to predominantly cellobiose indicates that it is a cellodextrinase. In *E. coli* the CelA enzyme is translocated into the periplasm. The lack of a typical signal sequence, and the results of transposon Tn*pho*A mutagenesis experiments, indicated that the enzyme is not secreted by a leader peptide.

2.1 Introduction.

R. flavefaciens has at least two types of cellulose degrading enzyme systems. One is at the level of β-1,4-glucanase activity and acts primarily on cellulose polymers. The second is at the level of cellodextrinase activity and acts mainly on cellulooligosaccharides (Rasmussen et al, 1988). The sequences of cellodextrinases from *Butyrivibrio fibrisolvens* H17c (Ced1; family "E") and *Pseudomonas fluorescens* subsp *cellulosa* (CelC; family "F") have been reported (Berger et al., 1990; Ferreira et al., unpublished data from GenBank). A cellodextrinase enzyme with an exotype function also exists in *F. succinogenes* (Huang and Forsberg, 1987). The cloning and expression in *E. coli* of a cellulase gene from *R. flavefaciens* FD-1 was reported previously (Barros and Thomson, 1987). Due to its activity on CMC it was classified as an endoglucanase. However, in this chapter further studies on the enzyme encoded by plasmid pMEB200 reveal that it has cellodextrinase activity. It has high activity on pNPC and releases predominantly cellobiose from cellodextrins. This chapter describes the analysis of the DNA sequence and further analyses of the cloned gene product.

2.2 Materials and Methods.

2.2.1 Bacterial strains, plasmids and growth conditions.

R. flavefaciens FD-1 was obtained from Dr. M. P. Bryant, University of Illinois, Urbana-Champaign. Plasmid pMEB200 carries the *cel*A gene of *R. flavefaciens* DNA cloned into plasmid pEcoR251 (Barros and Thomson, 1987). The *E. coli* strains HB101 (*leu*B6 *trp*E38 *met*E70 *rec*A13 *sup*E44; Boyer et al., 1969), LK111 (*thr*1 *leu*6 *thi*1 *sup*E44 *ton*A2$r_k^-m_k^+$*lac*I^q *lac*Z M15 *lac*Y1; Zabeau and Stanley, 1982), K514λ (r_k^- m_k^+ derivative of *E. coli* C600; Wood, 1966), CSH23 [Δ(*lac pro*) *sup*E *spc thi* (F'*lac*$^+$ *pro*A$^+$, B$^+$); Miller, 1972], and JM103 [Δ(*lac pro*) *thi*1 *str*A *sup*E *end*A *sbc*B *hsd*R4 (F' *tra*D36 *pro*A,B *lac*I^q *lac*ZΔM15); Messing et al., 1981] were used as recipient strains for recombinant plasmids. The PhoA$^-$ *E. coli* strain CC118 [*ara*D139 Δ(*ara*, *leu*)7697 Δ*lac*X74 *pho*AΔ20 *gal*E *gal*k *thi* *rps*E *rpo*B *arg*Eam *rec*A1; Manoil and Beckwith, 1985] was used for the Tn*pho*A fusion experiments. Phage λb221*rex*::Tn*pho*A cI857 Pam3 was a gift from Dr. C. Manoil (Gutierrez et al., 1987). Plasmid pEcoR251 was a gift from M Zabeau (Biotechnology Business Development, Ghent, Belgium) and is a positive selection vector containing the *E. coli* EcoRI gene under the control of the lambda rightward promoter, the ampicillin resistance gene (Ap^r) and the pBR322 origin of replication. It was derived from the pCL plasmids described by Zabeau and Stanley (1982). The M13-derived Bluescript plasmid (Stratagene, San Dieo, Calif.)

was used for subcloning and nucleotide sequencing. *E. coli* strains were grown in Luria-Bertani (LB) medium (Maniatis et al., 1982) (Appendix A) and 100 µg/ml ampicillin was used for the selection of transformants. *R. flavefaciens* FD-1 was grown anaerobically at 39°C in non-rumen fluid medium (NRF) as described by Caldwell and Bryant (1966) except that 0.4 ml trace element solution and 1 ml vitamin solution were added per 100 ml. In addition, the resazurin and cysteine $HCl.H_2O$ were replaced by indigo carmine (0.0005% w/v) and dithiothreitol (0.02% w/v) as indicator and reducing agents respectively (Appendix A).

2.2.2 Media and Buffers. All media, buffers and solutions not described in the text are listed in Appendix A.

2.2.3 Molecular techniques. The isolation of plasmid DNA and transformation experiments were carried out according to Maniatis et al. (1982). Restriction enzymes were obtained from Boehringer Mannheim and used according to the manufacturer's specifications. Sequencing was done by the chain termination method of Sanger et al.(1977) using a Sequenase Kit (United States Biochemical Corporation, Cleveland, Ohio) (Appendix B). The nucleotide and deduced amino acid sequences were analysed using the Genepro software package (Version 4.20, Riverside Scientific).

2.2.4 Preparation of periplasmic and intracellular extracts. Strains were grown overnight at 37°C in 200 ml LB. NaCl and Tris-HCl (pH 7.3) were added to final concentrations of 33 mM. Incubation was continued for a further 10 min and cells were collected by centrifugation (6000 g for 5 min). The supernatants were stored at -20°C. The pellets were resuspended in Tris-HCl(33 mM, pH 7.3) at 10 ml/g wet weight of cells. An equal volume of TSE [33 mM Tris-HCl, pH 7.3, 40% (w/v) sucrose, 2 mM EDTA] was added. After 5 min at room temperature the cells were collected as before. They were then resuspended in ice-cold deionised water at 10 ml/g wet weight of cells. After 1 min, $MgCl_2$ was added to a final concentration of 1 mM. The cells were collected as before and the supernatants, which represented the periplasmic fraction, were stored at -20°C. To obtain the intracellular fraction, the cells were resuspended in 5 ml PC buffer (50 mM K_2HPO_4, 12.5 mM citric acid pH 6). The cell suspension was cooled on ice and disrupted by sonication on ice(10s bursts for a total of 200s) using an MSE Soniprep 150 sonicator. The debris was removed by centrifugation for 15 min at 27000 g and the extract stored at -20°C.

2.2.5 Enzyme assays. Cellulase activity was determined according to Deshpande et al. (1984). Enzyme samples were suitably diluted in PC buffer and 250 ul mixed with 250 ul pNPC (Sigma, St. Louis, Mo.; 3.4 mM in PC buffer). This was

incubated at 39°C for 30 min, and the reaction stopped by the addition of 500 ul sodium carbonate (14% w/v). The absorbance was measured at 405 nm. Enzyme activity was expressed as micromoles of pNP liberated per minute per milligram protein.

2.2.6 ***β-Galactosidase enzyme assays.*** β-Galactosidase activity was assayed according to the method of Pardee et al. (1959) as described by Miller (1972). One unit of enzyme was defined as the amount of enzyme that produced 1 nmole *o*-nitrophenol per min.

2.2.7 ***β-Lactamase enzyme assays.*** β-Lactamase activity was assayed according to the method of Sykes and Nordstrom (1972). One unit of β-lactamase activity was defined as the amount of enzyme that hydrolyzes Ap at the rate of 1 μmole per min.

2.2.8 Protein determination and polyacrylamide gel electrophoresis. Protein concentrations were measured by the dye-binding method of Bradford (1976) using bovine serum albumin as a standard (Appendix B).

2.2.9 Analysis of proteins synthesized in vitro and in vivo. *In vitro* protein synthesis and analysis was carried out using the prokaryotic DNA-directed translation kit (Amersham) according to the manufacturer's instructions

except that half the recommended quantities of all the components were used. Periplasmic and intracellular proteins, prepared as above, were separated by SDS-polyacrylamide gel electro-phoresis (SDS-PAGE; Laemmli, 1970) (Appendix B).

2.2.10 TnphoA mutagenesis. Transposon insertions into pMEB200 were obtained using phage λ::Tn*phoA*(Manoil and Beckwith, 1985) according to the method of De Bruijn and Lupski (1984) with the modifications reported by Scholle et al. (1989), except that early stationary phase cells were used for infection with the phage. *E. coli* CC118 cells (5 ml) containing pMEB200 were grown to early stationary phase and infected by phage λ:Tn*phoA* at a multiplicity of infection of approximately 1.0 and incubated at 30°C for 15 min. The mixture was transferred to 50 ml LB broth and incubated for a further 4 hours at 30°C. The cells were concentrated and selected on LB agar supplemented with Ap (100 μg/ml), Km (250 μg/ml) and 5-bromo-4-chloro-3-indolyl phosphate (80 μg/ml). After incubation at 30°C for two to three days, the blue alkaline phosphatase-positive colonies were pooled, plasmid DNA isolated and CC118 cells re-transformed on the same selection medium. Plasmids prepared from re-transformants were characterized by restriction enzyme mapping using standard procedures (Maniatis et al., 1982).

2.2.11 Detection of CMCase activity. Cells were grown in LB agar containing 0.5% (w/v) medium-viscosity CMC (Sigma, St. Louis, Mo.). After 24 to 48 h of growth, the colonies were removed with a filter paper disc and the plates stained with 0.2% (w/v) Congo red and destained with 1M NaCl (Teather and Wood, 1982). As the CMCase activitity of CelA is low, it was necessary to use 0.5% CMC, rather than the more commonly used 0.1%, and it was preferable to incubate the plates for 48 h.

2.2.12 Analysis of hydrolysis products of cellooligosaccharides by HPLC. Hydrolysis of cellopentaose (Merck) was determined by incubating the enzyme with 100 µl of the dextrin (10 mg/ml in PC) at 39°C. Samples were taken at various times and analyzed in a high-pressure liquid chromatography (HPLC) system equipped with a model 156 refractive index detector (Beckman). Separation was achieved on a C_{18} column (Waters Associates, Milford, MA) at room temperature. Filtered water was used as an eluant at a flow rate of 1.5 ml/min.

2.3 Results.

2.3.1 Nucleotide sequence of the insert on pMEB200. To obtain templates for nucleotide sequencing, plasmid pMEB200 carrying *cel*A was digested with *Pvu*II and *Pst*I and fragments 1, 2 and 3 cloned into Bluescript SK digested with *Eco*RV and *Pst*I (Fig. 2.1). Exonuclease III was used to generate two sets of overlapping deletions of opposite polarity (Henikoff, 1984) (Fig 2.2). The nucleotide sequence of the 3.6 kb insert of pMEB200 carried an open reading frame (ORF) which, from the ATG start codon to the stop codon (TGA), contained 1011 nucleotides encoding 336 amino acid residues (Fig. 2.3). No typical promoter sequences could be identified in this region by sequence inspection. The molecular weight of the predicted polypeptide coded for by this region was 38686 Da. The ORF was preceded by a ribosome-binding site (GAGG) 8 bp upstream of the start codon. There is an imperfect inverted repeat situated at the 3' end of the gene. The transcript would have the potential to form a stem-loop structure with a ΔG = -19.7 kcal/mol (Salser, 1977).

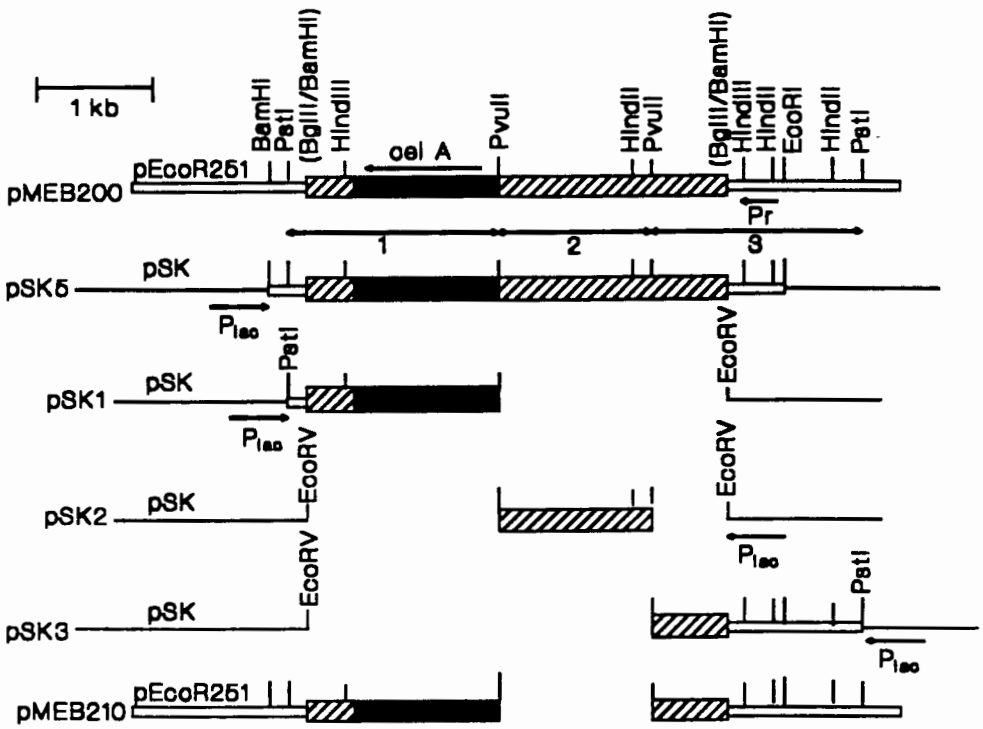


Fig. 2.1 Subcloning strategy for the *R. flavefaciens* FD-1 *celA* gene (▬▬). The insert (▨▨) on pEcoR251 (▭) was divided into 3 fragments which were independently subcloned into Bluescript pSK (——) plasmid for exonuclease III shortening. This generated 3 plasmids pSK1, pSK2 and pSK3. The *Bam*HI-*Eco*RI fragment of pMEB200 was subcloned onto pSK plasmid to generate pSK5. pMEB210 was generated by deleting the *Pvu*II fragment of pMEB200. P_r, lambda P_r promoter; P_{lac} , lac Z promoter.

2.3.2 ***Amino acid sequence homology.*** The amino acid sequence of the *celA* was compared with sequences in the GenEMBL data base using Fasta and TFASTA programs from the Genetics Computer Group sequence analysis package (Devereux et al., 1984). Amino acid sequence comparison of the CelA of *R. flavefaciens* FD-1, endoglucanase C of *C. thermocellum*

(Schwarz et al., 1988a) and endoglucanase 3 of *F. succinogenes* (McGavin et al., 1989) are shown in Fig. 2.4. Analysis of the aligned amino acid sequences showed 40.3% identity in the N-terminal regions of CelA and endoglucanase C of *C. thermocellum*, whereas the C-terminal regions showed no homology. It also showed 27.4% identity in the N-terminal region of CelA and the central region of endoglucanase 3 of *F. succinogenes*.

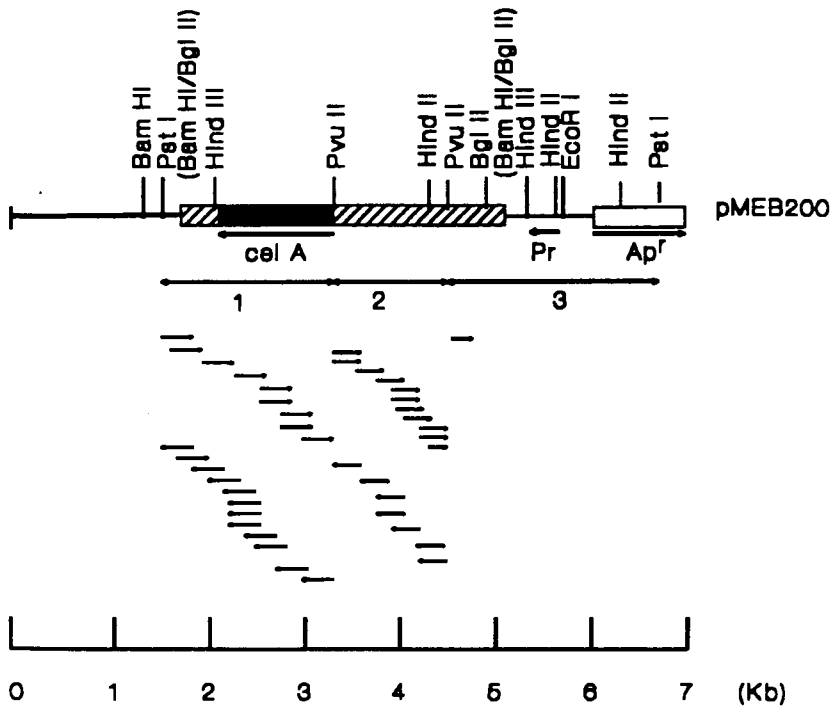


Fig. 2.2 Nucleotide sequencing strategy for the *cel*A gene. Vector DNA, (———); *cel*A gene open reading frame (ORF) (▬▬); *R. flavefaciens* insert DNA, (▨). Fragments 1, 2 and 3 were independently subcloned into Bluescript for exonuclease III shortening. The arrows represent the extent and direction of the templates generated.

```
-180 ATATTTTTATAATCCTTTGCGTTGTGTTCGCTCTTTAATACGCTCGGCCTTAACGCTTGC

-120 CATGTTCAATAAATTCTGTTCAAGAGACAATATTTTGCGTAAAGGCATCCCAAAATAAAC

-60  CTTCATAAATCGCCAGCTGTATATCAGAACAGTTAGTTCGGAGTTTTAGAGGTGAATATT
                 PvuII                                SD
1    ATGCTTAAAAGCAGAGGAATTATCAAGGGAATCAATCTGGGCGGATGGATGTCGCAGTGT
1    fM  L  K  S  R  G  I  I  K  G  I  N  L  G  G  W  M  S  Q  C
61   GACTACAGCCGTGAAAGACTGGATAACTTTGTAAAGGAAAACGACATAAAGCAGATAGCA
21    D  Y  S  R  E  R  L  D  N  F  V  K  E  N  D  I  K  Q  I  A
121  GACTGGGGATTCGACCACGTAAGACTTCCCATCGACTATAATATCGTTCAGAATAACGAC
41    D  W  G  F  D  H  V  R  L  P  I  D  Y  N  I  V  Q  N  N  D
181  GGAAGCGTTATCGAGGACGGCTATAACAGAATAGACAAGGTCGTTGAACTTTGCCGCAAA
61    G  S  V  I  E  D  G  Y  N  R  I  D  K  V  V  E  L  C  R  K
241  TATGGGCTGAAACTGGTCATCGACCTCCACAAGACAGCCGGTTTCTCATTTGATTTCGGC
81    Y  G  L  K  L  V  I  D  L  H  K  T  A  G  F  S  F  D  F  G
301  GAGCCTGAAAGCGGCTTTTTCGACAATAAAGAGTATCAGGAGCGCTTCTACATCCTGTGG
101   E  P  E  S  G  F  F  D  N  K  E  Y  Q  E  R  F  Y  I  L  W
361  GAAGAGATAGCAAGAAGATACGGCCACGACACCGACAATATCGTTTTTGAGCTTCTCAAT
121   E  E  I  A  R  R  Y  G  H  D  T  D  N  I  V  F  E  L  L  N
421  GAGGTGACCGATGAGGCGTTCATCGGCAAGTGGAACGAGATTTCGGATATCTGTATCGGA
141   E  V  T  D  E  A  F  I  G  K  W  N  E  I  S  D  I  C  I  G
481  AGGATTAGAAAAATAGCTCCCGAGGTGATTATTCTTCTCGGAAGCTACCACAATAATGCA
161   R  I  R  K  I  A  P  E  V  I  I  L  L  G  S  Y  H  N  N  A
541  GCGGACACAGTGCAGTTCCTCAATGCGCCGCACGATGACAGGGTAGTGTACAATTTCCAC
181   A  D  T  V  Q  F  L  N  A  P  H  D  D  R  V  V  Y  N  F  H
601  TGCTACGAACCGCTGAAATTCACTCATCAGGGTGCAACATGGACTCCCGATATCATCCCC
201   C  Y  E  P  L  K  F  T  H  Q  G  A  T  W  T  P  D  I  I  P
661  GGGGAGAGGATGAAATTCGAGGACAGCGAGACTTCCGAGGCGTATTTCGAGGAGCTTTTC
221   G  E  R  M  K  F  E  D  S  E  T  S  E  A  Y  F  E  E  L  F
721  TCCACAGCCATAAGCACGGCCGAGAAGTACGGCACTACTCTTTACTGCGGTGAATACGGT
241   S  T  A  I  S  T  A  E  K  Y  G  T  T  L  Y  C  G  E  Y  G
781  GTGATCGATGTCGTGTCCGCCGAGGATTCGCTGAAGTGGTTCAAGGTCATCAACAAGGTG
261   V  I  D  V  V  S  A  E  D  S  L  K  W  F  K  V  I  N  K  V
841  TTCAGCAAGCACGGCATTTCCCGTGAGTGTGGAATTACAAGGAAATGGACTTCGGCATTT
281   F  S  K  H  G  I  S  R  E  C  G  I  T  R  K  W  T  S  A  F
901  CCGACAGCAAGTACGACAGCAACCGTGAGAGATACTGAAGTATCTGTGATAAAAACAAGG
301   P  T  A  S  T  T  A  T  V  R  D  T  E  V  S  V  I  F  T  R
961  AAAGCACATAACTGCTGTACTCTTACAGAAGTTTTTACACGAGTCTATTGAGGGAAACCC
321   K  A  H  N  C  C  T  L  T  E  V  F  T  R  V  Y  *
1021 TTTTGAAAAAGGGTTCTCCCTCAAACTCCCTTCCTAAAACTTTCAGTATTAATTTTTGCC

1081 CATTATAGGGCGCACTTGAAAGAAATGACAAATTCTTACTTTACTTGTGTGCGCCCTATA
```

Fig. 2.3 Nucleotide sequence of the *celA* structural gene and flanking regions of *R. flavefaciens* DNA. The deduced amino acid sequence is given in single letter code. The ribosomal binding sequence (SD) is underlined with a double line. The inverted repeats including the TGA stop codon are underlined with a single line.

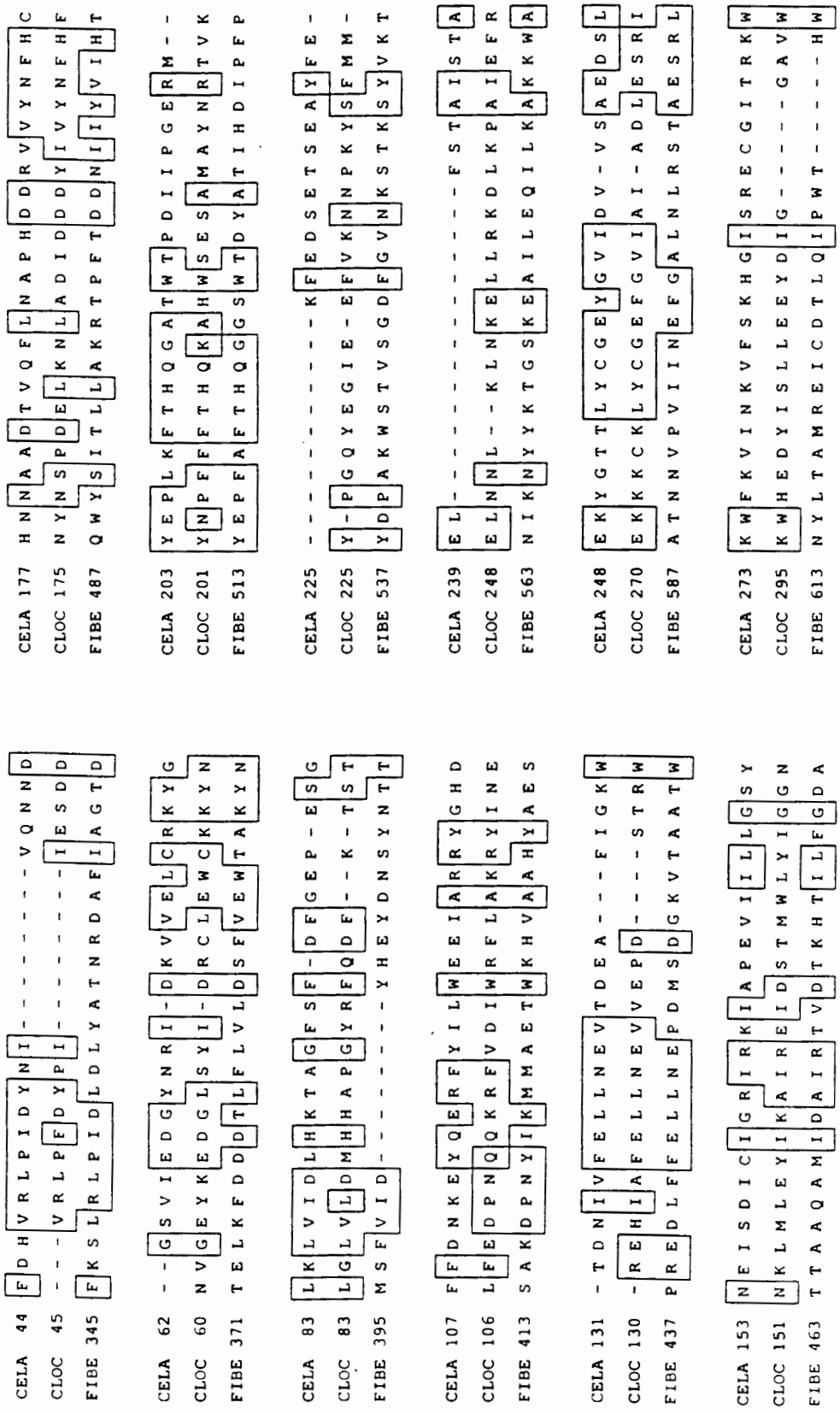


Fig. 2.4 Amino acid sequence homology of CelA of *R. flavefaciens* FD-1 (CELA), endoglucanase C (CelC) of *C. thermocellum* (CLOC) and endoglucanase 3 of *F. succinogenes* (FIBE). Regions of identity are boxed.

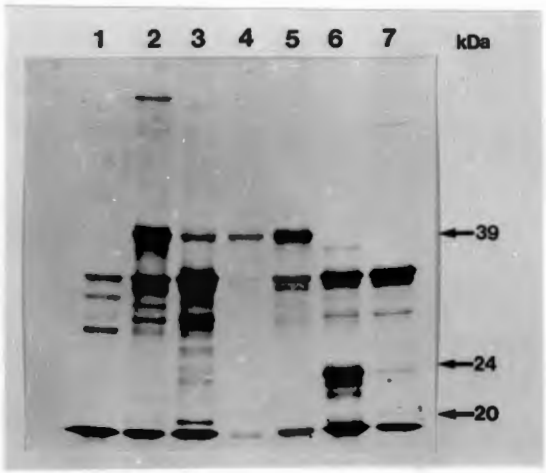


Fig. 2.5 SDS-polyacrylamide gel electrophoresis (SDS-PAGE) analysis of ^{35}S-labelled proteins synthesised *in vitro*. Lanes 1, pEcoR251; 2, pMEB200; 3, pSK1; 4, pSK5; 5, pMEB210; 6, pSK2; 7, pSK. The product of the *cel*A gene is the 39 kDa protein indicated.

2.3.3 Expression of celA in vitro. Analysis by SDS-PAGE of *in vitro* protein synthesis directed by pMEB200, revealed the presence of a major protein band with an apparent molecular weight of ca. 39 kDa (Fig. 2.5). A protein band with a similar molecular weight was also detected upon expression of pSK1 (fragment 1 of pMEB200 cloned into Bluescript), pSK5 (Bluescript carrying the large *BamH*I-*Eco*RI fragment of pMEB200 which contains the entire *R. flavefaciens* insert as well as some pEcoR251 DNA) and pMEB210 (which is pMEB200 with *Pvu*II fragment 2 deleted) (Fig 2.1). This protein band

was absent after expression of pSK2 (Bluescript carrying fragment 2 of pMEB200).

Table 2.1 Localization of enzyme activity in different cell fractions.

Strain	Plasmid	Enzyme	Enzyme activity[a]		
			Extra-cellular	Peri-plasmic	Intra-cellular
JM103	pMEB200	pNPCase	2 (3)	56 (89)	5 (8)
		β-lactamase	661(31)	1236 (59)	206(10)
JM103	pMEB210	pNPCase	0.4	0.6	4
		β-lactamase	260(18)	1140 (78)	69 (4)
LK111	pMEB200	pNPCase	42(22)	99 (51)	53(27)
		β-lactamase	333(56)	239 (40)	27 (4)
LK111	pMEB210	pNPCase	0	0	0
		β-lactamase	96(10)	802 (86)	40 (4)
CSH23	pMEB200	pNPCase	2 (1)	283 (87)	40(12)
		β-lactamase	201(32)	379 (61)	44 (7)
		β-galactosidase	1 (2)	0.3(<1)	62(98)

[a] Units *p*-nitrophenyl-β-D-cellobiosidase (pNPCase) are millimoles *p*-nitrophenol released per minute. Units β-lactamase are micromoles ampicillin hydrolysed per minute. Units β-galactosidase are millimoles *o*-nitrophenol released per minute. Percentages in parentheses.

2.3.4 Secretion and localization of CelA. Most of the CMCase activity (75%) of *E. coli* HB101 (pMEB200) has been found in the periplasmic fraction (Barros and Thomson, 1987). However, the amino acid sequence of the *cel*A gene product did not reveal a leader peptide such as is characteristic of many secreted proteins. Therefore the assays were repeated using pNPC as CelA has higher activity on this substrate than on CMC (1100 and 2.0 x 10^{-3} μmole min^{-1} mg^{-1} respectively). As controls for *E. coli*

intracellular and periplasmic enzymes, β-galactosidase and β-lactamase respectively were assayed. The data (Table 2.1) confirmed the previous results, although there was considerable strain variation in the level of CelA expression.

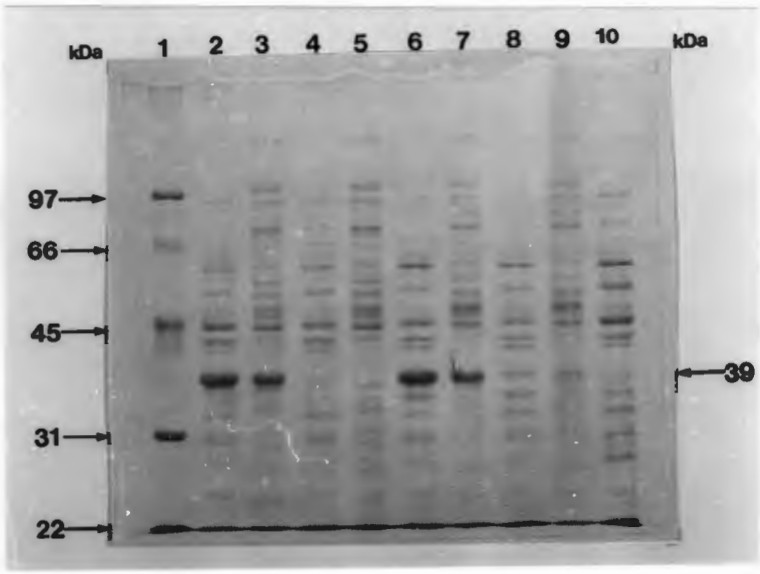


Fig. 2.6 SDS-PAGE analysis of periplasmic (even numbered lanes) and intracellular (odd numbered lanes) proteins of stationary phase cultures. Lanes 1, Molecular weight markers; 2 and 3, LK111 (pMEB200); 4 and 5, LK111 (pMEB210); 6 and 7, JM103 (pMEB200); 8 and 9, JM103 (pMEB210); 10, K514λ (pEcoR251). The 39 kDa CelA protein band can be detected in the periplasmic and intracellular fractions of the LK111 (pMEB200), JM103 (pMEB200) and JM103 (pMEB210), but is absent in LK111 (pMEB210). K514λ (pEcoR251) was used as a negative control.

