Analysis of genes and enzymes involved in the degradation of cellulose and proteins by *Butyrivibrio fibrisolvens* H17c

Eldie Berger

In partial fulfillment of the requirements for the degree of

Doctor of Philosophy in the Faculty of Science,

University of Cape Town.

CAPE TOWN

AUGUST 1990

CERTIFICATION OF SUPERVISOR

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

Signed by candidate

Professor D. R. Woods

Deputy Vice-Chancellor (Research)

University of Cape Town

Acknowledgements

I am greatly indebted to my supervisor, Professor D.R. Woods, for his enthusiastic guidance and encouragement.

I am grateful to Dr. Rod Mackie for his involvement at the beginning of this project; and to the Animal and Dairy Science Research Institute for financial support and the freedom to carry out these studies at the University of Cape Town.

Many thanks to the members of the "Anaerobe Unit" over the years for the relaxed working environment; and special thanks to Dr. Harold Zappe for dependable technical advice, and also to Harold and Joe Santangelo for laying low many computer "bugs". I appreciate greatly the many hours spent by Dr.'s Heide Goodman and Val Abratt, as well as Greg Blatch, on proofreading and critical review of the manuscript.

To my husband, Dave, your positive input and criticism (when necessary?) has been invaluable. Your insistence on taking "short breathers", jogging along the contours of Table Mountain, left the body exhausted, but the mind refreshed. Writing up in tandem has been great fun, especially finishing first!

Dedicated to my parents,
Ben and Elsa Strydom

Contents

Abstract

Butyrivibrio fibrisolvens H17c is a gram-negative obligate anaerobic bacterium found in the rumen of most ruminants. The aim of this thesis was to investigate the enzymes produced by *B. fibrisolvens* H17c involved in the degradation of cellulose, xylan, and protein.

A library of chromosomal DNA fragments from *B. fibrisolvens* H17c was established in the plasmid p*EcoR*251, an *Escherichia coli* positive selection vector. The library was screened for genes expressing cellulase, xylanase, and protease activity. Two genes expressing endo-β-1,4-glucanase and cellodextrinase activity were cloned in *E. coli* as host.

The gene expressing endo-β-1,4-glucanase activity (*end1*) was cloned on a recombinant plasmid pES400. The *end1* gene was located on a 6.8 kb DNA fragment and expressed from its own promoter in the *E. coli* host. It was shown that 64% of the endoglucanase activity was located in the periplasm of the *E. coli* host. Tn*phoA* mutagenesis indicated the presence of a functional *E. coli*-like signal peptide. The nucleotide sequence of *end1* was determined and the amino acid sequence (547 amino acids) deduced. The catalytic domain of End1 showed very good similarity to the catalytic domain of the *Clostridium thermocellum* EGE endoglucanase. End1 also has a non-catalytic domain similar to the binding domains of the CenA and Cex cellulases from *Cellulomonas fimi*

The gene expressing cellodextrinase activity (*ced1*) was cloned on a recombinant plasmid pES500. This gene was located on a 3.55 kb fragment and was also expressed from its own promoter in the *E. coli* host. The Ced1 enzyme was also exported to the periplasm of the *E. coli* host, but did not contain a functional *E. coli*-like signal peptide. The nucleotide sequence was determined and the deduced amino acid sequence (547 residues) showed high similarity to the catalytic domain of the *C. thermocellum* EGD endoglucanase. The proteins of End1 and Ced1 showed no similarity.

The End1 and Ced1 enzymes were characterized using a range of different substrates. The End1 enzyme showed optimal activity at pH 5.6 and 45°C. Optimal activity for the Ced1 enzyme was obtained at pH 6.6 and 50°C.

The proteolytic activity of *B. fibrisolvens* H17c was characterized using gelatin-SDS-PAGE. Ten bands of protease activity with apparent molecular weights ranging between 42 000 and 101 000 were detected at different stages during the growth cycle. The effect of protease inhibitors indicated that all ten protease bands were serine proteases. Optimal activity was observed between pH 6.0 to 7.5 and at a temperature of 50°C. The proteolytic activity of *B. fibrisolvens* H17c varied depending on the type of carbohydrate substrate in the medium, and was positively correlated with the growth rate.

Abbreviations

A	adenosine
A_{420}	absorbance at 420 nm
Ap	ampicillin
ATCC	American Type culture collection
ATP	adenosine 5'-triphosphate
bp	base pair
BSA	bovine serum albumin
C-	carboxy terminal (end of a protein)
CBH	cellobiohydrolase
Cm	chloramphenicol
CsCl	caesium chloride
CMC	carboxymethylcellulose
CMCase	carboxymethylcellulase
DMSO	dimethyl sulfoxide
dNTP	deoxynucleotide triphosphate
DNA	deoxyribonucleic acid
DNS	dinitrosalicylic acid
DP	degree of polymerization
DS	degree of substitution
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
LB	Luria-Bertani broth
MM	minimal medium
M10	*B. fibrisolvens* non-rumen fluid medium
MeUMb	methylumbelliferyl
min	minute(s)
mRNA	messenger RNA
M_r	relative molecular mass
N-	amino terminal (end of protein)
nt	nucleotides
4NPX	*p*-nitrophenyl-β-D-xylopyranoside

Abbreviations (cont.)

OD_{600}	optical density at 600 nm
ONPG	*o*-nitrophenyl-β-D-galactopyranoside
ONPP	*o*-nitrophenyl phosphate
ORF	open reading frame
p	plasmid
PAGE	polyacrylamide gel electrophoresis
PC	phosphate-citrate buffer
pHMB	*p*-hydroxymercuribenzoate, sodium salt
Pho	alkaline phosphatase
phoA	gene coding for alkaline phosphatase
PMSF	phenylmethylsulphonyl fluoride
pNPG	*p*-nitrophenyl-β-D-glucopyranoside
pNPC	*p*-nitrophenyl-β-D-cellobiose
pNPX	*p*-nitrophenyl-β-D-xylopyranoside
P_R	rightward promoter (phage lambda)
r	(superscript) resistance
RNA	ribonucleic acid
s	second(s)
SDS	sodium dodecyl sulfate
sp(p)	species
TAE	tris-acetate EDTA buffer
TCA	trichloroacetic acid
TEMED	*N,N,N',N'*-tetramethylethlenediamine
Tn	transposon
Tris	Tris(hydroxymethyl)aminomethane
U	units of enzyme activity
UV	ultraviolet (light)
v/v	volume/volume
w/v	weight/volume
XGal	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	lamda
μ	micro
φ	phi
δ	sigma

Chapter 1

General Introduction

Chapter 1

General Introduction

1.1 ASPECTS OF RUMINANT NUTRITION

1.1.1 Importance of ruminants as a food source.

Feedstuffs consumed by ruminants are all initially exposed to the fermentative activity in the rumen prior to gastric and intestinal digestion. Dietary polysaccharides and protein are generally degraded by the rumen organisms to produce specific end products which in turn provide nutrients for metabolism by the host animal. The anaerobic nature and other characteristics of the ruminal fermentation place an upper limit on the potential nutrient yield (Mackie and Kistner 1985). However, the ruminant animal provides a system for providing human food from materials which cannot be utilized directly by man. This places the ruminant in a central role in the total effective use of the world's nutrient resources. Phillips (1981) anticipated that the world's population by the turn of the century will be of the order of 6.5 billion, and approximately 5 billion of these people will live in countries with a food shortage. Thus, the prospects for achieving either an adequate level of production of animal products in relation to overall world needs, or a reasonable balance in supply of animal products between the developing and developed countries are not encouraging (Phillips 1981). The primary objective of research scientists working in the field of rumen function is, therefore, to advance animal production through better understanding of the mechanisms involved in digestion and metabolism in the rumen.

1.1.2 Energy metabolism in rumen bacteria.

All organisms require a continuous supply of energy, in order to maintain their structure, to grow, and to multiply. Until recently it was thought that most rumen anaerobes synthesized ATP exclusively by substrate-level phosphorylation (SLP). Hungate (1966) found that the majority of ruminal bacteria which ferment a variety

of carbohydrates were dependent on SLP for the major portion of ATP synthesized. However, it has been shown that many rumen bacteria possess components of electron transfer chains that are essential for ATP synthesis. Dawson et al. (1979) found that *Bacteroides ruminicola, Bacteroides succinogenes* (reclassified as *Fibrobacter succinogenes* subsp. *succinogenes*; Montgomery et al. 1988), and *Butyrivibrio fibrisolvens* could derive 33, 50, and 26% respectively, of their total molar growth yield/mole of carbohydrate fermented from electron transport reactions. These electron transfer linked phosphorylation mechanisms (ETP) could have a profound influence on the overall energetic efficiency of the ruminal fermentation (Mackie and Kistner 1985).

There are four biochemical reactions known to occur in ruminal bacteria which yield ATP via ETP. The most important of these is the reduction of CO_2 to CH_4, and the ruminal ecosystem contains at least three species of methanogens able to obtain ATP in this way, viz. *Methanobrevibacter ruminantium* (Smith and Hungate 1958), *Methanobacterium formicicum* (Bryant 1965), and *Methanomicrobium mobile* (Paynter and Hungate 1968). Sulphate-reducing bacteria (*Desulfovibrio* sp) are known to occur in the rumen (Howard and Hungate 1976), and it was shown that ATP is produced via an electron transfer chain and the reduction of SO_4^{2-} to S^{2-}. Nitrate reduction to nitrite has been reported in *Selenomonas ruminantium* and *Anaerovibrio lipolytica* (De Vries et al. 1974), and the reduction of fumarate to succinate was shown in *Wolinella succinogenes* (Wolin et al. 1961).

Ruminal organisms are subjected to periods of nutritional stress and starvation at different times in the feeding cycle when concentrations of free sugars drop to very low levels and insoluble polymers are being degraded very slowly. The survival capacity of ruminal bacteria is considerably less than that of other bacteria found in the soil or aquatic sediments (Mink and Hespell 1981a; 1981b). Results obtained (Hespell 1984) suggested that ruminal bacteria are unable to reduce their rate of metabolism in an environment with low levels of readily available organic nutrients. A proper understanding of the energy status of ruminal bacteria during growth and starvation is thus needed.

1.1.3 Nitrogen metabolism in the rumen

The amino acid requirements of the ruminant are provided by microbial synthesis in the rumen and from dietary protein that is not degraded in the rumen but is intestinally digestible (bypass protein). Enhancement of the portion which leaves the rumen can lead to improvements in the quantity and quality of nutrients available to the ruminant animal.

1.1.3.1 Ammonia assimilation in rumen bacteria. Depending upon the diet, 60-90% of the daily nitrogen intake by the ruminant is converted to ammonia, and 50-70% of bacterial nitrogen can be derived from ammonia (Mathison and Milligan 1970; Nolan and Leng 1972; Pilgrim et al. 1970). Bryant and Robinson (1962) found that 92% of ruminal bacterial isolates could utilize ammonia as the main source of nitrogen, while it was essential for growth of 25% of all isolates tested. The latter is consistent with conditions found in the ruminal micro-environment where little amino acid nitrogen is available for microbial growth due to rapid breakdown of amino acids to NH_3, CO_2, and volatile fatty acids. (Mackie and Kistner 1985).

Since ammonia is the preferred source of nitrogen for most rumen bacteria, enzymes for its assimilation are essential to the growth of most rumen organisms. Glutamate dehydrogenase (GDH) and the dual enzyme system glutamine synthetase (GS) and glutamate synthase (GOGAT) are the two most important routes by which ammonia may be assimilated (Hespell 1984). GDH activity was detected in nine species of rumen bacteria (Joyner and Baldwin 1966), GS activity has been detected in extracts of mixed rumen bacteria, but GOGAT has only been studied in *S. ruminantium* (Smith et al. 1980; 1981).

A considerable amount of nitrogen is potentially recycled between the rumen ammonia pool and micro-organisms (Fig. 1.1). Sources entering the ammonia pool include components soluble in the rumen liquid phase (peptides and amino acids), influx of nitrogen into the rumen via saliva, protozoan excretion of NH_3, the turnover of microbial protein, and endogenously secreted protein (Leng and Nolan 1984). Ammonia is irreversibly lost from the rumen fluid by incorporation into

microbial cells that pass out of the rumen, by absorption through the rumen epithelium, and in fluid passing out of the rumen.

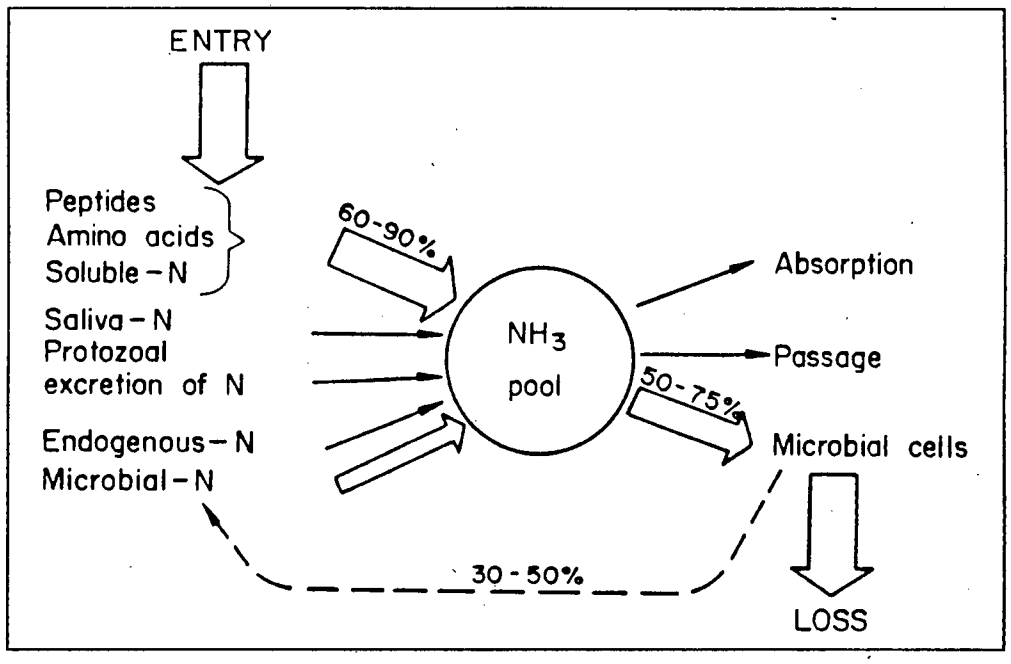


Fig. 1.1 Cycling of nitrogen through the ammonia pool of the rumen (after Leng and Nolan 1984).

1.1.3.2 Protein degradation in the rumen. The extent to which dietary proteins are degraded in the rumen and the proportion which escapes degradation and is subsequently hydrolyzed and absorbed in the lower digestive tract are recognized as important factors in the nutrition of ruminants (Fulghum and Moore 1963; Leng and Nolan 1984). Proteolysis in the rumen results in losses of high quality dietary proteins that would otherwise be directly digested and absorbed in the small intestine of the ruminant animal (Black and Tribe 1973). As a consequence, there is considerable interest in suppressing microbial catabolism of proteins, peptides, and amino acids in the rumen (Leng and Nolan 1984). Peptides seem to be particularly significant in this process and accumulate in ruminal fluid after feeding (Broderick et al. 1988; Chen et al. 1987). On the other hand, proteolysis during ruminal fermentation may benefit the host animal if the microbial protein synthesized is of higher biological value than the feed proteins (Fulgum and Moore 1963; Tamminga 1979). So, although much attention has been focussed on physical and chemical methods of controlling the rate of protein degradation in the rumen (Chalupa 1977;

Nugent et al. 1983; Tagari et al. 1962), the most direct and perhaps most effective means of decreasing the degradation of protein within the rumen may be through the selective inhibition of microbial proteases (Brock et al. 1982). Therefore, knowledge of the overall proteolytic activity in the rumen, as well as characterization of the proteolytic enzymes produced by individual organisms are of the utmost importance.

Protein degradation in the rumen is a composite of several microbial processes, including protein hydrolysis, peptide degradation, amino acid deamination, and the fermentation of amino acid carbon skeletons (Cotta and Hespell 1986). Much of the information on these microbial activities has been obtained from measurements of ruminal contents directly or from mixed bacterial populations obtained from the rumen. These studies have shown that several types of proteases are present in the rumen and that most of the proteolytic activity towards soluble proteins is associated with the bacterial cells (Brock et al. 1982). Cell-free rumen fluid and protozoa have been shown to contribute little proteolytic activity towards soluble proteins in the rumen (Kopecny and Wallace 1982; Nugent and Mangan 1981). However, the protozoa play an important role in the engulfment of bacteria and particulate matter and hence, in the degradation of insoluble proteins (Coleman 1980). The bacterial proteases in the rumen have been tentatively classified as serine, cysteine, and metalloproteases (Brock et al. 1982). A common problem in identifying the most important of a number of species of bacteria performing the same biochemical function in an ecosystem as complex as the rumen, is to decide whether bacterial numbers, specific activity, or some other criterion is of greatest significance (Wallace and Brammall 1985). Furthermore, the predominant proteolytic bacteria in the rumen will differ depending on diet (Brock et al. 1982). A combination of high specific activity and large numbers has been taken into account to implicate *B. amylophilus* (Blackburn and Hobson 1962), *B. ruminicola* (Blackburn and Hobson 1962; Hazlewood and Nugent 1978), and *B. fibrisolvens* (Fulghum and Moore 1963; Hazlewood et al. 1983) as the most important bacteria in the hydrolysis of protein in the rumen. It has been proposed by Kopecny and Wallace (1982) that rumen proteolytic activity may be attributable to Gram-negative bacteria alone, but Russell et al. (1981) identified the Gram-positive *Streptococcus bovis* as an important

proteolytic organism, and Hazlewood et al. (1983) found that appreciable numbers of proteolytic Gram-positive bacteria could be isolated from the rumen contents of cows fed either a hay-concentrates diet or fresh forage.

Proteases have been studied in only a few rumen bacteria. *B. ruminicola* R8/4 does not utilize ammonia as a source of nitrogen but assimilates nitrogen derived from peptides and protein added to the growth medium (Hazlewood and Nugent 1978). *B. ruminicola* subsp.. *ruminicola* strain 23 has been shown to require peptides of at least two and probably as many as five amino acid residues before the peptide nitrogen can be utilized (Pittman and Bryant 1964). However, Munn et al. (1983) found that at least some of the polypeptide products were taken up as individual amino acids. They showed that the presence of intact protein, presumably undergoing proteolysis, appeared to mediate the uptake of free amino acids. *B. ruminicola* R8/4, therefore, depends on the activity of its proteases to generate the peptide nitrogen required for growth (Hazlewood et al. 1981; Hazlewood and Nugent 1978). The proteolytic activity of *B. ruminicola* in batch culture was maximal and mostly (>90%) cell-associated during the mid-exponential growth phase (Hazlewood and Edwards 1981; Hazlewood et al. 1981). The proteolytic activity comprised a mixture of serine, cysteine, and aspartic proteases with the possibility that some of the activity is dependent on the presence of metal ions (Hazlewood and Edwards 1981).

Hullah and Blackburn (1971) found that peptides in the medium contributed to 17% of the protein synthesized by *B. amylophilus*. However, its obligate requirement for ammonia was not replaced by other forms of nitrogen. *B. amylophilus* H18 produces both cell-bound and cell-free proteolytic activity which consistently amounts to 80 and 20%, respectively, of the total activity during the exponential growth phase (Blackburn 1968a). The proportion of cell-free activity increases in stationary phase cultures (Lesk and Blackburn 1971) probably due to cell lysis. The protease activity of *B. amylophilus* was shown to contain one or two serine proteases with trypsin-like activity (Blackburn 1968b).

In contrast to the mostly cell-bound proteases of *B. ruminicola* and *B. amylophilus*, Cotta and Hespell (1986) found that the proteolytic activity of *B. fibrisolvens* was essentially extracellular. Characterization of the proteases produced by *B. fibrisolvens* H17c is discussed in Chapter 6.

1.1.4 Fibre digestion in the rumen.

Ruminant animals rely for their survival on microbial fermentation of plant structural carbohydrates. The animal itself cannot produce enzymes for the digestion of cellulose, hemicellulose (xylan) or pectin. The rumen microorganisms provide these enzymes, ferment the plant tissues eaten by the animal and generate volatile fatty acids and microbial cells which provide the bulk of the nutrients to the animal (Hungate 1966).

1.1.4.1 Lignocellulose degradation in the rumen. Lignocellulose is a complex material consisting of three main components which vary in proportion depending on the source. Cellulose accounts for 40-60%, hemicellulose 15-30%, and lignin 10-30% (Dekker and Lindner 1979). Lignin is the most recalcitrant of the three components, and its presence in association with the plant cell walls make it one of the most important factors limiting lignocellulose degradation in the rumen (Van Soest 1973). The lignin surrounds the cellulose fibril bundles and in mature tissues results in an encrustation believed to limit the access of cellulase enzymes (Wood 1985). Other limiting factors are the crystallinity of the cellulose which may limit the rate of digestion (Wood 1985), and access to the cellulose substrate may be physically limited by other cell wall components such as xylan, which may shield the cellulose from attack (Orpin 1988). The type of cell wall or structure of forage fragments will influence the digestion rate. Mesophyll cell walls are digested rapidly, followed by walls of phloem and epidermis, while the least digestible are the xylem vessels. Mechanical breakage of forage material by rumination is extremely important in exposing cell walls to bacterial attack in the rumen (Latham et al. 1978a; 1978b). Rumination was found to be responsible for 50-75% of the reduction in particle size of forage material in the rumen (Ulyat et al. 1985).

1.1.4.2 Microbiology of lignocellulose degradation in the rumen. Cellulolytic and xylanolytic rumen microorganisms (Hungate 1966; Bauchop and Clarke 1977) act together to hydrolyze plant cell wall polymers. In recent years there has been considerable interest in the part played by the protozoa and fungi in digestion of fibre in the rumen. Detailed work by Coleman (1978) showed that the ciliate protozoa support cellulolysis. Anaerobic fungi were first isolated from the rumen by Orpin (1975; 1977), and the rumen fungus *Neocallimastix frontalis* was shown to produce an extracellular cell-free cellulase system which contained endoglucanase and β-glucosidase activity and was able to degrade crystalline cellulose to a remarkable degree (Mountford and Asher 1983; Wood et al. 1986; Wood et al. 1988). It is, however, difficult to assess the importance of fungi in fibre digestion in the rumen as it is not possible at present to measure the size of the population present in the rumen (Van Gylswyk and Schwartz 1984).

The major part of fibre digestion in the rumen is carried out by the bacteria. Large fresh digesta fragments are colonized by motile butyrivibrios (which may be cellulolytic), selenomonads, and spirochetes (Latham 1980). The non-cellulolytic species are not directly important in lignocellulolysis at this stage, but may be of significance later in the removal of waste products and low molecular weight (M_r) saccharides (Orpin 1988). It was also postulated that these non-cellulolytic bacteria are attracted to the soluble carbohydrates in the cytoplasm of the plant cells (Latham 1980). However, these non-cellulolytic bacteria did not remain on the fragments (Van Gylswyk and Schwartz 1984) but moved away again before the passive adherence of the more common, but non-motile cellulolytic bacteria *Ruminococcus albus, Ruminococcus flavefaciens,* and *F. succinogenes* occurred (Groleau and Forsberg 1981; Ohmiya et al. 1985; Pettipher and Latham 1979). Using electronmicroscopy Akin (1980) has demonstrated that although many different morphological types of rumen bacteria adhere to forage fragments, coccoid bacteria are the most common. Morris and Cole (1987) found that the adhesion of *R. albus* appeared to be a prerequisite for effective cellulose degradation. However, adhesion was not necessarily followed by cellulolysis (Morris 1988). The close apposition of the adherent species to plant cell walls probably results in much of the low M_r hydrolysis products being directly available to the organism producing the

enzymes, but some products undoubtedly escape and can be utilized by other organisms which grow in close association with the adherent cellulolytic bacteria (Stanton and Canale-Parola 1980; Van Gylswyk and Schwarz 1984). Greve et al. (1978) found that cellulose degradation by *R. albus* was reduced by catabolite repression and, therefore, the removal of reaction products was essential to maintain cellulase production. The removal of fermentation products such as volatile fatty acids through the rumen wall, and lactate, succinate, and hydrogen via metabolism by other rumen bacteria is also believed to aid cellulolysis *in vivo* (Orpin 1988). It has been shown that co-cultures of cellulolytic species grown *in vitro* in the presence of species that can utilize the fermentation products increased cellulolysis (Dehority 1973; Latham and Wolin 1977; Scheifinger and Wolin 1973). The importance of these factors in controlling the rate or extent of ruminal lignocellulolysis is largely unknown and difficult to predict without more information being available concerning the kinetics of lignocellulose metabolism by mixed cultures of rumen organisms (Orpin 1988).

The hemicellulolytic bacteria in the rumen have received less attention than those digesting cellulose, although the amounts of hemicellulose and cellulose in forage are nearly equal (Van Gylswyk and Schwarz 1984). Unlike cellulose, hemicellulose does not have a homogeneous chemical composition. The predominant polymer (50% or more) is a pentosan or xylan composed of xylose with arabinose side chains. Depending on the given forage, smaller amounts of glucose, galactose, rhamnose, or glucuronic acid may be present (Dehority 1973). Overall, xylans represent a major energy source for microbial fermentation within ruminants and other forage-degrading animals (Hespell et al. 1987). Most of the cellulolytic rumen bacteria can also degrade hemicellulose (Coen and Dehority 1970; Dehority 1965; Hungate 1966; Kock and Kistner 1969; Morris and Van Gylswyk 1980), although some, like *F. succinogenes* and *R. flavefaciens* do not utilize the products of hydrolysis (Coen and Dehority 1970; Dehority 1965). Other prominent hemicellulose degrading bacteria are *B. ruminicola, R. albus,* and *B. fibrisolvens* (Coen and Dehority 1970; Hespell et al. 1987; Williams and Withers 1982).

Pectin is a major constituent of the middle lamella joining plant cells. Its breakdown is important in that it enables bacteria to penetrate between plant cells and so increase the surface of cell walls susceptible to attack (Van Gylswyk and Schwarz 1984). Most cellulolytic and hemicellulolytic bacteria can degrade pectin, but *F. succinogenes* and the ruminicocci use little or none of the uronic acids released (Gradel and Dehority 1972; Morris and Van Gylswyk 1980). Species that can both degrade and utilize the products produced include most strains of *B. fibrisolvens, B. ruminicola, Succinivibrio dextrinosolvens* and several types of spirochaetes (Van Gylswyk and Schwarz 1984; Ziolecki 1979; Ziolecki and Wojciechowicz 1980).

1.1.4.3 Important cellulolytic organisms in the rumen. *F. succinogenes* possesses an efficient cellulase system which has been extensively studied. Adherence of the bacterial cells to crystalline cellulose present in the culture medium (Forsberg et al. 1981) seems to be a prerequisite for cellulose hydrolysis to occur. Growth of *F. succinogenes* S85 in the presence of cellulose as a carbon source, showed that endoglucanase activity (approximately 70%) and some chloride-stimulated cellobiohydrolase activity were present in the extracellular culture fluid (Huang and Forsberg 1987). However, it was found that most of the cellobiohydrolase activity remained associated with the cells (Groleau and Forsberg 1981; Huang and Forsberg 1987; Huang et al. 1988). Cellodextrinase activity (capable of the successive removal of terminal cellobiose units from water-soluble cellodextrins) was located in the periplasmic space (Huang and Forsberg 1987). Two endoglucanases (EG1 and EG2) (McGavin and Forsberg 1988) and a cellobiosidase (Huang et al. 1988) were purified from *F. succinogenes* S85 culture supernatants, whereas a cellodextrinase enzyme was purified from the periplasmic fraction (Huang and Forsberg 1987). Two genes, one expressing cellodextrinase activity, and one expressing endoglucanase activity (*cel*-3) were cloned in *E. coli* and the nucleotide sequence of *cel*-3 was determined (McGavin et al. 1989).

Two other important cellulolytic species in the rumen are *R. albus* and *R. flavefaciens*. The cellulase systems of both *R. albus* and *R. flavefaciens* appear to be similar in many respects. Cellulolytic activity is constitutively expressed in both organisms and degradation of cellulose and cellobiose occur simultaneously (Hiltner and Dehority

1983; Pettipher and Latham 1979). *R. albus* produces both a high M_r complex similar to the *Clostridium thermocellum* cellulosome complex (see Chapter 1.2.2) (Stack and Cotta 1986; Stack and Hungate 1984) and a low M_r cellulase activity, with cellobiose being the main cellulose hydrolysis product in both cases (Wood et al. 1982). The cellulase system of *R. flavefaciens* contains endoglucanase (Pettipher and Latham 1979), exoglucanase (Gardner et al. 1987), and cellodextrinase (Rasmussen et al. 1988) activities but no β-glucosidase activity has been observed. Cellobiose and cellotriose are end products of the cellulase complex produced by *R. flavefaciens* (Rasmussen et al. 1988). Cellobiose is taken up by the cell and undergoes a phosphorolytic cleavage by cellobiose phosphorylase to yield glucose-1-phosphate and glucose, both of which are metabolized (Ayers 1958). The cloning of a number of cellulase genes in *E. coli* from both *Ruminococcus* species would be very helpful in identifying the different enzymes involved in cellulose degradation produced by these two organisms.

1.2 CELLULOSE DEGRADATION

1.2.1 The cellulose substrate.

The primary structure of cellulose consists of a linear polymer of up to 14 000 anhydroglucose units linked by β-1,4 linkages (Coughlan 1985). The average degree of polymerization (DP) in naturally occurring cellulose is 10 000, but can be as low as 15 (Eveleigh 1987). Each glucose unit is rotated 180 degrees with respect to its neighboring residues forming the basic repeating unit of cellobiose. The glucose polymers are oriented in parallel and arranged in a staggered fashion to form insoluble elementary fibrils in which the glucose chains are held together by hydrogen bonds (Alberts et al. 1983). The fibers contain a large proportion of highly compact, crystalline domains, which are separated by more amorphous regions. Inside plant cell walls, the fibers are embedded in a matrix composed of hemicellulose and lignin.

1.2.2 Cellulase systems.

The cellulase systems produced by cellulolytic organisms are usually complex, comprising a multiplicity of activities. Bacterial cellulases can be divided into three major classes: (1) endo–β-1,4–glucanases (endo–1,4–β–D–glucan 4-glucanohydrolase, EC 3.2.1.4), which cleave β–glucosidic bonds at random internal sites; (2) exoglucanases or exocellobiohydrolases (1,4–β–D–glucan cellobiohydrolase, EC 3.2.1.91), which remove cellobiose units from the non-reducing end of the polymer; (3) cellobiases (β–glucosidase, EC 3.2.1.21), which hydrolyze cellobiose to glucose (Coughlan 1985).

The classification of an enzyme is based on its activity on various substrates. (1) Endoglucanases hydrolyze soluble forms of cellulose such as carboxymethylcellulose (CMC) and hydroxyethylcellulose (HEC) (Gilkes et al. 1984a; Wirick 1968a; 1968b), but either fail to hydrolyze or exhibit weak activity on crystalline cellulose and other cellobiohydrolase specific substrates. A good test for endoglucanase activity is to measure the increase in specific fluidity of CMC as a function of the release of reducing sugars from the substrate (Gilligan and Reese 1954). A rapid increase in specific fluidity/unit increase in reducing sugars released, indicates randomly acting endoglucanases (Gilkes et al. submitted). (2) Cellobiohydrolases hydrolyze model substrates such as *p*–nitrophenyl–β–D–cellobioside (pNPC) and methylumbelliferyl–β–D–cellobioside (MUC) to aglycone and cellobiose (Deshpande et al. 1984; Van Tilbeurgh et al. 1982). They release cellobiose from crystalline cellulose and show weak activity against endoglucanase specific substrates. In contrast to endo–acting enzymes the exo–acting enzymes cause only a slow increase in specific fluidity. (3) β–Glucosidases release glucose from substrates such as cellobiose, *p*–nitrophenyl–β–D–glucoside (pNPG), methylumbelliferyl–β–D–glucoside (MUG), and salicin. (4) Cellodextrinases have been found in some bacteria (Huang and Forsberg 1987; Wang and Thomson 1990). They show weak activity on crystalline cellulose and are able to release cellobiose units from short cellulooligomers and pNPC.

Synergism between the three types of enzymes in fungal cellulases was explained by proposing that the endoglucanases attack the amorphous regions of cellulose fibers, thereby creating sites for the cellobiohydrolases which would then proceed to attack the crystalline regions of the fiber (Béguin 1990; Coughlan and Ljungdahl 1988; Wood et al. 1988). The β–glucosidases would perform the last step to prevent the build up of cellobiose, which inhibits the action of the cellobiohydrolases. The need for two cellobiohydrolases (CBHI and CBHII) in the *Penicillium pinophilum* cellulase system was explained by Wood et al. (1988) as being connected in some way with the stereochemistry of the cellulose chains and the fact that cellobiose is the repeating unit (Fig. 1.2).

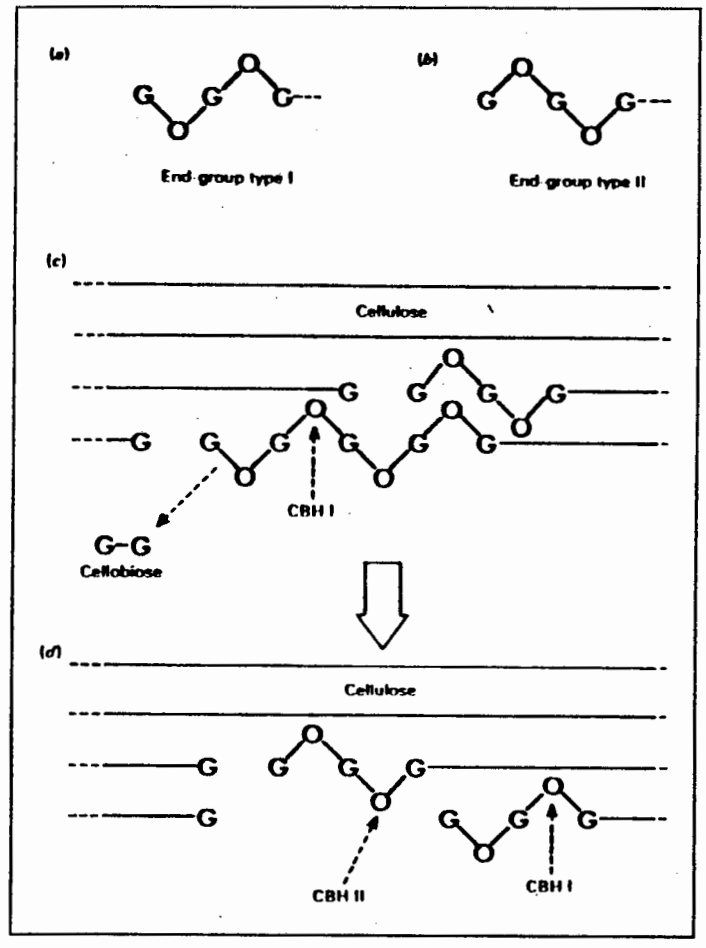


Fig. 1.2 Possible explanation of the synergism between cellobiohydrolases (CBH) I and II of *P. pinophilum* in solubilizing crystalline cellulose. Cellobiose is the repeating unit in cellulose, and theoretically there will be two types of non-reducing end groups (*a* and *b*) in the cellulose crystallite. These end groups will be held in position by non-covalent bonds and will require two different stereospecific cellobiohydrolases for hydrolysis. (*c*) shows CBHI action, and (*d*) shows CBHII attacking the new chain end exposed by CBHI action (from Wood et al. 1988)

Experimental evidence using the CBHI from *Trichoderma koningii* (Wood and McCrae 1972) and *Fusarium solani* (Wood and McCrae 1977) in synergism with the CBHII from *P. pinophilum* supported the hypothesis that the two CBH's attack two different stereospecific cellobiose units (Wood and McCrae 1986). However, the model is most likely an oversimplification, and the observations are still under discussion (Béguin 1990).

There seem to be two distinct patterns for cellulose degradation in bacteria (Béguin et al. 1988a). In the first system it appears as if "fungal-like" Procaryotes such as Actinomycetes and the related Corynebacteria (*Cellulomonas*) can degrade cellulose according to a mechanism similar to fungi, where the major part of the cellulase system is found in the medium and not bound to the cells (Béguin et al. 1988a; Coughlan and Ljungdahl 1988; Yablonsky et al. 1988). The second system occurs in many anaerobic bacteria like *C. thermocellum* and some of the rumen bacteria such as *F. succinogenes* and *R. albus*. The various cellulase components are found in tightly associated multimolecular complexes (cellulosomes) which are often quite stable and have high specific activity towards crystalline forms of cellulose such as Avicel and cotton (Johnson et al. 1982a). The entire cell surface is covered with a multitude of polycellulosomal protuberant organelles. Upon contact with the insoluble cellulose substrate they undergo extensive structural changes, such that the cellulosome clusters are able to attach to the surface of the cellulose substrate and the fibrous material can then structurally connect the cellulose-bound cellulosomes to the cell surface (Lamed and Bayer 1988). The cellulosomes are originally associated with the surface of the cells, where they mediate binding to the substrate, but after extended periods of growth on cellulose they are released into the growth medium (Forsberg et al 1981; Lamed and Bayer 1988; Wood et al. 1982). The best known example is the cellulosome complex of *C. thermocellum* (Lamed et al. 1983a; 1983b; Lamed and Bayer 1988). The *C. thermocellum* cellulosome comprises 14-18 different polypeptides which form a very stable extracellular structure, 18nm in diameter with a M_r of approximately $2x10^6$ kDa (Lamed et al. 1983b). The cellulosome shows high activity towards crystalline cellulose in the presence of Ca^{2+} and thiol-reducing agents (Johnson et al. 1982a) and is strongly inhibited in the presence of cellobiose (Johnson et al. 1982b). A cellulosome component (S_L) which

is devoid of CMCase activity and with an apparent M_r of 250 000 behaves as a binding factor to bind the cellulosome to the cellulose (Wu et al. 1988). The only evidence of the presence of cellobiohydrolase activity in the cellulosome is circumstantial (Béguin et al. 1988a). *C. thermocellum* is able to degrade crystalline cellulose very efficiently with cellobiose as the main product, but no cellobiohydrolase activity has been characterized. Mayer et al. (1987) proposed a model in which cellulose would be attacked simultaneously by regularly spaced catalytic subunits lining up along the cellulose molecules (Fig. 1.3).

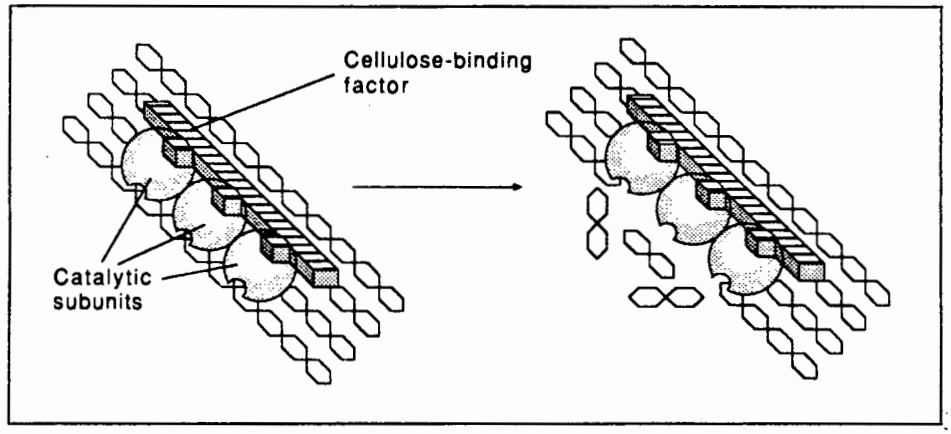


Fig. 1.3 Hypothetical model for the hydrolysis of cellulose by the cellulosome of *C. thermocellum*. Binding and alignment of the catalytic subunits at regular intervals along the cellulose fiber is assumed to be mediated by a cellulose-binding factor carrying reiterated anchoring sites for the enzymes. Simultaneous hydrolysis at multiple sites generates small fragments which no longer bind to the fiber lattice (from Mayer et al. 1987)

Electron micrographs of cellulosome preparations showing subunits aligned along a central axis are in agreement with this model (Mayer et al. 1987). It is interesting to note that Lamed et al. (1987) demonstrated by scanning electron microscopy (SEM) that the cell surfaces of a number of cellulolytic bacteria grown on cellobiose as substrate were covered with protruberant structures. It was shown that even a *Cellulomonas* sp. (enzymes mostly extracellular) had surface protruberances when grown on cellobiose but not when grown on glucose. These results seemed to indicate that the polycellulosomal surface organelles may be a general characteristic of cellulolytic bacteria, and it was concluded that the critical form of cellulase activity may not be the free floating cellulase enzymes, but rather the

surface-associated cellulase containing complexes which mediate contact with the insoluble substrate (Lamed et al. 1987).

Two factors which should not be neglected as sources of cellulase diversity are the glycosylation and proteolysis of cellulases. . Some bacterial cellulases are glycoproteins, once considered to be a rare event in Procaryotes (Gilkes et al. submitted). Bacteria producing at least some glycosylated cellulases are *Pseudomonas fluorescens* var. *cellulosa* (Yamane et al. 1970; Yamane and Suzuki 1988), *Cellulomonas* sp. (Béguin and Eisen 1978; Langsford et al. 1984), *C. thermocellum* (Gerwig et al. 1989; Ng and Zeikus 1981), and *F. succinogenes* (Huang et al. 1988). Glycoproteins may be subject to deglycosylation in the extracellular medium, and this may account for some of the changes observed in the cellulase profile during growth of *C. fimi* (Langsford et al. 1984). The function of glycosyl groups is still unclear, but they have been shown to protect two *C. fimi* cellulases against protease attack in the native host (Langsford et al. 1987b). Many bacterial cellulases retain activity after truncation by proteolysis (McGavin and Forsberg 1988; Owolabi et al. 1988) and this may also account for part of the variation in cellulase profiles observed during the growth of *C. fimi* and *Thermomonospora curvata* cultures (Langsford et al. 1984; Lupo and Stutzenberg 1988).

Rapid development in the molecular biology of cellulase genes and their products has opened new fields of investigation such as the organization of cellulase genes, their regulation at the molecular level, and the study of structural features required for enzyme activity (reviewed by Béguin 1990).

