

THE CLONING AND CHARACTERISATION OF AN
ENDOGLUCANASE AND AN ENDOXYLANASE FROM
CLOSTRIDIUM ACETOBUTYLICUM IN ESCHERICHIA
COLI

HAROLD ZAPPE

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

CAPE TOWN
FEBRUARY, 1988

The University of Cape Town has been given
the right to reproduce this thesis in whole
or in part. Copyright is held by the author.

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

CERTIFICATION OF SUPERVISORS

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

Professor D. R. Woods
Deputy Vice-Chancellor (Research)
University of Cape Town.
Director Microbial Research Units

Professor D. T. Jones
Head of Department of Microbiology
University of Cape Town.

ACKNOWLEDGEMENTS

I would like to thank my supervisors, Professor Dave Jones for advice and encouragement and Professor Dave Woods for his enthusiasm and helpful criticism during the course of this study.

I am grateful to Dr. M. Zabeau for the gift of the cloning vector pEcoR251 and to Dr. D. E. Eveleigh and Dr. N. Creuzet for preprints of their papers.

Many thanks to Wyn Jones for technical guidance in the nucleotide sequencing of the endoglucanase gene and to Mary-Lou Henderson for her kindness.

A very special thanks to my wife Cathy, for her unflinching support during the course of this study.

I am grateful for the financial support which I received from the C. S. I. R. and Sentrachem Ltd., South Africa.

CONTENTS

Acknowledgements	iii
Abstract	v
Abbreviations	viii
Chapter 1 General Introduction	1
Chapter 2 Construction and screening of a <u>Clostridium acetobutylicum</u> P262 genomic library in <u>Escherichia coli</u>	65
Chapter 3 Cloning and expression of a xylanase gene from <u>Clostridium acetobutylicum</u> P262 in <u>Escherichia coli</u>	86
Chapter 4 Cloning and expression of an endoglucanase gene from <u>Clostridium acetobutylicum</u> P262 in <u>Escherichia coli</u>	116
Chapter 5 Structure of the cloned endoglucanase gene from <u>Clostridium acetobutylicum</u> P262 and comparison with other endoglucanase genes	143
Chapter 6 General Discussion	171
Appendix A Maintenance of bacterial strains	179
Appendix B Media, buffers and solutions	180
Appendix C General Methods	197
Appendix D Salt requirements for restriction enzymes	209
Appendix E Vector restriction maps	210
Appendix F One- and three-letter codes used for amino acids	213
Appendix G Standard curves for the calculation of xylanase and endoglucanase activities	214
References	215

ABSTRACT

Clostridium acetobutylicum P262 is an endospore forming Gram-positive obligate anaerobe which has been used for the industrial production of acetone and butanol. Strains of C. acetobutylicum have been reported to exhibit some activity towards cellulosic and hemicellulosic substrates.

The aim of this thesis was to establish a genebank of C. acetobutylicum P262 DNA in Escherichia coli and to isolate and characterise genes encoding enzymes which show activity towards hemicellulose and cellulose.

A library of chromosomal DNA fragments from C. acetobutylicum P262 was established in the plasmid pEcoR251, an E. coli positive selection vector. The presence of inserts was confirmed by the recovery of plasmids, larger than the parental vector, in a random selection of clones. The library was screened for expression of C. acetobutylicum genes by complementation of nutritional markers in E. coli. Complementation of E. coli auxotrophic markers argG6 and hisG1 was observed. E. coli clones expressing glutamine synthetase, alcohol dehydrogenase, xylanase and endoglucanase activities were also isolated.

The xylanase gene from C. acetobutylicum P262 was cloned on a recombinant plasmid pHZ300 which enabled E. coli HB101 cells to produce intracellular xylanase activity. The xylanase gene was located on a 1.9 kb DNA fragment and was

expressed from its own promoter in the E. coli host. The cloned xylanase had an apparent M_r of approximately 28 000 and an isoelectric point of approximately 10. Optimum xylanase activity was obtained at pH 6.0 at 37 - 43°C, similar characteristics to a xylanase partially purified from the culture medium of C. acetobutylicum P262. Western blot analysis showed cross-reactivity between antibodies raised against the purified cloned enzyme and a polypeptide of the same M_r from C. acetobutylicum P262. The xylanase gene was not induced by xylan but was subject to catabolite repression by glucose and xylose in the E. coli host.

An endo- β -1,4-glucanase gene was isolated from the library of C. acetobutylicum P262 DNA fragments cloned in E. coli HB101. Of 15 000 colonies screened, six produced clear zones on Congo red stained CMC agar. Plasmids recovered from five of these colonies, showed identical restriction endonuclease digestion profiles and contained the same 7.2 kb DNA insert. This plasmid was designated pHZ120. A plasmid isolated from the remaining colony contained a 4.9 kb insert and was designated pHZ100. Restriction analysis showed that the 4.9 kb fragment was contained within the 7.2 kb fragment. The endoglucanase gene was shown to be expressed from its own promoter in E. coli cells and 75% of the endoglucanase activity was located in the periplasm of the E. coli host. Cellobiase activity was detected, but there was no activity against Avicel. The endoglucanase was partially purified from E. coli(pHZ100) cell extracts and migrated as two bands on non-denaturing protein gels. The endoglucanase showed optimal activity at

pH 6.0 and 50°C. The enzyme was unstable at higher temperatures with <1% of activity remaining after 1 h at 60°C. The endoglucanase gene was not induced by CMC or cellobiose but was subject to catabolite repression by glucose in the E. coli host.

The nucleotide sequence of the cloned C. acetobutylicum endo- β -1,4-glucanase gene was determined. The upstream region of the endoglucanase gene contained two putative extended promoter consensus sequences characteristic of Gram-positive bacteria. The putative ribosome binding site consisted of a TTG initiation codon and a strong Shine-Dalgarno complementarity containing five G-C bp situated seven bp upstream from the initiation codon. The complete amino acid sequence (448 residues) of the C. acetobutylicum endoglucanase was deduced, and comparisons were made with reported amino acid sequences of endoglucanases from other organisms. A high degree of homology was demonstrated between the C. acetobutylicum endoglucanase and enzymes from Bacillus strains. There was no extensive homology between the C. acetobutylicum endoglucanase and enzymes from C. thermocellum.

Abbreviations

AB	acetone-butanol
ADH	alcohol dehydrogenase
AHG	anhydroglucose
A _m	molecular activity
Ap	ampicillin
ATCC	American Type culture collection
ATP	adenosine 5'-triphosphate
bp	base pair(s)
cAMP	adenosine 3':5'-cyclic monophosphate
CAP	catabolite activator protein
CBM	<u>Clostridium</u> basal medium
CBQ	cellobiose:quinone oxidoreductase
CI	crystallinity index
Cm	chloramphenicol
CsCl	caesium chloride
C-terminal	carboxy terminal (end of a protein)
CMC	carboxymethylcellulose
CMCase	carboxymethylcellulase
d	day(s)
DEP	diethyl pyrocarbonate
DNase	deoxyribonuclease
DP	degree of polymerisation
DS	degree of substitution
DTT	1,4-dithio-L-threitol
EDTA	ethylenediaminetetra-acetic acid
Em	erythromycin
GARP	goat anti rabbit IgG conjugated horseradish peroxidase
h	hour(s)
HEC	hydroxyethylcellulose
HEWL	hen egg-white lysozyme
IEF	isoelectric focusing
IPTG	isopropyl β-D-thiogalactopyranoside
kb	kilobase pairs
Km	kanamycin
LB	Luria-Bertani medium
MES	2(N-Morpholino)ethanesulfonic acid
MeUmb	methylumbelliferyl
min	minute(s)
M _r	relative molecular mass
NAD	nicotinamide adenine dinucleotide
NADP	nicotinamide adenine dinucleotide phosphate
4NPX	p-nitrophenyl-β-D-xylopyranoside
N-terminal	amino terminal (end of a protein)

ONPG	<u>o</u> -nitrophenyl- β -D-galactopyranoside
ORF	open reading frame
p	plasmid
PAGE	polyacrylamide gel electrophoresis
PC	phosphate-citrate (buffer)
pI	isoelectric point
pI _{-app}	apparent isoelectric point
r	resistance (superscript)
RBB	remazol brilliant blue
RBS	ribosome binding site
RF	replicative form DNA (of M13 bacteriophage)
RNAse	ribonuclease
s	second(s)
SDS	sodium dodecyl sulphate (laural sulphate)
Sm	streptomycin
SS	single stranded (DNA)
TCA	trichloroacetic acid
TNP-CMC	trinitrophenyl-carboxymethylcellulose
Tris	Tris(hydroxymethyl)aminomethane
Tween	polyoxyethylene sorbitan monolaurate
UV	ultraviolet (light)
X-gal	5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside
::	novel joint

Chapter 1

General Introduction

Biomass as a renewable resource. Plant biomass is the most abundant renewable resource on earth with an estimated 1.8×10^{12} tonnes in existence and approximately 1×10^{11} tonnes synthesised annually by photosynthesis (Coughlan, 1985a). Plant biomass forms an integral part of our lives providing us with food, fibres, construction materials, and fuel. Cellulose as the major component of plant biomass is the most abundant molecule with estimates of $4 - 5 \times 10^{10}$ tonnes synthesised annually, or 70 kg produced per person per day (Lutzen *et al.*, 1983). Vast quantities of lignocellulosic wastes are generated annually by a variety of agricultural, industrial and domestic activities and various estimates have been reported (e.g. Bisaria and Ghose, 1981; Wiegel, 1982; Klyosov, 1986) (Table 1.1).

Table 1.1. Estimated annual production of solid cellulose wastes in the USA (from Wiegel, 1982). The quantities do not reflect the quantities of hemicellulose, the other major utilisable component of lignocellulose.

Waste	million tons	estimated cellulose content (%)	million tons
Agriculture and food	400	60	240
Manure	200	50	100
Urban refuse	150	45	68
Wood	60	55	33
Industrial	45	33	15
Municipal sewage solids	15	33	5
Miscellaneous organic waste	70	25	17
Total:	940		478

In nature large numbers of organisms are involved in the complex biodegradation of lignocellulose. This process can however be too time consuming for practical applications such as the production of humus or compost. Co-cultures of cellulolytic and nitrogen fixing organisms on buried agricultural wastes have been suggested as a method to expedite degradation and reduce the need for expensive fertilizers (Veal and Lynch, 1984).

The enormous quantities of lignocellulosic wastes produced and the fact that they are renewable means that they have considerable potential as a fermentation substrate for the production of food, fuel and chemical feedstocks. However, the major hurdle in the use of these materials as a substrate for fermentation is their resistance to hydrolysis. The crystalline structure of cellulose prevents the access of water, let alone any biological agent that could facilitate hydrolysis. This makes cellulases about 100 fold less efficient at hydrolysing cellulose than are amylases at hydrolysing starch (Saddler, 1986).

Structure of lignocellulose.

Lignocellulose is a complex material consisting of three main components which vary in proportion depending on the source. The major component is cellulose (40 - 60%) with smaller but significant amounts of hemicellulose (15 - 30%) and lignin (10 - 30%) (Dekker and Lindner, 1979). Certain woods contain as much as 35% hemicellulose (Timell, 1967).

Lignin is the most recalcitrant of the three components and its chemical structure consists of a complex polymer constructed of phenyl-propane repeat units (Amer and Drew, 1980) which can show a considerable amount of variation depending on its origin. Three categories of lignin can be identified: softwood (gymnosperm) lignin, hardwood (angiosperm) lignin and grass lignin. Lignins are produced by enzyme-initiated dehydrogenative polymerisation of three primary precursors: trans-coniferyl, trans-sinapyl and trans-p-coumaryl alcohols (Sarkanen and Ludwig, 1971). The three types of lignin vary in their proportions of these precursors. Lignin is closely associated with cellulose and hemicellulose in plant material and hinders the biodegradation of these substances primarily by preventing accessibility of the degradative enzymes.

Hemicellulose is a collective term used for polysaccharides that are associated with cellulose in plants (Dekker, 1985). The hemicelluloses form a distinct group of polysaccharides and have been classified according to the sugar residue present in the polymer (Timell, 1967; Aspinall, 1970; Wilkie, 1979). Three types predominate: polymers of galactose (1,3- and 1,4- β -D-galactans), mannose (1,4- β -D-mannans) or xylose (1,4- β -D-xylans). Unlike cellulose, they usually occur as heteroglycans (Wilkie, 1983) having side chains containing different types of sugar residues. In some cases the side chains can consist of single residues linked to the main glycan backbone (Dekker, 1985). Xylan makes up the major component of hemicellulose and is second in abundance to cellulose (Biely, 1985). Some xylans are

acetylated making them significantly more soluble in water (Biely, 1985).

The primary structure of cellulose is well documented (Alberts et al., 1983; Coughlan, 1985b; Eveleigh, 1987). It consists of a linear polymer of up to 14 000 anhydroglucose (AHG) units linked by β -1,4 linkages (Coughlan, 1985b). The average degree of polymerisation (DP) in naturally occurring cellulose (i.e. untreated) is 10 000, but can be as low as 15 (Eveleigh, 1987). Each glucose unit is rotated 180 degrees with respect to its neighbouring residues forming the basic repeating unit of cellobiose. The glucose polymers are orientated in parallel and arranged in a staggered fashion to form insoluble elementary fibrils in which the glucose chains are held together by intra- and intermolecular bonds (Alberts et al., 1983). Both hydrogen bonds and van der Waals forces promote interchain binding (Rees et al., 1982). Bundles of these fibrils aggregate to form the insoluble cellulose fibres (microfibrils). The fibres contain areas of complete order (crystalline regions) and areas of lesser order (amorphous regions). Although the bonding forces are strong, there is seldom perfect fibril association and as a result few celluloses are >90% crystalline (Eveleigh, 1987). A few exceptions are cotton (DP about 10 000), cell walls of the alga Valonia macrophysa (Chanzy and Henrissat, 1985) and the cellulose ribbons produced by Acetobacter xylinum (White and Brown, 1981). Commercial preparations such as filter paper and microcrystalline cellulose are not highly crystalline (Coughlan, 1985b; Eveleigh, 1987).

The three primary constituents of lignocellulose are intimately associated to form the structure of the plant cell wall. The crystalline cellulose fibrils form the structural framework and are set in a largely amorphous carbohydrate matrix of hemicellulose (Fig. 1.1). Rigidity and binding (adhesive) properties are provided by the amorphous polyphenolic material of which lignin is the major component (Raven *et al.*, 1976).

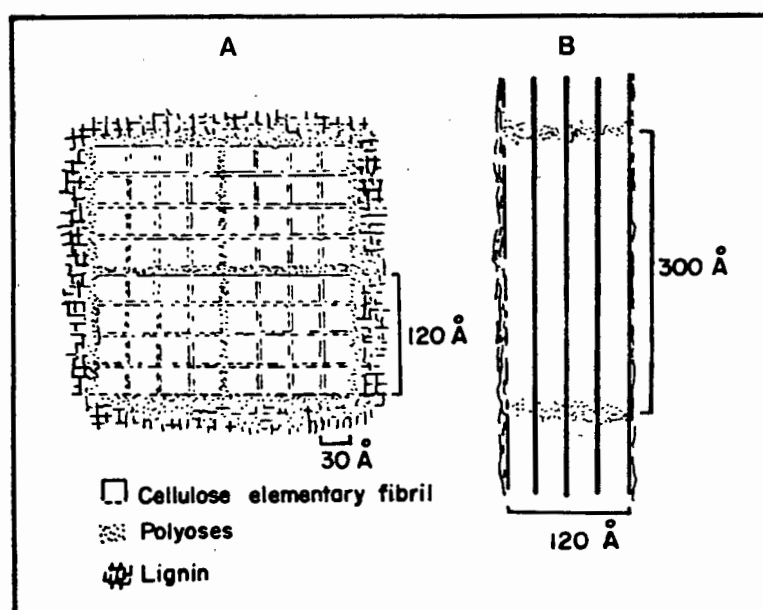


Fig. 1.1. Model of the ultrastructural organisation of the cell wall components in wood. A, cross section; B, longitudinal section (from Tsao, 1978).

The most obvious features of the plant cell structure are the primary and secondary cell walls. The primary wall contains 20 - 28% cellulose, while the secondary wall contains 45 - 55% cellulose. Three layers can often be distinguished in the secondary wall (S1, S2, S3) (Fig. 1.2). The layers differ in the orientation of their cellulose fibrils adding further structural strength to the cell wall. Lignin is found primarily in the middle lamella between the cell walls (Raven *et al.*, 1976) and therefore forms an

effective coating around the cellulose-hemicellulose matrix thus preventing access of hydrolytic enzymes (Fig. 1.2).

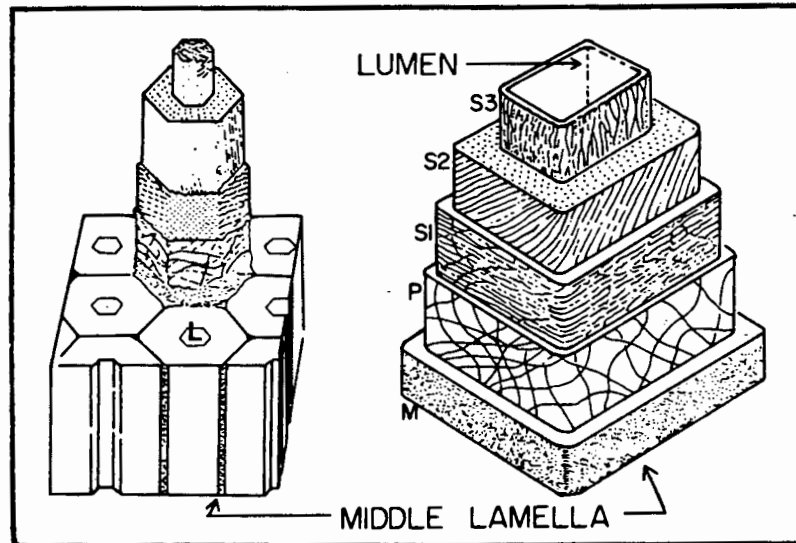


Fig. 1.2. Structure of a plant cell wall. M, middle lamella; P, primary wall; S1, S2, S3, three secondary wall layers; L, lumen (from Tsao, 1978).

Organisms involved in the degradation of lignocellulose.

A wide variety of microorganisms are known to produce enzymes which are capable of degrading naturally occurring lignocellulose. Some higher organisms have also been reported to produce cellulases and hemicellulases. The germinating seeds of plants have been shown to produce hemicellulases (see Dekker and Richards, 1976). These enzymes are involved in the release of storage carbohydrate during seed germination. Although many vertebrates are dependent on plant material as a source of food, hemicellulases could not be detected in the mucosal cells lining the alimentary canal of these animals or humans (see Dekker, 1985). These enzymes are produced by symbiotic bacteria and protozoa in ruminants and herbivores (see Dekker

and Richards, 1976), bacteria in the caecum (Bailey and MacRae, 1970) and bacteria in the human colon (Salyers et al., 1977). A number of studies have reported the presence of cellulases or hemicellulases in various invertebrate organisms including Coelenterata, Vermes, Arthropoda, Mollusca, Echinodermata, Crustacea (Gascoigne and Gascoigne, 1960). The origin of these enzymes however must be questioned as in some cases (e.g. Bivalva (shipworms)) it was later found that the source of the cellulase was symbiotic microorganisms (Waterbury et al., 1983).

In nature the degradation of lignocellulose is brought about almost entirely by microorganisms (e.g. see Dekker and Richards, 1976; Bisaria and Ghose, 1981). The environmental conditions under which the degradation occurs largely dictates the types of microorganisms that can be isolated. Generally fungi predominate in aerobic environments and bacteria in anaerobic environments, with certain exceptions.

Lignin degradation. Lignin is completely degraded by few microorganisms, although many can partially degrade the polymer (Crawford and Crawford, 1980). Fungi constitute the most active group of lignolytic microorganisms (Yang et al., 1980). White-rot (e.g. Polyporus verisicolor) and soft-rot fungi (e.g. Aspergillus fumigatus) have been identified as major lignin degraders, while brown-rot fungi (e.g. Lenzites trabea) show a preference for the carbohydrate components rather than lignin (see Amer and Drew, 1980). The ability to degrade lignin appears to depend both on the fungal

species, and the type of wood. The degradation of lignin by these fungi is dependent on the presence of oxygen and can only occur if a more easily metabolised carbon source is available simultaneously (Amer and Drew, 1980).

Certain bacteria such as Bacillus magaterium (see Crawford and Crawford, 1980) and certain Streptomyces species (Crawford, 1978) have also been reported to have lignolytic activity. However the role of bacteria in the degradation of lignin in nature is probably limited, at least in part, by the lack of ability to penetrate through the woody tissue when compared to fungal hyphae (Amer and Drew, 1980). A species of Xanthomonas has been reported to degrade lignin anaerobically, but as with fungi, a cosubstrate is required as well as NO_3^- as a terminal electron acceptor (Odier and Monties, 1978).

Hemicellulose degradation. Although hemicellulose consists of a number of different polysaccharides, xylan forms the major fraction. The most comprehensively studied of the hemicellulases are therefore the xylanases, and the occurrence of these enzymes is widespread (Dekker, 1985). Hemicellulose and cellulose always occur together in nature and, as would be expected, xylanases and cellulases can often be isolated from the same organisms. Xylanases have been isolated from many fungi and bacteria (see Dekker and Richards, 1976, Reilly, 1981; Biely, 1985; Dekker, 1985). Among the fungi, Reese et al. (1973) found a large number of xylanase producers from five genera: Aspergillus, Botryodiplodia, Penicillium, Pestalotia and Trichoderma.

Certain yeasts from the genera Aureobasidium, Cryptococcus and Trichosporon have also been reported to produce xylanases (Biely et al., 1978).

Xylanases are produced by many bacterial genera, most of which are anaerobes. These include rumen bacteria of various genera such as Bacteroides (Dehority, 1967; Forsberg et al., 1981), Butyrivibrio (Howard et al., 1960) and Ruminococcus (Dehority, 1967). Mesophilic and thermophilic Clostridium species have also been reported as xylanase producers (e.g. Lee et al., 1985b; Berenger et al., 1985). Aerobic xylanase producers include Bacillus and Streptomyces species (see Dekker and Richards, 1976) as well as various plant pathogens such as Agrobacterium, Corynebacterium, Erwinia, Flavobacterium, Pseudomonas and Xanthomonas (Hayward, 1977).

Cellulose degradation. Members of certain fungal genera have been identified as major producers of cellulases and have been extensively studied (see Coughlan, 1985b; Eveleigh, 1987). Fungal cellulase producers include members of the genera Trichoderma (e.g. Trichoderma reesei), Aspergillus (e.g. Aspergillus niger), Sporotrichum (e.g. Sporotrichum pulverulentum), and Penicillium (e.g. Penicillium citrinum). T. reesei has been extensively researched and mutants of this organism are the most prolific known producers of cellulase (Coughlan, 1985b; Eveleigh, 1987).

Bacterial cellulase producers appear to be divisible into two groups. Firstly, those that produce the complete complement of enzymes required for crystalline cellulose hydrolysis. These include members from the genera Acetivibrio (e.g. Acetivibrio cellulolyticus), Bacteroides (e.g. Bacteroides succinogenes), Cellulomonas (e.g. Cellulomonas fimi), Cellvibrio, Clostridium (e.g. Clostridium thermocellum), Pseudomonas (e.g. Pseudomonas fluorescens) and Ruminococcus (e.g. Ruminococcus albus). Actinomycetes such as Streptomyces griseus and thermophilic species of Thermomonospora (see Coughlan, 1985b) as well as Microbispora bispora (Waldron et al., 1986) have also been isolated as cellulase producers. Secondly, a number of bacterial species appear to produce only components of the cellulase system and are not able to hydrolyse crystalline cellulose (Gong and Tsao, 1979; Bisaria and Ghose, 1981). Clostridium acetobutylicum and Bacillus subtilis as well as other Bacillus species fall into this group (Fogarty and Griffin, 1973; Horikoshi et al., 1984; Lee et al., 1985a).

Cellulolytic organisms have been studied in much greater detail than organisms producing other lignocellulolytic enzymes. Attention has focused on the search for organisms that produce high levels of extracellular cellulase for industrial use. A list of fungi, bacteria and actinomycetes, which are either used for commercial cellulase production or are of potential commercial use, is given by Coughlan (1985b).

Enzymes involved in the degradation of lignocellulose.

Due to the complex nature of the substrate, large complexes of enzymes, many apparently acting synergistically, are required for hydrolysis (Coughlan, 1985b; Knowles et al., 1987; Eveleigh, 1987). In the industrial use of lignocellulose, the lignin can be removed by chemical (acid or alkali) pretreatment. However, pretreatment can result in the removal of all or part of the hemicellulose content, which together with the lignin, can account for approximately 50% of the lignocellulose (Dale, 1985). Efficient utilisation of lignocellulose would require the utilisation of the entire polysaccharide content of both the hemicellulose and cellulose of the raw material. Enzymes that can hydrolyse these polysaccharides have therefore been studied in detail. However the diversity and complexity of lignin chemistry has proved to be the major stumbling block to developing an understanding of lignin biodegradation. Adequate techniques for the chemical analysis of lignin and its degradation products also appear to be lacking (Amer and Drew, 1980).

Ligninases. Although a number fungi and bacteria are involved in the degradation of lignin, attempts to degrade lignin with spent culture broths has not been successful (Amer and Drew, 1980). Laccase (EC 1.10.3.2) and peroxidase are enzymes involved in lignin degradation (Ishihara, 1980), but neither of these enzymes alone can cause depolymerisation of lignin (see Amer and Drew, 1980). Laccase converts syringic acid and phenols to ortho-quinones

and phenoxy radicals in the presence of oxygen (Ishihara, 1980). The enzyme cellobiose:quinone oxidoreductase (CBQ) forms a link between cellulose and lignin degradation since its activity depends on the concurrent degradation of both these substances. CBQ oxidises cellobiose and cellobiose oligomers with the concomitant reduction of quinones and/or phenoxy radicals to form cellobiono-⁶-lactone (Ishihara, 1980). It has been suggested that the initial breakdown of lignin is not enzymic, but is achieved by the secretion of reduced oxygen species such as superoxide radicals by the organism (see Amer and Drew, 1980). Facts supporting this view include: the diversity of linkages in lignin which makes it unlikely that one or a few enzymes could cleave the polymer; lignin degradation is oxidative, the extent of degradation is increased by increasing the oxygen tension in the fermentation media; the cells must be in close proximity to the lignin to achieve degradation. Subsequent degradation of lignin appears to follow an exo-degradation mechanism i.e. the microorganisms are incapable of splitting the lignin polymer into intermediate polyphenolic moieties but attack it from the periphery (see Amer and Drew, 1980; Crawford and Crawford, 1980). Lignin model compounds such as ¹⁴C-synthetic lignin (dehydropolymers of coniferyl alcohol) have been used for the study of lignin degrading organisms. However organisms that can degrade them often cannot degrade natural lignin (Bisaria and Ghose, 1981).

Hemicellulases. The hemicellulase group of enzymes can be compared to other major hydrolase enzyme complexes such as the cellulases and amylases. Although these enzymes may

show some functional similarity with the other hydrolases, one would expect to find a greater variety of hemicellulases due to the greater complexity of hemicellulose (Reilly, 1981; Dekker, 1985). Hemicellulases (glycan hydrolases, EC 3.2.1) specifically degrade the glycans that make up the backbone chain of hemicelluloses and therefore include β -D-galactanases, β -D-mannanases, and β -D-xylanases (Dekker and Lindner, 1979; Dekker and Richards, 1976). Excluded from this group are the exoglycosidases such as the α and β -D-galactosidases, β -D-mannosidases, and β -D-xylosidases (Dekker, 1985), but their activity is required to achieve total hydrolysis of hemicellulose. The prefix "exo" refers to enzymes that remove residues (as mono- or disaccharides) from the non-reducing end of the polymer, and "endo" refers to enzymes that randomly cleave the polymer internally. In all studies relating to the specificity of a particular glucanase, the mode of action can only be meaningful if a homogeneous protein preparation is used since the enzymes show complex synergistic activity (Reilly, 1981; Dekker, 1985). Standard techniques of ion exchange chromatography, gel filtration, solvent and salt precipitation as well as preparative isoelectric focusing have been used to purify these enzymes. More recently, cloning of several xylanases has enabled the purification of the enzymes free of any other component of the system.

α -D-Galactanases degrade D-galactans and L-arabino-D-galactans. Two types of endogalactanases are recognised depending on the type of bond hydrolysed (either 1,4- β -D-galactosyl or 1,3- β -D-galactosyl bonds) (Dekker, 1985).

Although exogalactanases have not been unequivocally characterised (Dekker, 1985), an enzyme that possibly has both endo and exo activity has been isolated from B. subtilis (Labavitch et al., 1976). The endo-1,4- β -D-galactanases hydrolyse galactans randomly to produce galactose and galactose oligosaccharides. The endo-1,3- β -D-galactanases hydrolyse the 1,3 bonds of the arabinogalactans and liberate D-galactose and 1,3- and 1,6-linked β -D-galactose oligosaccharides (Dekker, 1985).

Both endo and exo type β -D-mannanases have been characterised. For example, an exomannanase was isolated from an Aeromonas species which hydrolysed 1,4- β -D-mannans from a green alga (Codium fragile) to mannobiose and oligosaccharides of DP 3 - 5. Mannobiose, aryl- β -D-mannopyranosides, and glucose and galactose containing mannans were not hydrolysed. The enzyme activity was consistent with exo-type hydrolase activity (mannobiohydrolase) since the rate of activity increased with an increase in DP, only linear mannans were degraded, the enzyme stalled at branch points, and the lowest DP product was mannobiose (Dekker, 1985).

The galactomannans produced commercially from guar (a legume grown for use as forage and seeds that yield guar gum) and carob (locust tree, Ceratonia siliqua) are the main substrates that have been used for studies on the mode of action the endomannanases (EC 3.2.1.78). The endomannanases generally degrade β -D-mannans to D-mannose and mannose oligosaccharides with DP 2 - 6 (Dekker, 1985). The specific

modes of action of endomannanases from bacterial, fungal and plant origin have been described in detail (Dekker and Richards, 1976).

Xylanases. Xylan is the major component of the hemicellulose group and is second in abundance to cellulose in nature. Therefore the xylans and xylanases have been studied in much greater detail than the other hemicelluloses and hemicellulases. There are many types of xylanases produced by a wide variety of fungal and bacterial species. Many xylanases are isozymes of one another showing similar specificities, but differing in amino acid and carbohydrate content, which can result in different enzyme stabilities, pH optima and isoelectric points. It is not known to what extent these enzyme differences are due to multiple genes or to post-transcriptional, post-translational or post-secretional modifications. Many xylanases have been described (Reilly, 1981; Dekker, 1985; Biely, 1985) and the enzymes have been generally grouped according to their specificities (Reilly, 1981):

1. β -xylosidases (EC 3.2.1.37) hydrolyse short xylooligosaccharides to xylose and have substantial transferase activity.
2. exoxylanases (often classed as β -xylosidases) produce xylose from xylan, the rate of production increasing with increasing DP. These enzymes have very little transferase activity.
3. endoxylanases (EC 3.2.1.8). Reilly (1981) divided the endoxylanases into classes based on their ability to cleave L-arabinose from xylan and by the end products

produced (either, mainly xylose and xylobiose or larger oligosaccharides). Endo- type xylanases are the only xylanases that have been unequivocally characterised (Dekker, 1985). Biely (1985) classed endoxylanases generally as enzymes cleaving the xylan backbone, but not adjacent to the units carrying a side chain. De-branching is carried out by separate enzymes which do not directly attack the xylan backbone but do aid in the degradation of xylan by removing side chains. They include α -L-arabinofuranosidase (EC 3.2.1.55), α -glucuronidase (EC 3.2.1.) and acetyesterase (EC 3.1.1.6) (Biely, 1985). Esterases act synergistically with endoxylanases to enhance the degradation of acetyl xylan (Biely, 1985; Biely *et al.*, 1986).

Regulation of xylanase synthesis. In non-cellulolytic bacteria and yeasts, xylanolytic enzymes appear to be inducible and can be catabolite repressed by xylose and glucose. Since xylan is unable to enter the cells, the inducer is likely to be a low molecular weight fragment such as xylobiose or xylotriose which are formed by low level constitutive secretion of the xylanases in a system similar to cellulase induction (Biely, 1985). Induction has also been achieved with synthetic alkyl and aryl β -D-xylosides in Streptomyces species and by methyl β -D-xyloside in yeast (Biely, 1985). Positional isomers of xylobiose such as 1,2- β -xylobiose have also been shown to induce xylanase production in yeast (Biely and Petrakova, 1984) in a manner analogous to 1,2- β -glucobiose (sophorose) induction of cellulases. Although xylose generally represses expression

of xylanase genes, it was found that xylose is a better inducer of xylanase expression than xylobiose in Bacillus pumilus (Panbangred et al., 1983b). The regulation mechanism may therefore be more complex and may not be similar for all systems. Regulation studies of xylanases are often complicated by the concurrent production of cellulases and cross specificity of these enzymes. The specificities for both enzyme types range from absolute for one polymer to equivalent specificity for both polymers. For example T. reesei QM9414 produces specific endoxylanases and endoglucanases (cellulases) with xylanase activity (Biely, 1985). Due to enzyme interactions such as synergism, characterisation of individual xylanases and cellulases is often hampered by the presence of low concentrations of other enzymes, leading to misinterpretation of the results. Recent reviews on xylanases specifically exclude all studies that involve characterisation of xylanases derived from non-homogeneous preparations (Reilly, 1981; Dekker, 1985).

Detection and assay of xylanase activity. Initial studies on the enumeration of hemicellulolytic organisms relied on the use of selective media (Leedle and Hespell, 1980). However, other non-hemicellulolytic organisms could grow by utilising the primary breakdown products. Zone formation on xylan-containing agar plates either by clearing of the opaque xylan-containing medium around the colony, or by the use of dye-polysaccharide interactions, are the commonly used methods of detection. Williams (1983) found that Gram's iodine gave consistently better results than Congo

red or Trypan blue when screening for hemicellulolytic organisms, although Congo red has been used successfully to detect clones expressing xylanase activity (Mondou et al., 1986; this study). Soluble chromogenic substrates such as Remazol brilliant blue-xylan have also been used for the detection of xylanase activity (Sipat et al., 1987). Xylanases have been classified by the products they produce from xylooligosaccharides. Xylanase activity is generally quantified by the release of reducing sugar from xylan or the use of soluble polysaccharides with covalently bound dyes as substrates (Biely et al., 1985b). Specific reaction products generated from xylooligosaccharides are usually examined by high pressure liquid chromatography (HPLC). β -xylosidase activity is assayed by using *p*-nitrophenol-xyloside as a substrate and monitoring the release of *p*-nitrophenol.

Mode of action of xylanases. Biely et al. (1981) proposed a lysozyme-like activity for the xylanase from Cryptococcus albidus based on results from the degradation of xylooligosaccharides. This mode of action has also been proposed for cellulases and is discussed in more detail in the section dealing with cellulases. Morosoli et al. (1986) determined the sequence of the first 72 amino terminal residues of the C. albidus endoxylanase and found weak homology with hen egg-white lysozyme (HEWL) in the regions surrounding the catalytic site. The nucleotide sequence of a xylanase gene from B. subtilis was determined by Paice et al. (1986). The deduced amino acid sequence exhibited over 50% identity with the deduced amino acid sequence of a

B. pumilus xylanase gene (Fukusaki et al., 1984). They also found that the amino acid composition of the cloned gene differed from a xylanase previously characterised from B. subtilis (Bernier et al., 1983) and concluded that the xylanases were encoded by different genes. The authors could find no apparent homology with HEWL or cellulases. Knowles et al. (1987) pointed out that the homologies demonstrated thus far between lysozyme and cellulases are too weak to draw any conclusions about their significance. The same conclusion could apply to the xylanases.

Cloning of xylanase genes. The characterisation of xylanases is complicated by the synergistic activity between these and other glycolytic enzymes. Homogeneity of the protein preparation is a prerequisite to ensure unambiguous results (Dekker, 1985). Cloning the gene into a non-glycolytic host would remove the need for rigorous protein purification and could simplify the characterisation of the enzyme. Paice and Jurasek (1984) investigated the potential of xylanases as industrial catalysts in the specific removal of hemicelluloses from hardwood pulps for possible application in dissolving pulp manufacturing. They concluded that the enzyme preparation must be free of cellulase activity to prevent cellulose degradation. One way to achieve this was to transfer a xylanase gene to a non-glycolytic host. These and other workers have demonstrated the feasibility of this idea by cloning xylanase genes from various organisms and demonstrating xylanase activity in the foreign host (Table 1.2).

Table 1.2. Cloned xylanase genes. Host organism is shown in parenthesis.