2.3.5 Analysis of proteins in vitro and in vivo. It was of interest to note that pNPCase activity was barely detectable in JM103 (pMEB210) and was absent in LK111 (pMEB210) (Table 2.1). This was in contrast to the *in vitro* data where pMEB210 directed readily detectable levels of a 39 kDa protein (Fig. 2.5, lane 5). The poor *in vivo* expression of

the *cel*A gene of pMEB210 was confirmed by SDS-PAGE analysis of the periplasmic and cytoplasmic proteins prepared from cultures of LK111 and JM103 carrying pMEB200 and pMEB210 (Fig. 2.6). The 39 kDa CelA protein band was not detected in LK111 (pMEB210) and was present in low amounts in JM103 (pMEB210).

2.3.6 TnphoA mutagenesis. λ::Tn*pho*A mutagenesis was used to determine whether the CelA protein contained an amino acid sequence that could act as a leader peptide. The leader sequence of the alkaline phosphatase gene on Tn*pho*A has been deleted and as export is required for activity of the enzyme, an in-frame sequence upstream of the transposition site has to function as a signal sequence for activity to be detectable. Tn*pho*A mutagenesis of pMEB200 resulted in the isolation of several recombinant plasmids conferring alkaline phosphatase activity on the Pho^- recipient *E. coli* CC118. Restriction enzyme analysis of a number of these Pho^+ clones revealed that all the Tn*pho*A inserts were in the vector (Fig. 2.7). To ensure that it was possible for Tn*pho*A to transpose into the cloned *R. flavefaciens* DNA, Km^r Ap^r Pho^- transformants were screened on CMC plates. $CMCase^-$ mutants were isolated and their plasmids subjected to retriction analysis. It was shown that the Tn*pho*A inserts were in the *cel*A gene (Fig. 2.7).

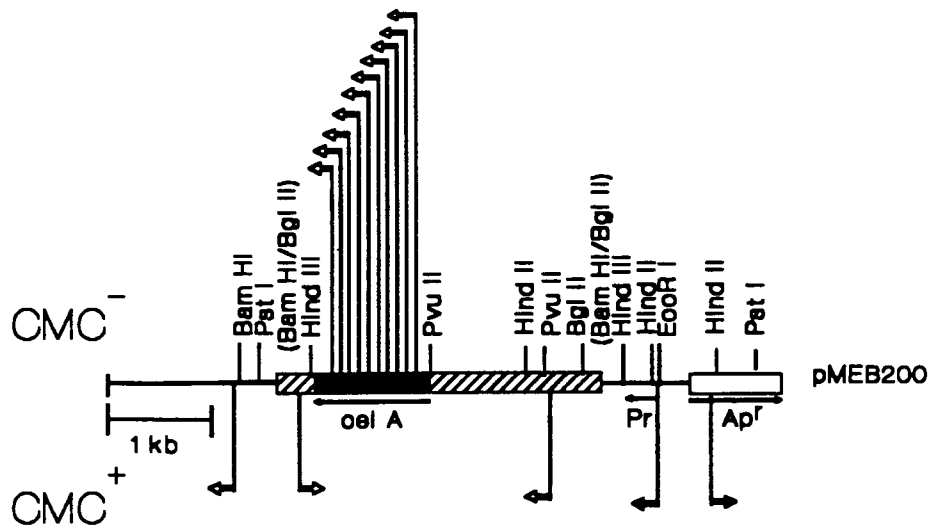


Fig. 2.7 Physical and genetic map of pMEB200::Tn*pho*A. The arrows show the orientation of Tn*pho*A insertions within the plasmid pMEB200. CMC^+ and CMC^- indicate positive and negative CMCase activities. Pho^+ (➡) and Pho^- (⇨) symbolize positive and negative alkaline phosphatase activities.

2.3.7 HPLC analysis of hydrolysis products of cellopentaose. HPLC analysis of the products of cellopentaose hydrolysed by CelA indicated that its degradation initially yielded predominantly cellobiose with some cellotriose and glucose (Fig. 2.8). After extending the reaction time, all the cellotriose was converted to cellobiose and glucose. The final product was predominantly cellobiose.

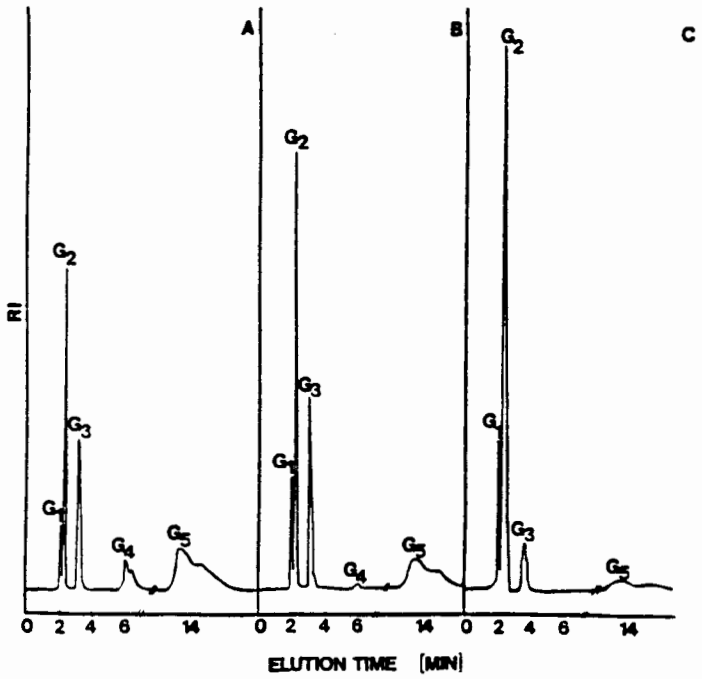


Fig. 2.8 Cellodextrinase activity of CelA on cellopentaose determined by HPLC analysis of degradation products. The reaction time of (A), (B) and (C) were 30 min., 60 min. and 90 min. respectively.

2.4 Discussion.

Nucleotide sequence analysis of the *cel*A gene of *R. flavefaciens* revealed a ribosomal binding site GAGG situated eight bp upstream of the initiation codon ATG. However no typical *E. coli* consensus promoter sequence could be found. The predicted amino acid sequence of the N-terminal region of CelA showed 40% identity with the N-terminal region of endoglucanase C of *C. thermocellum* and 27.4% identity with the central region of endoglucanase 3 of *F. succinogenes*. These three enzymes are grouped into the cellulase subfamily "A3" as defined by Henrissat et al. (1989) and Beguin (1990).

The protein which could be translated from this sequence would have a molecular weight of 38.7 kDa, in agreement with the 39 kDa protein bands found on expression of the *cel*A gene *in vivo* and *in vitro*. It has also been demonstrated by immunoassay that the purified CelA is identical in molecular weight and antigenicity to a protein produced by *R. flavefaciens* FD-1 (Brown et al., submitted). The N-terminal protein sequence has been determined and confirmed the amino acid sequence Met-Leu-Lys-Ser-Arg-Gly derived from the DNA sequence of the gene (Thomson, pers. comm.). HPLC analysis gave results consistent with an exo-type cellodextrinase or an endoglucanase able to act on cello-oligosaccharides. Small amounts of glucose could be detected during the

breakdown of cellopentaose. In a report by Petre et al. (1986) endoglucanase C of *C. thermocellum* was shown to have CMCase, pNPCase and cellodextrinase activities similar to CelA. However further analyses showed that it had higher activity on β-1,3- and β-1,4-linkages in larger substrates such as β-glucan and lichenan (Schwarz et al., 1988b). This is also the case with endoglucanase 3 of *F. succinogenes* (McGavin et al., 1989). Interestingly enough these two enzymes show lower homology with each other (25%) than either do to CelA (40% and 27% respectively) (McGavin et al., 1989; this study).

It was of interest to note that pNPCase activity was very low in JM103 (pMEB210) and absent in LK111 (pMEB210). SDS-PAGE analysis of cell extracts showed that this was the result of poor expression of the gene and not to any post-translational inactivation of the enzyme. Moreover, *in vitro* transcription and translation of pMEB210 showed that CelA could be produced under those conditions. Plasmid pMEB210, was derived from pMEB200 by the deletion of the *Pvu*II fragment 2 (about 1.45 kb), removing the 5' upstream region of the *cel*A gene. This presumably resulted in the deletion of the *cel*A promoter present on pMEB200. It is unlikely that lambda P_r is being used for the initiation of transcription of *cel*A in pMEB200 as this promoter is located about 2.25 kb upstream from the *cel*A structural gene in that plasmid. This is consistent with the Tn*pho*A

mutagenesis data, in which an insertion in the lambda P_r promoter and one 1.5 kb upstream of *cel*A between it and the lambda P_r promoter, both showed CMCase activity. In addition *cel*A was transcribed in pSK1 in which lambda P_r has been deleted and the direction of transcription is in the opposite orientation to the lac promoter of the vector pSK. Whatever sequence is acting as a promoter in the insert of pMEB200 appears to have been lost in the *Pvu*II deletion which generated pMEB210, although the gene can still be transcribed *in vitro*. The reason for this is not clear.

Most of the pNPCase activity encoded by pMEB200 and its derivatives was located in the periplasmic fraction. The amino acid sequence at the N-terminus of CelA does not, however, contain a typical signal sequence. Moreover, Tn*pho*A mutagenesis of pMEB200 resulted in alkaline phosphatase activity only when transposition occurred into vector DNA. That Tn*pho*A was able to transpose into the *cel*A gene in the same orientation was shown when 10 of 64 plasmids carrying inserts with a $CMCase^-$ $PhoA^-$ phenotype were analysed using restriction enzymes. Despite repeated attempts no $PhoA^+$ mutants were isolated in which Tn*pho*A was in the *cel*A gene. It has been reported that internal sequences can act as secretory signals in a *P. fluorescens* subsp. *cellulosa* cellulase (Hall et al., 1990). However CelA does not have such an internal sequence. It would therefore appear that CelA is being secreted by a mechanism

other than a signal peptide and that export of alkaline phosphatase across the cytoplasmic membrane by this mechanism may be lethal. This is reminiscent of the lethality to *E. coli* of β-galactosidase fused to a signal sequence (Bassford et al., 1979). The authors concluded that the enzyme "jammed" the export system resulting in cell death. It is possible that a similar situation occurs when alkaline phosphatase is fused to CelA.

In contrast, Gong et al. (1989) have found that the *F. succinogenes* periplasmic cellodextrinase was located intracellularly when the gene was cloned and expressed in *E. coli*. Scholle et al. (1989) have also shown that the *Vibrio alginolyticus* sucrase enzyme is secreted across the cytoplasmic membrane of *E. coli* by a mechanism other than the use of a typical signal sequence. The analysis of the sequences of more cellulases from rumen anaerobes may lead to an understanding of how these proteins are secreted.

CHAPTER 3

CLONING AND SEQUENCING OF AN ENDOGLUCANASE GENE

CHAPTER 3

CLONING AND SEQUENCING OF AN ENDOGLUCANASE GENE

3.0 Summary.

A gene which encodes a protein with both carboxymethyl cellulase and xylanase activity was cloned from *R. flavefaciens* FD-1 on an *E. coli/Bacillus subtilis* shuttle vector pEB1. The resulting plasmid, pWF1, carried a 3.6 kb DNA insert containing an open reading frame of 963 bp. This encoded an enzyme, CelE, of 320 amino acid residues with a calculated M_r of 35937 Da. CelE showed 11.6% identity and 55.3% similarity with the N-terminal catalytic region of a cellulase from the alkalophilic *Bacillus* sp. strain 1139. In order to obtain expression in *E. coli* the gene had to be transcribed from the lambda P_r promoter.

3.1 Introduction.

Cellulase and xylanase activities in various cellulolytic microorganisms can be associated in the same protein (MacKenzie et al., 1987; Kundu et al., 1988; Howard and White, 1990). Doerner and White (1990) have identified 18 endo-glucanase components in the extracellular fluid of cultures of *R. flavefaciens* FD-1, all of which also showed xylanase activity. Despite the abundance of cellulases in this organism the cloning of only four cellulase genes has been published. One encodes the cellodextrinase described in the previous chapter (Wang and Thomson, 1990), which may only have a minor role in cellulose degradation (Brown et al., submitted). The other three are two β-glucanase genes (*celB* and *celC*) and the *celD* gene encoding a bi-functional cellulase/xylanase enzyme (Howard and White, 1990; Doerner et al., 1992). Two xylanase genes and a cellulase gene have been cloned from the related strain *R. flavefaciens* 17 (Cunningham et al., 1991; Flint et al., 1991), while clones showing endo-glucanase, exo-glucanase and β-glucosidase activity from strain 186 have been identified by Huang et al. (1989). In addition, recombinant clones containing bi-functional cellulase/xylanase activities have been isolated from *R. albus* (Romaniec et al., 1989).

A number of gene libraries of *R. flavefaciens* FD-1 DNA in a variety of *E. coli* vectors have been screened but no stable

clones expressing cellulase activities other than the cellodextrinase and three cellulases mentioned above have been isolated (Doerner et al., 1992; Thomson, pers. comm.). This chapter describes the use of an *E. coli/B. subtilis* shuttle vector to clone an endoglucanase gene.

3.2 Materials and Methods.

3.2.1 Bacterial strains, plasmids. The *E. coli* strains MM294 (*sup*E44 *hsd*R *end*A1 *pro thi;* Meselson et al., 1968) and JM105 [*sup*E *end*A *sbc*B15 *hsd*R4 *rps*L1 *thi* Δ(lac proAB)], and the *E. coli/B. subtilis* shuttle vector pEB1 (Lin et al., 1990) were used for positive selection for inserted DNA fragments in E. coli. pEB1 was constructed by replacing the pBR322 fragment of pLP1202 (Robson and Chambliss, 1986; Ostroff and Pene, 1983) with pEcoR252, which is a derivative of pEcoR251 lacking the *Pst*I site in the β-lactamase gene (P. Janssen, pers. comm.). The structure of pEcoR251 (a gift from M. Zabeau, Plant Genetic Systems, Gent, Belgium) is similar to other plasmids utilizing inactivation of the *Eco*RI endonuclease gene as a selective marker (e.g. Cheng and Modrich, 1983) and has been described elsewhere (Zappe *et al.*, 1986). The suicide bacteriophage vector λ467::Tn5 (b221 *rex*::Tn5 *c*I857 *O*am29 *P*am80) was used for Tn5 mutagenesis (De Bruijn et al., 1984). The other bateria and plasmids used not described in here were as described in chapter 2.2.1

3.2.2 Media and Buffers. All media, buffers and solutions not described in the text are listed in Appendix A.

3.2.3 Preparation of chromosomal DNA. *R. flavefaciens* FD-1 chromosomal DNA was isolated as described by Berger *et al.*

(1989), and purified by CsCl-ethidium bromide density gradient centrifugation and ethanol precipitation (Sambrook et al., 1989).

3.2.4 Construction and screening of a R. flavefaciens FD-1 genomic library. *R. flavefaciens* chromosomal DNA was partially digested by *Sau*3A endonuclease and fragments of approximately 3 to 10 kb isolated by sucrose gradient centrifugation (Sambrook et al., 1989). These were ligated with pEB1 which had been digested with *Bgl*II endonuclease and used to transform competent *E. coli* MM294 cells (Meselson et al., 1968). Transformants were screened for cellulase activity by plating on LB agar containing 0.9% (w/v) medium-viscosity carboxymethyl cellulose (CMC; Sigma Chemical Co., St. Louis, Mo.). The endoglucanase positive clone was identified by a " halo " formation. An additional screen was performed by plating on LB agar, removing the colonies with Whatman No. 1 filter paper discs and flooding them with 4-methylumbelliferyl-β-D-cellobioside (MUC; Sigma Chemical Co., St. Louis, Mo.) according to the method of Van Tilbeurgh and Claeyssens (1985). The discs were incubated for 10 min. at 37°C and examined for fluorescent halos under UV light (254 nm).

3.2.5 Enzyme activity assays. Strains were screened for CMCase and xylanase activity by plating on LB agar containing 0.5% (w/v) medium-viscosity CMC or 0.1% xylan

(Sigma Chemical Co., St. Louis, Mo.) and staining with congo red according to the method of Teather and Wood (1982). An additional screen was performed by plating on LB and flooding the plates with 4-methylumbelliferyl-β-D-cellobioside (MUC; Sigma Chemical Co., St. Louis, Mo.) according to the method of Van Tilbeurgh and Claeyssens (1985). To screen liquid cultures, cell free extracts, prepared by French Press (pressure up to 600 psi) from stationary phase cells, were applied to wells (0.4 mm in diameter) in LB agar containing 0.5% CMC or 0.1% xylan, the plates stained with 0.1% (w/v) Congo red after 48 h of incubation at 37°C, and destained with 1M NaCl. For liquid assays cells were grown overnight at 37°C in 200 ml LB, harvested and resuspended in 5 ml 0.2 M phosphate buffer (pH 6.8). The cell suspensions were cooled on ice and extracts prepared in a French Press (pressure up to 600 psi). The cell debris was removed by centrifugation for 15 min at 27,000 g and the extracts stored at -20°C. Cellulase and xylanase activities were determined by the 3,5-dinitrosalicylic acid (DNS) method (Miller, 1959) using CMC (0.5% w/v) or xylan (0.25% w/v) as substrates. One unit of enzyme activity was defined as the amount of enzyme releasing 1 µmole reducing sugar per min. Activity on pNPC was as described in Chapter 2.2.5.

3.2.6 ***Southern blot analysis.*** Southern blot hybridization was performed as described by Sambrook et al. (1989). DNA

was transferred to Hybond-N^+ membrane (Amersham) as described by Reed and Mann (1985). The hybridization probe was nick-translated with (α-^{32}P) dCTP as described by Rigby *et al.* (1977). These methods are described in Appendix B.

3.2.7 DNA sequencing. To obtain templates for nucleotide sequencing the 3.6 kb *Hin*dIII fragment of plasmid pWF1 was subcloned into Bluescript SK (Stratagene, San Diego, California). Deletion, DNA sequencing and data analyses were performed as described in chapter 2.2.3.

3.2.8 Analysis of proteins synthesized in vitro. Proteins synthesized *in vitro* were analysed using the prokaryotic DNA-directed translation kit (Amersham) as described in chapter 2.2.9.

3.2.9 Detection of cellulase activity in gels. Proteins were separated by electrophoresis on 10% (w/v) polyacrylamide gels in the absence of SDS. Cellulase activity was detected by overlaying the gel with a polyacrylamide gel containing 0.5% (w/v) CMC, incubating in a humidified incubator at 39°C for 24 h, staining with 0.1% (w/v) Congo red, destaining with 1 M NaCl, and fixing with 5% acetic acid.

3.2.10 Tn5 mutagenesis. Tn5-mediated transposon mutagenesis was performed by the method of De Bruijn and

Lupski (1984), except that transposon-carrying JM105 cells containing the plasmid pWF1 were selected on LB plates containing 50 µg/ml Km and 100 µg/ml Ap. The cells were washed off, plasmids isolated and transformed into MM294 to screen for CMCase activity.

3.3 Results.

3.3.1 Cloning the celE gene and preparation of a restriction enzyme map of the plasmid pWF1. A gene library of *R. flavefaciens* FD-1 DNA was constructed by partial *Sau*3A digestion, ligation with *Bgl*II-digested pEB1 and transformation into *E. coli* MM294. It was screened on CMC-containing plates and by flooding colony-carrying filter paper discs with MUC. Of the 6000 clones tested, 9 had CMCase activity but none had activity on MUC. Preliminary restriction enzyme maps and xylan plate assays showed that five of the inserts carried the *cel*A gene previously cloned (Wang and Thomson, 1990), three carried the same 3.6 kb insert and also had xylanase activity, and one carried a different insert but only had CMCase activity. One of three CMCase/xylanase clones was called pWF1 and further characterized. A restriction map of the plasmid was prepared (Fig. 3.1).

3.3.2 Southern blot analysis. To ensure that the gene on pWF1 originated from *R. flavefaciens* FD-1 and to confirm that it was different from the *cel*A gene, FD-1 chromosomal DNA, pWF1 and pMEB200 were analyzed by Southern gel hybridization. A ^{32}P-labelled 3.5 kb *Hind*III-*Pst*I fragment of the insert on pWF1 was used as a probe. It hybridized to the same sized fragment in FD-1 and pWF1 DNA cut with the same two enzymes (Fig. 3.2). It did not hybridize to

pMEB200 DNA, proving that it was a different gene which originated from FD-1 (results not shown). To determine whether the gene was present in other rumen anaerobes, DNA from *R. albus* 22.08.6A, *B. fibrisolvens* H17c and *R. flavefaciens* 17 was subjected to Southern blot analysis using the same probe. No hybridization was found (results not shown).

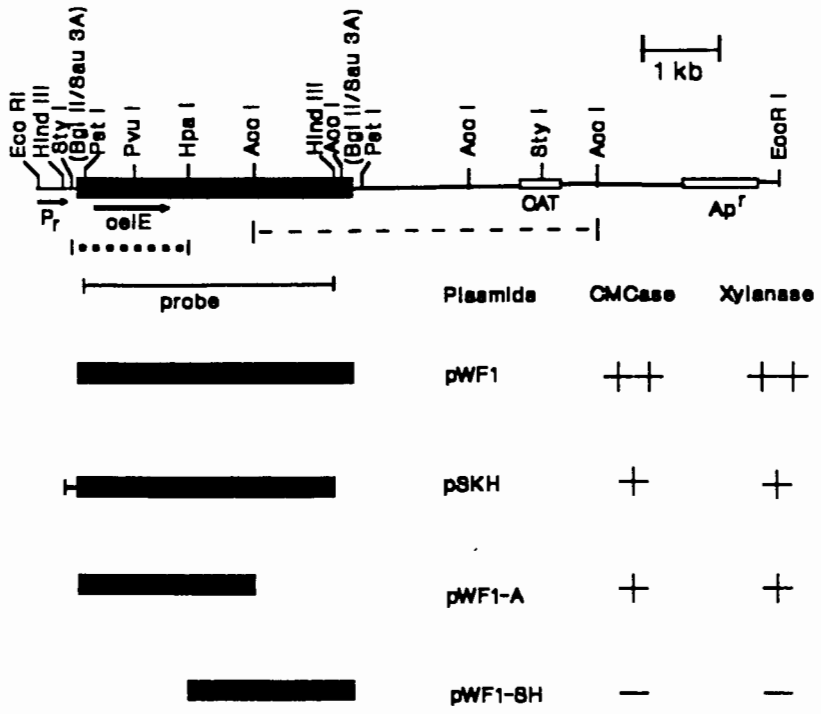


Fig. 3.1 Restriction map of pWF1 and the relevant deletions showing + or - CMCase and xylanase activities. Vector DNA (———); insert DNA (▬▬▬); 4.8 kb *Acc*I deletion in pWF1-A (------); 1.65 kb *Sty*I-*Hpa*I deletion in pWF1-SH (• • • •) and 3.8 kb *Hin*dIII fragment subcloned into bluescript SK to give pSKH. The 3.5 kb *Pst*I-*Hin*dIII fragment was used as probe for Southern hybridization. P_r, lambda P_r promoter. Arrow shows direction of transcription.

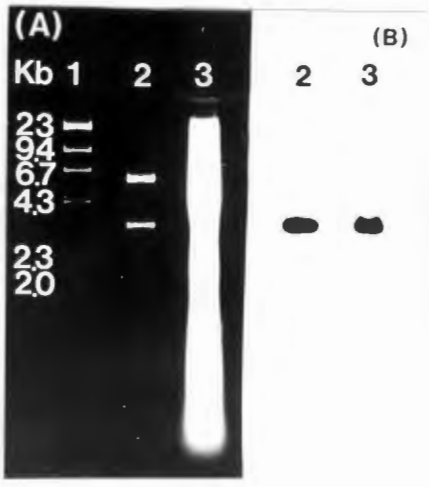


Fig. 3.2 Agarose gel (A) and Southern blot (B) of DNA isolated from pWF1 and *R. flavefaciens* FD-1 using a 3.5 kb *Hind*III-*Pst*I fragment of pWF1 as a probe. Lanes 1, Lambda *Hind*III M_r markers; 2, *Hind*III-*Pst*I digested pWF1; 3, *Hind*III-*Pst*I digested DNA from *R. flavefaciens* FD-1.

3.3.3 Enzyme activity. A cell free extract from *E. coli* MM294 (pWF1) was used to determine the effects of pH and temperature on CMCase and xylanase activities. The effect of pH was determined in different phosphate-citrate (pH 4.0 - 6.0), phosphate (pH 7.0 - 7.5) and glycine-NaoH (pH 8.6 - 10) buffers. The optima for both activities were 43.5°C and pH 7 (Fig. 3.3 A and B). Cell free extracts of *E. coli* MM294 strains carrying either pWF1 or pMEB200 were assayed on a variety of substrates, and showed clear differences in that pWF1 had low activity on pNPC but had activity on xylan (Table 3.1).

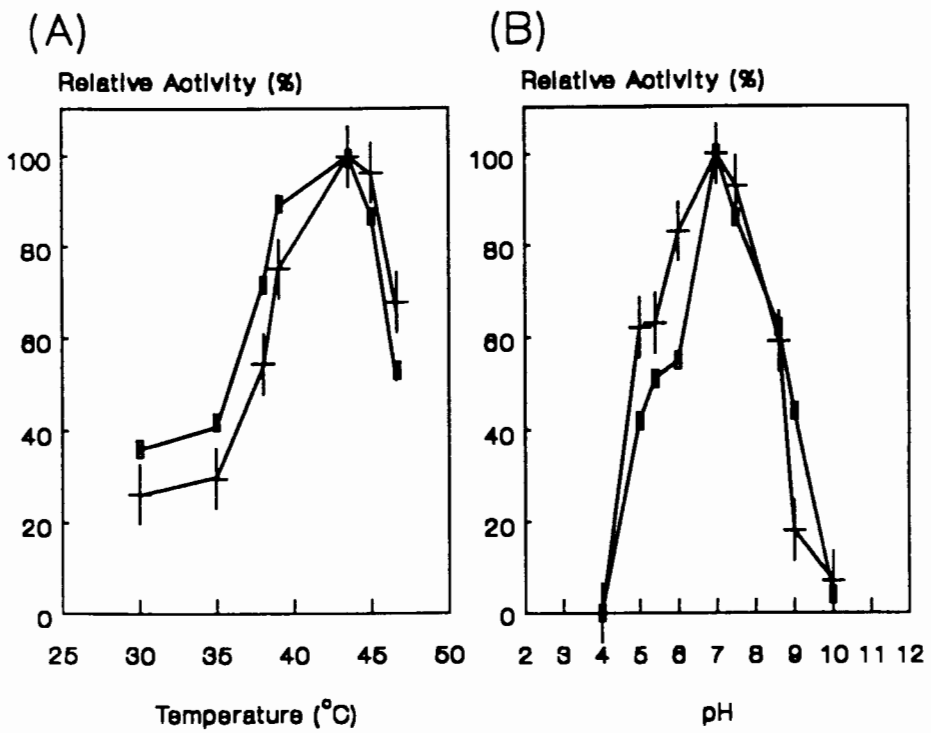


Fig. 3.3 Temperature (A) and pH (B) profiles of the endoglucanase (⊡) and xylanase (+) activities in cell free extracts of *E. coli* (pWF1).

3.3.4 Localization of the celE gene on pWF1. Various deletions of pWF1 were made to determine the localization of the *celE* gene (Fig. 3.1). pWF1-A, an *Acc*I deletion removing 1.3 kb of the insert and 3.5 kb of the vector, gave activity on both CMC and xylan albeit at somewhat reduced levels (Fig. 3.4; Table 3.1). pWF1-SH, a *Sty*I-*Hpa*I deletion removing 1.65 kb of pWF1 adjacent to the λP_r promoter, resulted in the total loss of CMCase and xylanase activity. It would therefore appear that the *celE* gene is located in the λP_r proximal region of the insert. To confirm this,

saturation Tn5 mutagenesis was performed on pWF1. A total of 33 CMCase$^-$ mutants were isolated and mapped. All were located within the λP_r promoter of the vector or in the adjacent region of the insert (Fig. 3.5). Thus the *celE* gene is probably being transcribed by the λP_r promoter. It was of interest to note that no Tn5 insertions occurred in the opposite end of the insert. The reason for this is unknown. Further evidence for the involvement of λP_r was obtained when a 3.8 kb *Hin*dIII fragment of pWF1 which removed the λP_r promoter but retained most of the insert, was subcloned into bluescript SK. Reduced levels of CMCase activity were seen in strains carring this plasmid, pSKH, despite it having a higher copy number than pWF1 (Fig. 3.4; Table 3.1).

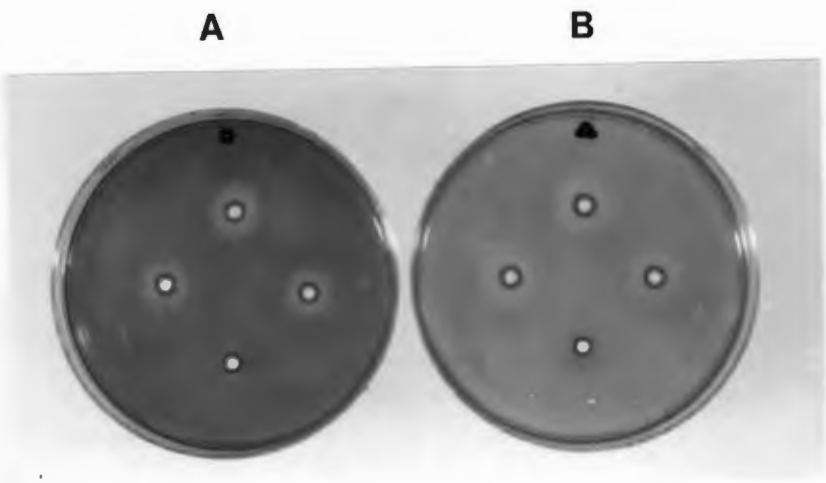


Fig. 3.4 CMCase (A) and xylanase (B) activity of MM294 carrying pWF1 (top well); pWF1-A (left well); pSKH (right well); pWF1-SH (bottom well).

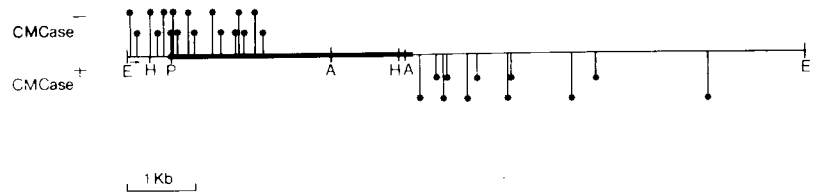


Fig. 3.5 Map of pWF1 showing the positions and CMCase activity of the various Tn5 insertions. E, *Eco*RI; H, *Hin*dIII; P, *Pst*I; A, *Acc*I. Arrow represents the λP_r promoter.

Table 3.1 Enzyme activity of plasmid-containing *E. coli* strains.