1.2.3. Cellulase synthesis and gene regulation in the native host.

There seem to be two mechanisms controlling synthesis and secretion of cellulases (Béguin 1990). In some systems the presence of high concentrations of an easily metabolizable carbon source like cellobiose will repress the synthesis of cellulases in some organisms such as *C. thermocellum* (Johnson et al. 1985) and *P. fluorescens* (Yamane et al. 1970) whereas in several other systems the synthesis of cellulases is

induced by the presence of cellobiose or sophorose (Coughlan 1985; Stewart and Leatherwood 1976).

Evidence indicates that cellulase synthesis is regulated at the level of mRNA transcription. Transcription analysis was used to investigate the molecular mechanisms which govern the expression of the *cenA*, *cenB*, *cenC*, and *cex* cellulase genes in *C. fimi* (Greenberg et al. 1987a; 1987b; Moser et al. 1989). The results showed that the carbon source provided during growth influences the levels of the gene-specific transcripts. All four *C. fimi* cellulase gene transcripts are induced in the presence of CMC and repressed at different levels in the presence of glycerol and glucose, with the *cex* gene being the most stringently regulated (reviewed by Miller et al. 1988). The *cenA* gene is transcribed from two overlapping promoters *cenAp1* and *cenAp2*, of which *cenAp1* is the stronger (Greenberg et al. 1987a). Both promoters appear to be induced in the presence of CMC, function weakly in the presence of glycerol, and are repressed in the presence of glucose. The *cex* and *cenC* genes are both transcribed from a single CMC inducible promoter which is repressed in the presence of glycerol and glucose (Greenberg et al. 1987a; Moser et al. 1989). Transcription of *cenB* is initiated from two tandem promoters: *cenBp1* and *cenBp2*. The *cenBp1* promoter is regulated by the carbon source, and the *cenBp2* promoter functions constitutively (Greenberg et al. 1987b).

T. reesei CBHI synthesis was also shown to be regulated at the level of mRNA transcription (El-Gogari et al. 1989) and the *T. reesei cbh1* mRNA was shown to initiate from multiple start sites as commonly found in eucaryotes (Teeri et al. 1987)

1.2.4 Cellulases from cloned genes.

Systems containing cloned cellulase genes have not yet provided practical alternatives to native cellulolytic organisms, if only for the lack of a complete set of cellulolytic enzymes. However, the complexities of bacterial cellulase systems and the difficulties in fractionating aggregates such as the cellulosome of *C. thermocellum* prompted the use of gene cloning as a method of analysis (Béguin et al. 1987). Clones often represent a useful source of cellulase components devoid of

contamination by other gene products (Béguin et al. 1988b; Schwarz et al. 1987; 1988), thereby allowing examination of their catalytic properties in crude extracts. Most of the cloning has initially been done in *E. coli* even though *E. coli* secretes few proteins (Béguin 1988b). However, a cloned gene can be transferred to a host of choice, motivated by factors such as the host's ability to secrete proteins into the extracellular medium, or because of closer evolutionary links which may allow for more efficient expression of the gene (Ghangas and Wilson 1987; 1988). The nucleotide sequence of a gene can be determined and the deduced amino acid sequence used to predict the two- and three-dimensional structures of the enzyme and to identify important domains. The level of expression of the gene can be increased dramatically by in-frame fusions between the gene of interest and a strong host promoter. The fusion of the *celD* gene of *C. thermocellum* to the *lacZ* promoter of pUC8 led to the over production of the enzyme in *E. coli* and subsequent crystallization of the enzyme was achieved (Millet et al. 1985; Joliff et al. 1986a; 1986b).

1.2.5 The structure and function of cellulases.

Nucleotide sequence analysis of cloned cellulase genes has revealed significant differences in the primary structure of the enzymes, and alignment of amino acids from enzymes with similar activity showed little overall homology. Henrissat et al. (1989) used hydrophobic cluster analysis (HCA) with great success and were able to classify more than 50 sequences of fungal and bacterial cellulases into six major families according to similarities in the catalytic domains of these enzymes (Table 1.1) (reviewed by Béguin 1990).

1.2.5.1 Catalytic domains. The removal of certain parts of a cellulase gene without loss of cellulase activity has been shown to be the case for many cellulases. The removal of 233 base pairs (bp) from the 3' end of the *C. thermocellum D* gene (Chauvaux et al. 1990) and 292 bp from the 3' end of the *C. thermocellum E* (Hall et al. 1988) gene without loss of enzyme activity, are two examples of catalytic "cores" which behave independently. Gilkes et al. (1988; 1989) showed that an extracellular *C. fimi* protease cleaved both *C. fimi* Cex and CenA enzymes cloned in *E. coli* in a

very specific manner without loss of activity against soluble substrates. The catalytic domains of the Cex and CenA enzymes can also function independently, and it was shown that a fusion protein between the two catalytic domains produced a bifunctional enzyme able to degrade endo- and exo- type substrates (Warren et al. 1987). The identification of similar catalytic domains in cellulase enzymes was investigated with HCA (Henrissat et al. 1989). This proved to be a powerful method as this approach can clearly detect similarities in the three dimensional folding of proteins of low sequence identity (eg. 10%). In addition HCA is effective in finding homologous domains which are separated by variable segments of widely differing sizes (Henrissat et al. 1988).

For HCA a sequence representation is derived from a two dimensional (2D) pattern for secondary structure predictions. The amino acid sequence, in one letter code, is drawn on a classical α-helical net (3.6 residues per turn) and duplicated (Gaboriaud et al. 1987). Clusters composed of adjacent hydrophobic amino acids not separated by prolines, are circled. The comparisons are based on the 2D topology and distribution of the hydrophobic clusters along the sequences. Therefore, the comparisons are not based essentially on the maximization of sequence identities, but on the successive correspondence of 2D structuring elements (hydrophobic clusters). An alignment of amino acid sequences is satisfactory when correspondence of successive clusters is respected. When hydrophilic segments situated between clusters are not conserved or differ in length, they are considered to correspond to loops and they do not dramatically influence protein folding. The amino acids considered as hydrophobic are: V,I,L,W,F,M, and Y. In an hydrophobic environment A and C are also sometimes considered as hydrophobic (Henrissat 1989).

Table 1.1 Families of homologous catalytic domains in cellulases and xylanases

Family	Sub family	Enzyme	Organism	Mode of action
A	A1	EGB	*Clostridium thermocellum*	EG
		ORFA	*Cellulomona flavigena*	EG
		EGB[1]	*Caldocellum saccharolyticum*	EG
	A2	EG	*Bacillus subtilis*	EG
		EGA	*Bacillus* sp. strain N-4	EG
		EGB	*Bacillus* sp. strain N-4	EG
		EGC	*Bacillus* sp. strain N-4	EG
		EGF	*Bacillus* sp. 1139	EG
		EG3A	*Bacillus circulans*	EG
		EG	*Clostridium acetobutylicum*	EG
		EGZ	*Erwinia chrysanthemi*	EG
	A3	EGC	*Clostridium thermocellum*	EG/LIC
		EG3	*Fibrobacter succinogenes* S85	EG/LIC
	A4	EGE	*Clostridium thermocellum*	EG/XYL
		EGH	*Clostridium thermocellum*	EG/XYL
		EGA	*Clostridium cellulolyticum*	EG/XYL
		EG	*Butyrivibrio fibrisolvens*	EG/XYL
		EGI	*Ruminococcus albus*	EG
		EG2	*Bacillus circulans*	EG
	A5	EGIII	*Trichoderma reesei*	EG
		EGI	*Schizophyllum commune*	EG
B	B1	EGA	*Cellulomonas fimi*	EG
		EG	*Streptomyces* sp. (KSM-9)	EG
		EGA	*Microbispora bispora*	EG
	B2	CBHII	*Trichoderma reesei*	CEL
C		CBHI	*Trichoderma reesei*	CEL
		CBHII	*Phanerochaete chrysospoium*	CEL
		EGI	*Trichoderma reesei*	EG
D	D1	EGA	*Clostridium thermocellum*	EG
	D2	EG	*Cellulomonas uda*	EG
		EGY	*Erwinia chrysanthemi*	EG

Table 1.1 continued

Family	Sub family	Enzyme	Organism	Mode of action
E	E1	CED	*Butyrivibrio fibrisolvens*	CED
		EGD	*Clostridium thermocellum*	EG
		EGA	*Pseudomonas fluorescens*	EG
	E2	EG	*Persia americana*	EG
		EGF	*Clostridium thermocellum*	EG
F		EXG	*Cellulomonas fimi*	EXO/XYL
		XYNZ	*Clostridium thermocellum*	XYL
		XYNA	*Bacillus* sp.	XYL
		XYN	*Cryptococcus albidus*	XYL
		EGB[2]	*Caldocellum saccharolyticum*	EXO
		XYNA	*Pseudomonas fluorescens*	XYL
BG	BGA	BGL	*Agrobacterium* sp. ATCC 21400	GLU
		BGL	*Caldocellum saccharolyticum*	GLU
		BGLA	*Clostridium thermocellum*	GLU
	BGB	BGL	*Kluyveromyces fragilis*	GLU
		BGL	*Candida pelliculosa*	GLU
		BGL	*Aspergillus wentii*	GLU
		BGL	*Schizophyllum commune*	GLU
		BGLB	*Clostridium thermocellum*	GLU

[1]COOH-terminal domain of protein
[2]NH_2-terminal domain of protein
Abbreviations: EG, endoglucanase; CEL, cellobiohydrolase; XYL, xylanase; EXO, exoglucanase; CED, cellodextrinase; GLU, β-glucosidase (from Béguin 1990)

Family A is the largest group and is subdivided into five subfamilies. There are considerable differences between members of subfamilies, however, alignment of family A enzymes revealed that they all share small conserved regions along the sequence. Apart from sharing similar patterns in the hydrophobic regions, certain residues were found to be identical in most or all of the sequences. The identification of conserved residues may provide valuable information towards identification of putative active sites in cellulase enzymes. If the amino acid sequence of *C. thermocellum* EGB (including the signal peptide) (Grépinet and Béguin 1986) is used as a point of reference then the following amino acids are found to be conserved in all family A enzymes: Arg_{98}, His_{155}, Asn_{203}, Glu_{204}, and

Glu_{363}. Pro_{205} is conserved in all enzymes except *C. thermocellum* EGC, His_{311} is conserved in all enzymes except *S. commune* EGI, and Pro_{100} is conserved in all enzymes except those of subfamily A1 (reviewed by Béguin 1990; Henrissat et al. 1989). All the members of family A described to date are endoglucanases, although they do differ in their specificity towards different substrates (Béguin 1990).

The three related endoglucanases in subfamily B1 are *C. fimi* CenA (Wong et al. 1986), *Streptomyces* sp. (KSM-9) EG (Nakai et al. 1988), and *M. bispora* EGA (Yablonsky et al. 1989). The *T. reesei* CBHII (Teeri et al 1987) cellobiohydrolase is the only member of subfamily B2 and shares more distant, but still significant homology with the three endoglucanases of B1 (Béguin 1990).

Family C contains three cellulases of fungal origin. Homology has previously been shown between the cellobiohydrolase CBHI and endoglucanase EGI of *T. reesei* (Pentillä et al. 1986). The cellobiohydrolase CBHI from *P. chrysosporium* was found to be highly similar (approximately 70% identity) to the CBHI from *T. reesei* (Simms et al. 1988).

C. thermocellum EGA (Béguin et al. 1985) is the only member of subfamily D1. The *C. uda* EG and *E. chrysanthemi* EGY endoglucanases have been shown to be more closely related (Béguin 1990) and are members of subfamily D2.

The *C. thermocellum* EGD (Joliff et al. 1986a) and *P. fluorescens* var. *cellulosa* EGA (Hall and Gilbert 1988) are members of subfamily E1. A number of hydrophobic amino acid clusters with similar shapes, sizes and orientations were found throughout both sequences (Henrissat et al. 1989). Subfamily E2 contains another *C. thermocellum* endoglucanase EGF, which is more closely related to a cellulase EG found in the ripening of the avocado fruit *P. americana* (Béguin 1990).

The members of family F are mostly xylanases. It is, however, interesting to note that the *C. fimi* EXG exoglucanase is also a member of this family. It has previously been shown that this enzyme has apart from exoglucanase activity also high activity towards xylan (Gilkes et al. 1984).

The different β-glucosidases share similarities and are classified as members of the BG family which is divided into two subfamilies BGA and BGB (Béguin 1990).

Only three cellulolytic enzymes could not be classified among any of the above families (Béguin 1990). The great variety of cellulase cores are therefore derived from a limited number of basic patterns. It is to be expected that the active sites should share some degree of similarity as all cellulases hydrolyze the same bond. It is generally accepted that enzymatic hydrolysis of glycosidic bonds proceeds through general acid catalysis, usually promoted by aspartic or glutamic acid residues. This leads to an intermediate carbonium ion, which is stabilized by a negatively charged group (aspartate, glutamate) or a histidine residue (Zvelebil and Sternberg 1988). Since active and/or substrate binding sites are often highly conserved, Henrissat et al. (1989) have listed the conserved Asp, Glu, and His residues in the homologous regions. This led to useful information in the case of family A, but since the other families contain only a few members each, uncertainty remains in identifying potential catalytic residues (Henrissat et al. 1989). Several of the major families contain enzymes with different substrate specificities (Table 1.1). Therefore, assuming that such enzymes share similar overall structure and active residues, this suggests that the variety of enzyme specificities required for the synergistic hydrolysis of cellulose may be due to subtle changes within similar sequences and not necessarily due to totally different enzymes (Béguin 1990).

1.2.5.2 Non-catalytic domains. In most cellulases the catalytic domain is coupled to a non-catalytic domain by a short segment (10-30 residues) highly enriched in hydroxyl amino acids and proline, which has been termed a hinge (Knowles et al. 1987). Béguin (1990) identified four types of conserved, non-catalytic domains amongst the cloned cellulases and xylanases (Table 1.2). It is interesting to note that the non-catalytic domains of cellulases and xylanases from one species all belong to one domain type, unlike the catalytic domains where the enzymes from one species belong to different families. This may indicate that the non-catalytic domains were added to the catalytic domain at a relatively late stage during evolution (Béguin 1990).

The *C. fimi* type non-catalytic domain was identified as a substrate binding domain and has been well characterized. Both *C. fimi* cellulases, CenA and Cex, bind strongly to cellulose (Gilkes et al. 1984; Langsford et al. 1984). Comparison of the predicted amino acid sequences (O'Neill et al. 1986a; Wong et al. 1986) showed that both enzymes contain a conserved Pro-Thr (PT) box and a non-catalytic binding domain of approximately 100 amino acid residues (Miller et al. 1988). The Cex binding domain is located at the COOH terminus and the CenA binding domain is positioned at the NH_2 terminus (Warren et al. 1987a). The substrate binding function of the *C. fimi* cellulase domains was established by proteolysis experiments. While bound to cellulose, the unglycosylated versions of CenA and Cex produced by recombinant *E. coli* clones were cleaved by a *C. fimi* extracellular serine protease at the COOH end of the PT-box, and this prevented binding of the cellulase to the cellulose substrate. However, the glycosylated counterparts (isolated from *C. fimi* culture medium) were not cleaved by the *C. fimi* protease (Langsford et al. 1987b). The generated catalytic domains retained activity on soluble substrates, but showed reduced binding to cellulose (Gilkes et al. 1988). Furthermore, distinct fragments devoid of cellulase activity but capable of binding independently to cellulose have been observed. These binding domains offer interesting prospects for the construction of fusion proteins. The fusion of the CenA binding domain to the alkaline phosphatase protein (Greenwood et al. 1989) and the Cex binding domain to a β-glucosidase enzyme of an *Agrobacterium* sp. (Ong et al. 1989), illustrated the use of the binding domains to immobilize an enzyme on an inexpensive inert support, such as cellulose, or to use them for affinity purification of a protein.

The *T. reesei* cellulases EGI, EGII, CBHI, and CBHII (Knowles et al. 1988; Paice et al. 1984) displayed no overall similarity, either with one another or with enzymes from other organisms. However, each of the enzymes contained a similar non-catalytic domain coupled to the catalytic domain with a PT-box, and these non-catalytic domains were also identified as binding domains (Van Tilbeurgh et al. 1986; Tomme et al. 1988). The *T. reesei* cellulases CBHI and CBHII displayed bifunctional organization closely resembling that of the *C. fimi* CenA and Cex enzymes (Knowles et al. 1987). Removal of these binding domains led to reduced hydrolysis of Avicel

and this reflected the critical role played by the binding domain in the hydrolysis of crystalline cellulose (Van Tilbeurgh et al. 1986). Thus, even though the *C. fimi* and the *T. reesei* enzymes do not resemble one another at the sequence level, they seem to have undergone convergent evolution towards the same structural organization with a substrate binding domain which is distinct from the catalytic domain (Béguin et al. 1988a).

The *C. thermocellum* EGA, EGB, EGD, EGE, EGH, and XYNZ enzymes all contain a highly conserved polypeptide of 65-69 residues. This region is characterized by the presence of two segments of 24 amino acids showing a high degree of similarity (Béguin et al. 1988b). The conserved region is located at the COOH terminal of the EGA, EGB, EGD, and EGH, enzymes, and in the middle of the EGE and XYNZ enzymes (Table 1.2). The presence of a duplicated element led to the hypothesis that a reiterated binding site was needed for the binding of two adjacent subunits in the substrate (Béguin et al. 1985). Alternatively, the two segments could serve as anchorage domains to some scaffolding protein in the cellulosome (possibly the cellulose binding factor S_L) (Béguin et al. 1988a). This model will account for the fact that the only *C. thermocellum* cellulase which does not contain the conserved segments, EGC, is not associated with the cellulosome (Béguin 1990). The first 12 residues of the conserved segments bear significant resemblance to the Ca^{2+} binding sites (Kretsinger 1980) of various Ca^{2+} binding proteins, and this led to the question whether these segments are perhaps involved in Ca^{2+} binding. However, Chauvaux et al. (1990) showed that the removal of the conserved segments did not affect the binding of *C. thermocellum* EGD to Ca^{2+}.

Similar *B. subtilis* type non-catalytic domains were found at the COOH terminus of the *B. subtilis* EG (MacKay et al. 1986), the *B. circulans* EG1 (ref. in Béguin 1990), and between the two catalytic domains of the *C. saccharolyticum* EGB (Saul et al. 1989) enzymes. However, the function of this domain has not been elucidated. Proteolytic processing of the non-catalytic domain was observed for the *B. subtilis* EG and the *B. circulans* EG1 enzymes in both *E. coli* and *B. subtilis* as hosts. However, no processing occurred with the *C. saccharolyticum* EGB enzyme (Saul et al. 1989). The fact that the non-catalytic domain was processed in the native host

argued against a role in cellulose degradation (Béguin 1990). Removal of the non-catalytic domain from the *B. circulans* EG1 (ref. in Béguin 1990) enzyme led to a marked reduction in exported EG1 enzyme from a *B. subtilis* host, suggesting that this domain may be involved in secretion of the cellulase enzyme (Table 1.2).

Table 1.2. Conserved, non-catalytic domains of cellulase and xylanase enzymes.[a]

Domain type	Function	Organism	enzyme family	Position in the protein
T.reesei	Substrate binding	*T.reesei* CBHI	C	COOH end
		T.reesei CBHII	B	NH_2 end
		*T.reesei*EGI	C	COOH end
		*T.reesei*EGIII	A	NH_2 end
		P.chrysosporium CBHI	C	COOH end
C.fimi	Substrate binding	*C.fimi* EXO	F	COOH end
		C.fimi EGA	B	NH_2 end
		M.bispora EGA	B	NH_2 end
		B.fibrisolvens EG	A	COOH end
		P.fluorescens EGA	E	COOH end
		P.fluorescens EGB	?	NH_2 end
		P.fluorescens XYNA	F	COOH end
C.thermocellum	Substrate binding ?	*C.thermocellum* EGA	D	COOH end
		C.thermocellum EGB	A	COOH end
		C.thermocellum EGD	E	COOH end
		C.thermocellum EGE	A	Middle
		C.thermocellum EGH	A	COOH end
	Binding to cellulosome ?	*C.thermocellum* XYNH	F	Middle
		C.cellulolyticum EGA	A	COOH end
B.subtilis	Secretion ?	*B.subtilis* EG	A	COOH end
		B.circulans EGI	?	COOH end
		C.saccharolyticum EGB	A/F	Middle

[a]After Béguin (1990)

The recent developments in the molecular biology of cellulases have shown what a powerful tool recombinant DNA technology has provided for the investigation of structure, function, and regulation of cellulolytic enzymes. A complete understanding of the degradation of cellulose still faces many problems, but the prospects are likely to improve, particularly when considering the rate at which knowledge has accumulated in the past few years (Béguin et al. 1988b). However,

the mechanism of cellulose degradation in the rumen is still unclear and control of this system via genetic manipulation remains a major challenge.

1.3 GENETIC MANIPULATION OF RUMEN BACTERIA

Both the potential and the limitations involved in the genetic manipulation of rumen bacteria are great. To date, the major genetic engineering accomplishments have been associated with industrial microbiology where the engineered organisms are mostly grown under carefully controlled conditions with antibiotics as selective agents. However, in natural environments like the rumen, energy sources are usually limiting and competition is very intense. Factors like substrate affinity, maintenance energy expenditure, resistance to toxic substances, attachment to solid surfaces and the ability to tolerate periods of nutrient starvation can be of critical importance to the survival of organisms in the rumen (Russell and Wilson 1988). These physiological factors are of little significance in industrial fermentations where the energy source is provided in excess and potential competitors are excluded.

Two major constraints for the establishment of a microbial species in the rumen are the short turnover time in the rumen and competition with other organisms (Patterson 1989). The microflora of the rumen, through a process of intense selection, have adapted very well to their environment. Therefore, new genetically modified non-rumen organisms introduced to the rumen may have little chance of surviving. Enumeration studies have shown that non-rumen bacteria such as *E. coli* do not compete well in the rumen microbial ecosystem (Russell and Wilson 1988) and native rumen bacteria are therefore the organisms of choice for genetic manipulation of this system.

The *in vivo* transfer of genetic material to rumen bacteria has been very limited. Teather (1985) obtained a few transformants from conjugative transfer of the broad host range plasmid RP4 from *E. coli* to a mutant *B. fibrisolvens* strain resistant to the antibiotic streptomycin. *B. fibrisolvens* sphaeroplasts prepared by growth in the presence of the antibiotic vancomycin were exposed to plasmid DNA in the

presence of PEG and some transformants were obtained, but the procedure was extremely inefficient due to poor regeneration of the protoplasts, and the transformants were unstable (Teather 1985; Hazlewood and Teather 1988). For genetic engineering to be successful in these bacteria, suitable vectors, hosts, markers, and transfer systems are needed. The most promising vectors for ruminal bacteria are endogenous plasmids. Native plasmid DNA has been isolated from the following rumen species: *B. fibrisolvens* (Mann et al. 1986; Teather 1982), *R. albus* (reference in Béguin 1990), *B. ruminicola* (Flint and Stewart 1987), and *Selenomonas ruminantium* (Dean et al. 1989; Martin and Dean 1989). A small cryptic plasmid, pOM1 (2.8 kb), was isolated from *B. fibrisolvens* (Mann et al. 1986) and could be very useful in the construction of a shuttle vector between *E. coli* and *B. fibrisolvens*. Flint et al. (1988) showed that tetracycline resistance in *B. ruminicola* 223/M2/7 was associated with a 19.5 kb plasmid. Ohmiya et al. (reference in Béguin 1990) made an important observation concerning cellulose degradation by *R. albus* F-40. Changes in the plasmid profile of this organism were correlated with adaptation of this bacterium on glucose, cellobiose, or ball-milled cellulose. Cellulose adapted cells contained two plasmids pRAB (7.4 kb) and pRAC (15.3 kb). Cellobiose adapted cells contained only pRAC, but were able to re-adapt to cellulose, with subsequent recovery of pRAB. Cells adapted on glucose did not harbor either pRAB or pRAC, but instead, contained a third plasmid pRAG (7.7 kb). The mechanism of conversion is still unclear, however, the results suggest that pRAB carries some key function(s) for the expression of the *R. albus* cellulase system (Ohmiya reviewed in Béguin 1990).

Manipulation of rumen organisms is at an early stage, and the development of genetically engineered microorganisms to use in the improvement of efficiency of animal production remains a major challenge.

1.4 IMPORTANCE OF *B. FIBRISOLVENS* IN THE RUMEN

The genus *Butyrivibrio* is composed of gram-negative anaerobic bacteria that ferment a wide variety of carbohydrates with the production of large amounts of butyric acid. Members of this genus occur widely in nature and are among the most

numerous bacteria isolated from the rumen of animals fed on a variety of rations (Roché et al. 1973). They have been isolated from the rumens of cattle, sheep, Zebu cattle, goats, Alaskan reindeer, Svalbard reindeer, and bison, as well as from faecal material obtained from rabbits, horses, and humans (Brown and Moore 1960; Bryant and Small 1956; Dehority 1966; Dehority and Grubb 1976; Margherita and Hungate 1963; Orpin et al. 1985; Varel and Dehority 1989). Phase microscopy showed *Butyrivibrio* cells to be short, curved rods that are between 1.5 and 6.0 μm in length and 0.4 and 0.8 μm in diameter. Short chains of two to three cells are occasionally seen, and the cells are generally motile in wet mounts (Cheng and Costerton 1977). Motility is by means of monotrichous flagellation, with the flagellum attached terminally or, less frequently, subterminally. All *Butyrivibrio* strains were classified as a single species, *B. fibrisolvens* (Buchanan and Gibbons 1974). However, Moore et al. (1976) described a new species *Butyrivibrio crossatus*, which has lophotrichous flagella, produces no gas, and ferments only a few carbohydrates.

An interesting feature of *Butyrivibrio* is its cell envelope structure. Bacteria of this genus are universally described as gram-negative, but Hewett et al. (1976) and Sharpe et al. (1975) isolated lipoteichoic acid and glycerol teichoic acid (compounds typical of gram-positive cell walls) from *B. fibrisolvens*. Electron microscopy of sections of *B. fibrisolvens* cells showed a gram-positive type cell wall that does not contain the trilamellar membrane profile typical of gram-negative bacteria (Costerton 1970; Cheng and Costerton 1977). The only characteristic of these cell walls that may explain the definite gram-negative staining reaction of the cells, is their extreme thinness (12-18 nm) as compared with that of most gram-positive bacteria (30-50 nm) (Cheng and Costerton 1977). Staining with ruthenium red showed the presence of knob-like structures at the surface of *B. fibrisolvens* cells (Cheng and Costerton 1977). Knobs were also observed in natural populations of rumen bacteria attached to forage cell walls (Akin 1976), and it has been proposed that these knobs are involved in adhesion of cells to cellulose fibers, and cell-to-cell adhesion (Cheng and Costerton 1975)

The different *B. fibrisolvens* strains show great phenotypic variability and are capable of fermenting a wide variety of carbohydrates. This heterogeneity has prompted

several studies which have attempted to classify isolates into groups based on one or more phenotypic characteristic(s). Shane et al. (1969) divided cellulolytic strains into two groups based on differences observed in the patterns of short-chain volatile fatty acids produced or consumed during growth. Roché et al. (1973) tried to exploit differences in nutritional requirements, and immunochemical properties were investigated as a means to differentiate between isolates (Hazlewood et al. 1986; Margherita and Hungate 1963; Margherita et al. 1964). Most *B. fibrisolvens* strains produced extracellular polysaccharides (EPS) when grown on a defined medium (Stack 1988). Analysis showed that several unusual sugars were constituents of these EPS, and the presence of different sugars was used to classify the *B. fibrisolvens* strains into five major groups. *B. fibrisolvens* strains 49 and H17c contained an additional unknown acidic sugar of the lactyl ether type and were classified as the only members of group IV type B (Stack 1988). DNA hybridization and G + C base composition were used to examine 39 *B. fibrisolvens* strains for DNA relatedness (Mannarelli 1988). *B. fibrisolvens* strains H17c and 49 were shown to be 96% related and the G + C content was 42% for both strains. The DNA hybridization results indicate that the strains presently classified in the species *B. fibrisolvens* actually comprise a number of distinct species and possibly several genera (Mannarelli 1988).

All *B. fibrisolvens* strains isolated from the rumen are saccharolytic and individual strains have been isolated which are cellulolytic (Shane et al 1969), xylanolytic (Dehority 1966), amylolytic (Cotta 1988), pectinolytic (Dehority 1969), and lipolytic (Hazlewood and Dawson 1979). In addition, a number of *B. fibrisolvens* strains have been isolated which produce extracellular proteases (Fulgham and Moore 1963).

The predominant proteolytic species in the rumen are found to be members of the major genera of saccharolytic rumen bacteria *B. amylophilus, B. ruminicola,* and *B. fibrisolvens.* However, in contrast to the former two organisms, virtually nothing was known of the proteolytic activity of *B. fibrisolvens* at the beginning of this project. Subsequently, Cotta and Hespell (1986) investigated the distribution of proteolytic activity among *B. fibrisolvens* strains, and described general properties of the proteolytic activity produced by strain 49. The majority of *B. fibrisolvens* strains investigated were proteolytic and several of these strains produced levels of activity

similar to those obtained for *B. amylophilus* or *B. ruminicola.* The proteolytic activity of *B. fibrisolvens* strain 49 was extracellular, was produced constitutively, and was characteristic of a serine protease (Cotta and Hespell 1986). However, these studies involved total activity produced by *B. fibrisolvens* and gave no indication of the number of proteases responsible for proteolytic activity of *B. fibrisolvens.*

B. fibrisolvens is the most important rumen bacterium involved in digestion of xylans (hemicelluloses) (Hespell et al. 1987). Almost all *B. fibrisolvens* strains are able to degrade xylans extensively, and most of these strains are able to utilize the breakdown products. Examination of xylanase and xylobiase activities in selected strains showed that xylanase activity was predominantly extracellular and xylobiase activity mostly associated with the cells (Hespell et al. 1987).

Only a few *B. fibrisolvens* strains have been identified as cellulolytic strains (Dehority 1966; Hobson and Wallace 1982). *R. albus, R. flavefaciens,* and *F. succinogenes* are generally regarded as the most important bacteria involved in the digestion of cellulose in the rumen, based on their numbers and the nature and activity of their cellulases. However, under certain feeding conditions, such as sheep fed low quality forage, *B. fibrisolvens* was found to play a more important role than previously thought (Kock and Kistner 1969; Shane et al. 1969). In the rumen of the high-arctic Svalbard reindeer, *B. fibrisolvens* was isolated as the major culturable cellulolytic bacterium in both summer and winter, representing 66 and 52% of the cellulolytic population, respectively (Orpin et al. 1985). Margherita and Hungate (1963) found that *B. fibrisolvens* was the most abundant cellulolytic bacterium in semi-starved zebu cattle in Kenya, and 38% of the cellulolytic bacteria isolated from the rumens of Alaskan reindeer were *B. fibrisolvens.* This suggests that *B. fibrisolvens* strains may be more resilient, and therefore more able to survive cycles of nutritional abundance and starvation of the host animal than other cellulolytic bacteria (Orpin et al. 1985). It is therefore surprising that so little attention has been given to the characterization of the enzymes produced by this important rumen organism.

B. fibrisolvens H17c is an important *Butyrivibrio* strain. It is very versatile in its ability to degrade different polysaccharides such as cellulose (Dehority 1966), xylan (Dehority 1966; Hespell et al. 1987), and starch (Cotta 1988). *B. fibrisolvens* H17c is also an important proteolytic organism of the rumen (Cotta and Hespell 1986) and was therefore, the organism of choice for this study. The aim of this study was to investigate the enzymes produced by *B. fibrisolvens* H17c involved in degradation of cellulose, xylan, and protein. This may help to gain a better understanding of the role this important organism has to play in the rumen. The few studies done to date all involved crude *B. fibrisolvens* extracts providing information about total enzyme activity, but nothing was known of the number of enzymes involved or characteristics of single enzymes. There was thus a great need for the isolation and characterization of single enzymes. DNA recombinant technology provides a powerful tool for studying single genes and their products at the molecular level. Knowledge of the genes at molecular level may lead to an understanding of gene regulation in this anaerobic rumen organism thus opening the way to manipulation of expression. Information on whether the gene structures are more similar to gram-negative or gram-positive organisms would be interesting in this organism which has a cell wall similar to gram-positive organisms but stains gram-negative. The approach taken in this study was therefore the construction of a *B. fibrisolvens* H17c genebank which could be propagated and screened in *E. coli* for genes whose products are involved in the degradation of cellulose, xylan, and protein.

Chapter 2

Construction and screening of a *Butyrivibrio fibrisolvens* H17c genomic library in *Escherichia coli*

Chapter 2

Construction and screening of a *Butyrivibrio fibrisolvens* H17c genomic library in *Escherichia coli*

2.0 Summary

A *B. fibrisolvens* H17c genomic library was constructed in *E. coli* using the plasmid pEcoR251 as a positive selection vector. Analysis of a random selection of clones showed the *B. fibrisolvens* library to contain approximately 8000 clones with inserts ranging from 4-10 kb in size. The library was screened for the expression of carboxymethylcellulase, xylanase and protease genes in *E. coli*. Two carboxymethylcellulase positive clones were isolated.

2.1 Introduction

Recombinant DNA technology has enabled researchers to manipulate genetic material from different organisms and to isolate and study individual genes. Molecular cloning essentially requires the ligation of DNA fragments from the organism of interest with DNA of a vector which is then transferred to a suitable host. The vector must be able to replicate in the host, and propagation of the DNA fragment(s) of interest should then occur. A variety of techniques are available for the construction of DNA genomic libraries with a high probability of containing the entire genome as fragments in the vector. The choice of a vector-host system in the construction of a genomic library depends on a number of factors, such as the size of the DNA fragments to be cloned, ease of selection of recombinants, and considerations of gene expression. These factors have all been extensively reviewed by Zappe, Ph.D thesis, Microbiology Dept, University of Cape Town (1988).

At the time of this study, no genes had been cloned from *B. fibrisolvens*, and nothing was known about the expression of genes from this organism in other hosts. Vector systems which function in *E. coli* have been extensively studied, and *E. coli* was chosen as a host system for the construction of a *B. fibrisolvens* H17c genomic library. The expression of foreign Gram-negative *B. fibrisolvens* genes in the Gram-negative *E. coli* host is also more likely to succeed than when using a Gram-positive host. Gram-positive mRNA appears to be more readily translated in a Gram-negative system rather than the reverse, presumably due to the less stringent constraints on the Shine-Dalgarno complementarity in Gram-negative bacteria (McLaughlin et al. 1981).

The vector used was the plasmid pEcoR251, a gift from M. Zabeau (Biotechnology Business Development, Ghent, Belgium). It is a positive selection vector containing the *E. coli Eco*RI gene under the control of the λ rightward promoter, the ampicillin resistance gene, and the pBR322 origin of replication. The *Eco*RI gene product, when expressed at high levels by the λ promoter, is lethal unless insertionally inactivated, or regulated by the plasmid pCI857^{ts}, a pACYC derivative which contains a

temperature sensitive λ repressor gene and confers kanamycin resistance (Remaut et al. 1983).

The construction of a *B. fibrisolvens* H17c genomic library and the screening for expression of CMCase, xylanase, and protease genes in *E. coli* is described in this Chapter.

2.2 Materials and Methods

2.2.1 Bacterial strains and plasmids. *B. fibrisolvens* H17c isolated by Dehority (1966) was obtained from Dr R.B. Hespell (Department of Animal Science, University of Illinois, Urbana) and maintained at -70°C on slopes of M10 medium (Appendix A). *E. coli* strains (C600, K514, and LK111) were used as recipient strains for recombinant plasmids (Appendix C). *E. coli* strains were maintained at -70°C in the presence of 15% (v/v) glycerol. A restriction map of the plasmid vector pEcoR251 is given in Appendix E.

2.2.2 Media and growth conditions. All media, buffers, and solutions not described in the text are listed in Appendix A. *B. fibrisolvens* was grown in modified M10 medium as described by Strydom et al. (1986). Samples (1 ml) of overnight modified M10 cultures were centrifuged, washed with one-quarter strength Ringer solution, resuspended in 10 ml of modified M10 medium and incubated at 37°C under stringent anaerobic conditions. *E. coli* strains were grown in Luria-Bertani (LB) medium (Maniatis et al. 1982) at 37°C. Plasmids were maintained by selection in the presence of the appropriate antibiotic.

2.2.3 Isolation of chromosomal DNA from *B. fibrisolvens* H17c. An adaptation of the method of Marmur (1961) was used for the isolation of chromosomal DNA. *B. fibrisolvens* chromosomal DNA was prepared from 500 ml overnight (16h) cultures which were harvested by centrifugation (10 000 *g* for 10 min), resuspended in 4 ml Solution A (10 mM TRIS-HCl, pH 8.0, 25% w/v sucrose) containing 5 mg/ml lysozyme, and mixed gently at 37°C for 30 min. Two milliliters of ice-cold Solution B (0.5 M EDTA, pH 8.0) were added and the mixture was kept on ice for 5 min. Four ml Solution C (10 mM TRIS-HCl, pH 7.5, 1 mM Na_2EDTA, 2% w/v SDS), containing 5 mg/ml proteinase K were added. The mixture was kept at 20°C for 10 min. CsCl (1 g/ml) and EtBr (250 µg/ml) were added, and the mixture was centrifuged at 15 000 *g* for 30 min. The refractive index of the supernatant was adjusted to 1.396, and the DNA was purified by isopycnic CsCl/EtBr ultracentrifugation. The DNA band visualized under UV light (350nm) was removed by puncturing the bottom of the Quickseal tube and allowing the fluid to drip through slowly. The EtBr was

removed by extraction (3x) with an equal volume of NaCl-saturated isopropanol (Appendix A). The DNA was precipitated from the CsCl solution by the addition of two volumes of water followed by an equal volume of isopropanol. The DNA was pelleted by centrifugation in an Eppendorf microfuge for 15 min. The pellet was resuspended in 0.5 ml of TE buffer. The DNA concentration was determined by monitoring the absorbance at 260 nm, where one absorbance unit is equivalent to 50 µg/ml DNA (Maniatis et al. 1982).

2.2.4 Preparation of vector DNA. The vector pEcoR251 was prepared by digesting 10 µg of a large scale plasmid preparation (Appendix B) with 20 units of *Bgl*II restriction endonuclease for 1 h at 37°C. A small sample of approximately 300 ng of the DNA was analyzed by agarose gel electrophoresis to confirm complete digestion.

2.2.5 Partial *Sau*3A digestion and recovery of chromosomal DNA from DEAE-cellulose paper. *B. fibrisolvens* DNA was partially digested in a two-fold dilution series with *Sau*3A endonuclease (Maniatis et al. 1982) to determine the optimum DNA - enzyme ratio to produce fragments 4 - 10 kb in size. Approximately 500 µg of DNA was digested and the reaction was terminated with the addition of DNA sample loading solution (Appendix A). The sized DNA was separated by agarose gel electrophoresis (Appendix B). After electrophoresis the stained gel was viewed under UV light (350 nm), a slit was made in the gel just below the 4 kb band, and a strip of DE 81 DEAE-cellulose paper (Whatman) was inserted according to the method by Dretzen et al. (1981). Further electrophoresis transferred the 4 - 10 kb fragments onto the DEAE-cellulose paper. The DNA was eluted by incubating the paper strip in an Eppendorf tube with 300 µl elution buffer [20 mM TRIS-HCl (pH 7.5), 1 mM EDTA, 1.5 M NaCl] for 2 h at 37°C and then centrifuging for 10 min. The elution was repeated three times. The pooled supernatant samples were extracted with equal volumes of butanol (3x) and then precipitated with an equal volume of isopropanol. The DNA was resuspended in 20 µl of TE buffer, and the DNA concentration determined.

2.2.6 Ligation of vector and insert DNA. pEcoR251 is a positive selection vector based upon insertional inactivation of the *Eco*RI gene, which codes for a lethal product. A DNA concentration of 5 pmole/ml was used to optimize recombinant formation (Zappe, Ph.D Thesis, Microbiology Dept, University of Cape Town, 1988) and an insert-vector ratio of 1:1 was used. The ligation reaction was carried out in a sterile Eppendorf tube containing DNA, 5 µl ligation buffer (10x) (Appendix A) and H_2O to 50 µl. T4 DNA ligase (0.25 units) was added and the reaction placed at 15°C for 12 - 16 h. The reaction was stopped by phenol extraction and DNA precipitation. The DNA was resuspended in 50 µl of TE buffer.

2.2.7 Transformation of *E. coli* K514 competent cells for the construction of a genebank. *E. coli* K514 competent cells were prepared as described in Appendix B. Ten transformation mixes were prepared by adding 2 µl of ligated DNA to 100 µl of competent cells in sterile Falcon tubes on ice. After 10 min on ice the cells were heat shocked for 5 min at 42°C to initiate DNA uptake. The cells were diluted with LB broth (1 ml) and incubated at 42°C for 30 min to allow expression of the ampicillin resistance (Ap^r) marker on the vector. The transformation mixes were pooled, diluted (1/20) into prewarmed LB broth containing Ap (100 µg/ml), and incubated at 37°C with vigorous aeration for 1 h to allow expression of the *Eco*RI gene (killing transformants containing religated vector). Correct functioning of the cloning vector was followed by plating transformants of control DNA, namely unrestricted pEcoR251 and pBR325 (5 ng each), before and after the 1 h incubation. pEcoR251 showed a drop in transformants of more than 50 fold while pBR325 showed an increase of one-two fold. Following the expression step the cells were collected by centrifugation (6 000 *g*), resuspended in LB broth (2 ml), and plated (100 µl/plate) onto LB agar containing Ap (100 µg/ml). Approximately 500 Ap^r colonies were pooled and recombinant pEcoR251 plasmid DNA was isolated. The DNA obtained from each pool was resuspended in 1 ml TE buffer and stored at -70°C.

2.2.8 Isolation of *B. fibrisolvens* genes expressing CMCase and xylanase activity. For the detection of CMCase positive clones, plasmid DNA from the gene library (2 µl/pool) was used to retransform *E. coli* C600 competent cells (100 µl/pool), and transformants were selected on LB agar containing Ap (100 µg/ml) and medium

viscosity CMC (1 g/l) (Sigma No. C4888, DS 0.7). Colonies were lifted off the plates with Whatman No.1 filter paper discs, and the plates were rinsed to remove any remaining cells before staining with Congo red (0.1% w/v; 15 min), followed by destaining with 1 M NaCl (Teather and Wood 1982) until zones appeared. Colonies showing CMCase activity were identified by a clear zone around and beneath the colony. For the isolation of xylanase positive clones the same method was used, except that the plates contained oat spelt xylan (1 g/l) (Sigma No. X-0376; Lot No. 14F-0421) as substrate.