Organism and gene	Reference
Alkalophilic <u>Aeromonas</u> species xylanase (E. coli)	Kudo <u>et al.</u> , 1985
Alkalophilic <u>Bacillus</u> species xylanase (E. coli)	Honda <u>et al.</u> , 1985
<u>Bacillus pumilus</u> xylanase (E. coli)	Panbangred <u>et al.</u> , 1983a
<u>Bacillus subtilis</u> xylanase (E. coli) B-xylosidase (E. coli)	Bernier <u>et al.</u> , 1983 Bernier and Desrochers, 1985
<u>Bacteroides succinogenes</u> xylanase (E. coli)	Sipat <u>et al.</u> , 1987
<u>Clostridium acetobutylicum</u> xylanase (E. coli)	Zappe <u>et al.</u> , 1987 (this study)
<u>Streptomyces</u> species xylanase (Streptomyces)	Iwasaki <u>et al.</u> , 1986
<u>Streptomyces lividans</u> xylanase (Streptomyces)	Mondou <u>et al.</u> , 1986

Cellulases. Cellulase is a collective name for a set of enzymes all of which are required to degrade cellulose. These enzymes have generally been divided into three classes: 1. endo- β -1,4-glucanases (endo-1,4- β -D-glucan 4-glucanohydrolase, EC 3.2.1.4), which cleave the glucose polymer 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. cellobiase (β -glucosidase, EC 3.2.1.21), which hydrolyses cellobiose to glucose (see

Coughlan, 1985b). Another class of exoglucanases that removes glucose units from the non-reducing end of the glucose polymer (EC 3.2.1.74) has been isolated from culture filtrates of Penicillium funiculosum (Wood and McCrae, 1982) and may be produced by T. reesei (Marsden et al., 1982). Further enzymic and non-enzymic factors which are directly or indirectly involved in the hydrolysis of cellulose were reviewed by Coughlan (1985b). The existence of non-enzymic microfibril generating factors, equivalent to the C₁ component of the Reese cellulose hydrolysis model (Reese et al., 1950), is still open to debate (see Coughlan, 1985b; Eveleigh, 1987; Enari and Niku-Paavola, 1987).

Multiplicity of cellulase enzyme forms. The occurrence of multiple cellulase enzyme forms of the same type is well documented (e.g. Coughlan, 1985b; Eveleigh, 1987). Multiplicity of enzyme forms appears to be due to several factors, some or all of which can occur in a given system. This makes it difficult to assess the significance of any one factor and complicates the study of cellulases. Many enzyme forms differ only slightly in their substrate specificity resulting in an apparent duplication of function. Enzyme multiplicity or heterogeneity appears to be due to a number of factors including: the effects of limited proteolysis, the formation of multi enzyme complexes, interaction of cellulases with other proteins or glycoproteins and multiple genes. Factors such as the methods of purification and storage of cellulolytic enzymes, which can result in spontaneous modification, may lead to

apparent multiple forms of a single enzyme (Enari and Niku-Paavola, 1987).

Limited proteolysis has been implicated in the modification (Mandels et al., 1961) and release of fungal cellulases. Trypsin was found to release cellulases from the cell walls of T. reesei (Kubicek, 1981) and proteolytic enzymes were shown to be involved in the activation of cellulases from Phanerochaete chrysosporium (then called Sporotrichum pulverulentum) (Eriksson and Pettersson, 1982). Nakayama et al. (1976) found that when an endoglucanase, isolated from a commercial preparation of T. reesei cellulase, was treated with a protease, it yielded a number of different forms which differed slightly in their substrate specificity. Paice et al. (1984) purified two endo- β -1,4-glucanases from Schizophyllum commune (EG 1 and EG 2) and found that they showed similar properties. The amino terminal amino acid sequences were identical except that EG 1 had an initial 16 residue alanine rich sequence. The authors suggested that extracellular proteolytic cleavage of EG 1 could be a reason for endoglucanase heterogeneity in this organism. Labudova and Farkas (1983) monitored proteolytic and endoglucanase enzyme activity in samples of T. reesei QM9414 cultures. As many as six different forms of endoglucanases were separated by isoelectric focusing (IEF). The authors concluded that the multiple forms of endoglucanases found could not result solely from post-translational modifications, such as proteolysis and glycosylation, since the enzyme forms were present in the culture from a very early stage and only low protease

activity was detected. Furthermore, glycosylated Trichoderma cellulases contained mainly neutral sugars in their carbohydrate moieties (Shoemaker and Brown, 1978) which would not appreciably change the proteins' isoelectric points. However, as pointed out by Coughlan (1985b), the authors did not consider the possibility of acid proteases since in uncontrolled fermentations the pH of the culture could be as low as pH 2. Dunne (1982) and Kammel and Kubicek (1985) could however not demonstrate proteolytic modifications of T. reesei cellulases indicating that the different enzyme forms were generated before release into the medium, providing further evidence that post-translational and/or postsecretional modifications are not the only processes involved in multiple form generation.

Cellulases have been shown to form multienzyme complexes which can appear homogeneous by one purification method and show heterogeneity by another. Furthermore the complexes can consist of a variable number of different enzymes which share a common isoelectric point (pI) after IEF purification (apparent pI , pI_{app}), rather than the true pI of a single enzyme (Sprey and Lambert, 1983). The characteristics of the complex would be different to those of the individual components. This form of enzyme multiplicity has been called macroheterogeneity (Sprey and Lambert, 1984; Eveleigh, 1987). Infidelity in transcription and variable glycosylation have been implicated in the occurrence of cellulase multiplicity in S. commune (Willick and Seligy, 1985) and T. reesei (Alurralde and Ellenreider, 1984). A purified T. reesei cellobiohydrolase was found to contain a

non-covalently bound carbohydrate moiety, removal of which resulted in an increase of enzyme activity (Alurralde and Ellenreider, 1984).

Growth conditions and the age of the culture have also been shown to affect the number of cellulases present. T. reesei produced only certain cellulases depending on the growth conditions and enzyme multiplicity increased with the age of the culture. There was also a concomitant rise in the protease activity (Gritzali and Brown, 1979; Gong et al., 1979). These results are in contrast to those reported by Labudova and Farkas (1983). Their results indicated that multiple forms of the individual enzymes appeared very early in the T. reesei culture medium and they ascribed this discrepancy to the better resolving power of IEF over the ion exchange chromatography method used by Gritzali and Brown (1979) and Gong et al. (1979).

Wood (1981) postulated that, for stereochemical reasons, at least two types of both endo- and exoglucanases would be required for hydrolysis of cellulose. This view was supported by evidence that two cellobiohydrolases from T. reesei (CBH I and CBH II) were found to be immunologically unrelated (Gritzali and Brown, 1979; Fagerstam and Pettersson, 1980) and the first 20 amino acids of their amino terminal ends showed no homology (Pettersson et al., 1981) suggesting that the proteins were encoded by different genes.

The occurrence of multiple cellulase genes appears to be an important source for enzyme multiplicity. The amino acid sequence of CBH I was determined (Fagerstam et al., 1984) and both of these genes (CBH I and CBH II) have been cloned and the DNA sequenced (Teeri et al., 1983; Shoemaker et al., 1983; Teeri et al., 1987; Chen et al., 1987). The gene products showed very little overall amino acid homology but do however show some interesting common features (Teeri et al., 1987; Knowles et al., 1987).

Among the organisms from which cellulase genes have been cloned (Table 1.3), and the nucleotide sequences determined (Table 1.4), C. thermocellum appears to have a particularly large number of unrelated cellulase genes. Millet et al. (1985) reported the cloning of 10 distinct DNA fragments from C. thermocellum coding for cellulases. Two of the fragments coded for the genes previously isolated (celA and celB) (Cornet et al., 1983), another five coded for previously unknown endoglucanases and three may code for cellobiohydrolase genes. The DNA fragments differed in their restriction profiles. As stated by the authors, the genebanks constructed were unlikely to represent more than one third of the C. thermocellum genome, and yet an unexpectedly large number of cellulase genes were isolated suggesting that the total number of cellulase genes in C. thermocellum may be high. It had previously been suggested that the number of endoglucanase genes alone may be as high as 21 (see Millet et al., 1985). Schwarz et al. (1985) cloned genes from another strain of C. thermocellum and isolated endoglucanase genes that differed in their

EcoRI restriction patterns to those reported by Millet et al. (1985). Romaniec et al. (1987) cloned a further 13 unique DNA fragments from C. thermocellum which were able to direct the synthesis of endoglucanases in E. coli. Gilkes et al. (1984a, 1984b) reported the cloning of two endoglucanases and one exoglucanase from C. fimi and showed that the genes shared no overall homology.

Table 1.3. Cloned cellulase genes.

Organism and gene(s)	Reference
Bacteria	
<u>Agrobacterium</u>	
1 β -glucosidase	Wakarchuk <u>et al.</u> , 1986
<u>Bacillus</u> species (alkalophilic)	
1 β -1,4-glucanase	Fukumori <u>et al.</u> , 1986a
2 β -1,4-glucanases	Sashihara <u>et al.</u> , 1984
<u>Bacillus amyloliquefaciens</u> ^a	
1 endo-1,3-1,4- β -gluc. ^a	Borriss <u>et al.</u> , 1985
<u>Bacillus subtilis</u>	
1 endo-1,3-1,4- β -gluc. ^a	Cantwell and McConnell, 1983
1 endo-1,3-1,4- β -gluc. ^a	Hinchliffe, 1984
1 endoglucanase	Koide <u>et al.</u> , 1986
1 endoglucanase	Park and Pack, 1986
1 β -1,4-glucanase	Robson and Chambliss, 1986
<u>Bacteroides succinogenes</u>	
1 endoglucanase	Crosby <u>et al.</u> , 1984
<u>Caldocellum saccharolyticum</u>	
1 β -glucosidase	Love and Streiff, 1987
<u>Cellulomonas fimi</u>	
1 endoglucanase	Whittle <u>et al.</u> , 1982
1 endoglucanase	Gilkes <u>et al.</u> , 1984a
1 cellobiohydrolase	Gilkes <u>et al.</u> , 1984a
<u>Clostridium acetobutylicum</u>	
1 endoglucanase	Zappe <u>et al.</u> , 1986

a gluc = glucanase

b var. cellulosa

c results not published but reported as personal communication in Knowles et al. (1987).

Table 1.3 continued.

Organism and gene(s)	Reference
<u>Clostridium thermocellum</u>	
2 endoglucanases	Cornet <u>et al.</u> , 1983b
1 endoglucanase	Schwarz <u>et al.</u> , 1986
7 endoglucanases	Millet <u>et al.</u> , 1985
3 exoglucanases	Millet <u>et al.</u> , 1985
2 B-glucosidases	Schwarz <u>et al.</u> , 1985
13 endoglucanases	Romaniec <u>et al.</u> , 1987
<u>Escherichia adecarboxylata</u>	
1 B-glucosidase	Armentrout and Brown, 1981
<u>Erwinia chrysanthemi</u>	
>1 endoglucanases	van Gijsegem <u>et al.</u> , 1985
1 endoglucanase	Barras <u>et al.</u> , 1984
1 endoglucanase	Kotoujansky <u>et al.</u> , 1985
1 endoglucanase	Boyer <u>et al.</u> , 1987
<u>Microbispora bispora</u>	
5 endoglucanase	see Knowles <u>et al.</u> , 1987 ^C
2 B-glucosidases	see Knowles <u>et al.</u> , 1987 ^C
<u>Pseudomonas fluorescens</u> ^b	
1 endoglucanase	Wolff <u>et al.</u> , 1986
1 endoglucanase	Lejeune <u>et al.</u> , 1986
<u>Thermomonospora YX</u>	
1 endoglucanase	Collmer and Wilson, 1983
Fungi	
<u>Aspergillus niger</u>	
1 B-glucosidase	Penttila <u>et al.</u> , 1984
<u>Candida pelliculosa</u>	
1 B-glucosidase	Kohchi and Toh-e, 1985
<u>Kluyveromyces fragilis</u>	
1 B-glucosidase	Raynal and Guerineau, 1984
<u>Phanerochaete chrysosporium</u>	
1 cellobiohydrolase	see Knowles <u>et al.</u> , 1987 ^C
<u>Schizophyllum commune</u>	
1 endoglucanase	see Knowles, <u>et al.</u> , 1987 ^C
<u>Trichoderma reesei</u>	
1 cellobiohydrolase	Shoemaker <u>et al.</u> , 1983
	Teeri <u>et al.</u> , 1983
1 cellobiohydrolase	Chen <u>et al.</u> , 1987
	Teeri <u>et al.</u> , 1987
1 endoglucanase	Penttila <u>et al.</u> , 1986
	van Arsdell <u>et al.</u> , 1987
1 endoglucanase	see Knowles <u>et al.</u> , 1987 ^C

Table 1.4. Cellulase genes for which nucleotide sequences have been determined. In one case (a), the protein sequence was determined.

Organism (plasmid/gene)	Reference
Bacteria	
<u>Bacillus</u> species (alkalophilic)	
pFK1	Fukumori <u>et al.</u> , 1986a
pNK1, pNK2	Fukumori <u>et al.</u> , 1986b
<u>Bacillus amyloliquefaciens</u>	
pEG1	Hofemeister <u>et al.</u> , 1986
<u>Bacillus subtilis</u>	
pJG83	Murphy <u>et al.</u> , 1984
pC6.5, pC6.3	MacKay <u>et al.</u> , 1986
pLG4002	Robson and Chambliss, 1987
<u>Cellulomonas fimi</u>	
<u>cenA</u>	Wong <u>et al.</u> , 1986
<u>cex</u>	O'Neill <u>et al.</u> , 1986
<u>Clostridium acetobutylicum</u>	
pHZ117	Zappe <u>et al.</u> , 1988
<u>Clostridium thermocellum</u>	
<u>celA</u>	Beguin <u>et al.</u> , 1985
<u>celB</u>	Grepinet and Beguin, 1986
<u>celD</u>	Joliff <u>et al.</u> , 1986
<u>Erwinia chrysanthemi</u>	
<u>celZ</u>	Guiseppi <u>et al.</u> , 1988
Fungi	
<u>Candida pelliculosa</u>	
-	Kohchi and Toh-e, 1985
<u>Trichoderma reesei</u>	
CBH I	Fagerstam <u>et al.</u> , 1984 ^a
CBH I	Shoemaker <u>et al.</u> , 1983
CBH II	Chen <u>et al.</u> , 1987
	Teeri <u>et al.</u> , 1987
EG I	Penttila <u>et al.</u> , 1986
	van Arsdell <u>et al.</u> , 1987

Cellulase gene architecture. Nucleotide sequence analysis of cloned cellulase genes has revealed significant differences in the primary structure of the enzymes, even when comparing genes of similar activity derived from the same organism. Little overall amino acid homology was demonstrated between three endoglucanases from C. thermocellum (EG A, EG B, and EG C) (Joliff et al., 1986) or between three of the four endoglucanases from T. reesei. Knowles et al. (1987) compared the amino acid sequences of the cellulases. In all four T. reesei glucanases (CBH I, CBH II, EG I, EG III) there was a region of about 30 amino acids which exhibited about 70% homology. These regions were defined as tail regions of the proteins (Knowles et al., 1987) and were located at the carboxy terminal (C-terminal) (EG I, CBH I) or amino terminal (N-terminal) (EG III, CBH II) end of the protein (Fig. 1.3). A similar arrangement was identified from the deduced protein sequence of two glucanases from C. fimi (EXG and ENG). A region showing approximately 50% homology was found at the N-terminal end of ENG and at the C-terminal end of EXG (Fig. 1.3). No homology was demonstrated between the tail regions T. reesei and C. fimi endoglucanases (Knowles et al., 1987). However, in both cases the tail regions were linked to the rest of the protein (the hydrolytic domain) by a region, rich in hydroxyl amino acids and proline, which has been termed a hinge (Knowles et al., 1987). The three sequenced endoglucanases from C. thermocellum also revealed an interesting structure. The proteins all had a direct repeat of 24 amino acids at their C-terminal end (Joliff et al., 1986). The hinge region was present in EG B, less

well defined in EG A, and could not be demonstrated in EG D (Fig. 1.3) (Knowles et al., 1987). The hinge-tail arrangement was also demonstrated for two glucosidases from A. niger (G1, G2), although only G1 had a tail region (Knowles et al., 1987).

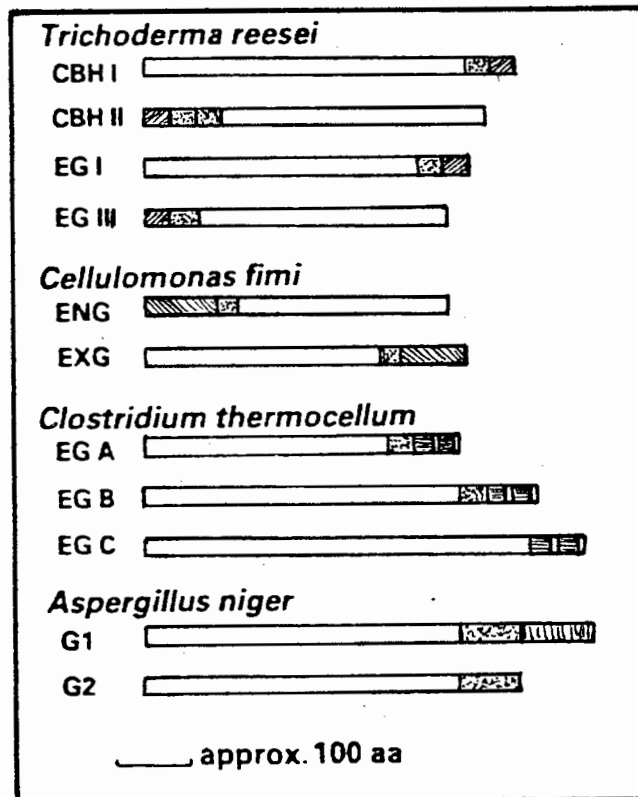


Fig. 1.3. Schematic structure of different cellulase and B-glucosidase genes. Each enzyme consists of a catalytic domain probably linked via a flexible "hinge" region to a "tail" domain. The drawing is only roughly to scale. [White box], hydrolytic domain; [Stippled box], "hinge"; [Diagonal lines box] "tail" (T. reesei); [Cross-hatched box], "tail" (C. fimi); [Horizontal lines box], "tail" (C. thermocellum); [Vertical lines box], "tail" (A. niger) (from Knowles et al., 1987).

Of the three structural enzyme regions described by Knowles et al. (1987), the tail region does not appear to be as clearly defined in endoglucanases isolated from organisms, such as B. subtilis, that are not truly cellulolytic (i.e. not able to degrade crystalline cellulose). MacKay et al. (1986) found that one third of the C-terminal end of a

B. subtilis endoglucanase could be deleted without loss of function. This is in agreement with the idea that the tail region is involved in binding to insoluble substrates. Joliff et al. (1986) suggested that the homologous terminal regions of the three sequenced C. thermocellum endoglucanase genes could be involved in binding the enzyme to cellulose, a process apparently not required for hydrolysis of soluble substrates since the cloned bacterial endoglucanases do not lose their activity towards these substrates when deleted in these regions. Another possibility is that the tail region of these genes could act as an attachment region involved in the formation of the cellulosome complex.

Removal of the homologous tail region from the endo- and exoglucanase of C. fimi did not affect their activity towards soluble substrates (Wong et al., 1986). The tail region is located at the C-terminal end of the exoglucanase and at the N-terminal end of the endoglucanase from C. fimi (Knowles et al., 1987) (Fig. 1.3). Removal of 12 amino acids from the hydrolytic domain (C-terminal end) of the C. fimi endoglucanase did however abolish activity (Wong et al., 1986). The T. reesei homologous domains can also be removed by limited proteolysis, resulting in a reduction of activity against crystalline cellulose, but not against soluble cellulose (Knowles et al., 1987).

A hypothetical model of a cellulolytic enzyme was presented by Knowles et al. (1987) based on the amino acid sequence data and low angle X-ray diffraction studies (Schmuck et al., 1986) (Fig. 1.4). Knowles et al. (1987) predicted

that all enzymes that hydrolyse solid carbohydrates, such as starch, hemicellulose and cellulose, will be shown to have the same domain structure. As described above, the enzyme consists of hydrolytic, hinge and tail domains. The hinge and tail domains can be located at the N-terminal or C-terminal end of the polypeptide (Fig. 1.4). Post-secretional proteolytic cleavage of cellulases has been proposed as one of the reasons for the multiplicity of cellulases within a given system. Cleavage of the enzyme in the glycosylated hinge region could alter its substrate specificity thus generating derivatives able to hydrolyse the more soluble oligosaccharides released as hydrolysis proceeds (Knowles et al., 1987). These substrate specificity changes (i.e. enzyme modifications) could occur at various stages in the growth cycle.

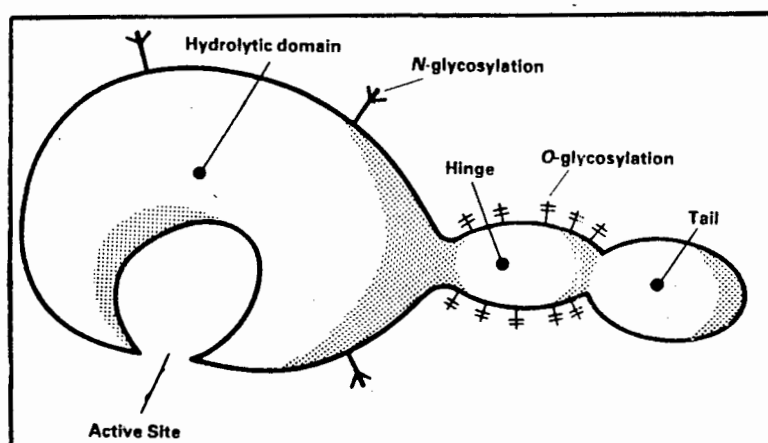


Fig. 1.4. A hypothetical model of a cellulolytic enzyme based on amino acid sequence data and X-ray diffraction studies (from Knowles et al., 1987).

The role of glycosylation in cellulase function. Cellulases are one of the few examples of glycosylated proteins in bacteria while most glycosylation in fungal (e.g. T. reesei) cellobiohydrolases occurs in the conserved terminal regions

as shown by Knowles et al. (1987). These observations would indicate that glycosylation has some importance in cellulase activity. However no functional role or significance has yet been attributed to glycosylation of cellulases. Endoglucanases of both fungal and bacterial origin which have been cloned in E. coli are expressed in an unglycosylated form. Alternatively when T. reesei endoglucanases were expressed in yeast, they were found to be overglycosylated (van Arsdell et al., 1987; see Knowles et al., 1987). In each case the activity towards soluble substrates was not affected. However it is possible that glycosylation affects substrate binding since purified overglycosylated T. reesei CBH II showed a decrease in activity towards β -glucan and the enzyme was slightly impaired in its affinity towards crystalline cellulose (see Knowles et al., 1987).

Glycosylation does not appear to be involved in secretion of cellulases since treatment of S. commune or T. reesei with tunicamycin (an inhibitor of asparagine-linked N-glycosylation) and 2-deoxyglucose (impairs N- and O-glycosylation) had little effect on secretion of cellulases from these organisms (Willick and Seligy, 1985; Merivouri et al., 1985; Murphy-Holland and Eveleigh, 1985). Glycosylation may be implicated in enzyme stability but conflicting results have been reported. Murphy-Holland and Eveleigh (1985) reported that partial deglycosylation had no major effects on the thermal or protease stability of T. reesei cellulases. However Merivouri et al. (1985) postulated for T. reesei cellulases that N-glycosylation was

important for thermal and protease stability while O-glycosylation was required for secretion activity.

Non-covalently bound carbohydrate may also have a role in cellulase activity and/or activation. Alurralde and Ellenreider (1984) purified a cellobiohydrolase from T. reesei and found that it was complexed with non-covalently bound carbohydrate which could be removed by repeated gel filtration with a concomitant rise in enzyme activity.

Enzymic mechanism of cellulose saccharification. Reese et al. (1950) originally proposed the $C_1:C_X$ model of cellulolysis whereby a factor (C_1) converted crystalline cellulose to a form more accessible to endoglucanases (C_X). A mechanism of enzymic cellulose saccharification has since been proposed based on the findings of many research groups (see Coughlan, 1985b). The model is based on studies of the T. reesei cellulase system. The first step has been termed amorphogenesis (equivalent to the C_1 component of the Reese model), a process whereby the crystalline cellulose is made more accessible to the hydrolytic enzymes. How this process occurs and whether only hydrolytic enzymes are involved is unclear and conflicting results have been reported (Coughlan, 1985b; Enari and Niku-Paavola, 1987; Eveleigh, 1987). Koenigs (1975) found that H_2O_2 and Fe^{2+} caused extensive depolymerisation of cellulose resulting in a product that was more accessible to cellulases. Griffin et al. (1984) isolated a factor from T. reesei cellulase preparations that generated short fibres from filter paper

but did not show reducing sugar activity. The factor was found to contain iron and produced the fibres much more rapidly than H_2O_2 and Fe^{2+} . However Chanzy et al. (1983) and others (Sasaki et al., 1979; Fagerstam and Pettersson, 1980) have shown that T. reesei cellobiohydrolases alone, can degrade highly crystalline cellulose.

The general model for the mode of action of cellulases is based on the synergistic action of endoglucanases, exoglucanases and glucosidases, resulting in the production of soluble oligomers, cellobiose and glucose. The endoglucanases break the β -1,4-glycosidic bonds randomly in the amorphous regions generating sites for the exocellulases, which remove cellobiose units from the non-reducing ends of the chains. Cellobiose is an inhibitor of cellulase activity (Shewale, 1982) and this disaccharide is removed by the action of β -glucosidase.

This model does not however define the manner of synergism of enzyme activity. Not only has synergism been demonstrated between endo- and exoglucanases (Wood and McCrae, 1972; Ladisch et al., 1983; Henrissat et al., 1985), but also between the two exoglucanases of T. reesei (CBH I and CBH II) (Fagerstam and Pettersson, 1980). The exo-exo synergism could be due to the stereospecific classes of enzymes predicted by Wood (1981). Some modes of interaction of cellulases have been elucidated using electron microscopy. Chanzy and Henrissat (1985) examined the effects of T. reesei EG II and CBH II on V. macrophysa cellulose using this technique. Addition of CBH II to the

crystalline cellulose resulted in degradation from one end with the formation of pointed tips. This demonstrated visually that the glycan chains are stripped off sequentially and that all the strands lie in the same direction. These results appear to discount the possibility of folding (Manley, 1964; see Bisaria and Ghose, 1981) or antiparallel packing (French, 1978) of the glycan chains within the cellulose microfibril. Addition of EG II resulted in degradation of amorphous regions and further unidirectional stripping from these new sites. Previously White and Brown (1981) could find no unidirectional degradation of crystalline cellulose using CBH I of T. reesei, presumed by Chanzy and Henrissat (1985) to be due to the more endoglucanase-like character of CBH I (Chanzy and Henrissat, 1983; van Tilbeurgh et al., 1982). White and Brown (1981) did however demonstrate synergistic activity between CBH I and EG IV on A. xylinum crystalline cellulose ribbons, but did not show unidirectional angular thinning of the cellulose microfibrils.

If synergism is based on the above model, then optimal synergism should occur at a high ratio of CBH:EG (Eveleigh, 1987), which was observed for CBH II combined with either EG I or EG II (95:5) (Chanzy and Henrissat, 1985). However, optimal synergism of CBH I with EG I was obtained at 1:1 ratios, depending on the substrate (Chanzy and Henrissat, 1985), indicating that synergism is more complex than the model proposes. Wood (1985) suggested that EG and CBH interact on the substrate surface, EG cleaving initially and then CBH immediately acting on the exposed reducing end to

prevent reclosure of the initial incision. Synergism is most marked when using highly crystalline substrates and does not occur with soluble substrates (Wood and McCrae, 1979).

There is mounting evidence that cellobiohydrolases may be the key components in fungal hydrolysis of crystalline cellulose. Similar T. reesei cellobiohydrolases isolated by different research groups were found to hydrolyse only crystalline cellulose and not soluble substituted forms (Halliwell and Griffin, 1973; Berghem et al., 1975; Nummi et al., 1983; Chanzy et al., 1983) whereas endoglucanases from T. reesei were shown to hydrolyse only soluble forms of cellulose (Niku-Paavola et al., 1985). Antigenically unrelated T. reesei cellobiohydrolases have been shown to act synergistically in the hydrolysis of crystalline cellulose (Fagerstam and Pettersson, 1980). This and other evidence prompted Enari and Niku-Paavola (1987) to propose a new model for the hydrolysis of crystalline cellulose by fungal cellulases. They propose that cellobiohydrolases CBH I and CBH II attack the cellulose to produce cello-dextrins and cellobiose. The cello-dextrins are reduced to cellobiose by endoglucanases I and II and β -glucosidase then reduces cellobiose to glucose.

Catalytic mechanism of cellulases. The proposed mechanism of cellulase catalysis is based on that of HEWL which also cleaves glycosidic bonds. The mode of action of HEWL involves acid catalysis by donation of a proton from the carboxyl group of a conserved Glu-35 residue and carbonium

ion stabilisation by a conserved Asp-52 residue (Blake et al., 1967; Vernon, 1967). It was shown that bacteriophage T4 and goose lysozyme have the same active site features as HEWL (Matthews et al., 1981; Grutter et al., 1983). Vernon and Banks (1963) originally suggested that all glucosidases function in a similar manner and a model was proposed by Yaguchi et al. (1983). The following evidence has been reported in support of a lysozyme-like activity for cellulases: 1. Limited amino acid sequence identity between various lysozymes and fungal cellulases in the active site region (apparent conserved Glu, Asn and Asp residues) (Yaguchi et al., 1983; Paice et al., 1984; Teeri et al., 1987); 2. The importance of carboxyl groups in cellulase activity (Legler and Bause, 1973; Clarke and Yaguchi, 1985). Although the homology is weak, the importance of the specific amino acids can now be tested by site-directed mutagenesis (Teeri et al., 1987).

Adsorption/ de-adsorption. Unlike most substrate-enzyme interactions, the cellulolytic system requires the interaction of soluble enzymes with an insoluble substrate. It has been shown that cellulases adsorb to the substrate and the factors affecting adsorption have been studied in detail (for reviews see Gilbert and Tsao, 1983; Coughlan, 1985b). The importance of an adsorption/ de-adsorption process must be seen in context with industrial enzymic saccharification of cellulosic materials. It has been estimated that the cost of cellulase can account for 60% of the total cost of saccharification (Wilke et al., 1976). Recovery of active enzyme after saccharification is

therefore of prime importance. Adsorption is affected by the type of substrate, its purity, pretreatment, and the enzyme substrate ratio (Roltz, 1986b). For example lignin can bind cellulases in an unproductive manner (Deshpande and Eriksson, 1984). The substrate also changes during hydrolysis (becomes more crystalline) and the products of hydrolysis are released, both of which can affect adsorption (see Coughlan, 1985b). More effective hydrolysis of pretreated lignocellulose has been obtained by the use of an ultrasound field. The explanation given was that ultrasound affected the adsorption/desorption mechanism by causing microturbulence which increased the enzymic rate by stimulating surface-enzyme migration and thus preventing non-productive enzyme-substrate binding (Roltz, 1986a).

Cell structure associated multiple enzyme complexes. For an organism to effectively degrade insoluble cellulose and be able to harvest the resultant products to the exclusion of other organisms, it must be in close contact with the substrate. Fungi achieve this by enveloping the substrate with a mycelium and secreting cellulases. In white rot fungi cytoplasmic organelles associated with a sheath around the fungal hyphae were probably involved in cellulose degradation (see Roltz, 1986b). T. reesei cellulases appear to be associated with pleiomorphic subcellular vesicles (Glenn et al., 1985). In plant pathogenic fungi, production of cellulases facilitates the penetration of the mycelium into the host cells. High level production of extracellular cellulases is characteristic of fungi.

Many bacteria produce a polysaccharide containing glycocalyx which is associated with the retention of extracellular enzymes and as a means of binding the cell to an insoluble substrate like cellulose (Costerton et al., 1981). Studies on the bacterial degradation of crystalline cellulose have led to the suggestion that the cells must bind to the substrate to effect degradation (Akin, 1980), although Kauri and Kushner (1985) demonstrated for cellulolytic soil bacteria that cell-substrate contact was not essential for degradation. However, it would appear that the rumen bacteria B. succinogenes, R. flavefaciens and R. albus must bind to the cellulose to effect degradation (Kudo et al., 1987). It has been shown that the bacterial surface is important for degradation to occur and anything that affects it is likely to affect cellulolytic activity. Stack and Hungate (1984) showed that 3-phenylpropanoic acid did not affect the growth of R. albus, but increased the production of the extracellular capsular structures and resulted in increased cellulolytic activity (Stack and Cotta, 1986). B. succinogenes cellulase enzymes have also been shown to be associated with membrane fractions (Groleau and Forsberg, 1983).

The anaerobic cellulolytic bacterium C. thermocellum possess an extracellular structure, the cellulosome, which is involved in cellulose degradation. It is a multisubunit structure with \underline{M}_r 2.1×10^6 and consists of at least 14 polypeptides (ranging from \underline{M}_r $48 - 210 \times 10^3$), of which eight showed endoglucanase activity. Only one subunit (S1, \underline{M}_r 210×10^3) was antigenically active (Lamed et al., 1983),

but showed no detectable cellulase activity. The cellulosomes appeared in an extracellular and cell-associated form (Lamed et al., 1983). They bind to cellulose and their cellulolytic characteristics are similar to those described for crude cellulase preparations characterised from this organism (Lamed et al., 1985). When wild type C. thermocellum was grown with cellobiose as the sole source of carbon, the cellulosomes were not distributed evenly on the cell surface, but were located on cell surface protuberances. The protuberances were absent in a mutant lacking the ability to bind to cellulose (Bayer and Lamed, 1986). In cellulose-grown wild type cells the protuberances were transformed into an amorphous or fibrous network which bound the cell to cellulosome clusters which in turn coat the cellulose surface (Bayer and Lamed, 1986). The amorphous network may be cell derived and serves as a channel to transport degradation products from the cellulosomes to the cell surface (Lamed and Bayer, 1986).

A correlation between the formation of cell surface structures and the production of cellulases has been demonstrated for a wide range of cellulolytic bacteria and all cellulolytic bacteria tested showed immunochemical cross-reactivity with anticellulosome antibody (Lamed et al., 1987). The occurrence of multienzyme complexes appears to be a common feature in cellulolytic bacteria and serves to enhance contact between the bacterium and the substrate and at the same time provide entrapment and probably stability for the cellulolytic enzymes (Roltz, 1986b).

Regulation of cellulase activity: Induction, repression and inhibition of cellulases. The synthesis of cellulases is generally induced in the presence of cellulose and repressed in the presence of glucose. This conclusion holds for fungi, bacteria and actinomycetes (Coughlan, 1985b). The mechanism of induction has not been established but the current view is that low level constitutive production of cellulase (basal synthesis) initially hydrolyses cellulose in the medium. Gong et al. (1979) showed that T. reesei constitutively produces low levels of cellulase, that the basal level is independent of carbon source and is active under starvation conditions. The products are taken up and converted to an actual inducer. Sophorose has been reported to induce production of only certain endo- and exoglucanases in T. reesei (Sternberg and Mandels, 1979) and is therefore not considered to be "the" inducer of cellulases. Cellobiose at high concentrations and lactose have also been shown to induce cellulases (see Coughlan, 1985b).

Synthesis appears to be regulated by an induction-repression system and the mechanism of regulation of synthesis is similar to other known inducible enzyme systems. In certain prokaryotes such as E. coli certain conditions must be met before a catabolite repressed gene can be expressed. Glucose concentrations must be low and therefore cyclic AMP (cAMP) will be high. The complex of cAMP and catabolite activator protein (CAP) bind to the DNA and increases the rate of transcription from adjacent promoters. An inducer must also be present to bind the repressor protein and thus free the operator. The mechanism of transcription of fungal

genes is not fully understood and may be more complicated than in bacterial cells (Coughlan, 1985b). There is evidence that cAMP is not involved in cellulase biosynthesis in P. fluorescens var. cellulosa or T. reesei (see Coughlan, 1985b). However Fennington et al. (1984) claim that cAMP is involved in the regulation of cellulase biosynthesis in Thermomonospora.

Detection of cellulase activity. Rapid and simple methods for the detection of cellulase activity are obviously indispensable in screening for more efficient cellulolytic organisms, improvement of existing strains by mutation and in the isolation of cloned cellulase genes.

Early screening techniques involved incorporation of cellulose in agar plates and the detection of cellulase producing organisms by the production of zones of hydrolysis or small craters around the colonies (Tansey, 1971; Montenecourt and Eveleigh, 1977). A correlation between zone size and ability to hydrolyse cellulosic substrates was observed using these techniques.

Soluble forms of cellulose such as hydroxyethyl- and carboxymethylcellulose (HEC and CMC respectively) are currently extensively used as model substrates for the detection of endoglucanase activity. Each AHG unit of the polymer contains three hydroxyl groups (positions 2,3 and 6) where substitution can occur. If all hydroxyl groups were substituted, then the preparation would have a degree of substitution (DS) of 3.00 (Wirick, 1968a).