Plasmid	Substrate	Activity (μmole min^{-1} mg^{-1})
pWF1	CMC	2.5×10^{-3}
	Xylan	8.4×10^{-4}
	Avicel	ND[a]
	Laminarin	ND
	Lichenan	1.6×10^{-4}
	Barley β-glucan	2.3×10^{-4}
	Salicin	ND
	pNPC	1.3×10^{-4}
	pNPG[b]	ND
	pNPX[c]	ND
pMEB200	CMC	2.0×10^{-3}
	Xylan	ND
	pNPC	1100
pWF1-A	CMC	8.4×10^{-4}
	Xylan	8.0×10^{-4}
pWF1-SH	CMC	ND
	Xylan	ND
pSKH	CMC	6.4×10^{-4}
	Xylan	6.7×10^{-4}

[a] Not detectable
[b] *p*-Nitrophenyl-β-D-glucopyranoside
[c] *p*-Nitrophenyl-β-D-xylopyranoside

3.3.5 Identification of the CMCase encoded by pWF1. Analysis by SDS-PAGE of the synthesis *in vitro* of proteins encoded by pWF1 revealed the presence of three protein bands with apparent M_r of ca. 45 kDa, 35 kDa and 24 kDa which were not synthesized by the vector alone (Fig. 3.6, lanes 3 and 4). Plasmid pWF1-A in which 1.3 kb of the insert and 3.5 kb of the vector have been deleted, did not synthesize any of these proteins, but instead directed the synthesis of a ca. 60 kDa protein (Fig. 3.6, lane 2). This deletion appears, therefore, to have created a fusion protein, which may explain the decreased levels of activity seen in cells carrying pWF1-A. Plasmid pWF1-SH, in which a 3.5 kb fragment of pWF1 adjacent to the lambda P_r promoter has been deleted, directs the synthesis of the ca. 45 kDa and 24 kDa proteins (Fig. 3.6, lane 1). As there is no CMCase or xylanase activity in cells carrying this plasmid, it would appear that the 35 kDa protein has this activity. In an attempt to prove this CMC zymograms were prepared. Non-denaturing PAGE gels containing cell extracts from *E. coli* strains carrying pWF1 and its deletion derivatives, overlayed with CMC-containing polyacrylamide, revealed the presence of a band of activity in strains carrying pWF1 and pWF1-A (Fig. 3.7). This band was absent in strains carrying pWF1-SH. Unfortunately the activity proved to be sensitive to SDS as no activity could be detected when the active band was cut out and subjected to SDS-PAGE Zymograms.

However, the results were consistent with the finding that the insert encodes a single protein with CMCase activity.

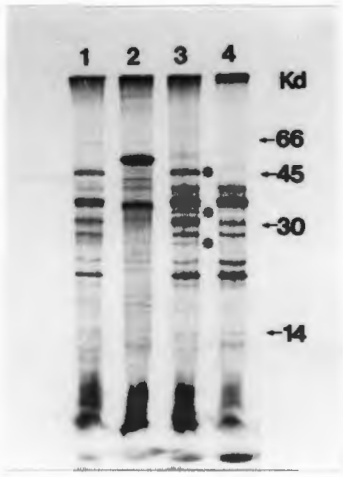


Fig. 3.6 SDS-PAGE analysis of proteins encoded *in vitro*. Lanes 1, pWF1-SH; 2, pWF1-A; 3, pwF1 and 4, pEB1. *, proteins encoded by pWF1.

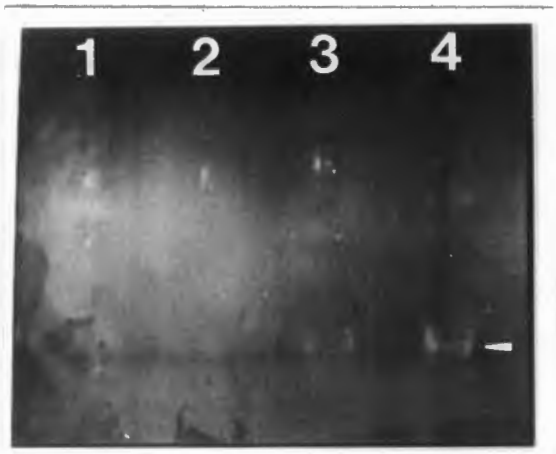


Fig. 3.7 Non-denaturing PAGE CMCase zymograms of *E. coli* strain MM294 carrying lanes 1, pSKH; 2, pWF1-SH; 3, pWF1-A and 4, pWF1. Arrowhead shows position of CMCase band.

```
      AGATCCAGAGCCTTTTTACAAGTGCTTGTCTTATTTGGATGGAAAGCCATATATGACAG

-180 AGGTCATTGCGTGTGCTGCAGAAAAAGTACGTTTGGAATTAAAGTAAGTATGCTTAATAT

-120 AAAGAACCGATAATTCATACTCATAAAACGCCAAGAGCGTTCAATTCATAGTCAGAGGTG

-60  ATACTAATGTGGATTTTAAATAATGATGTATAGAAAAACAATTTTGATTGGAGGGCTATT
                                                        SD
1    ATGCGAAAAAGAATAACTTCAATGATAACAGTAATTTCATTGACCTTATGCGGAACTATA
1    fM  R+ K+ R+ I  T  S  M  I  T  V  I  S  L  T  L  C  G  T  I
61   GGTACTTATCCACACAATGCTAATGCATCAGCGAAAGAAGCGTTCGAGGATATATCTAAA
21    G  T  Y  P  H  N  A  N  A  S  A  K  E  A  F  E  D  I  S  K
121  AGATCAGAGTGTACTAAAGTATTTGATAACGATAACGGAACATACACCGCTTATTCAAAT
41    R  S  E  C  T  K  V  F  D  N  D  N  G  T  Y  T  A  Y  S  N
181  ACTGCGCCAATTCACTATCTTAATAATGATGAGTGGAAAGAAATTGACAATACATTAATT
61    T  A  P  I  H  Y  L  N  N  D  E  W  K  E  I  D  N  T  L  I
241  GAGGACAGCGATGATTATTACAGAAACAAAGATAACTCTTTCAATATTTATATTCCAAAG
81    E  D  S  D  D  Y  Y  R  N  K  D  N  S  F  N  I  Y  I  P  K
301  GAATATTCGCTTGGTAAGAATATAAAAAATCCTGTTATAATGAATTATGACACATTTAGT
101   E  Y  S  L  G  K  N  I  K  N  P  V  I  M  N  Y  D  T  F  S
361  CTGTCAACGTCAATAACAGACATTAAGATGCCTGATAGACCTGAAGAAGAAACGTTTTCT
121   L  S  T  S  I  T  D  I  K  M  P  D  R  P  E  E  E  T  F  S
421  GAGGGAGTATATGCCGATATCAATAATGAGGTTTATACTACAATTGATAATTCAGCAATA
141   E  G  V  Y  A  D  I  N  N  E  V  Y  T  T  I  D  N  S  A  I
481  GCCGAAGGAATGAAAGCGGCTTTAAAGAAGTCAGCATCAATGGCGACCTATCATTCGATC
161   A  E  G  M  K  A  A  L  K  K  S  A  S  M  A  T  Y  H  S  I
541  GTTGAAGATATTGATCTTGATATAGCTGTGCATAACGCTTCTGTCAGCGAGTCGATAATC
181   V  E  D  I  D  L  D  I  A  V  H  N  A  S  V  S  E  S  I  I
601  ATCAATAAGTTAGAATCTTTACCGGAAACGATCACATATTCTATTTCGGTTGAAGACGCA
201   I  N  K  L  E  S  L  P  E  T  I  T  Y  S  I  S  V  E  D  A
661  ATTATTAAGAAAACTGAAGACAACAGATTGCAGCTTATCAAAGATAATGAAGCAGTGCTC
221   I  I  K  K  T  E  D  N  R  L  Q  L  I  K  D  N  E  A  V  L
721  ATACTATCCCCGTTTACGATCAGCGATTCATCTGAAAACGTCAATGTCATGCAGGTAGAA
241   I  L  S  P  F  T  I  S  D  S  S  E  N  V  N  V  M  Q  V  E
781  TATGATCTTACTGAATCTGAAGAAGGATATGAAGTTACTCTCTATCCGGCAGAAACAGTT
261   Y  D  L  T  E  S  E  E  G  Y  E  V  T  L  Y  P  A  E  T  V
841  AATCGCATGAGCAGCCCTGTTATGCCATTATCTTTAGGCTCGGAGTATTCATACGACAGA
281   N  R  M  S  S  P  V  M  P  L  S  L  G  S  E  Y  S  Y  D  R
901  CCATTTTCAACAGTTTACAATAGCCAGGCTTCGCTTCATCCGTCTATAATAATAATTACA
301   P  F  S  T  V  Y  N  S  Q  A  S  L  H  P  S  I  I  I  I  T
961  TAAAAGTAGGAAATGAAGCTAATAATAGTTATCATACTATAGGATCGTTTCTCGAAGATT
      *      ------------------>          <----------

1021 TTTCTTTTTATGGTCCATACGTTAGAATAGTAGATTCGACATTTTATTATTATGTAACTT

1081 CCTTAAGTACA
```

Fig. 3.8 Nucleotide sequence of the CMCase gene on pWF1 from *R. flavefaciens* FD-1. The derived amino acid sequence is given in the one-letter code from position 1 to 960 (320 residues). The Shine-Dalgarno (SD) sequence is underlined. The positively charged amino acids at the start of the putative signal sequence are indicated by plus signs, and the hydrophobic region is doubly underlined. The inverted repeat is indicated by arrows.

3.3.6 Nucleotide sequence of the insert on pWF1. The nucleotide sequence of the 3.6 kb insert of pWF1 carried an open reading frame (ORF) which, from the ATG start codon to the stop codon (TAA), contained 963 nucleotides encoding 320 amino acid residues (Fig. 3.8). The molecular mass of the predicted polypeptide coded for by this region was 35.9 kDa which agrees with the observed M_r of 35 kDa. Although the sequence has a possible ribosomal binding site (GGAGGG) 5 bp upstream of the putative ATG initiation codon, as was found in the *celA* gene it does not have any identifiable promoter sequences. The stop codon is followed by an inverted repeat sequence. The potential stem/loop structure that could be formed in the resultant mRNA had a $\Delta G = -8$ Kcal mol^{-1} (calculated according to Salser, 1977). This is therefore unlikely to be a rho-independent transcription termination sequence. The putative signal sequence was 29 or 31 amino acid residues long, with three positively charged amino acids near the N-terminal end followed by a 15 amino acid hydrophobic domain and a potential A-X-A leader peptidase cleavage site after the twenty-ninth or thirty-first amino acid. The hydropathy plot of the CelE signal sequence was analyzed by computer using Genepro program (Fig. 3.9). The predicted amino acid sequence of CelE was compared with sequences in the GenEMBL data base using FASTA programmes from the Genetics Computer Group sequence analysis package (Devereux *et al.*, 1984). It showed 11.6% identity and 55.3% similarity in a 293 amino acid overlap with the N-terminal

catalytic region of a cellulase gene from the alkalophilic *Bacillus* sp. strain 1139 (Fukumori et al., 1986; Gilkes et al., 1991; Fig. 3.10).

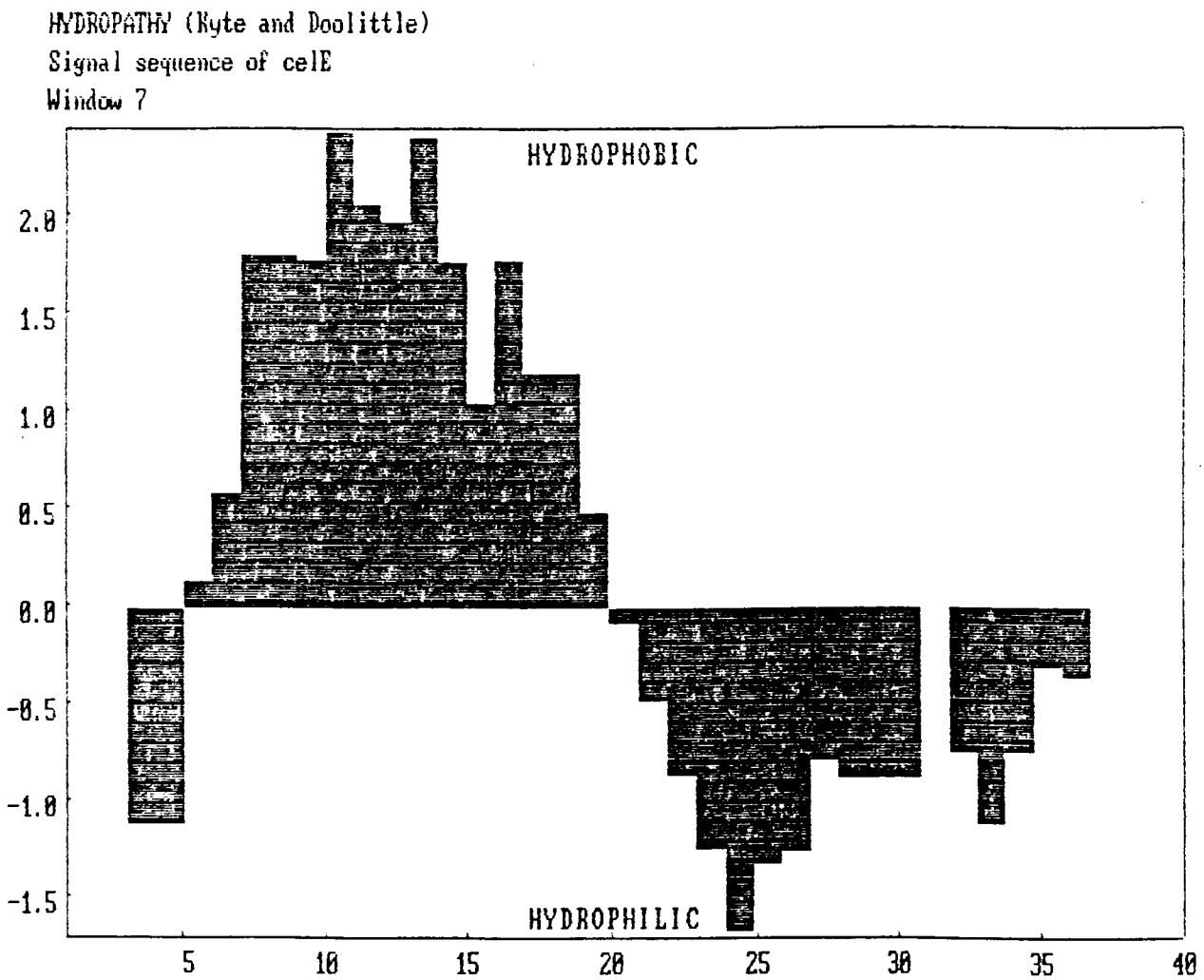


Fig. 3.9 Hydropathy plot of the signal peptide sequence of CelE from *R. flavefaciens* FD-1.

```
                10        20        ↓30       40        50
CelE     MRKRITSMITVISLTLCGTIGTYPHNANASAKEAFEDISKRSECTKV-FDNDNGTYTA
         :||::: :|: | |:|   :: :| : :|:::: :::::  : ::|   :::|:
Ba1139 MMLRKKTKQLISSI-LILVLLLSLFPTALAAEGNTREDNFKHLLGNDNVKRPSEAGALQL
              10        20       30↑       40        50

        60        70        80          90       100       110
CelE   YSNTAPIHYLNNDEWKEIDNTLIEDSDDYYRN--KDNSFNIYIPKEYSLGKNIKNPVIMN
        : :::: ::::: |   ::: ::: :::::  :|||:::  :::::: :: |: :: ::
Ba1139 QEVDGQMTLVDQHGEKIQLRGMSTHGLQWFPEILNDNAYKA-LANDWE-SNMIRLAMYVG
      60        70        80        90       100         110

         120       130       140       150       160       170
CelE   YDTFSLSTSITDIKMPDRPEEETFSEGVYADINNEVYTTID-NSAIAEGMKAALKKSASM
        :::: :::::: ::  :  : ::::::|: :: :|::: | :::: :| :: ::: |::
Ba1139 ENGYASNPELIKSRV-IKGIDLAIENDMYVIVDWHVHAPGDPRDPVYAGAEDFFRDIAAL
       120       130       140       150       160       170

          180       190       200       210       220       230
CelE   ATYHSIVEDIDLDIAVHNASVSESIIINKLESLPETITYSISVEDAIIKKTEDNRLQLIK
        : :: :   :|: :  ::: ::: | |: |:  :: :|: :: : ::::::::  ::|
Ba1139 YPNNPHI-IYELANEPSSNNNGGAGIPNNEEGWNAVKEYADPIVE-MLRDSGNADDNIII
        180       190       200       210       220       230

          240       250       260       270       280       290
CelE   DNEAVLILSPFTISDSSENVNVMQVEYDLTESEEGYEVTLYPAETVNRMSSPVMPLSLGS
        :::    :|   :|:: : :    : :: ::::: ::: ||:|| |: :: ||:
Ba1139 VGSPNWSQRPDLAADNPIDDHHTMYTVHFYTGSHAASTESYPPETPNSERGNVMSNTRYA
          240       250       260       270       280       290

          300       310       320
CelE   EYSYDRPFSTVYNSQASLHPSIIIIT

Ba1139 LENGVAVFATEWGTSQANGDGGPYFDEADVWIEFLNENNISWANWSLTNKNEVSGAFTPF
          300       310       320       330       340       350
```

Fig. 3.10 Amino acid sequence homology of the endoglucanase (CelE) from *R. flavefaciens* FD-1 and the endoglucanase from alkalophilic *Bacillus* sp. strain 1139 (Ba1139) (Fukumori et al., 1986). The amino acids are identified by the single-letter code, and regions of identity and similarity are represented by | and : symbols respectively. The putative signal peptidase cleavage sites are shown by arrows.

Table 3.2 Codon usage and frequencies in the *celA* gene and *celE* gene from *R. flavefaciens* FD-1 and comparison with *E. coli*[a], *B. subtilis*[b] and cellulase genes from *Ruminococcus*[c].

Residue & Codon	*celA* Total(%)[d]	*celE* Total(%)	Rf17*xynA* Total(%)	Rf17*endA* Total(%)	Ra*celA* Total(%)	Ra*celB* Total(%)	Ra*bglT* Total(%)	Ra*end* Total(%)	Ecohigh Total(%)	Ecolow Total(%)	Bsu Total(%)
Gly GGG	2(9)	0(0)	3(4)	0(0)	0(0)	1(3)	1(1)	1(4)	13(2)	10.33(15)	293(15)
GGA	8(36)	5(63)	12(17)	6(26)	3(11)	6(19)	19(21)	9(35)	3(0)	7.07(11)	606(31)
GGT	4(18)	2(25)	37(54)	8(35)	6(21)	11(35)	30(34)	8(31)	365(59)	17.37(26)	404(21)
GGC	8(36)	1(13)	17(25)	9(39)	19(68)	13(42)	39(44)	8(31)	238(38)	32.20(48)	622(32)
Glu GAG	19(68)	8(29)	19(86)	17(61)	13(46)	19(56)	35(53)	24(65)	108(22)	22.20(42)	618(30)
GAA	9(32)	20(71)	3(14)	11(39)	15(54)	15(44)	31(47)	13(35)	394(78)	30.60(58)	1416(70)
Asp GAT	9(38)	16(70)	18(32)	15(44)	20(53)	12(34)	35(54)	17(61)	149(33)	32.73(59)	1034(64)
GAC	15(63)	7(30)	38(68)	19(56)	18(47)	23(66)	30(46)	11(39)	298(67)	22.53(41)	576(36)
Val GTG	9(39)	2(12)	1(3)	3(14)	1(4)	2(10)	23(35)	3(13)	93(16)	17.40(29)	434(24)
GTA	4(17)	4(24)	14(48)	6(29)	11(48)	9(45)	19(29)	10(43)	146(26)	12.53(21)	414(23)
GTT	5(22)	8(47)	13(45)	9(43)	6(26)	9(45)	16(24)	10(43)	289(51)	17.33(29)	548(30)
GTC	5(22)	3(18)	1(3)	3(14)	5(22)	0(0)	8(12)	0(0)	38(7)	13.33(22)	422(23)
Ala GCG	4(24)	5(26)	0(0)	1(3)	2(7)	9(26)	14(16)	4(11)	161(26)	25.87(28)	501(24)
GCA	8(47)	6(32)	17(34)	13(42)	13(48)	15(44)	43(49)	18(47)	173(28)	25.13(27)	673(32)
GCT	1(6)	6(32)	33(66)	9(29)	7(26)	5(15)	22(25)	14(37)	212(35)	16.80(18)	546(26)
GCC	4(24)	2(11)	0(0)	8(26)	5(19)	5(15)	9(10)	2(5)	62(10)	24.[illegible](27)	366(18)
Lys AAG	11(52)	7(41)	35(97)	17(77)	10(53)	18(75)	38(73)	17(77)	111(26)	18.33(37)	588(27)
AAA	10(48)	10(59)	1(3)	5(23)	9(47)	6(25)	14(27)	5(23)	320(74)	31.27(63)	1623(73)
Asn AAT	10(59)	18(72)	71(33)	11(30)	11(42)	7(28)	22(52)	10(37)	19(6)	15.73(50)	699(53)
AAC	7(41)	7(28)	146(67)	26(70)	15(58)	18(72)	20(48)	17(63)	274(94)	15.53(50)	612(47)
Met ATG	3(100)	9(100)	17(100)	17(100)	7(100)	14(100)	29(100)	13(100)	170(100)	20.67(100)	729(100)
Ile ATA	7(24)	14(41)	3(12)	4(13)	13(54)	17(63)	26(60)	19(68)	1(0)	8.87(13)	291(15)
ATT	7(24)	13(38)	4(15)	3(10)	1(4)	1(4)	3(7)	4(14)	70(17)	35.20(50)	966(50)
ATC	15(52)	7(21)	19(73)	24(77)	10(42)	9(33)	14(33)	5(18)	345(83)	26.00(37)	657(34)

Table 3.2-Continued

Thr	ACG	2(9)	4(15)	2(6)	1(3)	0(0)	0(0)	12(19)	3(11)	25(7)	14.60(28)	365(24)
	ACA	10(43)	10(38)	19(53)	19(61)	11(46)	8(38)	22(34)	17(61)	14(4)	5.67(11)	694(45)
	ACT	8(35)	9(35)	9(25)	5(16)	2(8)	6(29)	12(19)	4(14)	130(35)	10.60(20)	271(18)
	ACC	3(13)	3(12)	6(17)	6(19)	11(46)	7(33)	18(28)	4(14)	206(55)	20.87(40)	213(14)
Trp	TGG	7(100)	1(100)	71(100)	13(100)	13(100)	10(100)	13(100)	10(100)	55(100)	13.87(100)	300(100)
Cys	TGT	4(50)	1(50)	3(43)	1(20)	1(33)	1(25)	8(50)	2(67)	22(49)	7.27(52)	87(44)
	TGC	4(50)	1(50)	4(57)	4(80)	2(67)	3(75)	8(50)	1(33)	23(51)	6.80(48)	110(56)
End	TGA	1(100)	0(0)	1(100)	1(100)	0(0)	0(0)	0(0)	1(100)	0(0)	0.00(0)	17(19)
	TAG	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0.00(0)	14(15)
	TAA	0(0)	1(100)	0(0)	0(0)	1(100)	1(100)	1(100)	0(0)	0(0)	0.00(0)	60(66)
Tyr	TAT	6(40)	15(79)	15(42)	8(42)	9(56)	9(50)	18(55)	8(47)	51(25)	19.07(64)	682(67)
	TAC	9(60)	4(21)	21(58)	11(58)	7(44)	9(50)	15(45)	9(53)	157(75)	10.87(36)	339(33)
Phe	TTT	6(30)	5(71)	3(15)	1(6)	2(20)	2(17)	13(29)	3(27)	51(24)	27.00(62)	802(67)
	TTC	14(70)	2(29)	17(85)	17(94)	8(80)	10(83)	32(71)	8(73)	166(76)	16.53(38)	389(33)
Ser	AGT	1(5)	1(3)	3(6)	1(4)	1(6)	3(10)	9(19)	4(17)	9(3)	11.40(18)	202(10)
	AGC	8(42)	6(16)	16(30)	6(26)	3(18)	10(32)	16(33)	7(30)	71(20)	9.00(14)	444(23)
	TCG	4(21)	6(16)	2(4)	4(17)	4(24)	4(13)	2(4)	6(26)	14(4)	14.67(23)	193(10)
	TCA	1(5)	13(35)	16(30)	4(17)	3(18)	4(13)	3(6)	1(4)	7(2)	10.67(16)	397(20)
	TCT	1(5)	10(27)	10(19)	3(13)	1(6)	5(16)	11(23)	3(13)	120(34)	9.87(15)	467(24)
	TCC	4(21)	1(3)	6(11)	5(22)	5(29)	5(16)	7(15)	2(9)	131(37)	9.27(14)	251(13)
Arg	AGG	5(28)	0(0)	0(0)	0(0)	2(13)	1(8)	10(20)	2(14)	1(0)	4.00(5)	126(11)
	AGA	8(44)	6(75)	9(53)	7(30)	5(31)	7(54)	13(27)	7(50)	0(0)	9.00(12)	350(30)
	CGG	0(0)	0(0)	0(0)	3(13)	1(6)	0(0)	3(6)	0(0)	1(0)	7.47(10)	155(13)
	CGA	1(6)	1(13)	0(0)	4(17)	0(0)	0(0)	3(6)	0(0)	2(1)	7.47(10)	110(9)
	CGT	2(11)	0(0)	7(41)	5(22)	4(25)	3(23)	8(16)	3(21)	290(74)	21.93(30)	222(19)
	CGC	2(11)	1(13)	1(6)	4(17)	4(25)	2(15)	12(24)	2(14)	96(25)	23.13(32)	198(17)

Table 3.2-Continued

Gln	CAG	6(100)	3(100)	129(100)	17(94)	3(100)	4(100)	21(91)	5(83)	233(86)	26.20(64)	537(45)
	CAA	0(0)	0(0)	0(0)	1(6)	0(0)	0(0)	2(9)	1(17)	37(14)	14.93(36)	653(55)
His	CAT	2(22)	3(60)	2(40)	3(30)	3(33)	3(50)	5(45)	5(71)	18(17)	21.47(69)	441(68)
	CAC	7(78)	2(40)	3(60)	7(70)	6(67)	3(50)	6(55)	2(29)	85(83)	9.87(31)	204(32)
Leu	TTG	0(0)	2(11)	1(3)	1(3)	1(5)	0(0)	2(3)	0(0)	18(3)	9.47(10)	387(15)
	TTA	0(0)	7(37)	0(0)	0(0)	0(0)	2(7)	0(0)	2(7)	12(2)	14.00(15)	586(23)
	CTG	7(37)	1(5)	7(18)	3(10)	8(38)	9(31)	29(38)	14(50)	480(83)	40.60(43)	587(23)
	CTA	0(0)	1(5)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	2(0)	4.47(5)	166(6)
	CTT	8(42)	6(32)	12(32)	9(31)	9(43)	12(41)	33(43)	8(29)	25(4)	14.20(15)	623(24)
	CTC	4(21)	2(11)	18(47)	16(55)	3(14)	6(21)	12(16)	4(14)	38(7)	10.67(11)	245(9)
Pro	CCG	3(38)	4(31)	2(10)	1(5)	1(8)	0(0)	2(7)	1(6)	190(77)	14.80(38)	372(38)
	CCA	0(0)	5(38)	2(10)	0(0)	2(17)	1(6)	1(3)	0(0)	36(15)	8.93(23)	208(21)
	CCT	1(13)	4(31)	11(55)	7(32)	1(8)	3(18)	10(34)	9(53)	19(8)	7.87(20)	311(32)
	CCC	4(50)	0(0)	5(25)	14(64)	8(67)	13(76)	16(55)	7(41)	1(0)	7.00(18)	88(9)

[a] *E. coli* high (Ecohigh) and low expression (Ecolow) genes from GenBank codon frequency file (1992).
[b] *B. subtilis* (Bsu) 105 genes from GenBank codon frequency file (1992).
[c] Xylanase gene (Rf17*xynA;* Zhang and Flint, 1992) and endoglucanase gene (Rf17*end;* Cunningham et al., 1991) from *R. flavefaciens* 17; endoglucanase genes (Ra*celA*, Ra*celB* and Ra*end;* Poole et al., 1990; Ohmiya et al., 1989) and β-glucosidase gene (Ra*bglT;* Ohmiya et al., 1990) from *R. albus*.
[d] % use of one codon relative to all others within one amino acid group is shown in parentheses. Single codon groups, such as Met or Trp, thus have a value of 100.

3.3.7 Codon usage. Codon usage by bacteria reflects the G+C content of their genomic DNA and their requirement for speed of translation (Grantham et al., 1981; Klump and Maeder, 1991). The codon usage in the *cel*A and *cel*E genes of *R. flavefaciens* FD-1 was compared with *E. coli* (Genbank codon frequency file, 1992), *B. subtilis* (Genbank codon frequency file, 1992) and cellulase genes from other *Ruminococcus* spp. (Ohmiya et al., 1989; Ohmiya et al., 1990; Poole et al., 1990; Cunningham et al., 1991; Zhang and Flint, 1992) (Table 3.2).

The third position of codons of *cel*A was analyzed and showed a preference for G or C while *cel*E showed a preference for A or T. In keeping with this the G+C content of the *cel*A nucleotides within the ORF (47.4%) is higher than *cel*E (35.4%).

The codon frequencies of *cel*A and *cel*E genes of *R. flavefaciens* FD-1 were compared with genes of *E. coli*, *B. subtilis* and cellulase genes of other *Ruminococcus* by computer using the CORRESPOND program from the GCG sequence analysis package (Devereux et al., 1984). The distance squared values show the relatedness of codon frequencies; the lower the value the higher the relatedness (Table 3.3). The *cel*A and *cel*E genes showed different codon frequencies. Correlation (distance squared value < 3.4) was found between *cel*A and the *R. flavefaciens* 17 endoglucanase, *R. albus*

endoglucanase, *R. flavefaciens* 17 xylanase, *R. albus* β-glucosidase and *E. coli* low expression genes. Correlation was also found between *celE* and the *R. albus* β-glucosidase, *B. subtilis* and *E. coli* low expression genes. The values also indicate the relatedness among the *R. albus* genes which show the least similarity to genes from *R. flavefaciens*.

Table 3.3 Correspondence analysis on frequencies by the usage of 64 codons among *Bacillus subtilis*, *E. coli* and *Ruminococcus*.

Gene	Distance squared value[a]										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
(1) Bsu	0	4.6	0.6	4.0	1.7	3.4	4.0	5.2	4.0	2.7	5.9
(2) Ecohigh		0	3.4	4.0	6.7	3.9	4.7	4.4	5.6	4.4	5.0
(3) Ecolow			0	2.5	2.3	2.5	3.3	3.7	2.7	2.0	4.4
(4) *celA*				0	5.2	3.4	3.6	1.6	2.1	3.1	2.6
(5) *celE*					0	3.6	3.8	6.7	5.0	2.9	6.8
(6) Ra*celA*						0	0.8	3.4	3.8	0.8	4.7
(7) Ra*celB*							0	3.2	3.3	0.8	3.9
(8) Rf17*endA*								0	2.3	3.6	1.2
(9) Ra*end*									0	2.8	2.1
(10) Ra*bglT*										0	4.2
(11) Rf17*xynA*											0

[a] Distances between messengers are constructed such that neighbouring mRNAs correspond to similar codon usages. The formula is defined as

$$d^2\ (i_1 i_2) = \sum_{j=1}^{64} (f_{i_1 j} - f_{i_2 j})^2\ ,$$ where $f_{i_1 j}$ and $f_{i_2 j}$ are the frequencies of codon j in mRNA i_1 and i_2. That is, the square of the distance between two messengers is the sum of the squares of their differences in use of each of the 64 codons.

It is of interest to note that the difference between *celA* and *celE* is relatively high.