2.2.9 Screening for *B. fibrisolvens* genes expressing protease activity. Plasmid DNA (2 µl/pool) was used to retransform either competent *E. coli* K514 or *E. coli* K514[pCI857^{ts}] cells. Plasmid pCI857^{ts} was maintained by selection with kanamycin (Km) (15 µg/ml). Transformants were selected for their ability to form haloes of clearing on LB agar containing skim milk (1% w/v), Ap (100 µg/ml) and Km (15 µg/ml) at 30^{o}C or 37^{o}C.

2.3 Results

2.3.1 *B. fibrisolvens* H17c genomic library. Approximately 8 000 *E. coli* K514 colonies containing *B. fibrisolvens* H17c recombinant clones were obtained from one experiment. These were divided into 16 pools containing approximately 500 colonies each. The number of recombinants required to obtain a library which would represent 99.9% of the genome of an organism, can be calculated using the following formula:

$$P = 1-(1-f)^N$$

or

$$N = \frac{\ln(1-P)}{\ln(1-f)}$$

where P is the probability that a given unique DNA sequence is present in a collection of N recombinants, and where f is the fraction of the total genome that each insert represents (Clarke and Carbon 1976). The average procaryotic genome size is 6×10^6 bp (Starr et al. 1981) and for this library the average insert size was 6 kb. For a 99.9% probability of cloning a specific fragment, the number of clones required was 4 600. A random selection of clones isolated from the genebank was tested for the presence of insert DNA, and 18 out of 20 clones showed plasmids larger than the vector.

2.3.2 Isolation of *B. fibrisolvens* genes expressing CMCase activity. *E. coli* C600 was transformed with recombinant pEcoR251 from the 16 plasmid pools, and five clones expressing CMCase activity were isolated from the 20 000 colonies screened. Plasmid DNA was isolated from the five transformants expressing CMCase activity and the plasmid origin of CMCase activity was confirmed by retransformation of *E. coli* C600. Two of the five clones proved to be very unstable and were lost after retransformation. *Pst*I restriction analysis (Fig. 2.2) indicated that pES400 and pES500 contained different DNA inserts, while pES300 seemed to be a deletion of pES400. The zones of CMC hydrolysis around *E. coli* C600[pES400] colonies were much bigger than the zones around *E. coli* C600[pES500] colonies (Fig. 2.1). CMCase activity was always associated with transformation to Ap^r. Neither *E. coli* C600 nor

E. coli C600[pBR322] used as controls gave zones on LB agar containing CMC. Molecular studies of the recombinant plasmids are described in Chapters 3 and 4, and characterization of the cloned gene products is described in Chapter 5.

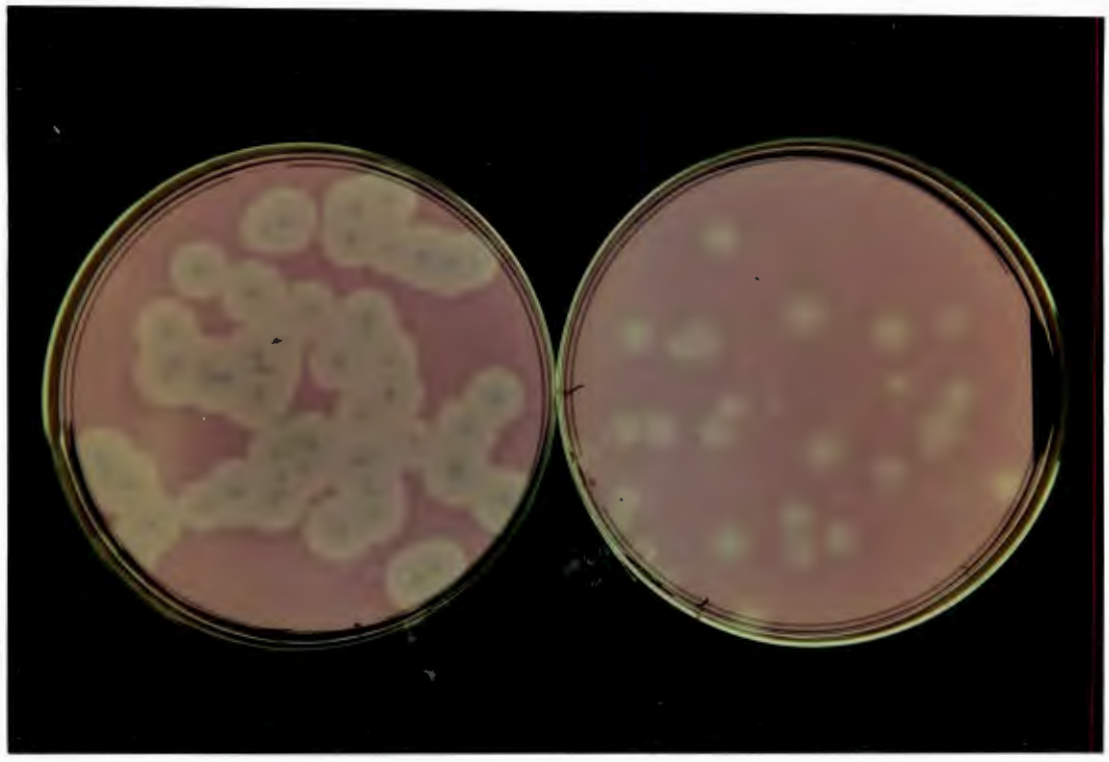


Fig. 2.1. LB agar plates, containing CMC and stained with Congo Red, showing zones of CMC hydrolysis produced by (1) *E. coli* C600[pES400] and (2) *E. coli* C600[pES500].

2.3.3 Isolation of *B. fibrisolvens* H17c genes expressing xylanase activity. *E. coli* C600 was transformed with p*Eco*R251 recombinant plasmid pools and plated onto LB agar containing oat spelt xylan. From the 10 000 colonies screened four showed zones of xylan hydrolysis. Retransformation of *E. coli* C600 by the four transformants confirmed the plasmid origin of xylanase activity. Plasmid DNA from the four clones expressing xylanase activity and the three clones expressing CMCase activity was digested with *Pst*I restriction endonuclease and analyzed by agarose gel electrophoresis (Fig. 2.2). The four clones expressing xylanase activity (pEX1, pEX2, pEX3, and pEX4) and one clone expressing CMCase activity (pES400) showed identical digestion patterns. Studies discussed in Chapter 5 will show that this clone is an endo–β–1,4–glucanase with xylanase activity.

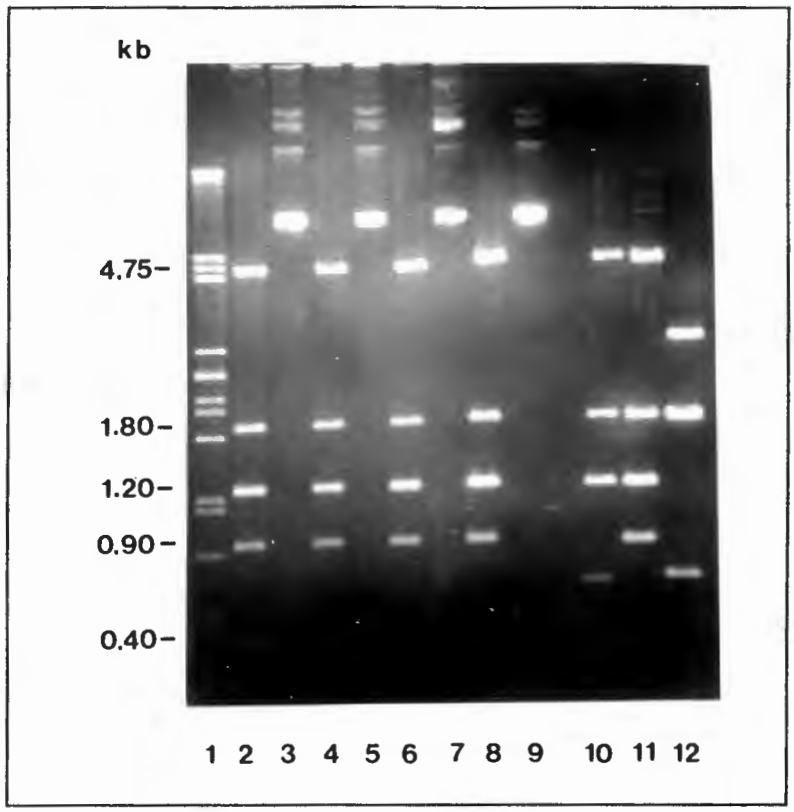


Fig. 2.2. Restriction analysis of plasmids expressing xylanase or CMCase activities. The plasmids (cut) were digested with *Pst*I restriction endonuclease. Lane 1; λDNA (cut), lane 2; pEX1 (uncut), lane 3; pEX1 (cut), lane 4; pEX2 (uncut), lane 5; pEX2 (cut), lane 6; pEX3 (uncut), lane 7; pEX3 (cut), lane 8; pEX4 (uncut), lane 9; pEX4 (cut), lane 10; pES300 (cut), lane 11; pES400 (cut), lane 12; pES500 (cut).

2.3.4 Screening for *B. fibrisolvens* H17c protease genes. *E. coli* K514 was transformed with recombinant pEcoR251 plasmid pools and plated onto LB agar plates containing Ap and skim milk. After two days of incubation at 37°C 20 colonies showed very small zones. Plasmid isolation gave low yields. Retransformation of *E. coli* C600 resulted in abnormal looking colonies which appeared like volcanoes. Lysis appeared to occur after approximately 8 h of growth. After retransformation zone formation was erratic. Some of the colonies produced no zones, while others produced small zones, and it was not possible to isolate any stable zone producing colonies. To determine whether this abnormal phenotype was regulated by the λ rightward promoter *E. coli* K514[pCI857^{ts}] was transformed with plasmids from these putative clones. Expression from the λ promoter can be

controlled by the product of the temperature sensitive λ repressor gene present on pCI857^{ts} (Remaut et al. 1983). The same abnormal colony morphology was displayed at 30^{o}C and 37^{o}C for each retransformant, and no stable genes expressing protease activity were isolated.

2.4 Discussion

Chromosomal DNA has not previously been isolated from *B. fibrisolvens*. A modification of the method by Marmur (1961) was used with great success. A *B. fibrisolvens* H17c genomic library containing 8 000 clones, with insert sizes ranging from 4-10 kb, was constructed.

Screening of the *B. fibrisolvens* genebank for CMCase and xylanase producing clones resulted in the isolation of two distinct CMCase genes from *B. fibrisolvens* in *E. coli*. Plasmids isolated from the four clones expressing xylanase activity were identical to pES400, the plasmid isolated from one of the clones expressing CMCase activity. Screening for protease expressing recombinants gave no clones. Twenty putative clones expressing protease activity turned out to be lethal or very unstable. The colonies showed diffuse haloes and appeared to lyse at an early stage of the growth cycle. Retransformation of the putative protease clones into *E. coli* K514[pCI857^{ts}] and growth at 30^{o}C did not alter the aberrant phenotypes. It would thus appear that the λ promoter was not involved in expression of these phenotypes and that the genes responsible were expressed off *B. fibrisolvens* H17c insert promoters. It was not possible to detect protease activity in the supernatant or in the cell free extract of liquid cultures of these clones. From these results it was not possible to say whether these were genes expressing proteases or not.

If the expression of *B. fibrisolvens* proteases cloned in *E. coli* from either a native- or a vector promoter is lethal, different techniques for screening the library will be necessary. A possible route to follow is the utilization of a phage vector system, such as one of the phage λ based systems (eg. λgt11 or λZAP, Short et al. 1988; Young and Davis 1983). The advantage of these systems is that multiplication of a recombinant phage is dependant upon phage directed replication within the initially transformed bacterial cell, followed by infection of a lawn of background bacterial cells to produce a plaque, while multiplication of a recombinant plasmid is solely dependent on replication within the initially transformed bacterial cell, and therefore dependent on the growth rate of that cell and its progeny. It is therefore probable that the yield of recombinant phage DNA molecules containing protease

genes and resultant active but lethal proteases could be greater within a plaque than that obtained from a colony harboring a recombinant plasmid containing a protease gene. This increase in sensitivity would enable identification of a recombinant phage clone containing a protease gene by either detection of the nucleic acid with a radiolabelled protease gene DNA probe or by detection of the protease itself.

For nucleic acid detection one could use a DNA probe derived from a heterologous gene expressing protease activity , or a homologous DNA probe constructed via "reverse genetics". This would involve purification of the *B. fibrisolvens* H17c protease enzyme, digestion of the protein into peptides, purification of a peptide for amino acid sequencing, and the design of an oligonucleotide probe (approximately 30 mer). The nucleotide sequences of five *B. fibrisolvens* H17c genes have been determined in the Microbiology Department at UCT. These are three genes expressing cellulase activity (*end1*, *ced1*, and *bglA*) (Berger et al. 1989; Berger et al. 1990; Lin et al. 1990) and two genes expressing amylase activity (*amy1* and *bgb1)* (Rumbak personal communication). These data could be used to determine *B. fibrisolvens* codon usage bias which will be of great help in the synthesis of an oligonucleotide probe.

Detection of the protease itself could be done by incorporation of skim milk in the sloppy overlay with the subsequent identification of zones of hydrolysis around the plaques, or the protease could be identified with antibodies produced from *B. fibrisolvens* protease extracts. Moses et al. (1989) raised specific antiserum to a purified 38 kDa serine protease from *Bacteroides nodosus*. The antiserum was used in a colony immunoassay to screen an *E. coli* genomic DNA library for the presence of protease clones. An *E. coli* clone expressing a 50 kDa immunoreactive polypeptide was identified. This clone showed no protease activity because the gene was truncated. This may be another route to follow to isolate lethal *B. fibrisolvens* protease genes in *E. coli*.

Chapter 3

Cloning in *Escherichia coli* of a *Butyrivibrio fibrisolvens* endo–β–1,4–glucanase (*end1*) gene: DNA sequence analysis, and comparison of the deduced amino acid sequence with other cellulases.

Chapter 3

Cloning in *Escherichia coli* of a *Butyrivibrio fibrisolvens* endo–β–1,4–glucanase (*end1*) gene: DNA sequence analysis, and comparison of the deduced amino acid sequence with other cellulases.

3.0 Summary

A *B. fibrisolvens* endo-β-1,4-glucanase gene (*end1*) was cloned on a recombinant plasmid pES400 which enabled *E. coli* C600 cells to produce clear zones on Congo Red stained CMC LB agar plates. The *end1* gene was located on a 6.8 kb DNA fragment. The nucleotide sequence of the cloned *B. fibrisolvens end1* gene was determined. The open reading frame (ORF) of this gene consisted of 1641 base pairs (bp) which encoded a protein of 547 amino acids with a calculated M_r of 61 077. The putative ribosome binding site (GGAGGT) was upstream of an ATG start codon. Three putative consensus promoter sequences were identified upstream of the ribosome binding site and a stop codon (TAA) downstream of the ORF was followed by a Ω independent terminator sequence. Tn*phoA* mutagenesis indicated the presence of a functional *E. coli*-like signal peptide. The End1 protein was divided into two distinct domains by a region rich in proline and threonine amino acids (PT-box). The catalytic domain showed homology with the endoglucanase EGE enzyme from *C. thermocellum* and the non-catalytic domain showed homology with the binding domains of the CenA endoglucanase and Cex exoglucanase enzymes from *C. fimi*.

3.1 Introduction

All organisms that are able to degrade crystalline cellulose secrete cellulase systems of varying complexity. These systems contain enzymes with different specificities and modes of action which act synergistically to hydrolyze cellulose. The study of cellulase genes and their products at the molecular level will help with the classification of these genes, the understanding of their regulation, and the study of structural features required for enzyme activity.

In recent years there has been a rapid development in the molecular biology of cellulase genes and their products, so that to date there are approximately 50 cloned fungal and bacterial cellulases of which the nucleotide sequences have been determined (Béguin 1990). The majority of genes cloned from cellulolytic organisms code for endoglucanases as they are the easiest to screen for. The largest number of cellulase-encoding genes was found in *C. thermocellum* (15 endoglucanase and two β-glucosidase genes) (Gräbnitz and Staudenbauer 1988; Hazlewood et al. 1988) while four endoglucanase genes and one β-glucosidase gene were cloned from a closely related *Clostridium* sp. (Sakka et al. 1989). Approximately 10 different endoglucanases have been cloned from a number of *R. albus* strains (reviewed by Béguin 1990). Only a few exoglucanases have been characterized in bacteria, and only two have been positively identified as cellobiohydrolases, one from *R. flavefaciens* (Gardner et al. 1987) and one from *C. fimi* (O'Neill et al. 1986). The latter has been shown to act synergistically with endoglucanases to degrade crystalline cellulose (Creuzet et al. 1983).

The majority of cloned cellulases have a signal sequence which is responsible for partial or total transport of the enzymes to the periplasm of *E. coli* (Béguin 1990). The study of protein secretion in procaryotes has been greatly facilitated by the use of gene fusions. A particularly useful system is that of alkaline phosphatase fusions (Hoffman and Wright 1985). Alkaline phosphatase (PhoA) is normally found in the periplasm of *E. coli* and its activity is absolutely dependent upon secretion from the cytoplasm. Hoffman and Wright (1985) deleted the promoter and signal peptide-encoding regions of the *phoA* gene so that expression and activation of the

enzyme requires that it be fused in the correct reading frame to DNA that contains a promoter, translational start site and a complete signal peptide encoding region. The Tn*phoA* system of Manoil and Beckwith (1985; 1986) allows the fusion of *phoA* to a target gene by making use of the random insertion properties of Tn5. Fusions of the endoglucanase End1 to alkaline phosphatase would indicate whether or not End1 could provide a functional signal sequence resulting in the export of the fusion protein in *E. coli*.

The nucleotide sequencing of an endo–β–1,4–glucanase gene (*end1*) from the *B. fibrisolvens* H17c genebank is described in this chapter. At the time of this study, no other genes from *B. fibrisolvens* had been sequenced. The deduced amino acid sequence of End1 is compared with other reported cellulase sequences, and Tn*phoA* mutagenesis proved the presence of a signal sequence functional during expression in *E. coli*.

3.2 Materials and Methods

3.2.1 Bacterial strains, plasmids, and phage. The plasmid pES400 (Chapter 2.3.2) was used as the primary source of DNA for plasmid mapping, templates for DNA shortening, and DNA deletion. The plasmid vectors pUC18 and pUC19 (Messing 1983) were used for the preparation of templates for DNA sequencing. *E. coli* strains LK111 and C600 were used as hosts for these recombinant plasmids (Appendix C). The 2.85 kb *Hin*dIII DNA fragment from pES400 (Fig. 3.1), was ligated to *Hin*dIII digested pBR325 (Appendix E) DNA. The resultant recombinant plasmid, pBn420, was used to transform the *E. coli* strain CC118 (PhoA$^-$) (Manoil and Beckwith 1985) for the Tn*phoA* fusion experiments. Phage λ::Tn*phoA* is *b*221 *c*I857 *Pam3* with Tn*phoA* in or near *rex* and was a gift from C. Manoil (Gutierrez et al. 1987).

3.2.2 Media, buffers, and enzymes. All media and buffers not described in the text are given in Appendix A. Restriction endonucleases, T4 DNA ligase and S1 nuclease were purchased from Boehringer Mannheim Biochemicals. Exonuclease III was obtained from Bethesda Research Laboratories.

3.2.3 Growth conditions. *E. coli* C600 strains containing recombinant plasmids were maintained on LB medium with CMC (0.1% w/v) and Ap (100 μg/ml). *E. coli* LK111 harboring recombinant pUC plasmids were plated on LB agar with 5–bromo–4–chloro–3–indolyl–β–D–galactopyranoside (XGal) (40 μg/ml) and Ap (100 μg/ml).

3.2.4 Restriction endonuclease mapping and Southern blot analysis. Plasmid DNA was prepared by the alkali–lysis method of Ish–Horowicz and Burke (1981). The small–scale (miniprep) and large–scale (maxiprep) methods of plasmid isolation are described in Appendix B. *B. fibrisolvens* chromosomal DNA isolation was as described in Chapter 2.2.3. Plasmid mapping was accomplished by the use of single- and double-enzyme digests. General methods for restriction endonuclease digestion are given in Appendix B.

Chromosomal DNA from *B. fibrisolvens* H17c was digested with *Stu*I, *Bgl*II-*Stu*I, and *Bst*EII-*Stu*I restriction endonucleases, and resolved by electrophoresis in an agarose gel (0.8% w/v) in Tris-acetate buffer. Samples of *Bst*EII-*Stu*I and *Bgl*II-*Stu*I endonuclease digested pES400, and *Stu*I endonuclease digested pES500 (Fig. 4.1) were resolved on the same gel. The DNA was transferred to Hybond-N hybridization membrane (Amersham) as described by Reed and Mann (1985). Plasmid pES400, nick-translated with [α-^{32}P]dCTP, was used as a hybridization probe (Rigby et al. 1977). These methods are described in Appendix B.

3.2.5 Sequencing strategy. Restriction mapping and subcloning showed that the endoglucanase gene was located on a 2.85 kb *Hind*III restriction fragment of pES400 (Fig. 3.1). This fragment was subcloned from pES400 into pUC19 to give the recombinant plasmid pES420 (Fig. 3.1). This plasmid was used to generate overlapping deletions from one orientation of the gene, using the exonuclease III shortening technique (Henikoff 1984; 1987). Subcloning using available restriction endonuclease sites was used to obtain templates for sequencing from the opposite orientation (Fig. 3.3). The whole insert of pES420 was sequenced at least once in each orientation.

3.2.6 Exonuclease III shortening of pES420 and cloning of shortened fragments. Unidirectional digestion of *Sac*I-*Xba*I restriction endonuclease digested pES420 using exonuclease III was carried out by an adaptation of the method of Henikoff (1984; 1987). After digestion, the linearised plasmid (10 µg) was precipitated with isopropanol, resuspended in Exo-buffer (100 µl) (Appendix A), and equilibrated at 37°C (5 min). Eleven microfuge tubes containing ice cold S1 nuclease mixture (25 µl) (Appendix A) were prepared immediately before starting the shortening reaction. A sample (9 µl) linearised plasmid in Exo buffer was removed to a microfuge tube containing S1 nuclease mixture before the shortening reaction was initiated by the addition of exonuclease III (300 units). Samples (9 µl) were removed at 30 s intervals and added to the 10 remaining microfuge tubes containing S1 nuclease mixture. The microfuge tubes were incubated at room temperature for 30 min while the S1 nuclease digested any single stranded DNA present. The S1 nuclease reaction was then stopped by addition of S1 stop solution (3.4 µl)

(Appendix A) to each microfuge tube and incubation at 70°C for 10 min. The extent of shortening was checked by electrophoresis of approximately 200 ng DNA (8 µl) from every second tube on an agarose gel (0.8% w/v) in Tris-acetate buffer. The exonuclease III generated ends were filled in by the addition of DNA polymerase I (Klenow) (0.5 units per tube) in Klenow buffer (Appendix A), incubation at room temperature for 3 min, followed by a further incubation of 5 min in the presence of a mixture of dNTP's (0.125 mM each, A, C, G, and T). The shortened DNA was religated by the addition of ligation mixture (120 µl) (Appendix A) to each tube, at 15°C for 4 h. Competent *E. coli* C600 cells were transformed with the ligation mixtures and selection was on LB agar plates containing Ap (100 µg/ml).

3.2.7 Nucleotide sequencing. Recombinant plasmids resulting from either exonuclease III shortening, or from the subcloning of desired restriction fragments (Fig. 3.3), were analyzed by restriction endonuclease mapping before the preparation of CsCl purified plasmid DNA for sequencing (Appendix B). The preparation of template DNA, the primer annealing and sequencing reactions, and the gel electrophoresis and autoradiography, are described in Appendix B.

3.2.8 Sequence analysis. The DNA and amino acid sequences were analyzed with a VAX computer using the GCG sequence analysis package version 6.1 (Devereux et al. 1984) and associated databases. A table of one- and three- letter codes for amino acids is given in Appendix D.

3.2.9 Isolation and sequencing of transposon (Tn*phoA*) insertions. Transposon insertions into pBn420 were obtained using an adaption of the protocol of Gutierrez et al. (1987). *E. coli* CCII8[pBn420] was grown overnight in LB containing $MgSO_4$ (10 mM), maltose (0.2% w/v), and Ap (100 µg/ml). A sample (1 ml) was mixed with phage λ::Tn*phoA* (100 µl) at a multiplicity of infection of approximately one, and incubated at 30°C for 15 min. Ten aliquots of this mixture were diluted 1:10 in LB and incubated at 30°C for 4 h to allow outgrowth of the phage. Samples (100 µl) were plated onto LB plates containing CMC (0.1% w/v), Km (250 µg/ml), Ap (100 µg/ml), and 5-bromo-4-chloro-3 indolyl-phosphate p-toluidine (XP) (40 µg/ml). After two days the dark blue colonies (alkaline phosphatase positive)

were re-streaked onto fresh plates and incubated overnight at 37°C. Plasmid DNA was prepared by the miniprep method (Appendix B) and the DNA was used to retransform *E. coli* CC118 competent cells. DNA maxipreps were prepared from pure blue (Pho^+) retransformants. This DNA was used for restriction endonuclease mapping and sequencing.

Nucleotide sequencing was utilized to determine the exact position and orientation of insertions. The method was as described in Chapter 3.2.7. A 15 bp synthetic oligonucleotide primer was a gift from I. Parker (UCT Medical School, Cape Town). The primer (5'-AAACGGCGAGCACCG-3'), was complementary to the DNA sequence corresponding to nucleotide positions 126-140 of the *phoA* gene, and was used to sequence across the junction of the fusions between the *phoA* and *end1* genes.

3.2.10 Alkaline phosphatase enzyme assay. Alkaline phosphatase activity was assayed according to Brickman and Beckwith (1975). A sample (2 ml) of an overnight culture was pelleted in an Eppendorf microfuge, and the cell pellet was resuspended in TRIS-HCl buffer (1 M, pH 8.0) to give duplicate samples of 1 ml each. The OD_{600} was recorded, the cell sample was mixed with chloroform (1 drop) and SDS (40 µl; 0.1% w/v), and equilibrated for 5 min at 37°C. The alkaline phosphatase substrate (4 mg/ml ONPP in 1 M Tris-HCl, pH 8.0) (100 µl) was added and the time taken for the sample to turn yellow was recorded. The reaction was terminated by the addition of KH_2PO_4 (1 M; 100 µl) and the samples were kept on ice until the A_{420} was measured. An arbitrary unit of alkaline phosphatase activity was calculated according to the formula of Brickman and Beckwith (1975):

$$U = \frac{OD_{420} \times \text{dilution}}{OD_{600} \times \text{time}} \times 100$$

Sample fractions collected from *E. coli* CC118[pBn420] were used as negative controls.

3.3 Results

3.3.1 Cloning of an endo-β-1,4-glucanase gene from *B. fibrisolvens* H17c. The transformation of the *B. fibrisolvens* H17c genebank into *E. coli* C600 produced colonies with zones of hydrolysis on LB agar containing CMC and Ap (Chapter 2.3.2). Each of the zone forming colonies contained a pEcoR251 recombinant plasmid and retransformation experiments showed that endoglucanase activity was always associated with transformation to Ap^r. Two clones, one producing large zones (pES400) and one producing small zones (pES500) on CMC LB agar were isolated. Preliminary restriction endonuclease mapping showed them to be different (Fig. 3.1; Fig. 4.1). The cloning, sequencing, and amino acid homology of the cloned gene product on pES400 will be discussed in this chapter, and that of pES500 in Chapter 4.

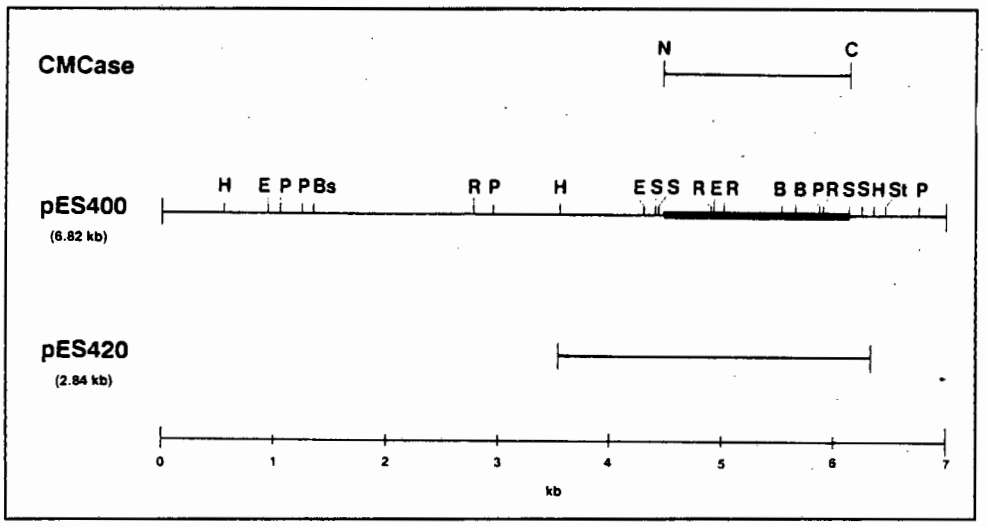


Fig. 3.1. Restriction maps of pES400 and pES420. The *end1* gene is indicated by the bold line with N and C representing the amino- and carboxy- terminal ends. The cleavage sites of restriction enzymes *Eco*RI(R), *Hin*dIII(H), *Eco*RV(E), *Bam*HI(B), *Pvu*II(P), *Stu*I(St), *Bst*EII(Bs), and *Ssp*(S) are shown.

Restriction endonuclease mapping of pES400 revealed a 6.8 kb DNA insert in pEcoR251 (Fig. 3.1). The subcloning of a 2.85 kb *Hin*dIII restriction endonuclease fragment into pUC19 (pES420) retained full endoglucanase activity (Fig. 3.1). The

same *Hind*III restriction endonuclease fragment was subcloned into pUC18, in reverse polarity to the *lac*Z promoter in pUC19 with no loss of endoglucanase activity, indicating that the endoglucanase gene was expressed from its own promoter in *E. coli*.

3.3.2 Origin of the DNA insert. The origin of the 6.8 kb DNA insert was determined by Southern blotting and DNA hybridization, using [^{32}P]-labelled pES400 as a probe.

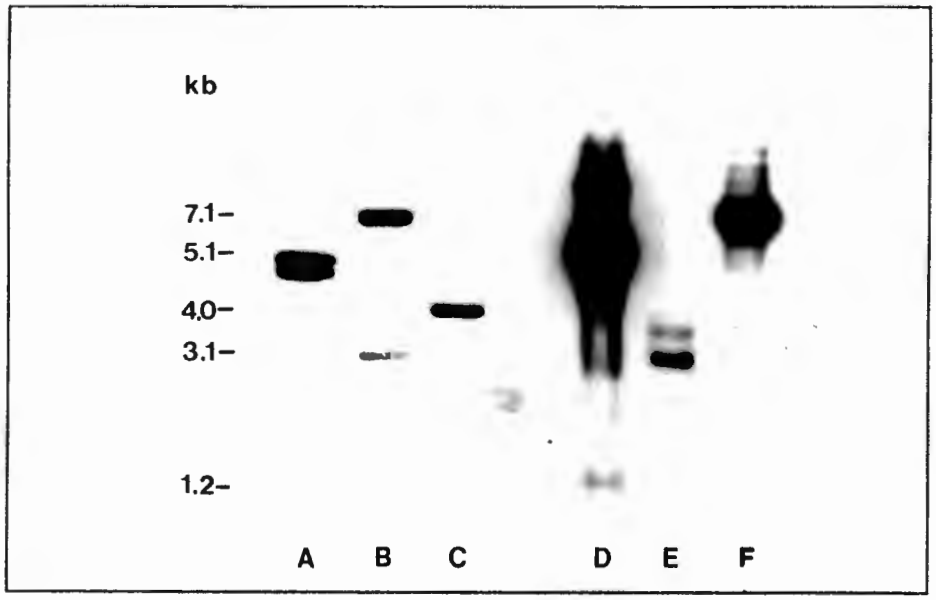


Fig. 3.2. Autoradiograph of ^{32}P labelled pES400 hybridized to restriction endonuclease digests of: lane A; pES400(*Bst*EII-*Stu*I), lane B; pES400(*Bgl*II-*Stu*I), lane C; pES500(*Stu*I), lanes D, E, and F; *B. fibrisolvens* chromosomal DNA digested with: lane D; *Bst*EII-*Stu*I, lane E; *Bgl*II-*Stu*I, and lane F; *Stu*I.

A 5.09 kb *Bst*EII-*Stu*I fragment internal to the 6.8 kb insert of pES400 hybridized to a *B. fibrisolvens* H17c fragment of the same size confirming the origin of this insert (lane D, Fig. 3.2). The two other hybridization signals at 8.3 kb and 1.2 kb in lane D represent hybridization between pES400 and the two *B. fibrisolvens* H17c chromosomal fragments which flank the 5.09 kb *Bst*EII-*Stu*I fragment. Lanes E and F contained *Bgl*II-*Stu*I, and *Stu*I endonuclease digested *B. fibrisolvens* H17c chromosomal DNA, and the hybridization signals confirmed the restriction map depicted in Fig. 3.1. Lane C contained pES500 (plasmid from clone producing small zones) digested with *Stu*I endonuclease. The pES500 insert has two internal *Stu*I

fragments (approximately 1.0 and 2.1 kb, respectively), which did not hybridize to pES400. Lanes A and B were positive controls of endonuclease digested pES400. Labelled pEcoR251 did not hybridize to *B. fibrisolvens* DNA (data not shown).

3.3.3 Preparation of clones for sequencing. The exonuclease III shortening method was successfully used to obtain overlapping shortened fragments for the sequencing of pES420 insert DNA in one direction. The other direction was covered by a range of subclones in pUC vectors. The extent of the sequence obtained from each clone is depicted in Fig. 3.3.

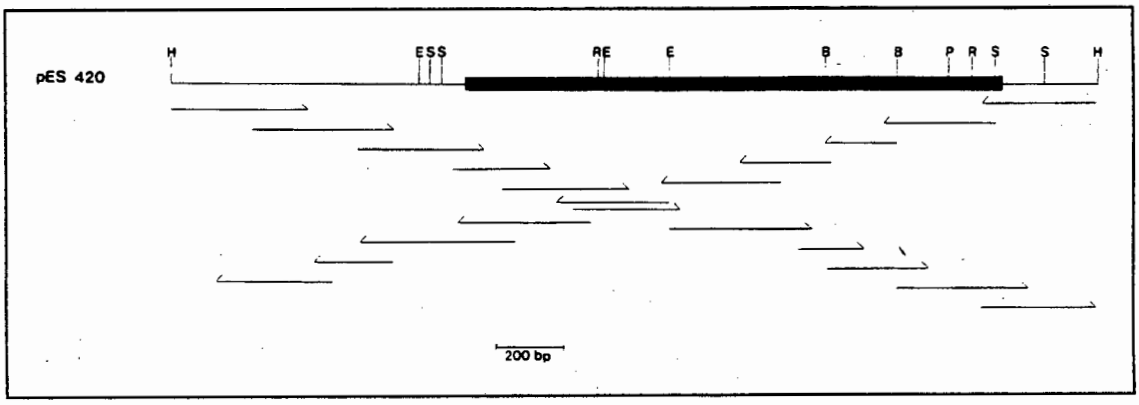


Fig. 3.3. Sequencing strategy for the *Hind*III fragment containing the *end1* gene from *B. fibrisolvens* H17c. The *end1* gene is indicated by the bold line and the thin line indicates *B. fibrisolvens* DNA flanking the gene. The arrows indicate the extent of sequence obtained from each clone. The cleavage sites of restriction enzymes *Eco*RI(R), *Hind*III(H), *Eco*RV(E), *Bam*HI(B), *Pvu*II(P), and *Ssp*(S) are shown.

3.3.4 Nucleotide sequence of the *end1* gene. The nucleotide sequence of the 2.85 kb DNA fragment coding for CMCase activity contained a single ORF which, from the presumptive start codon (ATG) to the stop codon (TAA), contained 1641 nucleotides encoding 547 amino acid residues (Fig. 3.4). The calculated M_r of the polypeptide encoded by the ORF was 61 077.

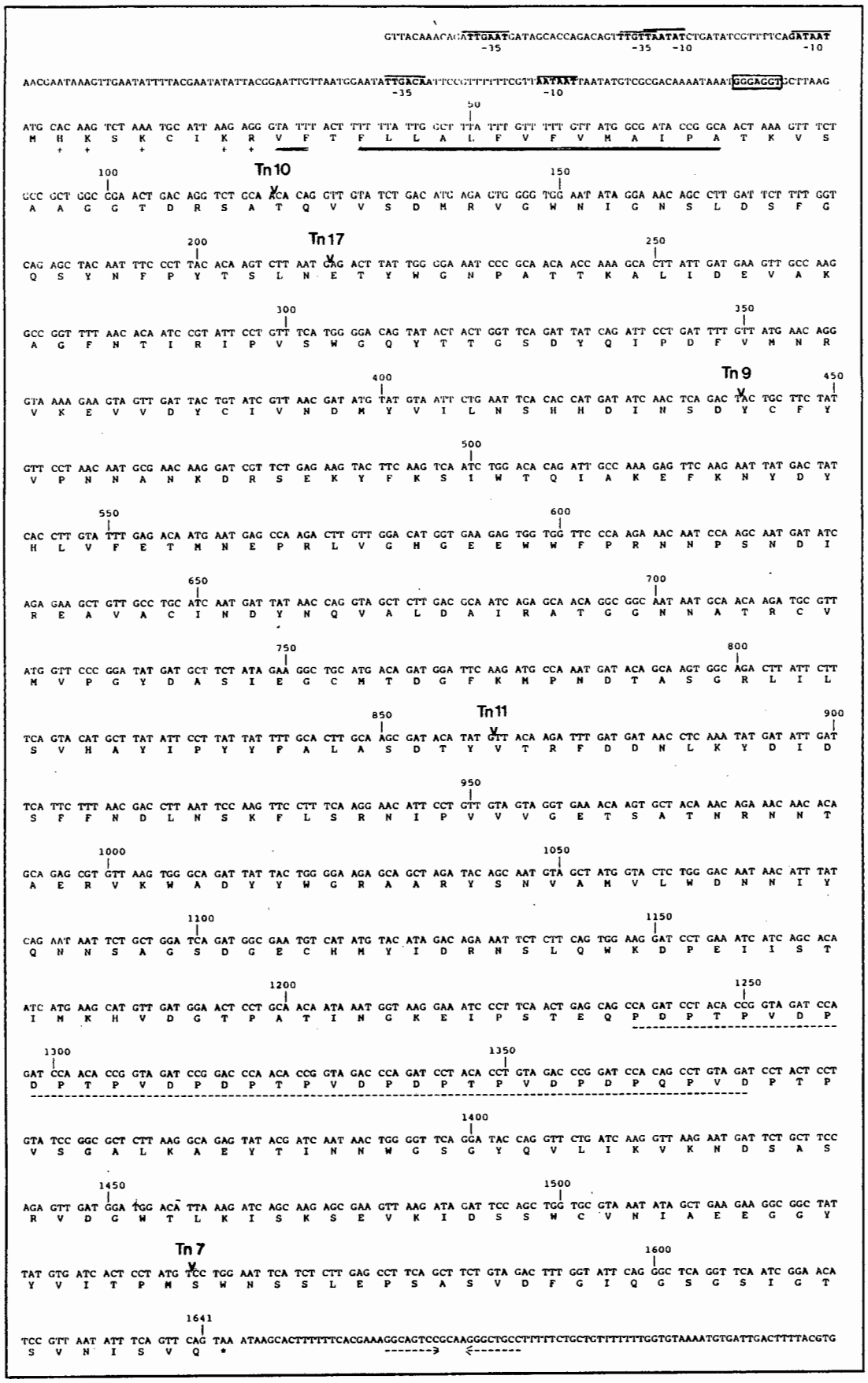


Fig. 3.4. Complete nucleotide sequence of the endo-β-1,4-glucanase (*end1*) gene from *B. fibrisolvens* H17c. The derived amino acid sequence is given in the one-letter code from position 1 to 1641 (547 residues). Three putative -10 and -35 promoter sequences are overlined. The Shine-Dalgarno (SD) sequence is boxed. The positively charged amino acids at the start of the putative signal sequence are indicated by plus signs, and the hydrophobic region is underlined twice. The five direct repeats are underlined once. The putative rho-independent terminator is indicated by arrows. The position of Tn*phoA* inserts are shown ( ).

The region upstream of the ORF contained three consensus putative promoter sequences (Rosenberg and Court 1979). A putative *E. coli*-like ribosome binding site was identified (GGAGGT) 8 bp upstream of the ATG initiation codon. The TAA stop codon was followed by a nucleotide sequence of dyad symmetry between nucleotides 1666 and 1685 (Fig. 3.4). This region could function as a rho-independent terminator sequence (Rosenberg and Court 1979), consisting of two complementary inverted repeats which can form a stem of 8 bp. The hairpin structure had a ΔG value of -18.4 kcal/mol (Salser 1977) and was followed by a stretch of 5 T's. It has been reported that the average G+C ratio for the *B. fibrisolvens* H17c genome was 42% (Mannarelli 1988). The G+C content of the *B. fibrisolvens end1* gene was 41.3% and that of the upstream regulatory region 30%.

3.3.5 Putative signal sequence and Tn*phoA* fusions. The amino terminal end of the *end1* gene product contained a typical signal peptide (MacKay et al. 1986). A short stretch of five positively charged amino acids was followed by a 16 amino acid hydrophobic domain (Fig. 3.4), and a predicted cleavage site after the S_{30}-A_{31}-A_{32} sequence, in agreement with the rules of von Heijne (1983; 1985). Tn*phoA* mutagenesis was used to prove that the signal peptide of End1 is functional in *E. coli*.