Endoglucanases apparently cleave the polymer only at positions where three or more adjacent unsubstituted AHG units occur. The DS therefore greatly affects the activity of the enzyme towards the substrate and a preparation with a DS = 2.45 is essentially resistant to enzymic cleavage (Wirick, 1968b). Commercially available CMC preparations used for the detection of endoglucanase activity have a DS of approximately 0.7.

Dye binding techniques in conjunction with soluble cellulose derivatives have been developed for detection of cellulase activity. Congo red in particular has been widely used in the detection of cellulosic and hemicellulosic bacteria (Teather and Wood, 1982; Williams, 1983) as well as in the isolation of cloned cellulase genes. Congo red interacts with unsubstituted β -D-glycans with a DP > 5 (Wood, 1980). CMC with a DS of 0.79 contains only 0.036% of unsubstituted areas of this length (Wirick, 1968b). The result is an extremely sensitive detection system which is specific for endoglucanases (Gilkes et al., 1984a). Biely et al. (1985a) described the use of Remazol brilliant blue (RBB) dyed xylan and Ostazin brilliant red dyed HEC as substrates for the detection and assay of xylanase and endoglucanase respectively. Chromophoric compounds such as 4-methylumbelliferyl (MeUmb) glycosides have also proved useful in the detection and analysis of cellulase activity (van Tilbeurgh et al., 1982; van Tilbeurgh et al., 1985). MeUmb- β -D-glucose can be used to detect β -glucosidase activity (Lejeune et al., 1986).

A number of detection methods have been developed specifically for the isolation of cloned cellulase genes. The dye-binding techniques are selective for certain cellulase activities (e.g. endoglucanases using Congo red). More complex methods are required for the detection of other cellulase activities. An effective method was used for the cloning of a T. reesei cellobiohydrolase (CBH I) gene (Shoemaker et al., 1983; Teeri et al., 1983). Two types of cDNA probe were prepared from polyadenylated mRNA isolated from T. reesei mycelia: one from cultures induced and the other from cultures uninduced for cellulase production. A genomic library of T. reesei DNA was screened using the induced probe and counterselected with the uninduced probe. Those clones hybridising strongly to the induced probe were used to select for corresponding fungal mRNAs by hybrid formation. In vitro translation of these mRNAs and subsequent immunoprecipitation of the proteins with cellulase specific antibodies enabled the identification of the cloned CBH I gene. Theoretically all genes efficiently induced during growth on a particular substrate, could be isolated using this differential hybridisation technique (Knowles et al., 1987). Alternatively, a method of direct immuno-detection of clones expressing antigenically active polypeptides can be employed (Whittle et al., 1982; Kemp and Cowman, 1981).

Assay of cellulase activity. Substrates and methods used in the measurement of cellulase activity are summarised in Table 1.5.

Table 1.5. Substrates, enzymes, products formed and assay techniques for cellulolytic enzymes. EG, endoglucanase; CBH, cellobiohydrolase, β -gluc, β -glucosidase; cel, complete cellulase preparations.

Substrate	Enzyme	Result	Assay/measure
Soluble substrates			
CMC, TNP-CMC, HEC, RBB-HEC, B-glucan	EG	reducing sugar decrease in DP	biochemical viscosity staining
Cellooligosaccharides	EG	reducing sugar oligosaccharides	biochemical HPLC
	CBH	cellobiose	
4-MeUmb- ¹ and p-nitrophenyl ² -oligosaccharides	EG	chromophore	fluorescence ¹ absorbance ²
	CBH	oligosaccharides chromophore	HPLC fluorescence ¹ absorbance ²
4-MeUmb-cellobioside	β -gluc	chromophore	fluorescence
cellobiose	β -gluc	dextrose	glucose oxidase
p-Nitrophenyl- β -D-glucoside	β -gluc	p-nitrophenol	absorbance
Insoluble substrates			
<u>Valonia macrophysa</u> and <u>Acetobacter xylinum</u> cellulose	EG, CBH	reducing sugar visual degradation	cyrstalline biochemical electron microscopy
cotton ³ , Avicel ⁴ , filter paper ⁵ , amorphous cellulose ⁶ , H ₃ PO ₄ -swollen cellulose ⁷ , dyed Avicel ⁸	cel	reducing sugar decrease in DP release of dye	biochemical ³⁻⁷ turbidity ³⁻⁶ staining ⁴ absorbance ⁸

The measurement of cellulase activity is complicated by the heterogeneity of both the substrate and the enzyme. Substrates may be soluble or insoluble and can be amorphous,

crystalline or a mixture of both. In some cases the substrate may also contain relatively high proportions of hemicellulose and lignin. Furthermore, the substrate is altered during the reaction with the amorphous regions being preferentially degraded resulting in the substrate becoming more crystalline. The variety of substrates that have been used and the methods adopted to measure cellulase activity as well as the definition of enzyme units, have led to much debate (see Coughlan, 1985b; Enari and Niku-Paavola, 1987). The requirement for consistency in comparative studies has led to the widespread use of a number of standardised "model" substrates.

Assay of total cellulase activity. Crude cellulase preparations normally consist of a mixture of endo- and exoglucanases and β -glucosidase all acting synergistically to achieve hydrolysis. These preparations show an initial fast rate of enzyme activity (release of reducing sugars) representing attack of the amorphous regions of the substrate by endoglucanases. In the absence of exoglucanases this activity then ceases. With exoglucanases present, the activity continues, but at a reduced rate (Coughlan, 1985b). The initial rate of activity is therefore not a measure of total cellulase activity and has led to unfounded claims about levels of total cellulolytic activity (see Mandels, 1975). As a result assay methods usually involve long incubation times.

A practical and widely used assay of total cellulase activity was devised by Mandels et al. (1976) in which

activity is assayed by measuring the quantity of enzyme required to effect a 4% hydrolysis of filter paper (50 mg) in one hour at optimum temperature and pH. An improvement of this method by the use of commercially available filter-paper antibiotic discs was suggested by Montenecourt et al. (1978). In this assay the relative proportions of the enzymes in the cellulase preparation, and particularly the concentration of β -glucosidase, can affect the result (Bailey, 1981; Joglekar et al., 1983). Furthermore, when estimating the activity of a cellulase preparation on a substrate to be used for industrial saccharification, the only worthwhile approach is to assay the rate of hydrolysis, of the actual substrate to be used, over a long period. Guidelines for the assay of cellulolytic activity have been established (Ghose, 1984). The range of substrates commonly used for estimating total cellulase activity is shown in Table 1.5. An assay for total cellulase activity was also described by Canevascini (1985) using acid swollen cellulose as a substrate. This assay is based on the microdetermination of cellobiose produced during the reaction using a fungal cellobiose dehydrogenase to oxidise the cellobiose and ferricyanide as the final electron acceptor. The removal of cellobiose is continuous thus preventing feedback inhibition of the cellulases, interference by β -glucosidase is prevented by glucono- δ -lactone inhibition and residual glucose does not interfere with the assay. Oxidation of cellodextrins produced by the endoglucanases is apparently as effective as cellobiose, making the assay applicable to these enzymes and exoglucanases or a combination of both such as in crude

preparations. A disadvantage of this method however is that the cellobiose dehydrogenase used in the assay, which is prepared from a cellulolytic Sporotrichum species, must be free of the cellulase activity.

Assay of specific enzymes of the cellulase system. Due to the heterogeneity of insoluble substrates, various model substrates are used in the assay of these enzymes (Table 1.5).

Soluble forms of cellulose such as CMC and HEC are commonly used substrates for endoglucanases and activity is determined by the release of reducing sugar equivalents or a reduction in the viscosity of the solution. Almin et al. (1975) described a viscosimetric assay for the hydrolysis of CMC which enables the expression of endoglucanase activity as molecular activity (A_m). The release of dye from substrates such as RBB-CMC (Biely et al., 1985a; 1985b) and Ostazin brilliant red H3B-HEC (Farkas et al., 1985), as well as the release of a bound chromophore such as trinitrophenyl-CMC (TNP-CMC) (Huang and Tang, 1976) have been used for the detection and assay of endoglucanase activity. More detailed analysis of the mode of action of T. reesei endoglucanases was achieved using 4-MeUmb oligosaccharides (DP 2 - 6) (van Tilbeurgh et al., 1982; 1985).

Exocellobiohydrolase and exoglucanase activity is assayed by the release of reducing sugars from acid swollen cellulose or by assaying directly for the formation of cellobiose or glucose. Exoglucanases show activity against acid-swollen

cellulose, but not against CMC. Oligosaccharide substrates with a bound chromophore such as TNP and 4-MeUmb have been used for studying the specificity of cellulases. van Tilbeurgh et al. (1982; 1985) showed differences in the activity of T. reesei CBH I and CBH II using 4-MeUmb oligosaccharides.

The use of modified substrates apply to purified preparations of these enzymes, since mixtures with other cellulolytic enzymes can have synergistic effects. Sometimes the modified substrates produce unexpected results (van Tilbeurgh et al., 1982) and, as pointed out by Eveleigh (1988), should be used with caution.

β -glucosidase activity is generally measured by the release of *p*-nitrophenol from *p*-nitrophenol- β -D-glucoside or by the release of glucose (Coughlan, 1985b). Reducing sugar is usually estimated by any of a number of reported methods (e.g. Nelson, 1952; Miller, 1959; Lever, 1977; Mullings and Parish, 1984). Dilution of the enzyme preparation can apparently affect the results (Cauchon and Le Duy, 1984). Two commonly used methods (Nelson, 1952; Miller, 1959) are not specific for sugars and can therefore be affected by other reducing compounds in the assay. The dinitro-salicylate (DNS) assay (Miller, 1959) can result in a marked over estimation of the extent of hydrolysis especially when impure lignocellulosic substrates are used (Breuil and Saddler, 1984). The methods are adequate when using purified substrates and enzymes (Coughlan, 1985b). A reliable method for determining the hydrolysis products is

by use of HPLC (Honda, 1984), although this method also appears to have limitations (Coughlan, 1985b).

Enzyme purification. Due to the complex mixture of enzymes involved in cellulose hydrolysis, purification of the individual components is a prerequisite for a better understanding of the entire system and the contribution of each component. Moreover, absolute purification is essential for the characterisation of individual cellulase components, since even low level contamination with other components of the cellulase complex can cause synergistic effects which can seriously affect the results. However in practice, such high degrees of purification have been difficult to achieve.

Standard protein purification techniques such as precipitation, gel filtration, ion-exchange chromatography and preparative isoelectric focusing have been employed in purification protocols. In some cases apparently homogeneous enzyme preparations have been shown by other methods to be a complex of different enzymes (Shewale and Sadana, 1981; Sprey and Lambert, 1983). Sprey and Lambert (1983) cautioned that purified fungal cellulases should be considered to be complexes and not single enzymes, unless homogeneity has been assessed by methods such as the titration curve technique or urea-octylglucoside treatment before polyacrylamide gel electrophoresis (PAGE). Treatment of cellulase complexes with sodium dodecyl sulphate (SDS) followed by SDS-PAGE was not sufficient to split the enzyme complex. Sprey and Lambert (1983) showed that a single

protein band of T. reesei cellulase could be resolved into six proteins by preparative IEF. The complex contained β -glucosidase, xylanase and weak cellulase activity.

Affinity chromatography was successfully used to separate T. reesei endoglucanases from cellobiohydrolases (van Tilbeurgh et al., 1984). T. reesei CBH I and II selectively bound to p -aminobenzyl 1-thio- β -D-cellobioside coupled to Affigel-10 (Biorad). Separation was achieved by elution with lactose for CBH I and cellobiose for CBH II. Another affinity technique (Nummi et al., 1981) and a technique using monoclonal antibodies in an immunomatrix (Riske et al., 1986) have also been described for the purification of cellulolytic enzymes. These techniques are highly specific and are the methods of choice (Enari and Niku-Paavola, 1987). Purification techniques and the related problems have recently been reviewed by Enari and Niku-Paavola (1987). One approach, now used extensively in the characterisation of the individual cellulase components, is to clone the respective genes into a non-cellulolytic host. Further studies on the interaction of the proteins, optimal ratios for activity etc. can then be determined.

Utilisation of lignocellulosic wastes.

These wastes are currently used in a variety of ways. They can simply be burned as fuel or ploughed back into the lands for soil conditioning. Some agricultural wastes are used as animal feed while others are finding application in the mushroom cultivation industry (Tseng and Luong, 1984).

Apart from the direct uses and partial treatments of lignocellulosic wastes for example to enrich the feed value (Avgerinos and Wang, 1980), the major potential use is in the fermentation industry for the production of food, chemicals and fuels. The raw material can be saccharified to produce a sugar syrup and then fermented by microorganisms to products such as alcohols, neutral solvents and organic acids. This approach is not new and saccharification of Southern pine mill waste using acid treatment at high temperatures was first carried out before the First World War. The resultant syrup was fermented to produce ethanol (see Klyosov, 1986).

Currently most industrial saccharification methods involve the use of fungal cellulases. In particular T. reesei has been the subject of intense research and has, by mutation and selection, been considerably improved for high level production of cellulases (see Eveleigh, 1987). However the crystallinity of the cellulose fraction and a lignin coating make a large part of the material inaccessible to enzymic degradation. As a result a number of methods of pretreatment have been developed for the conversion of lignocellulosic materials to forms suitable for hydrolysis to fermentable syrups. The methods include physical, chemical and biological processes. Dale (1985) reviewed a wide variety of pretreatment processes that have been developed, all of which increase the surface area available for chemical or enzymic treatment. The most effective pretreatments exploited to date are steam explosion and treatment with alkali (Dale, 1985; Saddler, 1986). As

pointed out by Dale (1985), many of these pretreatment methods have not been tested beyond the laboratory stage and it is not known whether they will be economically viable or even in some cases, where pretreatment is followed by acid hydrolysis, whether the resultant sugars are fermentable. Measurement of the effectiveness of a pretreatment method is also important where enzyme saccharification is used. It was thought that the extent of crystallinity (crystallinity index, CI) was the main factor in determining the efficacy of a particular method and that a reduction in CI meant increased susceptibility to enzymic treatment (e.g. Ryu et al., 1982). However it has been found that the degree of swelling or an increase in pore volume (Puls et al., 1985; Grethlein, 1985) are better indications of enzyme susceptibility. It does appear that some form of pretreatment will be an essential part of a commercially viable saccharification process.

Production of alcohols and neutral solvents. The production of ethanol from cellulose utilising cellulases derived from T. reesei and S. cerevisiae has been demonstrated (Blotkamp et al., 1978; Savarese and Young, 1978). However, the conversion of lignocellulose to its constituent sugars results in a syrup containing appreciable amounts of pentose sugars, especially xylose. A disadvantage of fermentation technology using Saccharomyces cerevisiae or Zymomonas mobilis is that these organisms do not utilise pentoses (see Buchholz et al., 1987). Some yeasts such as Pachysolen tannophilus can produce ethanol from pentoses although the

yield is low and the process appears to require the presence of oxygen (Schneider et al., 1985).

Various strains of the genus Clostridium have been shown to produce solvents from lignocellulose wastes. C. thermocellum has the ability to degrade cellulose (Avicel, compression-milled corn stover) (Johnson et al., 1982), and the direct conversion of lignocellulosic substrates to ethanol has been examined (Saddler et al., 1981; Saddler and Chan 1982; 1984). The efficiency of ethanol production decreased with increasing substrate concentration and when substrates with a high hemicellulose content were used. Various co-culture systems have been proposed to overcome these difficulties. Co-cultures using C. thermocellum and other clostridia such as Clostridium thermosaccharolyticum and Clostridium thermohydrosulphuricum (Ng et al., 1981; Saddler and Chan, 1984) or in sequential fermentation with Z. mobilis (Saddler et al., 1981) have been investigated. The co-culture systems produced more ethanol than with the mono-culture system, but yields remained low, especially with natural substrates such as steam-exploded wood (Yu et al., 1985b). Yu et al. (1985a) showed that when C. thermocellum is grown on lignocellulosic substrates both hexose and pentose accumulate in the medium. They proposed another sequential co-culture system using K. pneumoniae as the second organism. This organism was chosen for its ability to ferment the pentose and hexose sugars and produce butanediol (Yu and Saddler, 1982). The co-culture systems have advantages over the saccharification processes in that separate enzyme production is eliminated.

Certain of the saccharolytic clostridia are well documented for their ability to produce neutral solvents from a variety of hexose and pentose sugars under appropriate conditions and some have been used extensively at the industrial level for the production of acetone and butanol. The recognised strains of solvent producing clostridia are generally classified according to the types and proportions of solvents they produce (Table 1.6). The solvents produced by C. acetobutylicum are all important industrial chemicals and form the feedstocks of many industrial processes. Moreover, the fermentation technology is well established (see Jones and Woods, 1986).

Table 1.6. Solvent producing strains of Clostridium.

Strain	Solvents (ratio)	Reference
<u>C. acetobutylicum</u>	butanol, acetone ethanol (6:3:1)	Jones and Woods, 1986
<u>C. beijerinckii</u> (<u>C. butylicum</u>)	butanol, isopro- panol, ethanol (6:3:1)	George and Chen, 1983
<u>C. auranti-</u> <u>butyricum</u>	butanol, acetone, isopropanol	George and Chen, 1983
<u>C. tetanomorphum</u>	butanol, ethanol (1:1)	Gottwald <u>et al.</u> , 1984

Acetone-butanol fermentation. The acetone-butanol (AB) fermentation by C. acetobutylicum has been widely reported in the literature and recently Jones and Woods (1986) comprehensively reviewed all aspects of the fermentation, from its historical background to the cessation of the industrial fermentation.

The decline of the AB fermentation was due to biological limitations and various economic factors. Butanol toxicity restricted the levels of solvents that could be produced by batch fermentation to approximately 2%. The fermentation suffered from technical difficulties: yields were sometimes low and the fermentation was susceptible to infection, especially by bacteriophages.

High capital costs, costs of recovery of low levels of solvents and treatment of large quantities of effluent all contributed to making the fermentation unprofitable in competition with the petrochemical industry. However, the substrate cost, which accounted for approximately 60% of the total production costs, was the decisive factor in determining the economic viability of the AB fermentation.

Alternative substrates for the AB fermentation. The ability of saccharolytic clostridia to ferment a wide variety of carbohydrates (Robinson, 1922; Compere and Griffith, 1979), stimulated research into the use of alternative cheaper substrates (see Jones and Woods, 1986). Studies have included the use of substrates containing mono- and disaccharides, pentoses and various polysaccharides. Varying degrees of success have been achieved using wastes such as sulphite liquor (Wiley et al., 1941), cheese whey (Ennis and Maddox, 1985) and apple pomace (Voget et al., 1985). Apart from the conventional polysaccharide substrates like maize, wheat, rye and millet, other starch containing crops such as potatoes, rice, jawari, bajra and

tapioca have also been investigated with some success (Prescott and Dunn, 1959).

Jerusalem artichokes contain high quantities of carbohydrate in the form of fructans. Pretreatment with either acids or inulinase (to destroy the inulinic structure) results in a hydrolysate that, with the addition of only ammonia, can be fermented to solvents. This substrate has been tested to pilot plant level (Marchal et al., 1985).

C. acetobutylicum is capable of utilising the sugars resulting from the hydrolysis of cellulose and hemicellulose and 17 g/l of solvents on various combinations of sugars has been achieved (Compere and Griffith, 1979; Mes-Hartree and Saddler, 1982). Xylose is the predominant sugar resulting from the hydrolysis of hemicellulose, and the ability of C. acetobutylicum to ferment xylose to solvents has been demonstrated (e.g. Ounine et al., 1983). More recent studies have shown that C. acetobutylicum can partially degrade hemicellulose and has limited cellulase activity (Lee et al., 1985b; 1985a).

Direct fermentation of lignocellulosic wastes. The combination of solventogenic bacteria with other organisms able to hydrolyse cellulose could be a cost effective method of producing solvents since there is no requirement for a separate enzyme producing facility. However the results to date have not been encouraging. The direct fermentation of cellulose by a co-culture of C. acetobutylicum and C. cellulolyticum has been demonstrated (Fond et al., 1983;

Petitdemange et al., 1983), but the main product was butyric acid and not solvents. Similar results were obtained with co-cultures of C. acetobutylicum and C. thermocellum (Yu et al., 1985b).

Another possibility would be the direct conversion of lignocellulose to solvents by C. acetobutylicum, by enhancing the organism's inherent lignocellulolytic properties by genetic manipulation. C. acetobutylicum is one of a number of organisms which have been reported to produce components of the cellulolytic /hemicellulolytic systems (Bisaria and Ghose, 1981). Allcock and Woods (1981) reported that C. acetobutylicum P270 (a subculture of P262) produced an inducible carboxymethylcellulase (CMCase) and a β -glucosidase, although the levels of activity were low. Lee et al. (1985a; 1985b) screened 20 solvent producing C. acetobutylicum strains for cellulolytic and xylanolytic activity. Two strains had endoglucanase and β -glucosidase activity. Most of the strains (18) had xylanolytic activity and showed xylanase, xylopyranosidase and arabinofuranosidase activities. C. acetobutylicum P262 also showed xylanase activity (Zappe et al., 1987; this study).

Genetic manipulation of C. acetobutylicum.

The ability to genetically manipulate C. acetobutylicum is limited at present, but recent developments are improving the situation. Mutagenesis of this organism using certain chemical agents such as ethyl methanesulfonate has proved successful and many useful mutants were readily obtained

(Long et al., 1984; Bowring and Morris, 1985; Jones et al., 1985). Mutagenesis has been utilised for the improvement of industrial C. acetobutylicum strains (Jones and Woods, 1986).

The development of genetic transfer systems for C. acetobutylicum have proved to be a greater challenge. Bacteriophages, which have caused problems in industrial AB fermentations, are potential cloning vectors. The occurrence of lysogeny (Hongo et al., 1968) and phage induction (Ogata and Hongo, 1979) has been demonstrated although transduction has not been reported (see Jones and Woods, 1986). It may be feasible to develop bacteriophage vectors analogous to the ϕ 105 B. subtilis vectors.

Simple transformation procedures have not yet been developed for C. acetobutylicum and chemical treatments or a natural process to induce DNA uptake (competence) have not been identified for this organism. However, protoplast transformation, similar to systems described for B. subtilis, has been demonstrated. These procedures have involved the development of protoplasting and protoplast regeneration systems for C. acetobutylicum P262 (Allcock et al., 1982) and C. acetobutylicum N1-4080 (Reysset et al., 1987).

Transformation of C. acetobutylicum protoplasts is hampered by nuclease activity produced by this organism (Lin and Blaschek, 1984). Different approaches have been used to overcome this problem. Reid et al. (1983) transformed

bacteriophage DNA into protoplasts of C. acetobutylicum P262J, a strain showing reduced DNase activity. Lin and Blaschek (1984) used a heat inactivation step to inhibit DNase activity, but attempts to apply this method to other strains of C. acetobutylicum have not been successful, indicating that other factors may be involved (see Jones and Woods, 1986). A method of transformation could be developed by trapping plasmids in liposomes, thus providing protection against nucleases, and utilising existing protoplast fusion techniques (Jones et al., 1985) to transform the cells.

Conjugation also appears to be a promising system of gene transfer to C. acetobutylicum. Oultram and Young (1985) showed conjugal transfer of pAMB1 (a self transmissible broad host range plasmid isolated from Streptococcus faecalis) from Streptococcus lactis to C. acetobutylicum at high efficiency (4.1×10^{-5} to 1.4×10^{-3}) and from B. subtilis to C. acetobutylicum at low efficiencies. Successful transfer, albeit at a lower frequency, from C. acetobutylicum back to S. lactis and B. subtilis was also demonstrated. Reyset and Sebald (1985) demonstrated conjugal transfer of plasmid borne antibiotic resistance markers from streptococci to C. acetobutylicum. Three plasmids carrying combinations of erythromycin (Em), chloramphenicol (Cm), streptomycin (Sm) and kanamycin (Km) resistance determinants were transferred (pAMB1 Em^r; pIP501 Em^r Cm^r; pJH4 Em^r Sm^r Km^r). The Cm marker was the only resistance determinant not expressed in C. acetobutylicum. Recently Oultram and Young (1987) constructed a small non-conjugative plasmid pOD1 which becomes established in pAMB1-

containing B. subtilis cells in the form of a cointegrate (pAMP1::pOD1) which can then be transferred to and maintained in C. acetobutylicum with the same frequency as pAMP1. Derivatives of pOD1 containing unique restriction enzyme sites were constructed and could be used for the transfer of foreign genes into C. acetobutylicum via B. subtilis.

C. acetobutylicum would appear to be an excellent choice for genetic manipulation as an industrial fermentation organism utilising lignocellulosic waste materials as the substrate. The organism has been widely utilised for industrial fermentation processes and the technology is available. It can utilise all the major carbohydrates present in hydrolysed lignocellulose for growth and solvent production. It produces a number of industrially important solvents as opposed to yeast which produces only ethanol. It produces at least some of the necessary enzymic components of the lignocellulolytic complex. Eveleigh (1988) envisaged the construction of the "perfect" cellulolytic microbe containing fungal cellobiohydrolases, high specific activity C. thermocellum endoglucanases (Ng and Zeikus, 1981) and end-product resistant M. bispora cellobiase (Waldron and Eveleigh, 1986). As a potential candidate for such constructions, C. acetobutylicum exhibits a number of advantages when compared to yeast and Z. mobilis, which exhibit no lignocellulolytic activity and have a substrate range which is limited to hexose monomers (glucose) and disaccharides (fructose and sucrose) (see Buchholz et al., 1987).

The study of bacterial cellulolytic systems such as that of C. thermocellum has revealed a very complex system (the cellulosome) in which more than 21 different enzymes are thought to be involved (e.g. Coughlan et al., 1985). The question could be raised as to whether it would be possible to transfer such a complex system to another organism. The nitrogen fixation (nif) system also presents a genetic system of equivalent complexity to the cellulase system, however this has not deterred efforts to transfer the nif system to other organisms. Alternatively, it may not be necessary to transfer the entire system in order to obtain efficient cellulose hydrolysis in an industrial production system. The C. thermocellum cellulase complex appears to have evolved to deal with cellulose hydrolysis in nature where selection would occur for systems that ensure that the hydrolysis products released from the insoluble substrate would be available to the organism itself to the exclusion of others. Fungi by contrast, envelope the substrate with mycelia and can therefore effect hydrolysis by secreting extracellular cellulases. Studies with T. reesei cellulases have shown that a limited number of cellulases can efficiently hydrolyse cellulose, even though the optimum enzyme ratios have not been established (see Eveleigh, 1987).

If the direct fermentation of plant biomass by C. acetobutylicum is to be attempted, the characterisation of its lignocellulolytic system would be of importance. The characterisation can be attempted at different levels. Lee

and co-workers (Lee et al., 1985a; 1985b) have reported on the physiological characteristics of the cellulolytic and hemicellulolytic systems of C. acetobutylicum, and have purified and characterised some of the enzymes involved (Lee et al., 1987; Lee and Forsberg, 1987). The present study was aimed at gaining further insight at the molecular genetic level. A genebank of C. acetobutylicum DNA was established in the well characterised E. coli system and then screened for genes involved in cellulose and hemicellulose degradation, as well as genes involved in solventogenesis and nitrogen regulation. The study of the genes at the molecular level may contribute to the understanding of gene regulation in this Gram-positive anaerobe thus opening the way to manipulation of expression using techniques such as site-directed mutagenesis.

Chapter 2

Construction and screening of a Clostridium acetobutylicum P262 genomic library in Escherichia coli.

2.0 Summary.

A library of chromosomal DNA fragments from C. acetobutylicum was established in the plasmid pEcoR251, an E. coli positive selection vector. The presence of inserts was confirmed by the recovery of plasmids, larger than the parental vector, in a random selection of clones. The library was screened for expression of C. acetobutylicum genes by complementation of nutritional markers in E. coli. Complementation of E. coli auxotrophic markers argG6 and hisG1 was observed and was always associated with the acquisition of ampicillin resistance. The library was screened using a variety of methods and clones expressing glutamine synthetase (Usdin et al., 1986), alcohol dehydrogenase (Youngleson et al., 1988), xylanase (Zappe et al., 1987) and endoglucanase (Zappe et al., 1986) activities were isolated.

2.1 Introduction.

The development of recombinant DNA technology has enabled researchers to manipulate the genetic complement of an organism at the molecular level. Naturally occurring self-replicating DNA molecules and bacteriophages have been exploited in the development of vectors. Molecular cloning essentially requires the in vitro ligation of DNA fragments with a vector capable of replication in the chosen host and the transfer of the ligated DNA to the host organism. These techniques enable the propagation of specific DNA fragments (clones) in a foreign host. Techniques have been developed for the construction of DNA fragment libraries with a high probability that the entire genome of an organism would be represented.

The choice of the vector-host system in the construction of genomic libraries depends on a number of factors including the size of the DNA fragments to be cloned, ease of selection of recombinants, considerations of gene expression and the experimental objectives. Vector-host systems have been developed for a wide variety of organisms including Bacillus (Ehrlich et al., 1982), Pseudomonas (Sakaguchi, 1982), Streptomyces (Kieser et al., 1982), yeast (Beggs, 1978; Hinnen et al., 1978) and various mammalian (Muzyczka, 1980) and plant systems (Howell, 1982). Vector systems which function in E. coli have been developed to the greatest extent and include various bacteriophage derivatives (Williams and Blattner, 1979) and cosmids (Hohn

and Collins, 1980; Ish-Horowicz and Burke, 1981; Little and Cross, 1985), which are used to clone relatively large DNA fragments (20 - 40 kilobases (kb)), plasmids for cloning DNA fragments of up to approximately 15 kb, and single-stranded DNA bacteriophages such as the M13 based vectors (Messing, 1983) used to clone small DNA fragments for sequencing. A comprehensive list of vectors, their uses and their characteristics has been published (Pouwels et al., 1985).

Plasmids remain the most widely used vectors but have some disadvantages in the construction of genomic libraries, particularly relating to eukaryotes. Firstly, larger plasmids (>15 kb) were found to have much reduced transformation frequencies (Kushner, 1978). Secondly, due to the large eukaryotic genome, a large number of clones would be required to obtain a high probability of cloning any specific fragment. For example, a Drosophila plasmid library would require approximately 46 000 clones with an average insert size of 15 kb to have a 99% probability of containing a specific sequence (Clarke and Carbon, 1976). Thirdly, some eukaryotic genes are spread over large fragments (the $\alpha 2$ collagen gene was found to be spread over 38 kb (Vogeli et al., 1980)). As a result λ substitution vectors and cosmids were developed to enable the cloning of larger DNA fragments (23 kb for λ derivatives (Maniatis et al., 1982) and 40 kb for cosmids (Hohn and Collins, 1980)). These vectors are particularly useful when a hybridization probe is available to screen for a particular sequence, since the library would be relatively small (Hohn and Collins, 1980). Selection of the recombinant molecules

is achieved by in vitro packaging and infection of a suitable host (Sternberg et al., 1977; Hohn, 1979; Rosenberg et al., 1985)

Plasmid based libraries tend to be limited by the smaller size of the insert which can be incorporated into the vector (Kushner, 1978). However, very efficient cloning of moderately sized fragments (4 - 10 kb) can be achieved. When preparing a genomic library of a prokaryote, the number of clones involved is not excessively high. The number of recombinants required to obtain a library which would represent 99% of the genome of an organism, can be calculated by applying the formula:

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

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

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 method by which the DNA fragments are prepared must also be considered. Complete digestion by a 6 base pair (bp) recognition restriction enzyme could result in the sequence of interest being excluded from the range of fragments chosen as it may contain too many or too few sites for the enzyme. If expression is of prime importance then a more random method of fractionation must be used to avoid possible restriction of the gene of interest. The DNA can be randomly sheared or partially digested by a 4 bp recognition enzyme (Maniatis et al., 1978).

Plasmid cloning vectors have been developed to an extent where almost any desired feature is available (Pouwels et al., 1986) or can be constructed by combining various regions of existing plasmids. For example, there are general purpose cloning vectors such as pBR322 (Bolivar et al., 1977) and plasmids containing bacterial or bacteriophage promoters (De Boer et al., 1983; Botterman and Zabeau, 1985) that control the expression of cloned genes. Some of these expression vectors are specifically designed for high level expression of cloned genes for product harvesting (Remaut et al., 1983).

However, irrespective of the vector system utilized, selection of the recombinants from a background of reassembled parental vectors dictates the efficiency of the experiment. The most efficient cloning strategies are those which reduce the parent vector background. Since bacteriophage T4 ligase (T4 ligase) will not ligate DNA fragments lacking a 5' phosphate group (Weiss et al., 1968), dephosphorylation of the vector can substantially reduce the parent vector background.

Selection of recombinants at the transformation stage is the most commonly used method and all plasmids carry some form of selectable marker. Initially plasmid cloning vectors such as pBR322 used negative selection by insertional inactivation of an antibiotic resistance marker. However each putative recombinant colony had to be screened for sensitivity towards the specific antibiotic, which is

tedious when screening the thousands of colonies that constitute a library. A method of positive selection for pBR322 recombinants enabling direct selection of tetracycline sensitivity was later developed by Bochner et al. (1980) and improved by Kiel et al. (1987). Positive selection vectors are simpler to use and more efficient than other types of vectors or biochemical selection methods, especially in the construction of genomic libraries (Kuhn et al., 1986). These vectors typically rely upon the derepression of an antibiotic-resistance function (Roberts et al., 1980; Nilsson et al., 1983; Nikolnikov et al., 1984), inactivation of a dominant function conferring cell sensitivity to an antibiotic or metabolite (Dean, 1981; Hennecke et al., 1982; Ahmed, 1984; Burns and Beacham, 1984), removal of a lethal DNA sequence (Hagan and Warren, 1982) or inactivation of a lethal gene (Schumann, 1979; O`Conner and Humphreys, 1982; Cheng and Modrich, 1983; Vernet et al., 1985; Kuhn et al., 1986). The plasmid pEcoR251 used in this study is of the last type. It was derived from the pCL plasmids described by Zabeau and Stanley (1982) and was constructed by M. Zabeau. It consists of the E. coli EcoRI gene under the control of the λ rightward promoter, the ampicillin resistance gene from pBR322 and the pBR322 origin of replication. The EcoRI gene product, when expressed at high levels by the λ promoter is lethal unless insertionally inactivated, or regulated by a resident λ prophage. Effective temperature sensitive regulation can be achieved with plasmid pcI857, a colE1 compatible plasmid carrying the λ cI857 allele coding for a temperature-sensitive repressor protein (Remaut et al.,

1983) and conferring kanamycin resistance. The EcoRI gene has single restriction enzyme sites for HindIII, BglII and SstI. The BglII site can be used to clone fragments generated by BglII, BamHI, MboI, or Sau3A restriction endonucleases.

Since the cloning of C. acetobutylicum genes had not previously been reported, nothing was known about the regulation and expression of genes from this organism in other hosts. Therefore the E. coli based host system was chosen over other bacterial systems for its ease of manipulation and propagation of large amounts of cloned DNA. The expression in E. coli of foreign DNA from both Gram-positive and Gram-negative bacteria has been demonstrated on numerous occasions and results from other studies on C. thermocellum (Cornet et al., 1983; Millet et al., 1985) and various Bacillus species (Fukumori et al., 1986a; MacKay et al., 1986) were encouraging evidence that C. acetobutylicum genes may well be expressed in E. coli. Furthermore, 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 work described in this chapter involved the construction of a C. acetobutylicum DNA fragment library and screening of the library for gene expression by complementation of mutations or detection of enzyme activities not found in E. coli.

2.2 Materials and Methods.

2.2.1 Bacteria and plasmids. C. acetobutylicum strain P262 was obtained from National Chemical Products (NCP), Germiston, South Africa, and has been described previously (Jones et al., 1982). E. coli strains HB101 (Boyer and Roulland-Dossoix, 1969), JC1553 (Clark and Margulies, 1965) and ET8051 (Pahel and Tyler, 1979) were used as recipient strains for recombinant plasmids. The E. coli strains were obtained from the strain collection maintained in the Department of Microbiology at U.C.T. The strains, their relevant genetic markers, and references are listed in Table 2.1. The plasmid vector pEcoR251 was a gift from M. Zabeau, Plant Genetic Systems, Ghent, Belgium. A restriction map and the structure of pEcoR251 is given in Appendix E.

Table 2.1. Bacterial strains and relevant genetic markers.

Strain	Relevant Genotype	Reference
<u>C. acetobutylicum</u> P262	wild type	Jones <u>et al.</u> , 1982
<u>E. coli</u> HB101	<u>leuB6</u> <u>trpE38</u> <u>metE70</u> <u>recA13</u> <u>supE44</u>	Boyer and Roulland- Dussoix, 1969.
<u>E. coli</u> JC1553	<u>leuB6</u> <u>hisG1</u> <u>argG6</u> <u>recA1</u> <u>supE44</u>	Clark and Margulies, 1965.
<u>E. coli</u> ET8051	<u>glnA-glnG</u>	Pahel and Tyler, 1979.