3.4 Discussion.

Using the *E. coli/B. subtilis* vector pEB1, the *celE* gene from *R. flavefaciens* FD-1 was cloned and analyzed. The gene showed no cross-hybridization to DNA isolated from *R. flavefaciens* 17, *R. albus* 22.08.6A or *B. fibrisolvens* under conditions in which strong hybridization to FD-1 DNA was evident. This is not too surprising as the relatedness between *R.flavefaciens* FD-1 and 17 has not been accurately determined. Moreover there is often not sufficient similarity at the DNA level even between enzymes from the same cellulase family for hybridization to occur. Southern hybridization with genomic DNA also confirmed that the insert was cloned without deletion or rearrangement.

Although the 35 kDa endoglucanase had activity on CMC and xylan, it had barely detectable activity on pNPC. This is in contrast to the conclusion of Rasmussen et al. (1988) that pNPC hydrolysis is representative of both substrates for exoglucanase and endoglucanase activities in *R. flavefaciens* FD-1. It could be, however, that the activity of this enzyme is a minor one in *R. flavefaciens* FD-1 and was not detected in their assays.

Deletion analysis of pWF1 and subcloning of the insert into Bluescript SK indicated that the gene was transcribed from the λP_r promoter. Further evidence was obtained by Tn5

mutagenesis as inserts in the λP_r promoter resulted in loss of enzyme activities. This was confirmed when the nucleotide sequence revealed the absence of a consensus promoter sequence. A strong promoter in the vector is often required to obtain good expression in *E. coli* of genes from rumen anaerobes (Berger et al., 1990; Lin and Thomson, 1991; Thomson, unpublished observations).

Comparison of the deduced amino acid sequence of the *celE* gene with other cellulases showed that the only similarity was to an endoglucanase of the alkalophilic *Bacillus* sp. strain 1139 (EGF) which belongs to cellulase subfamily "A2" as defined by Henrissat et al. (1989) and Beguin (1990). However the homology with the N-terminal catalytic region of EGF was very low with only 11.6% identity and 55.3% similarity. This is not sufficient to be significant and it would therefore appear that CelE belongs to a new family of cellulases.

Most well-characterized extracellular proteins of Gram-positive strains appear to be synthesized as precursors with N-terminal signal sequences very similar in size and organization to Gram-negative secreted periplasmic proteins (Huang and Schell, 1992). Thus proteins secreted by Gram-positive bacteria can usually be translocated across the membrane of Gram-negative bacteria (Goldstein et al., 1990; Schatz and Bechwith, 1990; Huang and Schell, 1992).

Most signal peptides are cleaved immediately after a hydrophobic core and at a leader peptidase cleavage site A-X-B, where the position A may include Ala, Leu, Val and Ile, and position B may include Ala, Gly and Ser (Perlman and Halvorson, 1983). Although a cleavage site for the *R. flavefaciens* endoglucanase signal sequence was not determined, a predicted cleavage site could be after A_{27}-N_{28}-A_{29} or A_{29}-S_{30}-A_{31} sequences. In order for the enzyme to catalyze cellulose degradation it is presumably secreted by *R. flavefaciens* FD-1 via this leader sequence.

CHAPTER 4

TRANSCRIPTIONAL REGULATION OF THE CELLODEXTRINASE AND ENDOGLUCANASE GENES IN *R. FLAVEFACIENS* FD-1

CHAPTER 4

TRANSCRIPTIONAL REGULATION OF THE CELLODEXTRINASE AND ENDOGLUCANASE GENES IN *R. FLAVEFACIENS* FD-1

4.0 Summary.

R. flavefaciens FD-1 was grown in cellobiose, cellotriose or cellulose. Both *cel*A and *cel*E genes were used as probes to analyze transcription in *R. flavefaciens* grown under different conditions. Transcription of both genes was induced when cellulose was added to cells growing on cellobiose. This induction continued after cellulose depletion and after cell division had ceased. Transcription of both genes was also induced by cellotriose although the effect was not as pronounced as induction by cellulose and was greater for the *cel*A gene, than for the *cel*E gene.

4.1 Introduction.

In an early report on the regulation of cellulases in *R. flavefaciens* and *R. albus*, Fusee and Leatherwood (1972) concluded that cellobiose appeared to repress enzyme synthesis. However further studies reported that cellulases were produced constitutively and that while cellobiose inhibited enzyme activity it did not repress enzyme synthesis (Smith et al., 1973; Pettipher and Latham, 1979b; Hiltner and Dehority, 1983). More recently, Gardner et al. (1987) reported that cellobiose inhibited the catalytic activity of an exoglucanase from *R. flavefaciens* FD-1. Rasmussen et al. (1988) also showed that methylcellulose, which is resistant to biological degradation in anaerobic environments, acts as a competitive inhibitor of endo- and exocellulase but not of cellodextrinase activity.

During a study of gene regulation in *R. flavefaciens* 17 Flint et al. (1991) showed that expression of two xylanase genes was induced by xylan. Recently, Doerner et al. (1992) reported that cellulose induced the expression in *R. flavefaciens* FD-1 of an endoglucanase and a bifunctional cellulase/xylanase gene but not of the *cel*A gene or a gene encoding an enzyme with CMCase and xylanase activity. However it was unclear how insoluble cellulose induces the transcription of cellulase genes.

It has been reported that the main products of cellulose degradation by *R. flavefaciens* are cellotriose and cellobiose (Rasmussen et al., 1988). Transport mechanism(s) exist in this organism for the uptake of cellobiose and a similar mechanism may exist for cellotriose as cells can utilize cellobiose and cellotriose (Rasmussen et al, 1988; Helaszek and White, 1991). This chapter is the result of an investigation as to whether cellobiose and cellotriose affected the transcription of cellulase genes. *CelA* and *celE* genes were used as probes to investigate RNA levels in cells grown on cellobiose, cellulose and cellotriose. The choice of cellobiose was based on the findings that *R. flavefaciens* appears to lack an extracellular β-glucosidase, is unable to metabolize extracellular glucose, but grows well on cellobiose (Rasmussen et al., 1988).

4.2 Materials and Methods.

4.2.1 Media, buffers, and growth conditions. Late logarithmic phase cultures of *R. flavefaciens* grown in cellobiose (0.1% w/v) were diluted 1:100 in fresh cellobiose medium and grown for 36 hours. Ball-milled filter paper (0.1% w/v), cellobiose (0.01% or 0.1% w/v), or cellotriose (0.01% or 0.1% w/v) were added and the cultures grown for 40 hours. Upon depletion of the cellulose, cellobiose (0.1% w/v) was added to the ball-milled filter paper cultures and incubation continued for a further 24 hours. Control cells were grown in cellobiose medium alone. Growth condition were as described in chapter 2.

4.2.2 RNase-free work. Chemicals and reagents used for RNA work were purchased solely for this purpose. All new glassware used for RNA work was baked overnight at 180°C. When appropriate, solutions were treated with 0.2% (v/v) diethylpyrocarbonate (DEPC), allowed to stand overnight at 37°C, and autoclaved for 15 min at 15 lb/sq in on the liquid cycle. All plastic was sterilized by autoclaving after 0.2% DEPC pretreatment. Tanks used for electrophoresis of RNA were treated with detergent solution overnight, rinsed completely in sterile water, and dried with ethanol.

4.2.3 Bacterial growth measurements. Cells were counted using a counting chamber under phase contrast microscopy.

Optical density readings at 600 nm were taken on a Coring colorimeter 252 under anaerobic condition.

4.2.4 Isolation of RNA from R. flavefaciens. Cell samples were taken at various time intervals during growth and RNA isolated according to a modification of the method of Aiba et al, (1981). Cells were harvested and resuspended in 0.02M sodium acetate (pH 5.5), 0.5% SDS and 1 mM EDTA. After addition of equilibrated phenol, the mixture was incubated at 60°C for 5 min with gentle shaking. After centrifugation the aqueous phase was re-extracted with phenol and phenol/chloroform. The RNA was precipitated by adding three volumes of ethanol to the aqueous phase and chilling at -70°C for 30 min. The RNA precipitate was collected by centrifugation and redissolved in the same acetate/SDS buffer, while the aqueous phase was extracted with phenol and phenol/chloroform. Ethanol precipitation was performed at -70°C overnight. After centrifugation, the precipitate was collected and redissoved in DNAse I buffer (0.1 M sodium acetate, 5 mM $MgSO_4$, pH 5.0) and incubated with RNase inhibitor and RNase-free DNAseI (final concentration 1 U/µl; Boehringer Mannheim) at 37°C for 60 min. The mixture was extracted with phenol and phenol/chloroform and ethanol precipitated at -70°C overnight. The RNA banding patterns were checked by formaldehyde/agarose gel electrophoresis and RNA concentration determined by measuring the optical density at

260 nm. Concentration was calculated as 1 OD_{260} = 40 µg/ml RNA.

4.2.5 Northern blots and RNA dot blots. RNA Northern and dot blotting were performed according to Amersham's recommended procedures. Samples of each *R. flavefaciens* RNA preparation were incubated at 65°C for 5 min in three volumes of the sample solution (66% deionized formamide, 2.6 M formaldehyde, 20 mM MOPS, 5 mM sodium acetate and 1 mM EDTA), chilled on ice, and then one volume of cold 20 X SSC was added. For Northern blotting, 20 µg samples of each RNA preparation were electrophoresed in 1% agarose-6.3% formaldehyde gels at 20 to 40 mA with recirculation of runnning buffer. Gels were transferred onto Hybond-N^+ membranes by alkali blotting procedures with the appropriate RNA alkali transfer buffer (0.05 M NaOH for 3 hours). The sizes were estimated by reference to the migration of DNA molecular weight markers (lambda *Hin*dIII). For dot blotting, 5 µg samples were applied to Hybond-N^+ membranes using a dot blotting apparatus. The RNA was fixed onto the membrane by alkaline fixation (0.05 M NaOH for 5 minutes). Prehybridization and hybridization were performed essentially according to the Amersham's recommended procedures. Prehybridization was performed in a shaking water bath at 65°C for 1 hour.

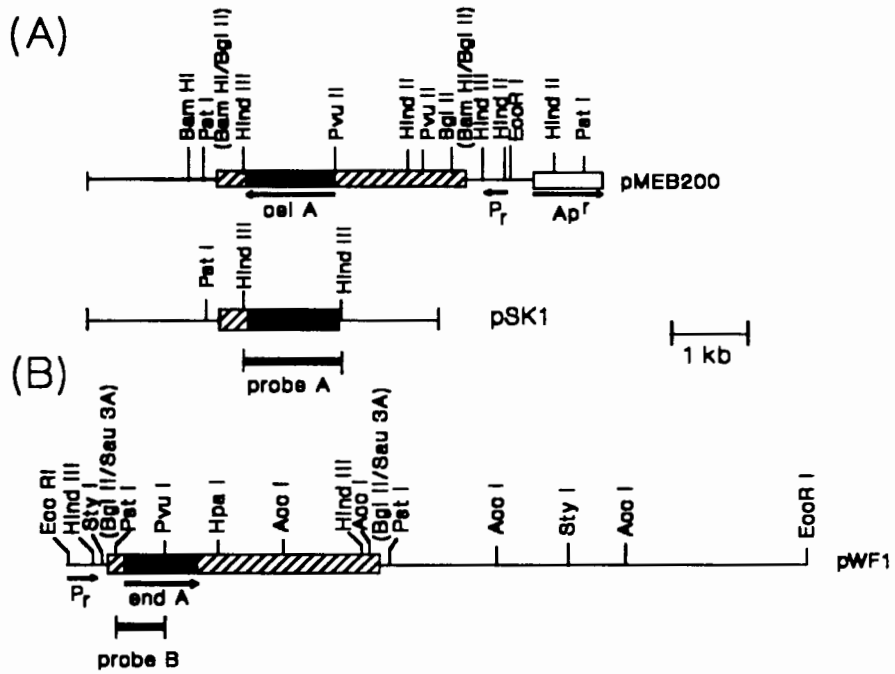


Fig. 4.1 Restriction map of (A) pMEB200 and (B) pWF1 showing the position of the hybridization probes.

The *cel*A probe was a 1.2 kb *Hind*III fragment of pSK1 and the *cel*E probe was a 0.7 kb *Pst*I-*Pvu*I fragment of pWF1 (Fig. 4.1). The sizes were estimated by reference to the migration of DNA molecular weight markers (lambda *Hind*III). After electroelution (Sambrook, et al., 1989), the fragments were labelled with α-^{32}P-dCTP using a nick translation kit (Amersham). The membranes were hybridized with 10^6 dpm of denatured probes ($6x10^7$ dpm/pmol of DNA) at 65°C for 20 hours. Blots were washed in 2 x SSC, 0.1% SDS at 65°C for 10 min. Time exposures of northern blot of *cel*A and *cel*E were 78 hours and 136 hours respectively. Time exposures of

northern blot of *cel*A and *cel*E were 96 hours and 240 hours respectively. After autoradiography, the films were scanned with a custom-built densitometer for quantitation (Retief et al., 1987).

4.3 Results.

4.3.1 Effect of carbon sources on growth. *R. flavefaciens* FD-1 was grown in non-rumen fluid (see Appendix A) with cellobiose as sole carbon source. During the late logarithmic phase of growth (36 hours), cellulose was added and cell numbers increased from $8x10^8$ to $1.6x10^9$ until microscopic examination of samples indicated that the cellulose had been depleted (76 hours). At this stage cellobiose was added and a further slight increase in cell numbers occurred to $2.1x10^9$ (88 hours) (Fig. 4.2A). In a subsequent experiment, *R. flavefaciens* FD-1 was grown in cellobiose for 36 hours, after which cellotriose (0.01% or 0.1%) was added (Fig. 4.2B).

4.3.2 Visualization of the growth phases. Insoluble cellulose degradation by *R. flavefaciens* FD-1 was apparent in Hungate tubes (Fig. 4.3A) and was characterized by a bright yellow pigment. Photomicrographs were taken of the cultures at various time intervals. At 36 hours the cells were present as chains (Fig. 4.3B). At 48 hours, 12 hours after the addition of cellulose, many diplococci could be seen adhering to cellulose fibres (Fig. 4.3C; Fig. 4.4A). By 60 hours most of the cellulose had been degraded and cells were present as diplococci and short chains, with some adhering to the remaining cellulose fibres (Fig. 4.4B).

After 76 hours, insoluble cellulose could no longer be detected (Fig. 4.4C).

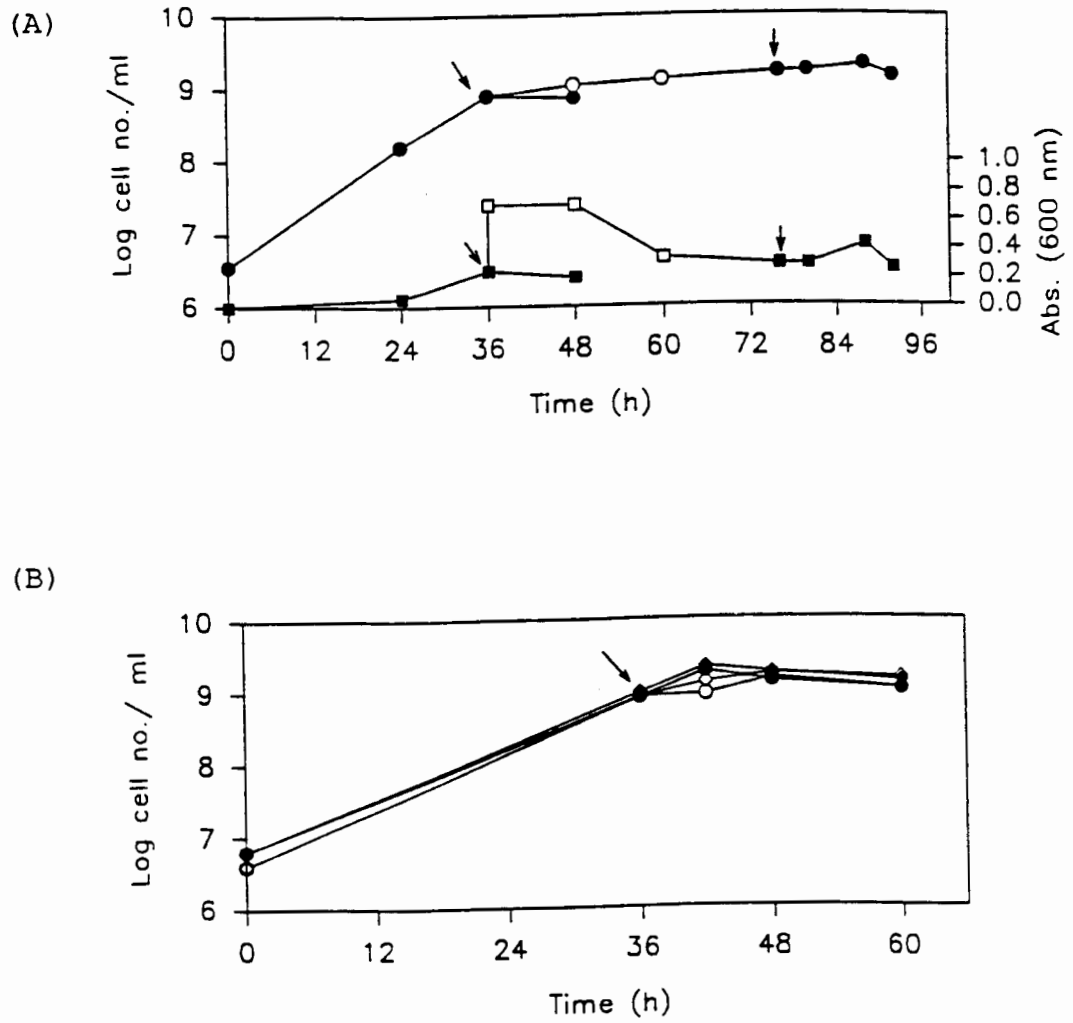


Fig. 4.2 (A) Growth of FD-1 cells measured by log. cell numbers (circles) under phase constrast microscopy and detected by OD_{600} (squares) under anaerobic conditions. Arrows indicate the addition of 0.1% cellulose (36 h) and 0.1% cellobiose (76 h). Cells growing in cellobiose (filled symbols); or cellulose (open symbols). Samples were taken at the times indicated by the symbols. (B) Growth of FD-1 cells measured by log. cell numbers (circles and diamonds) under phase constrast microscopy. Arrows indicate the addition of cellobiose (circle; 36 h) or cellotriose (diamond; 36 h) at concentrations of 0.01% (open symbols) or 0.1% (filled symbols). Samples were taken at the times indicated by the symbols.

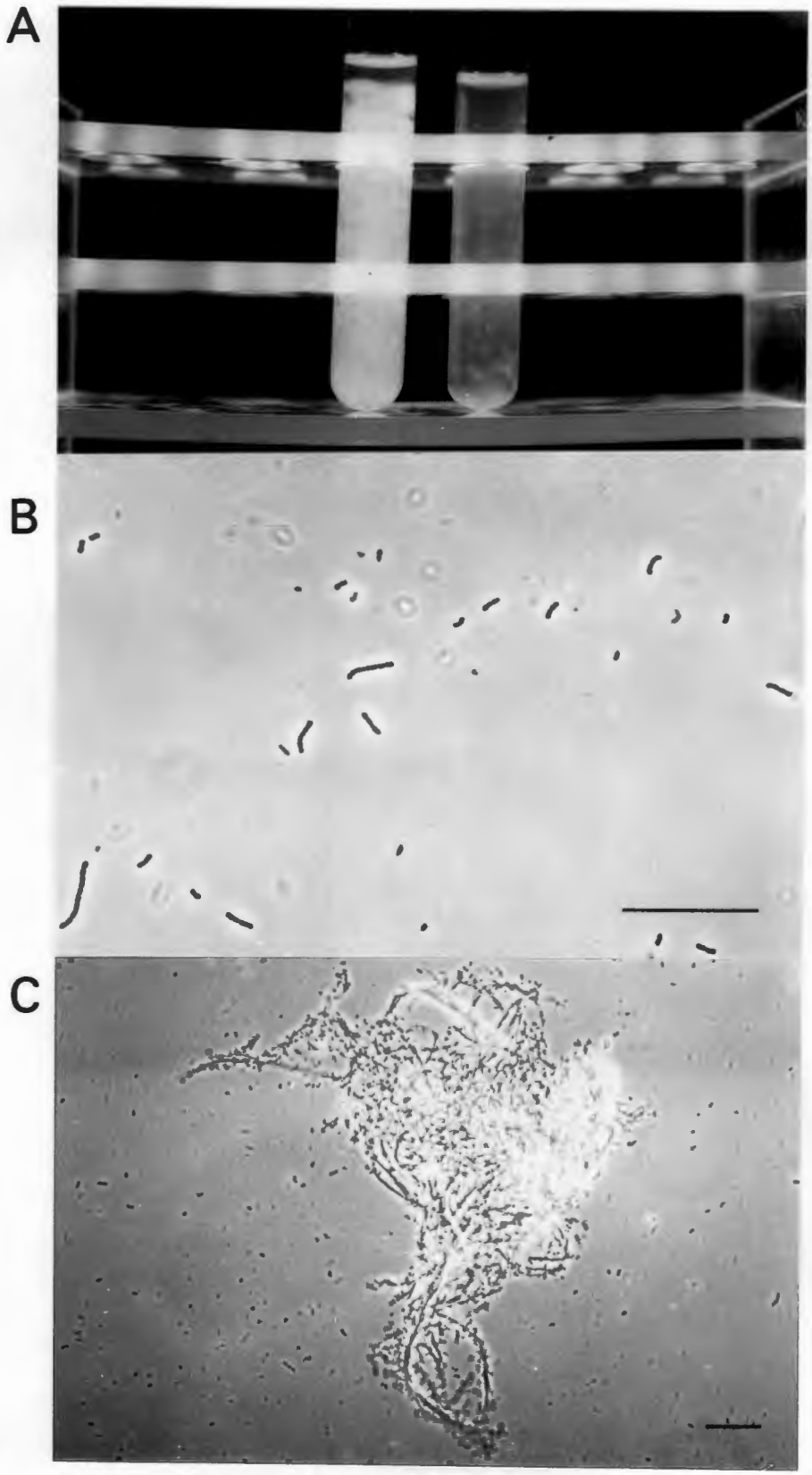


Fig. 4.3 Cellulose degradation by *R. flavefaciens* FD-1 (A) in Hungate tubes at 0 h (left tube) and 24 h (right tube), and by visualisation of the growth under phase contrast micrography at (B) 36 h in cellobiose and (C) 48 h, 12 h after the addition of cellulose. Bar = 20 μm.

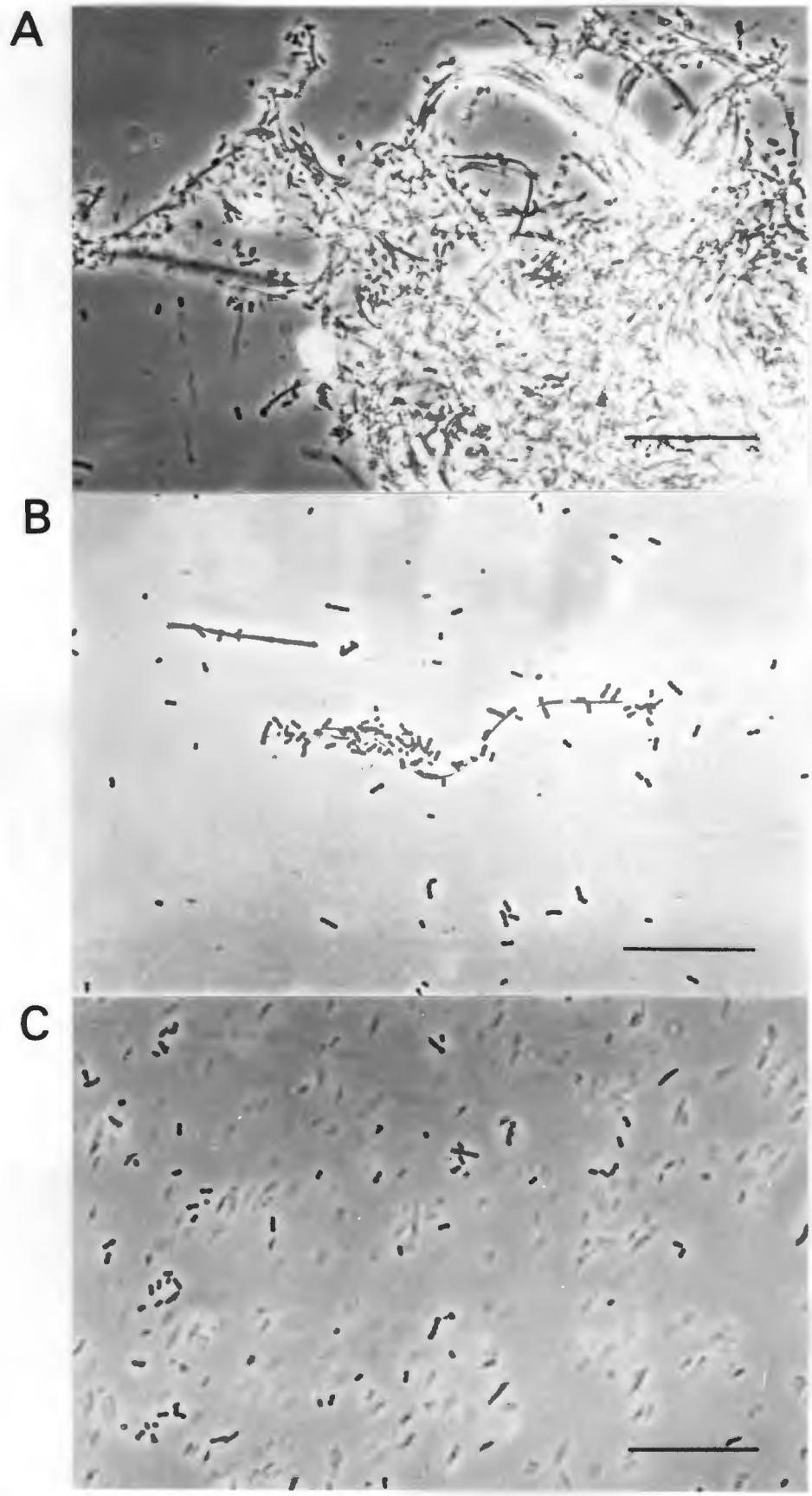


Fig. 4.4 Visualisation of the growth of *R. flavefaciens* FD-1 under phase contrast micrography at (A) 48 h, 12 h after the addition of cellulose; (B) 60 h, 24 h after the addition of cellulose and (C) 76 h, 40 h after the addition of cellulose. Bar = 20 μm.

4.3.3 Regulation of celA and celE gene expression in R. flavefaciens FD-1. Northern blots of RNA isolated from cells grown in cellobiose for 36 hours showed the absence of *cel*A and *cel*E specific transcripts. However, after the addition of cellulose and growth for a further 40 hours, *cel*A and *cel*E specific transcripts could be detected (Fig. 4.5). The *cel*A and *cel*E probes hybridized to mRNA transcripts from *R. flavefaciens* FD-1 were approximately 1100 base pairs long which is comparable to the sizes estimated from both sequences. RNA dot blot hybridization (Fig. 4.6A and B) showed that RNA homologous to *cel*A and *cel*E was present at very low levels during the first cellobiose growth phase. Upon addition of cellulose there was a slow increase in the levels of both RNA transcripts, which reached a peak at 76 hours when cellulose was depleted and cellobiose was added. Thereafter the levels of both RNA transcripts declined somewhat but were still readily detectable for a further 16 hours. In the culture grown only on cellobiose both RNA transcripts remained at a very low level throughout the 48 hours period (Fig. 4.6C), by which stage the cells had reached stationary phase (Fig. 4.2A).

4.3.4 Cellotriose induction. When cellotriose was substituted for cellulose, similar slow increases in RNA homologous to *cel*A and *cel*E were detected although the levels of induction were lower than those induced by

cellulose (Fig. 4.7). In addition, while induction of *celA* was detectable in cells grown in 0.01% cellotriose (Fig. 4.7A and C), increases in RNA homologous to *celE* were only detectable during growth in 0.1% cellotriose (Fig. 4.7B and D).

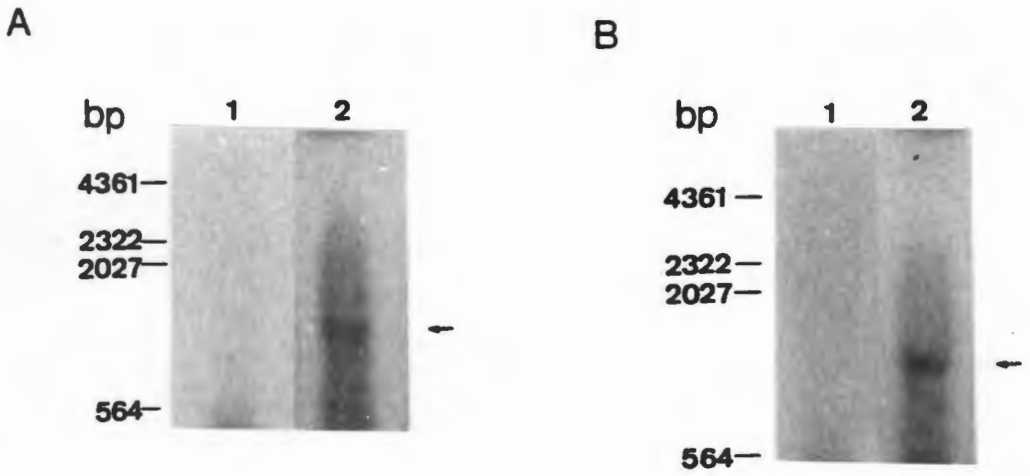


Fig. 4.5 Northern blot analysis using *celA* (A) and *celE* (B) as probes. RNA isolated after 36 h growth in cellobiose (lane 1) and after addition of cellulose and further 40 h growth (lane 2). Molecular weight markers in base pairs are indicated on the left.