Transformation of *E. coli* CC118 cells with plasmids carrying Tn*phoA* resulted in the isolation of six blue alkaline phosphatase positive colonies. Restriction endonuclease analysis of the plasmids showed that in five of these plasmids the transposon inserted into the *end1* gene, conferring alkaline phosphatase activity, but no CMCase activity on the *E. coli* host. One recombinant plasmid conferred both CMCase and alkaline phosphatase activity, and was found to contain a Tn*phoA* insertion just downstream of the signal peptide of the Tc^r gene of the pBR325 vector portion of pBn420. The exact point of insertion within the *end1* gene was determined from nucleotide sequencing (Fig. 3.4) using the primer to the *phoA* gene. Alkaline phosphatase activity was determined from overnight *E. coli* CC118 cells containing the recombinant plasmids. All the clones with Tn*phoA* insertions in the

correct frame and orientation expressed alkaline phosphatase activity, and there was a decrease in activity the further downstream the insert was from the signal sequence (Table 3.1).

Table 3.1. Levels of alkaline phosphatase activity produced by *E. coli* CC118 carrying fusion plasmids.

[plasmid]	Alkaline phosphatase activity
pBn420	-
pTnϕ10	187
pTnϕ17	125
pTnϕ9	98
pTnϕ11	80
pTnϕ7	36

3.3.6 Amino acid sequence alignments. Amino acid alignment studies showed homology between the *C. thermocellum celE* (Hall et al. 1988) and *B. fibrisolvens* H17c *end1* (Berger et al. 1989) gene products. Analysis of the aligned amino acid sequences showed homology over approximately 400 amino acids. Identity over this area was 37.1% and when amino acid replacement by conserved amino acids was taken into account, the similarity of this region was increased to 60.2% (Fig. 3.5). Homology studies with the *Bacillus*-like endoglucanases (Zappe et al. 1988) showed 27% identity and 39% similarity over approximately 195 amino acids. *C. fimi cenA* (Wong et al. 1986) and *cex* (O'Neill et al. 1986a) gene products showed less than 10% identity over the full length of the *B. fibrisolvens* End1 protein.

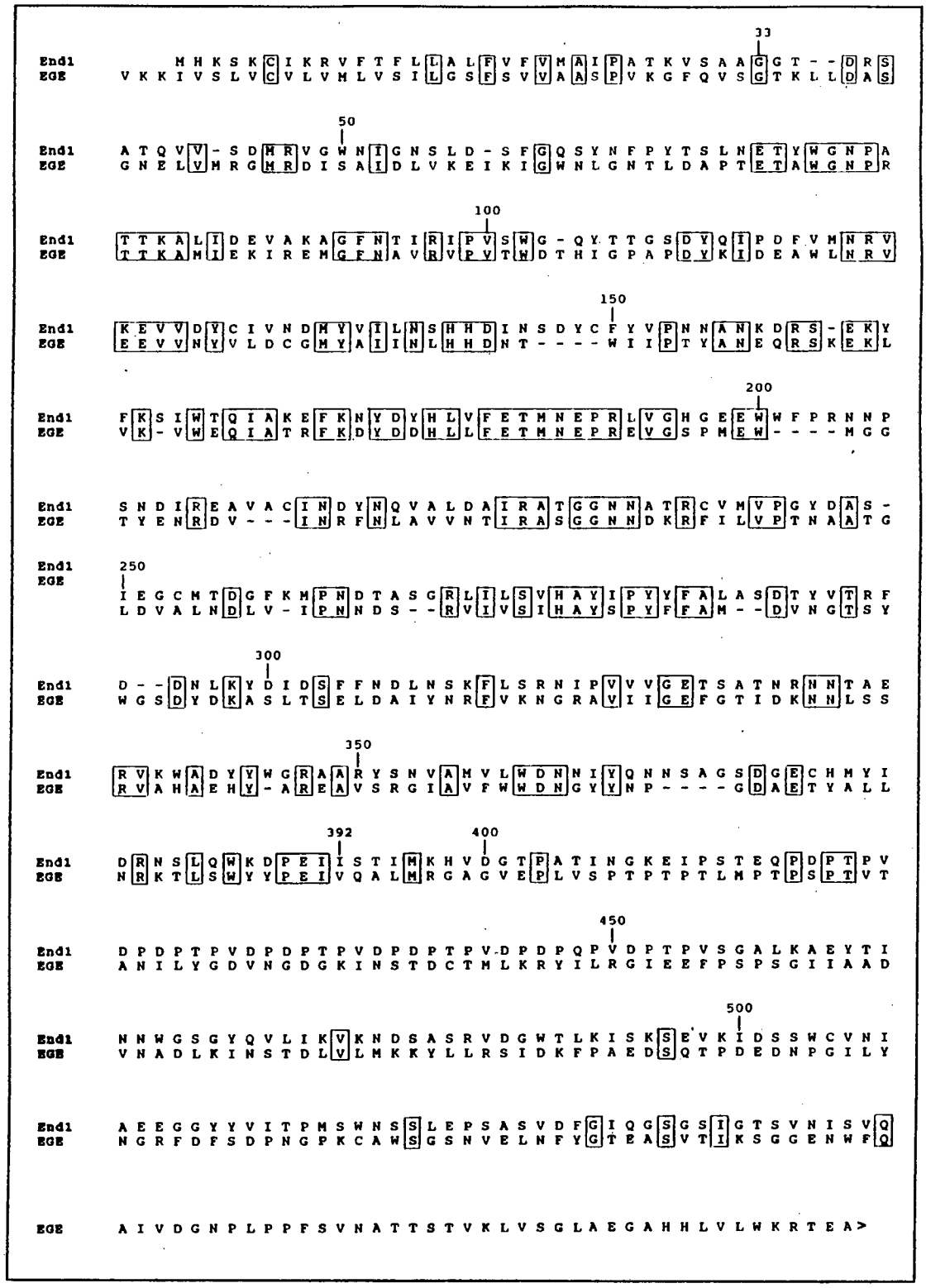


Fig. 3.5. Amino acid sequence alignment between the *B. fibrisolvens* H17c End1 and *C. thermocellum* EGE endoglucanases (Hall et al. 1988). The amino acids are identified by the single-letter code and regions containing identical amino acids are boxed.

An interesting structural feature of the End1 enzyme is the five perfect repeats (P-D-P-T-P-V-D) between amino acids 412-447. The *C. fimi cenA* gene product has 22 amino acids which are either T or P between amino acids 143-165 (Wong et al. 1986), while the *C. fimi cex* gene product has a P-T-P-T-P-T-T repeat (O'Neill et al. 1986a) which showed approximately 50% similarity with the five direct repeats P-D-P-T-P-V-D of the *B. fibrisolvens* End1 endoglucanase. Warren et al. (1987a) has shown that the primary structures of the *C. fimi* CenA and Cex proteins showed striking similarities in their overall protein architecture. A 112 amino acid amino-terminal region of CenA and a 108 amino acid carboxy-terminal region of Cex showed 50.5% identity and 63.5% similarity (Fig. 3.6).

```
End1   < S G - - A L K A E Y T I - N N W G S G Y Q V   470
CenA   < Q A A P G C R V D Y A V T N Q W P G G F G A    65
Cex    < S G P A G C Q V L W G V - N Q W N T G F T A   422

End1     L I K V K N D S A S R V D G W T L K I S K S   492
CenA     N V T I T N L - G D P V S S W K L D W T Y T    86
Cex      N V T V K N T S S A P V D G W T L T F S F P   444

End1     E V - K I D S S W C V N I A E E G G Y Y V I   513
CenA     A G Q R I Q Q L W N G T A S T N G G Q V S V   108
Cex      S G Q Q V T Q A W S S T V T Q S G S A V T V   466

End1     T P M S W N S S L E P S - A S V D F G I Q G   534
CenA     T S L P W N G S I - P T G G T A S F G F N G   129
Cex      R N A P W N G S I - P A G G T A Q F G F N G   487

End1   S - - G S I G T S V N I S V Q *                 547
CenA   S W A G S N P T P A S F S L N G T T C T G >     150
Cex    S H T G T N A A P T A F S L N G T P C T V >     508
```

Fig. 3.6. Amino acid sequence alignment between the carboxy-terminal region of *B. fibrisolvens* H17c End1, the amino-terminal region of *C. fimi* CenA (Wong et al. 1986), and the carboxy-terminal region of *C. fimi* Cex (O'Neill et al. 1986a). The amino acids are identified by the single-letter code and regions containing similar amino acids are boxed.

In both enzymes this conserved region was separated from a larger, non conserved region by a P-T box. It has been proposed that the non conserved regions comprised catalytic domains (Langsford et al. 1987), and the conserved region a binding domain (Gilkes et al. 1988). The P-T box at the carboxy-terminal of *B. fibrisolvens* End1 divided the protein into 2 distinct domains, a large domain homologous to *C. thermocellum* EGE, and a small domain of 95 amino acids. Alignment of this small domain (amino acids 452-547) of *B. fibrisolvens* End1 with the conserved domains of the CenA and Cex proteins resulted in 25.8 and 27.1% identity and 51.6 and 52.1% similarity, respectively (Fig. 3.6).

3.4 Discussion

Hydrophobic cluster analysis (HCA) was employed to classify the *B. fibrisolvens* H17c End1 endoglucanase as a member of the subfamily A4 (Béguin 1990). Family A is the largest family and is divided into 5 subfamilies which include cellulases from Gram-positive and Gram-negative, aerobic and anaerobic bacteria, as well as from the fungus *T. reesei*. Within each subfamily, similarities are usually sufficient to enable detection by classical comparisons (>25% identities) (Henrissat et al. 1989). This has been shown to be true for *B. fibrisolvens* End1 and *C. thermocellum* EGE (37%) (Fig. 3.5). The amino acid sequences of other proteins in this subfamily (*C. thermocellum* EGH, *C. cellulolyticum* EGA, *R. albus* EGI, and *B. circulans* EG2) (Yagüe et al. submitted; Faure et al. 1990 in press; personal communication with Béguin in Béguin 1990) have not been published as yet, and it was not possible to compare them. Similarity found between End1 and the *Bacillus*-like endoglucanases (Zappe et al. 1988) stretched over 190 amino acids and was 39%. The *Bacillus*-like endoglucanases belong to the subfamily A2 (Béguin 1990) and some similarity with End1, a member of the subfamily A4 is to be expected. Alignment of all A-type enzymes revealed small conserved regions which are distributed along the sequence. In *B. fibrisolvens* End1 the conserved amino acids were: Arg_{97}, Pro_{99}, His_{141}, Asn_{188}, Glu_{189}, Pro_{190}, His_{273}, and Glu_{321} (reviewed by Béguin 1990). Based upon the similarities of the substrates, it has been hypothesized that the active site of cellulases might resemble that of lysozyme, with carboxylic acid-bearing side chains participating in catalysis (Yaguchi et al. 1983). They proposed that EGI from the fungus *Schizophyllum commune* might act by a mechanism similar to that of lysozyme. However, the lysozyme-like region was not generally conserved in family A enzymes and was not present in End1. Most cellulases carry large numbers of acidic amino acids, and tentative active site identifications based on lysozyme-like sequences must therefore be viewed with caution. As families A to F each showed different structural frameworks, it would be more reasonable to base hypotheses concerning putative active sites on residues that are conserved within each family (Béguin 1990). All true cellulolytic bacteria (capable of degrading crystalline cellulose) produced enzymes belonging to different families.

This suggests that interspecific exchange of cellulase genes has occurred on a wide basis in the course of evolution and that cellulolytic organisms might have developed by collecting genes of various origins.

In most cellulases, the region required for catalytic activity is flanked by a non-catalytic domain which is distinct from the catalytic domain. This domain has been shown to be involved in the binding of *C. fimi* (Langsford et al. 1987; Gilkes et al. 1988) and *T. reesei* (Van Tilbeurgh et al. 1986; Tomme et al. 1988) enzymes to the cellulose substrate. In the *C. thermocellum* enzymes this domain may be involved in substrate binding or serve as an anchorage domain in the cellulosome, and in a number of *Bacillus* species this domain may be involved in secretion (Chapter 1.2.5.2). Comparisons of different enzymes have revealed that several distinct types of non-catalytic domains are associated with catalytic cores from different families. Based on the presence of different non-catalytic domains Béguin (1990) divided the cellulolytic enzymes into four types (see Table 1.2). End1 was classified as having a *C. fimi* type substrate binding domain. It is interesting to note that in all organisms analyzed so far, only one type of conserved non-catalytic domain was found in each species. This is in contrast with the catalytic domains within a species which often belong to different families. The non-catalytic domains were probably added to the core enzymes at a relatively late stage in evolution (Béguin 1990).

The catalytic and non-catalytic domains are often separated by a stretch of amino acids (10-30 residues) which is highly enriched in P and hydroxyl amino acids. In *T. reesei* cellulases EGI (Penttilä et al. 1986), EGIII (Saloheimo et al. 1988), CBHI (Shoemaker et al. 1983), and CBHIII (Chen et al. 1987) the conserved domain was separated from the catalytic domain by one or two P-T-S rich segments. Knowles et al. (1987) reported that the junction region between the homologous N-terminal and the non-homologous C-terminal regions of the endoglucanase enzymes from a number of *Bacillus* species was characterized by an area rich in hydroxyl amino

acids and P. *C. fimi* CenA and Cex enzymes as well as *B. fibrisolvens* End1 contained P-T boxes 22-39 residues in size (Wong et al. 1986; O'Neill et al. 1986a; Berger et al. 1989). The P-T box is similar in structure to the hinge region of IgA_1 immunoglobulins (Fragione et al. 1972), and may function as such.

Chapter 4

Cloning in *Escherichia coli* of a cellodextrinase gene (*ced1*) from *Butyrivibrio fibrisolvens* H17c: DNA sequence analysis, and structural comparison of the deduced gene product with other cellulases.

Chapter 4

Cloning in *Escherichia coli* of a cellodextrinase gene (*ced1*) from *Butyrivibrio fibrisolvens* H17c: DNA sequence analysis, and structural comparison of the deduced gene product with other cellulases.

4.0 Summary

A gene expressing cellodextrinase activity in *E. coli* was cloned from *B. fibrisolvens* H17c DNA on a 3.55 kb *Sau*3A restriction fragment in the recombinant plasmid pES500. The nucleotide sequence of a 2.314 kb DNA segment containing the *ced1* gene was determined. The Ced1 enzyme consisted of an ORF of 1641 bp which coded for a protein of 547 amino acids with a calculated M_r of 61 023. A GTG start codon was preceded by a putative Shine-Dalgarno (SD) sequence (GGAGGG) and a weak putative promoter consensus sequence. The stop codon (TAA) was not followed by a termination sequence. Homology was demonstrated between the *B. fibrisolvens* Ced1 cellodextrinase and the *C. thermocellum* endoglucanase D, *P. fluorescens* var. *cellulosa* endoglucanase, and a cellulase from the avocado fruit (*P. americana*).

4.1 Introduction

Most cellulase genes cloned in *E. coli* produce detectable levels of cellulase (endoglucanase or cellobiosidase) activity, and in many cases expression of the cellulase genes in *E. coli* is from native promoters (Béguin 1990). However, it has been shown that the fusion of some of the cellulase genes with promoters and ribosomal binding sites from highly expressed genes in *E. coli* result in much higher levels of cellulase expression (Béguin 1990, O'Neill et al. 1986b, Schwarz et al. 1987). An in-frame translational fusion between the *C. thermocellum celD* gene (Millet et al. 1985) and the *lacZ* promoter of pUC8 resulted in a 1000-fold increase of the *C. thermocellum* EGD enzyme (Joliff et al. 1986a). The *C. thermocellum* EGD enzyme precipitated as cytoplasmic granules from which the purified protein was readily crystallized (Joliff et al. 1986b). The crystals were suitable for high resolution X-ray diffraction analysis (Juy et al. 1988).

In common with other genes which code for secreted proteins, the N-terminal ends of many cloned cellulases have been found to contain a signal peptide and the proteins are partially or entirely transported to the periplasm of *E. coli* (Béguin 1990). However, the presence of a N-terminal, hydrophobic signal peptide does not seem to be a prerequisite for the transfer of a protein from the cytoplasm. Wang and Thomson (1990) found that the *R. flavefaciens* CelA cellodextrinase was transported to the periplasm without a typical procaryotic signal sequence, and Hall et al. (1989) found that deletion of 263 amino acids from the 5' end of the *P. fluorescens* var *cellulosa xynA* gene did not prevent the secretion of the xylanase enzyme into the periplasm of *E. coli*.

This chapter deals with the cloning and sequencing of a gene expressing cellodextrinase activity from *B. fibrisolvens* H17c in *E. coli*. The recombinant plasmid pES500 containing the cellodextrinase gene was originally isolated from LB agar plates containing CMC as a substrate. Comparison of the deduced amino acid sequence with other reported sequences showed good

similarity between the *B. fibrisolvens* Ced1 enzyme and the catalytic domain of the members of the E1 subfamily (*C. thermocellum* EGD and *P. fluorescens* EGA) (Berger et al. 1990; Hall and Gilbert 1988; Joliff et al. 1986a). Similarity was also found between the *B. fibrisolvens* Ced1 enzyme and a member from the E2 subfamily (the avocado fruit *P. americana* EG cellulase) (Tucker et al. 1987).

4.2 Materials and Methods

4.2.1 Bacterial strains, plasmids, and phage. The plasmids pES500 and pES520 were used as primary sources of DNA for plasmid mapping and templates for subcloning and DNA deletions. The plasmid vectors pUC18 and pUC19 (Messing 1983) were used for the preparation of templates for DNA sequencing. *E. coli* strains LK111 and C600 (Appendix C) were used as hosts for these recombinant plasmids. The 2.1 kb *Stu*I DNA restriction fragment from pES530 (Fig. 4.1) was ligated to pBR325 (*Eco*RV digested) DNA. The resultant recombinant plasmid, pBn530 was used to transform *E. coli* CC118 (*phoA*-) (Manoil and Beckwith 1985) for Tn*phoA* fusion experiments (Chapter 3.2.9).

4.2.2 Media, buffers and growth conditions. All media and buffers not described in the text are given in Appendix A. Growth conditions are as described in Chapter 3.2.3.

4.2.3 Restriction mapping and Southern blot analysis. Plasmid DNA was prepared according to the method of Ish-Horowicz and Burke (1981) and mapping of pES500 was accomplished by the use of single- and double-enzyme digests using the methods for restriction endonuclease digestions described in Appendix B.

Chromosomal DNA from *B. fibrisolvens* H17c was digested with *Stu*I endonuclease, and resolved by electrophoresis in an agarose gel (0.8% w/v) in Tris-acetate buffer. The DNA was transferred to Hybond-N hybridization membrane (Amersham) as described by Reed and Mann (1985) (Appendix B). Plasmid pES500 nick-translated with [α-^{32}P]dCTP was used as a hybridization probe (Rigby et al. 1977) (Appendix B).

4.2.4 Sequencing strategy and nucleotide sequencing. Restriction mapping and subcloning showed that the cellodextrinase gene was located on a 2.314 kb *Bgl*II deletion fragment of pES520 (Fig. 4.1). Subcloning of fragments into pUC18 and pUC19 using available restriction sites was used to obtain templates for sequencing (Fig. 4.3). The templates were sequenced and the sequences were analyzed as

described in Chapter 3.2.8. The whole insert of pES520 was sequenced at least once in each orientation.

4.2.5 Transposon (Tn*phoA*) mutagenesis. Transposon insertions into pBn530 were obtained using an adaptation of the protocol of Gutierrez et al. (1987). The method, isolation, and nucleotide sequencing of transposon insertions were described in Chapter 3.2.9.

4.3 Results

4.3.1 Cloning of a cellodextrinase gene (*ced1*) from *B. fibrisolvens* H17c. *E. coli* C600 was transformed with recombinant pEcoR251 plasmid DNA from the *B. fibrisolvens* genomic library (Chapter 2.2.7) and screened on 0.1% CMC-LB agar plates. A recombinant plasmid pES500 was isolated from a colony which produced a small halo on CMC-LB agar. Restriction enzyme analysis showed this plasmid to differ from pES400 isolated from the same *B. fibrisolvens* gene library (Chapter 2.3.2). CMCase activity of pES500 was always associated with transformation to Ap^R. Restriction endonuclease mapping of pES500 revealed a 3.55 kb DNA insert in pEcoR251 (Fig. 4.1).

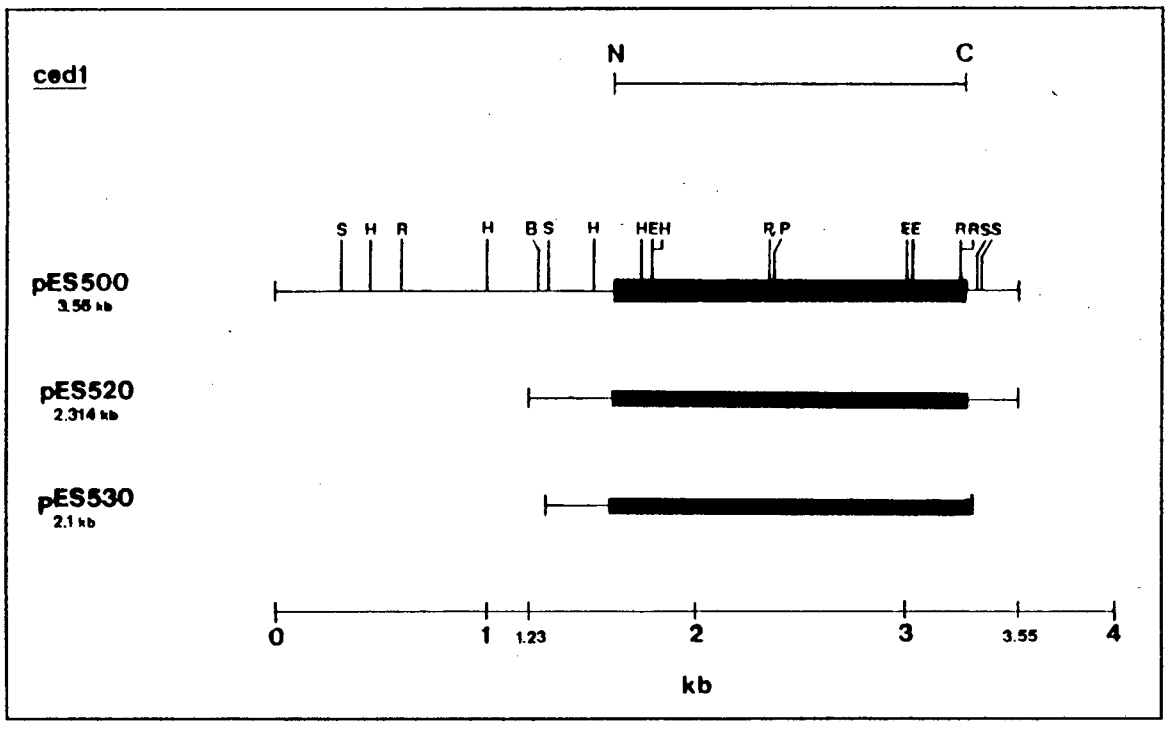


Fig. 4.1. Restriction endonuclease map of the *B. fibrisolvens* cellodextrinase gene (*ced1*). The cleavage sites of restriction enzymes *Eco*R1 (R), *Hin*dIII (H), *Eco*RV (E), *Pvu*II (Pv), *Pst*I (P), *Bgl*II (B) and *Stu*I (S) are shown. The *ced1* gene (1.641 kb) is indicated by the heavy line with N and C representing the amino- and carboxy-terminal ends of the encoded cellodextrinase enzyme.

A *Bgl*II-*Bam*HI restriction endonuclease deletion of the 3.55 kb insert resulted in the construction of pES520 (2.314 kb), and the subcloning of a *Stu*I restriction endonuclease fragment into pUC19 resulted in pES530 (2.1 kb). Both inserts on pES520 and pES530 were able to express CMCase activity (Fig. 4.1). Subcloning of

an *Eco*RI-*Bam*HI restriction fragment into pUC19 (*Eco*RI-*Bam*HI restriction endonuclease digested) resulted in pES540 (1.85 kb) which showed no CMCase activity. The insert DNA was subcloned in both orientations in pUC and pEcoR251 vectors. Since CMCase activity was detected independently of the orientation of the insert in the vector, it was concluded that in the *E. coli* host the enzyme activity was regulated by a promoter internal to the cloned insert. However, expression of CMCase activity from the internal promoter was weak and was enhanced when it was in tandem with an additional external promoter on the vector (data not shown).

4.3.2 Origin of the DNA insert. The origin of the cloned DNA fragment in pES500 was determined by hybridization of [α-^{32}P]dCTP labelled pES500 to *Stu*I restriction endonuclease digested *B. fibrisolvens* chromosomal DNA. *Stu*I endonuclease digestion of pES500 produced two internal insert fragments (1.0 and 2.1 kb) which hybridized to two identical sized *B. fibrisolvens* DNA fragments (Fig. 4.2, lane A).

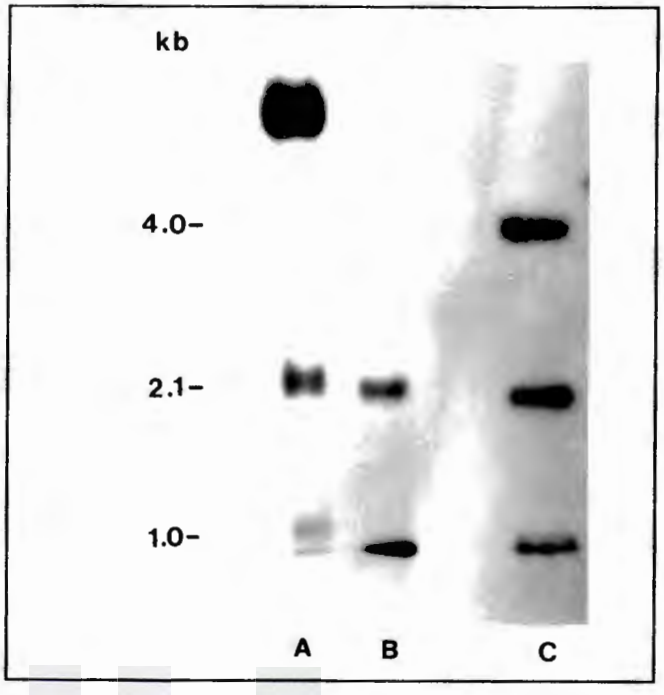


Fig. 4.2. Autoradiograph of ^{32}P-labelled pES500 hybridized to *Stu*I restriction endonuclease digests of: Lane A; *B. fibrisolvens* chromosomal DNA (*Stu*I), lane B; *B. fibrisolvens* chromosomal DNA (*Bgl*II-*Stu*I), and lane C; pES500 (*Stu*I).

The presence of a hybridization signal at >11 kb represents hybridization between pES500 and a *B. fibrisolvens* H17c chromosomal fragment which flanks the internal *Stu*I fragments on one side (Fig. 4.2, lane A). The *Stu*I restriction site on the other side of the internal *Stu*I fragments is only 86 bp from the end of the insert (Fig. 4.1).

The absence of a fourth hybridization signal in lane A (Fig. 4.2) may be the result of too weak a signal between this 86 bp fragment of the probe and the corresponding *B. fibrisolvens Stu*I fragment, or hybridization with a small fragment which ran off the gel. The *Bgl*II-*Stu*I digested *B. fibrisolvens* H17c chromosomal digest in lane B confirmed the restriction map depicted in Fig. 4.1. Lane C was a positive control of *Stu*I endonuclease digested pES500.

4.3.3 Preparation of clones for sequencing. Clones used for sequencing of the *ced1* gene were obtained from subcloning of restriction fragments into pUC18 and pUC19 using available restriction sites. The extent of the sequence obtained from each clone is depicted in Fig. 4.3.

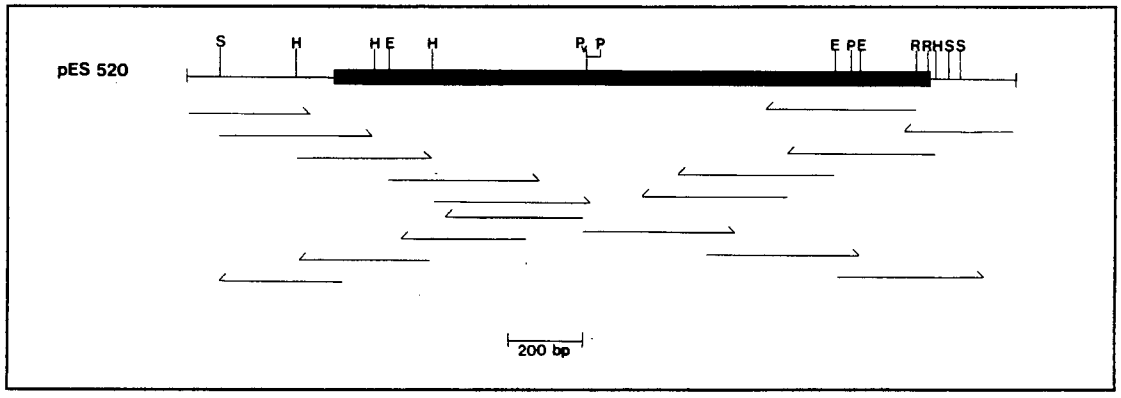


Fig. 4.3. The sequencing strategy for the 2.314 kb fragment containing the *ced1* gene from *B. fibrisolvens* H17c. The bold line indicates the position of the *ced1* gene, and the thin line represents the *B. fibrisolvens* DNA which flanks the gene in pES500. The cleavage sites of restriction enzymes *Eco*R1 (R), *Hin*dIII (H), *Eco*RV (E), *Pvu*II (Pv), *Pst*I (P) and *Stu*I (S) are shown. The arrows indicate the extent of sequence obtained from each clone.

4.3.4 Nucleotide sequence of the *ced1* gene. The nucleotide sequence of the functional 2.314 kb insert in pES520 was determined and coded for a single ORF, which from the presumptive start codon (GTG) to the stop codon (TAA) contained 1 641 nucleotides encoding 547 amino acid residues (Fig. 4.4). The calculated M_r of the predicted polypeptide encoded by the ORF was 61 023. The region upstream of the ORF contained a putative consensus promoter sequence (Fig. 4.4) with an inter-region space of 20 bp. Promoters with inter-region spaces of up to 20 bp were reported to be partially functional in *E. coli* (Hawley and McClure 1983). A putative SD sequence (GGAGGG) was located 6 bp upstream of the initiation codon (GTG) (Fig. 4.4).

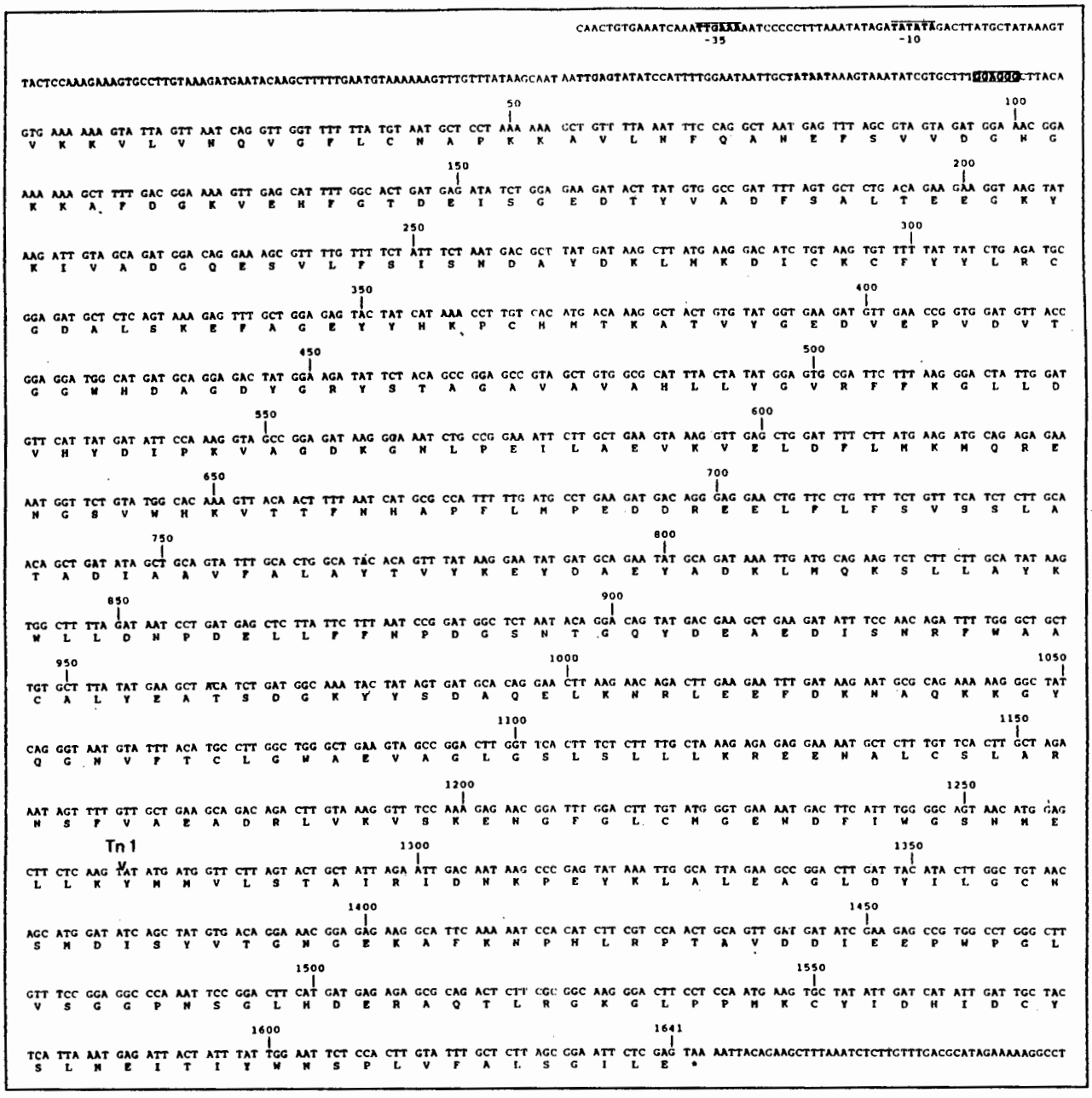


Fig. 4.4. Complete nucleotide sequence of the cellodextrinase (*ced1*) gene from *B. fibrisolvens* H17c. The derived amino acid sequence is given in the one-letter code from position 1 to 1 641 (547 residues). Putative -10 and -35 promoter sequences are overlined. The Shine-Dalgarno (SD) sequence is boxed. The position of the Tn*phoA* insert on pTn5ϕ1 is indicated (v).

The GC content of the *B. fibrisolvens ced1* gene was 40% and that of the upstream regulatory region 34%. The codon usage reflects the expected low GC content (42%) of *B. fibrisolvens* H17c DNA (Mannarelli 1988).

4.3.5 Tn*phoA* fusions. Since the majority of the Ced1 activity was located in the periplasm (Chapter 5.3.2) it was important to determine whether the *ced1* gene contained an N-terminal sequence that could code for a signal peptide typical of protein export in *E. coli*. k::Tn*phoA* mutagenesis (Manoil and Beckwith, 1985) was utilized to detect a signal peptide. The alkaline phosphatase gene fused to Tn*phoA* needs a signal peptide for export. As export is essential for alkaline phosphatase activity, an in-frame sequence upstream of the transposition site has to function as a signal peptide for activity to occur. Tn*phoA* mutagenesis of pBn530 to determine whether the cellodextrinase enzyme contained a functional signal peptide resulted in the isolation of 7 recombinant plasmids conferring alkaline phosphatase activity on the Pho$^-$ recipient strain *E. coli* CC118. Restriction enzyme analysis of the 7 clones revealed that all the Tn*phoA* fragments were inserted downstream of the functional tetracycline signal peptide in the vector. To ensure that it was possible for Tn*phoA* to transpose into the gene expressing cellodextrinase activity, six Ap$^+$, Pho$^-$, and CMC$^-$ mutants were isolated, plasmid DNA extracted, and the nucleotide sequences across the *phoA/ced1* junction points were determined. One recombinant plasmid pTn5ϕ1 was found with the Tn*phoA* insert downstream, in the right orientation and in-frame of a putative signal sequence (Fig. 4.4), indicating that this signal sequence was not able to facilitate the secretion of *phoA* in *E. coli*.

4.3.6 Amino acid sequence homology. Amino acid alignment studies showed homology between the *B. fibrisolvens* Ced1, *C. thermocellum* EGD (Joliff et al. 1986a), *P. fluorescens* var. *cellulosa* EGA (Hall and Gilbert 1988), and *P. americana* EG (Tucker et al. 1987) enzymes (Table 4.1).

Table 4.1. Identity (I) and similarity (S) between the cellulases of *B. fibrisolvens* Ced1 (BF), *C. thermocellum* EGD (CT) (Joliff et al. 1986a), *P. fluorescens* EGA (PF) (Hall and Gilbert 1988), and *P. americana* EG (PA) (Tucker et al. 1987).

	BF %I (%S)	CT %I (%S)	PF %I (%S)	PA %I (%S)
BF	-	35 (59)	29 (52)	22 (46)
CT	35 (59)	-	29 (53)	22 (46)
PF	29 (52)	29 (53)	-	22 (46)
PA	22 (46)	22 (46)	22 (46)	-

The best similarity was found between *B. fibrisolvens* Ced1 and *C. thermocellum* EGD (Fig. 4.5). The proteins of *C. thermocellum* EGD (649 amino acids) and *P. fluorescens* EGA (962 amino acids) are larger than that of the *B. fibrisolvens* Ced1 (547 amino acids) protein. Similarity stretched over the full length of the *B. fibrisolvens* Ced1 protein, but only over approximately the first 600 amino acids of the *C. thermocellum* EGD and *P. fluorescens* EGA proteins. Henrissat et al. (1989) used hydrophobic cluster analysis (HCA) to classify cellulases into six distinct families, some of which were further subdivided into subfamilies (Béguin 1990). The *C. thermocellum* EGD (Joliff et al. 1986a) and *P. fluorescens* var. *cellulosa* EGA (Hall and Gilbert 1988) enzymes were classified as members of the subfamily E1 (Henrissat et al. 1989) and the cellulase from the avocado fruit (*P. americana*) (Tucker et al. 1987) was classified as a member of the subfamily E2 (Béguin 1990).

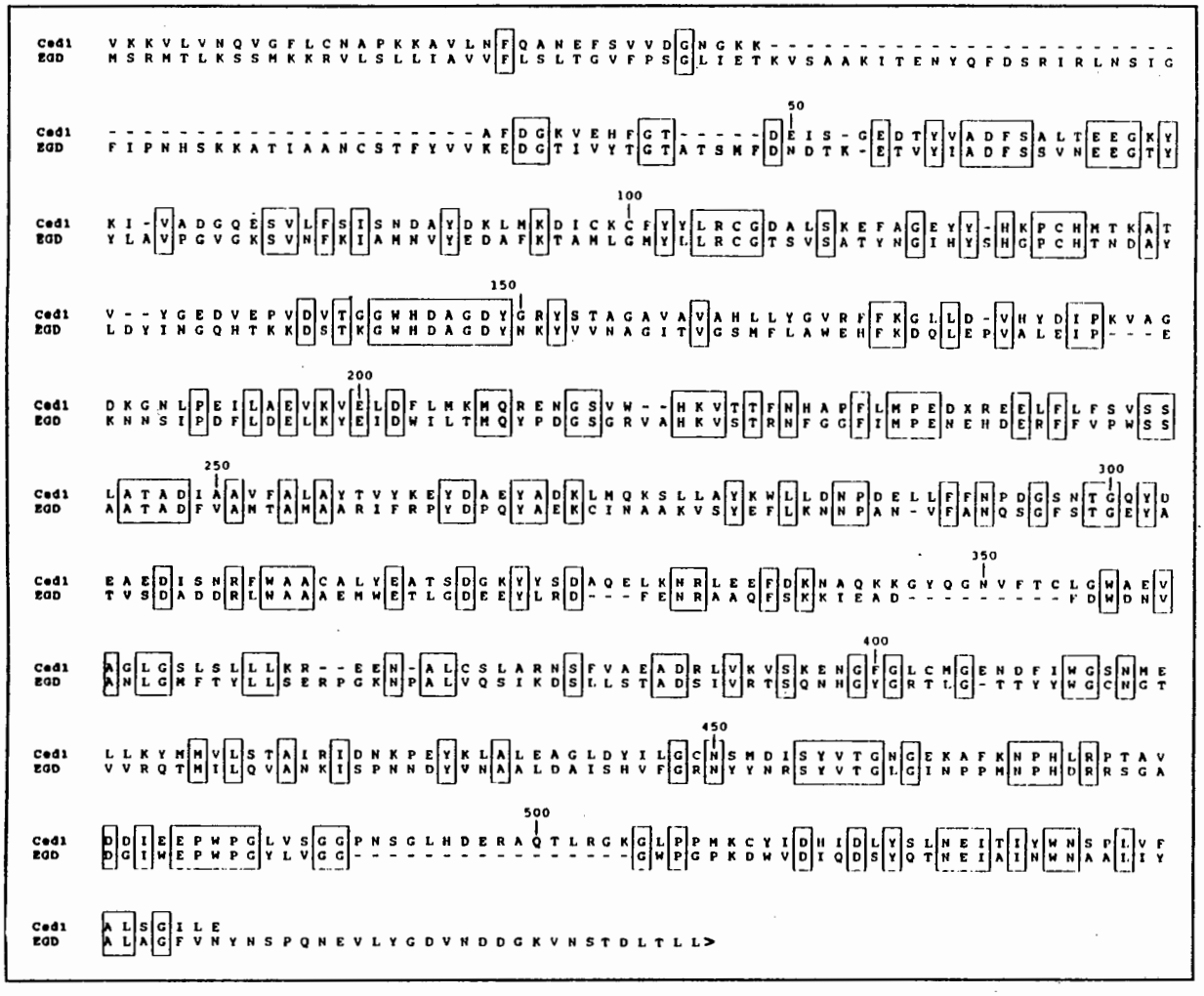


Fig. 4.5. Amino acid sequence alignment of the *B. fibrisolvens* Ced1 and the *C. thermocellum* EGD (Joliff et al. 1986a) proteins. The amino acids are identified by the single-letter code and regions containing identical amino acids are boxed.

Alignment studies of the cellodextrinase Ced1 with the endoglucanase End1 of *B. fibrisolvens* H17c (Berger et al. 1989) and endoglucanases EGA, EGB, EGC, and EGE of *C. thermocellum* (Béguin et al. 1985, Grépinet and Béguin 1986, Schwarz et al. 1988a, Hall et al. 1988) showed less than 12% identity. Alignment with a number of other endoglucanases showed very little identity (<10%).