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

2.2.3 Growth conditions. C. acetobutylicum was grown in buffered Clostridium Basal Medium (CBM) (O'Brien and Morris, 1971) (Appendix B) as described by Allcock et al. (1982). C. acetobutylicum was grown and protoplasts were prepared under stringent anaerobic conditions (Allcock et al., 1982). E. coli was grown in Luria-Bertani (LB) medium (Maniatis et al., 1982) and in minimal A medium (Miller, 1972) supplemented with specific carbon sources and amino acids (Appendix B). For the isolation of clones expressing xylanase or endoglucanase activity LB agar contained 0.1% (w/v) oat spelt xylan (Sigma X-0376; Lot 14F 0421) or 0.5% (w/v) CMC (low viscosity; Sigma C-8758; Lot 77C-0334; DS 0.7; DP about 1 100) respectively.

2.2.4 Isolation of DNA from C. acetobutylicum. C. acetobutylicum cells were protoplasted using a scaled-up version of the method reported by Allcock et al. (1982). The DNA was extracted using a method based on that described by Marmur (1961), with modifications to overcome the high nuclease activity exhibited by C. acetobutylicum (Lin and Blaschek, 1984). C. acetobutylicum cultures (1 l) at OD_{600} 0.5 - 0.6 were harvested by centrifugation (10 000 x g) and resuspended in 40 ml CBM containing 10% (w/v) sucrose, 12.5 mM $CaCl_2$, 12.5 mM $MgCl_2$ and lysozyme (5 mg/ml). The development of protoplasts was monitored microscopically until approximately 90% protoplasts were obtained. SDS (2%, w/v, final concentration) and EDTA (0.1 M, final

concentration) were added under aerobic conditions and the lysed culture immediately extracted with an equal volume of phenol/chloroform/isoamylalcohol (25:24:1). The aqueous phase was re-extracted, and then extracted twice with water-saturated ether. The aqueous phase was then dialysed against TE buffer (Appendix B) at 4°C. Heat treated ribonuclease (RNase) (Appendix B) was added (100 µg/ml, final concentration) and the DNA solution incubated at 37°C for 30 min. The DNA solution was extracted once with phenol equilibrated with TE buffer, and twice with ether. The DNA was then precipitated (Appendix C), resuspended in 2 ml TE buffer and stored at 4°C. 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).

2.2.5 Partial digestion and sucrose gradient fractionation of C. acetobutylicum DNA. C. acetobutylicum DNA was partially digested in a 2-fold dilution series with Sau3A endonuclease (Maniatis et al., 1982) to ascertain the optimum DNA to enzyme ratio to produce 4 - 10 kb fragments. Approximately 400 µg of DNA was then digested, phenol extracted, precipitated and resuspended in TE buffer. The DNA was heated to 60°C for 10 min prior to loading on a sucrose gradient (10 - 40%, w/v in 20 mM Tris-Cl (pH 8.0), 5 mM EDTA, and 1 M NaCl). DNA samples were centrifuged at 26 000 rpm for 24 h at 20°C in a Beckman SW28 rotor. The gradient was fractionated (0.5 ml) and samples (10 µl) of every third fraction analysed by agarose gel electrophoresis (Appendix C). Fractions containing fragments of 5 - 7 kb

were pooled, the DNA precipitated and resuspended in TE buffer.

2.2.6 Preparation of vector DNA. The vector pEcoR251 was prepared by digesting 10 µg of the plasmid with 20 units of BglII endonuclease for 1 h at 37°C. An aliquot of the DNA was analysed by agarose gel electrophoresis to confirm complete digestion.

2.2.7 Ligation of vector and insert. pEcoR251 is a positive selection vector and the EcoRI gene product was lethal unless insertionally inactivated. Therefore 5' dephosphorylation of the vector was unnecessary and an insert-vector ratio of 1:1 was used. A DNA concentration of 5 pmole/ml was used to optimize recombinant formation (M. Zabeau, pers. comm.). The ligation reaction was carried out in a sterile microfuge tube containing DNA, 5 µl ligation buffer (10X) (Appendix B) and water to 50 µl. T4 ligase (0.25 units) was added and the reaction placed at 15°C for 12 - 15 h. The reaction was stopped by phenol extraction and DNA precipitation (Appendix C), and the DNA was resuspended in 50 µl of TE buffer.

2.2.8 Transformation of E. coli HB101 competent cells for genebank construction. E. coli HB101 competent cells were prepared as described in Appendix C. Twenty transformation mixes were prepared by dispensing 2 µl of the ligated DNA and 100 µl of competent cells into sterile Falcon tubes on ice. After 10 min on ice the cells were heat-shocked (42°C, 5 min) to initiate DNA uptake, diluted with 1 ml of LB

medium, 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 then pooled, diluted (1/20) into prewarmed (37°C) LB medium containing ampicillin (Ap) (100 µg/ml), and incubated at 37°C with vigorous shaking for 1 h to allow expression of the EcoRI gene (counter selection of parental plasmids). The cells were collected by centrifugation (6 000 x g), resuspended in 4 ml LB medium, and plated on LB agar (100 µl/plate) containing Ap (100 µg/ml). Controls included unrestricted pBR322 and pEcoR251 (5 ng each), transformed as above and correct functioning of the cloning vector was monitored by plating 200 µl of each diluted expression mix before and after the 1 h incubation period. After 1 h incubation the pBR322 control showed a 1 - 2 fold increase, whereas the pEcoR251 control showed a >50 fold reduction, in the number of transformants.

2.2.9 Storage of recombinant plasmids. Plasmid DNA was prepared from pools of approximately 500 transformed Ap^R colonies using the maxiprep extraction method (Appendix C). The DNA pellets were resuspended in 1 ml TE buffer without performing the CsCl gradient centrifugation step. Aliquots of 250 µl were stored at -70°C.

2.2.10 Transformation for specific clone isolation. Competent E. coli cells were transformed (Appendix C) with recombinant pEcoR251 plasmids (2 µl of each recombinant plasmid pool) and the resultant putative clones analysed for plasmid linked activity of the selected gene. Genes

complementing nutritional markers of E. coli JC1553 (hisG1 and argG6) were isolated by selection on minimal A medium with Ap and appropriate nutritional supplements. A clone complementing the glutamine synthetase (glnA) mutation of E. coli ET8051 was isolated by selection on minimal A agar containing 0.1% (w/v) $(\text{NH}_4)_2\text{SO}_4$ as the sole nitrogen source. Clones expressing alcohol dehydrogenase (ADH) activity were isolated by transforming the genebank into E. coli HB101 and plating on LB agar with Ap. All Ap^r colonies were replica plated onto fresh LB agar containing Ap and LB agar containing Ap and 200 mM allyl alcohol (Rogers, 1986). Colonies unable to grow aerobically on plates containing allyl alcohol were assayed for plasmid linked ADH activity.

Clones expressing endoglucanase activity were isolated by transforming E. coli HB101 and plating on LB agar with Ap and containing medium viscosity CMC (0.5%, w/v). Colonies were replica plated, and the original plates washed free of any remaining bacteria before staining with Congo red (0.1%, w/v; 15 min) , followed by destaining with 1 M NaCl for 15 min (Teather and Wood, 1982). Colonies showing endoglucanase activity were identified by a clear zone around and beneath the colony. Colonies expressing xylanase activity were isolated using the same method, except that the plates contained oat spelt xylan (0.1%, w/v) instead of CMC.

2.3 Results.

2.3.1 DNA extraction from C. acetobutylicum P262.

C. acetobutylicum produced nuclease activity that interfered with the preparation of DNA which resulted in poor yields and highly degraded DNA. Reid et al. (1983) used a mutant strain of C. acetobutylicum P262 (P262J), which showed reduced nuclease activity, in transformation experiments with bacteriophage DNA. Urano et al. (1983) reported the presence of a nuclease which interfered with plasmid purification from Clostridium butyricum. Blaschek and Klacik (1984) reported that irreproducible recovery of plasmid DNA was attributed to nuclease activity in C. perfringens. C. acetobutylicum P262 culture medium was sampled during the vegetative growth phase to assay for the presence of a nuclease. Fig 2.1 shows λ DNA (2 μ g) incubated (15 min, 37°C) with samples (15 μ l) of culture medium sampled at OD₆₀₀ 0.02 (at inoculation), 0.08, 0.15 and 0.4 (lanes 2 - 5). At OD₆₀₀ 0.4 the λ DNA was totally degraded.

A simple method was used to overcome the nuclease activity for the preparation of total cellular DNA. Lysis of protoplasts with SDS followed by immediate extraction with phenol inactivated the nuclease and enabled purification of high M_r DNA (>50 kb) (Fig. 2.1 lanes 6 and 7). The nuclease inhibitor diethyl pyrocarbonate (DEP) (Fedorcsak and Ehrenberg, 1966) (0.2%, v/v) has also been used as an inhibitor of the nuclease produced by C. perfringens

(Blaschek and Klacik, 1984) and C. acetobutylicum P262 (Usdin, pers. comm.).



Fig. 2.1. Agarose gel analysis of λ DNA (2 μ g) incubated with 20 μ l of C. acetobutylicum culture medium sampled during vegetative growth. Lane 1, control; lanes 2 - 5, samples taken at OD₆₀₀ 0.02, 0.08, 0.15 and 0.4 respectively. Lanes 6 and 7 are different preparations of purified high M_r C. acetobutylicum DNA; lane 8, undigested λ DNA.

2.3.2 Isolation of recombinant DNA clones. E. coli HB101 was transformed with ligated DNA and plated on LB agar containing Ap. Approximately 15 000 colonies were obtained from two experiments. Assuming an average prokaryotic genome size of 6×10^6 bp (Starr et al., 1981; Mandelstam et al., 1982), and an average insert size of 6 kb, then for a 99% probability ($P = 0.99$) of cloning a specific fragment, the number of clones required (N) was calculated by substituting in the formula (Clarke and Carbon, 1976):

$$N = \frac{\ln(1-P)}{\ln(1-f)} = \frac{\ln(1-0.99)}{\ln 1 - \frac{6 \times 10^3}{6 \times 10^6}} = \approx 4600$$

where f is the fraction of the total genome that each insert represents.

2.3.3 Integrity of the genebank. The integrity of the genebank was assessed by isolation of plasmids, larger than the parental vector, from a random selection of clones. Fig. 2.2 shows an agarose gel of 18 recombinant plasmids. The plasmids were digested with PstI endonuclease, releasing a 1.8 kb vector fragment which confirmed the pEcoR251 origin of the recombinants. Inserts of various sizes were present in all the plasmids except in lanes 3, 5, 6, and 10. In some cases the inserts contained internal PstI restriction sites (lanes 2, 9, 11, 12, 15, 16, 19 and 20).

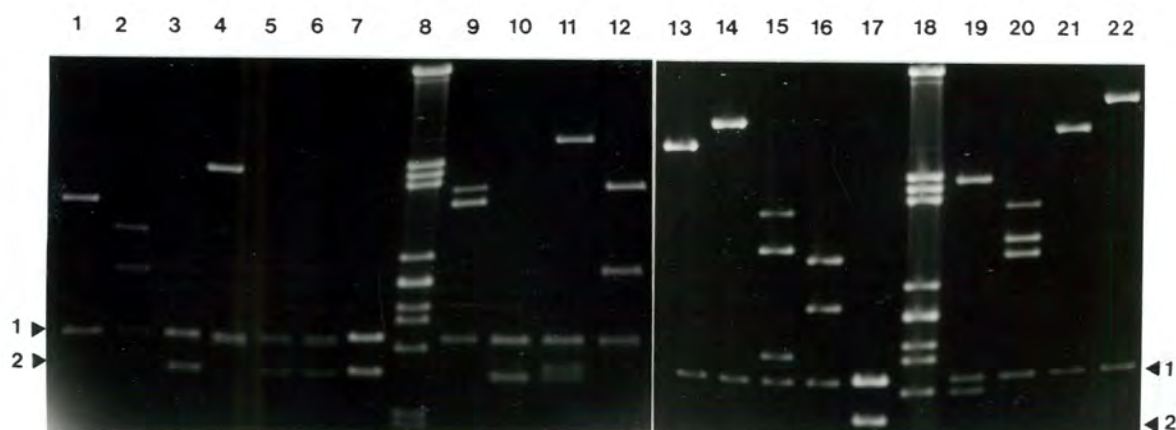


Fig. 2.2. Agarose gel analysis of a random selection of recombinant plasmids. All the plasmids and the vector (lanes 7 and 17) were digested with PstI endonuclease to release an internal vector fragment (1→). The smaller vector fragment (2→) contained the BglIII cloning site and increased in size according to the size of the insert (lanes 1, 4, 13, 14, 21, and 22). The inserts in lanes 2, 9, 11, 12, 15, 16, 19 and 20 contained internal PstI restriction sites. In lanes 3, 5, 6, and 10 no insert is apparent. DNA standards (lanes 8 and 18) were λ DNA digested with PstI endonuclease.

2.3.4 Complementation of E. coli amino acid auxotrophs. E. coli JC1553 Arg⁺ and His⁺ transformants were isolated on supplemented minimal medium lacking either arginine or

histidine, and containing Ap. Plasmid DNA was isolated and used to retransform the E. coli auxotroph. In retransformation experiments complementation of either the Arg⁻ or His⁻ phenotype was always associated with transformation of the recipient strain to Ap^r. Hybridisation experiments, using the recombinant plasmids as probes, showed that the insert DNA originated from C. acetobutylicum (results not shown). Restriction maps of these clones are shown in Fig. 2.3. No further analysis was performed on these clones.

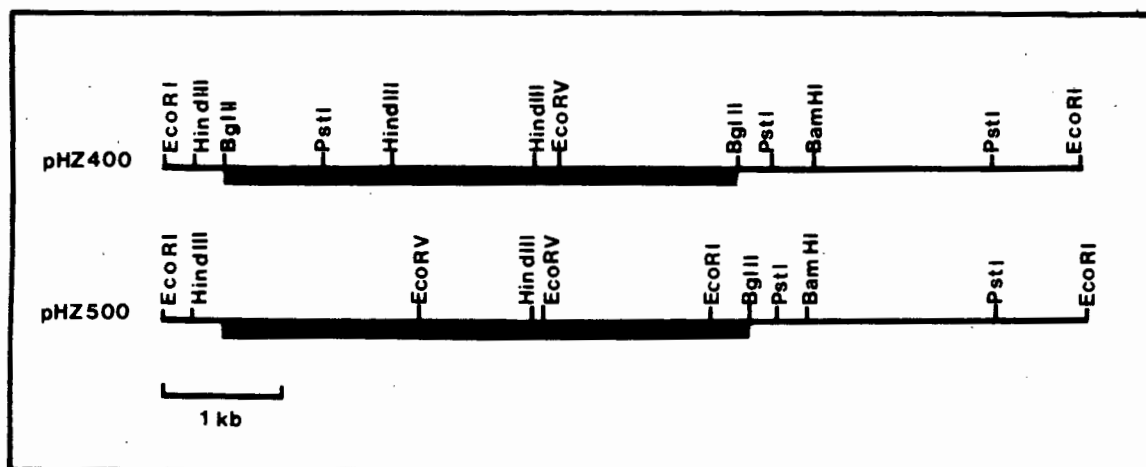


Fig. 2.3. Restriction maps of clones complementing the argG6 (pHZ400) and hisG1 (pHZ500) mutations of E. coli JC1553. The thick line represents C. acetobutylicum DNA.

2.3.5 Isolation of the C. acetobutylicum glnA gene. Four transformants from one pool of recombinant plasmid DNA complemented the GlnA⁻ phenotype of E. coli ET8051. All the plasmids isolated from the GlnA⁺ Ap^r transformants had inserts of the same size and the same PstI endonuclease digestion profile. This clone has been characterised (Usdin et al., 1986) and the C. acetobutylicum glnA gene has been sequenced (Janssen et al., 1988).

2.3.6 Isolation of a C. acetobutylicum ADH gene. Allyl alcohol has been used as a "suicide substrate" (Rando, 1974) for the selection of ADH deficient mutants of E. coli (Lorowitz and Clark, 1982) and C. acetobutylicum (Durre et al., 1986; Rogers, 1986). E. coli HB101 was unaffected by 200 mM allyl alcohol when grown aerobically, but was inhibited by 50 mM allyl alcohol when grown anaerobically (Youngleson et al., 1988). Susceptibility to allyl alcohol has been shown to correlate with the expression of ADH activity in E. coli (Lorowitz and Clark, 1982). Aldehydes in which the keto group is conjugated with a double or triple bond are potent protein alkylating agents. When the corresponding alcohols (allyl, propargyl, butynyl) are incubated with ADH in the presence of suitable co-factors (NAD or NADP) they are converted to lethal alkylating aldehydes. This phenomenon was used to select clones of E. coli HB101 harbouring putative C. acetobutylicum P262 ADH genes. An E. coli HB101 Ap^r transformant was isolated which was unable to grow on allyl alcohol plates under aerobic conditions. The cells contained a pEcoR251 recombinant plasmid which was designated pCADH100. Characterisation of the clone expressing C. acetobutylicum P262 ADH activity has been reported (Youngleson et al., 1988).

2.3.7 Isolation of a C. acetobutylicum xylanase gene. E. coli HB101 was transformed with pEcoR251 recombinant plasmid pools and plated on LB agar containing oat spelt xylan. A single clone was isolated which showed a zone of hydrolysis when the plate was stained with Congo red. Retransformation experiments confirmed the plasmid origin of

xylanase activity. Characterisation of this clone is described in Chapter 3.

2.3.8 Isolation of a C. acetobutylicum β -1,4-endoglucanase gene. E. coli HB101 was transformed with recombinant pEcoR251 plasmid pools and six transformants were isolated which produced a clear zone beneath the colony on LB agar containing CMC. Each transformant contained a pEcoR251 recombinant plasmid, and the plasmid origin of the endoglucanase activity was confirmed by retransformation of E. coli HB101. Endoglucanase activity was always associated with transformation to Ap^r. Fig. 2.4 shows the zones formed on LB agar containing CMC and stained with Congo red. E. coli HB101 or E. coli HB101 transformed with pBR322 as controls did not produce zones on CMC plates. Characterisation of the recombinant plasmids and the cloned gene product is described in Chapter 4. Sequencing and related studies of the gene are described in Chapter 5.

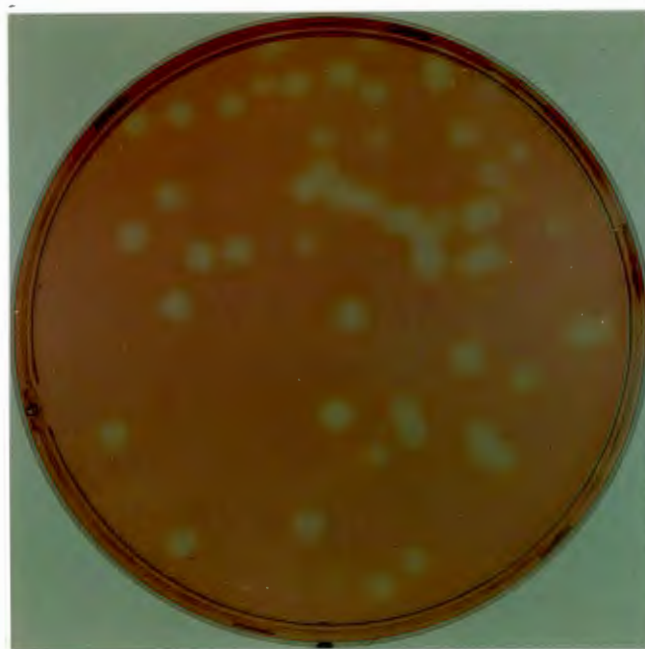


Fig. 2.4. Agar plate, containing CMC and stained with Congo red, showing zones of hydrolysis produced by E. coli HB101(pHZ100) clones expressing endoglucanase activity.

2.4 Discussion.

The preparation of high M_r DNA from C. acetobutylicum was unsuccessful using general extraction methods based on that of Marmur (1961) since this organism produced nuclease activity during vegetative growth. A number of bacterial genera including Clostridium, Staphylococcus and Serratia are known to produce extracellular nucleases (Stern and Warrack, 1964; Timmis and Winkler, 1973; Erickson and Deibel, 1973; Swiatek et al., 1987). Blaschek and Klacik (1984) found that the nuclease of C. perfringens was located primarily in the cell wall fraction (72.2%) and was released during protoplast formation. They attributed the erratic recovery of plasmid DNA from this strain to the cell wall-compartmentalised deoxyribonuclease (DNase) activity and found much improved plasmid recovery when a nuclease inhibitor such as DEP was used (Fedorcsak and Ehrenberg, 1966). The presence of nuclease activity has also been reported to cause problems with plasmid extraction from C. butyricum (Urano et al., 1983) and in transformation of C. acetobutylicum with plasmid (Lin and Blaschek, 1984) and bacteriophage (Reid et al., 1983) DNA. If a preparation of C. acetobutylicum P262 protoplasts was lysed with SDS and left at room temperature or 4°C for more than a few minutes, there was a rapid decrease in the viscosity of the solution and concomitant loss of high M_r DNA, indicating that the nuclease was SDS resistant. Extraction with phenol directly after SDS mediated protoplast lysis was employed as an effective measure to ensure the preparation of high M_r DNA. Blaschek and Klacik (1984) found when using DEP, that the

cleared lysates were extremely viscous, indicative of high concentrations of undegraded chromosomal DNA.

Two genomic libraries of C. acetobutylicum P262 DNA were constructed using pEcoR251 resulting in a total of approximately 15 000 clones which contained inserts ranging in size from 3 - 10 kb. Plasmid DNA was prepared from a random selection of clones, digested with PstI endonuclease and examined by gel electrophoresis (Fig. 2.4). Of 18 clones examined, four did not show the presence of an insert. The reason for the apparent ability of some parental plasmids to escape the lethal effects of the overexpressed ECORI gene product was not established. Presumably a low percentage of the vector DNA was damaged resulting in the disruption of the ECORI gene. However 78% of the random selection of clones examined did contain inserts of the expected size range. If this percentage is applied to the number of Ap^r colonies obtained (ca. 15 000), then approximately 12 000 would contain inserts, sufficient for a representative library.

The genomic library reported in this study has already yielded six C. acetobutylicum genes which were expressed in E. coli. The study of these genes has enabled the first characterisation of C. acetobutylicum structural genes and their regulatory regions. The genomic library is being utilised in attempts to isolate other C. acetobutylicum genes, particularly those involved in substrate utilisation, nitrogen metabolism and acid and solvent pathways.

Chapter 3

Cloning and expression of a xylanase gene from Clostridium acetobutylicum P262 in Escherichia coli.

3.0 Summary.

A xylanase gene from C. acetobutylicum P262 was cloned on a recombinant plasmid pHZ300 which enabled E. coli HB101 cells to produce intracellular xylanase activity. The xylanase gene was located on a 1.9 kb DNA fragment. The cloned xylanase had an apparent M_r of approximately 28 000 and an isoelectric point of approximately 10. Optimum xylanase activity was obtained at pH 6.0 at 37 - 43°C. Comparison with a xylanase partially purified from the culture medium of C. acetobutylicum P262 showed that the enzymes had similar characteristics and Western blot analysis showed cross-reactivity between antibodies raised against the purified cloned enzyme and a polypeptide of the same M_r from C. acetobutylicum P262.

3.1 Introduction.

Numerous studies have demonstrated that C. acetobutylicum is able to produce solvents from pentoses (see Jones and Woods, 1986). However, the ability of C. acetobutylicum to directly ferment hemicellulose to solvents has only recently been investigated.

Compere and Griffith (1979) showed that several freshly isolated strains of Clostridium species could produce solvents directly from xylan added to complex media. C. acetobutylicum has been shown to produce components of the hemicellulolytic enzyme system and xylanolytic activity in particular appears to be widespread among Clostridium species. Lee et al. (1985b) investigated the xylanolytic activity of various C. acetobutylicum strains. Of 20 strains tested, 17 were able to hydrolyse larch xylan. Two strains, NRRL B527 and ATCC 824, were examined further as these strains also exhibited endoglucanase activity.

The major hydrolysis products from larch xylan after incubation for 24 h with crude enzyme (concentrated from spent culture broth) from C. acetobutylicum ATCC 824 were xylose, xylobiose and xylotriose and 12% of the xylan was hydrolysed. The hydrolysis products were indicative of the activity of an endoxylanase and a β -xylosidase (Lee et al., 1985b). Cultures grown in oat spelt xylan also showed arabinofuranosidase activity. Under continuous culture conditions with pH maintenance (pH 5.2 or 6.0) and xylan as

the sole carbon source, C. acetobutylicum ATCC 824 utilised approximately 50% of xylan. The pH for optimum activity of the crude xylanase was pH 5.8 - 6.0, somewhat higher than the optimum pH for solvent production by this organism (pH 4.3) (Bahl et al., 1982; Huang et al., 1985). However, Holt et al. (1984) showed that C. acetobutylicum could produce solvents from glucose at neutral pH in the presence of acetic and butyric acid. Lee et al. (1987) suggested that these conditions could be used for xylan hydrolysis.

Lemmel et al. (1985) examined the ability of C. acetobutylicum ATCC 39236 to directly ferment hemicellulose (larch wood xylan) to solvents. They found that although a significant amount of hemicellulose was consumed, C. acetobutylicum ATCC 39236 was unable to completely ferment the substrate to solvents. The authors showed that the main product of larch xylan hydrolysis using crude enzyme (concentrated from spent culture broth) was xylose and concluded therefore that the primary xylan degrading enzyme of C. acetobutylicum ATCC 39236 was an exoxylanase. These results differed from those reported by Lee et al. (1985b) for C. acetobutylicum ATCC 824. The enzyme differences reported for the two strains could be due to strain differences or differences in culture conditions (Lemmel et al., 1985).

In each case the strain of C. acetobutylicum appeared to lack certain of the enzymes necessary for complete hydrolysis of xylan. Apart from β -xylosidase and exoxylanase, at least four types of endoxylanases have been

described (see Reilly, 1981). They differ in their hydrolysis products and whether they can cleave L-arabinosyl initiated branch points. The enzyme, arabinofuranosidase removes α -1,3-linked arabinose residues attached to the xylan polymer (Biely, 1985).

Subsequently, two endoxylanases (Xyn A and Xyn B) and one β -xylosidase have been purified and characterised from C. acetobutylicum ATCC 824 (Lee et al., 1987; Lee and Forsberg, 1987). Xyn A had an M_r of 65 000, hydrolysed larch xylan randomly, and produced xylooligosaccharides of DP 2 - 6 as end products. Xyn B had a M_r of 29 000, also hydrolysed larch xylan randomly, and produced xylotri- and xylobiose as end products. Xyn A showed activity towards CMC, acid swollen cellulose and lichenan (a glucose polymer of alternating β -1,4 and β -1,3 linkages). Xyn B showed no activity towards CMC or acid swollen cellulose, but was active against lichenan, but not laminarin (a glucose polymer of β -1,3 linkages). Apart from the differences in activity and general properties, the two xylanases also showed no antigenic relatedness indicating that the enzymes were encoded by two separate genes (Lee et al., 1987). The β -xylosidase isolated from C. acetobutylicum ATCC 824 showed typical β -xylosidase activity, producing xylose from xylooligosaccharides of DP 2 - 6 (Lee and Forsberg, 1987).

The studies discussed above have dealt with xylan hydrolysis and the characterisation of the xylanolytic enzymes produced by C. acetobutylicum with the aim of converting xylan to solvents at high efficiency. In order to obtain more

detailed information relating to the regulation and structure of the xylanase genes in C. acetobutylicum, a xylanase gene was cloned allowing characterisation of the gene at the molecular level. Xylanases have been cloned from Bacillus (Panbangred et al., 1983a; Bernier et al., 1983; Honda et al., 1985) and from Streptomyces (Iwasaki et al., 1986) with the aim of producing high quantities of purified enzyme for application in immobilised enzyme systems. The cloning of a xylanase gene from the rumen bacterium B. succinogenes has also been reported (Sipat et al., 1987).

Cloning a gene into a non-xylanolytic host also enables enzyme characterisation in isolation, since enzyme purity is of importance due to synergistic effects found between xylanases (Reilly, 1981). Knowledge of the structure of genes and their regulatory elements could enable alteration for high level expression in the original or a in a foreign host, for production of the enzymes for use in industrial enzymic saccharification processes.

The work described in this Chapter describes the cloning of a xylanase gene from C. acetobutylicum P262, its expression in E. coli, partial characterisation, and purification from E. coli. The general enzyme characteristics were determined in parallel with a xylanase partially purified from the culture medium of C. acetobutylicum P262.

3.2 Materials and Methods.

3.2.1 **Bacteria and plasmids.** Xylanase was purified from C. acetobutylicum P262 culture medium. E. coli HB101 (Boyer and Roulland-Dussoix, 1969) and E. coli LK111 (Zabeau and Stanley, 1982) were used as recipient strains for recombinant plasmids. The relevant genotypes are given in Chapters 2 and 5 respectively. E. coli HB101(pHZ300) was used for the preparation of the cloned xylanase.

3.2.2 **Media and Buffers.** All media, buffers and solutions not described in the text are listed in Appendix B. Chromatographic media were prepared according the manufacturer's instructions.

3.2.3 **Growth conditions.** Growth conditions for E. coli and C. acetobutylicum were described in Chapter 2 (2.2.3). For the production of xylanase by C. acetobutylicum, CBM broth contained oat spelt xylan (0.5%, w/v) (Sigma No. X-0376; Lot 14F-0421) and xylose (0.5%, w/v) instead of glucose. Xylanase producing E. coli colonies were isolated on LB agar containing xylan (0.1%, w/v). For β -galactosidase assays in E. coli, LB medium was supplemented with 1 mM isopropyl β -D-thiogalactopyranoside (IPTG). For β -lactamase assays, LB medium contained Ap (100 μ g/ml).

3.2.4 **Isolation of a C. acetobutylicum P262 xylanase gene.** A colony showing a zone of hydrolysis on LB agar containing xylan was isolated as described in Chapter 2. The

recombinant plasmid isolated from this colony was designated pHZ300.

3.2.5 Southern transfer, nick-translation and DNA hybridisation. Total cellular DNA from C. acetobutylicum was digested with BglII, EcoRV and XbaI endonucleases, resolved by electrophoresis in agarose gels (0.7%, w/v) in Tris-acetate buffer, and transferred to Gene Screen hybridisation membrane (New England Corp.) as described by Smith and Summers (1980). pHZ300, nick-translated with [α -³²P]dATP, was used as a hybridisation probe (Rigby et al., 1977). These methods are described in Appendix C.

3.2.6 Preparation of cell extracts. Cell-free extracts were prepared from 100 ml overnight cultures of E. coli HB101 (containing various plasmids) grown in LB medium with Ap (100 μ g/ml). The cells were collected by centrifugation (8 000 x g; 5 min), washed, resuspended in 5 ml 50 mM citrate-phosphate buffer (pH 6.0) and then disrupted by sonication on ice (10 s bursts for a total of 100 s). The extract was clarified by centrifugation (27 000 x g) for 15 min at 4°C.

Periplasmic extracts were prepared by the osmotic shock procedure of Willis et al. (1974). An overnight culture of E. coli HB101(pHZ300) was diluted (1/20) into fresh LB medium (100 ml) containing Ap (100 μ g/ml) and IPTG (1 mM) and incubated at 37°C. At OD₆₀₀ 0.6, NaCl and Tris-Cl (pH 7.3) were added to a final concentration of 33 mM. After 10 min the cells were collected by centrifugation

(6 000 x g; 5 min) 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 the cells were collected as before, resuspended in ice cold deionised water (20 ml/g wet wt. cells) and MgCl₂ (1 mM final concentration) was added within 1 min. The cells were then collected and used to prepare the cytoplasmic extract using the method described above for cell-free extracts. The supernatant was retained as the periplasmic fraction. Cell extracts were stored at -20°C.

3.2.7 Large scale preparation of cell free extracts. Cell-free extracts were prepared from 4 X 400 ml overnight cultures of E. coli HB101(pHZ300) grown in LB medium with Ap (100 µg/ml). The cells were washed and resuspended in 25 ml 50 mM Tris-MES buffer (pH 7.5) and then disrupted by sonication on ice (10 s bursts for a total of 200 s). Another 25 ml of buffer was added and the extract clarified by centrifugation at 27 000 x g for 15 min at 4°C. The extract was recentrifuged at 100 000 x g for 60 min at 4°C.

3.2.8 Enzyme assays.

3.2.8.1 Xylanase enzyme assays. Xylanase activity was assayed by incubating 50 µl of appropriately diluted samples with 1 ml of oat spelt xylan (Sigma No. X-0376; Lot No. 14F-0421) suspension (1%, w/v) in 50 mM citrate-phosphate buffer (pH 6.0) for 10 min at 40°C. Control experiments showed that the release of xylose equivalents under these conditions was linear for the first 30 min of the reaction.

The release of xylose equivalents was detected by the dinitrosalicylic acid reagent (DNS) for reducing sugars (Miller, 1959). After the 10 min incubation period, 2 ml of DNS solution (Appendix B) was added and the samples placed in a boiling waterbath for 5 min. Residual insoluble material was removed by centrifugation (5 000 x g; 5 min). The samples were diluted 1/4 in water and the OD₅₁₀ was measured. One unit of activity was defined as the amount of enzyme that liberates 1 umole xylose equivalents in one min and specific activity as units/mg protein. Activity against CMC was assayed by substituting xylan with medium viscosity CMC (Sigma C4888, degree of substitution 0.7). For comparative assays all samples were diluted to the same degree before assaying (Cauchon and LeDuy, 1984; Khan et al., 1986).

3.2.8.2 β -xylosidase enzyme assays. β -xylosidase activity was measured using p-nitrophenyl- β -D-xylopyranoside (4NPX) according to the method of Panbangred et al. (1983b). The reaction mixture, composed of 1 ml substrate solution (1 mg/ml 4NPX in 50 mM phosphate buffer pH 7.0) and 1 ml appropriately diluted sample, was incubated at 37°C for 10 min. The reaction was stopped by the addition of 2 ml 0.4 M Na₂CO₃ and the OD₄₀₅ of the released p-nitrophenyl was measured. One unit of enzyme was defined as the amount of enzyme that produced 1 μ mole p-nitrophenyl/min.

3.2.8.3 β -galactosidase enzyme assays. β -galactosidase activity was measured according to the method of Pardee et al. (1959) as described by Miller (1972). For the assay

0.25 ml samples of cell extract (cytoplasmic or periplasmic) were diluted with an equal volume of Z buffer (Appendix B) and equilibrated at 28°C. The reaction was initiated by addition of 0.1 ml substrate (13 mM o-nitrophenyl- β -D-galactoside (ONPG) in 0.25 M sodium phosphate buffer, pH 7.0). After 5 min at 28°C the reaction was stopped with 0.25 ml Na₂CO₃ (14%, w/v) and the OD₄₂₀ of the released o-nitrophenyl was measured. One unit of enzyme was defined as the amount of enzyme that produced 1 nmole o-nitrophenol/min.

3.2.8.4 β -lactamase enzyme assays. β -lactamase activity was assayed according to the method of Sykes and Nordstrom (1972). The reaction mixture contained 1 ml starch solution (0.2%, w/v), 1 ml substrate (0.1 mM sodium-Ap) and 0.9 ml buffer (0.1 M potassium phosphate buffer, pH 5.9). Both the starch solution and the substrate were made in the above buffer. The reaction was equilibrated at 30°C (5 min) and initiated by the addition of 0.1 ml appropriately diluted sample. After 20 min at 30°C the reaction was stopped with trichloroacetic acid (TCA) (1%, w/v) and the OD₆₂₀ was measured. The OD₆₂₀ at the start of the reaction was 1.0 - 1.2 and the sample was sufficiently diluted to ensure that the OD₆₂₀ was >0.1 after the 20 min incubation period. One unit of β -lactamase activity was defined as the amount of enzyme that hydrolyses Ap at the rate of 1 μ mole/min.

3.2.9 Purification of xylanase produced by E. coli HB101(pHZ300) and C. acetobutylicum P262. General protein purification and column chromatographic methods described by

Scopes (1982) were followed and all column chromatography was done at 4°C. Cell-free extract (50 ml) was prepared as described above. The crude sample was applied to a Sephadex G25 (medium) desalting column (19 cm² X 25 cm) equilibrated in Tris-MES buffer (50 mM, pH 7.0) and 20 ml fractions were collected (flow rate 5.7 ml/h/cm²). Fractions showing xylanase activity were pooled and applied to a CM52 (Whatman) column (5 cm² X 18 cm) equilibrated with the same buffer (flow rate 30 ml/h/cm²). Under these conditions the xylanase bound to the column whereas most of the other proteins appeared in the void volume. The absorbance (280 nm) was monitored and when this value returned to the base line the column was washed with a linear gradient of 0 - 0.3 M KCl in Tris-MES buffer and 5 ml fractions were collected. Xylanase activity was eluted at approximately 0.25 M KCl.