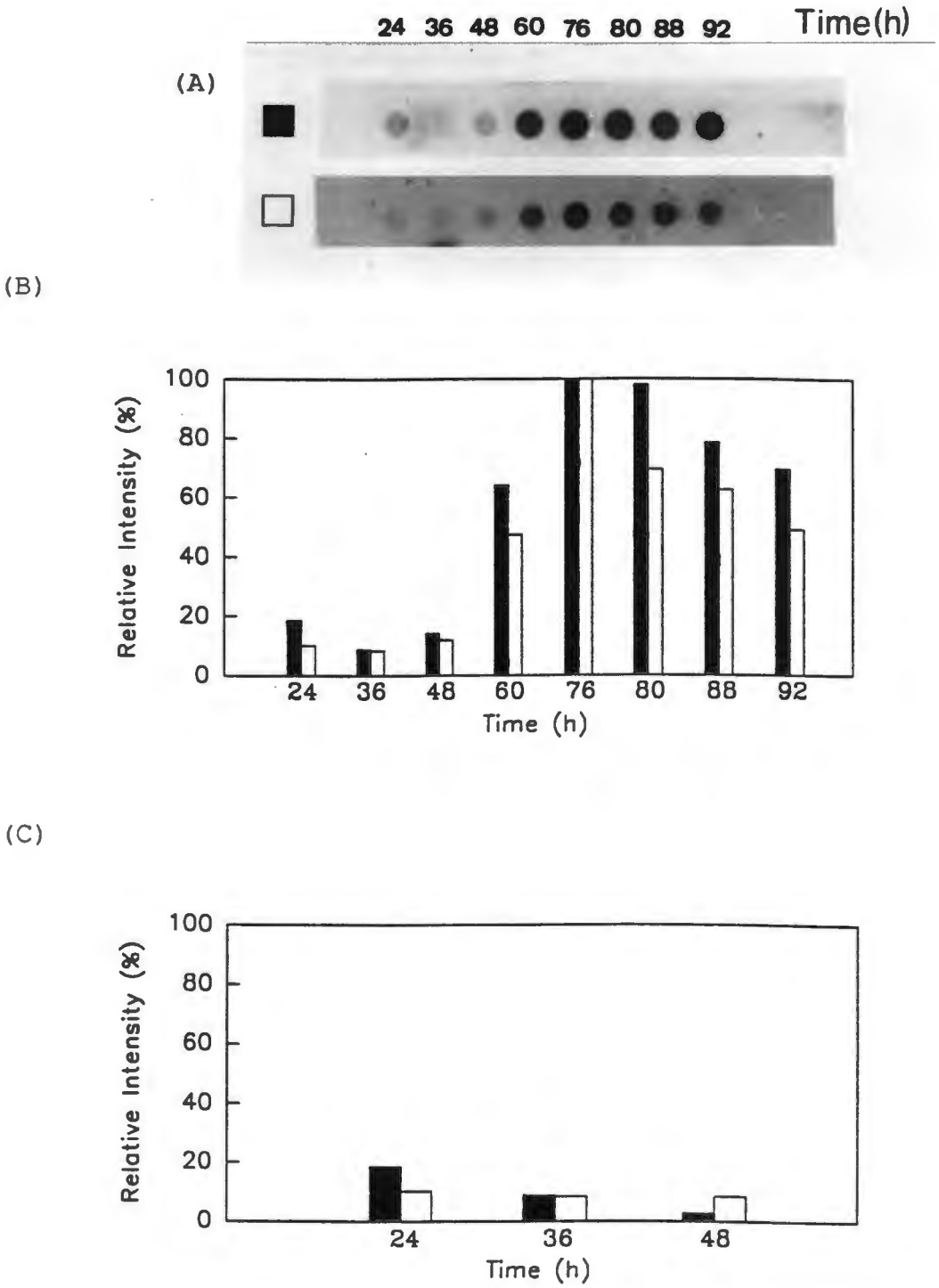


Fig. 4.6 Dot blots (A) and relative intensities (B, C) of RNA extracted from *R. flavefaciens* FD-1 cells grown in 0.1% cellobiose with the addition at 36 h of 0.1% ball-milled filter paper and at 76 h of 0.1% cellobiose (B), or 0.1% cellobiose only (C) and probed with *cel*A (■) or *cel*E (□). RNA was extracted from cells at the indicated time intervals.

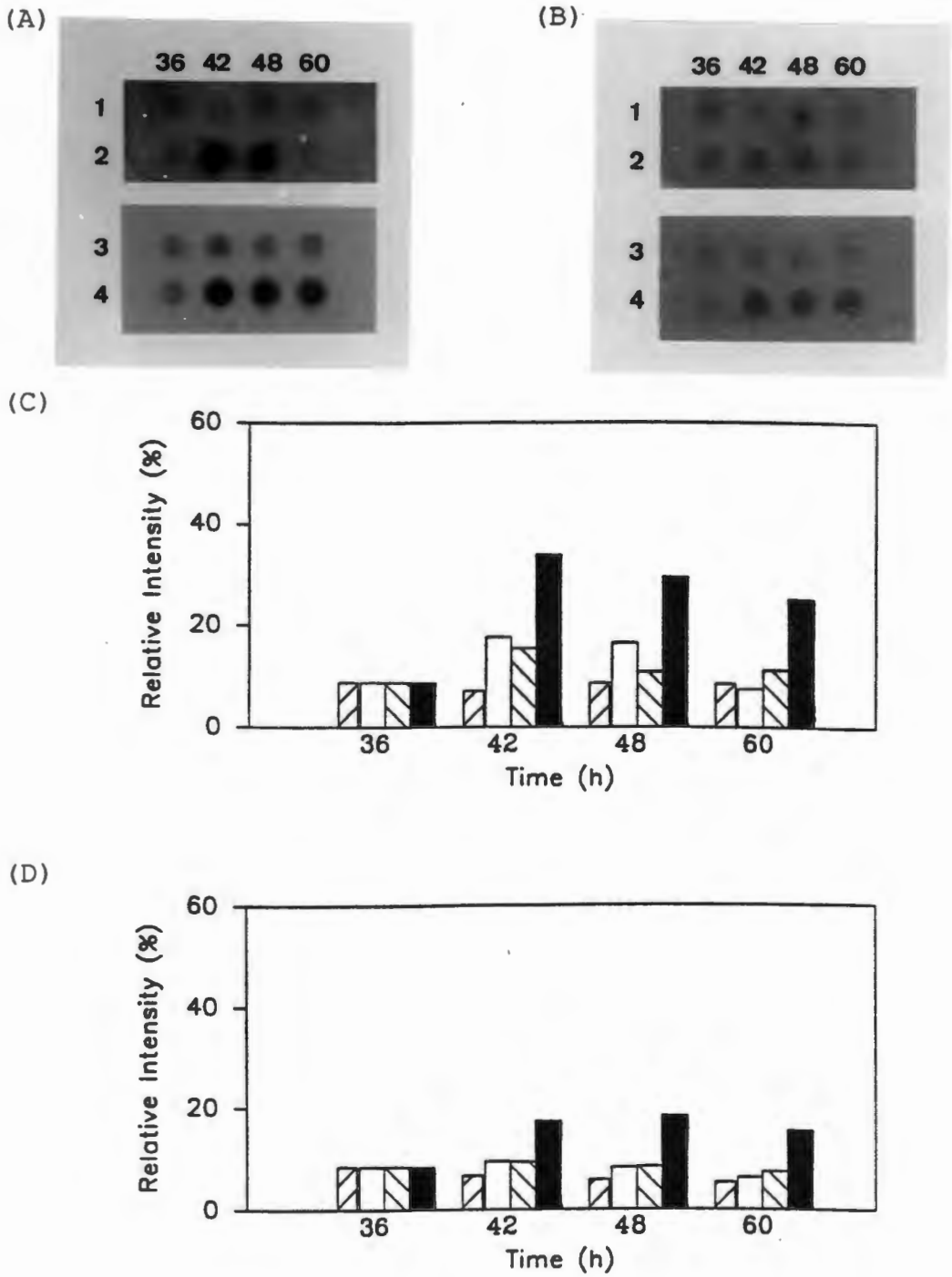


Fig. 4.7 Dot blots (A, B) and relative intensities (C, D) of RNA extracted from *R. flavefaciens* FD-1 cells grown in 0.1% cellobiose with the addition at 36 h of 0.01% cellobiose (1, ▨); 0.01% cellotriose (2, □) ; 0.1% cellobiose (3, ▧); 0.1% cellotriose (4, ■) and probed with *celA* (A, C) and *celE* (B, D). RNA was extracted from cells at the indicated time intervals.

4.4 Discussion.

A low molecular weight yellow substance promotes the binding of endoglucanase to cellulose during growth of *C. thermocellum* on cellulose (Ljungdahl and Eriksson, 1985). It has been reported that many strains of *R. flavefaciens* produce yellow to yellow-orange pigments especially when grown on cellulose (Bryant, 1986). Such pigmentation was also observed in this study, and it may be speculated that its function is to enhance cellulose degradation.

R. flavefaciens FD-1 was seen to grow as chains of cells in cellobiose but as diplococci or short chains in cellulose, tending to adhere to the insoluble substrate. It has been suggested that export of enzymes from their intracellular locations may occur via a secretory mechanism mediated by contact with cellulose (Berg and Pettersson, 1977; McGavin et al., 1990; Hrmova et al, 1991). It has also been found that the adhesion of *R. albus* appeared to be a prerequisite for effective cellulose degradation (Morris and Cole, 1987), although adhesion was not necessarily followed by cellulolysis (Morris, 1988). Evidence has been obtained that cellulose with higher levels of binding is digested more readily by *Ruminococcus* (Leatherwood, 1973). It may be speculated that *R. flavefaciens* FD-1 adhesion to cellulose represents a competitive advantage for the cells which can utilize the end products *in situ*.

It is unclear why most of the *R. flavefaciens* FD-1 cells form chains during stationary phase. One explanation is that, being a non-motile bacterium, by forming chains they can swim and hunt more widely because of their contacting surface and rotating angle, especially during food shortages. The second explanation is that for better competition *R. flavefaciens* FD-1 cells join together as chains during food shortages to save energy, to share the food with each other in a short distance and short time, and thus increase their survival (Stewart and Bryant 1988).

The data presented here suggest that expression of both *celA* and *celE* is induced by growth of *R. flavefaciens* FD-1 on cellulose. This is in contrast to the conclusion drawn by Doerner et al., (1992) that *celA* is constitutively expressed in cellobiose-grown cultures. However their data were not quantitative and RNA was extracted eight hours after the addition of cellulose. This study showed that induction by cellulose takes from 12 to 24 hours and suggests that neither *celA* nor *celE* is constitutively expressed. However, there are presumably constitutively expressed cellulases which, upon addition of cellulose, degrade it sufficiently to produce cellodextrins which can be taken up by the cells and induce transcription of these genes. The uptake of cellotriose by *R. flavefaciens* FD-1 has been postulated by Helaszek and White (1991).

To test this, cellotriose was substituted for cellulose. The pattern of induction by cellotriose of both the *celA* and *celE* genes was similar to that by cellulose, except that the level of induction was considerably lower. This is in keeping with the hypothesis that induction by cellulose requires its breakdown by constitutive enzymes to cellodextrins, including cellotriose, which may then be taken up by cells to induce gene expression. Addition of cellotriose instead of cellulose allowed more rapid uptake and induction of *celA* and *celE*, but the lower levels of induction could have a number of explanations. These include the fact that cellotriose is a metabolizable substrate and the possibility that cellodextrins other than cellotriose can be taken up and act as inducers.

In similar studies on induction of cellulase gene expression in *Cellulomonas fimi*, Greenberg et al. (1987a) have shown that the *cenA* gene which encodes an endoglucanase is transcribed at low levels when the cells are grown in glycerol, at high levels in carboxymethyl cellulose (CMC) but not at all in glucose. In contrast, the *cex* gene, which encodes an exoglucanase, is transcribed only during growth in CMC. They concluded that the *cenA* gene was being expressed at a low constitutive level during growth on glycerol and its function, and that of other constitutively expressed endoglucanases, is presumably to generate low

molecular weight cellulose-specific degradation products which can act as inducers once a suitable substrate is encountered. They ascribed the lack of transcription during growth in glucose to indicate repression at the transcriptional level by this readily assimilated carbon source. In contrast to the *cen*A gene, the *cen*B gene, which also encodes an endoglucanase, was transcribed during growth in glucose, albeit at a low level (Greenberg et al., 1987b).

During growth of *R. flavefaciens* in cellobiose alone, levels of both RNA transcripts remained low. This is unlikely to be due to catabolite repression as cellobiose and cellotriose are the final breakdown products of cellulose by *R. flavefaciens* (Rasmussen et al., 1988) and cellobiose will therefore be present throughout growth on cellulose. Moreover, addition of cellobiose after depletion of cellulose did not result in a marked decrease in transcription of the two genes. It is likely that the RNA detected after cellulose depletion is due to continued induction of transcription by the cellodextrins present in the medium. It is unlikely to represent stable RNA as, although mRNA turnover has not been measured in *R. flavefaciens*, it would probably not be stable for 16 h. Thus induction of *cel*A and *cel*E expression continues after cessation of cell division.

A number of cellulase genes have been shown to be induced by sophorose (Mandels et al., 1962; Hrmova et al., 1986; El-Gogary et al., 1989; Hrmova et al., 1991). This compound has not been tested as, in order to produce sophorose from cellulose, a microorganism must hydrolyse cellulose to glucose and then transglycosylate the glucose (Kubicek, 1987). As there are no reports of β-glucosidase production by *R. flavefaciens* and the end-products of cellulose breakdown are cellobiose and cellotriose (Rasmussen et al., 1988; Helaszek and White, 1991), it is unlikely that this bacterium could produce sophorose from cellulose.

CHAPTER 5

GENERAL DISCUSSION

Cellulolytic bacteria are potentially very useful for the bioconversion of lignocellulosic wastes. In spite of the progress made in recent years, relatively little is known about cellulase genes and mechanisms of action of cellulases of the important rumen cellulolytic bacterium, *R. flavefaciens* FD-1. To date five cellulase genes, the cellodextrinase (*cel*A) and β-glucanase (*cel*E) reported here, two β-glucanase (*cel*B and *cel*C) and a bifunctional β-glucanase/xylanase (*cel*D), have been cloned from this strain (Howard and White; 1990; Wang and Thomson, 1990; Doerner et al., 1992; Wang et al., submitted). Only two of the sequences, those of *cel*A and *cel*E, have been submitted to the Genbank database.

CelA, the cellodextrinase of *R. flavefaciens* FD-1, is in the same subfamily as endoglucanase C of *C. thermocellum* and endoglucanase 3 of *F. succinogenes*. The latter two have some cellodextrinase activity but their preferred substrates are larger polymers with β-1,4 and β-1,3 linkages (Schwarz et al., 1988b; McGavin et al., 1989). Three dimensional analysis of the enzymes is needed to compare their structures, active site(s) and substrate binding site(s). CelE does not show significant homology to any of the more than 60 cellulase and xylanase genes searched in the GenBank

GenBank database. It therefore appears to belong to a new family of such enzymes.

Although CelA lacks a signal sequence much of the enzyme is found in the *E. coli* periplasm. The question therefore arises concerning its location in *R. flavefaciens* FD-1. If the combined action of cellodextrinase and other cellulase components is required for effective hydrolysis of cellulose, all these enzymes might have to be exposed to the substrate outside the cell. However if degradation of cellulose to cellodextrins and the breakdown of cellodextrins are separate processes, and cellodextrins can be taken up by the cell, it is possible that the cellodextrinase could function in the cytoplasm. To determine the enzyme's localization, immunofluorescent labelling of antibodies against CelA could be used combined with transmission electron microscopy. This approach has been used to visualize the localization of the cellulolytic enzyme system in *C. thermocellum* (Nolte and Mayer, 1989) and *F. succinogenes* subsp. *succinogenes* S85 (Huang and Forsberg, 1990; McGavin et al., 1990)

To assist purification of CelA from the periplasm a signal sequence could be spliced into the 5' end region of the *cel*A sequence using the convenient *Ssp*I restriction enzyme site. This blunt-end cleavage site is located only three base pairs upstream of the *cel*A structural gene. The three-

dimensional structure of the catalytic core of cellobiohydrolase CBHII from *T. reesei* (Rouvinen et al., 1990) and endoglucanase CelD from *C. thermocellum* (Juy et al., 1992) have been determined. Studies are in progress to crystallize and determine the three dimensional structure of the CelA enzyme.

One approach to analyzing the functions of the individual components of a cellulase system is to mutate the genes and determine the effect on cellulose degradation. This is very difficult with an organism such as *R. flavefaciens* which produces a number of different cellulases. However gene cloning may be used to construct mutants by gene replacement. A cloned gene may be mutated in a heterologous host and the mutation recombined into the chromosome of *R. flavefaciens* FD-1. This, however, requires the establishment of a DNA transfer system between it and either *E. coli* or a Gram-positive bacterium such as *B. subtilis*, or the development of methods such as electroporation for transfer of the mutated gene.

R. flavefaciens FD-1 produces a yellow pigment when grown on cellulose (Bryant, 1986). To determine the role of this pigment in cellulose degradation, biochemical analysis and characterization of the pigment would be advantageous followed by molecular analysis of the genes involved in pigment production may be helpful.

Induction of *celA* and *celE* by cellotriose was found to occur at a lower level than induction by cellulose. This could be due to the fact that cellotriose is metabolised by *R. flavefaciens* or it may be that cellodextrins other than cellotriose can be taken up by the cells and act as inducers. Similar experiments to the ones performed here should therefore be done with higher molecular weight cellodextrins.

The study presented here was aimed at broadening our understanding at the molecular level of cellulose degradation by *R. flavefaciens* FD-1. To increase this knowledge further more cellulase genes need to be cloned and their regulation studied.

APPENDIX A

MEDIA, BUFFERS AND SOLUTIONS

APPENDIX A

MEDIA, BUFFERS AND SOLUTIONS

All media, buffers and solutions were sterilized by autoclaving at 121°C for 20 min unless otherwise indicated. Heat labile substances were sterilized by filtration through 0.22 µm membrane filters (Millipore).

A.1 Media

A.1.1 Non-Rumen Fluid Medium [a] for *R. flavefaciens* FD-1

Component	Percentage in medium [b]
Cellobiose,cellotriose or cellulose	0.1
Mineral solution [c] (V/V)	3.75
Hemin [c]	0.0001
Indigo carmine	0.0005
Yeast extract [c]	0.05
Trypticase [c]	0.2
Na_2CO_3 [c]	0.4
Dithiothreitol	0.02
Volatile Fatty Acid [c] (V/V)	0.31
Trace elements solution [d] (V/V)	0.4
Vitamine solution [e] (V/V)	1

[a] Prepared under 100% CO_2; final pH 6.8.

[b] Weight/volume unless otherwise indicated.

[c] Solution from Caldwell and Bryant (1966).

[d] Trace elements solution was dissolved in (0.2 M) HCl and contained (in g per 500 ml) : $MnSO_4.H_2O$, 0.56; $CoCl_2.6H_2O$, 0.183; $FeSO_4.7H_2O$, 0.183; $CuSO_4$, 0.01; $ZnSO_4.7H_2O$, 0.178; $Al_2(SO_4)_3.15H_2O$, 0.02; H_3BO_3, 0.01; $Na_2MoO_4.2H_2O$, 0.01175; $Na_2SeO_3.5H_2O$, 0.00329; Na_2SiO_3, 0.00154; $NiCl_2.6H_2O$, 0.00298; $Na_2WO_4.2H_2O$, 0.00042; nitrilotriacetic acid, 1.5.

[e] Vitamine solution was dissolved in 15 mg/ml NH_4HCO_3 solution and contained (in g per 100 ml): thiamine.2HCl, 0.02; Ca-pantothenate, 0.02; pyridoxamine.2HCl, 0.02; pyridoxine.HCl, 0.02; p-aminobenzoic acid, 0.001; cyanocobalamine, 0.0001; nicotinamide, 0.02; D-biotin, 0.001; folic acid, 0.0005; phenylpropanoic acid, 0.0375; riboflavin, 0.02.

[f] The dithiothreitol, trace elements solution, vitamin solution, volatile fatty acid and Na_2CO_3 were filter sterilized (0.22 µm pore size) under N_2 and aseptically added to sterile medium.

A.1.2 Luria-Bertani medium (LB)

Bacto tryptone	16 g
Yeast extract	5 g
NaCl	5 g
Distilled water	1000 ml

Solid media contained 1.5% (w/v) agar.

A.1.3 YT medium (X2)

Bacto tryptone	16 g
Yeast extract	10 g
NaCl	5 g
Distilled water	1000 ml

A.2 Media additives

Media were cooled to 50°C before addition of antibiotics, XGal, XP, or IPTG. Plates containing these additives were stored for no longer than one week at 4°C.

A.2.1 Antibiotics

Antibiotic stock solutions were as follows:

Ampicillin	100 mg/ml water
Kanamycin	62.5 mg/ml water

All antibiotics were filter sterilized and stored at -20°C.

A.2.2 IPTG (Isopropyl-b-D-thio-galactopyranoside)

IPTG (100 mM)	23.4 mg
Distilled water	1 ml

The solution was stored in aliquots at -70°C.

A.2.3 XGal (5-bromo-4-chloro-3-indolyl-b-galactoside)

XGal	0.2 g
DMSO	10 ml

The solution was stored at -70C.

A.2.4 XP (5-bromo-4-chloro-3-indolyl phosphate)

XP	80 mg
DMSO	1 ml

The solution was stored at -70°C.

A.3 Buffers and solutions

A.3.1 ATP (10X) (Maniatis et al., 1982)

Adenosine triphosphate	30 mg
Distilled water	5 ml

Adjusted pH to 7.0 with 0.1N NaOH before making up to 5 ml. Stored in 100 ml aliquots at -70°C. Discarded remainder once defrosted.

A.3.2 Azocasein

Azocasein	0.8 g
Phosphate buffer (0.1 M, pH 6.8)	100 ml

Stirred overnight to dissolve azocasein

A.3.3 Bradford solution (Bradford, 1976)

Coomassie Brilliant Blue (G-250)	100 mg
Ethanol (95%)	50 ml

Dissolved, then added 100ml phosphoric acid (85%). Diluted to final volume of 1 l. Filtered through Whatman GF/C filter paper. Stored in dark bottle.

A.3.4 Denhardt's solution (10X) (Maniatis et al., 1982)

Ficoll	1 g
Polyvinylpyrrolidone	1 g
BSA (Fraction V)	1 g
Distilled water	to 100 ml

The solution was stored in aliquots (10 ml) at -20°C.

A.3.5 Dinitrosalicylic acid solution (DNS) (Miller, 1959)

3,5 Dinitrosalicylic acid	10.6 g
NaOH	19.8 g
Rochelle salts (Na K Tartrate)	306 g
Phenol	7.6 ml
Na-meta bisulphite	8.3 g
Distilled water	1416 ml

The dinitrosalicylic acid, NaOH, and Rochelle salts were dissolved completely in water before adding the other constituents and dissolving each in turn. The phenol was melted at 50°C. A 3 ml sample was titrated to the end-point with 5 - 6 ml HCl (0.1 M) using phenolphthalein as an end-point indicator. However, if less HCl was required then solid NaOH was added to the DNS solution at the rate of 2 g/ml of HCl less than 5 ml, and the titration repeated. The DNS solution was stored in a dark bottle under N_2.

A.3.6 DNA loading solution (6X)

Bromophenol blue	0.25	g
Sucrose	40	g
Distilled water	to 100	ml

The solution was stored at 4°C

A.3.7 EDTA (0.5 M, pH 8.0) (Maniatis et al., 1982)

EDTAω$2H_2O$	168.1	g
Distilled water	to 1000	ml

EDTA was only dissolved when pH was adjusted to 8.0

A.3.8 Ethidium bromide solution (2,7-diamino-10-ethyl-9-phenyl-phenanthridinium bromide)

A solution of 10 mg/ml was made in distilled water and stored in a dark bottle.

A.3.9 Exo-nuclease III shortening solutions (Henikoff, 1984)

A.3.9.1 Exo buffer

Tris-HCl (1 M, pH 8.0)	660	ml
$MgCl_2$ (0.1 M)	66. 4	ml
Distilled water	9.27	ml

A.3.9.2 Klenow mixture

Tris-HCl buffer (0.1 M, pH 8.0)	3	ml
$MgCl_2$ (1 M)	6	ml
Distilled water	20	ml

A.3.9.3 Ligase mixture

Ligase buffer	144	ml
Distilled water	1440	ml

A.3.9.4 S1 buffer (10X)

KOAc (3 M)	1.1	ml
NaCl (5 M)	5	ml
Glycerol	5	ml
$ZnSO_4$	30	mg

A.3.9.5 S1 mixture

S1 buffer (10X)	41	ml
Distilled water	259	ml
S1 nuclease (60 U.)	1.5	ml

A.3.9.6 S1 Stop

Trisma base (no HCl)	0.3 M
EDTA (pH 8.0)	0.05 M

A.3.10 Isopropanol (salt saturated)

Isopropanol was saturated with aqueous 5M NaCl, 10 mM Tris-HCl and 1 mM EDTA (pH 8.0).

A.3.11 Ligase dilution buffer

The buffer was made according to the following table and stored at -20°C:

Stock solution	Final conc.	Volume
Tris-HCl (1 M, pH 7.6)	20 mM	0.2 ml
EDTA (0.5 M, pH 8.0)	1 mM	2 ml
DTT (0.5 M)	5 mM	10 ml
KCl (1 M)	60 mM	0.6 ml
Glycerol	44 % (v/v)	4.4 ml
Distilled water		4.788 ml

A.3.12 Ligation buffer (10X)

The buffer was made according to the following table and stored in aliquots at -70°C:

Stock solution	Final conc.	Amount
Tris-HCl (1 M, pH 7.6)	66 mM	0.66 ml
$MgCl_2$ (1 M)	6 mM	66 ml
ATP (0.1 M)	1 mM	0.1 ml
DTT	0.1 mM	15.4 mg
Distilled water		0.174 ml

A.3.13 Phenol (TE saturated)

Phenol (200 g, Merck) was melted at 65°C and 0.3 g of 8-hydroxyquinoline was added. The phenol was extracted three times with TE (10X) or until the pH of the aqueous phase was approximately pH 7.6. The phenol was stored under TE (1X) at -20°C.

A.3.14 Prehybridization solution

SSC buffer (6X)	100	ml
SDS	0.5	g
Denaturated salmon sperm DNA	1	ml
Denhardt's solution (50X)	10	ml
EDTA (0.5 M, pH 8.0)	2	ml

A.3.15 Restriction enzyme buffers (10X)

Stock solution	Final conc.
Tris-HCl (1 M, pH 7.9)	0.1 M
MgCl2 (1 M)	0.1 M
DTT (0.5 M)	10 mM
BSA (10 mg/ml)	1 mg/ml
Glycerol	44 % (v/v)
NaCl (5 M)	0, 50, 100 or 150 mM

The buffers were made using the following table and stored at -20°C:

Stock solution	Salt concentration (mM)			
	0	50	100	150
Tris-HCl	1 ml	1 ml	1 ml	1 ml
$MgCl_2$	1 ml	1 ml	1 ml	1 ml
DTT	0.2 ml	0.2 ml	0.2 ml	0.2 ml
BSA	1 ml	1 ml	1 ml	1 ml
Glycerol	4.4 ml	4.4 ml	4.4 ml	4.4 ml
NaCl	-	1 ml	2 ml	87.7 mg
H_20	2.4 ml	1.4 ml	0.4 ml	2.4 ml

A.3.16 Restriction enzyme dilution buffer

The buffer was made according to the following table and stored at -20°C:

Stock solution	Final Conc.	10 ml
Tris-HCl (1 M, pH 7.5)	10 mM	0.1 ml
NaCl (5 M)	50 mM	0.1 ml
Distilled water		5.3 ml
Filter sterilize this solution and then add the following constituents:		
2-Mercaptoethanol		7 ml
Gelatin (10 mg/ml)		0.1 ml
Glycerol		4.4 ml

A.3.17 SDS-Polyacrylamide gel electrophoresis (Laemmli, 1970)

A.3.17.1 Acrylamide-bis-acrylamide stock solution

Acrylamide	29.2	g
Bis-acrylamide	0.8	g
Distilled water	to 100	ml

The solution was filtered through Whatman's paper (No. 1) and stored in a dark bottle at room temperature.

A.3.17.2 Resolving gel buffer

Tris (1.5 M)	18.17	g
Distilled water	to 100	ml

Adjusted pH to 8.8 (approximately 5.5 ml conc. HCl)

A.3.17.3 Stacking gel buffer

Tris-HCl (0.5 M)	6.06	g
Distilled water	to 100	ml

Adjusted pH to 6.8 (approximately 5.5 ml conc. HCl).

A.3.17.4 Reservoir buffer (10X)

Tris base (0.25 M)	15.15	g
Glycine (0.192 M)	72.05	g
10% SDS (0.1%, w/v)	50	ml
Distilled water	to 5000	ml

The pH should be approximately 8.5.

A.3.17.5 Coomassie Blue staining solution

Coomassie Blue R250 (0.25%,w/v)	2.5	g
Destaining solution	1000	ml

The solution was stirred vigorously to dissolve the dye and then filtered through Whatman's paper (No. 1)

A.3.17.6 Destain solution

Acetic acid	250	ml
Methanol	750	ml
Distilled water	1500	ml

A.3.17.7 Sample treatment buffer

Stacking gel buffer	2.5 ml
SDS (10%)	4 ml
Glycerol	2 ml
2-Mercaptoethanol	1 ml
Distilled water	0.5 ml

A.3.17.8 Polyacrylamide gel preparation table

Stock solution	Stacking gel	Resolving gel (%)	
		10.0	7.5
Acrylamide	2.5ml	10.0 ml	7.5 ml
Resolving gel buffer	-	3.75ml	3.75ml
Stacking gel buffer	5.0ml	-	-
10% SDS	0.2ml	0.3 ml	0.3 ml
Ammonium persulfate (1.5%, w/v solution)	1.0ml	1.5 ml	1.5 ml
Water	11.3ml	14.45ml	16.95ml
TEMED	15 ml	15 ml	15 ml

A.3.17.9 Solutions for CMC-PAGE

Phosphate buffer (0.2 M)

$NaH_2PO_4.H_20$	13.8	g
Distilled water	to 500	ml

Adjust pH with NaOH to pH 6.8

CMC solution

CMC	0.6	g
Phosphate buffer (0.2M,pH6.8)	100	ml

The solution was sterilized by autoclaving

Triton X-100 (2.5%)

Triton X-100	50	ml
Distilled water	to 2000	ml

Congo red

Congo red	1	g
Distilled Water	to 1000	ml

Adjusted pH to approximately 12.0 with NaOH

NaCl (1 M)

NaCl	116.88	g
Distilled water	2000	ml

Adjusted pH to approximately 12.0 with NaOH

Acetic acid (5%)

Acetic Acid	100	ml
Distilled water	to 2000	ml

A.3.17.10 Preparation table of overlay with incorporated CMC

Stock solution	Overlay
0.2 M phosphate pH 6.8	17.4 ml
CMC (0.5%, w/v)	5.0 ml
Acrylamide	7.5 ml
TEMED	15 ml
10% Ammonium persulphate	150 ml

A.3.18 Salmon sperm DNA

A 10 mg/ml solution was made in TE buffer. The DNA solution was sonicated at full power (20 microns) for 10 min in a MSE Soniprep sonicator. The solution was aliquoted and stored at -20°C.

A.3.19 Sodium acetate (3 M, pH 5.2)

Sodium acetate$\omega 3H_2O$	4.08	g
Distilled water	to 10	ml

Adjusted pH with glacial acetic acid. Autoclaved.

A.3.20 SSC (20X)

NaCl (3 M)	175.3	g
Sodium citrate (0.3 M)	88.2	g
Distilled water	to 1000	ml

Adjusted pH to 7.0 with NaOH (10 N). Autoclaved.

A.3.21 Tris-acetate buffer (50X)

Trisbase	242	g
Acetic acid	57.1	ml
EDTA (0.5 M, pH 8.0)	100	ml
Distilled water	to 1000	ml

A.3.22 TE (Tris-EDTA) buffer (100X)

Tris-HCl (pH 7.6)	121	g
EDTA (0.5 M, pH 8.0)	200	ml
Distilled water	to 1000	ml

Autoclaved and diluted with sterile water before use.

A.3.23 TSB solution

LB	150	ml

pH to 6.1 with 2 drops conc. HCl.

PEG 4000	15	g
$MgSO_4$ (1 M)	1.5	ml
$MgCl_2$ (1 M)	1.5	ml

Dispensed in 20 ml aliquots and autoclaved. Added DMSO (1 ml) and glucose (0.5 M, 400 ml) immediately before use.