4.4 Discussion

The *ced1* gene was expressed from an internal promoter region in *E. coli* since enzyme activity was obtained when the insert was in both orientations in pUC and pEcoR251 vectors. However, expression from the internal promoter was weak and was enhanced when it was in tandem with an additional promoter of the vector. The GTG initiation codon was identified by the calculated and observed *Mr* (Chapter 5.3.3) of the *ced1* gene product and the presence of a ribosome binding site. It was extremely unlikely that an internal ATG codon was utilized since the first internal ATG codon was 276 nucleotides from the selected GTG start codon which would mean the production of a protein which lacked 92 amino acids. This did not correlate with the observed *Mr* (Chapter 5.3.3) of the enzyme and the internal ATG codon was not preceded by a ribosome binding site. There was no upstream ATG codon in phase with the open reading frame of the *ced1* gene. Tn*phoA* mutagenesis showed that the Ced1 enzyme did not have a typical procaryotic signal peptide which was able to facilitate the secretion of *pho*A in *E. coli*. Other examples of export across the cytoplasmic membrane without a signal peptide are the *Vibrio alginolyticus* sucrase enzyme (Scholle et al. 1989) and the *R. flavefaciens* CelA cellodextrinase enzyme (Wang and Thomson 1990).

Homology between the *C. thermocellum* EGD (Joliff et al. 1986a) and *P. fluorescens* EGA (Hall and Gilbert 1988) enzymes spanned nearly 600 amino acids which corresponded with the catalytic domain of each protein, and they were classified accordingly as members of the E1 subfamily (Henrissat et al. 1989). However, the C-terminal domains of the two proteins were different in size and showed no homology. The *C. thermocellum* EGD protein (Joliff et al. 1986a) contained a typical *C. thermocellum* non-catalytic domain (Béguin 1990), while the *P. fluorescens* EGA protein (Hall and Gilbert 1988) contained two regions rich in the hydroxyl amino acid Ser and a *C. fimi* type non-catalytic binding domain (Béguin 1990). It was also shown that deletion of these C-terminal domains did not affect enzyme activity (Chauvaux et al. 1989; Hall and Gilbert 1988). The function of these non-homologous C-terminal domains are most likely that of substrate binding domains, or in the case of the *C. thermocellum* EGD it may serve as an anchorage

domain to some scaffolding protein in the cellulosome (Béguin 1990; Lamed and Bayer 1988). Amino acid alignment of the *B. fibrisolvens* Ced1 and *C. thermocellum* EGD (Joliff et al. 1986a) enzymes showed very good similarity (59%) over the full-length of the Ced1 protein (547 amino acids), while similarity between the *B. fibrisolvens* Ced1 and *P. americana* EG (Tucker et al. 1987) (member of E2 subfamily) enzymes was only 46%, suggesting that Ced1 is a member of the E1 subfamily. The Ced1 protein does not seem to have a distinct non-homologous binding domain. Removal of 39 bp from the 3' end of the *ced1* gene resulted in complete inactivation of the Ced1 enzyme indicating that the whole protein is required for enzyme activity. The absence of a *C. fimi*-type binding domain in the *B. fibrisolvens* Ced1 protein was unexpected as the *B. fibrisolvens* End1 protein had a *C. fimi*-type binding domain (Chapter 3.3.6) and in all organisms studied so far only one type of conserved non-catalytic domain was found in each species (Béguin 1990). However, the *C. thermocellum* EGC (Schwarz et al. 1988a) protein also did not contain a *C. thermocellum*-type non-catalytic domain and the protein was found not to be involved with the *C. thermocellum* cellulosome (Béguin 1990). Neither the *B. fibrisolvens* Ced1 enzyme nor the *C. thermocellum* EGC enzyme had a P-T rich "hinge" region. Both enzymes seemed to contain only a catalytic domain. The absence of a distinct substrate binding domain may indicate that these enzymes need to form only loose associations with the substrate for hydrolysis to take place.

Chapter 5

Characterization and substrate specificity of the *Butyrivibrio fibrisolvens* H17c endoglucanase End1 and the cellodextrinase Ced1 expressed in *Escherichia coli* C600.

Chapter 5

Characterization and substrate specificity of the *Butyrivibrio fibrisolvens* H17c endoglucanase End1 and the cellodextrinase Ced1 expressed in *Escherichia coli* C600.

5.0 Summary

The majority of the *B. fibrisolvens* End1 and Ced1 enzyme activities were located in the periplasm of the *E. coli* host. The End1 enzyme showed optimal activity at pH 5.6 and 45°C. Optimal activity of the Ced1 enzyme was at pH 6.6 and 50°C. The End1 enzyme was most active against endoglucanase specific substrates lichenan and CMC, and End1 was therefore classified as an endo–β–1,4–glucanase. Ced1 was classified as a cellodextrinase as it was most active in hydrolyzing cellodextrins yielding cellobiose or cellobiose and glucose as endproducts, released cellobiose from pNPC, showed very low levels of activity against recalcitrant cellulose and was not inhibited by methylcellulose.

5.1 Introduction

The majority of cloned cellulase genes code for endo–β–1,4–glucanase enzymes (Béguin 1990). These enzymes generally show high levels of activity against soluble substrates such as CMC, lichenan, and barley β–glucan, whereas most of them show little activity against crystalline cellulose and are only able to degrade cellodextrins to a limited extent or not at all. Many procaryotic endoglucanases and xylanases show activity against pNPC, indicating exo–type activity, but this chromophoric substrate has been shown to be nonspecific for enzyme type (Deshpande et al. 1984, Ohmiya et al. 1982) and caution is necessary in interpreting these results. The presence of low levels of pNPCase activity is therefore to be expected with endoglucanase enzymes.

Genes expressing cellodextrinase activity in *E. coli* have been cloned from *F. succinogenes* S85 (Gong et al. 1989) and *R. flavefaciens* FD1 (Barros and Thomson 1987, Wang and Thomson 1990). These enzymes degrade cellodextrins to produce cellobiose or cellobiose and glucose as end products, and show little or no activity on CMC or crystalline cellulose. Cellodextrinases have been confused in the literature with exoglucanases (1,4–β–D–glucan cellobiohydrolases). Both enzymes remove cellobiose units from the non-reducing end of the polymer, but an exoglucanase enzyme acts mainly on crystalline cellulose whereas a cellodextrinase enzyme acts mainly on cellooligosaccharides (Huang et al. 1988).

The recalcitrance of methylcellulose to degradation and its widespread inhibitory activity make this compound a useful tool in the study of bacterial adherence to cellulose and of the enzymology of cellulose degradation. Rasmussen et al. (1988) used methylcellulose to explore the mechanisms of cellulose degradation in *R. flavefaciens* FD1. Methylcellulose inhibited cellulase activity in a dose-dependent manner, as measured by pNPC hydrolysis. In contrast, methylcellulose did not inhibit hydrolysis of cellodextrins. These data suggested that methylcellulose acted as a nonhydrolyzed, competitive inhibitor of those enzymes displaying endo–β–1,4–glucanase activity. An endoglucanase from *F. succinogenes* was also inhibited by the presence of methylcellulose (Groleau and Forsberg 1983). These

endoglucanases might also be involved in attachment as methylcellulose inhibits attachment of ruminal bacteria to cellulose and plant materials (Kudo et al. 1987; Minato and Suto 1981).

The effect of temperature and pH on the End1 and Ced1 enzyme activities are discussed in this chapter. The activity of both enzymes against a range of substrates was analyzed and this information was used to classify the enzymes as an endo–β–1,4–glucanase (End1) and a cellodextrinase (Ced1).

5.2 Materials and Methods

5.2.1 Bacterial strains and plasmids. *B. fibrisolvens* H17c, *E. coli* C600[pES420] and *E. coli* C600[pES530] were used for the preparation of cell extracts for enzyme assays. *E. coli* C600[pUC19] was used as a control.

5.2.2 Media and buffers. All media, buffers and solutions not described in the text are given in Appendix A.

5.2.3 Preparation of cell extracts. All extracts were prepared from *E. coli* strains grown in LB + Ap, and from *B. fibrisolvens* H17c grown in modified M10 medium. Cell-free extracts (CFE) were prepared from 200 ml overnight cultures of *E. coli* C600 (containing various plasmids) grown in LB broth with Ap (100 μg/ml). The cells were collected by centrifugation (8 000 *g*; 5 min), washed, resuspended in 5 ml phosphate–citrate (PC) buffer (50 mM K_2HPO_4; 14 mM citric acid; pH 5.6 or pH 6.6) and disrupted by sonication on ice (10 s bursts for a total of 200 s). The extract was clarified by centrifugation (27 000 *g*; 15 min; 4°C) and the supernatant was retained as the cell–free extract.

Periplasmic extracts were prepared according to the osmotic shock procedure of Willis et al. (1974). An overnight culture of *E. coli* C600[pES420] or *E. coli* C600[pES520] was diluted (1/1000) into fresh LB broth and Ap (100 μg/ml) and grown to an OD_{600} of 1.0. Cells from 100 ml of culture were removed by centrifugation (6000 *g*; 5 min) and the supernatant fraction stored at -20°C. NaCl and Tris–Cl (pH 7.3) were added to the remaining 100 ml of the culture to give a final concentration of 33 mM. Incubation was continued for a further 10 min and the cells were harvested and resuspended in Tris–Cl (33 mM, pH 7.3) (10 ml/g wet wt. cells). An equal volume of TSE (33 mM Tris–Cl, 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 deionized water (20 ml/g wet wt. cells) and $MgCl_2$ (1 mM final concentration) was added within 1 min. The cells were fractionated by centrifugation and the supernatant was retained as the periplasmic fraction. To obtain the cytoplasmic fraction, the cells

were resuspended in 5 ml PC buffer (50 mM K_2HPO_4; 12.5 mM citric acid; pH 6.6) and the cell suspension was cooled on ice and disrupted by sonication on ice (10 s bursts for a total of 200 s) using a MSE (Soniprep 150) sonicator. The cell debris was removed by centrifugation for 15 min at 27000 *g* and the supernatant was retained as the cytoplasmic extract.

Extracellular supernatant samples were concentrated by filtration using an Amicon ultrafiltration cell (model 8050; Amicon Corp., Danvers, Mass.) fitted with an Amicon PM–10 membrane. All the extracts were stored at -20°C.

B. fibrisolvens H17c was grown on modified M10 medium and the CFE was prepared as above.

5.2.4 Protein determination and SDS polyacrylamide gel electrophoresis (SDS–PAGE). Concentrations of proteins in crude protein preparations were determined according to the method of Bradford (1976) using BSA (fraction V) as a standard (Appendix B). Proteins from cell-free extracts of *B. fibrisolvens* and *E. coli* C600 cultures containing pES420, pES530 or pUC19 were separated on SDS–PAGE gels (Laemmli 1970) (Appendix B). For the detection of CMCase activity a modified method of Sharma and Sandhu (1986) was used. Samples (0.5 ml) were preincubated with solubilization mix (Laemmli 1970) for 1 h at 39°C and run on SDS-PAGE gels containing CMC (0.1% w/v) in the resolving gel.

5.2.5 Enzyme assays.

5.2.5.1 Cellulase assays. Endo–β–1,4–glucanase activity was assayed by the release of glucose equivalents as detected by the dinitrosalicylic (DNS) reagent for reducing sugars (Ghose 1987). Appropriately diluted samples (0.25 ml) were incubated with 0.25 ml of the substrate CMC (2% w/v; Sigma No. C4888; DS 0.7), lichenan (1% w/v), laminarin (1% w/v), or oat spelt xylan (2% w/v)) in 50 mM PC buffer (pH 5.6 or pH 6.6) for 30 min at 45°C. The samples were vortexed and incubated for 30 min at the optimum enzyme temperature 45°C. DNS (1.5 ml) was added and the samples were boiled for 5 min. The samples were diluted 1/6 in deionized water and the OD_{540} was measured. Enzyme activity was expressed as

International Units (IU) where 1 IU is equal to 1 µmol/min of liberated hydrolysis product. The calculation of the IU was determined according to the formula approved by the Commission on Biotechnology (Ghose 1987).

Viscosimetric assays performed by P. du Preez and D. Barry (AECI Research and Development Department) measured endoglucanase activity. Viscosity was measured by the addition of undiluted enzyme (6.3 ml) to the CMC substrate (14 ml; 2% w/v; 7 MP) in PC buffer (pH 6.3) containing sodium azide (0.02% w/v) and the mixture was incubated at 39°C in a Brookfield cone plate viscometer. The time of outflow of the reaction mixture was measured at different time intervals and the specific viscosity was calculated. The reducing sugar assay was carried out using 1 ml of the mixture removed at different time intervals.

Enzyme activity against Avicel PH-102 (FMC Corp.) was assayed by incubating 0.25 ml enzyme solution with 0.25 ml Avicel solution (1% w/v) in PC buffer. Enzyme activity against filter paper (Whatman no. 1) was assayed by incubating 25 mg strips of filter paper in 0.5 ml PC buffer with 0.5 ml enzyme solution. The samples were incubated for 24 h at 45°C and then assayed for reducing sugars as mentioned above.

Enzyme activity was measured against *p*-nitrophenyl-β-D-cellobioside (pNPC) (3.4 mM) (Deshpande et al. 1984), *p*-nitrophenyl-β-D-Xylopyranoside (pNPX) (6 mM), salicin [2-(hydroxymethyl) phenyl-β-D-glucopyranoside] (1% w/v) and *p*-nitrophenyl-β-D-glucopyranoside (pNPG) (6 mM) in PC buffer (pH 5.6 or pH 6.6). The assay mixture containing appropriately diluted enzyme (250 µl) and substrate (250 µl) was vortexed and incubated at 45°C for 30 min. The assay was terminated by adding an equal volume of sodium carbonate (14% w/v), and absorbance was measured at 405 nm. One unit of enzyme activity was defined as the amount of enzyme releasing 1 µmole of *p*-nitrophenol/min.

The effect of methylcellulose on cellulase activity was measured with pNPC as the substrate (Rasmussen et al. 1988). Methylcellulose (Sigma No. M-0262, 400 cP) was added to enzyme assays at final concentrations of 0.04 and 0.4% w/v.

Hydrolysis of cellobiose, cellotriose, cellotetraose, cellopentaose, and cellohexaose (Merck) was determined by incubating the enzyme (20 µl, 5.5 µg of protein) with 100 µl of the cellodextrin (10 mg/ml in PC buffer) at 45°C. Samples were taken at various times and analyzed in a high pressure liquid chromatography (HPLC) system equipped with a mode 156 refractive index detector (Beckman). Separation was achieved on a C18 column (Waters Associates, Milford, MA). The column was held at room temperature and filtered distilled water was used as an eluant at a flow rate of 1.5 ml/min.

5.2.5.2 β-Galactosidase assay. β-Galactosidase activity was assayed using an adaptation of the method by Pardee et al. (1959) as described by Miller (1972). For the assay 100 µl of cell extract (cytoplasmic, periplasmic, or supernatant) was diluted with an equal volume of Z-buffer (Appendix A) and equilibrated at 28°C. The reaction was initiated by the addition of 40 µl of the substrate *o*-nitrophenyl-β-D-galactopyranoside (ONPG) (13 mM) in sodium phosphate buffer (0.25 M, pH 7.0). The mixture was incubated for 5 min at 28°C and the reaction was stopped with the addition of Na_2CO_3 (100 µl; 14% w/v) and diluted with water (1.2 ml) before measuring the OD_{420} of the released *o*-nitrophenol. One unit of enzyme was defined as the amount of enzyme that released 1 nM *o*-nitrophenol/min.

5.2.5.3 β-Lactamase assay. β-Lactamase activity was assayed according to the method of Sykes and Nördstrom (1972). The reaction mixture contained starch solution (1 ml; 0.2% w/v), the Ap substrate (1 ml; 0.1 mM sodium-Ap), and buffer (0.9 ml; 0.1 M potassium phosphate, pH 5.9). Both the starch solution and the Ap substrate were made in the above buffer. The reaction was equilibrated at 30°C and initiated by the addition of appropriately diluted enzyme (0.1 ml). The reaction was stopped after 20 min incubation at 30°C with the addition of trichloroacetic acid (TCA) (1 ml; 1% w/v) and the OD_{620} was measured. One unit of β-lactamase activity was defined as the amount of enzyme that hydrolyzes Ap at the rate of 1 µmole/min.

5.3 Results

5.3.1 Effect of pH and temperature on the *B. fibrisolvens* End1 and Ced1 enzymes produced by *E. coli*. The effect of pH on the CMCase activity of CFE from *E. coli* C600 containing pES420 (End1) or pES530 (Ced1) was determined. The enzymes were assayed in different citrate (pH 3.0 - pH 6.5) and phosphate (pH 6.5 - pH 8.5) buffers. The End1 enzyme showed maximal activity between pH 5.0 and pH 7.5 with a peak at pH 5.6 (Fig. 5.1). This was in contrast to the Ced1 enzyme which showed a sharp peak at pH 6.6 (Fig. 5.1).

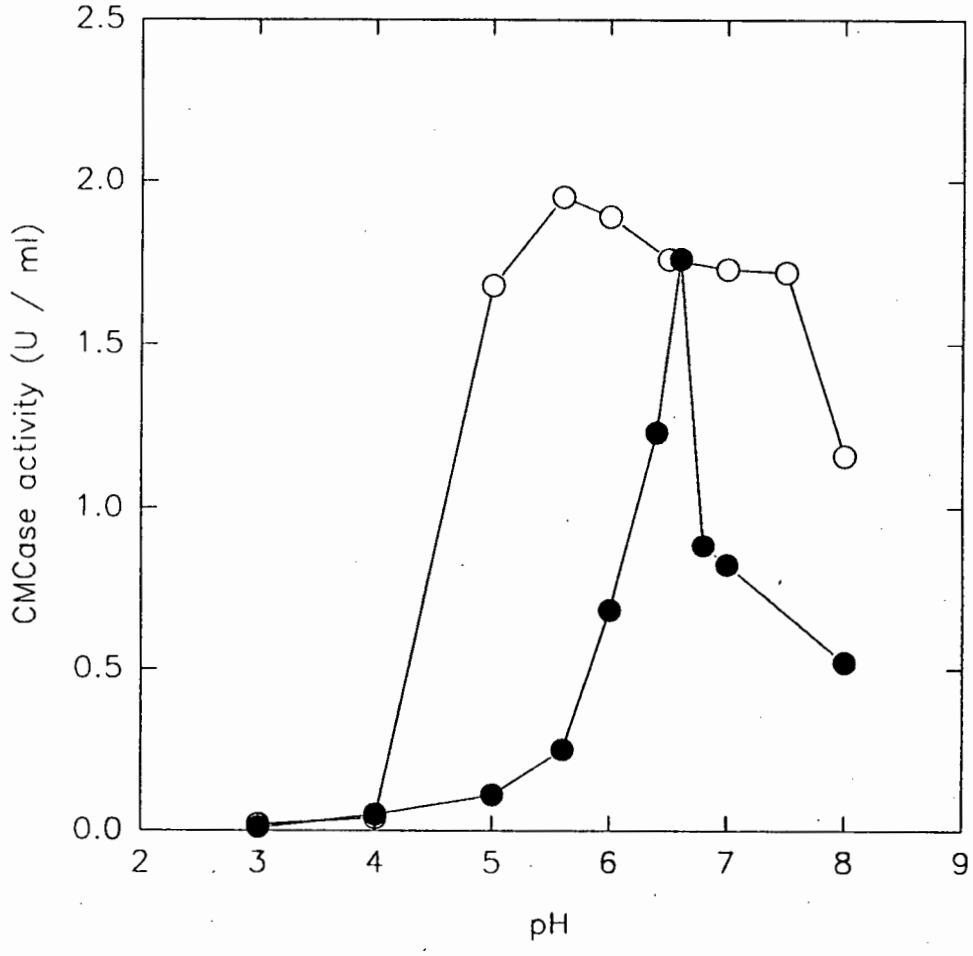


Fig. 5.1. pH activity profiles of the endoglucanase End1 and the cellodextrinase Ced1 from CFE of *E. coli* C600[pES420] and *E. coli* C600[pES530], respectively. Activity was assayed using substrate dissolved in the different buffers and the enzyme solutions were diluted in buffer of the specific pH. End1, (○); Ced1, (●).

The effect of temperature on the End1 and Ced1 enzymes was examined by performing the CMCase activity assay at various temperatures. The End1 enzyme showed maximum activity at 45°C whereas Ced1 showed maximum activity at 50°C (Fig. 5.2). Both enzymes showed a sharp decrease in activity at temperatures higher than 50°C, and the activity of End1 showed a sharper drop than that of Ced1 for temperatures lower than 45°C. The stability of the two enzymes was assessed by holding the enzymes for 60 min at various temperatures and then assaying the residual activity at 45°C. Both enzymes were stable at temperatures up to 45°C, but after 60 min at 50°C End1 had lost 55% of its activity and Ced1 85% (Fig. 5.2). From these results the parameters were chosen for enzyme assays. The pH for enzyme assays involving End1 was pH 5.6 and for Ced1 it was pH 6.6. The assay temperature for both enzymes was chosen as 45°C.

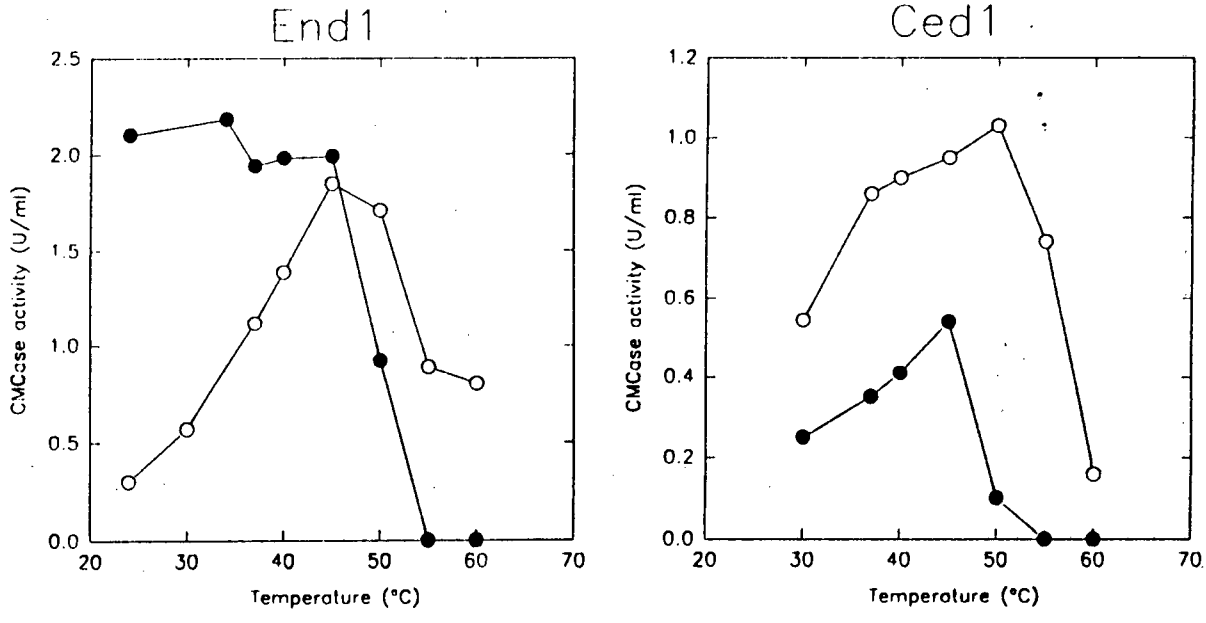


Fig. 5.2. Optimum temperature for activity (o), and temperature stability (●) profiles of the endoglucanase End1 and the cellodextrinase Ced1 enzymes from CFE of *E. coli* C600[pES420] and *E. coli* C600[pES530], respectively.

5.3.2 Localization of the *B. fibrisolvens* endoglucanase and cellodextrinase activity in *E. coli* C600 containing pES420 or pES530. Localization of the *B. fibrisolvens* End1 and Ced1 activities in *E. coli* C600 cells showed that periplasmic samples contained the majority (64 and 55%, respectively) of the activity (Table 5.1). The distribution of both enzymes reflected that of β-lactamase which is a periplasmic enzyme in *E. coli*.

The majority (92-94%) of the control cytoplasmic enzyme β-galactosidase was located in the cytoplasm (Table 5.1). The distribution of the control cytoplasmic and periplasmic enzymes was as expected.

Table 5.1. Distribution of endoglucanase (End1) and cellodextrinase (Ced1) activity in *E. coli* C600[pES420] and *E. coli* C600[pES530] cells, respectively.

Fractions	Total activity (units)[a]		
	CMCase (%)	β-galactosidase (%)	β-lactamase (%)
Supernatant			
End1	0.222 (15)	70 (6)	1263(11)
Ced1	0.477 (17)	14 (4)	900(12)
Periplasmic fraction			
End1	0.972 (64)	21 (2)	10078(84)
Ced1	1.524 (55)	5 (2)	5808(80)
Cytoplasmic fraction			
End1	0.315 (21)	1000(92)	744(6)
Ced1	0.768 (28)	290 (94)	547(8)

a Units CMCase; μmoles glucose released per min
Units β-galactosidase; nmoles ONP released per min
Units β-lactamase; μmoles Ap hydrolyzed per min

5.3.3 SDS-PAGE analysis. The proteins in CFE from *B. fibrisolvens* H17c, *E. coli* C600[pES420] (End1), *E. coli* C600[pES530] (Ced1), and *E. coli* C600[pUC19] (control) cultures were analyzed by SDS-PAGE (Fig. 5.3). There was no visible difference between the protein profiles of *E. coli* C600[pES420] and the control *E. coli* C600[pUC19], but the *E. coli* C600[pES530] protein profile showed the presence of a major protein band with an apparent M_r of approximately 61 000. This major band was also present in the concentrated supernatant from *E. coli* C600[pES530] cultures (Fig. 5.3), but a similar band was not present in the *B. fibrisolvens* CFE.

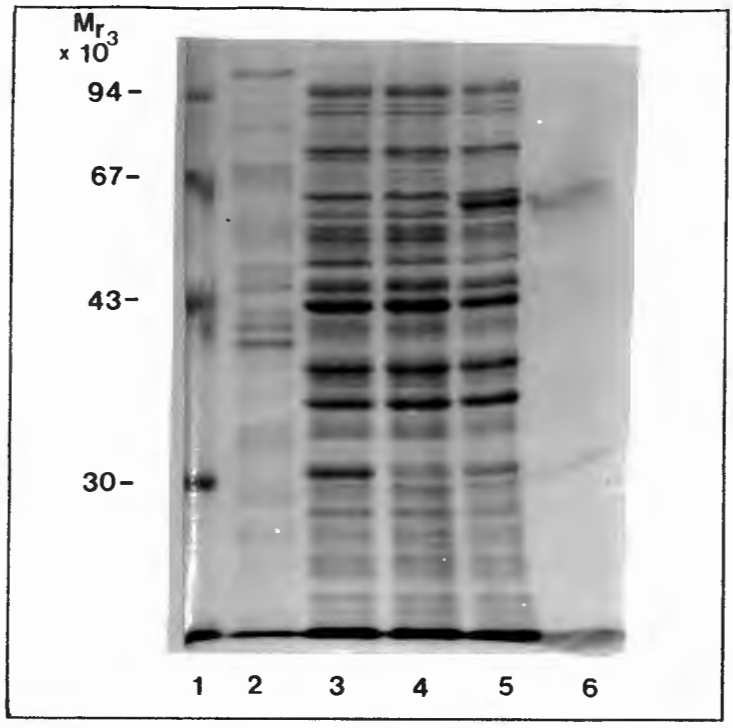


Fig. 5.3 SDS-PAGE analysis. Lane 1, protein standards; lane 2, *B. fibrisolvens* CFE (30 µg prot.); lane 3, *E. coli* C600[pUC19] CFE (50 µg); lane 4, *E. coli* 600[pES420] CFE (50 µg); lane 5, *E. coli* C600[pES530] CFE (50 µg); lane 6, *E. coli* 600[pES530] concentrated supernatant (15 µg). Protein standards were phosphorylase b (94 000), albumin (67 000), ovalbumin (43 000) and carbonic anhydrase (30 000).

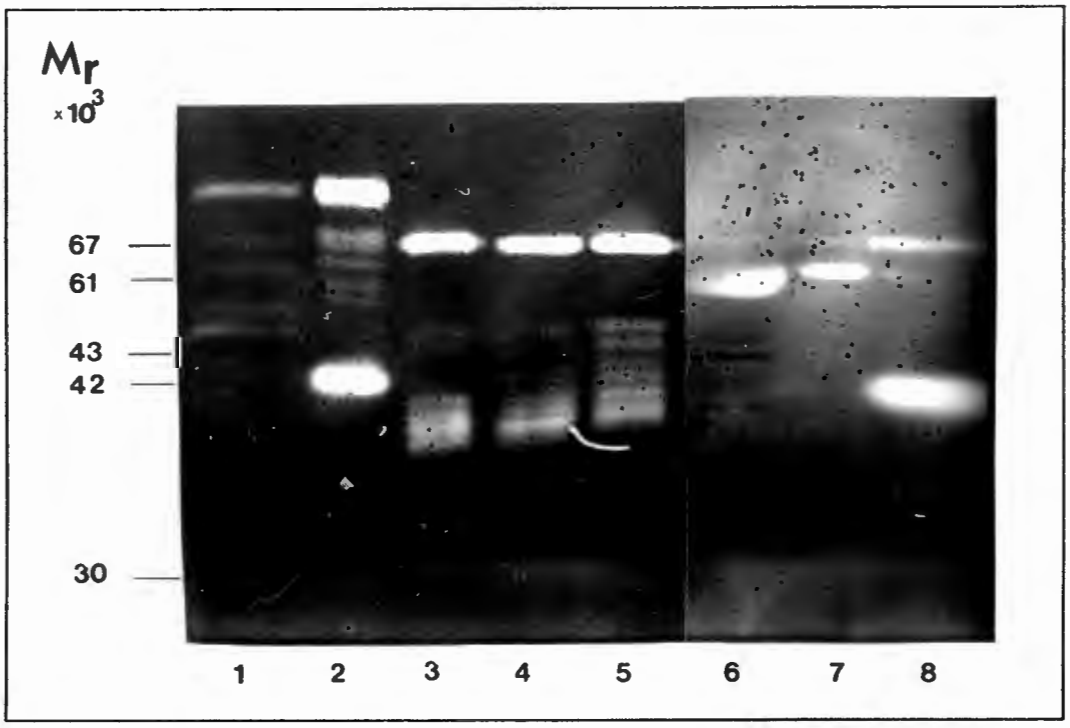


Fig. 5.4 Zymogram from different *B. fibrisolvens* and *E. coli* cell extracts. Lane 1, *B. fibrisolvens* supernatant (2 µg); lane 2, *B. fibrisolvens* CFE (5 µg); lane 3, *E. coli* C600[pES420] CFE (6 µg); lane 4, *E. coli* C600[pES420] periplasmic extract (6 µg); lane 5, *E. coli* C600[pES420] supernatant (3 µg); lane 6, *E. coli* C600[pES530] CFE (50 µg); lane 7, *E. coli* C600[pES530] periplasmic extract (30 µg); lane 8, *E. coli* C600[pES530] concentrated supernatant (20 µg).

Activity of the proteins in cell-free, periplasmic, and supernatant extracts from *B. fibrisolvens*, *E. coli* C600[pES420], and *E. coli* C600[pES530] were analyzed by SDS-PAGE with incorporated CMC (Fig. 5.4). A comparison of lanes containing the *B. fibrisolvens* supernatant (lane 1) and CFE (lane 2) showed a number of different bands. From these results it was impossible to speculate on the number of different cellulases produced by *B. fibrisolvens* H17c. Activity bands in the supernatant might be due to degradation products or protein processing during secretion into the supernatant. It is interesting to note that there were no major activity bands in the *B. fibrisolvens* extracts corresponding to that of the *E. coli* extracts containing the End1 or Ced1 enzymes (lanes 3, 5, 6, and 8). *E. coli* C600 containing the endoglucanase End1 cell-free (lane 3), periplasmic (lane 4), and supernatant (lane 5) extracts showed the presence of a major activity band with an apparent M_r of approximately 67 000. *E. coli* C600 containing Ced1 cell-free (lane 6) and periplasmic (lane 7) extracts showed a major activity band with an apparent M_r of approximately 61 000 whereas the concentrated supernatant showed a major activity band with an apparent M_r of 42 000.

5.3.4 Substrate specificity of *B. fibrisolvens* End1 and Ced1 as expressed in *E. coli* C600 cell-free extracts. The enzyme activities of the *B. fibrisolvens* End1 and Ced1 enzymes were measured against a range of substrates (Table 5.2). The enzyme in CFE of *E. coli* C600 harboring *end1*, was very active against lichenan, a mixed linkage polysaccharide (β1-4,β1-3), and CMC (β1-4). Activity was also observed against the soluble fraction of oat spelt xylan and a low level of activity was observed against laminarin (β1-3). The enzyme released reducing sugar during a prolonged incubation with recalcitrant forms of cellulose (Avicel and filter paper), but the level of activity was very low (Table 5.2). End1 was able to cleave the aglucone bond in pNPC suggesting that End1 had some exoglucanase activity. The enzyme showed very little activity against the aryl-glucosides pNPG and Salicin.

The enzyme in *E. coli* C600 containing *ced1* CFE showed a different activity profile to that of End1. CMCase activity of Ced1 was approximately 50% of that of End1 (Table 5.2) and pNPC activity of Ced1 was 3-fold higher than End1, indicating that Ced1 is a more exo-acting enzyme than End1. Background activity using *E. coli* cultures as controls was negligible.

Table 5.2 Activity of CFE from *E. coli* C600 containing *B. fibrisolvens* End1 or Ced1 against various substrates

Substrate (major linkage)	Specific activity (units/mg prot.) End1	Ced1
Filter paper	0.0007	0.0006
Avicel	0.0003	0.0005
Carboxymethylcellulose (β1-4)	0.881	0.38
Lichenan (β1-4\β1-3)	1.232	0.08
Laminarin (β1-3)	0.003	0.01
Oat spelt xylan (β1-4)	0.021	0.03
Salicin (β1-4)	<0.0001	<0.0001
p-nitrophenyl-β-D-cellobioside (β1-4)	0.05	0.14
p-nitrophenyl-β-D-xyloside (β1-4)	<0.0001	<0.0001
p-nitrophenyl-β-D-glucoside (β1-4)	<0.0001	<0.0001

The kinetics of viscosity decrease versus reducing sugar release from CMC was determined for the Ced1 and End1 enzymes (Fig. 5.5). The decrease of viscosity was less when CMC was treated with Ced1 than with End1. The gradient for Ced1 was 5.8×10^{-3} and for End1 1.4×10^{-2}, indicating that End1 acts in a more random fashion than Ced1 (Bhat et al. 1989).

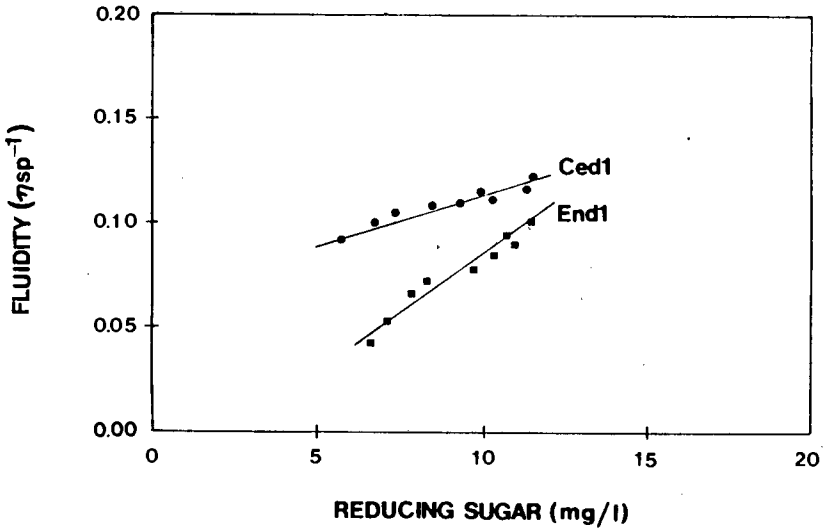


Fig. 5.5 Plot of increase in fluidity *versus* the release of reducing sugars for the hydrolysis of CMC by *B. fibrisolvens* End1 and Ced1 enzymes. End1, (■); Ced1, (●).

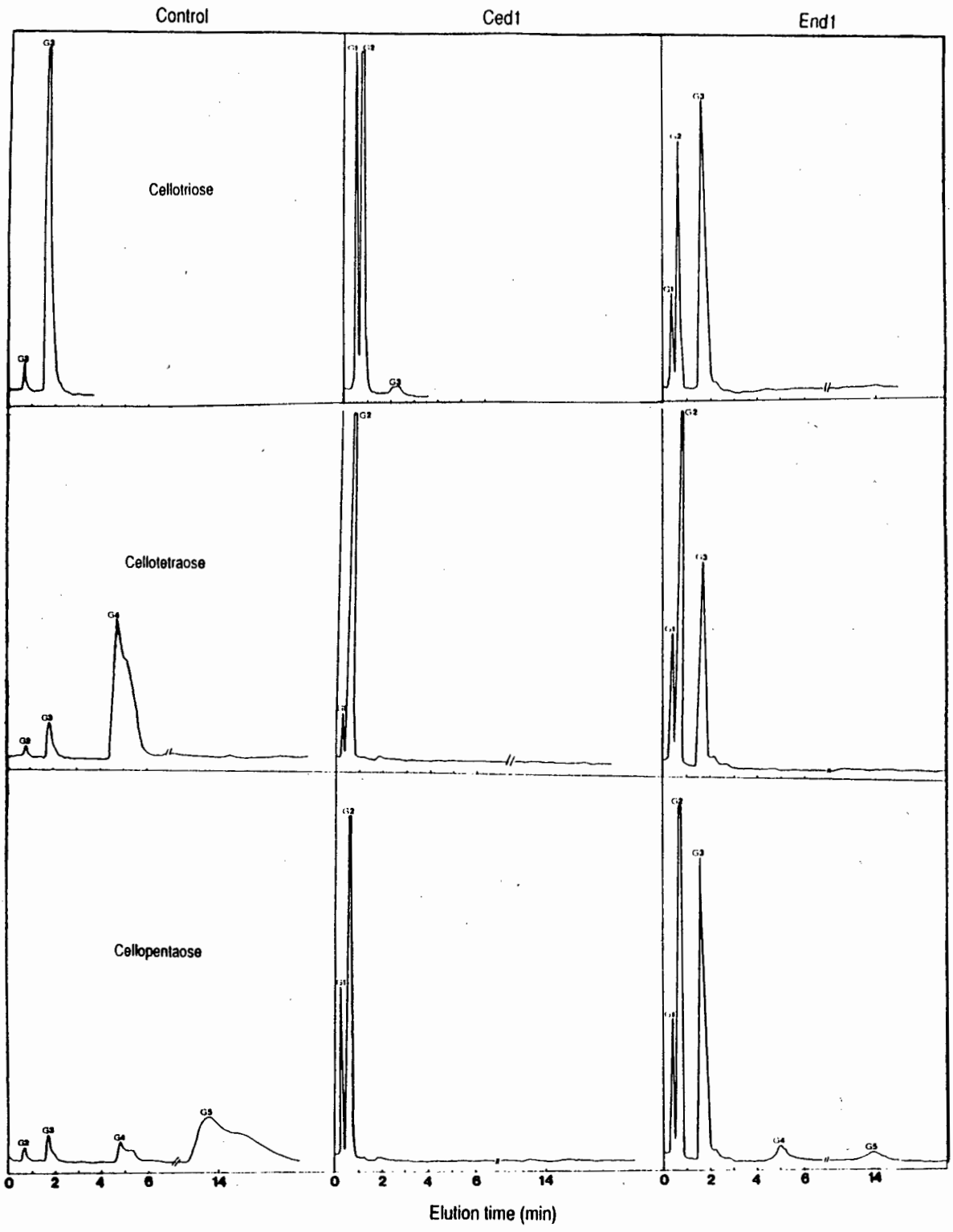


Fig. 5.6 Products of cellodextrins hydrolyzed by *B. fibrisolvens* End1 and Ced1. The enzymes were incubated with cellodextrins (1%) and incubated at 45°C for 24 h (End1) and 5 min (Ced1). The products produced were determined by HPLC. Abbreviations: G5, cellopentaose; G4, cellotetraose; G3, cellotriose; G2, cellobiose; G1, glucose.

Hydrolysis of cellodextrins by Ced1 and End1 was analyzed in a HPLC system (Fig. 5.6). Analysis of the degradation of cellotetraose and cellohexaose by the Ced1 enzyme revealed that cellobiose units were rapidly released and accumulated as the final degradation product. Cellotriose and cellopentaose were hydrolyzed by Ced1 to cellobiose and glucose. Cellobiose was not further degraded. The Ced1 enzyme completely degraded the cellodextrins within 5 min whereas under the same conditions the End1 enzyme was not able to degrade the cellodextrins completely over 24 h. The End1 enzyme degraded cellotetraose, cellopentaose and cellohexaose randomly to cellotriose, cellobiose and glucose. Cellotriose was degraded at a much slower rate to cellobiose and glucose, and End1 was not able to hydrolyze cellobiose.

Rasmussen et al. (1988) showed that the addition of methylcellulose to the assay mixture inhibited endoglucanase but not cellodextrinase activity. When methylcellulose was added (0.04 and 0.4% [wt/v], 400 cP) to the pNPC assay, activity was inhibited by 1.5 and 8.0% respectively for Ced1 and 59 and 63% respectively for End1.

5.3 Discussion

Localization studies showed that the majority of enzyme activity (64 and 55%, respectively) of *B. fibrisolvens* End1 and Ced1 enzymes expressed in *E. coli* was in the periplasm. This indicates that export of the cloned product through the cytoplasmic membrane occurs. Kopecny (1986) reported that the cellulases of *B. fibrisolvens* UC142 comprised extracellular and cell-associated enzymes. End1 was shown to contain a functional signal peptide when expressed in *E. coli,* (Chapter 3.3.5) whereas the amino terminal end of the Ced1 polypeptide did not contain a typical procaryotic signal peptide which was able to facilitate the secretion of *phoA* in *E. coli* (Chapter 4.3.5). Other examples of cloned gene products in *E. coli* which involve export across the cytoplasmic membrane and where typical signal peptides are absent are the *V. alginolyticus* sucrase enzyme (Scholle et al. 1989) and the *R. flavefaciens* CelA cellodextrinase enzyme (Wang and Thomson 1990). The sucrase enzyme was localized entirely in the cytoplasm of the parental *V. alginolyticus,* whereas in *R. flavefaciens* cellulase activity was located extracellularly, either cell-associated or in the supernatant fluid (Pettipher and Latam 1979). However, Gong et al. (1989) found that the *B. succinogenes* periplasmic cellodextrinase was located intracellularly when the gene was cloned and expressed in *E. coli.* Sequence analysis of more cellulases from rumen bacteria may lead to an understanding of how these proteins are secreted.