Xylanase was partially purified from the culture medium of C. acetobutylicum cells grown anaerobically (28 h at 34°C) in CBM (4 l) containing xylan (0.5%, w/v) and xylose (0.5%, w/v) instead of glucose. The proteins were precipitated from the culture medium by the addition of two volumes of ice cold acetone and holding at -20°C overnight. The precipitated proteins were collected by centrifugation (10 000 x g) and resuspended in 150 ml 50 mM Tris-MES buffer. All further purification steps were similar to those used for the preparation of the cloned xylanase.

Enzyme preparations were concentrated by filtration using an Amicon ultrafiltration cell (model 8050; Amicon Corp.,

Danvers, Mass.) fitted with an Amicon PM-10 membrane. Purified protein samples were stored at -20°C .

3.2.10 Protein determinations and polyacrylamide gel electrophoresis. Concentrations of proteins in crude and purified protein preparations were determined by the method of Lowry et al. (1951) using BSA (fraction V) as a standard (Appendix C). Proteins were separated by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) by the method of Laemmli (1970) (Appendix C).

3.2.11 Preparation of antibodies. A rabbit was immunised by intramuscular injection of approximately 200 μg of xylanase (0.5 ml plus 0.5 ml Freund's complete adjuvant) purified from E. coli HB101(pHZ300) cells on days 1, 8 and 14. Serum was collected on days 27, 34, and 41.

3.2.12 Western blotting of proteins. Western blotting from SDS-PAGE gels onto nitrocellulose membranes was done by the method of Towbin et al. (1979). Proteins were electroblotted onto a nitrocellulose membrane in Blot buffer (Appendix B) at 100 V (constant voltage) for 2 h using a Hoefer Transphor (model TE 42; Hoefer Scientific Instruments, San Francisco). The antibody binding and the development of bands using a goat anti-rabbit serum conjugated to horseradish peroxidase (GARP) was done according to the method of Rybicki and von Wechmar (1982), except that Tween-20 was used instead of NP-40. Non-specific binding of antibody to the filter was prevented by soaking the membrane in Blocking buffer (Appendix B)

overnight with gentle shaking. All further steps were done at 37°C with gentle shaking. The antibody was allowed to bind to the proteins by soaking the filter for 1.5 h in dilute serum (1/5 in Blocking buffer). The filter was rinsed (3 times, 10 min each) in wash saline (Appendix B) and then soaked (1.5 h) in GARP (bio-yeda Ltd., Israel) (1/1500 dilution in Blocking buffer without Na-azide). The rinsing step was repeated and the bands developed by soaking the filter in Horseradish peroxidase substrate solution (Appendix B).

3.2.13 Isoelectric focusing. The xylanase protein was isoelectrically focused in agarose gels as described in the Pharmacia manual (Isoelectric focusing: principles and methods, 1982). The gel contained 0.17 g IsoGel agarose (Marine Colloids, FMC Corp.) (1%, w/v), 1.99 g sorbitol (10%, w/v), 0.8 ml glycerol and 14.4 ml H₂O. After melting the agarose in the water, 1.05 ml Servalyt^R 3 - 10 (Serva No. 42940) carrier ampholyte was added and the gel cast. Samples (10 - 15 µl) were applied to the gel using applicator strips (Serva No. 42937) and electrophoresed at 4°C for 10 min at 1 Watt (constant power) followed by 10 min at 10 Watt, after which the sample applicator strip was removed, and electrophoresis continued for 45 min at 15 Watt. Protein test mixture 9 (Serva No. 39206) was used as a standard. An activity gel overlay technique was performed as described in Chapter 4 (4.2.13) except that the agarose gel contained 0.1% xylan (w/v) instead of CMC.

3.3 Results.

3.3.1 Cloning of a xylanase gene from C. acetobutylicum in E. coli. A colony producing a zone of hydrolysis on LB agar containing xylan and Ap was isolated as described in Chapter 2 (2.3.7). The E. coli HB101 cells from this colony contained a pEcoR251 recombinant plasmid, which was designated pHZ300. The pEcoR251 origin of the plasmid was confirmed by release of a 1.8 kb vector derived fragment after PstI endonuclease digestion. The plasmid origin of the xylanase activity was confirmed by retransformation of E. coli HB101. All resultant colonies showed xylanase activity which was always associated with the transformation to Ap^r. E. coli HB101 or E. coli HB101 carrying pBR322 or other pEcoR251 recombinant plasmids did not produce zones on plates containing xylan.

3.3.2 Restriction endonuclease mapping of pHZ300. A restriction endonuclease map of pHZ300 was obtained by complete single and double digestion with a number of restriction enzymes (Fig. 3.1). The C. acetobutylicum DNA insert was 5.4 kb in length and one of two possible BglII sites was regenerated during cloning of the Sau3A fragment. Three deletion plasmids were generated by digestion with BglII, EcoRV and XbaI endonucleases and were designated pHZ301, pHZ302 and pHZ303 respectively. The BglII and EcoRV endonuclease deletion plasmids expressed xylanase activity whereas the XbaI endonuclease deletion plasmid did not. A 1.9 kb HindIII - EcoRV fragment was subcloned into pUC18 and

pUC19 (see Chapter 5 for techniques) (Fig. 3.1, pHZ318 and pHZ319) and both these plasmids expressed xylanase activity.

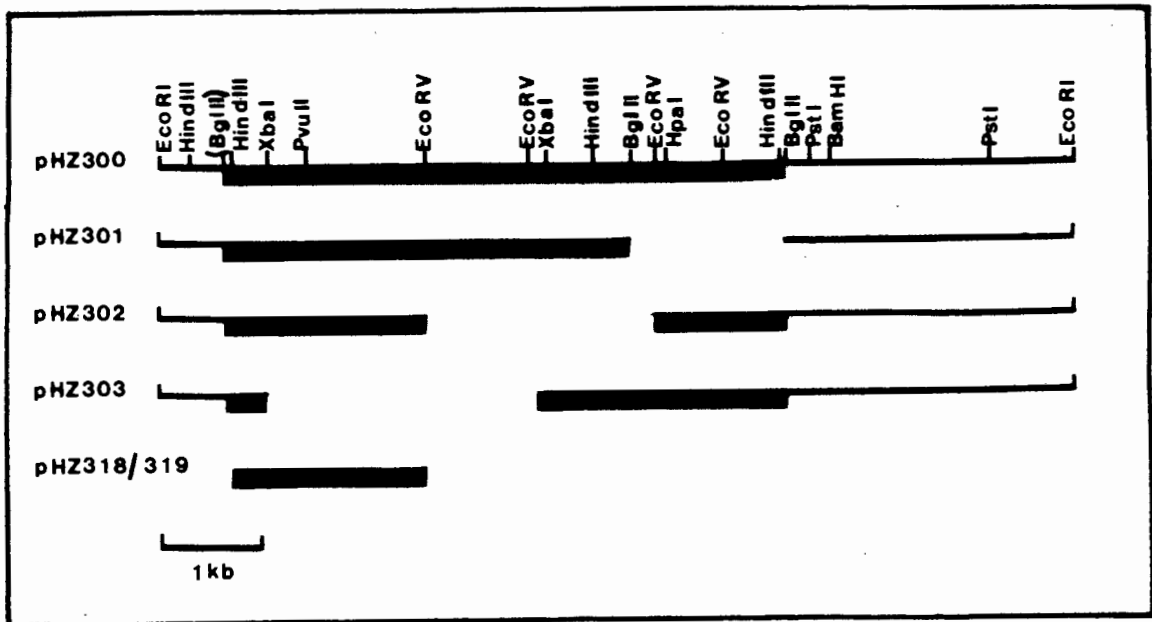


Fig. 3.1. Restriction endonuclease map and deletion derivatives of pHZ300. The thick line represents *C. acetobutylicum* DNA. The vector BglII site in parenthesis was lost during cloning. The HindIII - EcoRV fragment was subcloned into pUC18 (pHZ318) and pUC19 (pHZ319). Xylanase activity was lost in pHZ303.

3.3.3 Origin of the DNA insert. The origin of the cloned fragment in pHZ300 was determined by hybridisation of [α - 32 P]dATP labelled pHZ300 to restriction endonuclease digests of *C. acetobutylicum* chromosomal DNA (Fig. 3.2). The number of hybridising bands in lanes 1 - 3 confirmed the restriction map of the insert depicted in Fig. 3.1. pHZ300 hybridised to two fragments of *C. acetobutylicum* BglII endonuclease digested DNA (Fig. 3.2, lane 1). pHZ300 has one internal BglII site. The other BglII site (Fig. 3.1) was generated by the ligation of the BglII endonuclease cut vector and the Sau3A endonuclease cut insert DNA. EcoRV endonuclease digestion of pHZ300 produced two internal insert fragments of similar size (1.2 and

1.0 kb; Fig. 3.1) which hybridised to identically sized C. acetobutylicum DNA fragments (Fig. 3.2, lane 2). XbaI endonuclease digestion of pHZ300 produced one internal fragment of 2.7 kb (Fig. 3.1) which also hybridised to an identically sized C. acetobutylicum DNA fragment (Fig. 3.2, lane 3). The expected third hybridisation band in this lane appears to coincide with this band (comparison of band intensities). The faint hybridisation bands in Fig. 3.2, lane 3 are presumably due to incomplete digestion of the C. acetobutylicum chromosomal DNA. pEcoR251 did not hybridise to C. acetobutylicum chromosomal DNA (result not shown).

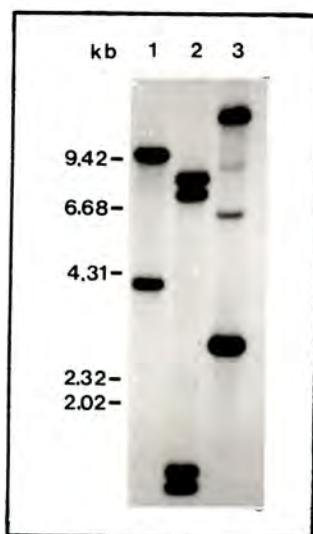


Fig. 3.2. Autoradiograph of ^{32}P -labelled pHZ300 hybridised to endonuclease digests of C. acetobutylicum DNA cleaved with BglIII (lane 1), EcoRV (lane 2) and XbaI (lane 3).

3.3.4 Expression of the C. acetobutylicum xylanase gene in E. coli HB101(pHZ300) cells. The xylanase activities of E. coli HB101 containing the original recombinant plasmid pHZ300, and the subclones pHZ318 and pHZ319 are shown in Table 3.1. The xylanase activity was expressed in both the subclones demonstrating that the gene was expressed independently of the orientation of the fragment. The

fragment subcloned into pUC18 gave consistently higher (ca. 2-fold) levels of activity.

Table 3.1. Xylanase activity of E. coli HB101 harbouring various plasmids.

Plasmid	Host vector	Insert size	Sp. act. ^a
pBR322	----	----	ND ^b
pHZ300	pEcoR251	5.4 kb	5.1
pHZ318	pUC18	1.9 kb	9.6
pHZ319	pUC19	1.9 kb	5.6

a Specific activity expressed as units/mg protein

b ND - not detected

To determine whether the expression of the xylanase gene was subject to metabolite regulation, E. coli HB101(pHZ300) cells were grown in LB medium containing either glucose (1%, w/v), xylose (1%, w/v) or xylan (0.1%, w/v). The activities of the xylanase produced under these conditions are shown in Table 3.2. The addition of xylan to the medium did not significantly alter the levels of xylanase activity. However cells grown in LB medium containing glucose or xylose showed a 5-fold and 2-fold reduction in activity respectively compared to activities obtained without these carbohydrates present.

Table 3.2. Xylanase activity of E. coli HB101(pHZ300) cells grown in the presence of various carbohydrates.

Medium	Sp. act. ^a
LB	5.1
LB + glucose	0.94
LB + xylose	2.4
LB + xylan	5.5

a Specific activity expressed as units/mg protein

β -xylosidase activity could not be detected in E. coli HB101(pHZ300) cell extracts and the cloned xylanase did not show any activity against CMC.

3.3.5 Localisation of the xylanase activity in E. coli HB101(pHZ300). The localisation experiments indicated that 95 - 98% of the xylanase activity was present in the cytoplasm of the E. coli cells. Control experiments showed that 96 - 98.5% of the total β -galactosidase activity was located in the cytoplasm, and 90 - 94% of the total β -lactamase activity was present in the periplasm. The experiments were done in triplicate and the range of values is given. Xylanase activity was not detectable in the culture medium.

3.3.6 Purification and partial characterisation of xylanase from E. coli HB101(pHZ300) cell extracts and C. acetobutylicum culture medium. The results of a typical enzyme purification experiment are summarised in Table 3.3. The xylanase from E. coli HB101(pHZ300) was purified 34-fold in specific activity with a recovery yield of 29.8%. Fig. 3.3 shows the column chromatogram on CM52. The xylanase activity and protein peaks were coincident (fraction 108) and fractions 106 to 110 had the same specific activity. Similar results were obtained with the xylanase purified from C. acetobutylicum. The xylanase from C. acetobutylicum was purified 28-fold in specific activity with a recovery yield of 48%.

Table 3.3. Purification of the xylanase from E. coli HB101(pHZ300) and C. acetobutylicum P262.

Source	Fraction	Vol. (ml)	Tot. prot. (mg)	Sp. act. (U/mg)	Yield (%)
<u>E. coli</u>	crude ext.	50	800	4	100
	After CM52	34	7	136	29.8
<u>C. aceto-</u> <u>butylicum</u>	crude ext.	150	528	1.5	100
	After CM52	12.4	13.6	28	48

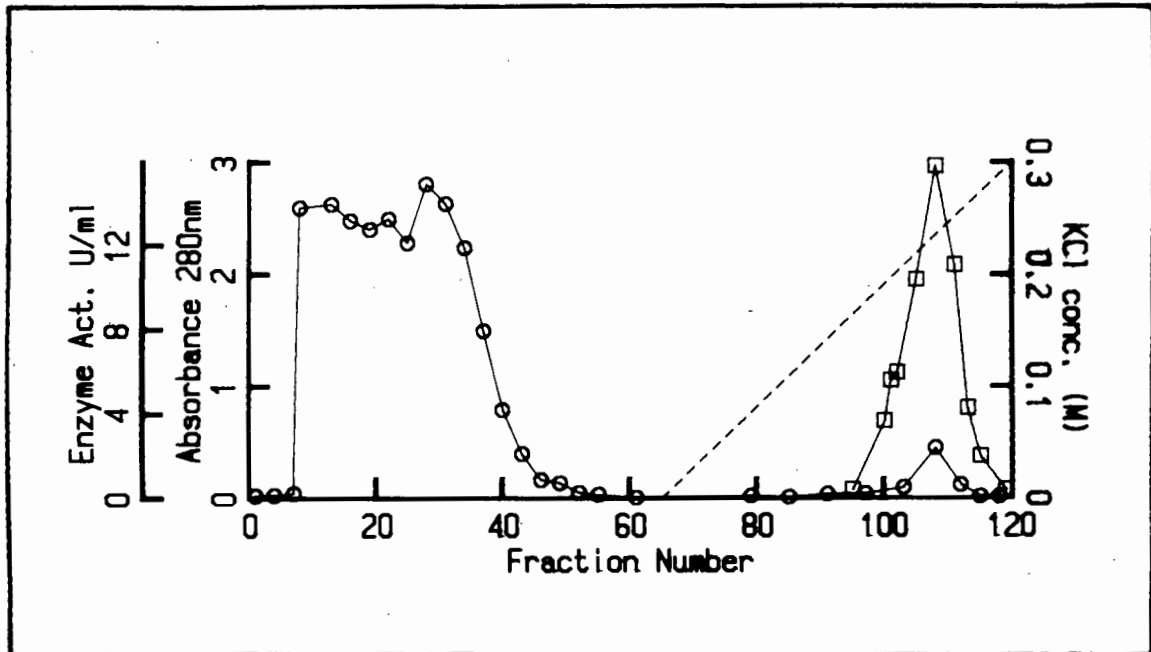


Fig. 3.3. Column chromatogram of the xylanase purified from E. coli HB101(pHZ300). Cell extract was applied to a CM52 column and elution was carried out with 50 mM Tris-Mes buffer pH 7.0 up to fraction 65, and then with a linear gradient of 0 - 0.3 M KCl in the same buffer. O , OD₂₈₀; □ , xylanase activity; --- , KCl concentration.

The xylanase purified from E. coli HB101(pHZ300) gave a single band on SDS-Page gels with apparent M_r of approximately 28 000 (Fig. 3.4A, lane 1). The partially purified xylanase from C. acetobutylicum showed a

polypeptide of the same M_r as the xylanase purified from E. coli HB101(pHZ300) (Fig. 3.4A, lane 2).

Western blot analysis showed cross-reactivity between the antiserum prepared against the cloned xylanase purified from E. coli(pHZ300) and a polypeptide of the same M_r in the partially purified preparation from C. acetobutylicum culture medium (Fig. 3.4B).

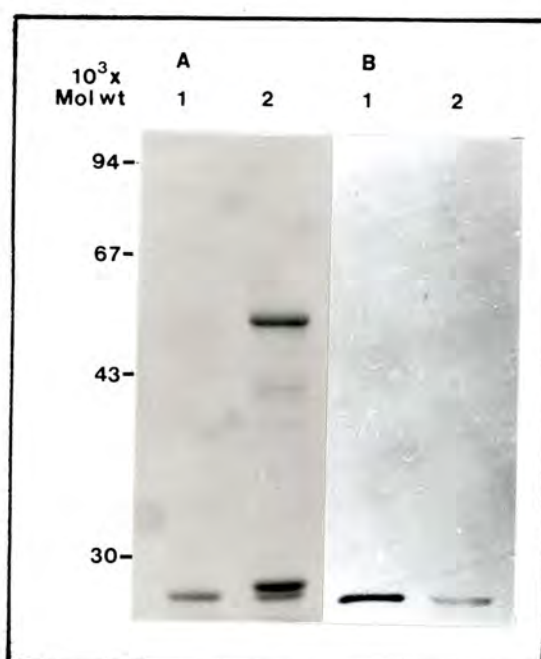


Fig. 3.4. (A) SDS-Page electrophoresis of xylanase purified from E. coli HB101(pHZ300) (lane 1) and partially purified from C. acetobutylicum (lane 2). Standards were phosphorylase b (94 000), albumin (67 000), ovalbumin (43 000) and carbonic anhydrase (30 000). (B) Western blot analysis of the xylanases. Antibodies raised against the cloned xylanase were reacted to the cloned xylanase (lane 1) and to C. acetobutylicum xylanase (lane 2).

The effects of pH and temperature on the xylanases purified from E. coli HB101(pHZ300) and C. acetobutylicum were determined and shown to be similar. The effect of pH on the xylanase enzymes (from E. coli HB101(pHZ300) and C. acetobutylicum) were determined in different citrate (pH 3.0 - 6.5) and phosphate (pH 7.0 - 8.0) buffers

(Fig. 3.5). The enzyme showed maximum activity between pH 5.5 and 6.5. The activity decreased sharply at lower and higher pH values with <25% remaining at pH 3.5 and at pH 7.5. The enzyme was stable at pH 3.5 and pH 8.5 since activity was restored when the pH was returned to pH 6.0 (the enzyme was held at the respective pH for 24 h before assaying activity at pH 6.0).

The optimum temperature for xylanase activity was between 37°C and 43°C (Fig. 3.6). Activity decreased sharply above 45°C, with >50% inactivation at 50°C and >90% inactivation at 65°C. Temperature stability was determined by holding the enzyme at various temperatures for 1 h and then assaying for residual activity at the optimum temperature (Fig. 3.7). The xylanase enzyme was relatively stable at 55°C for 1 h (approximately 80% activity remaining) but showed a sharp decrease in activity at higher temperatures.

The isoelectric point (pI) of the purified cloned xylanase enzyme was estimated from a gel and was approximately 10. The presence of the enzyme at this isoelectric point was confirmed by an activity gel overlay.

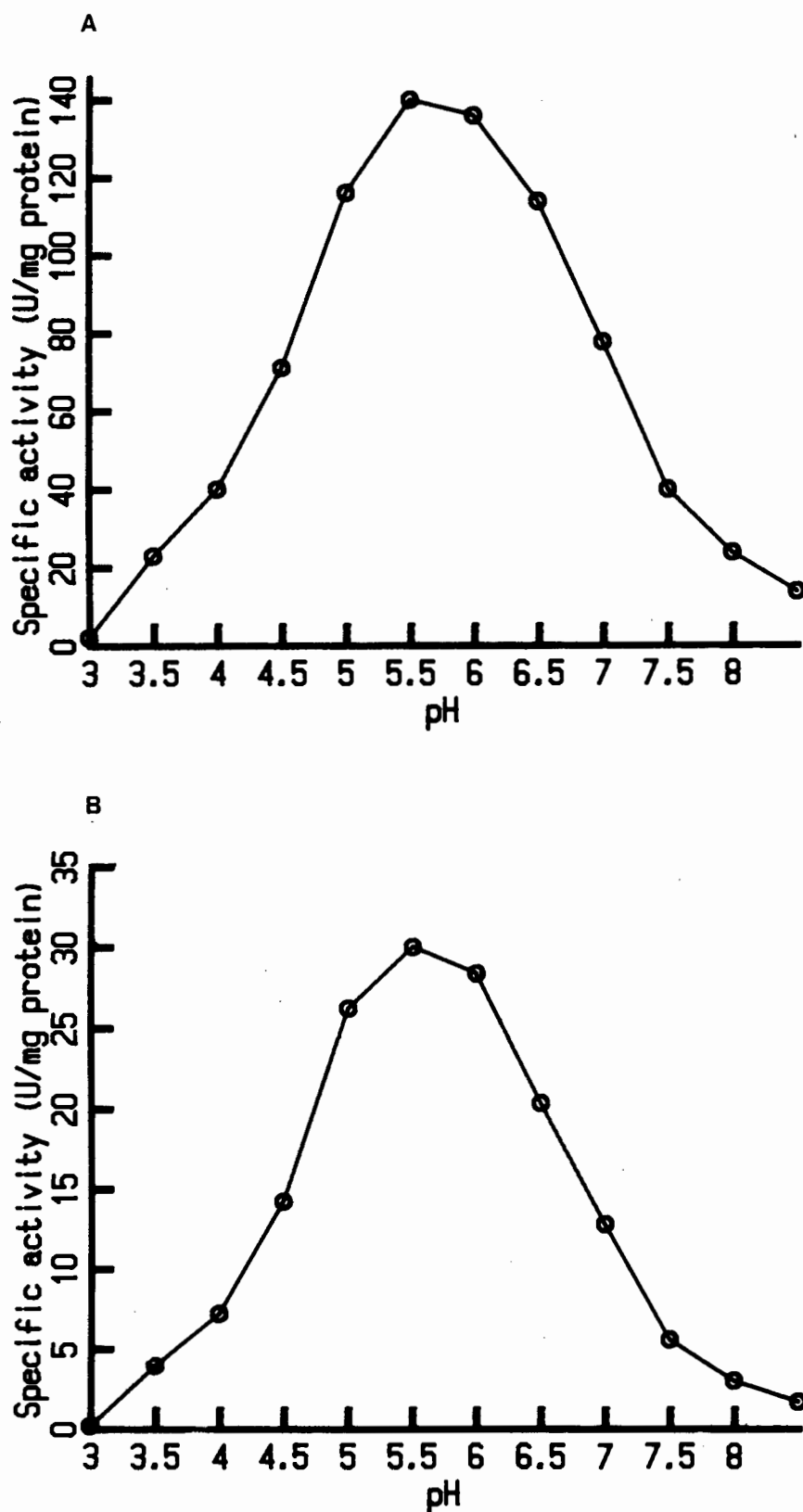


Fig. 3.5. pH activity profiles of the xylanase purified from *E. coli* HB101(pHZ300) (A) and from *C. acetobutylicum* (B). Enzyme solutions (50 μ l) were diluted in substrate solutions (1 ml) adjusted to the different pH values and the enzyme activity assayed at the particular pH.

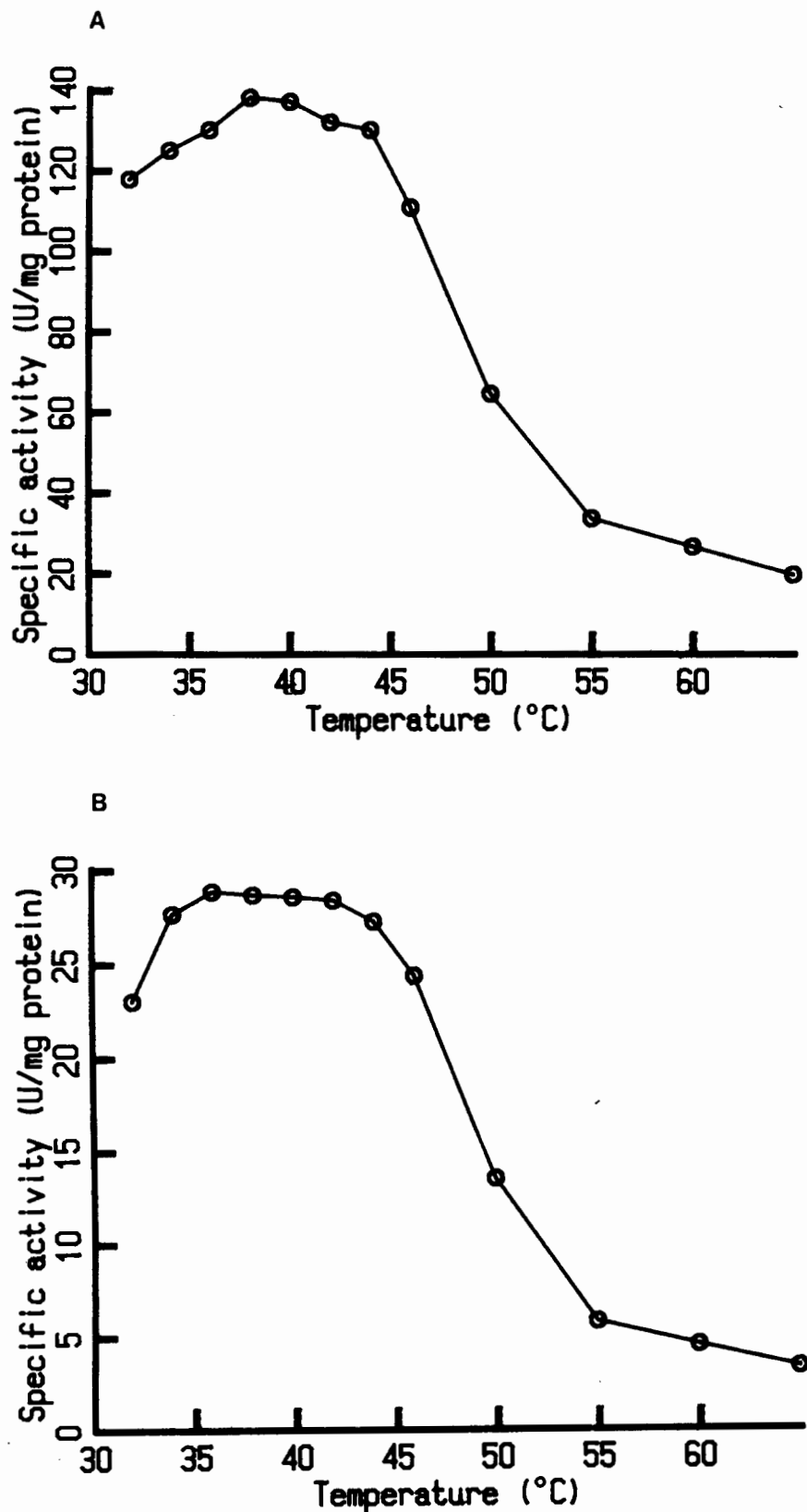


Fig. 3.6. Temperature activity profiles of the xylanase purified from *E. coli* HB101(pHZ300) (A) and from *C. acetobutylicum* (B). Enzyme solutions (50 μ l) were added to substrate (1 ml) prewarmed for 10 min at various temperatures and assayed at a particular temperature.

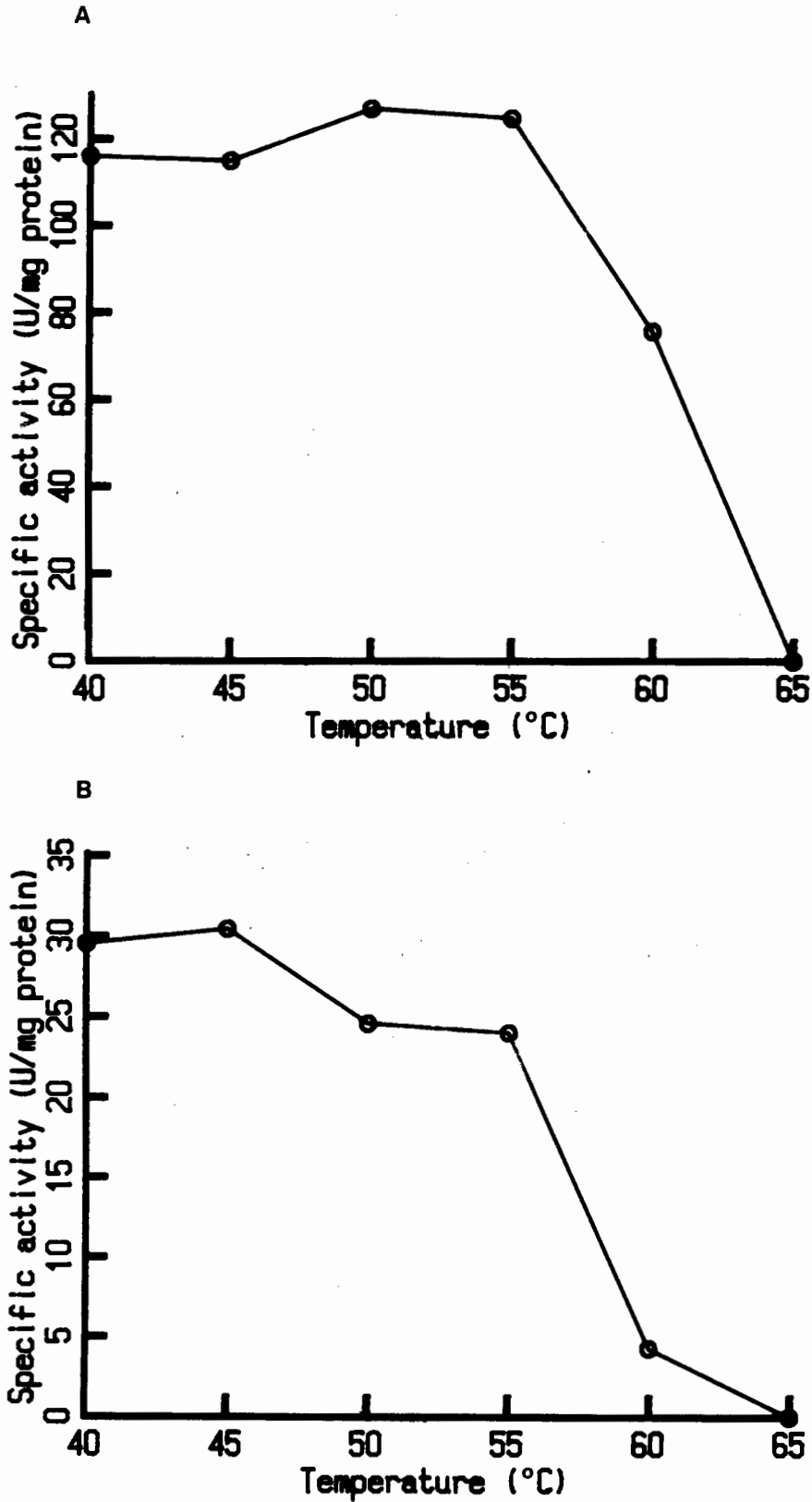


Fig. 3.7. Temperature stability profiles of the xylanase purified from *E. coli* HB101(pHZ300) (A) and from *C. acetobutylicum* (B). The xylanases were held at various temperatures for 1 h and then residual activity was assayed at 40°C.

3.4 Discussion.

Bernier et al. (1983) described a method of detecting xylanase producing E. coli colonies carrying recombinant plasmids by in situ cell lysis, which allowed the xylanase to diffuse into agar plates containing 1% xylan and produce zones. A more sensitive method was adopted in this study (Teather and Wood, 1982) which enabled direct selection of colonies expressing xylanase activity. The detection of false positives was avoided by first washing any remaining cells from the surface of the plates after replica plating and then staining with Congo red. The occurrence of zones in the agar was presumed to be due to cell leakage or lysis and the sensitivity of the detection method used since most of the xylanase activity was later found to be located in the cytoplasmic fraction of the cells. Remazol brilliant blue-xylan has also been used in the detection of xylanase producing clones (Mondou et al., 1986; Sipat et al., 1987).

The presence of a C. acetobutylicum xylanase gene in the cloned fragment was confirmed on the basis of 5 criteria: (1) Xylanase activity was not detected in E. coli HB101 cells alone or carrying other plasmids (pBR322 or other pEcoR251 recombinant plasmids) whereas 100% of colonies transformed with purified pHZ300 gave zones on LB xylan plates stained with Congo red; (2) hybridisation experiments showed that the fragment hybridised to specific C. acetobutylicum DNA restriction enzyme fragments which confirmed the restriction map of the fragment and vector DNA

did not hybridise to C. acetobutylicum DNA; (3) the expression of the xylanase gene in E. coli HB101(pHZ300) cells was demonstrated and the enzyme was purified from E. coli HB101(pHZ300); (4) xylanase partially purified from the culture medium of C. acetobutylicum P262 showed similar characteristics to those obtained for the cloned xylanase; (5) antigenic cross-reactivity between the antibodies raised against the cloned xylanase and a xylanase from the culture medium of C. acetobutylicum.

A 1.9 kb fragment of C. acetobutylicum DNA, subcloned from pHZ300, expressed xylanase activity when inserted in both pUC18 (pHZ318) and pUC19 (pHZ319) thereby demonstrating that the gene was probably expressed from its own promoter in the E. coli host. The reason for the elevated xylanase expression levels of pHZ318 was not established but could be the result of a fusion between the xylanase promoter and the lacZ promoter of the pUC vector. Certain promoter fusions have been shown to give elevated expression levels and this phenomenon has been exploited in the construction of expression vectors (De Boer et al., 1983).

To ascertain if pHZ300 encoded xylanase expression was subject to metabolic regulation in E. coli, cells harbouring the plasmid were grown in medium containing glucose, xylose or xylan. Addition of glucose or xylose resulted in a decrease of xylanase activity suggesting that gene expression was subject to catabolite repression by these carbohydrates. The addition of xylan to the medium did not significantly affect expression of the xylanase gene

suggesting that enzyme synthesis was not repressed nor inducible by xylan in the E. coli host. Sipat et al. (1987) cloned a xylanase gene from B. succinogenes in E. coli and found that xylanase expression was not affected by the addition of glucose or xylose to the medium (no catabolite repression) whereas addition of xylan resulted in a reduction of activity. The authors did not suggest a reason for this apparent anomaly. An endoglucanase gene cloned from the same organism was subject to glucose repression (Taylor et al., 1987).

Panbangred et al. (1983a) constructed a library of B. pumilus DNA in E. coli and in selecting for β -xylosidase activity, found a clone expressing both β -xylosidase and xylanase activity. Restriction analysis revealed two structural genes located on the same DNA fragment $1.95 - 3.6 \times 10^6$ Dalton apart and the authors presented evidence to suggest that the genes were controlled by the same operon in B. pumilus. The pHZ300 insert was 5.4 kb in length and subclones showed that the xylanase gene was located on a 1.9 kb fragment. In order to determine if a β -xylosidase gene was present on the 5.4 kb DNA fragment, cell extracts prepared from E. coli HB101(pHZ300) were assayed for β -xylosidase activity. β -xylosidase activity could however not be detected.

In the various strains of C. acetobutylicum which have been investigated, the endoglucanase and xylanase are excreted from the cell as extracellular enzymes (Allcock and Woods, 1981; Lee et al., 1985a; Lee et al., 1985b). Localisation

experiments showed that in E. coli HB101(pHZ300) cell extracts 98% of the xylanase activity occurred in the cytoplasmic fraction. This differed from the results obtained with the endoglucanase gene cloned from the same organism (Zappe et al., 1986; Chapter 4) where a large percentage of the activity was located in the periplasm. Panbangred et al. (1983a) also showed that the activity of the cloned B. pumilus xylanase gene was present in the cytoplasm of the E. coli host, whereas 40% of the activity of a xylanase gene cloned in E. coli from an alkalophilic Aeromonas spp. was found in the periplasm (Kudo et al., 1985). Periplasmic xylanase activity in E. coli was also reported for a cloned B. subtilis xylanase gene (Bernier et al., 1985).

The pH activity optima of xylanases characterised from various organisms including C. acetobutylicum (Lee et al., 1985b; Lee et al., 1987), Clostridium stercoararium (Berenger et al., 1985) and B. pumilus (Panbangred et al., 1983b) are all in the range of pH 6. However the optimum temperature for activity of the xylanase from C. acetobutylicum P262 was 37 - 43°C whereas an optimum of 65°C was reported for the crude xylanase preparation from C. acetobutylicum ATCC 824 (Lee et al., 1985b) and for a purified C. stercoararium xylanase (Berenger et al., 1985). B. pumilus xylanase had a temperature optimum of 45 - 50°C (Panbangred et al., 1983b).