A.3.24 Z-buffer (pH 7.0)

Na_2HPO_4 (60 mM)	16.1	g
$NaH_2PO_4\omega 2H20$ (40 mM)	5.5	g
KCl (10 mM)	0.75	g
$MgSO_4\omega 7H_20$ (1 mM)	0.246	g
2-Mercaptoethanol (0.05 M)	2.7	ml
Distilled water	to 1000	ml

APPENDIX B

GENERAL TECHNIQUES

APPENDIX B

GENERAL TECHNIQUES

B.1 Plasmid DNA preparations from *E. coli*

B.1.1 Small scale plasmid DNA isolation (miniprep). Plasmid DNA was isolated from a 5 ml overnight culture (LB + Ap, 100 µg/ml) as described by Ish-Horowicz and Burke (1981). Cells were collected in a 1.5 ml Eppendorf tube by centrifugation in an Eppendorf microfuge for 1 min. The pellet was resuspended in 100 µl Solution I (50 mM glucose; 25 mM Tris-HCl, pH 8.0; 10 mM Na_2ωEDTA). The sample was left at room temperature for 5 min, and then 200 µl of Solution II (0.2 M NaOH; 1% SDS, w/v) was added and mixed gently. The sample was placed on ice for 5 min, and then 150 µl pre-cooled Solution III (5 M KOAc, pH 4.8) was added and the mixture vortexed vigorously. The tube was placed on ice for a further 5 - 10 min. The sample was microfuged for 5 min and the supernatant fluid was transferred to a fresh tube. Two volumes of 95% ethanol were added, the sample was allowed to stand at room temperature for 5 min and was then spun for 15 min in a microfuge. The pellet was dried and resuspended in 50 µl TE buffer.

B.1.2 Large scale plasmid DNA isolation (maxiprep). 200 ml overnight LB broth culture, grown at 37°C in the presence of

the appropriate antibiotic, was harvested by centrifugation at 6,000 x g for 5 min. The cell pellet was resuspended in 4 ml solution I and incubated at room temperature for at least 5 min. Freshly prepared solution II (8 ml) was added, mixed gently for clear lysis and the sample put on ice for 5 min. Precooled solution III (6 ml) was added, the mixture vortexed vigorously and put back on ice for 10 min. The cellular debris was removed by centrifugation at 12,000 x g for 10 min. The supernatant was transferred into a clean tube. An equal volume of isopropanol was added to the fluid and the sample left at room temperature for at least 30 min. The pellet containing DNA was precipitated by centrifugation at 15,000 x g for 15 min. The pellet was washed with 70% ethanol and then resuspended in 5 ml TE buffer. Cesium chloride (1 g/ml, DNA solution) and ethidium-bromide (200 ml of 10 mg/ml) were added before a clearing spin at 27,000 x g for 15 min. The refractive index of the DNA solution was adjusted to 1.396. The DNA preparation was sealed in Quick seal tubes and spun in a Beckman Vti 65.2 rotor at 55,000 rpm at 15°C for 12 h. The plasmid DNA band was visualized by long wave UV light (350 nm), and isolated in the smallest volume as possible. EtBr from the resulting plasmid band was extracted with salt-saturated isopropanol (2 - 3 times) until no pink colour remained. The DNA was precipitated from the CsCl solution by the addition of four volumes of water followed by an equal volume of isopropanol, and centrifugation in an Eppendorf microfuge for 15 min. The

DNA pellet was resuspended in 200 µl TE buffer and the concentration was spectrophotometrically scanned between 220 and 310 nm by measuring the absorbance of a appropriate dilution. The concentration was determined by using the relationship A_{260} = 1 for 50 µg/ml double-stranded DNA.

B.2 Restriction endonuclease digestion. Restriction digestion were carried out using one of the four restriction buffers (Appendix A) according to the salt requirements of the particular enzyme. Reaction volumes were usually 20 µl containing one unit of restriction enzyme for 0.5 - 1 µg DNA. Digestions were performed at 37°C (most enzymes) for 1 h. Concentrated enzyme stocks were diluted to 1 or 2 units using universal restriction enzyme dilution buffer (Appendix A). For electrophoretic analysis, the digestions were terminated by the addition of 5 µl DNA loading solution (Appendix A) to the 20 µl digestion mixture. If the sample was to be used for ligation, the digestion was terminated by a phenol-chloroform extraction without loading solution followed by ethanol precipitation. The DNA solution was extracted with the addition of phenol (1/10 volume, TE-saturated) and an equal volume of chloroform : isoamyl-alcohol (24 : 1). The mixture was vortexed briefly, and the two phases were separated by centrifugation. The aqueous phase was extracted twice with water-saturated ether. The DNA was precipitated by the addition of 0.5 M NaCl and an equal volume of isopropanol, and collected by 15 min

centrifugation. After centrifugation the DNA pellet was washed with 70% ethanol and resuspended in TE buffer.

B.3 Agarose gel electrophoresis. Agarose gel electrophoresis was performed using a horizontal submerged gel system. Tris-acetate buffer (Appendix A) was used routinely. Sigma type II agarose was used at different concentrations. The amount of DNA loaded per lane also varied with the sizes and number of fragments but under normal circumstances about 200 ng of plasmid DNA was used. The gels were electrophoresed at 2 V/cm for 16 h. Gels were stained in electrophoresis buffer containing EtBr (0.5 mg/ml) for 15 - 30 min. DNA bands were visualized using a 254 nm transilluminator. A 310 nm transilluminator was used if the DNA was to be recovered from the gel.

Gels were photographed using a Polaroid CU-5 Land camera fitted with a red filter and a fixed focal length attachment. Polaroid type 667 film (ASA 3,000) was used with an exposure time of 1 - 2 sec. at f4.7. If a negative was required then a Polaroid type 665 film (ASA 64) with an exposure of 120 - 140 sec at f4.7 was used.

Fragment sizes were calculated by extrapolation from a standard curve of the mobility of lambda DNA fragments, digested with either *Hind* III or *Pst* I, plotted against the

$\log_{10}$ of their molecular weights. Approximately 0.2 - 1.0 µg of a lambda DNA digested was loaded per gel lane.

B.4 Ligation reaction. The methods of Maniatis et al. (1982) were generally used. Different vector to insert ratios were used depending on the sizes of the fragments to be cloned. Recircularization reactions for isolating deletion plasmids contained DNA concentrations of 1 pmole/ml or less. A DNA concentration in the order of 5 pmole/ml was used for recombination reactions. Ligation reactions containing DNA, ligation buffer (Appendix A) and water to the required volume were performed in sterile microfuge tubes. Sticky-end ligations were performed at room temperature for 3 h or at 15°C overnight using 0.1 - 0.25 U. of ligase , whereas blunt-end ligations were performed at room temperature for 2 - 20 h using 20 - 100X more ligase.

B.5 Subcloning protocol. The rapid subcloning protocol of Struhl (1985) was used. The DNA fragments were separated by electrophoresis through low melting point (LMP) agarose (1%) (SeaplaqueR) in Tris-acetate buffer (50 mM, pH 8.2). The gel was stained with EtBr after electrophoresis and the DNA fragments were viewed under UV light (310 nm) as briefly as possible. The desired fragments were excised using sterile scalpel blades in as small a volume as possible. The gel slices were melted at 70°C for 5 min in a microfuge tube and the required amounts (generally using 2 µl vector DNA, 8 µl

insert DNA) were added hot to the prepared ligation mixture containing ligation buffer, ligase and water (10 µl). The ligation was incubated at room temperature for 3 h. Before transformation of *E. coli* competent cells, the gel ligation reactions were melted at 70°C for 5 min, and then diluted with 4 volumes TSB solution (Appendix A).

B.6 The preparation and transformation of competent *E. coli* cells. *E. coli* cells were made competent for DNA uptake according to the method of Chung and Miller (1988). A 1/100 dilution of an overnight *E. coli* culture in LB was inoculated into 25 ml prewarmed LB and incubated at 37°C, with shaking, until the culture had reached early exponential phase (OD_{600} = 0.3 - 0.6). The cell culture was poured into a pre-cooled sterile SS-34 tube and the cells were harvested at 5,000 x g for 5 min at 4°C. The cell pellet was resuspended in 2.5 ml (1/10 volume) ice-cold transformation and storage buffer (TSB) and held on ice for 10 min. The *E. coli* cells (100 µl) were then mixed with DNA and held on ice for a further 30 min. TSB solution (0.9 ml) containing glucose (20 mM) was added to each transformation mixture and incubated at 37°C for 60 min, to allow expression of the plasmid-borne antibiotic marker.

B.7 Radioactive labelling of DNA probes. DNA probes were labelled with [α - ^{32}p] dCTP to high specific activity by nick-translation (Rigby et al., 1977). The reagents were

obtained in kit form Amersham and used according to the supplier's specifications. Contaminating nucleotides were removed from the radioactively labelled probe preparation using a Sephadex G-50 spin column as described by Maniatis et al., (1982). Radioactively labelled probes were stored in lead containers at -20°C. Probes were denatured by boiling (10 min) in a fume hood just before use.

B.8 DNA hybridization. DNA fragments resolved by agarose gel electrophoresis were transferred to a Hybond-N^+ hybridization membrane (Amersham) essentially by the protocol of Reed and Mann (1985). The use of a nylon transfer membrane allows the capillary transfer of DNA restriction fragments in alkali rather than in neutral, high ionic strength solvents (used in conventional Southern transfer), and eliminates the need for post-transfer fixation (Reed and Mann, 1985). After electrophoresis the gel was rinsed in 2 volumes of HCl (0.25 M) for 20 min at room temperature with gentle agitation, followed by a brief rinse in distilled water. The gel was then placed on top of 2 sheets of Whatman 3 MM filter paper (wetted with 0.4 N NaoH, and placed on top of an inverted gel-casting tray in a plastic box, such that the filter paper touched the base of the box, forming a wick), and was flooded with 50 - 100 ml of 0.4 N NaoH, and then immersed in distilled water, hybond-N^+ membrane was placed on top of the gel, and any air bubbles were removed. Three sheets of Whatman 3 MM filter

paper, wetted in 0.4 N NaoH, were laid onto the membrane, followed by a 10 cm thick layer of absorbent paper. A light weight was placed on top of this, and transfer left to continue overnight. After transfer, the membrane was removed and rinsed for 20 min in 2 X SSC (Appendix A). Hybridization and washing conditions were essentially according to Maniatis et al., (1982). The membrane was gently shaken in pre-hybridization solution (Appendix A) for 2 h at 68°C, while the probe was being prepared. The radioactively-labelled probe to be used was denatured by boiling for 10 min and was added to the pre-hybridization fluid. Hybridization was carried out with constant agitation at 68°C for 10 - 16 h. The filter was washed stringently at 68°C for 30 min, and the washing was terminated after checking the radioactivity by means of a Geiger-counter. The membrane was exposed to autoradiographic film (XAR-5) at -70°C.

B.9 Nucleotide sequencing.

B.9.1 Primer annealing reaction. Supercoiled DNA (6 - 10 µg, in TE buffer) was diluted to a final volume of 20 µl in distilled water. Alkaline denaturation in 0.2 N NaoH (5 min at room temperature) was followed by the addition of 5 µl of 3 M sodium acetate (pH 5.2), 25 µl of distilled water and 150 µl of ethanol. This mixture was chilled to -70°C, centrifuged at 4°C for 20 min in a microfuge and washed with

200 µl of ethanol (70%). The DNA pellet was dried and resuspended in a final volume of 10 ml sequencing buffer (40 mM Tris-HCl, pH 7.5; 20 mM $MgCl_2$; 50 mM NaCl) and 12 ng of primer. This mixture was annealed for 30 min at 40°C immediately prior to sequencing. The primers used were the forward sequencing primer as supplied in the sequenase DNA sequencing kit (US Biochemical Corp., Cleveland, Ohio) and the M13 reverse sequencing primer (Amersham).

B.9.2 Sequencing reactions. DNA sequencing was done by the dideoxy chain termination method of Sanger et al., (1977) according to the protocol of Tabor and Richardson (1987), using T7 DNA polymerase and a "sequenase" sequencing kit supplied by the US Biochemical Corporation, Cleveland, Ohio. The DNA chain was radiolabelled with [α - ^{35}S] dATP (1200 Ci/mmol; Amersham).

B.9.3 Gel electrophoresis and autoradiography. The sequencing reactions were analysed on standard 6% denaturing acrylamide urea sequencing gels. The composition and running conditions of the gels were as described in the Amersham M13 sequencing handbook. After electrophoresis the gels (0.3 mm thick) were dried onto Whatman No. 3 filter paper using a Dual Temperature Slab Gel Dryer (Model 1125B; Hoefer Scientific Instruments, San Francisco). Gels containing ^{35}S-labelled DNA were placed under XAR-5 autoradiographic film and exposed for 1-2 days. The

autoradiographs were developed using Kodak GBX X-ray developer and fixer.

B.10 SDS-polyacrylamide gel electrophoresis of proteins. SDS-polyacrylamide gels were prepared according to the method of Laemmli (1970) using a Hoefer gel apparatus (SE600) with 0.75 mm spacers assembled according to the manufacturer's specifications. All buffers and a preparation table for resolving and stacking gels are given in Appendix A. Propan-2-ol was layered on the gel to promote a sharp interface. After the gel had polymerized, the propan-2-ol was removed by rinsing with the stacking gel buffer, and the stacking gel solution was added.

Samples were prepared in sample treatment buffer (Appendix A) and placed in a boiling waterbath for 2 min before being loaded onto the gel. Electrophoresis was continued at 15 mA (constant current) per gel until the front dye had migrated to the end of the gel (4 to 5 h). After electrophoresis the gels were stained for 3 h in staining solution with gentle agitation, destained and dried.

B.11 Non-denaturing (native) polyacrylamide gel electrophoresis. Proteins were separated under non-denaturing conditions using conditions identical to SDS-PAGE with the exception that SDS was omitted from all buffers. Samples were combined with an equal volume of sucrose-dye

solution (50% [w/v] sucrose, 0.1% [w/v] bromophenol blue) without SDS and loaded directly on the gel. Samples were not boiled in this experiment.

B.12 Determination of protein concentrations. Protein concentrations in solutions were determined by the method of Bradford (1976). Assays were performed in triplicate using disposable cuvettes. The reaction contained protein solution (0.1 ml) and 2 ml Bradford solution (Appendix A). The mixture was left for 5 min at room temperature and the optical density of the reaction monitored at OD_{595}. Protein concentrations were calculated using a standard curve (BSA Fraction, 10 - 200 mg/ml). Protein samples were diluted to ensure that OD_{595} did not exceed 0.8.

APPENDIX C

BACTERIAL STRAINS, GENOTYPES AND REFERENCES

R. flavefaciens FD-1, which was a gift from Dr. M. P. Bryant, University of Illinois, Urbana-Champaign, was isolated from rumen fluid.

Strain	Genotype	Reference
E. coli JM103	Δ(*lac pro*) *thi*I *str*A *sup*E *end*A *sbc*B *hsd*R$^-$ F'*tra* 36 *pro*AB *lac*I Z M15	Messing et al. (1981)
E. coli HB101	*leu*B6 *trp*E38 *met*E70 *rec*A13 *sup*E44	Boyer and Roulland-Dussoix (1969)
E. coli MM294	*sup*E44 *hsd*R *end*A1 *pro thi*	Meselson & Yuan (1968)
E. coli K514λ	$r_K^- m_K^+$ *thr*-1 *leu*B6 *thi*-1 *supp*E44 *lac*Y1 *ton*A21	Wood (1966)
E. coli LK111	*lac*I^q *lac*Z ΔM15 *lac*Y$^+$ derivative of *E. coli* K514λ	Zabeau & Stanley (1982)
E. coli CC118	*ara*Δ139 Δ(*ara leu*)7697 Δ*lac*X174 *pho*AΔ20 *gal*E *gal*K *thi rps*E *rpo*B *arg*Eam *rec*A1	Manoil and Beckwith (1985)
E. coli CSH23	Δ(*lac pro*) *sup*E *spc thi*(F'*lac*$^+$ *pro*A$^+$, B$^+$)	Miller (1972)

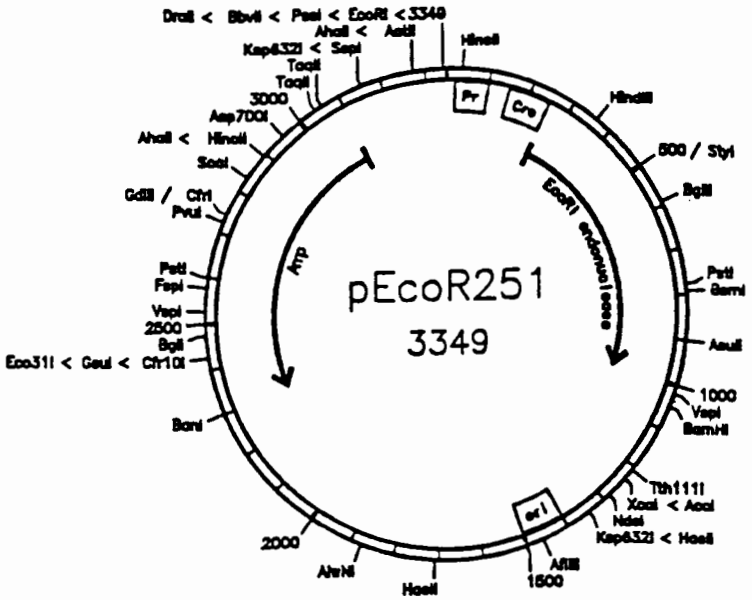
APPENDIX D

ONE- AND THREE-LETTER AND CODES USED FOR AMINO ACIDS.

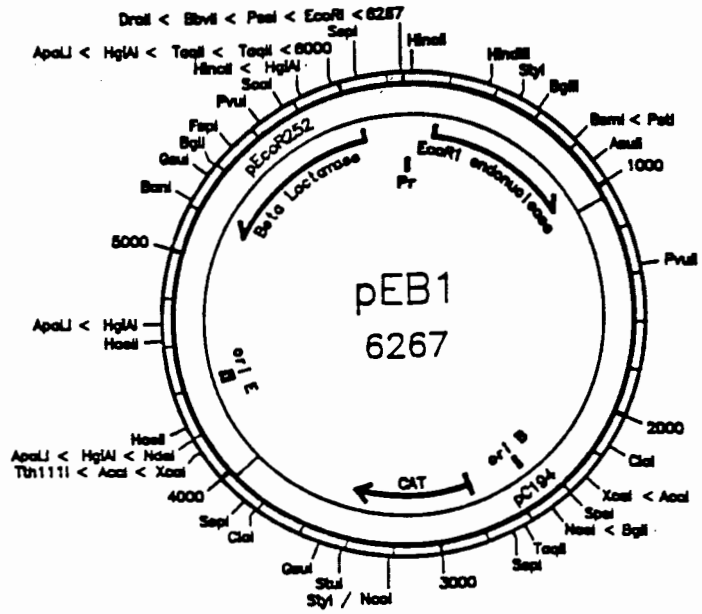
Amino acid	Code		Code	Amino acid
Alanine	Ala	A	A	Alanine
Arginine	Arg	R	C	Cysteine
Asparagine	Asn	N	D	Aspartic acid
Aspartic acid	Asp	D	E	Glutamic acid
Cysteine	Cys	C	F	Phenylalanine
Glutamine	Gln	Q	G	Glycine
Glutamic acid	Glu	E	H	Histidine
Glycine	Gly	G	I	Isoleucine
Histidine	His	H	K	Lysine
Isoleucine	Ile	I	L	Leucine
Leucine	Leu	L	M	Methionine
Lysine	Lys	K	N	Asparagine
Methionine	Met	M	P	Proline
Phenylalanine	Phe	F	Q	Glutamine
Proline	Pro	P	R	Arginine
Serine	Ser	S	S	Serine
Threonine	Thr	T	T	Threonine
Tryptophan	Trp	W	V	Valine
Tyrosine	Tyr	Y	W	Tryptophan
Valine	Val	V	Y	Tyrosine

APPENDIX E

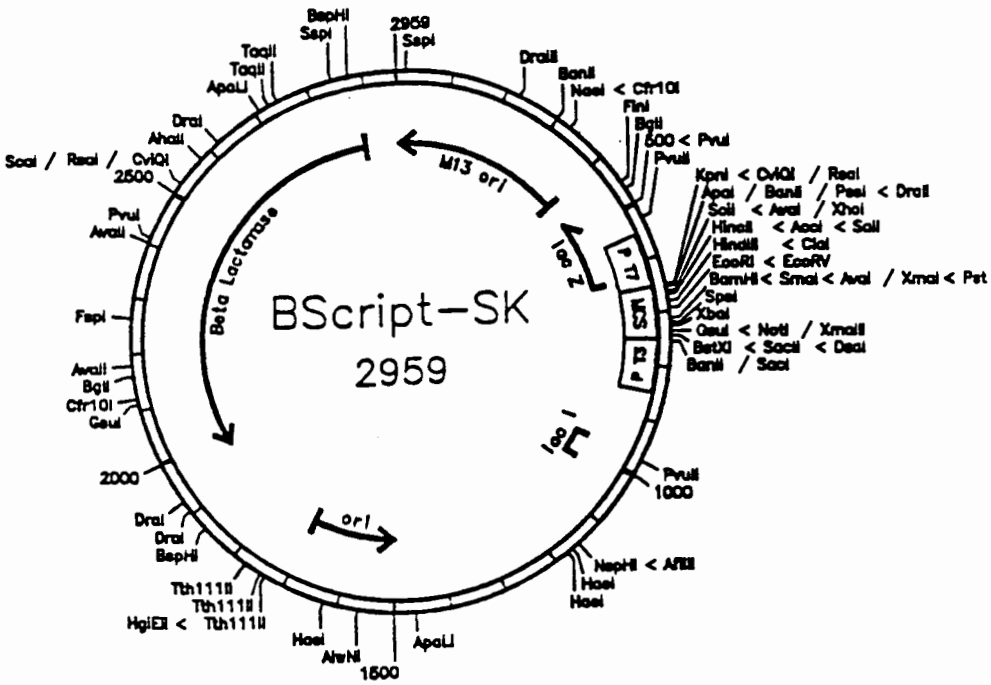
PLASMID VECTOR AND PHAGE MAPS



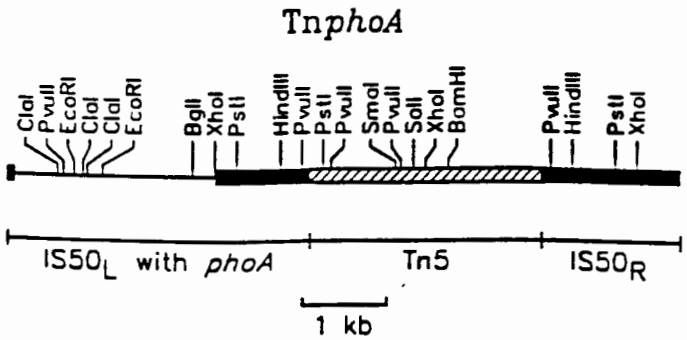
E.1 Restriction map of **pEcoR251** (Zabeau and Stanley, 1982).



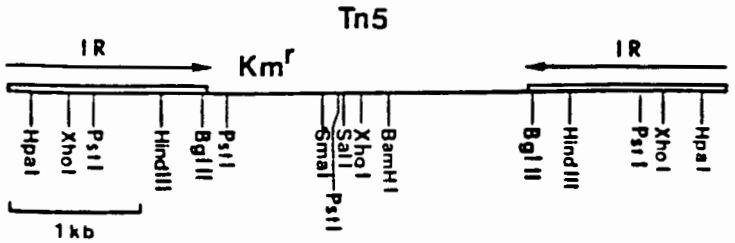
E.2 Restriction map of **pEB1** (H. Zappe, pers. comm.).



E.3 Restriction map of **Bluescript SK$^+$** (Stratagene, San Diego, California).



E.4 Partial restriction map of **TnphoA** (Manoil and Beckwith, 1985).



E.5 Partial restriction map of **Tn5** (de Bruijn and Lupski, 1984).

REFERENCES

Aiba H., S. Adhya, and B. de Crombrugghe. 1981. Evidence for two functional gal promoters in intact *Escherichia coli* cells. J. Biological Chemistry **256**:11905-11910.

Akin, D. E. 1980. Evaluation by electron microscopy and anaerobic culture of types of rumen bacteria associated with digestion of forage cell walls. Appl. Environ. Microbiol. **39**:242-252.

Albersheim, P. 1965a. p. 151-186. J. Bonner, and J. E. Varner (editors), Plant Biochemistry: The substructure and function of the cell wall. Academic Press, New York and London.

Albersheim, P. 1965b. p. 298-321. J. Bonner, and J. E. Varner (editors), Plant Biochemistry: Biogenesis of the cell wall. Academic Press, New York and London.

Amer, G. I., and S. W. Drew. 1980. Microbiology of lignin degradation. Annu. Rep. Fermentat. Processes **4**:67-103.

Ayers, W. A. 1958. Phosphorylation of cellobiose and glucose by *Ruminococcus flavefaciens*. J. Bacteriol. **76**:515-517.

Ayers, W. A. 1959. Phosphorolysis and Synthesis of cellobiose by cell extracts from *Ruminococcus flavefaciens*. J. Biological Chemistry **234**:2819-2822.

Bagnara-Tardif, C., C. Gaudin, A. Belaich, P. Hoest, T. Citard and J.-P. Belaich. 1992. Sequence analysis of a gene cluster encoding cellulases from *Clostridium cellulolyticum*. Gene **119**:17-28.

Barras, F., J. P. Chambost, and M. Chippaux. 1984. Cellobiose metabolism in *Erwinia*: Genectics study. Mol. Gen. Genet. **197**:486-490.

Barros, M. E. C., and J. A. Thomson. 1987. Cloning and expression in *Escherichia coli* of a cellulase gene from *Ruminococcus flavefaciens*. J. of Bacteriol. **169**:1760-1762.

Bassford, P. J. JR., T. J. Silhavy, and J. R. Beckwith. 1979. Use of the gene fusion to study secretion of maltose-binding protein into *Escherichia coli* periplasm. J. Bacteriol. **139**:19-31.

Bauchop, T. 1979. Rumen anaerobic fungi of cattle and sheep. Appl. Environ. Microbiol. **38**:148-158.

Beguin, P. 1990. Molecular biology of cellulose degradation. Annu. Rev. Microbiol. **44**:219-248.

Berg, B. and A. Von Hofsten. 1976. The ultrastructure of the fungus *Trichoderma viride* and investigation of its growth on cellulose. J. Appl. Bacteriol. **41**:395-399.

Berg, B., and G. Pettersson. 1977. Location and formation of cellulases in *Trichoderma viride*. J. Appl. Bacteriol. **42**:65-75.

Berg, B., B. V. Hofsten, and G. Pettersson. 1972. Electron microscopic observation on the degradation of cellulose fibres by *Cellvibrio fulvus* and *Sporocytophaga myxococcoides*. J. Appl. Bacteriol. **35**:215-219.

Berger, E., W. A. Jones, D. T. Jones, and D. R. Woods. 1990. Sequencing and expression of a cellodextrinase (*ced*1) gene from *Butyrivibrio fibrisolvens* H17c cloned in *Escherichia coli*. Mol. Gen. Genet. **223**:310-318.

Berger, E., W.A. Jones, D.T. Jones, and D.R. Woods. 1989. Cloning and sequencing of an endoglucanase (*end*1) gene from *Butyrivibrio fibrisolvens* H17c. Mol. Gen. Genet. **219**:193-198.

Biely, P., and E. Petrakova. 1984. Novel inducers of the xylan-degrading enzyme system of *Cryptococcus albidus*. J. Bacteriol. **160**:408-412.

Boyer, H. W., and D. Roulland-Dussoix. 1969. A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. J. Mol. Biol. **41**:459-472.

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.

Bronnenmeier, K., K. P. Rucknagel, and W. L. Staudenbauer. 1991. Purification and properties of a novel type of exo-1,4-β-glucanase (Avicelase II) from the cellulolytic thermophile *Clostridium stercorarium*. Eur. J. Biochem. **200**:379-385.

Bryant, M. P. 1973. Nutritional requirements of the predominant rumen cellulolytic bacteria. Fed. Proc. **32**:1809-1813.

Bryant, M. P. 1986. Genus *Ruminococcus* Sijpesteijn 1948, 152^{AL}, p. 1093-1097. In P. H. A. Sneath, N. S. Mair, M. E. Sharpe, and J. G. Holt (ed.), Bergy's manual of systematic bacteriology, vol. 2. The Williams & Wilkins Co., Baltimore.

Caldwell, R. C., and M. P. Bryant. 1966. Medium without rumen fluid for nonselective enumeration and isolation of rumen bacteria. Appl. Microbiol. **14**:794-801.

Cavedon, K., S. B. Leschine, and E. Canale-Parola. 1990. Characterization of the extracellular cellulase from a mesophilic *Clostridium* (strain C7). J. Bacteriol. **172**:4231-4237.

Chauvaux, S., P. Beguin, J. P. Aubert, K. M. Bhat, L. A. Gow, T. M. Wood, and A. Bairoch. 1990. Calcium-binding affinity and calcium-enhanced activity of *Clostridium thermocellum* endoglucanase D. Biochem. J. **265**:261-265.

Cheng, S.-C., and P. Modrich. 1983. Positive selection cloning vehicle useful for overproduction of hybrid proteins. J. Bacteriol. **154**:1005-1008.

Chesson, A., and C. W. Forsberg. 1988. Polysaccharide degradation by rumen microorganisms. p. 251-284. *In*: P. N. Hobson (ed.) The rumen microbial ecosystem, Elsevier Applied Science, London, New York.

Chung, C. T., and R. H. Miller. 1988. A rapid and convenient method for the preparation and storage of competent bacterial cells. Nucleic Acids Res. **16**:3580.

Cocconcelli, P. S., E. Ferrari, F. Rossi, and V. Bottazzi. 1992. Plasmid transformation of *Ruminococcus albus* by means of high-voltage electroporation. FEMS Micrbiol. lett **94**:203-208.

Coughlan, M. P. 1985. **3**:39-109. G. E. Russell (editor), Biotechnology and Genetic Engineering Reviews: The properties of fungal and bacterial cellulases with comment on their production and application. Intercept Ltd., Ponteland, Newcastle upon Tyne.

Cunningham, C., C. A. McPherson, J. Martin, W. J. Harris, and H. J. Flint. 1991. Sequence of a cellulase gene from the rumen anaerobe *Ruminococcus flavefaciens* 17. Mol. Gen. Genet. **228**:320-323.

Fierobe H.-P., C. Gaudin, A. Belaich, M. Loutfi, E. Faure, C. Bagnara, D. Baty, and J.-P. Belaich. 1991. Characterization of endoglucanase A from *Clostridium cellulolyticum*. J. Bacteriol. **173**:7956-7962.

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.

Flint, H. J., C. A. McPherson, and J. Martin. 1991. Expression of two xylanase genes from the rumen cellulolytic bacterium *Ruminococcus flavefaciens* 17 cloned in pUC13. J. Gen. Microbiol. **137**:123-129.

Fukumori, F., T. Kudo, Y. Narahashi, and K. Horikoshi. 1986. Molecular cloning and nucleotide sequence of the alkaline cellulase gene from the alkalophilic *Bacillus* sp. strain 1139. J. Gen. Microbiol. **132**:2329-2335.

Fusee, M. C. and J. M. Leatherwood. 1972. Regulation of cellulase from *Ruminococcus*. Can. J. Microbiol. **18**:347-353.

Gardner, R. M., K. C. Doerner, and B. A. White. 1987. Purification and characterization of an exo-β-1,4-glucanase from *Ruminococcus flavefaciens* FD-1. J. Bacteriol. **169**:4581-4588..

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.

Gilbert, H.J., J. Hall, G.P. Hazlewood, and L.M.A. Ferreira. 1990. The N-terminal region of an endoglucanase from *Pseudomonas fluorescens* subspecies *cellulosa* constitutes a cellulose-binding domain that is distinct from the catalytic centre. Mol. Microbiol. **4**:759-767.