Post-translational modification, such as glycosylation, of the cellulase enzymes in *B. fibrisolvens* but not in *E. coli* may explain why there were no corresponding activity bands between the *B. fibrisolvens* extracts and the extracts obtained from *E. coli* C600 containing either End1 or Ced1 (Fig. 5.4). Glycosylation was shown for the *C. fimi* CenA and Cex cellulases (Gilkes et al. 1984a, Langsford et al. 1987), and for some of the *C. thermocellum* cellulases (Morgenstern et al. 1987). If an enzyme was glycosylated in *B. fibrisolvens* then such an enzyme would migrate at a higher position than that of its counterpart expressed in *E. coli.* The major activity band with apparent M_r of approximately 67 000 observed in *E. coli* C600 containing End1 cell-free, periplasmic, and supernatant extracts was approximately 6 kDa larger than the calculated M_r of 61 077 (Chapter 3.3.4). This aberrant migration of End1 was

also found for the CenA endoglucanase of *C. fimi* (Gilkes et al. 1988) when expressed in *E. coli*. The apparent M_r of the CenA protein was 4.9 kDa larger than predicted by *cen*A. Gilkes et al. (1989) found that the P-T box region was responsible for the aberrant migration of CenA during SDS-PAGE. A 33-amino acid segment containing related sequences ($(E\text{-}P)_4$ and $(K\text{-}P)_6$) was responsible for similar behavior in the *E. coli* TonB protein, due to its highly extended sequence-imposed structure (Postle and Good 1983, Evans et al. 1988). Gilkes et al. (1989) then proposed that it was probable for the P-T box of the CenA protein to adopt an extended conformation which was responsible for the slower migration of the protein. The *B. fibrisolvens* End1 protein contains a 35-amino acid P-T box (Chapter 3.3.6) which may be responsible for the aberrant migration of the End1 protein during SDS-PAGE. The cell-free and periplasmic extracts of *E. coli* C600 containing Ced1 showed a major activity band with an apparent M_r of 61 000 which is in excellent agreement with the calculated M_r of 61 023 (Chapter 4.3.3). The concentrated supernatant extract from *E. coli* C600 expressing the Ced1 protein showed a major activity band with an apparent M_r of 42 000 and only a faint activity band at the expected M_r of 61 000. Cellulases can be very difficult to denature and "false" bands can be observed which migrate faster than expected. This problem is further enhanced by the fact that partially denatured enzymes renature preferentially to the totally renatured form, giving rise to activity bands which are not representative of their actual abundance. To test this theory a fully denatured supernatant sample was investigated after normal SDS-PAGE. The protein profile obtained from the concentrated supernatant sample (Fig. 5.3) had a major protein band with an apparent M_r of 61 000 as expected. There was only a very faint protein band present with an apparent M_r of 42 000 and this result confirmed that the activity band (Fig. 5.4) with an apparent M_r of 42 000 was most likely a "false" band. The concentration of the supernatant with an Amicon apparatus may have resulted in the concentration of other components which prevented the proper denaturation of the sample.

The substrate specificity profile of End1 was similar to that of endoglucanase EGA cloned from *B. fibrisolvens* strain A46 (Hazlewood et al. submitted) and the Cel endoglucanase from *F. succinogenes* (Taylor et al. 1987). Activity against soluble

substrates lichenan and CMC were relatively high, whereas only low levels of activity were observed against laminarin and oat spelt xylan. All three enzymes were able to cleave the aglucone bond of pNPC but showed very little or no activity against the aryl-glucoside pNPG and crystalline substrates such as Avicel and filter paper.

The *B. fibrisolvens* Ced1 enzyme rapidly released cellobiose units from cellodextrins. The final product from the degradation of cellotetraose and cellohexaose was cellobiose, and cellotriose and cellopentaose were degraded to cellobiose and glucose. Cellobiose was not hydrolyzed further. Ced1 was shown to release cellobiose from pNPC but activity against the crystalline substrates Avicel and filter paper was very low. These results suggested that Ced1 has cellodextrinase activity. The enzyme specificity profile was similar to that of a cellodextrinase purified (Huang and Forsberg 1988) and cloned (Gong et al. 1989) from *F. succinogenes*. The Ced1 enzyme was also similar to the *R. flavefaciens* CelA cellodextrinase. *R. flavefaciens* CelA released predominantly cellobiose from cellodextrins, has high activity on pNPC (Thomson, personal communication), and was isolated on CMC LB agar (Barros and Thomson 1987). Methylcellulose did not inhibit the *B. fibrisolvens* Ced1 cellodextrinase whereas it inhibited the *B. fibrisolvens* End1 endoglucanase providing further evidence that Ced1 is a cellodextrinase and End1 an endoglucanase.

Chapter 6

Detection and characterization of extracellular proteases in *Butyrivibrio fibrisolvens* H17c

Chapter 6

Detection and characterization of extracellular proteases in Butyrivibrio fibrisolvens H17c

6.0 Summary

The proteolytic activity of *B. fibrisolvens* H17c was characterized using gelatin-SDS-PAGE. Nine bands of protease activity with apparent M_r of approximately 101 000, 95 000, 87 000, 80 000, 76 000, 68 000, 63 000, 54 000, and 42 000 were detected after gelatin-SDS-PAGE of supernatants from exponential phase cultures. A tenth protease band with an apparent M_r of 32 000 was detected in stationary phase cells. The effect of protease inhibitors indicated that all ten protease bands were serine proteases. No precise optimal pH was observed, and the highest activities for all ten protease bands were in the pH range 6.0 to 7.5. The proteolytic activity of *B. fibrisolvens* H17c varied depending on the type of carbohydrate substrate in the medium and was positively correlated with the growth rate. However, a similar correlation was not observed with the addition of different nitrogen sources. The highest protease activity was obtained with the lowest growth rate which occurred with casamino acids as nitrogen source.

6.1 Introduction

In the rumen, proteins from ingested feedstuffs are extensively degraded (60-90%) along with plant cell polysaccharides (Leng and Nolan 1984). Information available on protein breakdown in the rumen has been mostly obtained from measurements of ruminal contents directly, or of mixed bacterial populations obtained from the rumen contents. Since these data represent the combined effects of many enzymes produced by a variety of microorganisms, available information is biochemically imprecise and difficult to interpret. However, such studies have shown that ruminal contents have several types of proteases present and that most proteolytic activity is associated with the bacteria in the rumen (Brock et al. 1982).

Important proteolytic species in the rumen are *B. amylophilus, B. ruminicola,* and *B. fibrisolvens.* The proteolytic activity of *B. amylophilus* has been studied in some detail (Blackburn 1968a; Blackburn 1968b). The proteolytic activity produced during exponential growth was cell-associated (80%) and cell-free (20%), but the cell-free activity increased in stationary phase cultures probably due to lysis of the cells. The proteolytic activity may consist of one or two trypsin-like proteases. *B. ruminicola* was shown to have proteolytic activity that was maximal and mostly (>90%) cell-associated during the mid-exponential growth phase (Hazlewood and Edwards 1981; Hazlewood et al. 1981). This proteolytic activity was sensitive to oxygen and may consist of a mixture of serine, thiol, and carboxyl proteases (Hazlewood and Edwards 1981). In contrast to the previous two organisms 95% or more of the proteolytic activity of *B. fibrisolvens* was found to be associated with the supernatant fluid and not the cells, and this enzyme activity was not sensitive to oxygen (Cotta and Hespell 1986). However, the above studies all dealt with overall proteolytic activity of these organisms. Since *B. fibrisolvens* is one of the predominant proteolytic bacterial species in the rumen (Cotta and Hespell 1986) the purpose of this study was to examine the variety and nature of extracellular proteases produced by *B. fibrisolvens* H17c using a gelatin-SDS-PAGE technique.

6.2 Materials and Methods

6.2.1 Bacterial strain, media, and growth conditions. The proteolytic *B. fibrisolvens* H17c strain has been described in Chapter 2.2.1. *B. fibrisolvens* H17c was maintained on non-rumen fluid medium 10 (M10) of Caldwell and Bryant (1966) as modified by Strydom et al. (1986) (Appendix A). To study the effect of different nitrogen compounds on the proteolytic activity of *B. fibrisolvens* H17c the defined medium of Cotta and Hespell (1986) was used as the minimal medium (MM) (Appendix A) supplemented with the different nitrogen compounds. Samples (1 ml) of overnight modified M10 cultures were centrifuged, washed with one-quarter strength Ringer solution, and resuspended in 10 ml of either modified M10 or MM and incubated at 37°C. Samples (1 ml) were removed at time intervals, centrifuged in an Eppendorf microfuge and the supernatants were assayed for protease activity.

All *B. fibrisolvens* cultures were grown anaerobically in batch cultures at 37°C. Growth was monitored spectrophotometrically by determining the optical densities of cultures at 600 nm.

6.2.2 Protease assays

6.2.2.1 Azocasein assay. Extracellular protease activity was assayed using the synthetic substrate azocasein (Sigma) at a concentration of 0.8% w/v in a phosphate buffer (0.1 M, pH 6.8) according to Brock et al. (1982). Supernatant samples (0.5 ml) were added to the azocasein solution (0.5 ml) and the mixture was incubated at 55°C for 3 h. The reaction was terminated by the addition of cold $HClO_4$ (1.5 M, 0.5 ml) and after 30 min at 4°C the precipitated protein was removed by centrifugation in an Eppendorf microfuge and the supernatant fluid (1 ml) was added to an equal volume of NaOH solution (0.1 M). The concentration of acid-soluble azopeptides in the resultant solution was determined spectrophotometrically at 440 nm. One optical density unit was equal to 320 µg/ml (Cotta and Hespell 1986). One unit of protease activity equaled 1 µg of azocasein digested per hour under the above conditions.

6.2.2.2 Gelatin-SDS-PAGE assay. Extracellular proteases produced by *B. fibrisolvens* H17c were characterized by PAGE in slabs containing SDS and gelatin (0.1% w/v) as a copolymerized substrate (Heussen and Dowdle 1980). Cultures were sedimented by centrifugation in an Eppendorf microfuge and supernatant samples (1 ml) were mixed with SDS (25% w/v, 0.1 ml) and glycerol (0.1 ml) and incubated at 37°C for 30 min. PAGE was carried out at 4°C in SDS-gelatin-polyacrylamide gels (pH 6.8) at a constant voltage (70 V) for 17 h. Each lane was loaded with 10 µl of sample from early exponential phase cultures, 5 µl from late exponential phase cultures, and 2.5 µl from stationary phase cultures. After electrophoresis the gels were washed in Triton X-100 (2.5% w/v) for 1 h at 4°C to remove the SDS and restore enzyme activity. After incubation in phosphate buffer (0.1 M, pH 6.8) for 3.5 h at 55°C bands of proteolytic activity were detected after staining with amido black (0.2% w/v). The proteolytic activity bands were revealed as clear areas which lacked gelatin. The M_r markers were obtained from BRL. Treated supernatant samples could be stored at -20°C for several months without noticeable loss of protease activity.

The detection of protease band profiles was also done after electrophoresis in standard SDS-PAGE gels (Laemmli 1970) which were overlaid with agarose gels (1%) containing gelatin (0.1%). After incubation the overlay (agarose-gelatin) gel was stained as above.

6.2.3 The effect of pH on protease activity. Early stationary phase *B. fibrisolvens* H17c cultures were used to test the effect of pH on extracellular protease activity. After gelatin-SDS-PAGE the gels were incubated in 0.1 M citric acid buffer (pH 4.5-5.5), 0.1 M phosphate buffer (pH 6.0-8.0), or glycine buffer (pH > 8.0) and the effect on the protease bands were investigated after staining.

6.2.4 Treatment of supernatants with protease inhibitors. The effect of different protease inhibitors on the activity of the proteases produced by *B. fibrisolvens* H17c were investigated. *B. fibrisolvens* H17c was grown to early stationary phase in modified M10 medium. The inhibitors were added to the gelatin gels during the Triton X-100 wash and the incubation step in phosphate buffer (0.1 M, pH 6.8). The

inhibitors used were phenylmethylsulphonyl fluoride (PMSF) (10 mM) dissolved in dimethylsulfoxide (DMSO), *p*-hydroxymercuribenzoate sodium salt (pHMB) (2.5 mM), tosyl-lysine chloromethyl ketone (TLCK) (1 mM), tosyl-phenyl-alanine chloromethyl ketone (TPCK) (1 mM), EDTA (50 mM), *o*-phenanthroline (2.5 mM), and soybean trypsin inhibitor (1 mM). Inhibition of protease activity by the divalent cations from $CaCl_2$ (1.25 mM), $MgSO_4$ (1.25 mM), and $ZnSO_4$ (1.25 mM), was also investigated. pHMB, TLCK, TPCK, and *o*-phenanthroline were obtained from Sigma, and PMSF was obtained from Schwarz Mann, Orangeburg, New York, USA.

6.2.5 Addition of different carbohydrates and nitrogen compounds to media. Different carbohydrates (cellobiose, fructose, galactose, glucose, maltose, mannose, rhamnose, sucrose, and xylose) were added to modified M10 medium (1% w/v) as carbon sources, and the effect on growth and protease activity of *B. fibrisolvens* H17c was determined. Cells were harvested at the end of the logarithmic growth phase and the supernatant was used for azocasein protease assays.

Growth and protease activity of *B. fibrisolvens* H17c cultured in glucose MM supplemented with different nitrogen compounds was investigated. The nitrogen compounds used were trypticase (0.1 and 2% w/v), casamino acids (0.1 and 2% w/v), NH_3 (45 mM), urea (45 mM), and $(NH_4)_2SO_4$ (45 mM).

6.3 Results

6.3.1 Characterization of proteases by gelatin-SDS-PAGE. The production of extracellular proteases by *B. fibrisolvens* H17c in modified M10 medium was determined using the gelatin-SDS-PAGE technique. Supernatant samples were obtained from different stages in the growth cycle and the proteins were separated on 10% acrylamide gels containing gelatin (Fig. 6.1).

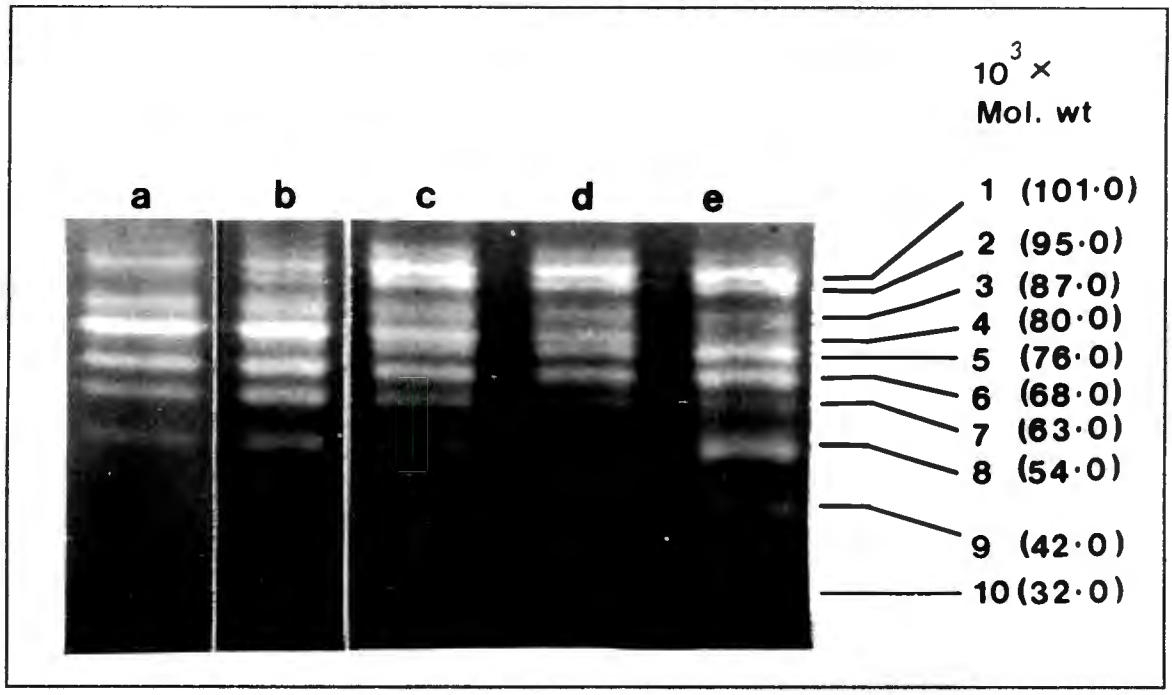


Fig. 6.1 Production of extracellular proteases in modified M10 medium during the growth cycle of *B. fibrisolvens* H17c as determined by gelatin-SDS-PAGE. Protease production was determined at the end of the lag growth phase - 0.15 OD_{600} (*a*), early exponential growth phase - 0.20 OD_{600} (*b*), mid exponential growth phase - 0.45 OD_{600} (*c*), late exponential growth phase - 0.60 OD_{600} (*d*), and stationary growth phase - 0.68 OD_{600} (*e*). The lines indicate the positions of protease bands 1 - 10 with their apparent M_r.

Six extracellular protease bands with apparent M_r of approximately 101 000, 87 000, 80 000, 68 000, 63 000, and 54 000 (protease bands 1, 3, 4, 6, 7, and 8, respectively) were produced by cells harvested during the early exponential growth phase (0.15, OD_{600}). During the late exponential phase (0.60, OD_{600}) nine extracellular protease bands were detected. These included the six protease bands which were present in

the early exponential growth phase and three additional protease bands with apparent M_r of approximately 95 000, 76 000, and 42 000 (protease bands 2, 5, and 9, respectively). In stationary phase cultures (0.68, OD_{600}) nine extracellular protease bands were detected. Protease band 4 (apparent M_r 80 000) disappeared and an additional protease band with an apparent M_r of 32 000 (protease band 10) was observed. The stationary phase extracellular protease band profile did not change in cultures incubated at 37°C for a further 24 h. Similar protease band profiles as shown in Fig. 6.1 were obtained with the overlay electrophoresis technique.

6.3.2 Influence of pH and temperature on protease activity produced by *B. fibrisolvens* H17c. Both the assay temperature and pH influenced the proteolytic activity of the supernatant fluid. Increasing the assay temperature from 30°C to 50°C resulted in a two-fold increase of proteolytic activity. At assay temperatures above 55°C there was a rapid decrease in activity with an increase in temperature, with less than 10% of the optimum activity (50°C) left at 70°C. The influence of pH on extracellular protease activity from *B. fibrisolvens* H17c after gelatin-SDS-PAGE was investigated by incubation of the gels in the appropriate buffers to give a pH range between pH 4.5 and pH 8.0. The activity of all ten protease bands was optimal between pH 6.0 and pH 7.5 (Fig. 6.2).

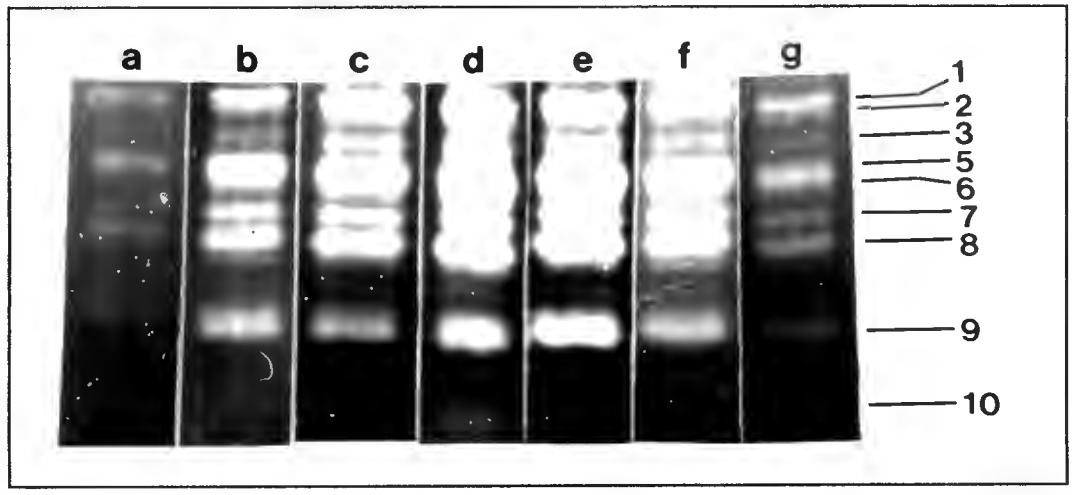


Fig. 6.2 Influence of pH on the activities of the extracellular protease bands produced by *B. fibrisolvens* H17c after gelatin-SDS-PAGE. The pH of the buffer solutions was: pH 4.5 (*a*), pH 5.5 (*b*), pH 6.0 (*c*), control pH 6.8 (*d*), pH 7.0 (*e*), pH 7.5 (*f*), and pH 8.0 (*g*)

6.3.3 Effect of protease inhibitors on protease activity in *B. fibrisolvens* H17c. A range of compounds known to inhibit various proteases was tested for their effects on the extracellular protease bands of *B. fibrisolvens* H17c after gelatin-SDS-PAGE (Fig. 6.3). All the protease bands were inhibited by the serine protease inhibitor PMSF which reacts with the hydroxyl group of a serine residue in the active site of the enzyme. Control gels showed that the organic solvent DMSO, used to dissolve the PMSF, did not affect the activities of the protease bands. No inhibition was found with TPCK or TLCK, inhibitors of serine proteases with trypsin- or chymotrypsin-like specificities. The protease bands were not inhibited by the sulphydryl reagent pHMB or by inhibitors of metallo- and carboxyl-proteases (EDTA and *o*-phenanthroline). The divalent cations from $CaCl_2$, $MgSO_4$, and $ZnSO_4$ did not affect the protease bands (results not shown).

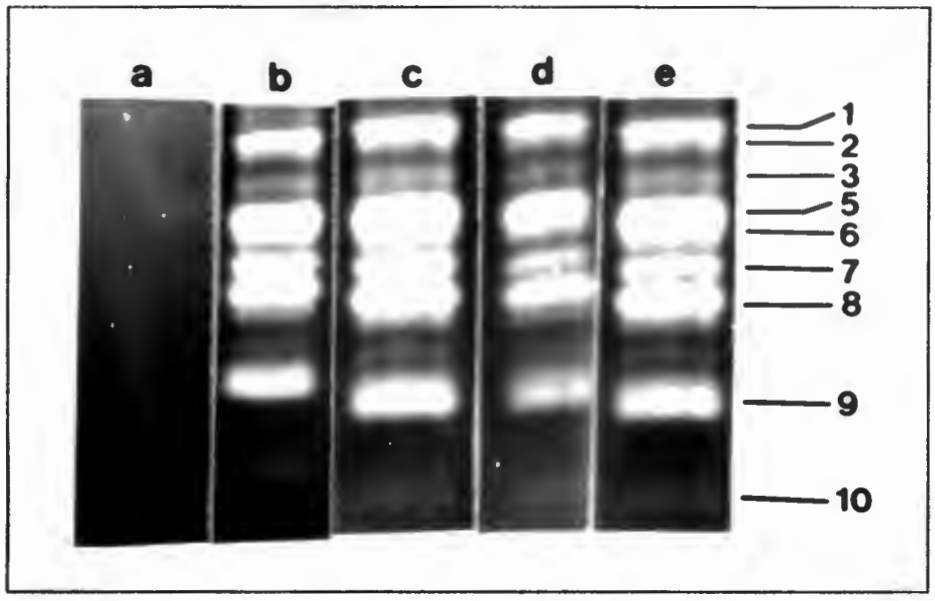


Fig. 6.3 Effect of inhibitors on the activities of the extracellular protease bands produced by stationary phase *B. fibrisolvens* H17c culture supernatants after gelatin-SDS-PAGE. Inhibitors added were: PMSF (10 mM) (*a*), *o*-phenanthroline (2.5 mM) (*b*), control without inhibitor (*c*), TLCK (1 mM) (*d*), and pHMB (2.5 mM) (*e*).

6.3.4 Effect of carbohydrates on the production and activity of the extracellular proteases produced by *B. fibrisolvens* H17c. The effect of various carbohydrates on the growth of *B. fibrisolvens* H17c and the activity of extracellular proteases produced was investigated. *B. fibrisolvens* H17c was able to grow on a wide range of carbohydrates that served as energy and carbon sources. The nature of the carbohydrate in the medium affected the growth rate of the cells and there was a tendency for rapidly growing cells to exhibit higher levels of proteolytic activity. Good growth supported high levels of protease activity (Table 6.1)

Table 6.1 Effect of carbohydrates (1% w/v) on the growth and proteolytic activity of *B. fibrisolvens* H17c

Carbohydrate	OD_{600}[a]	Activity[b]
Glucose	0.63	21.30
Cellobiose	0.64	18.11
Maltose	0.65	20.00
Sucrose	0.63	18.13
Fructose	0.55	13.87
Galactose	0.42	12.80
Mannose	0.12	0.85
Xylose	0.10	0.21
Rhamnose	0.10	0.32

[a] At time of harvest
[b] Activity in supernatant. One unit of protease activity equals 1 µg/h of azocasein digested. Values are from duplicate samples from two separate experiments.

6.3.5 Influence of nitrogen compounds on azocasein protease activity and growth of *B. fibrisolvens* H17c. Cotta and Hespell (1986) reported that 95% or more of the proteolytic activity of *B. fibrisolvens* strains (including *B. fibrisolvens* H17c) was in the supernatant fluid regardless of the stage of growth, and was not cell associated as was reported for *Bacteroides* species (Blackburn 1968a). The azocasein protease activity of supernatant samples from late exponential *B. fibrisolvens* H17c cultures grown in glucose MM supplemented with different nitrogen sources

was determined. The nitrogen source added to the MM had a marked effect on the growth of *B. fibrisolvens* H17c. Growth occurred in MM supplemented with casamino acids or trypticase but MM containing NH_3, urea, or $(NH_4)_2SO_4$ as sole nitrogen sources was unable to sustain growth of *B. fibrisolvens* H17c (Table 6.2).

Table 6.2 Effect of different nitrogen compounds on the growth and proteolytic activity of *B. fibrisolvens* H17c

Nitrogen source	Conc.	OD_{600}[a]	Activity[b]
Trypticase	0.1% (w/v)	0.36	5.12
	2.0% (w/v)	0.72	16.32
Casamino acids	0.1% (w/v)	0.36	5.9
	2.0% (w/v)	0.60	19.2
NH_3	45.0 (mM)	0.10	0.32
Urea	45.0 (mM)	0.10	0.33
$(NH_4)_2SO_4$	45.0 (mM)	0.05	0.11

[a] At time of harvest
[b] Activity in supernatant. One unit of protease activity equals 1 μg/h of azocasein digested. Values are from duplicate samples from two separate experiments.

Azocasein protease activity was determined throughout the growth cycle of *B. fibrisolvens* H17c cultures grown in modified M10 and in glucose MM supplemented with trypticase or casamino acids as nitrogen sources. In all three media extracellular azocasein protease activity was produced during the exponential growth phase. Azocasein protease activity did not accumulate during the stationary growth phase. Although the slowest growth rate and the lowest cell yield was obtained in MM containing casamino acids, this medium showed the highest azocasein protease activity (Fig. 6.4).

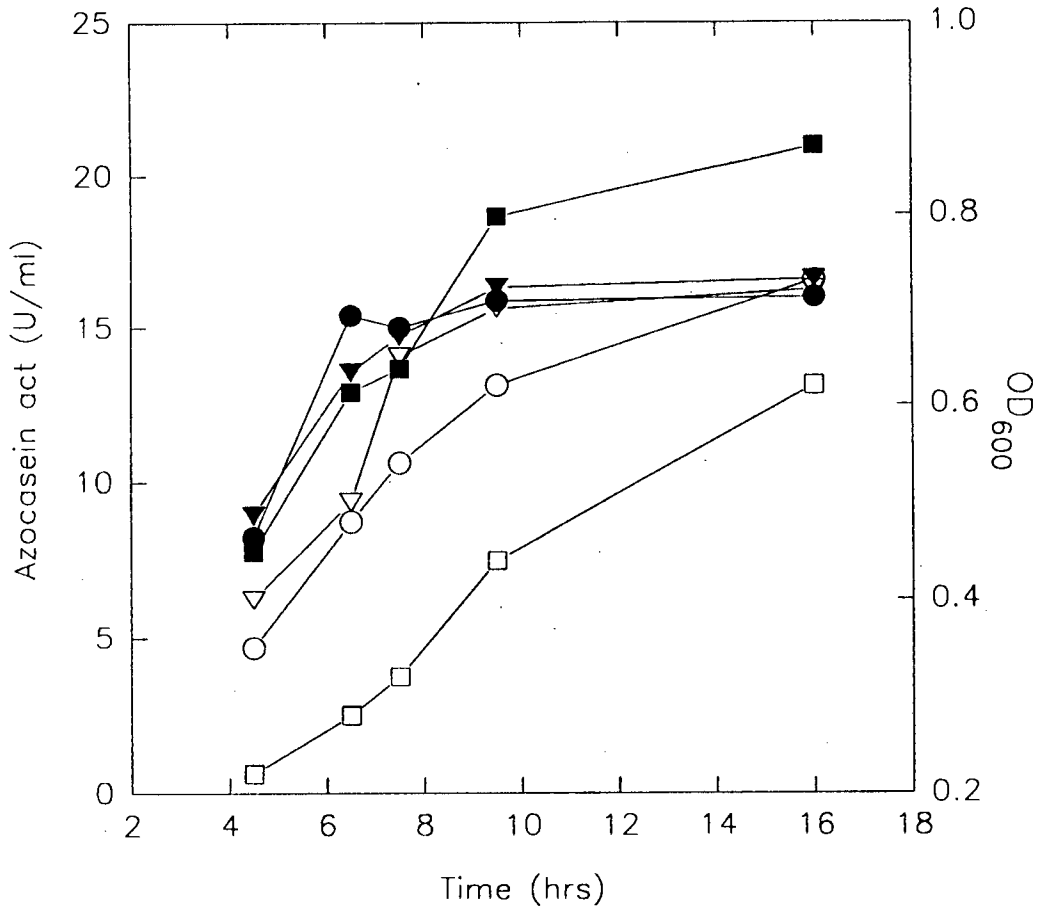


Fig. 6.4 Growth and production of extracellular azocasein protease activity by *B. fibrisolvens* H17c in modified M10 and glucose MM supplemented with casamino acids or trypticase. Growth in M10 (○), MM + casamino acids (□), and MM + trypticase (▽). Extracellular azocasein protease activity in M10 (●), MM + casamino acids (■), MM + trypticase (▼).

6.4 Discussion

The extracellular proteases produced by *B. fibrisolvens* H17c were characterized by gelatin-SDS-PAGE and ten bands of protease activity were detected. Nine of these protease bands were detected during the exponential growth phase and a protease band with the lowest apparent M_r of approximately 32 000 (protease band 10) was only detected once the cultures entered the stationary growth phase. Protease band 5 (M_r of approximately 76 000) appeared during late exponential phase and there was a concomitant disappearance of protease band 4 (M_r of approximately 80 000). The majority, therefore, of the protease activities of the various bands were relatively stable and did not change as the culture entered the stationary phase and after incubation for a further 24 h.

A feature of all the protease activity bands was their similar properties. The activities of all ten protease bands were inhibited by the serine protease inhibitor PMSF but not by TPCK or TLCK indicating that the *B. fibrisolvens* H17c protease activity bands represented serine proteases which lacked trypsin- or chymotrypsin-like specificities similar to that of the *B. subtilis* extracellular subtilisin proteases which are serine proteases and are not inhibited by TLCK or TPCK. Similar results were obtained for *B. fibrisolvens* 49 (Cotta and Hespell 1986).

The ten protease bands required similar conditions for optimal activity, suggesting that the bands of protease activity represented isoenzymes with different M_r. However, it was impossible to conclude from the gelatin-SDS-PAGE characterization of the extracellular proteases from *B. fibrisolvens* H17c whether the different M_r serine protease bands were different gene products or degradation products of one or a few gene products. The similarity in the enzymes did though favour the latter possibility. The different M_r serine protease bands were not due to anomalous migration patterns in the gelatin-SDS-PAGE technique, which involved electrophoresis of the SDS inactivated proteases in a supposedly immobilized substrate. Similar protease band profiles were obtained in overlay gels after electrophoresis without the gelatin substrate.

The proteolytic activity of *B. fibrisolvens* H17c was produced constitutively, but the activity level was subject to modulation by growth conditions. Growth on different carbohydrates could be positively correlated with proteolytic activity, whereas the addition of different nitrogen sources to glucose MM resulted in different levels of protease activity produced (Fig 6.1). Cotta and Hespell (1986) obtained similar results from *B. fibrisolvens* 49 and suggested that exogenous amino acids or peptides may also influence proteolytic activities. By comparison, the proteolytic activity of *B. amylophilus* was not induced or repressed by the presence of a variety of nitrogen sources or other nutrients in the medium, but the production of proteolytic activity did vary with the growth rate (Blackburn 1968a; Blackburn 1968b).

Members of the genus *Butyrivibrio* occur widely in nature and the strains show great phenotypic variability (Mannarelli 1988). *B. fibrisolvens* H17c and *B. fibrisolvens* 49 were isolated from different sources (Bryant and Small 1956; Dehority 1966) and it was therefore surprising to find that the results obtained from characterization of the proteolytic activity of *B. fibrisolvens* H17c were very similar to that obtained for *B. fibrisolvens* 49 (Cotta and Hespell 1986). Mannarelli (1988) examined the DNA relatedness of 39 strains classified as *B. fibrisolvens* and found that *B. fibrisolvens* H17c and *B. fibrisolvens* 49 were approximately 95% related. This therefore explains the high level of similarity found between the proteolytic activity of the two strains.

Cotta and Hespell (1986) attempted to purify the extracellular proteolytic activity of *B. fibrisolvens* 49 for biochemical characterization. These attempts were generally unsatisfactory, as low yields were obtained and little purification was achieved. Preliminary SDS-PAGE revealed a major protein at approximately 30 000 Da and several faint ones around 75 000 to 120 000 Da. Their purification attempts showed that the proteolytic activity was in association with a carbohydrate material. This material was found to be a high M_r polysaccharide but the nature of the association between the proteolytic enzyme(s) and the polysaccharide was unknown This polysaccharide association prevented them from purifying the protease(s) using conventional methods. Further characterization of the proteolytic activity of *B. fibrisolvens* H17c using recombinant DNA technology is therefore an attractive alternative.

Chapter 7

General Conclusions

This study was aimed at obtaining an understanding of the genes and enzymes involved in cellulose and protein degradation by the important rumen bacterium *B. fibrisolvens* H17c.

Chromosomal DNA was isolated from *B. fibrisolvens* H17c and used for the construction of a *B. fibrisolvens* gene library in the plasmid pEcoR251, an *E. coli* positive selection vector. The library was screened for genes expressing cellulase, xylanase, and protease activity. Two clones expressing cellulase activity were isolated. The methods employed in this study required *B. fibrisolvens* genes to be expressed and the products to be active in *E. coli*. The inability to isolate the genetic determinants responsible for *B. fibrisolvens* protease activity may be due to a deficiency in either of these two requirements. Alternatively, *B. fibrisolvens* proteases may be lethal when expressed from high copy number plasmids in *E. coli*.

DNA probes or antibody screening methods could provide a route for the identification and cloning of the protease encoding genes. Antibodies raised against alkaline serine protease enzymes from *B. fibrisolvens* H17c or other organisms could be used to detect *B. fibrisolvens* proteases synthesized in an enzymatically inactive but antigenically active form in *E. coli*. The nucleotide sequences of five *B. fibrisolvens* H17c genes have been determined in this department to date. These data could be used to compile a *B. fibrisolvens* H17c codon usage table for use in the synthesis of oligonucleotide probes. Alkaline serine proteases have conserved regions which could be of great help in the design of oligonucleotide probes for the isolation of *B. fibrisolvens* H17c genes encoding alkaline serine proteases which are enzymatically inactive or not synthesized in *E. coli*. Screening of a phage λ-based (using λ-ZAP or λgt11 vector systems, Short et al. 1988; Young and Davis 1983) *B. fibrisolvens* genebank transduced into *E. coli* either (1) for zones of hydrolysis on

skim milk plates, or (2) by plaque hybridization (with antibody or oligonucleotide probes) will be investigated.

The two cellulase genes isolated from the *B. fibrisolvens* H17c library, *end1* and *ced1*, encoded enzymatically active proteins in *E. coli*. The *end1* gene codes for an endo–β–1,4–glucanase, and the *ced1* gene encodes a cellodextrinase enzyme. Both enzymes were found predominantly in the periplasm of the *E. coli* host indicating that protein export occurred from the cytoplasm. End1 was shown to contain a signal peptide typical for procaryotes. However, Ced1 does not contain a typical signal sequence and the mechanism of export is unknown. It is important to investigate mechanisms for possible transfer of genetic material into *B. fibrisolvens*, as well as methods to stabilize subsequent transformants. Homologous recombination of mutant genes may enable one to determine how significant these two enzymes are in the degradation of cellulose by *B. fibrisolvens* H17c.

Both genes were expressed from their own promoters in *E. coli*. Future work to identify the active promoters could be carried out by fusion of the putative promoter regions to a promoter-less *lacZ* gene, and assaying for β-galactosidase activity. mRNA transcript mapping would indicate whether the same promoters are recognized in *B. fibrisolvens* H17c as in *E. coli*.

A very interesting aspect that emerged from the sequencing of the DNA upstream of the *end1* gene was the presence of another ORF 135 bp upstream of the *end1* start codon (data not shown). The predicted translation of the 3'end of this ORF comprised 245 amino acids which showed high similarity with the C-terminal region of a number of chemotaxis proteins (*E. coli* Tsr, Trg, Tar, and Tap) (Bollinger et al. 1984; Boyd et al. 1983; Krikos et al. 1983). As *B. fibrisolvens* is one of the rapid colonizers of fresh cellulosic material arriving in the rumen it is very tempting to speculate that this protein may be involved in the colonization of cellulosic material. Further studies will involve characterization of the putative chemotaxis protein, and investigation of its role in colonization and cellulose and protein degradation.

This study describes the cloning and sequencing of the first genes from *B. fibrisolvens* H17c, and therefore lays the foundation for genetic manipulation of this rumen microorganism. It is particularly significant that two different *B. fibrisolvens* cellulase genes were expressed in *E. coli*, the workhorse of recombinant DNA studies. The development of a method for the introduction of DNA into *B. fibrisolvens* cells, and the construction of plasmid vectors which replicate in *B. fibrisolvens* are essential pre-requisites for such applications. However, either *end1* or *ced1* represent convenient homologous genetic markers which could be carried on plasmid vectors in a non-cellulolytic *B. fibrisolvens* strain.

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). Polyacrylamide gel solutions and buffers are given in Appendix B.

A.1 Media

A.1.1 *B. fibrisolvens* minimal medium[a] (MM) (Cotta and Hespell 1986)

Ingredient	Amount/l
Glucose solution (20% w/v)[b]	20.0 ml
Mineral solution 1[b]	40.0 ml
Mineral solution 3[b]	40.0 ml
Trace mineral solution[d]	10.0 ml
Resazurin solution (0.1% w/v)	1.0 ml
Volatile fatty acid solution[b]	3.1 ml
Vitamin solution[e]	10.0 ml
Reducing agent solution (5.0% w/v)[f]	10.0 ml
Sodium carbonate solution (8.0% w/v)[b]	50.0 ml
Nitrogen source[g]	

[a]Prepared under CO_2; pH adjusted to 6.8 prior to autoclaving.

[b]See M10 medium for details

[c]Mineral solution 3 was prepared as mineral solution 2 (see M10 medium for details) except 5.58 g of Na_2SO_4 added in place of $(NH_4)_2SO_4$.

[d]Trace mineral solution contained/l:

Na_4EDTA	500 mg
$FeSO_4 \cdot 7H_2O$	200 mg
$MnCl_2 \cdot 4H_2O$	200 mg
$ZnSO_4 \cdot 7H_2O$	10 mg
H_3BO_3	30 mg
$CoCl_3 \cdot 6H_2O$	12 mg
$CuCl_2 \cdot 2H_2O$	1 mg
$NiCl_2 \cdot 6H_2O$	2 mg
$NaMoO_4 \cdot 2H_2O$	30 mg
Na_2SeO_3	50 mg

Keep mineral solutions over chloroform at 4°C.

[e]Vitamin solution contained/100 ml of 5 mM Hepes buffer (pH 7.5):

biotin	2.5 mg
folic acid	2.5 mg
calcium D-pantothenate	20 mg
nicotinamide	20 mg
riboflavin	20 mg
thiamine	20 mg
pyridoxine HCl	20 mg
p-aminobenzoic acid	2.5 mg
cyanocobalamin	2.5 mg

Solution was filter sterilized and stored at -20°C.

[f]Reducing agent solution contained cysteine HCl (5% w/v). Prepared under N_2 as a seperate sterile solution and added to cooled medium.

[g]Nitrogen sources were prepared under N_2 and added as individual sterile solutions to yield media with final concentrations of:

NH_3	45 mM
trypticase	0.1% or 2.0% w/v
casamino acids	0.1% or 2.0% w/v
urea	45 mM
$(NH_4)_2SO_4$	45 mM

A.1.2 *B. fibrisolvens* non-rumen fluid medium (M10)

Ingredient	Amount/l
Glucose	10.0 g
Cellobiose	0.5 g
Mineral solution 1[a]	38.0 ml
Mineral solution 2[b]	38.0 ml
Cysteine HCl (5% w/v)	10.0 ml
Na_2CO_3 (5% w/v)	50.0 ml
Resazurin (0.1% w/v)	1.0 ml
Trypticase	2.0 g
Yeast extract	0.5 g
Volatile fatty acid[c]	3.1 ml
Hemin[d]	10.0 ml
Distilled water	850.0 ml

[a]Mineral solution 1 contained/l:

K_2HPO_4	6.0 g

[b]Mineral solution 2 contained/l:

KH_2PO_4	6.0 g
NaCl	12.0 g
$(NH_4)_2SO_4$	6.0 g
$CaCl_2 \cdot 2H_2O$	1.6 g
$MgSO_4 \cdot 7H_2O$	2.5 g

[c]Volatile fatty acid solution contained:

acetic acid	17 ml
propionic acid	6 ml
butyric acid	4 ml
isobutyric acid	1 ml
n-valeric acid	1 ml
iso-valeric acid	1 ml
2-methylbutyric acid	1 ml

Solution is stable at 4°C.