The reported M_r values for various xylanases differ quite considerably. The xylanase from an alkalophilic Aeromonas spp. had an apparent M_r of 135 000 (Kudo et al., 1985),

whereas 3 xylanases purified from C. stercorarium had molecular weights of 44 000, 72 000 and 62 000 (Berenger et al., 1985). Streptomyces lividans xylanase had an apparent M_r of 43 000 (Mondou et al., 1986), the same as that reported for an alkalophilic Bacillus species C-125 (Honda et al., 1985). The C. acetobutylicum P262 xylanase had an apparent M_r of approximately 28 000 which was similar to that reported for B. pumilus (22 000) (Panbangred et al., 1983b) and B. subtilis (22 000) (Bernier et al., 1983).

Although the specific hydrolysis products of the C. acetobutylicum P262 xylanase were not determined, the enzyme properties are similar to those reported for Xylanase B of C. acetobutylicum ATCC 824 and Xylanase B had an M_r of 29 000 (Lee et al., 1987).

In general the pH and temperature optima and M_r values for the C. acetobutylicum xylanase appear to show the greatest similarity to those reported for Bacillus species. The endoglucanase from C. acetobutylicum P262 also showed similarity to those reported from various Bacillus species and these similarities were further confirmed by the extensive amino acid sequence homology between these proteins.

The possibility exists for improving the xylanolytic activity of C. acetobutylicum by modifying the genes for high level expression and re-introducing them into C. acetobutylicum using plasmid vectors (Lin and Blaschek,

1984). Xylanase genes from other organisms, complementing those already present in C. acetobutylicum, can also be introduced in this way. Alternatively, xylanase genes from C. acetobutylicum can be used to improve specific characteristics of other industrially important microorganisms.

Chapter 4

Cloning and expression of an endoglucanase gene from Clostridium acetobutylicum P262 in Escherichia coli.

4.0 Summary.

An endo- β -1,4-glucanase gene was isolated from a library of C. acetobutylicum P262 DNA fragments cloned in E. coli HB101. Of 15 000 colonies screened, six produced clear zones on Congo red stained CMC agar. Plasmids recovered from five of these colonies, showed identical restriction endonuclease digestion profiles and contained the same 7.2 kb DNA insert. This plasmid was designated pHZ120. A plasmid isolated from the remaining colony contained a 4.9 kb insert and was designated pHZ100. Restriction analysis showed that the 4.9 kb fragment was contained within the 7.2 kb fragment. The endoglucanase gene was shown, by deletion and subcloning, to be expressed from its own promoter in E. coli cells and 75% of the endoglucanase activity was located in the periplasm of the E. coli host. Cellobiase activity was detected, but there was no activity against Avicel. The endoglucanase was partially purified from E. coli(pHZ100) cell extracts and migrated as two bands on non-denaturing protein gels. The endoglucanase showed optimal activity at pH 6.0 and 50°C. The enzyme was unstable at higher temperatures with <1% of activity remaining after 1 h at 60°C. The endoglucanase gene was not induced by CMC or cellobiose but was subject to catabolite repression by glucose in the E. coli host.

4.1 Introduction.

Certain C. acetobutylicum strains have the ability to produce components of the cellulase complex but are unable to degrade crystalline cellulose (Allcock and Woods, 1981; Lee et al., 1985a). Allcock and Woods (1981) reported that C. acetobutylicum P270 (a subculture of strain P262 used in this study) produced an endoglucanase which was induced by an unidentified factor in a molasses fermentation medium. The enzyme could not be induced in CBM containing acid-swollen cellulose, Avicel, CMC, sucrose, glucose, cellobiose or xylose. It was also found that the addition of glucose to the fermentation medium did not affect the activity of the endoglucanase and its production was not repressed by glucose. A pH of 4.6 and temperature of 37°C were optimum for activity. This strain was able to degrade non-crystalline cellulose substrates such as CMC, but no activity was demonstrated against more crystalline substrates such as filter paper or Avicel.

Lee et al. (1985a) screened 21 solvent producing strains of Clostridium species for growth on a chemically defined medium with phosphoric acid-swollen cellulose as the sole carbon source. None of these strains were able to utilise the substrate. The strains were then tested for their ability to hydrolyse CMC and two strains, NRRL B527 and ATCC 824, were shown to produce zones of hydrolysis using the overlay and Congo red staining method of Teather and Wood (1982). It had been shown earlier that cellobiose

could induce the endoglucanase of strain B527, but attempts to induce cellulose utilisation with added cellobiose were not successful and growth ceased after the cellobiose was consumed (Lee et al., 1985a). Crude enzyme prepared from the culture broth of strain B527 was however able to hydrolyse acid-swollen cellulose and, to a lesser degree, Avicel. The possibility that the hydrolysis products were not fermentable was excluded by demonstrating the ability to culture the cells in acid-swollen cellulose previously hydrolysed with the crude enzyme preparation from the same organism. The endoglucanase activities reported for both ATCC 824 and NRRL B527 were found to be in the same range as those reported for other bacterial endoglucanases but substantially lower than those reported for T. reesei QM6a.

From the results reported by Lee et al. (1985a) it would appear that cellulolytic activity is not as widespread among solvent producing Clostridium strains as xylanolytic activity. The presence of at least some cellulolytic and xylanolytic activity in some strains provides a basis for the enhancement of the lignocellulosic degrading capability in this organism. In addition to the conventional techniques such as mutation and selection, genetic manipulation is likely to form an integral part in attempts to improve the existing cellulolytic characteristics of C. acetobutylicum so as to allow this organism to be used for solvent production from cellulosic wastes. In order to gain further information relating to the structure and regulation of the endoglucanase activity of C. acetobutylicum P262, the isolation and partial

characterisation in E. coli, of an endoglucanase gene from this organism was undertaken.

4.2 Materials and Methods.

4.2.1 **Bacteria and plasmids.** E. coli HB101 (Boyer and Roulland-Dussoix, 1969) was used as a host for plasmids for the preparation of cell extracts for enzyme assays. E. coli LK111 (Zabeau and Stanley, 1982) was used as the recipient strain for subcloning experiments involving pUC vectors. The relevant genotypes are given in Chapters 2 and 5 respectively. Endoglucanase was partially purified from E. coli HB101(pHZ100).

4.2.2 **Media and Buffers.** All media, buffers and solutions not described in the text are listed in Appendix B. Chromatographic media were prepared according to the manufacturer's instructions.

4.4.3 **Growth conditions.** Growth conditions for E. coli were described in Chapter 2 (2.2.3). Endoglucanase producing E. coli colonies were isolated on LB agar containing medium-viscosity CMC (0.5%, w/v) (Sigma no. C4888; degree of substitution 0.7). For β -galactosidase assays in E. coli LB medium was supplemented with 1 mM IPTG. For alkaline phosphatase assays cells were grown in the phosphate limiting medium of Garen and Levinthal (1960) (Appendix B).

4.2.4 **Isolation of a C. acetobutylicum P262 endoglucanase gene.** Six colonies showing zones of hydrolysis on LB agar containing CMC and stained with Congo red were isolated as described in Chapter 2.

4.2.5 Southern transfer, nick-translation and hybridisation of DNA. C. acetobutylicum total cellular DNA, prepared as described in Chapter 2, was digested with PstI, HindIII, EcoRI and BglII endonucleases. The restriction fragments were resolved by electrophoresis in agarose gels (0.7%, w/v) in Tris-acetate buffer and transferred to GeneScreen^R hybridisation membrane (New England Corp.) as described by Smith and Summers (1980) (Appendix C). pHZ100, nick-translated with [α -³²P]dATP, was used as a hybridisation probe (Rigby et al., 1977) (Appendix C).

4.2.6 Preparation of cell extracts. Cell-free extracts were prepared as described in Chapter 3 (3.2.6). Periplasmic extracts were prepared by a scaled-up version of the chloroform shock procedure of Ames et al. (1984). An overnight culture of E. coli HB101(pHZ100) was diluted (1/20) into fresh LB medium (100 ml) containing Ap (100 μ g/ml) and IPTG (1 mM) and incubated at 37°C. At OD₆₀₀ 1.0 the cells were collected by centrifugation (6 000 x g, 5 min). The cell pellet was resuspended in the residual broth by vortexing to form a paste and then 1 ml of chloroform was added. The cell-chloroform mixture was vortexed (1 min) and then left at room temperature for 15 min. Thereafter 20 ml of 10 mM Tris-HCl (pH 8.0) was added, the solution was mixed gently and the cells pelleted by centrifugation (10 000 x g, 10 min). The supernatant was retained as the periplasmic extract and the cell pellet was used to prepare the cytoplasmic extract as described (3.2.6).

4.2.7 Large-scale preparation of cell extracts. Cell extracts were prepared as described in Chapter 3 (3.2.8) but 10 mM Tris-MES buffer (pH 7.5) was used.

4.2.8 Preparation of acid-swollen cellulose. Acid-swollen cellulose was prepared by the method of Tansey (1971) with minor modifications. All steps were carried out at 4°C. Ortho-phosphoric acid (85%, 400 ml) was added slowly with vigorous stirring to 30 g of Whatman CF11 cellulose powder. After a further 2 h of stirring 2 l of distilled water was added (with stirring) and the mixture was filtered under suction through 3 layers of filter paper (Whatman No. 1). The acid-treated cellulose was washed on a suction filter with 3 l of distilled water and then suspended in 0.5 l of Na_2CO_3 (2%, w/v). The suspension was homogenised in a blender for 5 min, stored at 4°C for 12 h, washed, suspended in 6 l of distilled water, pelleted by centrifugation (10 000 x g, 5 min) and blended for a further 5 min. If the pH of the cellulose slurry was not equivalent to that of distilled water, further washing steps were carried out.

4.2.9 Enzyme assays.

4.2.9.1 Cellulase enzyme assays. Enzyme activity was assayed by the release of glucose equivalents as detected by the DNS reagent for reducing sugars (Miller, 1959) as described in Chapter 3 (3.2.8.1), using 1% (w/v) medium viscosity CMC as a substrate and omitting the centrifugation step. One unit of activity was defined as the amount of

enzyme that liberates 1 μ mole glucose equivalents in 1 min and specific activity as units/mg protein.

Endo- β -1,4-glucanase activity was assayed by incubating 50 μ l of appropriately diluted samples with 1 ml of medium-viscosity CMC (1%, w/v) (Sigma no. C4888; degree of substitution 0.7) in phosphate-citrate (PC) buffer (50 mM K_2HPO_4 , 14 mM citric acid, pH 6.0) for 10 min at 50°C. Preliminary experiments showed that the release of glucose equivalents under these conditions was linear in a defined range. Activity against xylan was assayed by substituting CMC with xylan (1%, w/v) in the above assay.

The presence of endoglucanase activity was also detected by the use of a rapid viscosimetric assay (Cornet et al., 1983; Lejeune et al., 1986). Appropriately diluted enzyme samples (50 μ l) were combined with 1 ml CMC solution (1.5%, w/v) in PC buffer and incubated at 50°C for various times. The reaction was stopped by heating to 75°C for 10 min. The relative viscosity of the mixture was determined by measuring the efflux time of a 0.2 ml sample through a 0.2 ml pipette at room temperature. A decrease in the efflux time with reference to a control was used as an indication of endoglucanase activity (Lejeune et al., 1986).

Enzyme activity against acid-swollen cellulose was assayed by incubating 0.8 ml cellulose suspension in PC buffer (20 mg dry weight/ml) with 0.2 ml enzyme solution for 30 min at 50°C. Enzyme activity against filter paper (Whatman no. 1) was assayed by incubating 50 mg strips of filter

paper in 0.8 ml PC buffer with 0.2 ml enzyme solution for 0.5 - 24 h at 40 and 50°C. Enzyme activity against Avicel PH-102 (FMC Corp.) was assayed by incubating 0.2 ml enzyme solution with 1 ml 0.5% Avicel suspension in PC buffer for 24 h at 40 and 50°C.

4.2.9.2 β -galactosidase enzyme assays. β -galactosidase activity was assayed according to the method of Pardee et al. (1959) as described in Chapter 3 (3.2.8.3).

4.2.9.3 Alkaline phosphatase enzyme assays. Alkaline phosphatase activity was assayed according to the method of Garen and Levinthal (1960). Phosphate limiting medium (400 ml) of Garen and Levinthal (1960) (Appendix B) was inoculated with E. coli HB101(pHZ100) (4 ml of an overnight culture previously washed in the same medium) and incubated at 37°C with shaking for 24 h. The cells were collected by centrifugation and cell extracts (periplasmic and cytoplasmic) prepared as described above (4.2.6).

Alkaline phosphatase activity was assayed by incubating 0.25 ml 2 M Tris-HCl (pH 8.0), 0.25 ml enzyme solution and 5 μ l substrate (20 mg/ml *p*-nitrophenyl phosphate in 1 M Tris-HCl (pH 8.0), final concentration 0.2 mg/ml) at room temperature for 2 h. The increase in the OD₄₁₀ of the sample as a result of the released *p*-nitrophenol was measured. Under the above conditions alkaline phosphatase activity is linearly dependent on the enzyme concentration in the range 0.001 - 1.0 OD₄₁₀ units and is independent of substrate concentration when it is in excess of 0.1 mg/ml

(Garen and Levinthal, 1960). One unit of alkaline phosphatase activity was defined as the amount of enzyme that releases 1 umole nitrophenol/min.

4.2.10 Cellobiase activity. The presence of cellobiase activity was determined by the production of acid and gas in peptone water (1 g peptone, 0.5 g NaCl/100 ml water) containing cellobiose (10 g/l), and by growth on minimal A medium containing cellobiose (2 g/l) as the sole carbon source.

4.2.11 Partial purification of the endo- β -1,4-glucanase produced by E. coli HB101(pHZ100). Cell extract (33 ml) from E. coli HB101(pHZ100) was applied to a desalting column as described in 3.2.9 with the exception that the column was equilibrated in 10 mM Tris-MES (pH 7.5). Fractions showing endoglucanase activity were pooled and applied to a DE52 (Whatman) column (5 cm² X 22 cm) equilibrated in the same buffer (flow rate 30 ml/h/cm²). The endoglucanase bound to the column under these conditions. The OD₂₈₀ was monitored and when this value returned to the base line, the column was washed with a linear gradient of 0 - 0.2 M KCl in the same buffer and 5 ml fractions were collected. Endoglucanase activity was eluted at approximately 0.1 M KCl. Further contaminating proteins were removed by applying the recovered active fraction to a CM52 (Whatman) column of similar dimensions and using 20 mM Tris-MES buffer (pH 6.5) at the same flow rate. Enzyme preparations were concentrated by ultrafiltration through an Amicon PM-10 membrane.

4.2.12 Protein determinations and polyacrylamide gel electrophoresis. Concentrations of proteins were determined by the method of Lowry et al. (1951) using BSA (fraction V) as a standard (Appendix C). Proteins were separated by SDS-PAGE (Laemmli, 1970) or by non-denaturing (native) PAGE (Appendix C).

4.2.13 Activity gel overlay. Non-denaturing PAGE gels were overlaid with an agar gel (1%, w/v) containing CMC (0.1%, w/v) in 50 mM PC buffer (pH 6.0) as described by Beguin (1983). The polyacrylamide gel was equilibrated in PC buffer (pH 6.0) (30 min at room temperature), dried for 10 min at 50°C, and the gels were sandwiched together between glass plates, sealed with parafilm, and incubated at 50°C for 30 min. After allowing the gel sandwich to cool (15 min, 4°C) activity bands in the overlay gel were developed by staining with Congo red (0.1%, w/v) (30 min) and destaining with 1 M NaCl until the activity bands were clearly visible. The agar gel was soaked in 5% (v/v) acetic acid which turned it dark blue and facilitated photography using a red filter.

4.3 Results.

4.3.1 Cloning of an endo- β -1,4-glucanase gene from C. acetobutylicum in E. coli. The C. acetobutylicum genebank was transformed into E. coli HB101 and six colonies, producing zones of hydrolysis on LB agar containing CMC and Ap, were isolated as described in Chapter 2 (2.3.8). Each of these colonies contained a pEcoR251 recombinant plasmid and retransformation experiments showed that endoglucanase activity was always associated with transformation to Ap^r. E. coli HB101 or E. coli HB101 transformed with pBR322 or any other pEcoR251 recombinant did not produce zones on Congo red stained CMC containing plates. Furthermore, cell extracts prepared from E. coli cells harbouring pEcoR251 recombinants and producing zones on CMC plates, caused a rapid decrease in the qualitative viscosimetric assay described above. Cell extracts from E. coli HB101, E. coli HB101(pBR322) or E. coli HB101 harbouring other pEcoR251 recombinants did not cause a reduction in the viscosity of the CMC solution.

4.3.2 Restriction endonuclease mapping of endo- β -1,4-glucanase encoding plasmids. Plasmid DNA isolated from one of these colonies contained a 4.9 kb DNA insert and was designated pHZ100. The other five colonies contained plasmids with identical PstI, HindIII and EcoRI restriction endonuclease digestion profiles and each contained a 7.6 kb DNA insert. A plasmid carrying the 7.6 kb fragment was designated pHZ120. A restriction endonuclease map of both

plasmids was obtained by single and double digestion with a number of restriction enzymes (Fig. 4.1).

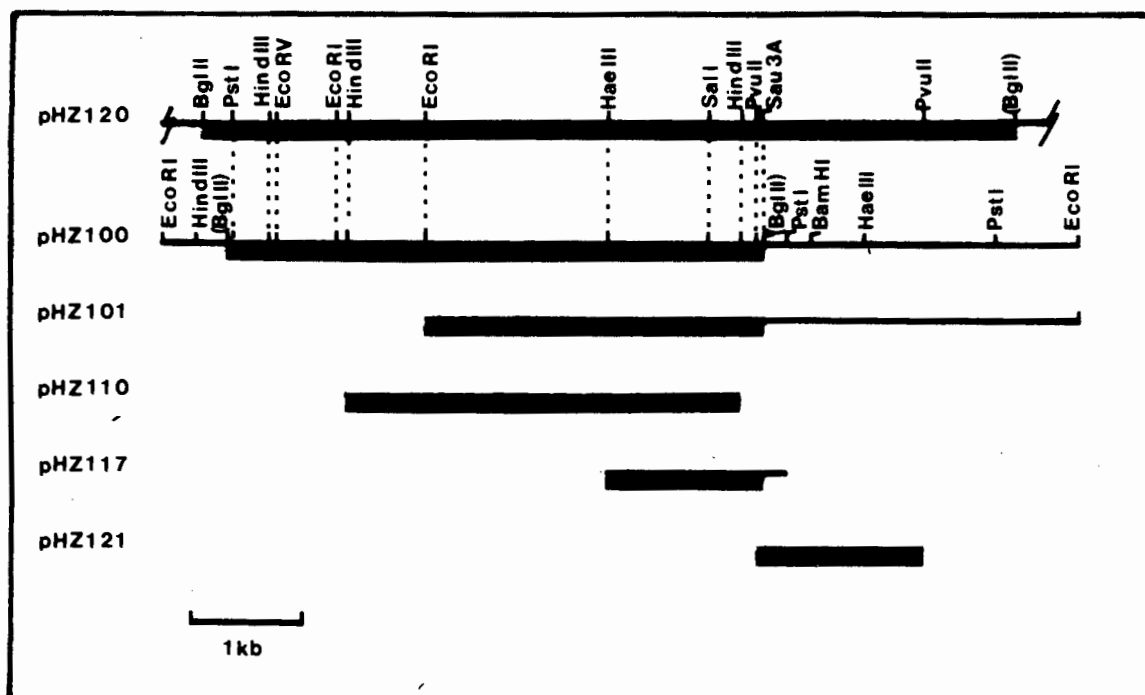


Fig. 4.1. Restriction maps of pHZ100, pHZ120 and deletion derivatives. The thick line represents *C. acetobutylicum* DNA, the thin line vector DNA. The vector region of pHZ120 is not shown in full. The *C. acetobutylicum* DNA fragments cloned into pHZ100 and pHZ120 were aligned to show the common sequences. The DNA fragment cloned into pHZ120 was 200 bp longer at one end and 2.2 kb longer at the other end compared to the pHZ100 fragment. The vector *Bgl*III sites shown in parenthesis were lost during cloning of the *Sau*3A fragments. Deletion plasmids pHZ101 and pHZ117 were able to direct the synthesis of an active endoglucanase whereas pHZ110 could not.

The DNA fragment cloned in pHZ100 was contained within the fragment cloned in pHZ120 since the restriction maps could be aligned as shown in Fig. 4.1. In pHZ100 neither of the two possible *Bgl*III sites was recovered (shown in parenthesis) whereas in pHZ120, one of the *Bgl*III sites was

present (Fig. 4.1). An EcoRI endonuclease deletion derivative of pHZ100 (pHZ101) retained endoglucanase activity as did subcloning of a 1.7 kb HaeIII - PstI fragment into pUC19 (pHZ117). However when a 2.9 kb EcoRI - HindIII fragment was subcloned into pBR322 (pHZ110) endoglucanase activity was lost. More detailed mapping and sequencing of the endoglucanase gene is described in Chapter 5.

4.3.3 Origin of the DNA insert. The origin of the cloned fragment in pHZ100 was determined by hybridisation of [α -³²P]dATP labelled pHZ100 to restriction endonuclease digests of C. acetobutylicum chromosomal DNA (Fig. 4.2). The number of hybridising bands in lanes 1 - 4 confirmed the restriction map depicted in Fig. 4.1. pHZ100 hybridised to one fragment of BglII endonuclease digested C. acetobutylicum DNA confirming the absence of BglII sites in pHZ100 (Fig. 4.2, lane 1). EcoRI endonuclease digestion of pHZ100 produced one internal insert fragment (0.8 kb) which hybridised to an identically sized C. acetobutylicum DNA fragment (Fig. 4.2, lane 2). Digestion of pHZ100 with HindIII endonuclease produced two internal insert fragments (3.2 and 0.6 kb). The larger fragment hybridised to a similarly sized C. acetobutylicum DNA fragment (Fig. 4.2 lane 3). The smaller fragment presumably migrated off the end of the gel. The faint bands hybridising in this lane (→) were presumably due to incomplete digestion of the C. acetobutylicum DNA. The PstI site in the pHZ100 insert mapped to the extreme end of the fragment which resulted in a very weak second hybridising band visible above the major

band (Fig. 4.2, lane 4). pEcoR251 DNA did not hybridise to C. acetobutylicum chromosomal DNA.

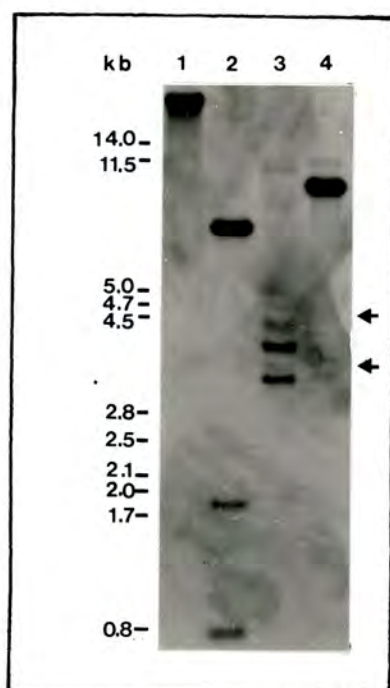


Fig. 4.2. Autoradiograph of ^{32}P -labelled pHZ100 hybridised to endonuclease digests of C. acetobutylicum chromosomal DNA cleaved with BglIII (lane 1), EcoRI (lane 2), HindIII (lane 3) and PstI (lane 4) restriction enzymes.

4.3.4 Expression of the endo- β -1,4-glucanase gene in E. coli HB101(pHZ100) and HB101(pHZ120) cells. The endoglucanase activities of E. coli HB101 containing pHZ100, pHZ120, pHZ101 and pHZ117 are shown in Table 4.1. pHZ101 was an EcoRI endonuclease deletion derivative of pHZ100. The deletion resulted in the loss of a fragment containing the vector λ promoter region and 1.7 kb of the insert DNA without affecting expression of the gene, indicating that the gene may be expressed from its own promoter in E. coli. Removal of 350 bp from the opposite end of the insert (pHZ110) resulted in the complete loss of endoglucanase

activity showing that the gene was located at this end of the insert. The 1.74 kb HaeIII - PstI fragment was the smallest fragment that could be subcloned and still retain endoglucanase activity (pHZ117). E. coli cells transformed with pHZ120 gave consistently lower endoglucanase activities than E. coli cells transformed with pHZ100.

Table 4.1. Endoglucanase activities of E. coli HB101 harbouring various plasmids.

Plasmid	Host vector	Insert size	Sp. act ^a
pBR322	----	----	ND ^b
pHZ100	pEcoR251	4.9 kb	0.336
pHZ120	pEcoR251	7.5 kb	0.157
pHZ101 ^c	pEcoR251	2.6 kb	0.384
pHZ110	pBR322	3.4 kb	ND ^b
pHZ117 ^d	pUC19	1.7 kb	0.387

a Specific activity expressed as units/mg protein

b ND - not detected

c this deletion also removed 590 bp of vector DNA including the λ promoter.

d the insert includes 180 bp of the EcoRI gene from pEcoR251.

To determine whether the expression of the endoglucanase gene was subject to metabolite regulation, E. coli HB101 cells harbouring either pHZ100 or pHZ120 were grown in LB medium containing either glucose (1%, w/v), cellobiose (1%, w/v) or CMC (0.1%, w/v). The activities obtained for cells harbouring both the plasmids are shown in Table 4.2. The addition of either cellobiose or CMC to the culture medium did not have a marked effect on the expression of endoglucanase activity. However addition of glucose to the culture medium had an inhibitory effect on the expression of the endoglucanase gene in cells harbouring both plasmids.

Table 4.2. Endoglucanase activity of E. coli HB101 cells harbouring either pHZ100 or pHZ120 and grown in the presence of various carbohydrates.

Plasmid	Medium	Sp. act ^a
pHZ100	LB	0.336
	LB + glucose	0.127
	LB + cellobiose	0.356
	LB + CMC	0.374
pHZ120	LB	0.157
	LB + glucose	0.067
	LB + cellobiose	0.159
	LB + CMC	0.156

a Specific activity expressed as units/mg protein

E. coli HB101 harbouring either pHZ100 or pHZ120 produced acid and gas in peptone water containing cellobiose, and grew on minimal medium with cellobiose as the sole carbon source, whereas E. coli HB101 or E. coli HB101(pBR322) did not. E. coli HB101(pHZ100) did not produce acid or gas in peptone water lacking cellobiose. The cellobiose activity exhibited by cells harbouring these plasmids was not examined further and will be part of a future investigation.

4.3.5 Localisation of the endoglucanase activity in E. coli HB101(pHZ100). Localisation of the endoglucanase activity in E. coli cell extracts indicated that periplasmic samples contained 75% of the total endoglucanase activity. Control experiments showed that 98.7% of the total β -galactosidase activity was located in the cytoplasm and 33% of the total alkaline phosphatase activity was present in the periplasmic extracts. Supernatant samples showed no detectable endoglucanase activity.

4.3.6 Partial purification and characterisation of the endo- β -1,4-glucanase from E. coli HB101(pHZ100). The results of a typical enzyme purification experiment are summarised in Table 4.3 and the level of purification achieved after column chromatography shown in the PAGE gel in Fig.4.3. The endoglucanase was purified approximately 20-fold in specific activity with a recovery yield of 31.5%.

Table 4.3. Partial purification of the endoglucanase from E. coli HB101(pHZ100).

Fraction	Volume (ml)	tot. prot. (mg)	Sp. act. ^a	Yield %
crude CFE	33.0	929	0.49	100
after DE52	7.5	23.3	8.51	43.4
after CM52	3.2	15.0	9.59	31.5

a Specific activity expressed as units/mg protein

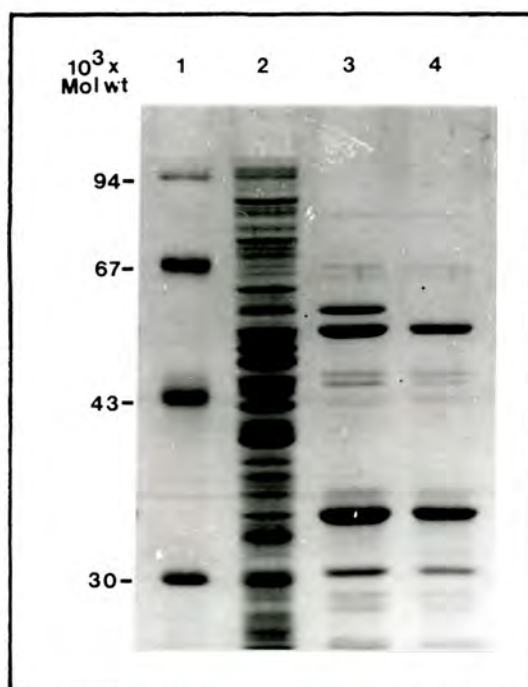


Fig. 4.3. SDS-PAGE analysis of the endoglucanase partially purified from E. coli HB101(pHZ100). Lane 1, protein standards; lane 2, crude CFE; lane 3, after DE52 column chromatography; lane 4, after CM52 column chromatography. Protein standards were phosphorylase b (94 000), albumin (67 000), ovalbumin (43 000) and carbonic anhydrase (30 000).

Endoglucanase activity was not recoverable after SDS-PAGE according to the method of Beguin (1983) using an activity gel overlay, as the enzyme was presumably irreversibly denatured by SDS.

The effects of pH on the endoglucanase produced by E. coli HB101(pHZ100) were determined. The enzyme was assayed in different citrate (pH 3.0 - 6.5) and phosphate (pH 7.0 - 8.5) buffers (Fig. 4.4). The enzyme showed maximal activity between pH 5.0 and pH 6.5 with a peak at pH 6.0. There was a sharp decrease in activity below pH 5.0 with 38% of maximal activity at pH 4.0 and no activity detectable at pH 3.5 under standard assay conditions. There was a similar decrease in activity in buffers above pH 6.5 with 13% of maximal activity present at pH 8.5. Enzyme stability at low (pH 3.5) and high (pH 8.5) pH was assessed by holding the enzyme at the respective pH for 24 h and then assaying for residual activity at pH 6.0. The endoglucanase was unstable at both pH levels with <5% of activity remaining after 24 h at pH 3.5 and <8% of activity remaining after the same time at pH 8.5.

The effects of temperature on the endoglucanase activity were examined by performing the assay at various temperatures (Fig. 4.5). The enzyme showed a broad range of near maximal activity between 46 and 54°C where activities were >95% of the optimum activity obtained at 50°C. The stability of the enzyme was assessed by holding the enzyme for 60 min at various temperatures and then assaying the residual activity at 50°C (Fig. 4.5). The enzyme was

unstable at temperatures above 50°C with 55% of activity remaining after 60 min at 55°C and <0.5% activity remaining after 60 min at 60°C.

The activity of the endoglucanase against acid-swollen cellulose, filter paper and Avicel was determined. Although low levels of activity (<0.05 μ /mg protein) against filter paper and acid-swollen cellulose was detected initially, no further release of glucose equivalents occurred after 2 h at either 40 or 50°C. No hydrolysis of Avicel was detected over 24 h at 40 or 50°C.

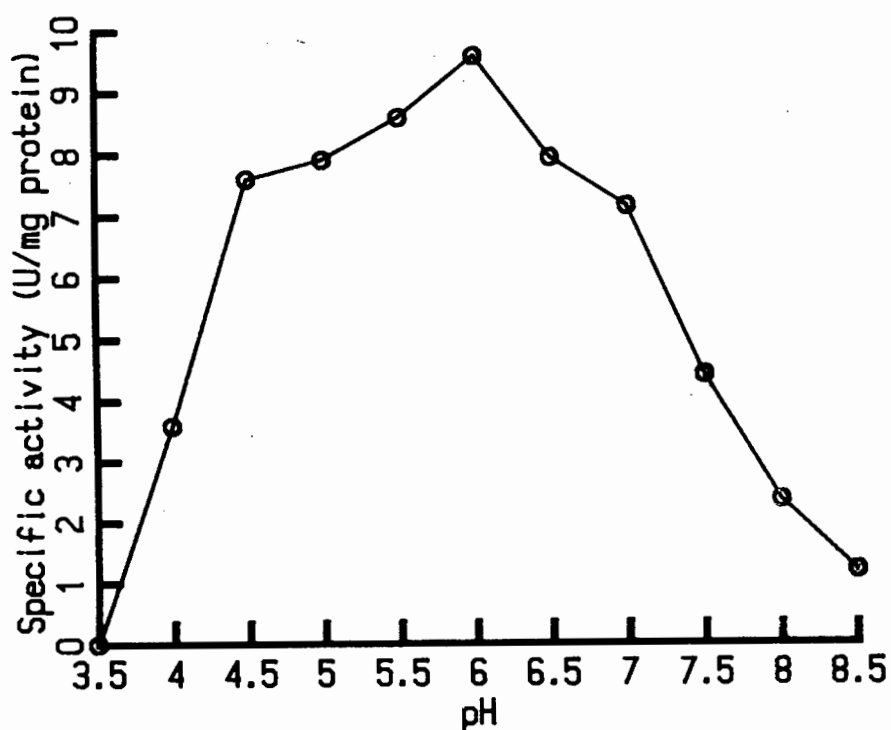


Fig. 4.4. pH activity profile of the endoglucanase partially purified from *E. coli* HB101(pHZ100). Activity was assayed using substrate (1 ml) in different pH buffers and enzyme solutions diluted in buffer at the specific pH.

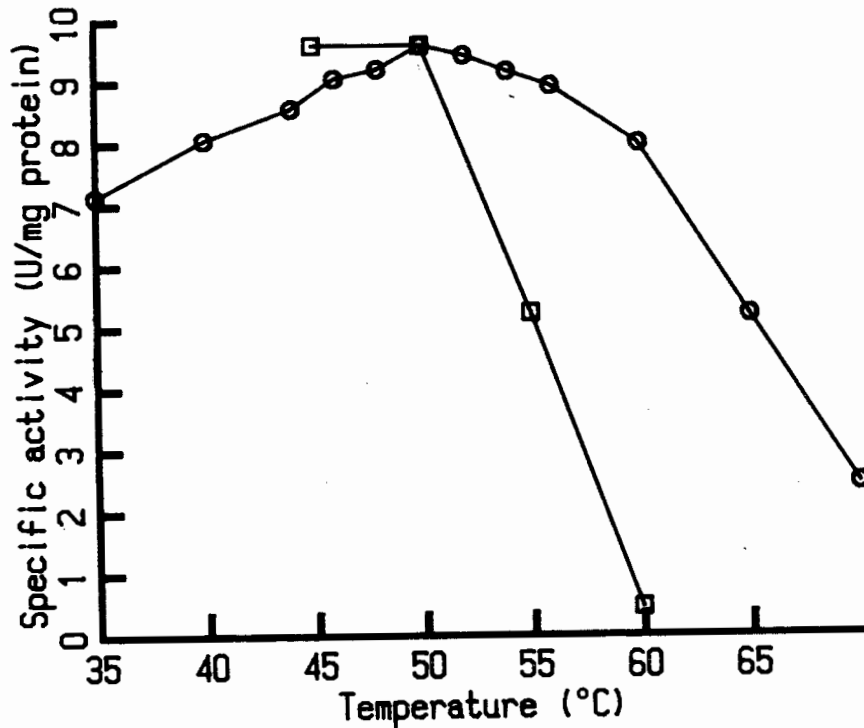


Fig. 4.5. Temperature activity (○) and stability (□) profiles of the endoglucanase partially purified from *E. coli* HB101(pHZ100). For temperature activity measurements, substrate (1 ml) was prewarmed to various temperatures and then enzyme (50 μ l) added and the activity assayed at the specific temperature. For the stability profile, enzyme samples were held at various temperatures for 60 min, and then residual activity was assayed at 50°C.

4.3.7 Activity gel overlay of non-denaturing PAGE gels.

The partially purified endoglucanase, when electrophoresed through a non-denaturing PAGE gel, resolved into two bands of activity visible on the overlay gel (Fig. 4.6, lane 1). The bands were not due to the cloning of two distinct endoglucanases on the same DNA fragment since the endoglucanase partially purified from cells harbouring pHZ117, a subclone of pHZ100, produced the same result (Fig. 4.6, lane 2). Sequencing of the fragment in pHZ117 (Chapter 5) showed the presence of only one open reading frame.



Fig. 4.6. Zymogram of the cloned *C. acetobutylicum* endoglucanase. Samples of the endoglucanase partially purified from *E. coli* cells harbouring either pHZ100 or pHZ117 were electrophoresed under non-denaturing conditions. The PAGE gel was then overlaid with an agar replica containing CMC, incubated as described in the methods, and bands of activity visualised by Congo red staining. Lane 1, endoglucanase from *E. coli* HB101(pHZ100) cells; lane 2, endoglucanase from *E. coli* HB101(pHZ117) cells.