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. Microbiological Reviews **55**:303-315.

Gilkes, N. R., R. A. J. Warren, R. C. Miller, Jr., and D. G. Kilburn. 1988. Precise excision of the cellulose binding domains from two *Cellulomonas fimi* cellulases by a homologous protease and the effect on catalysis. J. Biol. Chem. **263**:10401-10407.

Goldstein, J., S. Lehnhardt, and M. Inouye. 1990. Enhancement of protein translocation across the membrane by specific mutation in the hydrophobic region of the signal peptide. J. Bacteriol. **172**:1225-1231.

Gong, C.-S., L. F. Chen, M. C. Flickinger, and G. T. Tsao. 1981. Conversion of hemicellulose carbohydrates. Adv. Biochem. Eng. **20**:93-118.

Gong, J., and C. W. Forsberg. 1989. Factors affecting adhesion of *Fibrobacter succinogenes* subsp. *succinogenes* S85 and adherence-defective mutants to cellulose. Appl. Environ. Microbiol. **55**:3039-3044.

Gong, J., R. Y. C. Lo, and C. W. Forsberg. 1989. Molecular cloning and expression in *Escherichia coli* of a cellodextrinase gene from *Bacteroides succinogenes* S85. Appl. Environ. Microbiol. **55**:132-136.

Gosalbes, M. J., J. A. Perez-Gonzalez, R. Gonzalez, 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.

Grantham, R., C. Gantier, M. Gouy, M. Jacobzone, and R. Mercier. 1981. Codon catalog usage is a genome strategy modulated for gene expressivity. Nucleic Acids Res. **9**:r43-r74.

Greenberg, N. M., R. A. J. Warren, D. G. Kilburn, and R. C. Miller, Jr. 1987a. Regulation, initiation, and termination of the *cen*A and *cex* transcripts of *Cellulomonas fimi*. J. Bacteriol. **169**:646-653.

Greenberg, N. M., R. A. J. Warren, D. G. Kilburn, and R. C. Miller, JR. 1987b. Regulation and initiation of *cen*B transcripts of *Cellulomonas fimi*. J. Bacteriol. **169**:4674-4677.

Groleau, D., and C. W. Forsberg. 1983. Partial characterization of the extracellular carboxymethylcellulase activity produced by the rumen bacterium *Bacteroides succinogenes*. Can. J. Microbiol. **29**:504-517.

Gutierrez, C., J. Barondess, C. Manoil, and J. Beckwith. 1987. The use of transposon TnphoA to detect genes for cell envelope proteins subject to a common regulatory stimulus. J. of Mol. Biol. **195**:289-297

Hall, J., G. P. Hazlewood, N. S. Huskisson, A. J. Durrant, and H. J. Gilbert. 1990. Conserved serine-rich sequences in xylanase and cellulase from *Pseudomonas fluorescens* subspecies *cellulosa*: internal signal sequence and unusual protein processing. Mol. Microbiol. **3**:1211-1219.

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**:29-38.

Halliwell, G., and M. P. Bryant. 1963. The cellulolytic activity of pure strains of bacteria from the rumen of cattle. J. Gen. Microbiol. **32**:441-448.

Hamamoto, T., F. Foong, O. Shoseyov, and R. H. Doi. 1992. Analysis of functional domains of endoglucanases from *Clostridium cellulovorans* by gene cloning, nucleotide sequencing and chimeric protein construction. Mol. Gen. Genet. **231**:472-479.

Hazlewood, G. P., and R. M. Teather. 1988. The genetics of rumen bacteria, p 323-342. *In*: P. N. Hobson (ed.) The rumen microbial ecosystem, Elsevier Applied Science, London, New York.

Heale, J. B., and D. P. Gupta. 1971. The utilization of cellobiose by *Verticillum albo-atrum*. J. Gen. Microbiol. **63**:175-181.

Helaszek, C. T., and B. A. White. 1991. Cellobiose uptake and metabolism by *Ruminococcus flavefaciens*. Appl. Environ. Microbiol. **57**:64-68.

Henikoff S. 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene **28**:351-359.

Henrissat, B., H. Driguez, C. Viet, and M. Schulein. 1985. Synergism of cellulases from *Trichoderma reesei* in the degradation of cellulose. Biotechnol. **3**:722-726.

Henrissat, B., M. Claeyssens, P. Tomme, L. Lemesle, and J. P. Mornon. 1989. Cellulase families revealed by hydrophobic cluster analysis. Gene **81**:83-95.

Hiltner, P., and B. A. Dehority. 1983. Effect of soluble carbohyrates on digestion of cellulose by pure cultures of rumen bacteria. Appl. Environ. Microbiol. **46**:642-648.

Howard, G. T. and B. A. White. 1988. Molecular cloning and expression of cellulase genes from *Ruminococcus albus* 8 in *Escherichia coli* Bacteriophage λ. Appl. Enviro. Microbiol. **54**:1752-1755

Howard, G. T., and B. A. White. 1990. Cloning in *Escherichia coli* of a bi-functional cellulase/xylanase enzyme from *Ruminococcus flavefaciens* FD-1. Animal Biotechnology **1**:95-106.

Hrmova, M., E. Petrakova, and P. Biely. 1991. Induction of cellulose- and xylan-degrading enzyme systems in *Aspergillus terreus* by homo- and heterodisaccharides composed of glucose and xylose. J. Gen. Microbiol. **137**:541-547.

Hrmova, M., P. Biely, and M. Vrsanska. 1986. Specificity of cellulase and β-xylanase induction in *Trichoderma reesei* QM9414. Arch. Microbiol. **144**:307-311.

Hrmova, M., P. Biely, and M. Vrsanska. 1989. Cellulose-xylan- degrading enzymes of *Aspergillus terreus* and *Aspergillus niger.* Enzyme and Microbial Technology **11**:610-616.

Huang L. and C. W. Forsberg. 1990. Cellulose digestion and cellulase regulation and distribution in *Fibrobacter succinogenes* subsp. *succinogenes* S85. Appl. Environ. Microbiol. **56**:1221-1228.

Huang, C-M., W. J. Kelly, R. V. Asmundson, and P-L. Yu. 1989. Molecular cloning and expression of multiple cellulase genes of *Ruminococcus flavefaciens* strain 186 in *Escherichia coli.* Appl. Microbiol. Biotechnol. **31**:265-271.

Huang, J., and M. A. Schell. 1992. Role of the two-component leader sequence and mature amino acid sequences in extracellular export of endoglucanase EGL from *Pseudomonas solanacearum.* J. Bacteriol. **174**:1314-1323.

Huang, L., and C. W. Forsberg. 1987. Isolation of a cellodextrinase from *Bacteroides succinogenes.* Appl. Environ. Microbiol. **53**:1034-1041.

Huang, L., and C. W. Forsberg. 1990. Cellulose digestion and cellulase regulation and distribution in *Fibrobacter succinogenes* subsp. *succinogenes* S85. Appl. Environ. Microbiol. **56**:1221-1228.

Hungate, R. E. 1988. The ruminant and the rumen. p. 1-19. *In*: P. N. Hobson (ed.) The rumen microbial ecosystem, Elsevier Applied Science, London, New York.

Inglin, M., B. A. Feinberg, and J. R. Loewenberg. 1980. Partial purification and characterization of a new intracellular β-glucosidase of *Trichoderma reesei*. Biochemical J. **185**:515-519.

Ish-Horowicz, D. and J. F. Burke. 1981. Rapid and efficient cosmid cloning. Nucleic Acids Res. **9**:2989-2998.

Johnson, E. A., F. Bouchot, and A. L. Demain. 1985. Regulation of cellulase formation in *Clostridium thermocellum*. J. Gen. Microbiol. **131**:2303-2308.

Juy, M., A. G. Amit, P. M. Alzari, R. J. Poljak, M. Claeyssens, P. Beguin, and J.-P. Aubert. 1992. Three-dimensional structure of a thermostable bacterial cellulase. Nature **357**:89-91.

Kirk, T. K., M. Tien, P. J. Kersten, and M. D. Mozuch. 1986. Ligninase of *Phanerochaete chrysosporium*; mechanism of its degradation of the non phenolic aryglycerol β-aryl ether substructure of lignin. Biochem. J. **236**:279-287.

Klump, H. H., and D. L. Maeder. 1991. The thermodynamic basis of the genetic code. Pure Appl. Chem. **63**:1357-1366.

Knowles, J., P. Lehtovaara, and T. Teeri. 1987. Cellulase families and their genes. Trends in Biotechnol. **5**:255-261.

Kubicek, C. P. 1987. Involvement of a conidial endoglucanase and a plasma-membrane-bound β-glucosidase in the induction of endoglucanase synthesis by cellulose in *Trichoderma reesei*. J. Gen. Microbiol. **133**:1481-1487.

Kudo, H., K.-J. Cheng, and J. W. Costerton. 1987. Electron microscopic study of the methylcellulose-mediated detachment of cellulolytic rumen bacteria from cellulose fibres. Can. J. Microbiol. **33**:267-272.

Kundu, R. K., S. Dube, and D. K. Dube. 1988. Extracellular cellulolytic enzyme system of *Aspergillus japonicus*: 3. Isolation, purification, and characterization of multiple forms of endoglucanase. Enzyme Microb. Technol. **10**:100-109.

Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227**:680-685.

Lamed, R., and E. A. Bayer. 1988. The cellulosome of *Clostridium thermocellum*. Adv. Appl. Microbiol. **33**:1-46.

Latham, M. J., B. E. Brooker, G. L. Pettipher, and P. J. Harris. 1978. *Ruminococcus flavefaciens* cell coat and Adhesion to cotton cellulose and to cell walls in leaves of perennial ryegrass (*Lolium perenne*). Appl. Environ. Microbiol. **35**:156-165.

Leatherwood, J. M. 1973. Cellulose degradation by *Ruminococcus*. Federation Proceedings **32**:1814-1818.

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.

Lin L-L., and J.A. Thomson. 1991. Cloning, sequencing and expression of a gene encoding a 73 kDa xylanase enzyme from the rumen anaerobe *Butyrivibrio fibrisolvens* H17c. Mol. Gen. Genet. **228**:55-61.

Lin, E., and D. B. Wilson. 1987. Regulation of β-1,4-Endoglucanase synthesis in *Thermomonospora fusca*. Appl. Environ. Microbiol. **53**:1352-1357.

Lin, E., and D. B. Wilson. 1988. Transcription of the *cel*E gene in *Thermomonospora fusca*. J. Bacteriol. **170**:3838-3842.

Lin, L-L., E. Rumbak, H. Zappe, J. A. Thomson, and D. R. Woods. 1990. Cloning, Sequencing and analysis of expression of a *Butyrivibrio fibrisolvens* gene encoding a β-glucosidase. J. Gen. Microbiol. **136**:1567-1576.

Ljungdahl, L. G., and K.-E. Eriksson. 1985. **8**:237-299. K. C. Marshall(editors), Adv. Microb. Ecol.: Ecology of microbial cellulose degradation. PLenum Press. New York and London.

Luthi, E., D. R. Love, J. McAnulty, C. Wallace, P. A. Caughey, D. Saul, and P. L. Bergquist. 1990. Cloning, sequence analysis, and expression of genes encoding xylan-degrading enzymes from the thermophile *Caldocellum saccharolyticum*. Appl. Environ. Microbiol. **56**:1017-1024.

Luthi, E., N. B. Jasmat, R. A. Grayling, D. R. Love, and P. L. Bergquist. 1991. Cloning, sequence analysis, and expression in *Escherichia coli* of a gene coding for a β-mannanase from the extremely thermophilic bacterium *Caldocellum saccharolyticum*. Appl. Environ. Microbiol. **57**:694-700.

Mackenzie, C. R., D. Bilous, H. Schneider, and K. G. Johnson. 1987. Induction of cellulolytic and xylanolytic enzyme systems in *Streptomyces* spp. Appl. Environ. Microbiol. **53**:2835-2839.

Mandels, M., and E. T. Reese. 1960. Induction of cellulase in fungi by cellobiose. J. Bacteriol. **79**:816-826.

Mandels, M., F. W. Parrish, and E. T. Reese. 1962. Sophorose as an inducer of cellulase in *Trichoderma viride*. J. Bacteriol. **83**:400-408.

Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: A Laboratory Manual. New York: Cold Spring Harbour Laboratory.

Manoil C., and J. Beckwith. 1985. A transposon probe for protein export signals. Proc. Natl. Acad. Sci. USA **82**:8129-8133.

Matte, A. and C. W. Forsberg. 1992. Purification, characterization, and mode of action of endoxylanase 1 and 2 from *Fibrobacter succinogenes* S85. Appl. Environ. Microbiol. **58**:157-168.

Mayer, F., M. P. Coughlan, Y. Mori, and L. G. Ljungdahl. 1987. Macromolecular organization of the cellulolytic enzyme complex of *Clostridium thermocellum* as revealed by electron microscopy. Appl. Environ. Microbiol. **53**:2785-2792.

McGavin, M. J. Lam, and C. W. Forsberg. 1990. Regulation and distribution of *Fibrobacter succinogenes* subsp. *succinogenes* S85 endoglucanases. Appl. Environ. MNicrobiol. **56**:1235-1244.

McGavin, M. J., C. W. Forsberg, B. Bell, A. W. Crosby, D. Dignard, and D. Y. Thomas. 1989. Structure of the *cel*-3 gene from *Fibrobacter succinogenes* S85 and characteristics of the encoded gene product, endoglucanase 3. J. Bacteriol. **171**:5587-5595.

Meinke, A., N. R. Gilkes, D. G. Kilburn, R. C. Miller, JR., and R. A. J. Warren. 1991. Multiple domains in endoglucanase B (CenB) from *Cellulomonas fimi*: Functions and Relatedness to domains in other polypeptides. J. Bacteriol. **173**:7126-7135.

Meselson, M. and R. Yuan. 1968. DNA restriction enzyme from *E. coli*. Nature **217**:1110.

Messing, J., R. Crea, and P. H. Seeburg. 1981. A system for shotgun DNA sequencing. Nucleic Acids Res. **9**:309-321.

Miller, G. L. 1959. Use of dinitrosalicylic acid reagent for determination of reducing sugar. Anal. Chem. **31**:426-428.

Miller, J. H. 1972. Experiments in molecular genetics. Cold Spring Harbor laboratory, Cold Spring Harbor, NY.

Miller, P. B., H. Shen, N. R. Gilkes, D. G. Kilburn, R. C. Miller, Jr., A. G. Plaut, and R. A. J. Warren. 1992. Endoglucanase A from *Cellulomonas fimi* in which the hinge sequence of human IgA1 is substituted for the linker connecting its two domains is hydrolyzed by IgA proteases from *Neisseria gonorrhoeae*. FEMS Microbiol. Lett. **92**:199-204.

Morag, E., I. Halevy, E. A. Bayer, and R. Lamed. 1991. Isolation and properties of a major cellobiohydrolase from the cellulosome of *Clostridium thermocellum*. J. Bacteriol. **173**:4155-4162.

Morris, E. J. 1988. Characteristics of the adhesion of *Ruminococcus albus* to cellulose. FEMS Microbiology Letters **51**:113-118.

Morris, E. J., and O. J. Cole. 1987. Relationship between cellulolytic activity and adhesion to cellulose in *Ruminococcus albus*. J. Gen. Microbiol. **133**:1023-1032.

Morrison, M., R. I. Mackie, and B. A. White. 1992. Partial characterization of a DNA restriction endonuclease from *Ruminococcus flavefaciens* FD-1 and Its inhibition by site-specific adenine methylation. Appl. Environ. Microbiol. **58**:66-69.

Moser, B., N. R. Gilkes, D. G. Kilburn, R. A. J. Warren, and R. C. Miller, JR. 1989. Purification and characterization of endoglucanase C of *Cellulomonas fimi*, cloning of the gene, and analysis of in vivo transcripts of the gene. Appl. Environ. Microbiol. **55**:2480-2487.

Nolte, A. and F. Mayer. 1989. Localization and immmunological characterization of the cellulolytic enzyme system in *Clostridium thermocellum*. FEMS Micrbiol. Lett. **61**:65-72.

Ohmiya, K., M. Takano, and S. Shimizu. 1990. DNA sequence of a β-glucosidase from *Ruminococcus albus*. Nucleic Acids Res. **18**:671.

Ohmiya, K., T. Kajino, A. Kato, and S. Shimizu. 1989. Structure of a *Ruminococcus albus* endo-1,4-β-glucanase gene. J. Bacteriol. **171**:6771-6775.

Olsen, O., R. Borriss, O. Simon, and K. K. Thomsen. 1991. Hybrid *Bacillus* (1-3, 1-4)-β-glucanases: engineering thermostable enzymes by construction of hybrid genes. Mol. Gen. Genet. 225:177-185.

Ong, E., J. M. Greenwood, N. R. Gilkes, D. G. Kilburn, R. C. Miller, Jr., and R. A. J. Warren. 1989. The cellulose-binding domains of cellulases: tools for biotechnology. Trends in Biotech. **7**:239-243.

Orpin, C. G., and K. N. Joblin. 1988. The rumen anaerobic fungi. p. 129-150. *In*: P. N. Hobson (ed.) The rumen microbial ecosystem, Elsevier Applied Science, London, New York.

Orpin, C. G., S. D. Mathiesen, Y. Greenwood, and A. S. Blix. 1985. Seasonal changes in the ruminal microflora of the high-arctic Svalbard reindeer (*Rangifer tarandus platyrhynchus*). Appl. Environ. Microbiol. **50**:144-151.

Ostroff, G. R., and J. J. Pene. 1983. Molecular cloning with bifunctional plasmid vectors in *Bacillus subtilis*: Isolation of a spontaneous mutant of *Bacillus subtilis* with enhanced transformability for *Escherichia coli*-propagated chimeric plasmid DNA. J. Bacteriol. **156**:934-936.

Pardee, A.B., F. Jacob, and J. Monod. 1959. The genetic control and cytoplasmic expression of "inducibility" in the synthesis of β-galactosidase by *Escherichia coli*. J. Mol. Biol. **1**:165-178.

Patterson, H., R. Irvin, J. W. Costerton, and K. J. Cheng. 1975. Ultrastructure and adhesion properties of *Ruminococcus albus*. J. Bacteriol. **122**:278-287.

Perlman, D., and H. O. Halvorson. 1983. A putative signal peptidase recognition site and sequence in eukaryotic and prokaryotic signal peptidases. J. Mol. Biol. **167**:391-409.

Petre, D., J. Millet, R. Longin, P. Beguin, H. Girard and J.-P. Aubert. 1986. Purification and properties of the endoglucanase C of *Clostridium thermocellum* produced in *Escherchia coli*. Biochimie **68**:687-695.

Pettipher, G. L., and M. J. Latham. 1979a. Characteristics of enzymes produced by *Ruminococcus flavefaciens* which degade plant cell walls. J. Gen. Microbiol. **110**:21-27.

Pettipher, G. L., and M. J. Latham. 1979b. Production of enzymes degrading plant cell walls and fermentation of cellobiose by *Ruminococcus flavefaciens* in batch and continuous culture. J. Gen. Microbiol. **110**:29-38.

Poole, D. M., A. J. Durrant, G. P. Hazlewood, and H. J. Gilbert. 1991. Characterization of hybrid proteins consisting of the catalytic domains of *Clostridium* and *Ruminococcus* endoglucanases, fused to *Pseudomonas* non-catalytic cellulose-binding domains. Biochem. J. **279**:787-792.

Poole, D. M., G. P. Hazlewood, J. I. Laurie, P. J. Barker, and H. J. Gilbert. 1990. Nucleotide sequence of the *Ruminococcus albus* SY3 endoglucanase genes *cel*A and *cel*B. Mol. Gen. Genet. **223**:217-223.

Ramachandra, M., D. L. Crawford, and A. L. Pometto III. 1987. Extracellular enzyme activities during lignocellulose degradation by *Streptomyces* spp.: A comparative study of Wild-Type and genetically manipulated strains. Appl. Environ. Microbiol. **53**:2754-2760.

Rasmussen, M. A., R. B. Hespell, B. A. White, and R. J. Bothat. 1988. Inhibitory effects of methylcellulose on cellulose degradation of *Ruminococcus flavefaciens*. Appl. Environ. Microbiol. **54**:890-897.

Reed, K. C., and D. A. Mann. 1985. Rapid transfer of DNA from agarose gels to nylon membranes. Nucleic Acids Res. **13**:7207-7221.

Retief J. D., B. T. Sewell, and C. Von Holt. 1987. Nucleosome cores Assembled in vitro occupy two preferred frames flanking the histone H1 gene from *Psammechinus miliaris*. Biochemistry **26**:4449-4453.

Rho, D., M. Desrochers, L. Jurasek, H. Driguez, and J. Defaye. 1982. Induction of cellulasse in *Schizophyllum commune*: Thiocellulobiose as a new inducer. J. Bacteriol. **149**:47-53.

Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg. 1977. Labelling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. J. Mol. Biol. **113**:237-251.

Robson, L. M., and G. H. Chambliss. 1986. Cloning of the *Bacillus subtilis* DLG β-1,4-glucanase gene and its expression in *Escherichia coli* and *Bacillus subtilis*. J. Bacteriol. **165**:612-619.

Romaniec, M. P. M., K. Davidson, B. A. White, and G. P. Hazlewood. 1989. Cloning of *Ruminococcus albus* endo-β-1,4-glucanase and xylanase genes. Lett. Appl. Microbiol. **9**:101-104.

Rouvinen, J., T. Bergfors, T. Teeri, J. K. C. Knowles, T. A. Jones. 1990. Three-dimensional structure of cellobiohydrolase II from *Trichderma reesei*. Science **249**:380-386.

Royer, J. C. and J. P. Nakas. 1990. Interrelationship of xylanase induction and cellulase induction of *Trichoderma longibrachiatum*. Appl. Environ. Microbiol. **56**:2535-2539.

Salser, W. 1977. Globin mRNA sequences: analysis of base-pairing and evolutionary implications. Cold Spring Harbor Symp. Quant. Biol. **42**:985-1002.

Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular Cloning: A Laboratory Manual, 2nd Edition. New York: Cold Spring Harbour Laboratory Press.

Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463-5467.

Sasaki, T., T. Tanaka, S. Nakagawa, and K. Kainuma. 1983. Purification and properties of *Cellvibrio gilvus* cellobiose phosphorylase. Biochem. J. **209**:803-807.

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.

Saul, D. J., L. C. Williams, R. A. Grayling, L. W. Chamley, D. R. Love, and P. L. Bergquist. 1990. *cel*B, a gene coding for a bifunctional cellulase from the extreme thermophile *Caldocellum saccharolyticum*. Appl. Environ. Microbiol. **56**:3117-3124.

Schatz, P. J. and J. Beckwith. 1990. Genetic analysis of protein export in *Escherichia coli*. Annu.Rev. Genet. **24**:215-248.

Schmiz, K. L., B. Broll and B. John. 1983. Cellobiose phosphorylase(EC. 2.4.1.20) of *Cellulomonas*:occurrence, induction and its role in cellobiose metabolism. Arch. Microbiol. **135**:241-249.

Scholle R. R., S. M. Robb, F. T. Robb, and D. R. Woods. 1989. Nucleotide sequence and analysis of the *Vibrio alginolyticus* sucrase gene (*scr*B). Gene **80**:49-56.

Schwarz, W. H., S. Schimming, and W. L. Staudenbauer. 1988b. Degradation of barley β-glucan by endoglucanase C of *Clostridium thermocellum*. Appl. Microbiol. Biotech. **29**:25-31.

Schwarz, W. H., S. Schimming, K. P. Rucknagel, S. Burgschwaiger, G. Kreil, and W. L. Staudenbauer. 1988. Nucleotide sequence of the *celC* gene encoding endoglucanase C of *Clostridium thermocellum*. Gene **63**:23-30.

Shen, H., M. Schmuck, I. Pilz, N. R. Gilkes, D. G. Kilburn, R. C. Miller, Jr., and R. A. J. Warren. 1991. Deletion of the linker connecting the catalytic and cellulose-binding domains of endoglucanase A (CenA) of *Cellulomonas fimi* alters its conformation and catalytic activity. J. Biol. chem. **266**:11335-11340.

Shewale, J. G., and J. C. Sadana. 1978. Cellulase and β-glucosidase production by a basidiomycete species. Can. J. Microbiol. **24**:1204-1216.

Shoseyov, O., and R. H. Doi. 1990. Essential 170-kDa subunit for degradation of crystalline cellulose by *Clostridium cellulovorans* cellulase. Proc. Natl. Acad. Sci. USA **87**:2192-2195.

Shoseyov, O., M. Takagi, M. A. Goldstein, and R. H. Doi. 1992. Primary sequence analysis of *Clostridium cellulovorans* cellulose binding protein A. Proc. Natl. Acad. Sci. USA **89**:3483-3487.

Smith, W. R., I. Yu, and R. E. Hungate. 1973. Factors affecting cellulolysis by *Ruminococcus albus*. J. Bacteriol. **114**:729-737.

Sprey, B., and H. P. Bochem. 1991. Electron microscopic observations of cellulose microfibril degradation by endocellulase from *Trichoderma reesei*. FEMS microbiol. Lett. **78**:183-188.

Sternberg, D., and G. R. Mandels. 1979. Induction of cellulolytic enzymes in *Trichoderma reesei* by sophorose. J. Bacteriol. **139**:761-769.

Stewart, B. J., and J. M. Leatherwood. 1976. Derepressed synthesis of cellulase by *Cellulomonas*. J. Bacteriol. **128**:609-615.

Stewart, C. S., and M. P. Bryant. 1988. The rumen bacteria. p. 21-75. *In*: P. N. Hobson (ed.) The rumen microbial ecosystem, Elsevier Applied Science, London, New York.

Struhl, K. 1985. A rapid method for creating recombinant DNA molecules. Bio/Technol. **3**:452-453.

Sykes, R. B., and K. Nordstrom. 1972. Microiodometric determination of β-lactamase activity. Antimicrob. Agents. Chemother. **1**:94-99.

Tabor, S., and C. C. Richardson. 1987. DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. Proc. Natl. Acad. Sci. USA **84**:4767-4771.

Teather, R. M., and P. J. Wood. 1982. Use of Congo Red-polysaccharide interactions in enumeration and characterization of celulolytic bacteria from the Bovine Rumen. Appl. Env. Microbiol. **43**:777-780.

Umile C., and C. P. Kubicak. 1986. Constitutive plasma membrane bound β-glucosidase in *Trichoderma reesei*. FEMS Microbiol. Lett. **34**:291-195.

Utt, E. A., C. K. Eddy, K. F. Keshan, and L. O. Ingram. 1991. Sequencing and expression of the *Butyrivibrio fibrisolvens xyl*B gene encoding a novel bifunctional protein with β-D-xylosidase and α-L-arabinofuranosidase activities. Appl. Environ. Microbiol. **57**:1227-1234.

Vaheri, M., M. Leisola, and V. Kauppinen. 1979. Transglycosylation products of cellulase system of *Trichoderma reesei*. Biotechnol. Lett. **1**:41-46.

Van Gylswyk, N. O., K. Wejdemar, and K. Kulander. 1992. Comparative growth rates of various rumen bacteria in clarified rumen fluid from cows and sheep fed different diets. Appl. Environ. Microbiol. **58**:99-105.

Van Soest, P. J. 1973. The uniformity and nutritive availability of cellulose. Federation Proceeding **32**:1804-1808.

Van Tilbeurgh, H., and M. Claeyssens. 1985. Detection and differentiation of cellulase components using low molecular mass fluorogenic substrates. FEBS lett. **187**:283-288.

Wang, W., and J. A. Thomson. 1990. Nucleotide sequence of the celA gene encoding a cellodextrinase of *Ruminococcus flavefaciens* FD-1. Mol. Gen. Genet. **222**:265-269.

Wang, W., and J. A. Thomson. 1992. (erratum) Nucleotide sequence of the *cel*A gene encoding a cellodextrinase of *Ruminococcus flavefaciens* FD-1. Mol. Gen. Genet. **233**.492.

Ware, C. E., T. Bauchop, and K. Gregg. 1989. The isolation and comparison of cellulase genes from two strains of *Ruminococcus albus*. J. Gen. Microbiol. **135**:921-930.

Watanabe, T., W. Oyanagi, K. Suzuki, and H. Tanaka. 1990. Chitinase system of *Bacillus circulans* WL-12 and importance of chitinase A1 in chitin degradation. J. Bacteriol. **172**:4017-4022.

Weimer, P. J., A. D. French, and T. A. Calamari, JR. 1991. Differential fermentation of cellulose allomorphs by ruminal cellulolytic bacteria. Appl. Environ. Microbiol. **57**:3101-3106.

Weimer, P. J., J. M. Lopez-Guisa, and A. D. French. 1990. Effect of cellulose fine structure on kinetics of its digestion by mixed ruminal microorganisms in vitro. Appl. Environ. Microbiol. **56**:2421-2429.

Whistler, R. L., and E. L. Richards. 1970. Hemicelluloses, p. 447-469. *In*: W. Pigman and D. Horton (ed.), The carbohydrates-chemistry and biochemistry, 2nd ed., vol. 2A. Academic Press Inc., New york.

White, B. A., M. A. Rasmussen, and R. M. Gardner. 1988. Methylcellulose inhibition of exo-β-1,4-glucanase A from *Ruminococcus flavefaciens* FD-1. Appl. Environ. Microbiol. **54**:1634-1636.

Wilkie, K. C. B. 1983. Hemicellulose. Chemtech. p. 306-319.

Wong, K. K. Y., L. U. L. Tan, and J. N. Saddler. 1988. Multiplicity of β-1,4-Xylanase in microorganisms: Functions and applications. Microbiological Reviews **52**:305-317.

Wood, T. M. 1985. Properties of cellulolytic enzyme systems. Biochemical Society Transactions **13**:407-410.

Wood, W. A., and S. T. Kellogg. (Editors) 1988. Biomass Part A: Cellulose and Hemicellulose. Meth. Enzymol. Vol. **160**.

Wood, W. B. 1966. Host specificity of DNA produced by *Escherichia coli*: bacterial mutations affecting the restriction and modification of DNA. J. Mol. Biol. **16**:118-133.

Wynne, E. S., and J. M. Pemberton. 1986. Cloning of a gene cluster from *Cellvibrio mixtus* which codes for cellulase, chitinase, amylase, and pectinase. Appl. Environ. Microbiol. **52**:1362-1367.

Yague, E., P. Beguin, and J. P. Aubert. 1990. Nucleotide sequence and deletion analysis of the cellulase-encoding gene *cel*H of *Clostridium thermocellum*. Gene **89**:61-67.

Zabeau, M. and K. K. Stanley. 1982. Enhanced expression of cro-β-galactosidase fusion proteins under the control of P_R promoter of bacteriophage λ. EMBO J. **1**:1217-1224.

Zappe, H., D. T. Jones and D. R. Woods. 1986. Cloning and expression of *Clostridium acetobutylicum* endoglucanase, cellobiase and amino acid biosynthesis genes in *Escherichia coli*. J. Gen. Microbiol. **132**:1367-1372.

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.

Fierobe H.-P., C. Gaudin, A. Belaich, M. Loutfi, E. Faure, C. Bagnara, D. Baty, and J.-P. Belaich. 1991. Characterization of endoglucanase A from *Clostridium cellulolyticum*. J. Bacteriol. **173**:7956-7962.

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.

Flint, H. J., C. A. McPherson, and J. Martin. 1991. Expression of two xylanase genes from the rumen cellulolytic bacterium *Ruminococcus flavefaciens* 17 cloned in pUC13. J. Gen. Microbiol. **137**:123-129.