[d]Hemin solution:

KOH	0.28 g
ethanol (95%)	25.0 ml
hemin	0.10 g
distilled water	75 ml

Solution is stable at 4°C

All the ingredients were added and the pH adjusted to 6.8 prior to autoclaving.

A.1.3 Luria-Bertani medium (LB)

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

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

A.1.4 YT medium (x2)

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

For pUC recombinant selection, IPTG (0.1 ml) and XGal (0.8 ml) were added to 250 ml agar before pouring the plates.

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
Chloramphenicol	20 mg/ml ethanol (96%)
Kanamycin	62.5 mg/ml water
Tetracycline	12.5 mg/ml ethanol (50%)

All antibiotics were filter sterilized and stored at -20°C, except for Tc which was always made fresh

A.2.2 IPTG (isopropyl-β-D-thio-galactopyranoside)

IPTG (100mM)	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-β-galactoside)

XGal (2% w/v)	0.2 g
Dimethylformamide	10 ml

The solution was stored at -70°C.

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

Adjust pH to 7.0 with 0.1 N NaOH before making up to 5 ml. Store in 100 µl aliquots at -70°C. Discard remainder once defrosted.

A.3.2 Azocasein

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

Stir overnight to dissolve azocasein fibres. Autoclave.

A.3.3 Bradford solution (Bradford 1976)

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

Dissolve, then add 100 ml phosphoric acid (85%). Dilute to final volume of 1 l. Filter through Whatman GF/C filter paper. Store 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

Filter sterilize and store 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 the 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 five 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 DTT (1M)

DTT	0.618 g
Sodium acetate (0.01 M, pH 5.2)	4 ml

Filter sterilize

A.3.8 EDTA (0.5 M, pH 8.0) (Maniatis et al. 1982)

EDTA·$2H_2O$	168.1 g
Distilled water	to 1000 ml

EDTA will only dissolve when pH has been adjusted to 8.0. (Use approximately 20 g NaOH pellets for this purpose).

A.3.9 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.3.10 Exo-nuclease III shortening solutions (Henikoff 1987)

A.3.10.1 Exo buffer.

Tris/HCl (1 M, pH 8.0)	660 µl
$MgCl_2$ (0.1 M)	66.4 µl
Distilled water	9.27 ml

A.3.10.2 Klenow mixture.

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

A.3.10.3 Ligase mixture.

Ligase buffer	144 µl
Distilled water	1440 ml

A.3.10.4 S_1 buffer (10x).

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

A.3.10.5 S_1 mixture.

S_1 buffer (10x)	41 µl
Distilled water	259 µl
S_1 nuclease (60 U)	1.5 µl

A.3.10.6 S_1 stop.

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

A.3.11 Isopropanol (salt saturated)

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

A.3.12 Klenow (DNA polymerase I) buffer

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

Stock solution	Final conc.	/10 ml
Tris-Cl (1 M, pH 7.6)	0.1 M	1 ml
$MgCl_2$ (1 M)	0.1 M	1 ml
NaCl (5 M)	0.5 M	1 ml
2-mercaptoethanol	0.7 M	50 µl
Distilled water		6.95 ml

A.3.13 Ligase dilution buffer

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

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

A.3.14 Ligation buffer (10x)

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

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

A.3.15 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.3.16 Prehybridization solution

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

A.3.17 Restriction enzyme buffers (10x)

Stock solution	Final conc.
Tris-Cl (1 M, pH 7.9)	0.1 M
$MgCl_2$ (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.

	Salt concentration (mM)			
Stock solution	0	50	100	150
Tris-Cl	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_2O	2.4 ml	1.4 ml	0.4 ml	2.4 ml

A.3.18 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	10 mM	7 µl
Gelatin (10 mg/ml)	100 µg/ml	0.1 ml
Glycerol	44% (v/v)	4.4 ml

A.3.19 *Sma*I restriction endonuclease buffer (10x)

Stock solution	Final conc.	/10 ml
Tris-HCl (1 M, pH 8.0)	0.1 M	1 ml
KCl (1 M)	0.2 M	2 ml
$MgCl_2$ (1 M)	0.1 M	1.ml
DTT (0.5 M)	10 mM	0.2 ml
Glycerol	44% (v/v)	4.4 ml
Distilled water		1.4 ml

A.3.20 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. Immediately before use the DNA was denatured by boiling for 10 min followed by cooling on ice.

A.3.21 Sodium acetate (3 M, pH 5.2)

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

Adjust pH with glacial acetic acid. Autoclave.

A.3.22 SSC (20x)

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

Adjust pH to 7.0 with NaOH (10 N). Autoclave.

A.3.23 Tris acetate buffer (50x)

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

A.3.24 TE (Tris-EDTA) buffer (100x)

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

Autoclave and dilute with sterile water before use.

A.3.25 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

Dispense in 20 ml aliquots and autoclave. Add DMSO (1 ml) and glucose (0.5 M, 400 µl when necessary) immediately before use.

A.3.26 Z-buffer (pH 7.0)

Na_2HPO_4 (60 mM)	16.1 g
$NaH_2PO_4 \cdot 2H_2O$ (40 mM)	5.5 g
KCl (10 mM)	0.75 g
$MgSO_4 \cdot 7H_2O$ (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 preparations.

B.1.1 Small scale (miniprep).

Plasmid was isolated from a 5 ml overnight culture (LB + Ap, 100 μg/ml) as described by Ish-Horowicz and Burke (1981). Cells from a 1.5 ml sample of the culture were harvested by centrifugation in an Eppendorf microfuge tube for 1 min. The pellet was resuspended in 200 μl Solution I (50 mM glucose; 25 mM Tris-HCl, pH 8.0), incubated for 5 min at room temperature, and then 400 μl of Solution II (0.2 M NaOH, 1% (w/v) SDS) was added. The sample was vortexed briefly and placed on ice for 5 min, before the addition of 300 μl ice-cold Solution III (5 M KOAc, pH 4.8). The sample was vortexed briefly, and, after 5 min on ice, cellular debris and denatured chromosomal DNA were pelleted by centrifugation for 5 min. The supernatant (750 μl) was removed to a fresh tube and sedimented by centrifugation with an equal volume of isopropanol for 5 min. The pellet was resuspended in TE (600 μl), $NaClO_4$ was added (60 μl, 5 M), and the DNA was sedimented with an equal volume of isopropanol, washed with 70% ethanol, air dried and resuspended in 50 μl TE buffer.

B.1.2 Large scale (maxiprep).

A 200 ml culture was grown overnight at 37°C in the presence of the appropriate antibiotic. The cells were harvested by centrifugation at 6 000 *g* for 5 min and then resuspended in 4 ml Solution I. After 5 min at room temperature 8 ml Solution II was added, and the mixture was kept on ice for 5 min, before the addition of 6 ml ice cold Solution III. After a further 5 min on ice the cellular debris was removed by centrifugation at 12 000 *g* for 10 min. An equal volume of isopropanol was added to the supernatant and the DNA was precipitated by centrifugation at 27 000 *g* for 15 min. The pellet was washed with 70% ethanol and resuspended in 4.2 ml TE buffer, and purified by isopicnic CsCl-EtBr ultracentrifugation (Maniatis et al. 1982).

The plasmid preparation was prepared for ultracentrifugation by the addition of CsCl (1 mg/ml) and EtBr (0.5 ml of a 10 mg/ml stock). The solution was centrifuged at 27 000 *g* for 15 min to precipitate any remaining protein debris. The refractive index of the supernatant was adjusted to 1.396, the sample sealed in Beckman Quickseal ultracentrifuge tubes and centrifuged for 12 h at 55 000 rpm at 15°C in a Beckman Vti 65.2 rotor. The plasmid DNA band was visualized by long wave UV light (350 nm), and removed in the smallest volume possible. The EtBr was removed by extraction (3 times) with equal volumes of NaCl-saturated isopropanol. The DNA was precipitated from the CsCl solution by the addition of two volumes of water followed by an equal volume of isopropanol, and centrifugation in an Eppendorf microfuge for 15 min. The pellet was resuspended in 200 µl TE buffer and the concentration was determined spectrophotometrically by measuring the absorbance of 10 µl (diluted in TE) between 220 and 310 nm. The concentration was determined by using the relationship A_{260} = 1 for 50 µg/ml double-stranded DNA.

B.2 Restriction endonuclease digestion.

Restriction digests were carried out using one of the four restriction buffers (Appendix A) according to the salt requirements of the particular enzyme. The enzyme *Sma*I required a unique buffer (Appendix A). Digestion volumes were routinely 20 µl containing 300-500 ng DNA and one unit of restriction enzyme. Digestions were done 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 digestions. If the sample was to be used for ligation the digestion was terminated by a phenol-chloroform extraction. The DNA solution was extracted with the addition of phenol (1/10 volume, TE-saturated) and an equal volume of chloroform:isoamyl alchohol (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 5M $NaClO_4$ (1/10 volume), an equal volume of isopropanol, and 15 min centrifugation. If the DNA

concentration was less than 2 µg/100µl *E. coli* tRNA was added (2 µg) before precipitation. After centrifugation the pellet was washed with 70% ethanol and resuspended in TE buffer.

B.3 Agarose gel electrophoresis.

Agarose gel electrophoresis was carried out using a horizontal submerged gel system. Tris-acetate buffer (Appendix A) was used routinely. Sigma type II agarose was used at varying concentrations. The amount of DNA loaded/lane also varied with the sizes and number of fragments but under normal circumstances about 300 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 µg/ml) for 15-30 min. DNA bands were visualised 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 f 4.7 was used.

B.4 DNA ligation reactions.

DNA ligation reactions were of two basic types: recircularization of plasmids for the isolation of deletion clones (use low DNA concentrations, 1 pmole DNA/ml) and recombination reactions, for example in subcloning (use 5 pmole DNA/ml). DNA concentration was calculated using the formula 1 pmole = (0.662 x kb)µg.

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 3-20 h using 20-100 x 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%) (Seaplaque^R) in Tris-acetate buffer (50 mM, pH 8.2, no EDTA, no EtBr). The gel was stained with EtBr after electrophoresis and the DNA bands were viewed under UV light (310 nm), as briefly as possible. The desired bands 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 (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 gelled ligation reactions were melted at 70°C for 5 min, and then diluted with 4 volumes of 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) (2-4 h). The cell culture was poured into a pre-cooled sterile SS34 tube and the cells were harvested at 1000 *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) (Appendix A) and held on ice for 10 minutes. The *E. coli* cells (100 µl) were then mixed with DNA (routinely 50 ng) 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.

Unused cells could be stored at -70°C after rapid freezing in a dry ice/ethanol bath or liquid nitrogen and retained viability provided that the cells were thawed slowly on ice when needed.

B.7 Nucleotide sequencing

B.7.1 Primer annealing reaction.

The 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 μl of 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 15 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.7.2 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 "Sequanase" 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.7.3 Gel electrophoresis and autoradiography.

The sequencing reactions were analyzed 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 (0.2mm 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.8 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 suppliers specifications. Contaminating nucleotides were removed from the radioactively labelled probe preparation using a Sephadex G50 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 (5 min) in a fume hood just before use.

B.9 DNA hybridization.

DNA fragments resolved by agarose gel electophoresis 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. A sheet of Hybond-N, wetted by floating onto, and then immershed in, distilled water 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 4 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 4 h at 65°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 at 65°C overnight. The membrane was washed in 1 x SSC at 65°C, for 30 min, and after checking the radioactivity by means of a Geiger-counter, the washing was terminated and the membrane sealed in a plastic bag. The membrane was exposed to autoradiographic film (XAR-5) overnight at -70°C.

B.10 SDS Polyacrylamide gel electrophoresis (SDS-PAGE)

B.10.1 SDS-PAGE for the separation of proteins.

Discontinuous SDS-PAGE was done according to the method of Laemmli (1970), using a Hoefer SE600 vertical slab electrophoresis unit (Hoefer Scientific Instruments, San Fransisco, CA, USA). The 1.5 mm thick gel spacers were used. The resolving gel was prepared and degassed before pouring. Propan-2-ol was layered on the gel to promote a sharp interface. After the gel had polymerized (30 min), the propan-2-ol was removed by rinsing with the stacking gel buffer, and the stacking gel was cast.

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 35 mA (constant current)/gel until the dye front migrated to the end of the gel (four to five hours).

After electrophoresis the gels were stained for 3 h in coomassie blue staining solution with gentle agitation, destained and dried.

The acrylamide gels (10%) were prepared as follows:

Solution	Resolving gel	Stacking gel
Acrylamide solution	13.3 ml	2.66 ml
Resolving gel buffer	10.0 ml	-
Stacking gel buffer	-	5 ml
SDS	0.4 ml	0.2 ml
Distilled water	15.0 ml	12.2 ml
Ammonium persulphate	300 µl	100 µl
TEMED	100 µl	40 µl

B.10.2 SDS-PAGE with incorporated CMC.

A modified method of Sharma and Sandhu (1986) was used for the detection of CMCase activity. The 0.75 mm thick gel spacers were used and the stacking gels were prepared as described in Appendix B.10.1, but part of the water in the resolving gel was replaced with CMC solution at a final concentration of 0.1% w/v. The samples were prepared as described in section 5.2.4 and the proteins were electrophoresed at 100 V in the stacking gel and 70 V in the resolving gel. After electrophoresis the gels were washed in Triton X-100 (15 min; 2.5% v/v), incubated for 6 - 24 h in phosphate buffer (0.2 M; pH 6.6), stained for 1 h in Congo Red (0.1% w/v), destained in alkaline NaCl (1 M) until bands appeared, and transferred into 5% acetic acid to fix the bands and facilitate photography using a red filter.

The resolving CMC/acrylamide gels (7.5, 10, and 12%) were prepared as follows:

Solution	Resolving gel		
	7.5%	10%	12.5%
Acrylamide solution	10.0 ml	13.3 ml	16.7 ml
Resolving gel buffer	10.0 ml	10.0 ml	10.0 ml
SDS	0.4 ml	0.4 ml	0.4 ml
Distilled water	12.5 ml	9.3 ml	6.0 ml
CMC	6.7 ml	6.7 ml	6.7 ml
Ammonium persulphate	300 µl	300 µl	300 µl
TEMED	100 µl	100 µl	100 µl

B.10.3 SDS gelatin-PAGE.

Gelatin-PAGE was done essentially according to the method of Heussen and Dowdle (1980). The running conditions and treatment of the gels after electrophoresis are described in Chapter 6.2.2.2. The stacking gels were prepared as described in Appendix B.10.1. The resolving gelatin/acrylamide gels (8, 9, and 10%) were prepared as follows:

Solution	Resolving gel		
	8%	9%	10%
Acrylamide solution	12.0 ml	13.3 ml	15.0 ml
Resolving gel buffer	11.25 ml	11.25 ml	11.25 ml
Distilled water	16.8 ml	15.3 ml	13.8 ml
SDS	0.45 ml	0.45 ml	0.45 ml
Gelatin	4.5 ml	4.5 ml	4.5 ml
Ammonium persulphate	100 µl	100 µl	100 µl
TEMED	50 µl	50 µl	50 µl

The following buffers and solutions were used for SDS-PAGE:

B.10.4 SDS-PAGE solutions

B.10.4.1 General SDS-PAGE solutions

Acrylamide solution

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

Filter through Whatman's paper (No. 1) and store in dark bottle at 4°C.

Resolving gel buffer

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

Adjust pH to 8.8 (approximately 3.3 ml conc. HCl).

Stacking gel

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

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

Ammonium persulphate (10% w/v)
A fresh solution was made immediately before use.

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.

SDS (10%)

SDS	50 g
Distilled water	to 500 ml

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).

Destain solution

Acetic acid	250 ml
Methanol	750 ml
Distilled water	1500 ml

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

The solution was stored in aliquots at -20°C.

B.10.4.2 Solutions for CMC/PAGE

Phosphate buffer (0.2 M)

$NaH_2PO_4 \cdot H_2O$	13.8g
Distilled water	to 500 ml

Adjust pH with NaOH to pH 6.8.

CMC solution

CMC	0.6 g
Phosphate buffer (0.2 M, pH 6.8)	to 100 ml

Grind CMC with a small amount of buffer in a pestle and mortar to disperse evenly. Make up to volume, dispense into McCartney bottles in 10 ml aliquots and autoclave.

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

Adjust pH to approximately 12.0 with NaOH.

NaCl (1 M)

NaCl	116.88 g
Distilled water	2000 ml

Adjust pH to approximately 12.0 with NaOH.

Acetic acid (5%)

Acetic acid	100 ml
Distilled water	to 2000 ml

B.10.4.3 Solutions for Gelatin/PAGE

Gelatin solution (1%)

Gelatin	1 g
Distilled water	80 ml

Boil solution until the gelatin has dissolved, then make volume up to 100 ml with water and autoclave.

Amido Black stain

"Amidoswartz" (for electrophoresis)	1 g
Distilled water	to 50 ml

Mix for 10 min. Made up to 500 ml with destain. Stir for 30 min. Filter through Whatman's paper (No. 3). Stain can be re-used several times.

B.11 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 V; 10-200 μg/ml). Protein samples were diluted such that OD_{595} did not exceed 0.8.

Appendix C

E. coli strains, genotypes, and references

E. coli strain	Genotype/description	Reference/origin
C600	*supE*44 *hsdR thi*-1 *thr*-1 *leuB*6 *lacY*1 *tonA*21	Appleyard (1954)
CC118	*araD*139 Δ(*ara, leu*)7697 Δ*lacX*74 *phoA*Δ20 *galE galK thi rpsE rpoB* $argE_{am}$ *recA*1	Manoil and Beckwith (1985)
K514	*thr*-1 *leuB*6 *thi*-1 *supE*44 *lacY*1 *ton*A21 r_k^-, m_k^+ (C600 derivative)	Wood (1966)
LK111	lacI^q lacZΔM15 lacY$^+$ (K514 derivative)	Zabeau and Stanley (1982)

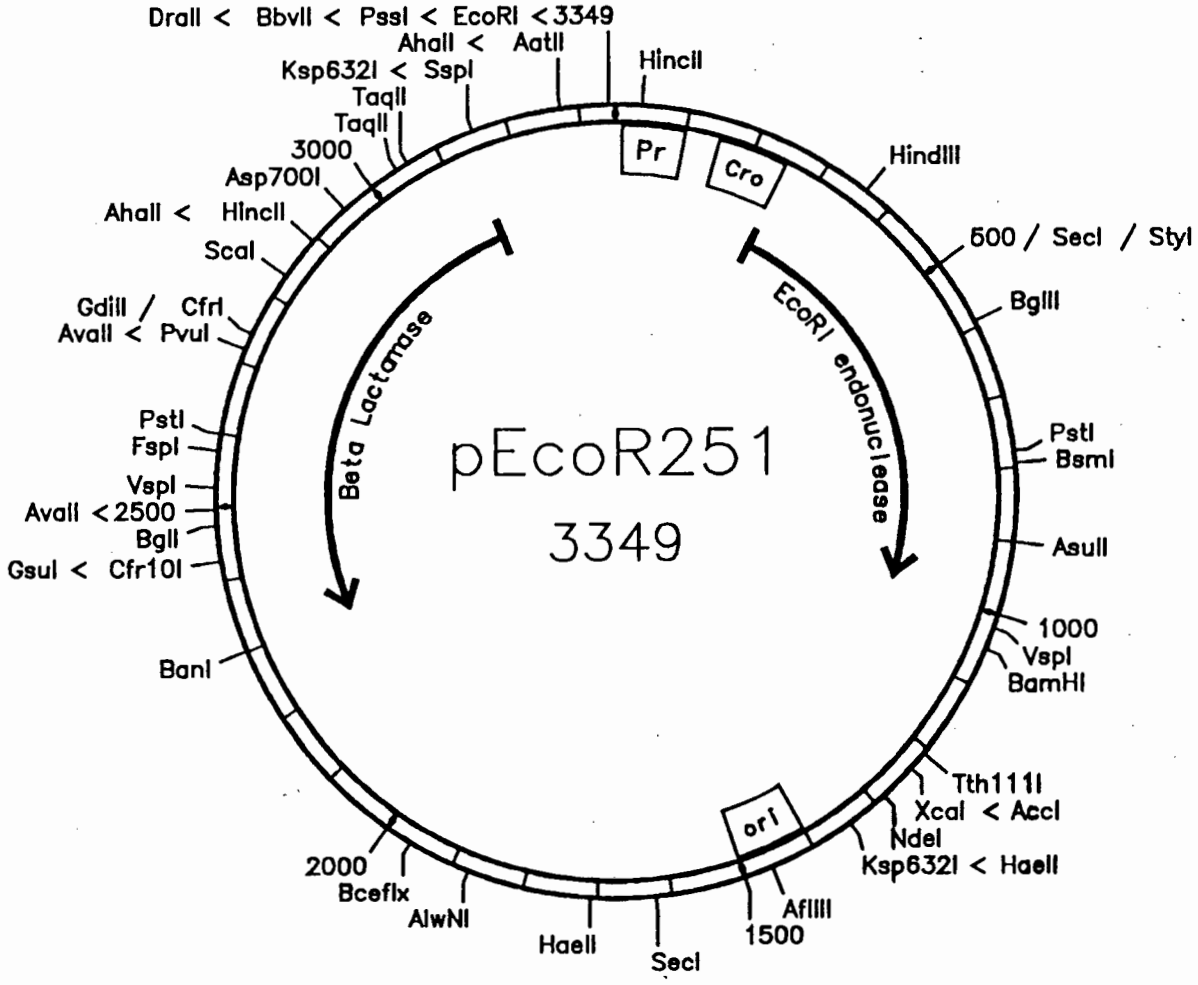
Appendix D

One- and three-letter codes used for amino acids

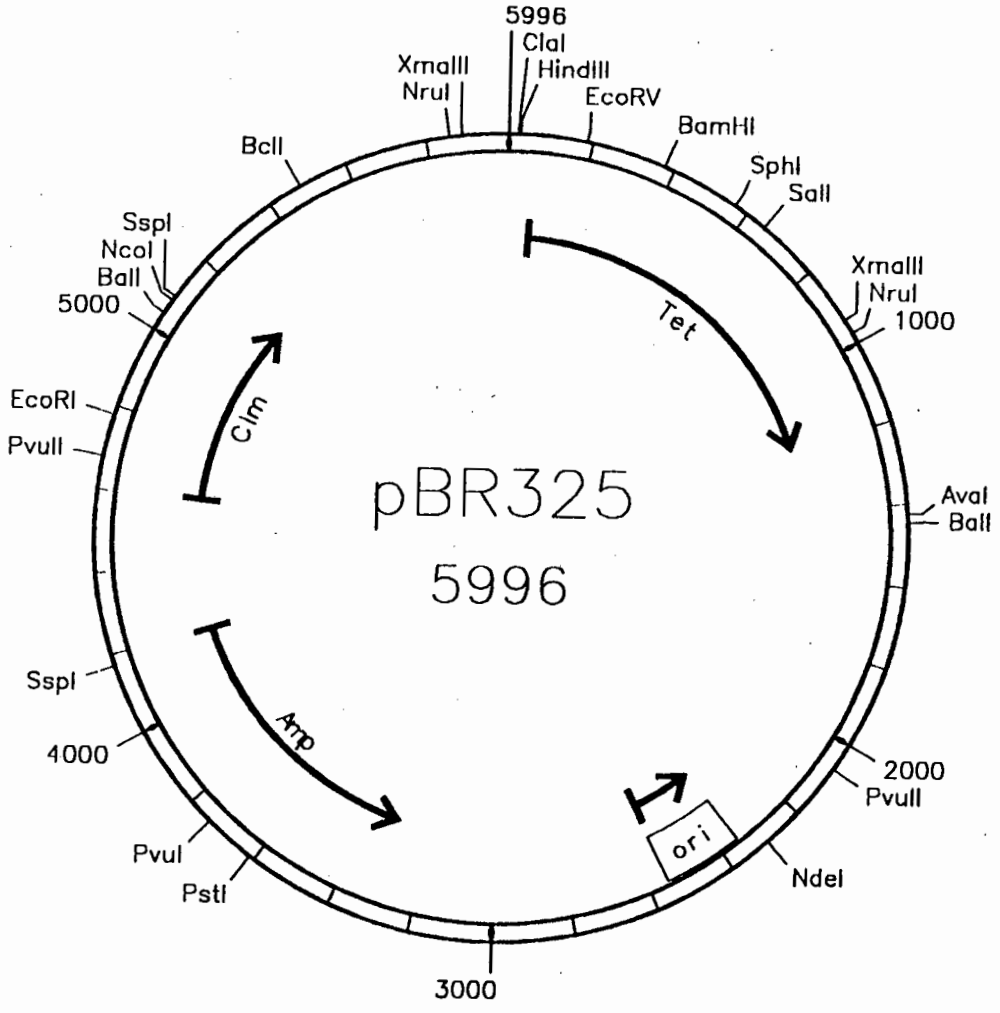
Amino acid	Codes		Code	Amino acid
Alanine	Ala	A	A	Alanine
Arginine	Arg	R	C	Cysteine
Asparagine	Asn	N	D	Aspartic acid
Aspartic acid	Asp	D	E	Glutamic acid
Cysteine	Cys	C	F	Phenylalanine
Glutamine	Gln	Q	G	Glycine
Glutamic acid	Glu	E	H	Histidine
Glycine	Gly	G	I	Isoleucine
Histidine	His	H	K	Lysine
Isoleucine	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

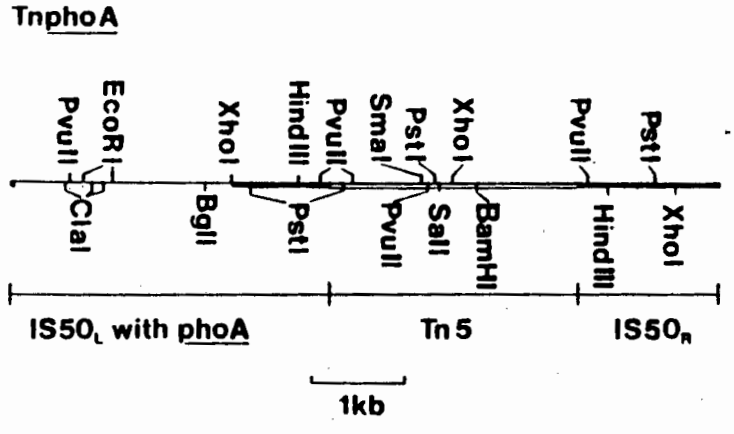
Plasmid vector and phage maps.



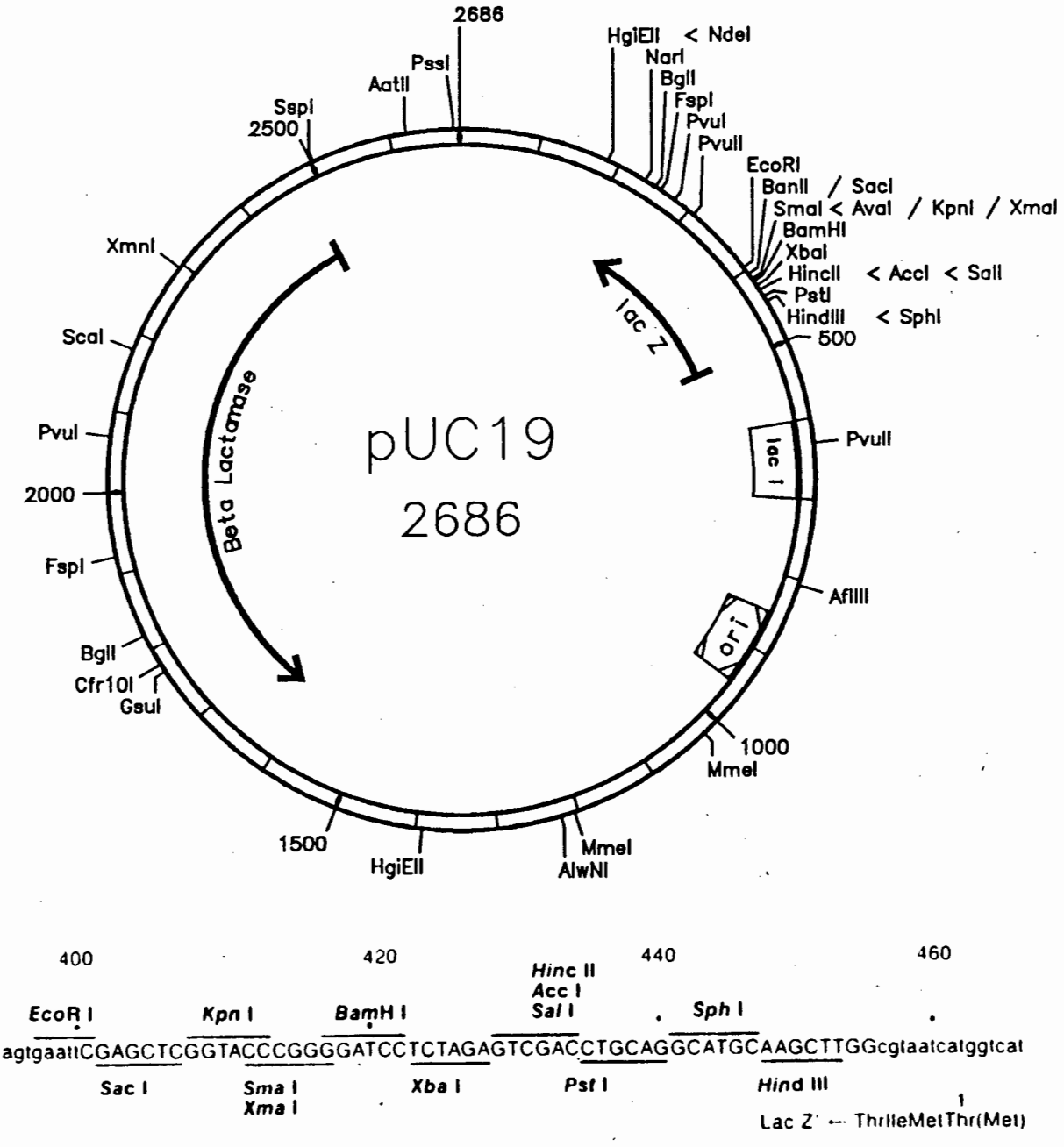
Reference: Zabeau and Stanley personal communication



Reference: Bolivar (1978)



Partial restriction map of Tn*phoA* showing position of Tn5 relative to the leftward and rightward insertion sequences (J Beckwith pers. comm.)



Restriction map of pUC18 and pUC19. The plasmids differ in the orientation of the multiple cloning site polylinker shown below the restriction map and the orientation shown corresponds to pUC19 Reference: Norrander et al. (1983)

References

Akin DE (1976) Ultrastructure of rumen bacterial attachment to forage cell walls. Appl Environ Microbiol **31**:562-568

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

Alberts B, Bray D, Lewis V, Raff M, Roberts K and Watson JD (1983) Molecular Biology of the Cell. Garland, New York

Appleyard RK (1954) Segregation of new lysogenic types during growth of a doubly lysogenic strain derived from *Escherichia coli* K12. Genetics **39**:440

Ayers WA (1958) Phosphorylation and synthesis of cellobiose by cell extracts from *Ruminococcus flavefaciens*. J Bacteriol **76**:515-517

Barros MEC and Thomson JA (1987) Cloning and expression in *Escherichia coli* of a cellulase gene from *Ruminococcus flavefaciens*. J Bacteriol **169**:1760-1762

Bauchop T and Clarke RTJ (1976) Attachment of the ciliate *Epidinium crawley* to plant fragments in the sheep rumen. Appl Environ Microbiol **32**:417-422

Béguin P (1990) Molecular biology of cellulose degradation. Annu Rev Microbiol (in press)

Béguin P, Cornet P and Aubert JP (1985) Sequence of a cellulase gene of the thermophilic bacterium *Clostridium thermocellum*. J Bacteriol **162**:102-105

Béguin P and Eisen H (1978) Purification and partial characterization of three extracellular cellulases from *Cellulomonas* sp. Eur J Biochem **87**:525-531

Béguin P, Gilkes NR, Kilburn DG, Miller Jr RC and O'Neill GP (1987) Cloning of cellulase genes. CRC Crit Rev Biotechnol **6**:129-162

Béguin P, Grépinet O, Millet J and Aubert J-P (1988a) Recent aspects in the biochemistry and genetics of cellulose degradation. In: G Durand et al. (eds) 8th Int Biotechnol Symp Proc **2**:1015-1019

Béguin P, Millet J, Grépinet O, Navarro A and Juy M (1988b) The *cel* (cellulose degrading) genes of *Clostridium thermocellum*. In: J-P Aubert et al. (eds) Biochemistry and genetics of cellulose degradation. Academic Press, London, FEMS Symp **43**:267-282

Berger E, Jones WA, Jones DT and Woods DR (1989) Cloning and sequencing of an endoglucanase (*end1*) gene from *Butyrivibrio fibrisolvens* H17c. Mol Gen Genet **219**:193-198

Berger E, Jones WA, Jones DT and Woods DR (1990) Sequencing and expression of a cellodextrinase (*ced1*) gene from *Butyrivibrio fibrisolvens* H17c in *Escherichia coli*. Mol Gen Genet (in press)

Bhat KM, McCrae SI and Wood TM (1989) The endo-(1,4)-β-D-glucanase system of *Penicillium pinophilum* cellulase: isolation, purification, and characterization of five major endoglucanase components. Carbohydr Res **190**:279-297

Black JL and Tribe DE (1973) Comparison of ruminal and abomasal administration of feed on the growth and composition of lambs. Aust J Agr Res **24**:763-773

Blackburn TH (1968a) Protease production by *Bacteroides amylophilus* H18. J Gen Microbiol **53**:27-36

Blackburn TH (1968b) The protease liberated from *Bacteroides amylophilus* strain H18 by mechanical disintegration. J Gen Microbiol **53**:37-51

Blackburn TH and Hobson PN (1962) Further studies on the isolation of proteolytic bacteria from the sheep rumen. J Gen Microbiol **29**:69-81

Bolivar F (1978) Construction and characterization of new cloning vehicles. Derivatives of plasmid pBR322 carrying unique *Eco*R1 sites for selection of *Eco*R1 generated recombinant DNA molecules. Gene **4**:121-136

Bollinger J, Park C, Harayama S and Hazelbauer GL (1984) Structure of the Trg protein: Homologies with and differences from other sensory transducers of *Escherichia coli*. Proc Natl Acad Sci USA **81**:3287-3291

Boyd A, Kendall K and Simon MI (1983) Structure of the serine chemoreceptor in *Escherichia coli*. Nature **301**:623-626

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem **72**:248-259

Brickman E and Beckwith J (1975) Analysis of the regulation of *Escherichia coli* alkaline phosphatase synthesis using deletions and ϕ80 transducing phages. J Mol Biol **96**:307-316

Brock FM, Forsberg CW and Buchanan-Smith JG (1982) Proteolytic activity of rumen microorganisms and effects of proteinase inhibitors. Appl Environ Microbiol **44**:561-569

Broderick GA, Wallace RJ and McKain N (1988) Uptake of small neutral peptides by mixed rumen microorganisms *in vitro*. J Sci Food Agric **42**:109-118

Brown DW and Moore WEC (1960) Distribution of *Butyrivibrio fibrisolvens* in nature. J Dairy Sci **43**:1570-1574

Bryant MP (1965) Rumen methanogenic bacteria. In: RW Dougherty et al. (eds) Physiology of digestion in the ruminant. Butterworths, Washington, DC, pp 411-422

Bryant MP and Robinson IM (1962) Some nutritional characteristics of predominant culturable ruminal bacteria. J Bacteriol **84**:605-614

Bryant MP and Small N (1956) The anaerobic monotrichous butyric acid producing curved rod-shaped bacteria of the rumen. J Bacteriol **72**:16-21

Buchanan RE and Gibbons N (eds) (1974). Bergey's manual of determinative bacteriology, 8th edn, Williams and Wilkins, Baltimore

Caldwell DR and Bryant MP (1966) Medium without rumen fluid for nonselective enumeration and isolation of rumen bacteria. Appl Microbiol **14**:794-801

Chalupa W (1977) Manipulating rumen fermentation. J Anim Sci **46**:585-599

Chauvaux S, Béguin P, Aubert JP, Bhat KM, Gow LA, Wood TM and Bairoch A (1990) Calcium-binding affinity and calcium-enhanced activity of *Clostridium thermocellum* endoglucanase D. Biochem J **265**:261-265

Chen CM, Gritzali M and Stafford DW (1987) Nucleotide sequence and deduced primary structure of cellobiohydrolase II from *Trichoderma reesei*. Biotechnol 5:274-278

Chen G, Russell JB and Sniffen CJ (1987) A procedure for measuring peptides in rumen fluid and evidence that peptide uptake can be a rate-limiting step in ruminal protein degradation. J Dairy Sci **70**:1211-1219

Cheng KJ and Costerton JW (1975) Ultrastructure of cell envelopes of bacteria of the bovine rumen. Appl Microbiol **29**:841-849

Cheng KJ and Costerton JW (1977) Ultrastructure of *Butyrivibrio fibrisolvens*: A Gram-positive bacterium? J Bacteriol **129**:1506-1512

Chung CT and Miller RH (1988) A rapid and convenient method for the preparation and storage of competent bacterial cells. Nucleic Acids Res **16**:3580

Clarke L and Carbon J (1976) A colony bank containing synthetic Col EI hybrid plasmids representative of the entire *Escherichia coli* genome. Cell **9**:91-99

Coen JA and Dehority BA (1970) Degradation and utilization of hemicellulose from intact forages by pure cultures of rumen bacteria. Appl Microbiol **20**:362-368

Coleman GS (1978) The metabolism of cellulose, glucose and starch by the rumen ciliate protozoan, *Eudiplodinium maggii*. J Gen Microbiol **107**:359-366

Coleman GS (1980) Rumen ciliate protozoa. Adv Parasitol **18**:121-173

Costerton JW (1970) The structure and function of the cell envelope of Gram-negative bacteria. Rev Can Biol **29**:299-316

Cotta MA (1988) Amylolytic activity of selected species of ruminal bacteria. Appl Environ Microbiol **54**:772-776

Cotta MA and Hespell RB (1986) Proteolytic activity of the ruminal bacterium *Butyrivibrio fibrisolvens*. Appl Environ Microbiol **52**:51-58

Coughlan MP (1985) The properties of fungal and bacterial cellulases with comment on their production and application. Biotechnol Genetic Engng Rev **3**:39-109

Coughlan MP and Ljungdahl LG (1988) Comparative biochemistry of fungal and bacterial cellulolytic enzyme systems In: J-P Aubert et al. (eds) Biochemistry and genetics of cellulose degradation. Academic Press, London, FEMS Symp **43**:11-30

Creuzet N, Berenger J-F and Frixon C (1983) Characterization of exoglucanase and synergistic hydrolysis of cellulose in *Clostridium stercorarium*. FEMS Microbiol Lett **20**:347 350

Dawson KA, Preziosi MC and Caldwell DR (1979) Some effects of uncouplers and inhibitors on growth and electron transport in rumen bacteria. J Bacteriol **139**:384-392

De Vries W, Van Wijck-kapteyn WMC and Oosterhuis SKH (1974) The presence and function of cytochromes in *Selenomonas ruminantium, Anaerovibrio lipolytica* and *Veillonella alcalescens*. J Gen Microbiol **81**:69-78

Dean RG, Martin SA and Carver C (1989) Isolation of plasmid DNA from the ruminal bacterium *Selenomonas ruminantium* HD4. Lett Appl Microbiol **6**:113-115

Dehority BA (1965) Degradation and utilization of isolated hemicellulose by pure cultures of cellulolytic bacteria. J Bacteriol **89**:1515-1520

Dehority BA (1966) Characterization of several bovine rumen bacteria isolated with a xylan medium. J Bacteriol **91**:1724-1729

Dehority BA (1969) Pectin-fermenting bacteria isolated from the bovine rumen. J Bacteriol **99**:189-196

Dehority BA (1973) Hemicellulose degradation by rumen bacteria. Fed Proc **32**:1819-1825

Dehority BA and Grubb JA (1976) Basal medium for the selective enumeration of rumen bacteria utilizing specific energy sources. Appl Environ Microbiol **32**:703-710

Dekker RFH and Lindner WA (1979) Bioutilization of lignocellulosic waste materials: A review. S Afr J Sci **75**:65-71

Deshpande MV, Erikson RE and Petterson LG (1984) An assay for selective determination of exo-1,4-B-glucanases in a mixture of cellulolytic enzymes. Anal Biochem **138**:481-487

Devereux J (1984) A comprehensive set of sequence analysis programs for the Vax. Nucleic Acids Res **12**:387-395

Dretzen G, Bellard M, Sassone-Corsi P and Chambon P (1981) A reliable method for the recovery of DNA fragments from agarose and acrylamide gels. Biochem J **112**:295-298

El-Gogary S, Leite A, Crivellaro O, Eveleigh DE and El-Dorry H (1989) Mechanism by which cellulose triggers cellobiohydrolase I gene expression in *Trichoderma reesei*. Proc Natl Acad Sci USA **86**:6138-6141

Evans JS, Levine BA, Trayer IP, Dorman CJ and Higgins CF (1988) Sequence imposed structural constraints in the TonB protein of *E. coli*. FEBS Lett **208**:211-216

Eveleigh DE (1987) Cellulase: a perspective. Phil Trans R Soc Lond **A321**:435-448

Faure E, Belaich A, Bagnara C, Gaudin C and Belaich J-P (1990) Sequence analysis of the *Clostridium cellulolyticum celCCA* endoglucanase gene. Gene (in press)

Flint HJ and Stewart CS (1987) Antibiotic resistance patterns and plasmids of ruminal strains of *Bacteroides ruminicola* and *Bacteroides multiacidus*. Appl Microbiol Biotechnol **26**:450-455

Flint HJ, Thomson AM and Bisset J (1988) Plasmid-associated transfer of tetracycline resistance in *Bacteroides ruminicola*. Appl Environ Microbiol **54**:855-860

Forsberg CW, Beveridge TJ and Hellstrom A (1981) Cellulase and xylanase release from *Bacteroides succinogenes* and its importance in the rumen invironment. Appl Environ Microbiol **42**:886-896

Fragione B and Wolfenstein-Todel C (1972) Partial duplication in the "hinge" region of IgA1 myeloma proteins. Proc Natl Acad Sci USA **69**:3672-367.