4.4 Discussion.

Allcock and Woods (1981) reported that C. acetobutylicum P270 (a subculture of P262) showed inducible extracellular CMCase (endoglucanase) activity which was not able to degrade more complex cellulosic substrates such as filter paper, acid-swollen cellulose or Avicel. In comparison with other bacteria and fungi the levels of endoglucanase obtained for C. acetobutylicum P270 were low. The optimum conditions for activity for the CMCase reported by Allcock and Woods (1981) was pH 4.6 at 37°C, whereas the cloned endoglucanase showed optimal activity at pH 6.0 and 50°C. It appears therefore that the endoglucanase cloned from C. acetobutylicum P262 could be a different enzyme from that studied by Allcock and Woods (1981) from C. acetobutylicum P270. However it is possible that the same enzyme has been cloned but that it is modified in some way in E. coli resulting in the enzyme exhibiting different pH and temperature optima in the two systems. Alternatively the C. acetobutylicum P270 endoglucanase could undergo posttranslational and/or postsecretional modifications (e.g. glycosylation or limited proteolysis), which do not occur in the E. coli host system, that could alter its characteristics (see Chapter 1). The presence of multiple non-homologous endoglucanase genes in a single organism is not uncommon and have been reported for (among others) C. thermocellum (Cornet et al., 1983; Millet et al., 1985) and C. fimi (Gilkes et al., 1984a). Two endoglucanases showing partial homology have been cloned from an alkalophilic Bacillus species (strain N-4) (Sashihara

et al., 1984). The characteristics of the endoglucanases from C. acetobutylicum P270 (Allcock and Woods, 1981) and the cloned endoglucanase from C. acetobutylicum P262 also differed from those reported for C. acetobutylicum ATCC 824 and NRRL B527 (Lee et al., 1985a).

The levels of endoglucanase expression reported for C. acetobutylicum P270 were very low and production of the enzyme was induced by an unidentified molecule present in the molasses medium (Allcock and Woods, 1981). Acid-swollen cellulose, Avicel, CMC, sucrose, glucose, cellobiose and xylose did not induce endoglucanase activity in CBM medium. Attempts to induce the endoglucanase activity in C. acetobutylicum P262 and to purify the enzyme from the culture medium (5 l) gave inconsistent results and the levels of activity were at the limit of detection. Antibodies raised against the partially purified cloned endoglucanase did not elicit cross-reactivity with the C. acetobutylicum P262 preparations presumably due to the very low levels of activity present in these preparations.

The origin of the endoglucanase activity encoded by the inserts cloned in pHZ100 and pHZ120 was however confirmed by hybridisation, and the E. coli host alone, or carrying other pEcoR251 recombinants did not produce zones of hydrolysis using the Congo red staining method. Production of zones using this method is very specific for endoglucanases and zone formation results from cleavage of the CMC into fragments, smaller than cellobiose, to which Congo red does not bind (Wood, 1980). Gilkes et al. (1984a) showed that

E. coli, carrying a cellulase gene cloned from C. fimi, produced zones of hydrolysis on CMC agar, but did not produce haloes by the Congo red method. On the basis of the specificity of the Congo red technique the authors suggested that the cloned enzyme was an exoglucanase rather than an endoglucanase. A qualitative viscosimetric assay also demonstrated the presence of endoglucanase activity produced by E. coli cells carrying pHZ100 or pHZ120.

Unlike the xylanase expressed in E. coli HB101(pHZ300) (Chapter 3), 75% of the endoglucanase activity was detected in the periplasmic space of the E. coli host, indicative of an active excretion mechanism for the cloned product across the cytoplasmic membrane. Various other endoglucanases, cloned and expressed in E. coli, have shown varying degrees of excretion both to the periplasmic space and to the external medium. Cornet et al. (1983) cloned two endoglucanase genes (celA and celB) from C. thermocellum into E. coli and found that the respective enzyme activities were about equally distributed between the cytoplasm and periplasm of the E. coli host, whereas the two endoglucanases cloned in E. coli from an alkalophilic Bacillus species showed a distribution similar to that found for the cloned C. acetobutylicum endoglucanase. Two endoglucanases cloned in E. coli from Thermomonospora showed 40% cytoplasmic and 30% periplasmic activity while the remaining 30% was excreted to the external medium (Collmer and Wilson, 1983). The authors suggested that E. coli must have a mechanism for secreting proteins into the medium,

though E. coli cells do not use the mechanism for many proteins.

Expression of the C. acetobutylicum endoglucanase by E. coli cells harbouring pHZ120 was consistently lower than the levels of activity obtained with cells harbouring pHZ100. The only apparent physical difference between the two clones is that pHZ120 contains a larger C. acetobutylicum DNA fragment, extended downstream of the endoglucanase gene (Fig. 4.1). This downstream region may affect activity of the gene. Removal of a region downstream of the glutamine synthetase gene cloned from C. acetobutylicum P262 was shown to affect gene expression (Janssen et al., 1988). Addition of glucose to the culture medium of E. coli cells harbouring either plasmid resulted in a degree of repression, whereas addition of either cellobiose or CMC did not adversely affect expression. Taylor et al. (1987) found a similar glucose repression effect with a B. succinogenes endoglucanase gene cloned in E. coli.

The C. acetobutylicum endoglucanase gene appeared to be expressed from its own promoter in E. coli cells since deletion of the λ promoter of pEcoR251 (deletion plasmid pHZ101) did not affect expression of the gene. Furthermore cloning the gene in pUC19 (pHZ117) in reverse polarity to the lacZ promoter in pUC19 (see Chapter 5) did not affect the expression of the endoglucanase gene.

The endoglucanase produced by E. coli(pHZ100) cells was SDS sensitive since attempts to recover activity after SDS-PAGE

using the activity gel overlay (zymogram) technique described by Beguin (1983) were unsuccessful. Beguin (1983) showed that the activity of certain C. thermocellum endoglucanases could be recovered after SDS-PAGE and the bands of activity visualised using the reported technique. Other endoglucanases have been shown to be SDS sensitive (Gilkes et al., 1984a; Beguin et al., 1983) and activity bands were visualised on non-denaturing gels using the same technique. The partially purified cloned C. acetobutylicum endoglucanase was electrophoresed under non-denaturing conditions and separated into two bands of activity (samples prepared from cells harbouring either pHZ100 or pHZ117). The coding capacity of the 1.74 kb fragment in pHZ117 was too small to code for two endoglucanases (calculated endoglucanase M_r is 49 000; Chapter 5). The two bands are probably the product of a single gene, one being the primary transcript and the other a processed form of the protein. Similar results were reported for the cloned celB gene product of C. thermocellum (Beguin et al., 1983) and for the expression of the Staphylococcus aureus A protein (Duggleby and Jones, 1983) in E. coli cells. Robson and Chambliss (1987) sequenced a B. subtilis endoglucanase gene and showed that the DNA had the coding capacity for an 85 000 M_r protein but the native protein purified from B. subtilis had an M_r of only 35 200. When the gene was expressed in E. coli both forms of the protein were present.

Chapter 5

Structure of the cloned endoglucanase gene from Clostridium acetobutylicum P262 and comparison with other endoglucanase genes.

5.0 Summary.

The nucleotide sequence of a cloned C. acetobutylicum endo- β -1,4-glucanase gene was determined. The upstream region of the endoglucanase gene contained two putative extended promoter consensus sequences characteristic of Gram-positive bacteria. The putative ribosome binding site consisted of a TTG initiation codon and a strong Shine-Dalgarno complementarity containing five G-C bp situated seven bp upstream from the initiation codon. The complete amino acid sequence (448 residues) of the C. acetobutylicum endoglucanase was deduced, and comparisons were made with reported amino acid sequences of endoglucanases from other organisms. A high degree of homology was demonstrated between the C. acetobutylicum endoglucanase and enzymes from Bacillus strains. There was no extensive homology between the C. acetobutylicum endoglucanase and enzymes from C. thermocellum.

5.1 Introduction.

Cellulases (1,4- β -D-glucan glucanohydrolase: EC 3.2.1.4) generally occur as multienzyme complexes in "true" cellulolytic organisms. However some components of the complex appear to occur in other bacteria which are unable to hydrolyse crystalline cellulose including C. acetobutylicum, B. subtilis and E. chrysanthemi (see Chapter 1).

The cloning of endo- β -1,4-glucanase genes from a variety of different organisms has facilitated the identification of components of the cellulase complex and has made possible a more detailed study of the characteristics and some aspects of the regulation and transport of these enzymes. Recently the nucleotide sequences of a number of endo- β -1,4-glucanases genes from different fungi, Gram-positive bacteria (C. thermocellum, C. fimi, B. subtilis and two alkalophilic Bacillus species), and a Gram-negative bacterium (E. chrysanthemi) have been reported (Table 1.4). Analysis of endoglucanase nucleotide sequences will assist in identifying functionally important regions of these enzymes and in elucidating evolutionary relationships between members of the cellulase gene family.

Very limited deduced protein sequence homology has been demonstrated between the sequenced cellulase genes from the "true" cellulolytic organisms although the genes appear to show a common structure (Knowles et al., 1987). On the other hand, endoglucanases from non-cellulolytic organisms

appear to show a much greater degree of protein sequence homology. Both the C. acetobutylicum endoglucanase gene sequence reported here and that of an E. chrysanthemi endoglucanase gene (Guiseppi et al., 1988) show extensive protein homology with endoglucanases from B. subtilis and alkalophilic Bacillus species.

This chapter describes the sequencing of the endoglucanase from C. acetobutylicum cloned in pHZ100 and the comparison of the of the nucleotide sequence and the deduced amino acid sequence with other reported endoglucanase sequences.

5.2 Materials and methods

5.2.1 **Bacteria and plasmids.** Plasmids pHZ101 and pHZ120 were used as primary sources of DNA. The plasmid vectors pUC18 and pUC19 were used for all subcloning for sequencing except for one fragment (SalI - PstI) which was subcloned into M13mp18 and M13mp19 vectors (Messing, 1983). E. coli LK111 (lacI⁻, lacZΔM15, lacY⁺, Zabeau and Stanley, 1982) was used as a host for pUC plasmids. E. coli LK111 is a derivative of E. coli K514 which in turn is a $r_k^- m_k^+$ derivative of E. coli C600 (Wood, 1966). E. coli JM103 (Δlac-pro, thi, strA, supE, endA, sbcB15, hsdR4, (F', traD36, proAB, lacI^{qz} M15); Messing et al., 1981) was used as a host for the M13 vectors.

5.2.2 **Media and buffers.** All media and buffers not described in the text are listed in Appendix B.

5.2.3 **Growth conditions.** E. coli LK111 harbouring pUC derived recombinant plasmids were plated on 2 x YT agar containing IPTG, 5-bromo-4-chloro-3-indolyl-β-galactoside (X-gal) and Ap (Appendix B). E. coli JM103 was maintained on minimal A medium (Appendix B). Both strains were grown in 2 x YT medium at 37°C (Appendix B) for the preparation of plasmid (pUC vectors) or replicative form (RF) (M13 vectors) DNA.

5.2.4 **Sequencing strategy.** Restriction mapping and subcloning (Chapter 4) showed that the endoglucanase gene was located on a HaeIII - PstI fragment which was subcloned

from pHZ101 into pUC19 (pHZ117, Fig. 4.1). The 0.74 kb SalI - PstI fragment was subcloned into M13mp18 and sequenced in one orientation using a commercially available forward primer (BRL) and one unique primer (a, Fig. 5.1). The fragment was subcloned in the other orientation in M13mp19 and sequenced from the SalI site over the HindIII site. Subcloning of a HindIII - PstI fragment into pUC19 enabled sequencing (from the HindIII site) over the Sau3A site at the insert-vector junction (Fig. 5.1). The 1.2 kb HaeIII - HindIII fragment was subcloned into pUC19 previously linearised with SmaI and HindIII endonucleases. The resultant recombinant plasmid contained the fragment between HindIII and EcoRI restriction sites derived from the pUC19 polylinker. This plasmid was linearised with either HindIII or EcoRI endonucleases and used to generate two ordered sets of BAL-31 exonuclease deletions of opposite polarity using a strategy similar to that described for fragments subcloned in M13 vectors (Poncz *et al.*, 1982; Dale *et al.*, 1985).

After the sequence of the functional HaeIII - PstI fragment had been determined, it was found that it did not contain an in-frame stop codon within the cloned fragment (transcription presumably terminated at the first in-frame stop codon within the vector DNA). This finding necessitated the partial sequencing of an overlapping PvuII fragment subcloned from pHZ120 (pHZ121, Fig. 4.1 and b in Fig 5.1).

Two regions of the gene, from position -357 to -283 and from position 1276 to 1441 (refer to the nucleotide sequence

shown in Fig. 5.2), were sequenced 3 times in one direction as shown in Fig 5.1 (b and c). The remainder of the sequence was determined using overlapping subclones of opposite polarity.

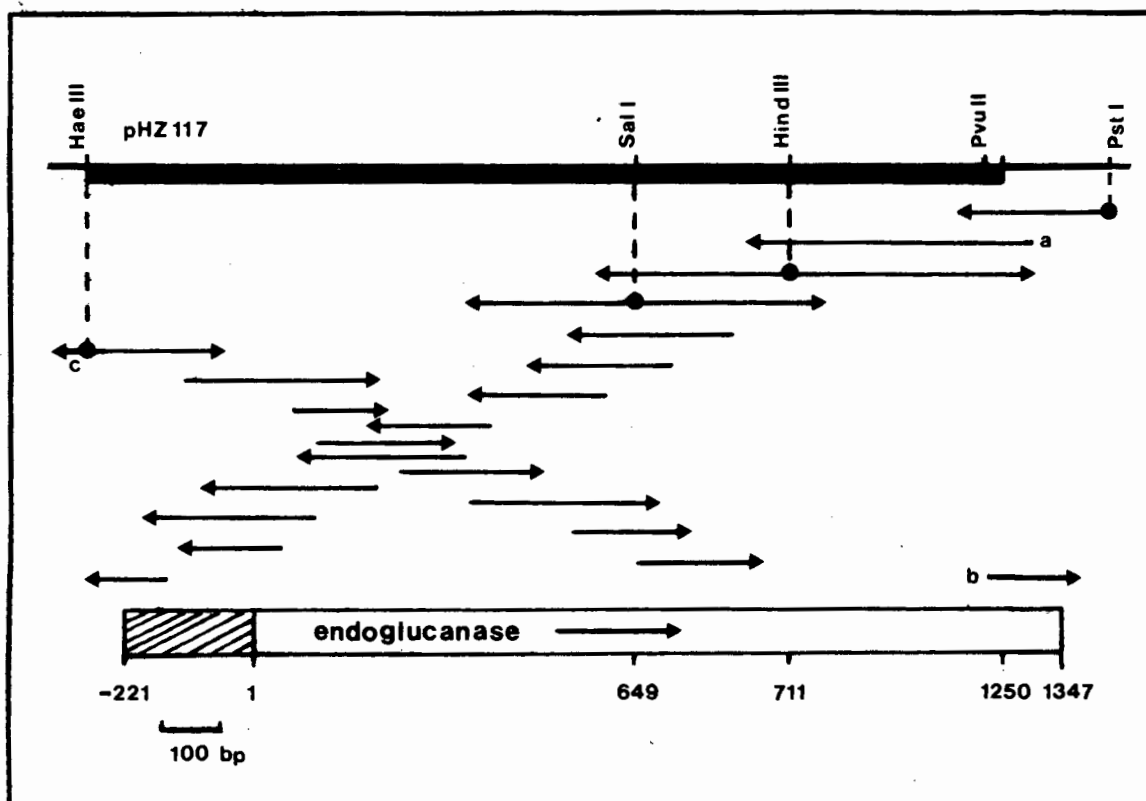


Fig. 5.1. Nucleotide sequencing strategy for the *C. acetobutylicum* endoglucanase gene. The thick line represents *C. acetobutylicum* DNA, the thin line vector DNA. The arrows depict the direction and extent of sequencing of subclones in pUC vectors. Those arrows not connecting to the restriction map with dashed lines were subclones generated by BAL-31 exonuclease deletion. (a) This Sequence was determined with the use of a unique primer. (b) Unidirectional sequencing of the PvuII fragment subcloned from pHZ120 (see Fig. 4.1). The box beneath the arrows represents the endoglucanase gene, the arrow within it, the direction of transcription. The hatched area represents the putative promoter region. The numbers below the box refer to bp numbers in the nucleotide sequence depicted in Fig. 5.2. Where the numbers coincide with restriction sites, the base number given is the first base (5') of the recognition sequence.

5.2.5 **BAL-31 exonuclease shortening of DNA cloned into pUC19.** A BAL-31 exonuclease reaction was prepared by combining linearised plasmid (10 μ g), 5 X BAL-31 reaction buffer (Appendix B) and water to a final volume of 140 μ l and equilibrating at 37°C (5 min). The reaction was initiated by the addition of BAL-31 exonuclease (Gray *et al.*, 1975) (10 μ l, 5 units; New England Biolabs.) and 25 μ l samples were removed at 1 min intervals to microfuge tubes containing 75 μ l TE buffer and 10 μ l TE buffer-saturated phenol. Each aliquot was further extracted with chloroform-isoamyl alcohol (24:1) (100 μ l) and the DNA was precipitated from the aqueous phase as described in Appendix C. The DNA pellet was resuspended in restriction buffer and digested with the appropriate restriction enzyme to release the shortened fragments. The samples were analysed by agarose gel electrophoresis and the results used to calibrate the BAL-31 exonuclease. Suitable digestion time intervals were selected that produced fragments of the desired size. The experiment was repeated at the required time intervals and the BAL-31 exonuclease shortened DNA used in ligations with pUC19 prepared as described below. It was found unnecessary to fill-in the BAL-31 generated ends as a sufficiently large fraction of the DNA molecules had fully base-paired ends (Gray *et al.*, 1975).

5.2.6 **Preparation of pUC19 vector and cloning of BAL-31 exonuclease shortened fragments.** pUC19 was digested with SmaI and either EcoRI or HindIII endonucleases (5 μ g for each combination) and electrophoresed on low melting point (LMP) agarose to remove the short polylinker fragment. The

DNA was viewed under long wave (310 nm) UV light after staining the gel (Appendix C), the pUC19 band excised, the gel slice melted, and used directly in ligations (Struhl, 1985; Appendix C) with BAL-31 shortened fragments. The ligations were diluted 5-fold with 0.1 M CaCl and used to transform competent E. coli LK111 (Appendix C). Recombinant plasmids isolated from Ap^r white colonies on X-gal plates and containing shortened fragments (confirmed by miniprep, restriction endonuclease analysis and agarose gel electrophoresis) were purified by the maxiprep method (Appendix C) prior to sequencing.

5.2.7 Transformation of E. coli JM103 with M13 vectors.

Competent E. coli JM103 cells were prepared as described in Appendix C. An E. coli JM103 culture was inoculated from an overnight culture such that these cells were in exponential phase to serve as lawn host cells when plating with the transformed cells. E. coli JM103 was transformed by heat-shocking as described in Appendix C. After the heat shock step, the transformed cells (in 100 μ l) were combined with 20 μ l IPTG (100 mM in distilled water), 40 μ l X-gal (2%, w/v in dimethylformamide), 200 μ l of exponential phase cells and 3 ml of H top agar (Appendix B) and plated on H agar plates (Appendix B). After the top agar had solidified, the plates were incubated overnight at 37°C. Exponential phase E. coli JM103 cells (5 ml culture) were infected with M13 phage originating from white plaques and incubated for 5 h at 37°C. The cells were collected by centrifugation and M13 RF DNA was prepared by the miniprep method (Appendix C). The presence of cloned fragments was confirmed as described

above (5.2.6). The culture medium containing the recombinant M13 phage was retained and used to prepare single-stranded (SS) DNA templates.

5.2.8 Preparation of template DNA. SS M13mp DNA was isolated and prepared to serve as the SS template for sequencing reactions as described in the BRL manual supplied with the sequencing kit. SS template derived from pUC recombinant plasmids was prepared as follows: pUC plasmid DNA (2 - 4 μg in 20 μl H_2O) was denatured by adding 2 μl of NaOH (2 M) and leaving at room temperature for 5 min. The reaction was neutralised by adding 5 μl of sodium acetate (3 M), diluted with 25 μl H_2O and the DNA precipitated by adding 150 μl cold (-70°C) ethanol and holding at -70°C for 5 min. The DNA was collected by centrifugation (30 min, 4°C) in a microfuge. The DNA pellet was washed with 70% cold (-20°C) ethanol, air dried and resuspended in 8 μl of water.

5.2.9 Primers and primer annealing reactions. Primers for reverse and forward sequencing for both pUC and M13 vectors were obtained from BRL. One unique primer was a gift from D. Botes (Dept. of Biochemistry, University of Cape Town, SA). The primer had the sequence 5'-GTGGCTCTCAGAGAGCA-3' and was synthesised in an Applied Biosystems 381A DNA synthesiser (Forster City, CA, USA) using their proprietary reagents and solvents. Primers for SS pUC DNA were annealed to pUC DNA by adding 1.4 μl 10 X DNA polymerase I buffer (Appendix B) and 4.5 μl primer (4 ng/ μl) to the 8 μl of SS DNA prepared above, and incubating for 15 min at 40°C . The

DNA was used directly for sequencing or stored at -70°C . M13mp priming reactions were performed according to the BRL manual using reagents supplied with the sequencing kit.

5.2.10 Sequencing reactions. DNA sequences were determined using the dideoxy chain termination procedure (Sanger *et al.*, 1977). Reagents for the DNA sequencing reactions were obtained in kit form (BRL kit M13 C/SS I) and used according to the manufacturer's specifications. The DNA chain was radiolabelled with $[\alpha\text{-}^{32}\text{P}]\text{dCTP}$ (3000 Ci/mmol) obtained from Amersham Int., UK).

5.2.11 Polyacrylamide gel electrophoresis. The sequencing reactions were analysed on denaturing polyacrylamide gels. Polyacrylamide wedge (0.4 mm spacers at top, 0.8 mm spacers at bottom) gels (8%) were cast in a BRL apparatus (BRL model SO) (34 X 40 cm). A sharktooth comb was used to form the wells (BRL 1045 SC, 24 teeth, 0.4 mm thick, 6 mm point-to-point). Although this comb had the capacity for 24 samples (6 templates), only 4 templates were routinely loaded in the central wells of the gel, to improve the resolution and minimise smiling effects. The composition of the gels and the running conditions were as described in the BRL manual supplied with the kit. After electrophoresis, gels were dried onto filter paper (Whatman No. 3) using a Dual Temperature Slab Gel Dryer (Model 1125B) (Hoefer Scientific Instruments, San Francisco).

5.2.12 Autoradiography. Gels containing ^{32}P -labelled DNA were placed under Cronex 4 film (Rontgen autoradiographic

film) in an X-Ray cassette and exposed for 5 - 24 h. The autoradiographs were developed using Kodak GBX X-Ray developer and fixer.

5.2.13 DNA sequence analysis. Manipulation of DNA sequences was done using an IBM microcomputer and the MicroGenie^R (Beckman) software package developed by Queen and Korn (1984). Single letter codes were used for amino acids. A table of three- and one-letter codes for amino acids is given in Appendix F.

5.3 Results.

5.3.1 Nucleotide sequence. The nucleotide sequence of the functional 1.74 kb HaeIII-PstI fragment, subcloned from pHZ101, was determined as described in the methods (Fig. 5.1). Although this fragment induced the synthesis of an enzymatically active protein in E. coli it did not contain an inframe translational stop codon and the protein was presumably terminated by a stop codon within the vector DNA. The extent of the open reading frame (ORF) in this fragment is shown in Fig. 5.2. This indicated that the original fragment of chromosomal DNA which was cloned in the plasmid pHZ100 did not contain the entire ORF of the endoglucanase gene. To determine the nucleotide sequence coding for the C-terminal region of the gene an additional PvuII fragment was subcloned from pHZ120 as described in the sequencing strategy (5.2.4). This fragment, extending downstream of the PvuII restriction site in pHZ100, contained the remainder of the ORF and the translational stop codon, and had a 25 bp overlap with the ORF from pHZ100. The ORF of the entire endoglucanase gene encoded a protein of 448 amino acid residues (Fig. 5.2).

The nucleotide sequence of the C. acetobutylicum endoglucanase gene did not contain an inframe ATG start codon for the initiation of translation of the ORF. Instead the putative ribosome binding site (RBS) resembled that reported for the β -lactamase gene from S. aureus (McLaughlan et al., 1981) and the putative RBS reported for the chloramphenicol acetyltransferase gene (cat-86) of B. pumilus (Harwood et al., 1983). The putative RBS consisted of a TTG initiation codon and a strong Shine-Dalgarno complementarity containing five G-C bp situated seven bp upstream of the initiation codon (Fig. 5.3). The UUG initiation codon of the S. aureus β -lactamase mRNA initiated protein synthesis with a methionine residue (McLaughlan et al., 1981).

<u>C. acetobutylicum</u>	5'	<u>A G G G G G</u>	T A T T A A C	<u>T T G</u>	3'
<u>B. pumilus</u>	5'	<u>A G G A G G</u>	A G A T A A A A	<u>T T G</u>	3'
<u>S. aureus</u>	5'	<u>G G A G G G</u>	T T T A T T	<u>T T G</u>	3'

Fig. 5.3 Putative ribosome binding sites of the endoglucanase gene from C. acetobutylicum and the B. pumilus cat-86 gene compared to that of the S. aureus β -lactamase gene. The Shine-Dalgarno region and start codon are overlined.

The entire ORF of the C. acetobutylicum endo- β -1,4-glucanase gene encoded a protein of 448 amino acid residues with a calculated M_r of 49,354. The (G+C) content of the nucleotides within the ORF (32.8%) was considerably higher than the (G+C) content of the nucleotides in the regulatory region upstream of the gene (19.9%). The average (G+C)

ratio for the whole genome of C. acetobutylicum is approximately 28% (Cummins and Johnson, 1971). As might be expected, the codon usage within the endoglucanase gene exhibited a strong bias towards the use of codons in which A and U predominated (Table 5.1).

Table 5.1. Codon usage and frequencies in the C. acetobutylicum endoglucanase gene. Stop codons are represented by a - symbol.

TTT F 17 3.8	TCT S 14 3.1	TAT Y 16 3.6	TGT C 2 0.4
TTC F 1 0.2	TCC S 1 0.2	TAC Y 6 1.3	TGC C 2 0.4
TTA L 12 2.7	TCA S 19 4.2	TAA - 1 0.2	TGA - 0 0.0
TTG L 3 0.7	TCG S 1 0.2	TAG - 0 0.0	TGG W 10 2.2
CTT L 4 0.9	CCT P 9 2.0	CAT H 5 1.1	CGT R 2 0.4
CTC L 2 0.4	CCC P 0 0.0	CAC H 1 0.2	CGC R 0 0.0
CTA L 0 0.0	CCA P 6 0.3	CAA Q 11 2.5	CGA R 0 0.0
CTG L 1 0.2	CCG P 0 0.0	CAG Q 6 1.3	CGG R 0 0.0
ATT I 10 2.2	ACT T 26 5.8	AAT N 27 6.0	AGT S 12 2.7
ATC I 2 0.4	ACC T 2 0.4	AAC N 4 0.9	AGC S 4 0.5
ATA I 23 5.1	ACA T 21 4.7	AAA K 24 5.4	AGA R 3 0.7
ATG M 11 2.5	ACG T 0 0.0	AAG K 8 1.8	AGG R 0 0.0
GTT V 12 2.7	GCT A 21 4.7	GAT D 20 4.5	GGT G 8 1.8
GTC V 2 0.5	GCC A 2 0.4	GAC D 5 1.1	GGC G 6 1.3
GTA V 10 2.2	GCA A 6 1.3	GAA E 11 2.5	GGA G 15 3.3
GTG V 0 0.0	GCG A 0 0.0	GAG E 2 0.4	GGG G 0 0.0

5.3.2 Putative promoter sequences. It was shown by deletion and subcloning (Chapter 4) that the cloned endoglucanase gene from C. acetobutylicum was expressed from its own promoter in E. coli. The region upstream of the ORF between nucleotides -221 and -175, contained a putative promoter sequence which consisted of a TTGTATT -35 region and a TACAAT -10 region separated by 16 nucleotides (Fig. 5.4). This promoter consensus sequence closely resembled the extended consensus sequence reported to be a feature of the σ^{43} RNA polymerase recognition site of

B. subtilis and other Gram-positive bacteria, including an AT rich upstream region (Graves and Rabinowitz, 1986). A second putative extended promoter consensus sequence was situated immediately upstream of the ribosome binding site between nucleotides -66 and -19 (Fig. 5.4). In view of the relative positions of the two putative promoters in relation to the putative ribosome binding site, it seems likely that the distal promoter consensus sequence is the functional promoter for mRNA transcription. The 5' upstream region of the cloned C. acetobutylicum endoglucanase gene did not contain any promoter consensus elements which have been reported as recognition sites for other Bacillus holoenzymes σ^{37} , σ^{32} , σ^{29} , or σ^{28} (Doi and Wang, 1986).

Gram-positive consensus	T A A A A A A	T T G A C A	a A	a T T G	T A T A A T A A t a t
EG A P1	T T T T A T A T A T A T	T T G T A T	T A A A T T A	- - - A C T T T A A T G T	T A C A A T G T T C T T A
EG A P2	A A T A A A A A A A A A	T T A T T A	A T G T A A A A A T A T A C T A A G T A	T A G A A T	A A T A G G G
GLN A P1	T A A A A A A A T C A C	T T G A T T	T C T T A A A A	- - - A A A A A G G G A A G	T A T A A T T T A G T T A
GLN A P2	T G A G C A A A G C G C G	T T C G A T	A G A A G T T	- - - T A T A C T T G T C	T A T T G T G C G C T T T
		-35			-10

Fig. 5.4. Gram-positive consensus promoter sequences. EG A P1 and P2 are the two putative endoglucanase promoters; GLN A P1 and P2 are promoters identified by lacZ fusions for the C. acetobutylicum glutamine synthetase gene (Janssen et al., 1988; Janssen, pers. comm.). Highly conserved bases (>70%) are indicated by overlined capital letters; >50% conservation are capital letters and >41% conservation are lower case letters (after Graves and Rabinowitz, 1986).

5.3.3 Putative signal sequence. It was shown that the endoglucanase activity occurred predominantly in the periplasm of E. coli cells containing the cloned gene (Chapter 4). In common with other genes which code for secreted proteins, the endoglucanase gene from C. acetobutylicum contained a putative signal sequence at the N-terminal end of the protein (Fig. 5.2 and Fig. 5.5) which was typical of the signal sequences reported for other Gram-positive bacteria (MacKay et al., 1986). The amino acid sequences of the positively charged and hydrophobic core regions of the signal sequences of the five Bacillus endoglucanases are conserved, although the hydrophilic regions of the enzymes from the alkalophilic strain N-4 are shorter than the others (see Fig. 5.6). The putative signal sequence of the endoglucanase from C. acetobutylicum contained a longer positively charged region with five lysine residues and a hydrophobic core which contained some conserved amino acid residues. The site of signal sequence processing in the endoglucanases from B. subtilis PAP115 and DLG occurred between two alanine residues following the hydrophobic core (MacKay et al., 1986; Robson and Chambliss, 1987). Many signal peptides are cleaved immediately after an A-X-A sequence (Pearlman and Halvorsen, 1983) and this holds true for the two B. subtilis endoglucanases and other B. subtilis genes (Fig. 5.5). Although a cleavage site for the C. acetobutylicum endoglucanase signal sequence was not determined, a predicted cleavage site would be after the $A_{36}-E_{37}-A_{38}$ sequence between A_{38} and A_{39} (Fig. 5.2 and Fig. 5.5). The signal sequence cleavage site has not been established for the endoglucanases from the alkalophilic

approximately 75% of the amino acid sequence showed extensive homology with the sequences of a number of other endo- β -1,4-glucanase enzymes, whereas the remaining C-terminal end of the protein exhibited no discernible homology with the other endoglucanase amino acid sequences.

Recently the nucleotide sequences of the genes encoding endoglucanases and exoglucanases from a number of organisms including C. thermocellum, C. fimi, T. reesei, E. chrysanthemi, B. subtilis and an alkalophilic Bacillus species have been reported (Table 1.4). The amino acid sequence of the endo- β -1,4-glucanase enzyme from C. acetobutylicum showed little overall homology with the endoglucanases encoded by the celA, celB and celD genes of C. thermocellum, the endoglucanase and exoglucanase encoded by the cenA and cex genes of C. fimi or the endoglucanase EG I from T. reesei (Table 5.2). There was also no discernible homology with the amino acid sequence encoded by the endo- β -1,3-1,4-glucanase gene from B. subtilis C120.

Table 5.2. Amino acid homology between the endo- β -1,4-glucanase from C. acetobutylicum and endoglucanases from other organisms. Refer to Table 1.4 for references.

Organism		Linkage specificity of endoglucanase	Number of AA residues	AA homology of protein sequence
<u>C. acetobutylicum</u>	P262	β 1,4	448	100.0%
<u>C. thermocellum</u>	CelA	β 1,4	488	11.9%
<u>C. thermocellum</u>	CelB	β 1,4	563	14.4%
<u>C. thermocellum</u>	CelD	β 1,4	649	15.3%
<u>C. fimi</u>	CenA	β 1,4	449	13.7%
<u>T. reesei</u>	EG I	β 1,4	437	14.5%
<u>B. subtilis</u>	C120	β 1,3 1,4	242	13.8%
<u>B. subtilis</u>	PAP115	β 1,4	499	43.4%
<u>B. subtilis</u>	DLG	β 1,4	499	44.6%
Alkalophilic <u>Bacillus</u> sp N-4(pNK1)		β 1,4	488	43.9%
Alkalophilic <u>Bacillus</u> sp N-4(pNK2)		β 1,4	409	48.4%
Alkalophilic <u>Bacillus</u> sp 1139		β 1,4	800	23.8%

However, comparison of the C. acetobutylicum endoglucanase amino acid sequence with two enzymes from the alkalophilic B. subtilis N-4 strain, the enzymes from B. subtilis strains PA115 and DLG (which exhibit 93.4% homology) and E. chrysanthemi, showed significant homology (Table 5.3). A much larger endo- β -1,4-glucanase from another alkalophilic Bacillus strain 1139 exhibited much less overall homology.

Table 5.3. Nucleotide and amino acid homology between the endo- β -1,4-glucanase from C. acetobutylicum and endoglucanases from Bacillus species.