Fukumori, F., T. Kudo, Y. Narahashi, and K. Horikoshi. 1986. Molecular cloning and nucleotide sequence of the alkaline cellulase gene from the alkalophilic *Bacillus* sp. strain 1139. J. Gen. Microbiol. **132**:2329-2335.

Fusee, M. C. and J. M. Leatherwood. 1972. Regulation of cellulase from *Ruminococcus*. Can. J. Microbiol. **18**:347-353.

Gardner, R. M., K. C. Doerner, and B. A. White. 1987. Purification and characterization of an exo-β-1,4-glucanase from *Ruminococcus flavefaciens* FD-1. J. Bacteriol. **169**:4581-4588..

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.

Gilbert, H.J., J. Hall, G.P. Hazlewood, and L.M.A. Ferreira. 1990. The N-terminal region of an endoglucanase from *Pseudomonas fluorescens* subspecies *cellulosa* constitutes a cellulose-binding domain that is distinct from the catalytic centre. Mol. Microbiol. **4**:759-767.

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. Microbiological Reviews **55**:303-315.

Gilkes, N. R., R. A. J. Warren, R. C. Miller, Jr., and D. G. Kilburn. 1988. Precise excision of the cellulose binding domains from two *Cellulomonas fimi* cellulases by a homologous protease and the effect on catalysis. J. Biol. Chem. **263**:10401-10407.

Goldstein, J., S. Lehnhardt, and M. Inouye. 1990. Enhancement of protein translocation across the membrane by specific mutation in the hydrophobic region of the signal peptide. J. Bacteriol. **172**:1225-1231.

Gong, C.-S., L. F. Chen, M. C. Flickinger, and G. T. Tsao. 1981. Conversion of hemicellulose carbohydrates. Adv. Biochem. Eng. **20**:93-118.

Gong, J., and C. W. Forsberg. 1989. Factors affecting adhesion of *Fibrobacter succinogenes* subsp. *succinogenes* S85 and adherence-defective mutants to cellulose. Appl. Environ. Microbiol. **55**:3039-3044.

Gong, J., R. Y. C. Lo, and C. W. Forsberg. 1989. Molecular cloning and expression in *Escherichia coli* of a cellodextrinase gene from *Bacteroides succinogenes* S85. Appl. Environ. Microbiol. **55**:132-136.

Gosalbes, M. J., J. A. Perez-Gonzalez, R. Gonzalez, 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.

Grantham, R., C. Gantier, M. Gouy, M. Jacobzone, and R. Mercier. 1981. Codon catalog usage is a genome strategy modulated for gene expressivity. Nucleic Acids Res. **9**:r43-r74.

Greenberg, N. M., R. A. J. Warren, D. G. Kilburn, and R. C. Miller, Jr. 1987a. Regulation, initiation, and termination of the *cen*A and *cex* transcripts of *Cellulomonas fimi*. J. Bacteriol. **169**:646-653.

Greenberg, N. M., R. A. J. Warren, D. G. Kilburn, and R. C. Miller, JR. 1987b. Regulation and initiation of *cen*B transcripts of *Cellulomonas fimi*. J. Bacteriol. **169**:4674-4677.

Groleau, D., and C. W. Forsberg. 1983. Partial characterization of the extracellular carboxymethylcellulase activity produced by the rumen bacterium *Bacteroides succinogenes*. Can. J. Microbiol. **29**:504-517.

Gutierrez, C., J. Barondess, C. Manoil, and J. Beckwith. 1987. The use of transposon TnphoA to detect genes for cell envelope proteins subject to a common regulatory stimulus. J. of Mol. Biol. **195**:289-297

Hall, J., G. P. Hazlewood, N. S. Huskisson, A. J. Durrant, and H. J. Gilbert. 1990. Conserved serine-rich sequences in xylanase and cellulase from *Pseudomonas fluorescens* subspecies *cellulosa*: internal signal sequence and unusual protein processing. Mol. Microbiol. **3**:1211-1219.

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**:29-38.

Halliwell, G., and M. P. Bryant. 1963. The cellulolytic activity of pure strains of bacteria from the rumen of cattle. J. Gen. Microbiol. **32**:441-448.

Hamamoto, T., F. Foong, O. Shoseyov, and R. H. Doi. 1992. Analysis of functional domains of endoglucanases from *Clostridium cellulovorans* by gene cloning, nucleotide sequencing and chimeric protein construction. Mol. Gen. Genet. **231**:472-479.

Hazlewood, G. P., and R. M. Teather. 1988. The genetics of rumen bacteria, p 323-342. *In*: P. N. Hobson (ed.) The rumen microbial ecosystem, Elsevier Applied Science, London, New York.

Heale, J. B., and D. P. Gupta. 1971. The utilization of cellobiose by *Verticillum albo-atrum*. J. Gen. Microbiol. **63**:175-181.

Helaszek, C. T., and B. A. White. 1991. Cellobiose uptake and metabolism by *Ruminococcus flavefaciens*. Appl. Environ. Microbiol. **57**:64-68.

Henikoff S. 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene **28**:351-359.

Henrissat, B., H. Driguez, C. Viet, and M. Schulein. 1985. Synergism of cellulases from *Trichoderma reesei* in the degradation of cellulose. Biotechnol. **3**:722-726.

Henrissat, B., M. Claeyssens, P. Tomme, L. Lemesle, and J. P. Mornon. 1989. Cellulase families revealed by hydrophobic cluster analysis. Gene **81**:83-95.

Hiltner, P., and B. A. Dehority. 1983. Effect of soluble carbohyrates on digestion of cellulose by pure cultures of rumen bacteria. Appl. Environ. Microbiol. **46**:642-648.

Howard, G. T. and B. A. White. 1988. Molecular cloning and expression of cellulase genes from *Ruminococcus albus* 8 in *Escherichia coli* Bacteriophage λ. Appl. Enviro. Microbiol. **54**:1752-1755

Howard, G. T., and B. A. White. 1990. Cloning in *Escherichia coli* of a bi-functional cellulase/xylanase enzyme from *Ruminococcus flavefaciens* FD-1. Animal Biotechnology **1**:95-106.

Hrmova, M., E. Petrakova, and P. Biely. 1991. Induction of cellulose- and xylan-degrading enzyme systems in *Aspergillus terreus* by homo- and heterodisaccharides composed of glucose and xylose. J. Gen. Microbiol. **137**:541-547.

Hrmova, M., P. Biely, and M. Vrsanska. 1986. Specificity of cellulase and β-xylanase induction in *Trichoderma reesei* QM9414. Arch. Microbiol. **144**:307-311.

Hrmova, M., P. Biely, and M. Vrsanska. 1989. Cellulose-xylan- degrading enzymes of *Aspergillus terreus* and *Aspergillus niger*. Enzyme and Microbial Technology **11**:610-616.

Huang L. and C. W. Forsberg. 1990. Cellulose digestion and cellulase regulation and distribution in *Fibrobacter succinogenes* subsp. *succinogenes* S85. Appl. Environ. Microbiol. **56**:1221-1228.

Huang, C-M., W. J. Kelly, R. V. Asmundson, and P-L. Yu. 1989. Molecular cloning and expression of multiple cellulase genes of *Ruminococcus flavefaciens* strain 186 in *Escherichia coli*. Appl. Microbiol. Biotechnol. **31**:265-271.

Huang, J., and M. A. Schell. 1992. Role of the two-component leader sequence and mature amino acid sequences in extracellular export of endoglucanase EGL from *Pseudomonas solanacearum*. J. Bacteriol. **174**:1314-1323.

Huang, L., and C. W. Forsberg. 1987. Isolation of a cellodextrinase from *Bacteroides succinogenes*. Appl. Environ. Microbiol. **53**:1034-1041.

Huang, L., and C. W. Forsberg. 1990. Cellulose digestion and cellulase regulation and distribution in *Fibrobacter succinogenes* subsp. *succinogenes* S85. Appl. Environ. Microbiol. **56**:1221-1228.

Hungate, R. E. 1988. The ruminant and the rumen. p. 1-19. *In*: P. N. Hobson (ed.) The rumen microbial ecosystem, Elsevier Applied Science, London, New York.

Inglin, M., B. A. Feinberg, and J. R. Loewenberg. 1980. Partial purification and characterization of a new intracellular β-glucosidase of *Trichoderma reesei*. Biochemical J. **185**:515-519.

Ish-Horowicz, D. and J. F. Burke. 1981. Rapid and efficient cosmid cloning. Nucleic Acids Res. **9**:2989-2998.

Johnson, E. A., F. Bouchot, and A. L. Demain. 1985. Regulation of cellulase formation in *Clostridium thermocellum*. J. Gen. Microbiol. **131**:2303-2308.

Juy, M., A. G. Amit, P. M. Alzari, R. J. Poljak, M. Claeyssens, P. Beguin, and J.-P. Aubert. 1992. Three-dimensional structure of a thermostable bacterial cellulase. Nature **357**:89-91.

Kirk, T. K., M. Tien, P. J. Kersten, and M. D. Mozuch. 1986. Ligninase of *Phanerochaete chrysosporium*; mechanism of its degradation of the non phenolic aryglycerol β-aryl ether substructure of lignin. Biochem. J. **236**:279-287.

Klump, H. H., and D. L. Maeder. 1991. The thermodynamic basis of the genetic code. Pure Appl. Chem. **63**:1357-1366.

Knowles, J., P. Lehtovaara, and T. Teeri. 1987. Cellulase families and their genes. Trends in Biotechnol. **5**:255-261.

Kubicek, C. P. 1987. Involvement of a conidial endoglucanase and a plasma-membrane-bound β-glucosidase in the induction of endoglucanase synthesis by cellulose in *Trichoderma reesei*. J. Gen. Microbiol. **133**:1481-1487.

Kudo, H., K.-J. Cheng, and J. W. Costerton. 1987. Electron microscopic study of the methylcellulose-mediated detachment of cellulolytic rumen bacteria from cellulose fibres. Can. J. Microbiol. **33**:267-272.

Kundu, R. K., S. Dube, and D. K. Dube. 1988. Extracellular cellulolytic enzyme system of *Aspergillus japonicus*: 3. Isolation, purification, and characterization of multiple forms of endoglucanase. Enzyme Microb. Technol. **10**:100-109.

Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227**:680-685.

Lamed, R., and E. A. Bayer. 1988. The cellulosome of *Clostridium thermocellum*. Adv. Appl. Microbiol. **33**:1-46.

Latham, M. J., B. E. Brooker, G. L. Pettipher, and P. J. Harris. 1978. *Ruminococcus flavefaciens* cell coat and Adhesion to cotton cellulose and to cell walls in leaves of perennial ryegrass (*Lolium perenne*). Appl. Environ. Microbiol. **35**:156-165.

Leatherwood, J. M. 1973. Cellulose degradation by *Ruminococcus*. Federation Proceedings **32**:1814-1818.

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.

Lin L-L., and J.A. Thomson. 1991. Cloning, sequencing and expression of a gene encoding a 73 kDa xylanase enzyme from the rumen anaerobe *Butyrivibrio fibrisolvens* H17c. Mol. Gen. Genet. **228**:55-61.

Lin, E., and D. B. Wilson. 1987. Regulation of β-1,4-Endoglucanase synthesis in *Thermomonospora fusca*. Appl. Environ. Microbiol. **53**:1352-1357.

Lin, E., and D. B. Wilson. 1988. Transcription of the *celE* gene in *Thermomonospora fusca*. J. Bacteriol. **170**:3838-3842.

Lin, L-L., E. Rumbak, H. Zappe, J. A. Thomson, and D. R. Woods. 1990. Cloning, Sequencing and analysis of expression of a *Butyrivibrio fibrisolvens* gene encoding a β-glucosidase. J. Gen. Microbiol. **136**:1567-1576.

Ljungdahl, L. G., and K.-E. Eriksson. 1985. **8**:237-299. K. C. Marshall(editors), Adv. Microb. Ecol.: Ecology of microbial cellulose degradation. PLenum Press. New York and London.

Luthi, E., D. R. Love, J. McAnulty, C. Wallace, P. A. Caughey, D. Saul, and P. L. Bergquist. 1990. Cloning, sequence analysis, and expression of genes encoding xylan-degrading enzymes from the thermophile *Caldocellum saccharolyticum*. Appl. Environ. Microbiol. **56**:1017-1024.

Luthi, E., N. B. Jasmat, R. A. Grayling, D. R. Love, and P. L. Bergquist. 1991. Cloning, sequence analysis, and expression in *Escherichia coli* of a gene coding for a β-mannanase from the extremely thermophilic bacterium *Caldocellum saccharolyticum*. Appl. Environ. Microbiol. **57**:694-700.

Mackenzie, C. R., D. Bilous, H. Schneider, and K. G. Johnson. 1987. Induction of cellulolytic and xylanolytic enzyme systems in *Streptomyces* spp. Appl. Environ. Microbiol. **53**:2835-2839.

Mandels, M., and E. T. Reese. 1960. Induction of cellulase in fungi by cellobiose. J. Bacteriol. **79**:816-826.

Mandels, M., F. W. Parrish, and E. T. Reese. 1962. Sophorose as an inducer of cellulase in *Trichoderma viride*. J. Bacteriol. **83**:400-408.

Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: A Laboratory Manual. New York: Cold Spring Harbour Laboratory.

Manoil C., and J. Beckwith. 1985. A transposon probe for protein export signals. Proc. Natl. Acad. Sci. USA **82**:8129-8133.

Matte, A. and C. W. Forsberg. 1992. Purification, characterization, and mode of action of endoxylanase 1 and 2 from *Fibrobacter succinogenes* S85. Appl. Environ. Microbiol. **58**:157-168.

Mayer, F., M. P. Coughlan, Y. Mori, and L. G. Ljungdahl. 1987. Macromolecular organization of the cellulolytic enzyme complex of *Clostridium thermocellum* as revealed by electron microscopy. Appl. Environ. Microbiol. **53**:2785-2792.

McGavin, M. J. Lam, and C. W. Forsberg. 1990. Regulation and distribution of *Fibrobacter succinogenes* subsp. *succinogenes* S85 endoglucanases. Appl. Environ. MNicrobiol. **56**:1235-1244.

McGavin, M. J., C. W. Forsberg, B. Bell, A. W. Crosby, D. Dignard, and D. Y. Thomas. 1989. Structure of the *cel*-3 gene from *Fibrobacter succinogenes* S85 and characteristics of the encoded gene product, endoglucanase 3. J. Bacteriol. **171**:5587-5595.

Meinke, A., N. R. Gilkes, D. G. Kilburn, R. C. Miller, JR., and R. A. J. Warren. 1991. Multiple domains in endoglucanase B (CenB) from *Cellulomonas fimi*: Functions and Relatedness to domains in other polypeptides. J. Bacteriol. **173**:7126-7135.

Meselson, M. and R. Yuan. 1968. DNA restriction enzyme from *E. coli*. Nature **217**:1110.

Messing, J., R. Crea, and P. H. Seeburg. 1981. A system for shotgun DNA sequencing. Nucleic Acids Res. **9**:309-321.

Miller, G. L. 1959. Use of dinitrosalicylic acid reagent for determination of reducing sugar. Anal. Chem. **31**:426-428.

Miller, J. H. 1972. Experiments in molecular genetics. Cold Spring Harbor laboratory, Cold Spring Harbor, NY.

Miller, P. B., H. Shen, N. R. Gilkes, D. G. Kilburn, R. C. Miller, Jr., A. G. Plaut, and R. A. J. Warren. 1992. Endoglucanase A from *Cellulomonas fimi* in which the hinge sequence of human IgA1 is substituted for the linker connecting its two domains is hydrolyzed by IgA proteases from *Neisseria gonorrhoeae*. FEMS Microbiol. Lett. **92**:199-204.

Morag, E., I. Halevy, E. A. Bayer, and R. Lamed. 1991. Isolation and properties of a major cellobiohydrolase from the cellulosome of *Clostridium thermocellum*. J. Bacteriol. **173**:4155-4162.

Morris, E. J. 1988. Characteristics of the adhesion of *Ruminococcus albus* to cellulose. FEMS Microbiology Letters **51**:113-118.

Morris, E. J., and O. J. Cole. 1987. Relationship between cellulolytic activity and adhesion to cellulose in *Ruminococcus albus*. J. Gen. Microbiol. **133**:1023-1032.

Morrison, M., R. I. Mackie, and B. A. White. 1992. Partial characterization of a DNA restriction endonuclease from *Ruminococcus flavefaciens* FD-1 and Its inhibition by site-specific adenine methylation. Appl. Environ. Microbiol. **58**:66-69.

Moser, B., N. R. Gilkes, D. G. Kilburn, R. A. J. Warren, and R. C. Miller, JR. 1989. Purification and characterization of endoglucanase C of *Cellulomonas fimi*, cloning of the gene, and analysis of in vivo transcripts of the gene. Appl. Environ. Microbiol. **55**:2480-2487.

Nolte, A. and F. Mayer. 1989. Localization and immmunological characterization of the cellulolytic enzyme system in *Clostridium thermocellum*. FEMS Micrbiol. Lett. **61**:65-72.

Ohmiya, K., M. Takano, and S. Shimizu. 1990. DNA sequence of a β-glucosidase from *Ruminococcus albus*. Nucleic Acids Res. **18**:671.

Ohmiya, K., T. Kajino, A. Kato, and S. Shimizu. 1989. Structure of a *Ruminococcus albus* endo-1,4-β-glucanase gene. J. Bacteriol. **171**:6771-6775.

Olsen, O., R. Borriss, O. Simon, and K. K. Thomsen. 1991. Hybrid *Bacillus* (1-3, 1-4)-β-glucanases: engineering thermostable enzymes by construction of hybrid genes. Mol. Gen. Genet. **225**:177-185.

Ong, E., J. M. Greenwood, N. R. Gilkes, D. G. Kilburn, R. C. Miller, Jr., and R. A. J. Warren. 1989. The cellulose-binding domains of cellulases: tools for biotechnology. Trends in Biotech. **7**:239-243.

Orpin, C. G., and K. N. Joblin. 1988. The rumen anaerobic fungi. p. 129-150. *In*: P. N. Hobson (ed.) The rumen microbial ecosystem, Elsevier Applied Science, London, New York.

Orpin, C. G., S. D. Mathiesen, Y. Greenwood, and A. S. Blix. 1985. Seasonal changes in the ruminal microflora of the high-arctic Svalbard reindeer (*Rangifer tarandus platyrhynchus*). Appl. Environ. Microbiol. **50**:144-151.

Ostroff, G. R., and J. J. Pene. 1983. Molecular cloning with bifunctional plasmid vectors in *Bacillus subtilis*: Isolation of a spontaneous mutant of *Bacillus subtilis* with enhanced transformability for *Escherichia coli*-propagated chimeric plasmid DNA. J. Bacteriol. **156**:934-936.

Pardee, A.B., F. Jacob, and J. Monod. 1959. The genetic control and cytoplasmic expression of "inducibility" in the synthesis of β-galactosidase by *Escherichia coli*. J. Mol. Biol. **1**:165-178.

Patterson, H., R. Irvin, J. W. Costerton, and K. J. Cheng. 1975. Ultrastructure and adhesion properties of *Ruminococcus albus*. J. Bacteriol. **122**:278-287.

Perlman, D., and H. O. Halvorson. 1983. A putative signal peptidase recognition site and sequence in eukaryotic and prokaryotic signal peptidases. J. Mol. Biol. **167**:391-409.

Petre, D., J. Millet, R. Longin, P. Beguin, H. Girard and J.-P. Aubert. 1986. Purification and properties of the endoglucanase C of *Clostridium thermocellum* produced in *Escherchia coli*. Biochimie **68**:687-695.

Pettipher, G. L., and M. J. Latham. 1979a. Characteristics of enzymes produced by *Ruminococcus flavefaciens* which degade plant cell walls. J. Gen. Microbiol. **110**:21-27.

Pettipher, G. L., and M. J. Latham. 1979b. Production of enzymes degrading plant cell walls and fermentation of cellobiose by *Ruminococcus flavefaciens* in batch and continuous culture. J. Gen. Microbiol. **110**:29-38.

Poole, D. M., A. J. Durrant, G. P. Hazlewood, and H. J. Gilbert. 1991. Characterization of hybrid proteins consisting of the catalytic domains of *Clostridium* and *Ruminococcus* endoglucanases, fused to *Pseudomonas* non-catalytic cellulose-binding domains. Biochem. J. **279**:787-792.

Poole, D. M., G. P. Hazlewood, J. I. Laurie, P. J. Barker, and H. J. Gilbert. 1990. Nucleotide sequence of the *Ruminococcus albus* SY3 endoglucanase genes *cel*A and *cel*B. Mol. Gen. Genet. **223**:217-223.

Ramachandra, M., D. L. Crawford, and A. L. Pometto III. 1987. Extracellular enzyme activities during lignocellulose degradation by *Streptomyces* spp.: A comparative study of Wild-Type and genetically manipulated strains. Appl. Environ. Microbiol. **53**:2754-2760.

Rasmussen, M. A., R. B. Hespell, B. A. White, and R. J. Bothat. 1988. Inhibitory effects of methylcellulose on cellulose degradation of *Ruminococcus flavefaciens*. Appl. Environ. Microbiol. **54**:890-897.

Reed, K. C., and D. A. Mann. 1985. Rapid transfer of DNA from agarose gels to nylon membranes. Nucleic Acids Res. **13**:7207-7221.

Retief J. D., B. T. Sewell, and C. Von Holt. 1987. Nucleosome cores Assembled in vitro occupy two preferred frames flanking the histone H1 gene from *Psammechinus miliaris*. Biochemistry **26**:4449-4453.

Rho, D., M. Desrochers, L. Jurasek, H. Driguez, and J. Defaye. 1982. Induction of cellulasse in *Schizophyllum commune*: Thiocellulobiose as a new inducer. J. Bacteriol. **149**:47-53.

Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg. 1977. Labelling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. J. Mol. Biol. **113**:237-251.

Robson, L. M., and G. H. Chambliss. 1986. Cloning of the *Bacillus subtilis* DLG β-1,4-glucanase gene and its expression in *Escherichia coli* and *Bacillus subtilis*. J. Bacteriol. **165**:612-619.

Romaniec, M. P. M., K. Davidson, B. A. White, and G. P. Hazlewood. 1989. Cloning of *Ruminococcus albus* endo-β-1,4-glucanase and xylanase genes. Lett. Appl. Microbiol. **9**:101-104.

Rouvinen, J., T. Bergfors, T. Teeri, J. K. C. Knowles, T. A. Jones. 1990. Three-dimensional structure of cellobiohydrolase II from *Trichderma reesei*. Science **249**:380-386.

Royer, J. C. and J. P. Nakas. 1990. Interrelationship of xylanase induction and cellulase induction of *Trichoderma longibrachiatum*. Appl. Environ. Microbiol. **56**:2535-2539.

Salser, W. 1977. Globin mRNA sequences: analysis of base-pairing and evolutionary implications. Cold Spring Harbor Symp. Quant. Biol. **42**:985-1002.

Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular Cloning: A Laboratory Manual, 2nd Edition. New York: Cold Spring Harbour Laboratory Press.

Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463-5467.

Sasaki, T., T. Tanaka, S. Nakagawa, and K. Kainuma. 1983. Purification and properties of *Cellvibrio gilvus* cellobiose phosphorylase. Biochem. J. **209**:803-807.

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.

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.

Schatz, P. J. and J. Beckwith. 1990. Genetic analysis of protein export in *Escherichia coli*. Annu.Rev. Genet. **24**:215-248.

Schmiz, K. L., B. Broll and B. John. 1983. Cellobiose phosphorylase(EC. 2.4.1.20) of *Cellulomonas*:occurrence, induction and its role in cellobiose metabolism. Arch. Microbiol. **135**:241-249.

Scholle R. R., S. M. Robb, F. T. Robb, and D. R. Woods. 1989. Nucleotide sequence and analysis of the *Vibrio alginolyticus* sucrase gene (*scr*B). Gene **80**:49-56.

Schwarz, W. H., S. Schimming, and W. L. Staudenbauer. 1988b. Degradation of barley β-glucan by endoglucanase C of *Clostridium thermocellum*. Appl. Microbiol. Biotech. **29**:25-31.

Schwarz, W. H., S. Schimming, K. P. Rucknagel, S. Burgschwaiger, G. Kreil, and W. L. Staudenbauer. 1988. Nucleotide sequence of the *celC* gene encoding endoglucanase C of *Clostridium thermocellum*. Gene **63**:23-30.

Shen, H., M. Schmuck, I. Pilz, N. R. Gilkes, D. G. Kilburn, R. C. Miller, Jr., and R. A. J. Warren. 1991. Deletion of the linker connecting the catalytic and cellulose-binding domains of endoglucanase A (CenA) of *Cellulomonas fimi* alters its conformation and catalytic activity. J. Biol. chem. **266**:11335-11340.

Shewale, J. G., and J. C. Sadana. 1978. Cellulase and β-glucosidase production by a basidiomycete species. Can. J. Microbiol. **24**:1204-1216.

Shoseyov, O., and R. H. Doi. 1990. Essential 170-kDa subunit for degradation of crystalline cellulose by *Clostridium cellulovorans* cellulase. Proc. Natl. Acad. Sci. USA **87**:2192-2195.

Shoseyov, O., M. Takagi, M. A. Goldstein, and R. H. Doi. 1992. Primary sequence analysis of *Clostridium cellulovorans* cellulose binding protein A. Proc. Natl. Acad. Sci. USA **89**:3483-3487.

Smith, W. R., I. Yu, and R. E. Hungate. 1973. Factors affecting cellulolysis by *Ruminococcus albus*. J. Bacteriol. **114**:729-737.

Sprey, B., and H. P. Bochem. 1991. Electron microscopic observations of cellulose microfibril degradation by endocellulase from *Trichoderma reesei*. FEMS microbiol. Lett. **78**:183-188.

Sternberg, D., and G. R. Mandels. 1979. Induction of cellulolytic enzymes in *Trichoderma reesei* by sophorose. J. Bacteriol. **139**:761-769.

Stewart, B. J., and J. M. Leatherwood. 1976. Derepressed synthesis of cellulase by *Cellulomonas*. J. Bacteriol. **128**:609-615.

Stewart, C. S., and M. P. Bryant. 1988. The rumen bacteria. p. 21-75. *In*: P. N. Hobson (ed.) The rumen microbial ecosystem, Elsevier Applied Science, London, New York.

Struhl, K. 1985. A rapid method for creating recombinant DNA molecules. Bio/Technol. **3**:452-453.

Sykes, R. B., and K. Nordstrom. 1972. Microiodometric determination of β-lactamase activity. Antimicrob. Agents. Chemother. **1**:94-99.

Tabor, S., and C. C. Richardson. 1987. DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. Proc. Natl. Acad. Sci. USA **84**:4767-4771.

Teather, R. M., and P. J. Wood. 1982. Use of Congo Red-polysaccharide interactions in enumeration and characterization of celulolytic bacteria from the Bovine Rumen. Appl. Env. Microbiol. **43**:777-780.

Umile C., and C. P. Kubicak. 1986. Constitutive plasma membrane bound β-glucosidase in *Trichoderma reesei*. FEMS Microbiol. Lett. **34**:291-195.

Utt, E. A., C. K. Eddy, K. F. Keshan, and L. O. Ingram. 1991. Sequencing and expression of the *Butyrivibrio fibrisolvens xyl*B gene encoding a novel bifunctional protein with β-D-xylosidase and α-L-arabinofuranosidase activities. Appl. Environ. Microbiol. **57**:1227-1234.

Vaheri, M., M. Leisola, and V. Kauppinen. 1979. Transglycosylation products of cellulase system of *Trichoderma reesei*. Biotechnol. Lett. **1**:41-46.

Van Gylswyk, N. O., K. Wejdemar, and K. Kulander. 1992. Comparative growth rates of various rumen bacteria in clarified rumen fluid from cows and sheep fed different diets. Appl. Environ. Microbiol. **58**:99-105.

Van Soest, P. J. 1973. The uniformity and nutritive availability of cellulose. Federation Proceeding **32**:1804-1808.

Van Tilbeurgh, H., and M. Claeyssens. 1985. Detection and differentiation of cellulase components using low molecular mass fluorogenic substrates. FEBS lett. **187**:283-288.

Wang, W., and J. A. Thomson. 1990. Nucleotide sequence of the celA gene encoding a cellodextrinase of *Ruminococcus flavefaciens* FD-1. Mol. Gen. Genet. **222**:265-269.

Wang, W., and J. A. Thomson. 1992. (erratum) Nucleotide sequence of the *cel*A gene encoding a cellodextrinase of *Ruminococcus flavefaciens* FD-1. Mol. Gen. Genet. **233**.492.

Ware, C. E., T. Bauchop, and K. Gregg. 1989. The isolation and comparison of cellulase genes from two strains of *Ruminococcus albus*. J. Gen. Microbiol. **135**:921-930.

Watanabe, T., W. Oyanagi, K. Suzuki, and H. Tanaka. 1990. Chitinase system of *Bacillus circulans* WL-12 and importance of chitinase A1 in chitin degradation. J. Bacteriol. **172**:4017-4022.

Weimer, P. J., A. D. French, and T. A. Calamari, JR. 1991. Differential fermentation of cellulose allomorphs by ruminal cellulolytic bacteria. Appl. Environ. Microbiol. **57**:3101-3106.

Weimer, P. J., J. M. Lopez-Guisa, and A. D. French. 1990. Effect of cellulose fine structure on kinetics of its digestion by mixed ruminal microorganisms in vitro. Appl. Environ. Microbiol. **56**:2421-2429.

Whistler, R. L., and E. L. Richards. 1970. Hemicelluloses, p. 447-469. *In*: W. Pigman and D. Horton (ed.), The carbohydrates-chemistry and biochemistry, 2nd ed., vol. 2A. Academic Press Inc., New york.

White, B. A., M. A. Rasmussen, and R. M. Gardner. 1988. Methylcellulose inhibition of exo-β-1,4-glucanase A from *Ruminococcus flavefaciens* FD-1. Appl. Environ. Microbiol. **54**:1634-1636.

Wilkie, K. C. B. 1983. Hemicellulose. Chemtech. p. 306-319.

Wong, K. K. Y., L. U. L. Tan, and J. N. Saddler. 1988. Multiplicity of β-1,4-Xylanase in microorganisms: Functions and applications. Microbiological Reviews **52**:305-317.

Wood, T. M. 1985. Properties of cellulolytic enzyme systems. Biochemical Society Transactions **13**:407-410.

Wood, W. A., and S. T. Kellogg. (Editors) 1988. Biomass Part A: Cellulose and Hemicellulose. Meth. Enzymol. Vol. **160**.

Wood, W. B. 1966. Host specificity of DNA produced by *Escherichia coli*: bacterial mutations affecting the restriction and modification of DNA. J. Mol. Biol. **16**:118-133.

Wynne, E. S., and J. M. Pemberton. 1986. Cloning of a gene cluster from *Cellvibrio mixtus* which codes for cellulase, chitinase, amylase, and pectinase. Appl. Environ. Microbiol. **52**:1362-1367.

Yague, E., P. Beguin, and J. P. Aubert. 1990. Nucleotide sequence and deletion analysis of the cellulase-encoding gene *cel*H of *Clostridium thermocellum*. Gene **89**:61-67.

Zabeau, M. and K. K. Stanley. 1982. Enhanced expression of cro-β-galactosidase fusion proteins under the control of P_R promoter of bacteriophage λ. EMBO J. **1**:1217-1224.

Zappe, H., D. T. Jones and D. R. Woods. 1986. Cloning and expression of *Clostridium acetobutylicum* endoglucanase, cellobiase and amino acid biosynthesis genes in *Escherichia coli*. J. Gen. Microbiol. **132**:1367-1372.

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.