Fukomori F, Sashihara N, Kudo T and Horikoshi K (1986) Nucleotide sequences of two cellulase genes from alkalophilic *Bacillus* sp. Strain N-4 and their strong homology. J Bacteriol **168**:479-485

Fulghum RS and Moore WEC (1963) Isolation, enumeration, and characteristics of proteolytic ruminal bacteria. J Bacteriol **85**:808-815

Gaboriaud C, Bissery V, Benchetrit T and Mornon J-P (1987) Hydrophobic cluster analysis. An efficient new way to compare and analyze amino acid sequences. FEBS Lett **224**:149-155

Gardner RM, Doerner KC and White BA (1987) Purification and characterization of an exo-β-1,4-Glucanase from *Ruminococcus flavefaciens* FD-1. J Bacteriol **169**:4581-4588

Gerwig GJ, DeWaard P, Kammerling JP, Vliegenthart JFG, Morgenstern E, Lamed R and Bayer EA (1989) Novel O-linked carbohydrate chains in the cellulase complex (cellulosome) of *Clostridium thermocellum*. J Biol Chem **264**:1027-1035

Ghangas GS and Wilson DB (1987) Expression of a *Thermomonospora fusca* endoglucanase E2 gene in *Streptomyces lividans* and *Bacillus subtilis*. Appl Environ Microbiol **53**:1470-1475

Ghangas GS and Wilson DB (1988) Cloning of the *Thermomonospora fusca* endoglucanase E2 gene in *Streptomyces lividans*: affinity purification and functional domains of the cloned gene product. Appl Environ Microbiol **54**:2521-2526

Ghose TK (1987) Measurement of cellulase activities. Pure Appl Chem **59**:257-268

Gilkes NR, Kilburn DG, Langsford ML, Miller RC Jr, Wakarchuk WW, Warren RAJ, Whittle DJ and Wong WKR (1984a) Isolation and characterization of *Escherichia coli* clones expressing cellulase genes from *Cellulomonas fimi*. J Gen Microbiol **130**:1377-1384

Gilkes NR, Kilburn DG, Miller Jr RC and Warren RAJ (1989) Structural and functional analysis of a bacterial cellulase by proteolysis. J Biol Chem **264**:17802-17808

Gilkes NR, Langsford ML, Kilburn DG, Miller Jr RC and Warren RAJ (1984b) Mode of action and substrate specificities of cellulases from cloned bacterial genes. J Biol Chem **259**:10455-10459

Gilkes NR, Warren RAJ, Miller Jr RC and Kilburn DG (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

Gilligan W and Reese ET (1954) Evidence for multiple components in microbial cellulases. Can J Microbiol **1**:90-107

Gong J, Lo RYC and Forsberg CW (1989) Molecular cloning and expression in *Escherichia coli* of a cellodextrinase gene from *Bacteroides succinogenes* S85. Appl Environ Microbiol **55**:132-136

Gräbnitz F and Staudenbauer WL (1988) Characterization of two β-glucosidase genes from *Clostridium thermocellum*. Biotechnol Lett **10**:73-78

Gradel CM and Dehority BA (1972) Fermentation of isolated pectin and pectin from intact forages by pure cultures of rumen bactera. Appl Microbiol **23**:332-340

Greenberg NM, Warren RAJ, Kilburn DG and Miller Jr RC (1987a) Regulation, initiation, and termination of the *cenA* and *cex* transcripts of *Cellulomonas fimi*. J Bacteriol **169**:646-653

Greenberg NM, Warren RAJ, Kilburn DG and Miller Jr RC (1987b) Regulation and initiation of *cenB* transcripts of *Cellulomonas fimi*. J Bacteriol **169**:4674-4677

Greenwood JM, Gilkes NR, Kilburn DG, Miller Jr RC and Warren RAJ (1989) Fusion to an endoglucanase allows alkaline phosphatase to bind to cellulose. FEBS Lett **244**:127-131

Grépinet O and Béguin P (1986) Sequence of the cellulase gene of *Clostridium thermocellum* coding for endoglucanase B. Nucleic Acids Res **14**:1791-1799

Greve LC, Labavitch JM and Hungate RE (1984) Muralytic activities of *Ruminococcus albus* 8. Appl Environ Microbiol **47**:1141-1145

Groleau D and Forsberg CW (1981) Cellulolytic activity of the rumen bacterium *Bacteroides succinogenes*. Can J Microbiol **27**:517-530

Groleau D and Forsberg CW (1983) Partial characterization of the extracellular endoglucanase produced by *Bacteroides succinogenes*. Can J Microbiol **29**:504-517

Gutierrez C, Barondess J, Manoil C and Beckwith J (1987) The use of transposon Tn*phoA* to detect genes for cell envelope proteins subject to a common regulatory stimulus: analysis of osmotically regulated genes in *Escherichia coli*. J Mol Biol **195**:289-297

Hall J and Gilbert HJ (1988) The nucleotide sequence of a carboxymethylcellulase from *Pseudomonas fluorescens* subsp. *cellulosa*. Mol Gen Genet **213**:112-117

Hall J, Hazlewood GP, Barker PJ and Gilbert HJ (1988) Conserved reiterated domains in *Clostridium thermocellum* endoglucanases are not essential for catalytic activity. Gene **69**:29-38

Hall J, Hazlewood GP, Huskisson NS, Durrant AJ and Gilbert HJ (1989) Homology of a xylanase and cellulase from *Pseudomonas fluorescens* subsp. *cellulosa*: internal signal sequence and unusual protein processing. Mol Microbiol **3**:1211-1217

Hawley DK and McClure WR (1983) Compilation and analysis of *Escherichia coli* promoter DNA sequences. Nucleic Acids Res **11**:2239-2255

Hazlewood GP and Dawson RMC (1979) Characteristics of a lipolytic and fatty acid-requiring *Butyrivibrio* sp. isolated from the ovine rumen. J Gen Microbiol **112**:15-27

Hazlewood GP and Edwards R (1981) Proteolytic activities of a rumen bacterium, *Bacteroides ruminicola* R8/4. J Gen Microbiol **125**:11-15

Hazlewood GP, Jones GA and Mangan JL (1981) Hydrolysis of leaf fraction 1 protein by the proteolytic rumen bacterium *Bacteroides ruminicola* R8/4. J Gen Microbiol **123**:223-232

Hazlewood GP and Nugent JHA (1978) Leaf Fraction 1 protein as a nitrogen source for the growth of a proteolytic rumen bacterium. J Gen Microbiol **106**:369-371

Hazlewood GP, Orpin CG, Greenwood Y and Black ME (1983) Isolation of proteolytic rumen bacteria by use of selective medium containing leaf fraction 1 protein (ribulosebisphosphate carboxylase). Appl Environ Microbiol **45**:1780-1784

Hazlewood GP, Romaniec MPM, Davidson K, Grepinet 0, Beguin P, Millet J, Raynaud O and Aubert J-P (1988) A catalogue of *Clostridium thermocellum* endoglucanase, β-glucosidase and xylanase genes cloned in *Escherichia coli*. FEMS Microbiol Lett **51**:231-236

Hazlewood GP and Teather RM (1988) The genetics of rumen bacteria. In: PN Hobson (ed) The rumen microbial ecosystem. Elsevier Applied Science, London, pp 323-341

Hazlewood GP, Theodorou MK, Hutchings A, Jordan DJ and Galfre G (1986) Preparation and characterization of monoclonal antibodies to a *Butyrivibrio* sp. and their potential use in the identification of rumen butyrivibrios, using an enzyme-linked immunosorbent assay. J Gen Microbiol **132**:43-52

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

Henikoff S (1987) Unidirectional digestion with exonuclease III in DNA sequence analysis. In: R Wu (ed) Methods and enzymology. Academic Press, New-York, pp 155-165

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

Henrissat B, Popineau Y and Kader JC (1988) Hydrophobic cluster analysis of plant protein sequences. A domain homology between storage and lipid transfer proteins. Biochem J **255**:901-905

Hespell RB (1984) Influence of ammonia assimilation pathways and survival strategy on rumen microbial growth. In: FMC Gilchrist and RI Mackie (eds) Herbivore Nutrition in the subtropics and tropics. The Science Press, Craighall, pp 346-358

Hespell RB, Wolf R and Bothast RJ (1987) Fermentation of xylans by *Butyrivibrio fibrisolvens* and other ruminal bacteria. Appl Environ Microbiol **53**:2849-2853

Heussen C and Dowdle EB (1980) Electrophoretic analysis of plasminogen activators in polyacrylamide gels containing sodium dodecyl sulfate and copolymerized substrates. Anal Biochem **102**:196-202

Hewett MJ, Wicken AJ, Knox KW and Sharpe ME (1976) Isolation of lipoteichoic acids from *Butyrivibrio fibrisolvens*. J Gen Microbiol **94**:126-130

Hiltner P and Dehority BA (1983) Effect of soluble carbohydrates on digestion of cellulose by pure cultures of rumen bacteria. Appl Environ Microbiol **46**:642-648

Hobson PN and Wallace RJ (1982) Microbial ecology and activities in the rumen. Crit Rev Microbiol **9**:253-320

Hoffman CS and Wright A (1985) Fusions of secreted proteins to alkaline phosphatase: An approach for studying protein secretion. Proc Natl Acad Sci USA **82**:5107-5111

Howard BH and Hungate RE (1976) *Desulfovibrio* of the sheep rumen. Appl Environ Microbiol **32**:598-602

Huang L and Forsberg CW (1987) . Isolation of a cellodextrinase from *Bacteroides succinogenes*. Appl Environ Microbiol **53**:1034-1041

Huang L and Forsberg CW (1988) Purification and comparison of the periplasmid and extracellular forms of the cellodextrinase from *Bacteroides succinogenes*. Appl Environ Microbiol **54**:1488-1493

Huang L, Forsberg CW and Thomas DY (1988) Purification and characterization of a chloride-stimulated cellobiosidase from *Bacteroides succinogenes* S85. J Bacteriol **170**:2923-2932

Hullah WA and Blackburn TH (1971) Uptake and incorporation of amino acids and peptides by *Bacteroides amylophilus*. Appl Microbiol **21**:187-191

Hungate RE (1966) The rumen and its microbes. Academic Press, New York

Ish-Horowicz D and Burke JF (1981) Rapid and efficient cosmid cloning. Nucleic Acids Res 9:2989-2998

Johnson EA, Bouchot F and Demain AL (1985) Regulation of cellulase formation in *Clostridium thermocellum*. J Gen Microbiol **131**:2303-2308

Johnson EA, Reese ET and Demain AL (1982a) Inhibition of *Clostridium thermocellum* cellulase by end products of cellulolysis. J Appl Biochem **4**:64-71

Johnson EA, Sakajoh M, Halliwell G, Madia A and Demain AL (1982b) Saccharification of complex cellulosic substrates by the cellulase system from *Clostridium thermocellum*. Appl Environ Microbiol **43**:1125-1132

Joliff G, Béguin P and Aubert J-P (1986a) Nucleotide sequence of the cellulase gene *celD* encoding endoglucanase D of *Clostridium thermocellum*. Nucleic Acids Res **14**:8605-8613

Joliff G, Béguin P, Juy M, Millet J, Ryter A, Poljak R and Aubert J-P (1986b) Isolation, crystallization and properties of a new cellulase of *Clostridium thermocellum* overproduced in *Escherichia coli*. Biotechnol **4**:896-900

Joyner AE and Baldwin RL (1966) Enzymatic studies of pure cultures of rumen microorganisms. J Bacteriol **92**:1321-1330

Juy M, Amit A, Alzari P, Béguin P, Millet J, Aubert J-P and Poljak RJ (1988) Crystallization and X-ray crystallographic study of the endoglucanase D of *Clostridium thermocellum*. In: J-P Aubert et al. (eds) Biochemistry and genetics of cellulose degradation. Academic Press,London, FEMS Symp **43**:379-383

Knowles J, Lehtovaara P and Teeri T (1987) Cellulase families and their genes. Tibtech **5**:255-261

Knowles J, Teeri TT, Lehtovaara P, Pentilla M and Saloheimo M (1988) The use of gene technology to investigate fungal cellulolytic enzymes. In: J-P Aubert et al. (eds) Biochemistry and genetics of cellulose degradation. Academic Press,London, FEMS Symp **43**:153-169

Kock SG and Kistner A (1969) Extent of solubilization of α-cellulose and hemicellulose of low-protein teff hay by pure cultures of cellulolytic rumen bacteria. J Gen Microbiol **55**:459-462

Kopecny J (1986) Cellulolytic enzymes of *Butyrivibrio fibrisolvens*. Abstract P B17-25, XIV International Congress of Microbiol, p 104

Kopecny J and Wallace RJ (1982) Cellular location and some properties of proteolytic enzymes of rumen bacteria. Appl Environ Microbiol **43**:1026-1033

Kretsinger R (1980) CRC Crit Rev Biochem 8:119-174

Krikos A, Mutoh N, Boyd A and Simon MI (1983) Sensory transducers of *E. coli* are composed of discrete structural and functional domains. Cell 33:615-622

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

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680-685

Lamed P, Setter E and Bayer EA (1983a) Characterization of a cellulose-binding, cellulase-containing complex in *Clostridium thermocellum.* J Bacteriol 156:828-836

Lamed P, Setter E, Kenig R and Bayer EA (1983b) The cellulosome - a discrete cell surface organelle of *Clostridium thermocellum* which exhibits separate antigenic, cellulose-binding and various cellulolytic activities. Biotechnol Bioeng Symp 13:163-181

Lamed R and Bayer EA (1988) The cellulosome of *Clostridium thermocellum.* Adv Appl Microbiol 33:2-41

Lamed R, Naimark J, Morgenstern E and Bayer EA (1987) Specialised cell surface structures in cellulolytic bacteria. J Bacteriol 169:3792-3800

Langsford ML, Gilkes NR, Singh B, Moser B, Miller Jr RC, Warren RAJ and Kilburn DG (1987) Glycosylation of bacterial cellulases prevents proteolytic cleavage between functional domains. FEBS Lett 225:163-167

Langsford ML, Gilkes NR, Wakarchuk WW, Kilburn DG, Miller Jr RC and Warren AJ (1984) The cellulase system of *Cellulomonas fimi.* J Gen Microbiol 130:1367-1376

Latham MJ (1980) Adhesion of rumen bacteria to plant cell walls. In: RCW Berkeley et al. (eds) Microbial Adhesion to surfaces. Ellis Horwood, Chichester, pp 339-350

Latham MJ, Brooker BE, Pettipher GL and Harris PJ (1978a) *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

Latham MJ, Brooker BE, Pettipher GL and Harris PJ (1978b) Adhesion of *Bacteroides succinogenes* in pure culture and in the presence of *Ruminococcus flavefaciens* to cell walls in leaves of perennial ryegrass (*Lolium perenne*). Appl Environ Microbiol 35:1166-1173

Latham MJ and Wolin MJ (1977) Fermentation of cellulose by *Ruminococcus flavefaciens* in the presence and absence of *Methanobacterium ruminantium.* Appl Environ Microbiol **34**:297-301

Leng RA and Nolan JV (1984) Nitrogen metabolism in the rumen. J Dairy Sci **70**:1072-1089

Lesk EM and Blackburn TH (1971) Purification of *Bacteroides amylophilus* protease. J Bacteriol **106**:394-402

Lin LL, Rumbak E, Zappe H, Thomson JA and Woods DR (1990) Cloning, sequencing and analysis of expression of a *Butyrivibrio fibrisolvens* gene encoding a β-glucosidase. J Gen Microbiol (in press)

Lupo D and Stutzenberg F (1988) Changes in endoglucanase patterns during growth of *Thermomonospora curvata* on cellulose. Appl Environ Microbiol **54**:588-599

Mackay RM, Lo A, Willick G, Zuker M, Baird S, Dove M, Moranelli F and Seligy V (1986) Structure of a *Bacillus subtilis* endo-β-1,4-glucanase gene. Nucleic Acids Res **14**:9159-9170

Mackie RI and Kistner A (1985) Some frontiers of research in basic ruminant nutrition. S Afr J Anim Sci **15**:72-85

Maniatis T, Fritsch EF and Sambrook J (1982) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor,New York

Mann SP, Hazlewood GP and Orpin CP (1986) Characterization of a cryptic plasmid (pOM1) in *Butyrivibrio fibrisolvens* by restriction endonuclease analysis and its cloning in *Escherichia coli.* Curr Microbiol **13**:17-22

Mannarelli BM (1988) Deoxyribonucleic acid relatedness among strains of the species *Butyrivibrio fibrisolvens.* Int J Syst Bacteriol **38**:340-347

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

Manoil C and Beckwith J (1986) A genetic approach to analyzing membrane protein topology. Science **233**:1403-1408

Margherita SS and Hungate RE (1963) Serological analysis of *Butyrivibrio* from the bovine rumen. J Bacteriol **86**:855-860

Margherita SS, Hungate RE and Storz H (1964) Variation in rumen *Butyrivibrio* strains. J Bacteriol **87**:1304-1308

Marmur J (1961) A procedure for the isolation of deoxyribonucleic acid from micro-organisms. J Mol Biol 3:208-218

Martin SA and Dean RG (1989) Characterization of a plasmid from the ruminal bacterium *Selenomonas ruminantium*. Appl Environ Microbiol **55**:3035-3038

Mathison GW and Milligan LP (1971) Nitrogen metabolism in sheep. Br J Nutr **25**:351-366

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

McGavin M and Forsberg CW (1988) Isolation and characterization of endoglucanases 1 and 2 from *Bacteroides succinogenes*. J Bacteriol **170**:2914-2922

McGavin MJ, Forsberg CW, Crosby B, Bell AW, Dignard D and Thomas DY (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

McLaughlin JR, Murray CL and Rabinowitz JC (1981) Unique features in the ribosome binding site sequence of the gram-positive *Staphylococcus aureus* β-lactamase gene. J Biol Chem **256**:11283-11291

Miller Jr RC, Gilkes NR, Greenberg NM, Kilburn DG, Langsford ML and Warren RAJ (1988) *Cellulomonas fimi* cellulases and their genes. In: J-P Aubert et al. (eds) Biochemistry and genetics of cellulose degradation. Academic Press, London, FEMS Symp **43**:235-248

Miller GL (1959) Use of dinitrosalicylic acid reagent for determination of reducing sugar. Anal Chem **31**:426-428

Miller JH (1972) Experiments in Molecular Genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor,New York

Millet J, Petre D, Béguin P, Raynaud O and Aubert J-P (1985) Cloning of ten distinct DNA fragments of *Clostridium thermocellum* coding for cellulases. FEMS Microbiol Lett **29**:145-149

Minato H and Suto T (1981) Technique for fractionation of bacteria in rumen microbial ecosystem. IV. Attachment of rumen bacteria to cellulose powder and elution of bacteria attached to it. J Gen Appl Microbiol **27**:21-31

Mink RW and Hespell RB (1981a) Long-term nutrient starvation of continuously cultured (glucose-limited) *Selenomonas ruminantium*. J Bacteriol **148**:541-550

Mink RW and Hespell RB (1981b) Survival of *Megasphaera elsdenii* during starvation. Curr Microbiol 5:51-56

Montgomery L, Flesher B and Stahl D (1988) Transfer of *Bacteroides succinogenes* (Hungate) to *Fibrobacter* gen. nov. as *Fibrobacter succinogenes* comb. nov. and description of *Fibrobacter intestinalis* sp. nov. Int J Syst Bacteriol **38**:430-435

Moore WEC, Johnson JL and Holdeman LV (1976) Emendation of *Bacteroidaceae* and *Butyrivibrio* and descriptions of *Desulfomonas* gen. nov. and ten new species in the genera *Desulfomonas, Butyrivibrio, Eubacterium, Clostridium,* and *Ruminococcus.* Int J Syst Bacteriol **26**:238-252

Moran CP Jr, Lang N, Le Grice SFJ, Lee G, Stephens M, Sonenshein AL, Pero J and Losick R (1982) Nucleotide sequences that signal the initiation of transcription and translocation in *Bacillus subtilis.* Mol Gen Genet **186**:339-346

Morgenstern E, Bayer EA, Vliegenthart JFG and Gerwig G (1987) Abstracts of FEMS Symposium: Biochemistry and genetics of cellulose degradation. Societe Francaise de Microbiologie, Paris, p 56

Morris EJ (1988) Characteristics of the adhesion of *Ruminococcus albus* to cellulose. FEMS Microbiol Lett **51**:113-118

Morris EJ and Cole OJ (1987) Relationship between cellulolytic activity and adhesion to cellulose in *Ruminococcus albus.* J Gen Microbiol **113**:1023-1032

Morris EJ and Van Gylswyk NO (1980) Comparison of the action of rumen bacteria on the cell walls from *Eragrostis teff.* J Agric Sci Camb **95**:313-323

Moser B, Gilkes NR, Kilburn DG, Warren RAJ and Miller Jr RC (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

Moses EK, Rood JI, Yong WK and Riffkin GG (1989) Molecular analysis of one of multiple protease-encoding genes from the prototype virulent strain of *Bacteroides nodosus.* Gene **77**:219-228

Mountfort DO and Asher RA (1983) Role of catabolite regulatory mechanisms in control of carbohydrate utilization by the rumen anaerobic fungus *Neocallimastix frontalis.* Appl Environ Microbiol **46**:1331-1338

Munn EA, Hazlewood GP and Graham M (1983) Uptake and incorporation of the products of proteolysis by the rumen bacterium *Bacteroides ruminicola* R8/4. Curr Microbiol 8:317-320

Nakai R, Horinouchi S and Beppu T (1988) Cloning and sequencing of a cellulase gene, *casA,* from an alkalophilic *Streptomyces* strain. Gene **65**:229-238

Ng TK, Ben-Bassat A and Zeikus JG (1981) Ethanol production by thermophilic bacteria: fermentation of cellulosic substrates by cocultures of *Clostridium thermocellum* and *Clostridium thermohydrosulfuricum*. Appl Environ Microbiol **41**:1337-1343

Nolan JV and Leng RA (1972) Dynamic aspects of ammonia and urea metabolism in sheep. Br J Nutr **27**:177-194

Norrander J, Kempe T, and Messing J (1983) Construction of improved M13 vectors using oligonucleotide-directed mutagenesis. Gene **261**:101-106

Nugent JHA, Jones WT, Jordan DJ and Mangan JL (1983) Rates of proteolysis in the rumen of the soluble proteins casein, Fraction I (18S) leaf protein, bovine serum albumin and bovine submaxillary mucoprotein. Br J Nutr **50**:357-368

Nugent JHA and Mangan JL (1981) Characteristics of the rumen proteolysis of Fraction 1 (18S) leaf protein from lucerne (*Medicago sativa* L). Brit J Nutr **46**:39-58

O'Neill G, Goh SH, Warren RAJ, Kilburn DG and Miller Jr RC (1986a) Structure of the gene encoding the exoglucanase of *Cellulomonas fimi*. Gene **44**:325-330

O'Neill G, Kilburn DG, Warren RAJ and Miller Jr RC (1986b) Overproduction from a cellulase gene with a high guanine plus cytosine content in *Escherichia coli*. Appl Environ Microbiol **52**:737-743

Ohmiya K, Shimizu M, Kurachi Y and Shimizu S (1985) Isolation and properties of β-glucosidase from *Ruminococcus albus*. J Bacteriol **161**:432-434

Ohmiya K, Shimizu M, Taya M and Shimizu S (1982) Purification and properties of cellobiosidase from *Ruminococcus albus*. J Bacteriol **150**:407-409

Ong E, Gilkes NR, Warren RAJ, Miller Jr RC and Kilburn DG (1989) Enzyme immobilization using the cellulose-binding domain of a *Cellulomonas fimi* exoglucanase. Biotechnol **7**:604-607

Orpin CG (1975) Studies on the rumen flagellate *Neocallimastix frontalis*. J Gen Microbiol **91**:249-262

Orpin CG (1977) The occurrence of chitin in the cell walls of the rumen organisms *Neocallimastix frontalis*, *Piromonas communis* and *Sphaeromonas communis*. J Gen Microbiol **99**:215-218

Orpin CG (1988) Genetic approaches to the improvement of lignocellulose degradation in the rumen. In: J-P Aubert et al. (eds) Biochemistry and genetics of cellulose degradation. Academic Press, London, FEMS Symp **43**:171-179

Orpin CG, Mathiesen SD, Greenwood Y and Blix AS (1985) Seasonal changes in the ruminal microflora of the high-arctic Svalbard reindeer (*Rangifer tarandus plytyrhynchus*). Appl Environ Microbiol **50**:144-151

Owalbi JB, Béguin P, Kilburn DG, Miller Jr RC and Warren RAJ (1988) Expression in *Escherichia coli* of the *Cellulomonas fimi* structural gene for endoglucanase B. Appl Environ Microbiol **54**:518-523

Paice MG, Desrochers M, Rho D, Jurasek L, Roy C, Rollin CF, De Miguel E and Yaguchi M (1984) Two forms of endoglucanase from the basidiomycete *Schizophyllum commune* and their relationship to other β-1,4-glucoside hydrolases. Biotechnol **2**:535-539

Pardee AB, Jacob F and Monod J (1959) The genetic control and cytoplasmic expression of "inducibility" in the synthesis of beta-galactosidase by E. coli. J Mol Biol **1**:165-178

Patterson JA (1989) Prospects for establishment of genetically engineered microorganisms in the rumen. Enzyme Microb Technol **11**:187-189

Paynter MJB and Hungate RE (1968) Characterization of *Methanobacterium mobilis* sp. n., isolated from the bovine rumen. J Bacteriol **95**:1943-1951

Penttilä M, Lehtovaara P, Nevalainen H, Bhikhabhai R and Knowles J (1986) Homology between cellulase genes of *Trichoderma reesei*: Complete nucleotide sequence of the endoglucanase I gene. Gene **45**:253-263

Pettipher GL and Latham MJ (1979) Characteristics of enzymes produced by *Ruminococcus flavefaciens* which degrade plant cell walls. J Gen Microbiol **110**:21-27

Phillips RW (1981) The upper limits of livestock production. In: New Prot Foods. Academic Press, New York **4**:5-12

Pilgrim AF, Gray FV, Weller RA and Belling GB (1970) Synthesis of microbial protein from ammonia in the sheep's rumen and the proportion of dietary nitrogen converted into microbial N. Br J Nutr **24**:589-598

Pittman KA and Bryant MP (1964) Peptides and other nitrogen sources for growth of *Bacteroides ruminicola*. J Bacteriol **88**:401-410

Postle K and Good RF (1983) DNA sequence of the *Esherichia coli tonB* gene. Proc Natl Acad Sci USA **30**:5235-5239

Rasmussen MA, Hespell RB, White BA and Bothast RJ (1988) Inhibitory effects of methylcellulose on cellulose degradation by *Ruminococcus flavefaciens*. Appl Environ Microbiol **54**:890-897

Reed KC and Mann DA (1985) Rapid transfer of DNA from agarose gels to nylon membranes. Nucleic Acids Res **13**:7207-7221

Remaut E, Tsao H and Fiers W (1983) Improved plasmid vectors with a thermoinducible expression and temperature-regulated runaway replication. Gene **22**:103-113

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

Roché C, Albertyn H, Van Gylswyk NO and Kistner A (1973) The growth response of cellulolytic acetate-utilizing and acetate-producing butyrivibrios to volatile fatty acids and other nutrients. J Gen Microbiol **78**:253-260

Rosenberg M and Court D (1979) Regulatory sequences involved in the promotion and termination of RNA transcription. Annu Rev Genet **13**:319-353

Russell JB, Bottje WG and Cotta MA (1981) Degradation of protein by mixed cultures of rumen bacteria: identification of *Streptococcus bovis* as an actively proteolytic rumen bacterium. J Anim Sci **53**:242-252

Russell JB and Wilson DB (1988) Potential opportunities and problems for genetically altered rumen microorganisms. Br J Nutr **118**:271-279

Sakka K, Furuse S and Shimada K (1989) Cloning and expression in *Escherichia coli* of the thermophilic *Clostridium* sp. F1 genes related to cellulose hydrolysis. Agric Biol Chem **53**:905-911

Saloheimo M, Lehtovaara P, Pentilla M, Teeri TT, Stahlberg J, Johansson G, Pettersson G, Claeyssens M, Tomme P and Knowles JKC (1988) EGIII, a new endoglucanase from *Trichoderma reesei*: and the characterization of both gene and enzyme. Gene **63**:11-21

Salser W (1977) Globin mRNA sequences: Analysis of base-pairing and evolutionary implications. CSH Symp Quant Biol **42**:985-1002

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

Saul DJ, Williams LC, Love DR, Chamley LW and Bergquist PL (1989) Nucleotide sequence of a gene from *Caldocellum saccharolyticum* encoding exocellulase and endocellulase activity. Nucleic Acids Res **17**:439-439

Scheifinger CC and Wolin MJ (1973) Propionate formation from cellulose and soluble sugars by combined cultures of *Bacteroides succinogenes* and *Selenomonas ruminantium*. Appl Microbiol **29**:480-483

Scholle RR, Robb SM, Robb FT and Woods DR (1989) Nucleotide sequence and analysis of the *Vibrio alginolyticus* sucrase gene (*scrB*). Gene **80**:49-56

Schwarz WH, Schimming S, Rucknagel KP, Burgschwaiger S, Kreil G and Staudenbauer WL (1988a) Nucleotide sequence of the *celC* gene encoding endoglucanase C of *Clostridium thermocellum*. Gene **63**:23-30

Schwarz WH, Schimming S and Staudenbauer WL (1987) High-level expression of *Clostridium thermocellum* cellulase genes in *Escherichia coli*. Appl Microbiol Biotechnol **27**:50-56

Schwarz WH, Schimming S and Staudenbauer WL (1988b) Degradation of barley β-glucan by endoglucanase C of *Clostridium thermocellum*. Appl Microbiol Biotechnol **29**:25-31

Shane BS, Gouws L and Kistner A (1969) Cellulolytic bacteria occurring in the rumen of sheep conditioned to low-protein teff hay. J Gen Microbiol **55**:445-457

Sharma S and Sandhu DK (1986) An improved method for detection of isozymes of endocellulases. Indian J Exp Biol **24**:732-733

Sharpe ME, Brock JH and Phillips BA (1975) Glycerol teichoic acid as an antigenic determinant in a gram-negative bacterium *Butyrivibrio fibrisolvens*. J Gen Microbiol **88**:355-363

Shoemaker S, Schweickart V, Ladner M, Gelfand D, Kwok S, Myambo K and Innis M (1983) Molecular cloning of exo-cellobiohydrolase I derived from *Trichoderma reesei* strain L27. Biotechnol **1**:691-696

Short JM, Fernandez JM, Sorge JA, and Huse WD (1988) λ ZAP: a bacteriophage λ expression vector with *in vivo* excision properties. Nucleic Acids Res **16**:7583-7600

Simms PFG, James C and Broda P (1988) The identification, molecular cloning and characterization of a gene from *Phanerochaete chrysosporium* that shows strong homology to the exo-cellobiohydrolase I gene from *Trichoderma reesei*. Gene **74**:411-422

Smith CJ, Hespell RB and Bryant MP (1980) Ammonia assimilation and glutamate formation in the anaerobe *Selenomonas ruminantium*. J Bacteriol **141**:593-602

Smith CJ, Hespell RB and Bryant MP (1981) Regulation of ammonia assimilatory enzymes and urease in *Selenomonas ruminantium*. Appl Environ Microbiol **42**:89-96

Smith PH and Hungate RE (1958) Isolation and characterization of *Methanobacterium ruminantium* n. sp. J Bacteriol **75**:713-718

Stack RJ (1988) Neutral sugar composition of extracellular polysaccharides produced by strains of *Butyrivibrio fibrisolvens*. Appl Environ Microbiol **54**:878-883

Stack RJ and Cotta MA (1986) Effect of 3-phenylpropanoic acid on growth and cellulose utilisation by cellulolytic ruminal bacteria. Appl Environ Microbiol **52**:209-210

Stack RJ and Hungate RE (1984) Effect of 3-phenylpropanoic acid on capsule and cellulases of *Ruminococcus albus* 8. Appl Environ Microbiol **48**:218-223

Stanton TB and Canale-Parole E (1980) *Treponema bryantii* sp. nov., a rumen spirochete that interacts with cellulolytic bacteria. Arch Microbiol **127**:145-156

Starr MP, Stolp H, Truper HG, Balows A and Schlegel HG (1981) The Prokaryotes. Volume I. Springer-Verlag, Berlin.

Stewart BJ and Leatherwood JM (1976) Derepressed synthesis of cellulase by *Cellulomonas*. J Bacteriol **128**:609-615

Struhl K (1985) A rapid method for creating recombinant DNA molecules. Biotechn 3:452-453

Strydom E, Mackie RI and Woods DR (1986) Detection and characterization of extracellular proteases in *Butyrivibrio fibrisolvens* H17c. Appl Microbiol Biotechnol **24**:214-217

Sykes RB and Nordstrom K (1972) Microiodometric determination of β-lactamase activity. Antimicrob Agents Chemother **1**:94-99

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

Tagari H, Ascarelli I and Bondi A (1962) The influence of heating on the nutritive value of soya-bean meal for ruminants. Br J Nutr **16**:237-241

Tamminga S (1979) Protein degradation in the forestomachs of ruminants. J Anim Sci **49**:1615-1630

Taylor KA, Crosby B, McGavin M, Forsberg CW and Thomas DY (1987) Characteristics of the endoglucanase encoded by a *cel* gene from *Bacteroides succinogenes* expressed in *Escherichia coli*. Appl Environ Microbiol **53**:41-46

Teather RM (1982) Isolation of plasmid DNA from *Butyrivibrio fibrisolvens*. Appl Environ Microbiol **43**:298-302

Teather RM (1985) Application of gene manipulation to rumen microflora. Can J Anim Sci **65**:563-574

Teather RM and Wood PJ (1982) Use of Congo red-polysaccharide interactions in enumeration and characterization of cellulolytic bacteria from the bovine rumen. Appl Environ Microbiol **43(4)**:777-780

Teeri TT, Lehtovaara P, Kauppinen S, Salovuori I and Knowles J (1987) Homologous domains in *Trichoderma reesei* cellulolytic enzymes: gene sequence and expression of cellobiohydrolase II. Gene **51**:43-52

Tomme P, Van Tilbeurgh H, Pettersson G, Van Damme J and Vandekerckhove J (1988) Studies of the cellulolytic system of *Trichoderma reesei* QM 9414. Eur J Biochem **170**:575-581

Tucker ML, Durbin ML, Clegg MT and Lewis LN (1987) Avocado cellulase: Nucleotide sequence of a putative full-length cDNA clone and evidence for a small gene family. Plant Mol Biol **9**:197-203

Ulyatt MJ, Dellow DW, John A, Reid CSW and Waghorn GC (1986) Contribution of chewing during eating and rumination to the clearance of digesta from the ruminoreticulum. In: A Dobson et al. (eds) Control of digestion and metabolism in ruminants. Prentice-Hall Inc., Englewood Cliffs, New York, pp 494-575

Van Gylswyk NO and Schwartz HM (1984) Microbial ecology of the rumen of animals fed high-fibre diets. In: FMC Gilchrist and RI Mackie (eds) Herbivore nutrition in the subtropics and tropics. The Science Press, Craighall, New York, p 359-377

Van Soest PJ (1973) The uniformity and nutritive availability of cellulose. Fed Proc **32**:1804-1808

Van Tilbeurgh H, Claeyssens M and De Bruyne CK (1982) The use of 4-methylumbelliferyl and other chromophoric glycosides in the study of cellulolytic enzymes. FEBS Lett **149**:152-156

Van Tilbeurgh H, Tomme P, Claeyssens M, Bhikhabhai R and Pettersson G (1986) Limited proteolysis of the cellobiohydrolase I from *Trichoderma reesei*. FEBS Lett **204**:223-227

Varel VH and Dehority BA (1989) Ruminal cellulolytic bacteria and protozoa from bison, bison hybrids, and cattle fed three levels of alfalfa. Appl Environ Microbiol **55**:148-155

Von Heijne G (1983) Patterns of amino acids near signal sequence cleavage sites. Eur J Biochem **133**:17-21

Von Heijne G (1985) Signal sequences: the limits of variation. J Mol Biol **184**:99-105

Wallace RJ and Brammall ML (1985) The role of different species of bacteria in the hydrolysis of protein in the rumen. J Gen Microbiol **131**:821-832

Wallace RJ and McKain N (1989) Analysis of peptide metabolism by ruminal microorganisms. Appl Environ Microbiol **55**:2372-2376

Wang W and Thomson JA (1990) Nucleotide sequence of the *celA* gene encoding a cellodextrinase of *Ruminococcus flavefaciens* FD1. Mol Gen Genet (in press)

Warren RAJ, Beck CF, Gilkes NR, Kilburn DG, Langsford ML, Miller Jr RC, O'Neill GP, Scheufens M and Wong WKR (1987a) Sequence conservation and region shuffling in an endoglucanase and an exoglucanase from *Cellulomonas fimi* Pro Struct Function Genet **1**:335-341

Warren RAJ, Gerhard B, Gilkes NR, Owalbi JB, Kilburn DG and Miller Jr RC (1987b) A bifunctional exoglucanase-endoglucanase fusion protein. Gene **61**:421-427

Williams AG and Withers SE (1982) The production of plant cell wall polysaccharide-degrading enzymes by hemicellulolytic rumen bacterial isolates grown on a range of carbohydrate substrates. J Appl Bacteriol **52**:377-387

Willis RC, Morris RG, Cirakoglu C, Schellenberg GD, Gerber NH and Furlong CE (1974) Preparation of the periplasmic binding proteins from *Salmonella typhimurium* and *Escherichia coli*. Arch Biochem Biophys **161**:64-75

Wirick MG (1968a) A study of the enzymic degradation of CMC and other cellulose ethers. J Polymer Sci **6**:1965-1974

Wirick MG (1968b) Study of the substitution pattern of hydroxyethylcellulose and its relationship to enzymic degradation. J Polymer Sci **6**:1705-1718

Wolin MJ, Wolin EA and Jacobs NJ (1961) Cytochrome-producing anaerobic vibrio, *Vibrio succinogenes* sp. n. J Bacteriol **81**:911-917

Wong WK, Gerhard B, Guo ZM, Kilburn DG, Anthony R, Warren J and Miller RC Jr (1986) Characterization and structure of an endoglucanase gene *cenA* of *Cellulomonas fimi*. Gene **44**:315-324

Wood DA (1985) In: KW Fuller and JR Gallon (eds) Plant products and the new technology. Ann Proc Phytochem Soc Europe **26**:295-309

Wood TM and McCrae SI (1972) The purification and properties of the C1 component of *T. koningii* cellulase. Biochem J **128**:1183-1192

Wood TM and McCrae SI (1977). In: TK Ghosh (ed) Proc Bioconversion Symp, Delhi, pp 111-114

Wood TM and McCrae SI (1986) Studies of two low-molecular weight endo-(1-4)-β-D-xylanases constitutively synthesized by the cellulolytic fungus *Trichoderma koningii*. Carbohydr Res **148**:321-330

Wood TM, McCrae SI, Wilson CA, Bhat KM and Gow LA (1988) Aerobic and anaerobic fungal cellulases, with special reference to their mode of attack on crystalline cellulose. In: J-P Aubert et al. (eds) Biochemistry and genetics of cellulose degradation. Academic Press, London, FEMS Symp **43**:30-52

Wood TM, Wilson CA, McCrae SI and Joblin KN (1986) A highly active extracellular cellulase from the anaerobic rumen fungus *Neocallimastix frontalis*. FEMS Microbiol Lett **34**:37-40

Wood TM, Wilson CA and Stewart CS (1982) Preparation of cellulase from the cellulolytic anaerobic rumen bacterium *Ruminococcus albus* and its release from the bacterial cell wall. Biochem J **205**:129-137

Wood WB (1966) Host specificity of DNA produced by *Escherichia coli*: Bacterial mutations affecting the restriction and modification of DNA. J Mol Biol **16**:118-133

Wu JHD, Orme-Johnson WH and Demain AL (1988) Two components of an extracellular protein aggregate of *Clostridium thermocellum* together degrade crystalline cellulose. Biochem **27**:1703-1709

Yablonsky MD, Bartley T, Elliston KO, Kahrs SK, Shalita ZP and Eveleigh DE (1988) Characterization and cloning of the cellulase complex of *Microbispora bispora*. In: J-P Aubert et al. (eds) Biochemistry and genetics of cellulose degradation. Academic Press,London, FEMS Symp **43**:249-266

Yablonsky MD, Elliston KO and Eveleigh DE (1989) The relationship between the endoglucanase *MbcelA* of *Microbispora bispora* and the cellulases of *Cellulomonas fimi*. In: MP Coughlan (ed) Enzyme systems for lignocellulose degradation. Elsevier Appl Sci,London and New York, pp 73-83

Yaguchi M, Roy C, Rollin CF, Paice MG and Jurasek L (1983) A fungal cellulase shows sequence homology with the active site of hen egg-white lysozyme. Biochem Biophys Res Comm **116**:408-411

Yamane K and Suzuki H (1988) Cellulases of *Pseudomonas fluorescens* var. *cellulosa*. Meth Enzymol **160**:200-210

Yamane K, Suzuki H and Nisizawa K (1970) Purification and properties of extracellular and cell-bound cellulase components of *Pseudomonas fluorescens* var. *cellulosa*. J Biochem **67**:19-35

Yanisch-Perron C, Vieira J and Messing J (1985) Improved M13 phage cloning vectors and host strains: Nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**:103-119

Young RA and Davis RW (1983) Efficient isolation of genes by using antibody probes. Proc Natl Acad Sci USA **80**:1194-1198

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

Zappe H, Jones WA, Jones DT and Woods DR (1988) Structure of an endo-B-1,4-glucanase gene from *Clostridium acetobutylicum* P262 showing homology with endoglucanase genes from *Bacillus* spp. Appl Environ Microbiol **54**:1289-1292

Ziolecki A (1979) Isolation and characterization of large treponemes from the bovine rumen. Appl Environ Microbiol **37**:131-135

Ziolecki A and Wojciechowicz M (1980) Small pectinolytic spirochetes from the rumen. Appl Environ Microbiol **39**:919-922

Zvelebil MJJM and Sternberg MJE (1988) Analysis and prediction of the location of catalytic residues in enzymes. Protein Eng **2**:127-138