Bacteria	Nucleotide homology		Amino acid homology		
	Whole gene	Amino terminal region	Whole gene	Amino terminal region	Carboxyl terminal region
<u>B. subtilis</u> PAPI15	49%	61%	43.4	60.1	16.3
<u>B. subtilis</u> DLG	49%	61%	44.6	61.0	16.3
Alkalophilic <u>Bacillus</u> sp N-4(pNK1)	57%	65%	43.9	59.5	16.7
Alkalophilic <u>Bacillus</u> sp N-4(pNK2)	57%	65%	48.4	61.0	20.2

The amino acid alignment of the endo- β -1,4-glucanases from the C. acetobutylicum P262, B. subtilis PAPI15, DLG, alkalophilic Bacillus N-4 and 1139 strains, and E. chrysanthemi is shown in Fig. 5.6. The aligned amino acid sequences of the endoglucanases showed a region of homology which extended for approximately 300 residues from

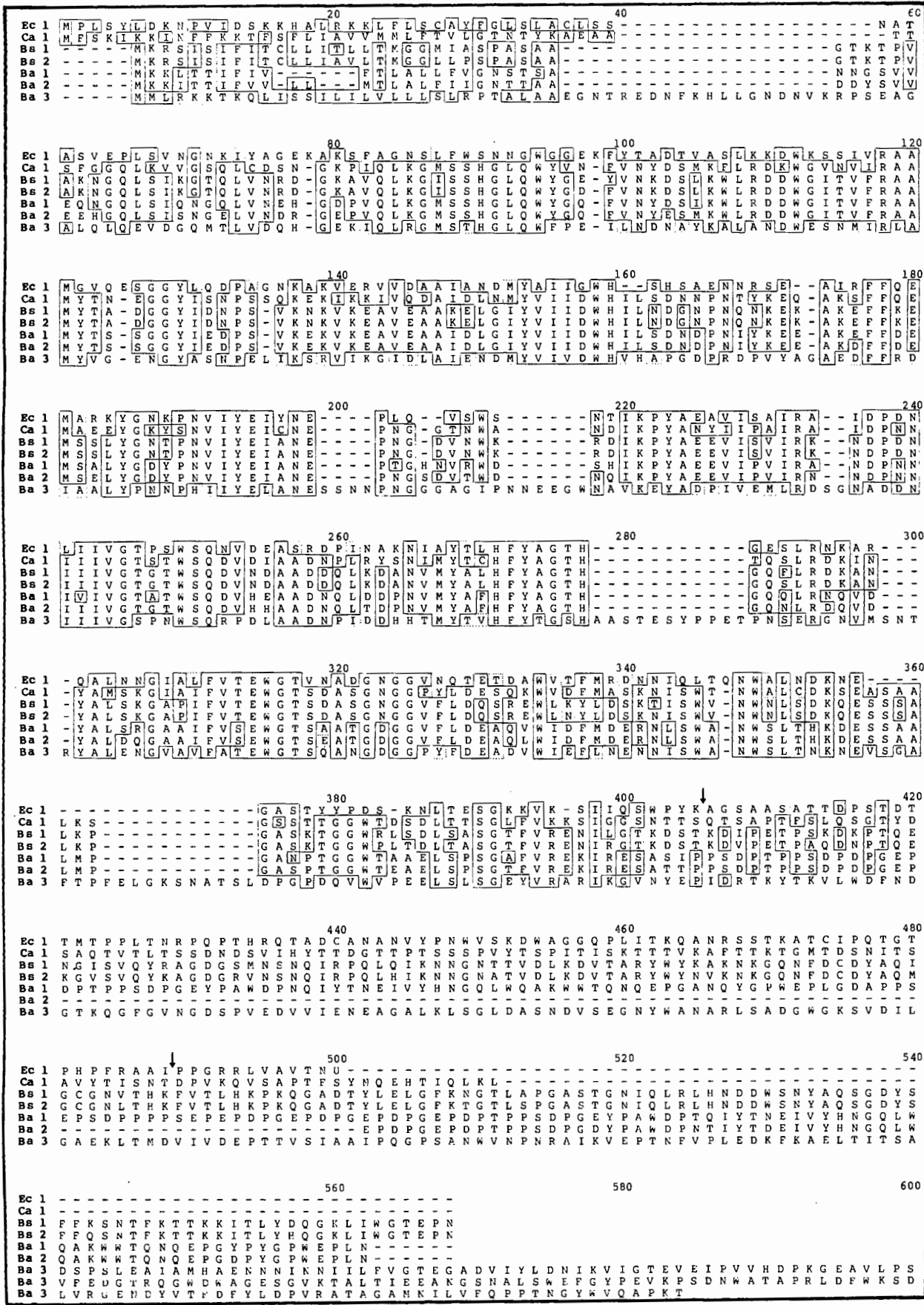


Fig. 5.6. Amino acid sequence alignments of the endo-β-1,4-glucanases from *E. chrysanthemi* (Ec 1), *C. acetobutylicum* P262 (Ca 1), *B. subtilis* PAP115 (Bs 1), *B. subtilis* DLG (Bs 2), Alkalophilic *Bacillus* species N-4 pNK1 (Ba 1), Alkalophilic *Bacillus* species N-4 pNK2 (Ba 2) and Alkalophilic *Bacillus* species 1139 (Ba 3). The amino acids are identified by the single letter code and regions of

identical homology are boxed with solid lines. Conserved amino acids are boxed by broken lines. The arrow at position 405 indicates the point to which BAL-31 digestion removed the C-terminal region of the B. subtilis PAP115 endoglucanase without affecting enzyme activity. The arrow at position 488 indicates the extent of the truncated active endoglucanase gene from C. acetobutylicum cloned in the recombinant plasmid pHZ100. Refer to Table 1.4 for references.

the end of the signal sequence (Fig. 5.6). In comparison with the other four endoglucanases (Table 5.3), this region of the C. acetobutylicum enzyme showed approximately 61 - 65% nucleotide and 60% amino acid sequence homology. If amino acid replacement by conserved amino acids is taken into account, the overall homology of this region is increased to 80%. Although the amino acid homology of the entire endoglucanase from the alkalophilic Bacillus 1139 strain was only 23%, the amino acid homology of the N-terminal region was 40%.

Previously MacKay et al. (1986) identified a short region situated between amino acids 82 and 141 which appeared to show some similarity to the amino acid sequences of the endo- β -1,4-glucanase encoded by the celB gene of C. thermocellum. These authors suggested that this region may contain some of the residues of the active site of the enzyme. Comparison with the aligned amino acid sequences of the endoglucanases from C. acetobutylicum and the Bacillus species confirmed the presence of two short regions situated between positions 85 - 97 and 115 - 143 which appeared to show some homology with the amino acid sequence of the endoglucanase encoded by the celB gene of C. thermocellum.

The C-terminal region of the first five Gram-positive endoglucanases (Ca 1 to Bs 2, Fig. 5.6) varied in length from approximately 20 to 160 amino acids (from position 400, Fig. 5.6). Comparison of the amino acid sequence of the C-terminal region of the endoglucanase from C. acetobutylicum with those from B. subtilis PAP115 and DLG showed no apparent homology. However, the C-terminal region of these two B. subtilis strains showed a high degree of homology. The C-terminal regions of the two endoglucanases from the alkalophilic N-4 strain were characterized by a number of short direct repeat sequences, including a P - D - P - Q - G sequence (Fukumori et al., 1986b). The C-terminal region of the larger protein contained a longer direct repeat sequence consisting of 60 amino acids. The endoglucanases from C. acetobutylicum, B. subtilis PAP115 and DLG did not contain repeat sequences in the C-terminal region. Repeated amino acid sequences are present in the C-terminal regions of three endoglucanases from C. thermocellum (Beguin et al., 1985; Grepinet and Beguin, 1986; Joliff et al., 1986) and the exoglucanase of C. fimi (O'Neill et al., 1986). Although the endoglucanases encoded by the celA, celB and celD genes of C. thermocellum did not show extensive homology over most of their length (Joliff et al., 1986) a conserved region of more than 60 amino acids containing an imperfect repeat of 24 amino acids was located at the C-terminal end of all three proteins. It has been suggested that the C-terminal repeated sequences may play a role in the binding of the enzymes to cellulose (Joliff et al., 1986). However, C-terminal repeated sequences do not appear to be implicated in the binding to cellulose of the

endoglucanases from C. acetobutylicum, B. subtilis PAP115, DLG and Bacillus N-4.

The cellulases from T. reesei and C. fimi have terminal domains linked to the remainder of the protein by a sequence rich in hydroxyl amino acids and proline. The C. acetobutylicum endoglucanase has a C-terminal region which is rich in hydroxyl amino acids and a few proline residues. Considering the endoglucanase structure model of Knowles et al. (1987), it appears that a hinge region, but no tail region, is present in the C. acetobutylicum endoglucanase.

Analysis of the nucleotide sequence of the C. acetobutylicum endoglucanase gene revealed that the fragment of DNA cloned in the recombinant plasmid pHZ100, which encoded a functional endoglucanase, lacked the last 24 amino acid residues. This C-terminal region was not essential for function. A similar finding was reported in the B. subtilis PAP115 endoglucanase where it was possible to delete approximately one third of the C-terminal amino acids (position 343, Fig. 4) without the loss of enzyme activity (MacKay et al., 1986). It is interesting to note that this corresponds to the start of the short direct repeat sequences at the beginning of the C-terminal region of the endoglucanase from the alkalophilic Bacillus strain N-4. A HindIII deletion derivative of pHZ100 (pHZ110; Fig. 4.1), which removed a further 120 C-terminal amino acids (about one third of the protein) was not able to direct the synthesis of a functional endoglucanase.

DNA sequence analysis of the endo- β -1,4-glucanase genes of B. subtilis PAP115 and DLG indicated that although the genes had a coding capacity for proteins with a M_r of > 50 000, the purified mature extracellular enzyme had a M_r of approximately 35 000. However, an intracellular protein with a M_r of 51 500 was isolated from cell extracts of E. coli transformed with a high copy number plasmid containing the endoglucanase gene from the DLG strain (Robson and Chambliss, 1987). These results suggest that in these strains considerable processing or degeneration of the precursor protein occurs before or during export. Amino terminal amino acid sequencing of the mature extracellular protein indicated that the C-terminal region was removed (MacKay et al., 1986; Robson and Chambliss, 1987). This appears to be the first report of C-terminal processing of a protein in B. subtilis. Fukumori et al. (1986a) presented some evidence to indicate that C-terminal processing of the endo- β -1,4-glucanase from the alkalophilic Bacillus 1139 strain may also occur. These results suggest that C-terminal processing to produce the mature extracellular protein may be a common feature of these endoglucanases. In contrast, the C-terminal end of the endoglucanase encoded by the cenA gene of C. fimi was essential for activity (Wong et al., 1986). However this would be expected if the architecture of the gene were taken into account (see Fig. 1.3): the C-terminal end of the C. fimi cenA gene forms part of the structural gene (hydrolytic domain) since removal of 12 amino acids abolished activity, while removal

of 76 amino acids from the N-terminal end (tail region) did not affect activity (Wong et al., 1986).

5.4 Discussion.

A comparison of the nucleotide sequences of six Gram-positive endo- β -1,4-glucanase genes, from C. acetobutylicum, B. subtilis and two alkalophilic Bacillus strains, indicated that they all contained typical δ^{43} promoter sequences (Doi and Wang, 1986; Graves and Rabinowitz, 1986). Graves and Rabinowitz (1986) compared the promoter sequences of 29 Gram-positive organisms and showed that apart from the -10 and -35 conserved regions, some additional conserved regions appeared to be present. A group of adenosine residues was present at position -41 to -45 with >50% conservation. Two other regions of conservation were found bordering the -10 region namely a TG pair at -16 and -15 and A and T residues between -7 and -3. Generally E. coli (Gram-negative) RNA polymerase can utilise Gram-positive promoters efficiently but not the reverse. The requirement of this "extended" promoter region in Gram-positive promoters, which is not present in E. coli promoters, may be a reason why Gram-negative genes are not generally expressed by Gram-positive organisms (Graves and Rabinowitz, 1986).

The amino terminal ends of the six enzymes showed typical Gram-positive signal sequences and the amino acid sequences of the enzymes from the four different species revealed the presence of a non-homologous C-terminal region of varying length which is probably absent in the mature extracellular enzyme.

Prior to the results presented here and the recent sequencing of and E. chrysanthemi endoglucanase, no close homology had been demonstrated between endoglucanases from different bacterial genera. Since there appeared to be little or no conservation of amino acid sequences between endoglucanases from different species and much greater conservation between enzymes from the same or closely related strains, it was postulated that the cellulolytic enzymes of the various microorganisms may have evolved independently (Joliff et al., 1986). The occurrence of a highly conserved protein sequence in the endoglucanases from C. acetobutylicum, B. subtilis, the alkalophilic bacilli and E. chrysanthemi, indicated that at least with this group of enzymes, this is not the case. The limited homology between the C. acetobutylicum endoglucanase and the three endoglucanases from C. thermocellum on the one hand, and the strong homology of the C. acetobutylicum endoglucanase with enzymes from the Bacillus group on the other, is interesting, and tends to support recent phylogenetic studies (Woese, 1987). These studies indicate the separation of the low (G+C) Gram-positive bacteria into a number of branches which delineate the thermophilic clostridia from the mesophilic clostridia and the bacilli (Woese, 1987). It is also interesting that the endoglucanases showing strong homology all originate from non-cellulolytic species which produce only one or a few endoglucanases, whereas the endoglucanases which show little homology belong to cellulolytic species which are known to produce a complex battery of cellulolytic enzymes.

Chapter 6

General discussion.

C. acetobutylicum has been extensively used in industrial fermentation systems, but relatively low solvent yields and high substrate costs have made the fermentation unprofitable. Renewed interest in the fermentation, using various lignocellulosic wastes as cheap substrates, has resulted in studies dealing with the ability of C. acetobutylicum to produce solvents from these wastes in mono- or co-culture systems. C. acetobutylicum does not have the ability to efficiently utilise lignocellulosic materials, but the possibility of genetic manipulation of the organism for improved solvent yields is being investigated (see Jones and Woods, 1986).

This study was aimed at establishing a genebank of C. acetobutylicum DNA for the isolation of genes involved in cellulose and hemicellulose degradation. The library was also screened for genes involved in nitrogen metabolism and solvent production.

A number of different approaches can be utilised to study the enzymes involved in cellulose and hemicellulose hydrolysis in C. acetobutylicum. Lee and co-workers (Lee et al., 1985a; 1985b) have reported a number of physiological studies aimed at the characterisation of the cellulolytic and hemicellulolytic enzymes produced by C. acetobutylicum. In this study an attempt was made to

characterise the genes involved in the production of these enzymes.

A library of chromosomal DNA fragments from C. acetobutylicum was established in the plasmid pEcoR251, an E. coli positive selection vector. A library of approximately 15 000 clones was generated with an average insert size of 6 kb. Calculation of the number of clones required for a 99% probability of cloning a specific gene showed that this number of clones was more than adequate. The presence of inserts was confirmed by the recovery of plasmids, larger than the parental vector, in a random selection of clones. Initially it was not known if genes from this Gram-positive anaerobe would be expressed in E. coli. Therefore the library was screened for direct expression of C. acetobutylicum genes by complementation of nutritional markers in E. coli. Complementation of E. coli auxotrophic markers argG6 and hisG1 was observed. The library was then screened for genes expressing cellulolytic and xylanolytic activity. Clones expressing glutamine synthetase and alcohol dehydrogenase activities were also isolated. This genebank has provided a useful basis for the molecular genetic study of C. acetobutylicum P262 and is currently being examined for clones expressing other activities involved in solventogenesis and sporulation. The methods of gene isolation are not limited to detection by expression, but also employ isolation by the use of DNA probes and antigenic methods. It is possible that certain genes may be expressed in an enzymatically inactive, but antigenically active, form in E. coli.

A xylanase gene was isolated from the genebank which enabled E. coli HB101 cells to produce intracellular xylanase activity. The cloned xylanase had an apparent M_r of approximately 28 000 and an isoelectric point of approximately 10. Optimum xylanase activity was obtained at pH 6.0 at 37 - 43°C. Comparison with a xylanase partially purified from the culture medium of C. acetobutylicum P262 showed that the enzymes had similar characteristics and Western blot analysis showed that they were antigenically related. The xylanase cloned from C. acetobutylicum was similar in characteristics to Xyn B described by Lee et al. (1987). Therefore the cloned gene is defined as the C. acetobutylicum P262 xynB gene and its product as Xyn B.

A number of further studies of the cloned xylanase can now be undertaken. The nucleotide sequence of the gene will be determined and compared to other xylanase genes. In view of the similarity between the endoglucanases from C. acetobutylicum and Bacillus it is possible that homology with other Bacillus xylanases will be demonstrated. It will be interesting to determine if the xylanase gene codes for a signal peptide and if so, to compare it to the sequence obtained for the putative endoglucanase signal peptide, since xylanase activity was found predominantly in the cytoplasm and endoglucanase activity predominantly in the periplasm of the E. coli host. A more detailed analysis of the specific activity of the xylanase can be determined by analysing the reaction products generated from xylooligosaccharides using HPLC analysis.

The endo- β -1,4-glucanase gene was isolated from the C. acetobutylicum library by isolation of two clones generated from independent cloning events. One of the clones (pHZ100) did not contain the entire gene, but still encoded an enzymatically active protein in the E. coli host. The endoglucanase encoded by pHZ100 is defined as the C. acetobutylicum P262 celA gene and its product EG A. Attempts to purify the endoglucanase from C. acetobutylicum P262 were not successful due to the very low levels of activity produced by this strain (Allcock et al., 1981). Lee et al. (1985a) found that C. acetobutylicum B527 endoglucanase activity was unstable in batch culture and showed that pH maintenance at pH 5.2 resulted in a 400-fold increase in activity. Application of a similar technique to C. acetobutylicum P262 may result in enhanced production of the endoglucanase. As with the xylanase, further analysis of the endoglucanase activity can be conducted by examining the hydrolysis products produced from cellooligosaccharides using HPLC. Purification of the protein to homogeneity could possibly be achieved by a combination of gel filtration, ion exchange and affinity chromatography.

The reducing sugar assay method employed in this study was used to partially characterise the endoglucanase and xylanase activities produced by the cloned genes and not with the intention of absolute quantitative analysis. It is clear that reducing sugar assay methods can be affected by various substances in the assay and can result in overestimates of activity. As stated by Coughlan (1985b), the reducing sugar assay methods are adequate when using

well defined substrates. HPLC analysis of the enzyme products is the procedure of choice although even this technique is not without problems (Coughlan, 1985b).

The characteristics of the cloned endoglucanase differed from those previously reported for this strain (Allcock et al., 1981) and for ATCC 824 (Lee et al., 1985a). This could be due to the phenomenon of enzyme multiplicity which can be caused by multiple genes or modification after translation or secretion. The enzyme may be processed differently in C. acetobutylicum and in E. coli. The occurrence of multiple genes for the cellulolytic and hemicellulolytic systems has been demonstrated for fungal and bacterial cellulase systems. Lee et al. (1987) have isolated two antigenically distinct xylanases from C. acetobutylicum ATCC 824, presumably produced by different genes. The number of non-homologous endoglucanase and xylanase genes present in organisms that do not produce the full battery of cellulolytic enzymes may not be as high as the "true" cellulolytic organisms (e.g. C. thermocellum) because the endoglucanases, for which the nucleotide sequence is known, show extensive amino acid homology (Bacillus species, C. acetobutylicum and E. chrysanthemi).

Most of the endoglucanase activity was detected in periplasmic extracts of E. coli cells, indicating that export of the protein from the E. coli cytoplasm occurred. A putative signal peptide was identified which conformed, in general structure, to many other signal peptides from both pro- and eukaryotes (Pearlman and Halvorsen, 1983). Protein

export activity of this region could be determined by an in-frame fusion with TnphoA. If the sequence compensates for the missing phoA signal sequence the fusion will produce alkaline phosphatase activity (Manoil and Beckwith, 1985). The pHZ300 encoded xylanase activity was detected in the cytoplasm of E. coli cells although the enzyme is, as with the endoglucanase, secreted by C. acetobutylicum into the extracellular medium. The reason for the inability of E. coli cells to export the protein to the periplasm is not known, but comparison of the N-terminal regions of the genes may reveal interesting differences in the signal peptide sequences.

Both genes were expressed from their own promoters in E. coli as demonstrated by deletion, subcloning and the reversal of orientation of restriction fragments. The nucleotide sequence of the upstream region of the endoglucanase gene contained two putative extended promoter consensus sequences characteristic of Gram-positive bacteria suggesting that C. acetobutylicum utilises typical δ^{43} type promoters, at least for genes transcribed during the vegetative growth phase. The promoter activity of this region can be tested by fusing the region to a promoter-less lacZ gene and assaying β -galactosidase activity, while the actual start of the mRNA transcript could be determined by a 3' to 5' primer extension experiment. Regulation of C. acetobutylicum genes also appears to occur in E. coli. Both the endoglucanase and xylanase genes were subject to glucose repression although xylose did not repress xylanase expression. Xylanase expression was enhanced in

C. acetobutylicum ATCC 824 cultures containing xylose (Lee et al., 1985b). The C. acetobutylicum P262 glutamine synthetase (GS) gene was found to be regulated by nitrogen levels in E. coli (Usdin et al., 1986; Janssen et al., 1988).

Previously little homology had been demonstrated between various cellulase genes. Some homology was demonstrated between cellulases of T. reesei and those of P. chrysosporium and S. commune (see Knowles et al., 1987) although most cellulases, even from the same organism, appeared to be very dissimilar. The complete amino acid sequence (448 residues) of the C. acetobutylicum endoglucanase was deduced, and comparisons were made with reported amino acid sequences of endoglucanases from other organisms. A high degree of homology was found between the C. acetobutylicum endoglucanase and endoglucanases from Bacillus strains but, no extensive homology with sequenced endoglucanases from C. thermocellum could be demonstrated. A recent report of the sequence of an E. chrysanthemi endoglucanase (Guisseppi et al., 1988) also revealed homology with Bacillus endoglucanases. Extensive homology between endoglucanases has so far only been demonstrated for organisms that are not considered to be truly cellulolytic and would indicate a common ancestor for these genes. It would be interesting to probe a C. thermocellum or T. reesei DNA library for the presence of a related gene.

This study has provided a foundation for the manipulation of C. acetobutylicum genes. Further characterisation and

manipulation of the genes described here may eventually lead to the enhancement of the cellulolytic and hemicellulolytic capabilities of C. acetobutylicum.

Appendix A

Maintenance of bacterial strains

A.1 C. acetobutylicum P262. C. acetobutylicum P262, which was used for DNA extraction and purification of xylanase, was maintained as a spore stock in distilled water at 4°C. Spore stocks were prepared as described by Long et al. (1983).

A.2 E. coli.

A.2.1 E. coli stock cultures. Stock cultures of E. coli cultures were maintained at -70°C. Cultures (5 ml LB medium, containing antibiotics as required) (Appendix B) were inoculated from overnight cultures (1/1 000 dilution) or from isolated colonies on LB plates, and incubated with aeration at 37°C to OD₆₀₀ 0.6. Sterile glycerol was added (15%, v/v final concentration), and aliquots (100 µl) were stored at -70°C.

E. coli JM103 was plated on minimal A medium (Appendix B). Isolated colonies were dispersed in 100 µl LB medium containing 15% (v/v) glycerol and stored at -70°C.

A.2.2 Competent E. coli cells. E. coli competent cells, prepared as described in Appendix C, were stored at -70°C in 0.1 M CaCl₂ containing 15% (v/v) glycerol.

Appendix B

Media, buffers and solutions

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

B.1 Media

B.1.1 Clostridial Basal Medium (CBM) (O'Brien and Morris, 1971).

Glucose	10 g
Casein hydrolysate	4 g
Yeast extract	4 g
Distilled water	972 ml

The following constituents were added from stock solutions as indicated.

MgSO ₄ ·7H ₂ O	(20%, w/v)	1 ml
MnSO ₄ ·4H ₂ O	(1%, w/v)	1 ml
FeSO ₄ ·7H ₂ O	(1%, w/v)	1 ml
p-Aminobenzoic acid	(0.1%, w/v)	1 ml
Biotin	(0.0002%, w/v)	1 ml
Thiamine HCl	(0.1%, w/v)	1 ml
NaHCO ₃	(10%, w/v)	10 ml
Cysteine HCl	(5%, w/v)	10 ml
Resazurin	(0.25 mg/ml)	2 ml

The stock solutions were not autoclaved and were stored at 4°C, except FeSO₄ which was stored at -20°C.

Glucose was replaced with 5 g xylose and 5 g of xylan was added to the broth when cultures were used for the purification of xylanase. The quantity of glucose was reduced to 1 g and 5 g of CMC was added to the broth for endoglucanase purification experiments. Agar (1.5% w/v) was added to CBM broth made as above without resazurin and with the exception that the Cysteine HCl and NaHCO₃ were added after the medium had been autoclaved and prior to pouring.

Liquid CBM medium was autoclaved and allowed to stand for 15 h under stringent anaerobic conditions in an anaerobic glove box (Forma Scientific Inc., Marietta, Ohio; atmosphere 70% N₂ : 20% CO₂ : 10 % H₂ (v/v/v)) before inoculation. Cysteine hydrochloride, added to all liquid anaerobic cultures, scavenged any residual oxygen.

Broth aliquots (10 ml), used for the preparation of overnight cultures of C. acetobutylicum, were prepared by boiling the medium for 15 min to drive off most of the oxygen, dispensing into Hungate tubes and perfused with H₂ and CO₂, before sterilisation by autoclaving.

B.1.2 Minimal A medium (Miller, 1972).**Solution 1**

Agar (Oxoid No. 1)	15 g
Distilled water	600 ml

Minimal salts solution (X5)

K_2HPO_4	52.5 g
KH_2PO_4	22.5 g
$(NH_4)_2SO_4$	5.0 g
$C_6H_5Na_3O_7 \cdot 2H_2O$ (Sodium citrate)	2.5 g
Distilled water	to 1000 ml

This solution was stored over chloroform at 4°C.

Solution 2

Minimal salts solution (X5)	200 ml
Distilled water	200 ml

Solutions 1 and 2 were autoclaved separately, cooled to 50°C, combined, and the remainder of the constituents added from stock solutions as indicated.

Carbon source	(20%, w/v)	10 ml
Amino acids	(0.2%, w/v for L-amino acids) (0.4%, w/v for DL-amino acids)	1 ml
$MgSO_4$	(20%, w/v)	1 ml
Vitamin B1	(0.5%, w/v)	1 ml

B.1.3 H plates

Bacto tryptone	10 g
NaCl	8 g
Agar	12 g
Distilled water	1000 ml

B.1.4 H top agar

Bacto tryptone	10 g
NaCl	8 g
Agar	8 g
Distilled water	1000 ml

B.1.5 Luria-Bertani medium (LB)

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

The pH of the medium was adjusted to pH 7.0 with 0.1 M NaOH before autoclaving. Solid media contained 1.5% (w/v) agar.

B.1.6 Phosphate limiting medium (Garen and Levinthal, 1960)

The medium was made according to the following table:

Stock solution	final conc.	per 1
Tris-Cl (1 M, pH 7.5)	120 mM	120 ml
CaCl ₂ (1 M)	0.2 mM	0.2 ml
KCl (1 M)	20 mM	20 ml
MgCl ₂ (1 M)	1 mM	1 ml
NaCl (5 M)	80 mM	16 ml
NH ₄ Cl (1 M)	20 mM	20 ml
Glycerophosphate (1 M)	0.14 mM	0.14 ml
Na ₂ SO ₄ (1 M)	0.5 mM	0.5 ml
Glucose (20%, w/v)	12 mM	11 ml
Bacto peptone	0.04%, w/v	0.4 g
FeCl ₃ (0.1 M)	2 uM	20 ul

The medium was sterilised by autoclaving. FeCl₃ was filter sterilised and added after autoclaving.

B.1.7 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 X-gal (0.8 ml) were added to 250 ml agar (50°C) before pouring the plates.

B.2 Media additives

Media were cooled to 50°C before addition of ampicillin, X-gal or IPTG. Plates containing these additives were stored for no longer than one week at 4°C.

B.2.1 Ampicillin (X1000)

Ampicillin (sodium salt; Sigma)	1 g
Distilled water	10 ml

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

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

IPTG (100 mM)	23.8 mg
Distilled water	1 ml

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

B.2.3 X-gal (5-bromo-4-chloro-3-indolyl- β -galactoside)

X-gal (2%, w/v)	0.2 g
Dimethylformamide	10 ml

The solution was stored at -70°C.

B.3 Buffers and Solutions

B.3.1 BAL-31 dilution/storage buffer. The buffer was made according to the following table and stored at -20°C :

Stock solution	final conc.	per 10 ml
Tris-Cl (1 M, pH 8.0)	20 mM	0.2 ml
CaCl ₂ (1 M)	5 mM	50 μl
MgCl ₂ (1 M)	5 mM	50 μl
EDTA (0.5 M, pH 8.0)	1 mM	20 μl
NaCl (5 M)	0.1 M	0.2 ml
Glycerol	44% (v/v)	4.4 ml
Distilled water		5.58 ml

B.3.2 BAL-31 reaction buffer (X5). The buffer was made according to the following table and stored at 4°C :

Stock solution	final conc.	per 10 ml
Tris-Cl (1 M, pH 8.0)	0.1 M	1 ml
CaCl ₂ (1 M)	60 mM	0.6 ml
MgCl ₂ (1 M)	60 mM	0.6 ml
EDTA (0.5 M, pH 8.0)	5 mM	0.1 ml
NaCl (5 M)	3 M	6 ml
Distilled water		1.7 ml

B.3.3 DNA polymerase I buffer (Klenow). The buffer was made according to the following table and stored at -20°C

Stock solution	final conc.	per 10 ml
Tris-Cl (1 M, pH 7.6)	0.1 M	1 ml
MgCl ₂ (1 M)	60 mM	1 ml
NaCl (5 M)	0.5 M	1 ml
2-mercaptoethanol	0.7 M	50 μl
Distilled water		6.95 ml

B.3.4 DNA sample loading solution (X6)

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

The solution was stored at 4°C .

B.3.5 Denhardt's solution (X10) (Denhardt, 1966)

Ficoll (1%, w/v)	1 g
Polyvinylpyrrolidone-40 (1%, w/v)	1 g
BSA (Fraction V) (1%, w/v)	1 g
Distilled water	to 100 ml

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

B.3.6 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 completely dissolved in the water before adding the other constituents and dissolving in turn. The phenol was melted at 50°C. A 3 ml sample was titrated to the end-point with 5 - 6 ml 0.1 M HCl 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 N₂.

B.3.7 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 at 4°C.

B.3.8 Folin-Lowry assay solution A

Na ₂ CO ₃	(2%, w/v)	1 g
Sodium citrate	(2%, w/v solution)	0.5 ml
CuSO ₄ ·5H ₂ O	(1%, w/v solution)	0.5 ml
Distilled water		50 ml

The solution was prepared immediately before use.

B.3.9 Isopropanol solution, salt saturated. Isopropanol was saturated with aqueous 5 M NaCl, 10 mM Tris-Cl and 1 mM EDTA (pH 8.5) (Maniatis *et al.*, 1982).

B.3.10 Ligase dilution buffer. The buffer was made according to the following table and stored at -20°C :

Stock solution	final conc.	per 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

B.3.11 Ligation buffer (X10). The buffer was made according to the following table and stored in aliquots at -70°C :

Stock solution	final conc.	per 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

B.3.12 Phenol (buffer 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 1 M Tris-Cl (pH 8.0) or until the pH of the aqueous phase was approximately pH 7.6. The phenol was stored under 0.1 M Tris-Cl (pH 8.0) at -20°C .

B.3.13 Restriction enzyme core buffers (X10).

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	0, 50, 100 or 150 mM

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

Stock solution	Salt concentration (mM)			
	0	50	100	150
Tris-Cl (pH 7.9)	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
Sterile water	2.4 ml	1.4 ml	0.4 ml	2.4 ml
NaCl (5 M)	-	1 ml	2 ml	87.7 mg

B.3.14 Restriction enzyme dilution buffer

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

Stock solution	final conc.	per 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
This solution was filter sterilized and then the following constituents added:		
2-mercaptoethanol	10 mM	7 μl
Gelatin (10 mg/ml)	100 $\mu\text{g/ml}$	0.1 ml
Glycerol	44% (v/v)	4.4 ml

B.3.15 SmaI restriction endonuclease buffer (X10)

Stock solution	final conc.	per 10 ml
Tris-Cl (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

B.3.16 SDS-Polyacrylamide gel electrophoresis reagents and gel preparation table (Laemmli, 1970).

Acrylamide-bis-acrylamide stock solution

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

The solution was filtered through Whatman's paper (No. 1) and stored in a dark bottle at 4°C.

Destain solution

Acetic acid	10%, v/v
Methanol	25%, v/v
Distilled water	65%, v/v

Running gel buffer

Tris-Cl (1.5 M, pH 8,8)	36.3 g
Distilled water	to 200 ml

Sample treatment buffer

Stacking gel buffer	2.5 ml
SDS (10%, w/v solution)	4 ml
Glycerol	2 ml
2-mercaptoethanol	1 ml
Distilled water	0.5 ml

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

Stacking gel buffer

Tris-Cl (0.5 M, pH 6.8)	3.0 g
Distilled water	to 50 ml

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).

Tank buffer

Tris base (0.25 M)	12 g
Glycine (0.192 M)	57.6 g
SDS (0.1%, w/v)	4 g
Distilled water	to 4000 ml

Polyacrylamide gel preparation table (10% gels)

Stock solution	Running gel	Stacking gel
Acrylamide	20 ml	2.66 ml
Running gel buffer	15 ml	-
Stacking gel buffer	-	5 ml
SDS (10%, w/v solution)	0.6 ml	0.2 ml
Distilled water	24.1 ml	12.2 ml
Ammonium Persulphate* (10%, w/v solution)	0.3 ml	0.1 ml
TEMED	20 ul	10 ul

* made immediately before use.

B.3.17 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 an MSE Soniprep sonicator. The solution was aliquotted and stored at -20°C. Immediately

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

B.3.18 SSC (X20)

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

The solution was adjusted to pH 7.0 with NaOH and sterilised by autoclaving.

B.3.19 TE (Tris-EDTA) buffer (X100)

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

The buffer was autoclaved before use and appropriately diluted into sterile distilled water.

B.3.20 Tris-acetate buffer (X50)

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

B.3.21 Western blotting buffers, reagents and solutions

Blocking buffer stock solution

Tris-Cl (pH 7.4)		1.21 g
NaCl		9 g
Distilled water		to 1000 ml

Blocking buffer

Blocking buffer stock solution	500 ml
Fat-free milk powder	10 g
Tween-20	0.25 ml
Sodium azide	0.1 g

Blot buffer

Tris base	(25 mM)	15 g
Glycine	(0.192 M)	72 g
Methanol	(20%, v/v)	1000 ml
Distilled water		to 5000 ml

GARP solution. (Goat anti rabbit IgG conjugated horseradish peroxidase; bio-Yeda Ltd., Israel, Cat. No. 3671).. The serum was diluted 1/1500 in blocking buffer without sodium azide immediately before use.

Horseradish peroxidase substrate solution

4-chloro-1-naphthol (Aldrich No. C5,780-4)	30 mg
Methanol	10 ml
Substrate buffer	50 ml
H ₂ O ₂	30 ul

Substrate buffer

Tris-Cl	(0.2 M, pH7.4)	6 g
NaCl	(0.2 M)	11.6 g
Distilled water		to 1000 ml

Wash saline

NaCl	(0.9%, w/v)	9 g
Tween-20	(0.05%, v/v)	0.5 ml
Distilled water		to 1000 ml

B.3.22 Z-buffer (pH 7.0)

Na_2HPO_4	(60 mM)	16.1 g
$\text{NaH}_2\text{PO}_4 \cdot 2\text{H}_2\text{O}$	(40 mM)	5.5 g
KCl	(10 mM)	0.75 g
$\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$	(1 mM)	0.246 g
Distilled water		to 1000 ml

Appendix C

General Techniques

- C.1 Preparation of E. coli plasmid DNA
 - C.1.1 Small scale (miniprep) method
 - C.1.2 Large scale (maxiprep) method
- C.2 Determination of DNA concentration
- C.3 Restriction endonuclease digestion
- C.4 Agarose gel electrophoresis
- C.5 Ligation reactions
- C.6 Rapid subcloning technique
- C.7 Preparation of competent E. coli cells
- C.8 Transformation of competent E. coli cells
- C.9 Radioactive labelling of DNA probes
- C.10 Transfer of DNA from agarose gels to nitrocellulose membranes.
- C.11 DNA hybridisation
- C.12 SDS-Polyacrylamide gel electrophoresis of proteins
- C.13 Non-denaturing (native) polyacrylamide gel electrophoresis
- C.14 Determination of protein concentrations

C.1 Preparation of E. coli plasmid DNA

C.1.1 Small scale method (miniprep). Plasmid was isolated from 4 ml overnight cultures (LB + Ap, 100 µg/ml) as described by Ish-Horowicz and Burke (1981). The cells were collected in a 2 ml microfuge tube by centrifugation (1 min, Eppendorf microfuge). The pellet was resuspended in 0.2 ml Solution I (50 mM glucose, 25 mM Tris-Cl pH 8.0), held at room temperature for 5 min, and then 0.4 ml of Solution II (0.2 M NaOH, 1% (w/v) SDS) was added. The mixture was vortexed briefly and placed on ice for 5 min and then 0.3 ml of precooled Solution III (5 M KOAc, pH4.8) was added and vortexed briefly. After 5 min on ice, precipitated cellular debris and denatured chromosomal DNA were pelleted by centrifugation for 5 min. Plasmid DNA was precipitated from the supernatant (0.9 ml) by addition of isopropanol (0.6 ml). After 2 min at room temperature the precipitate was collected by centrifugation (5 min). The pellet was washed with 70% ethanol, air dried, and resuspended in 0.1 ml TE buffer. Plasmid DNA prepared by this method was sufficiently pure for restriction enzyme digestion and optimal results were obtained if the plasmid DNA preparation was diluted 10 - 20 fold in the digestion reaction. Sufficient plasmid DNA was obtained to carry out at least 50 restriction enzyme digests.

C.1.2 Large scale method (maxiprep). Plasmid DNA was prepared from 800 ml overnight cultures (LB + Ap, 100 µg/ml) using a scaled-up version of the miniprep method. The

procedure was identical but the volumes of Solutions I, II, and III were increased 40 fold. Cellular debris was removed by centrifugation at 12 000 x g for 10 min and nucleic acid was pelleted at 27 000 x g for 15 min. The pellet obtained after isopropanol precipitation was resuspended in 8 ml TE buffer and the pH of the solution was adjusted to pH 7 by addition of 2.5 M Tris-base. Ethidium Bromide (EtBr) (0.5 ml of a 10 mg/ml stock) and CsCl (1 g/ml) were added and the solution was centrifuged (27 000 x g) for 10 min. The refractive index of the supernatant was adjusted to 1.396, the sample sealed in Beckman Quickseal ultracentrifuge tubes and centrifuged for 6 - 14 h at 55 000 rpm at 15°C in a Beckman Vti 65.2 rotor. The plasmid band was visualized by long wave UV light (350 nm) and removed in as small a volume as possible. If the DNA was to be used as a cloning vector it was purified by two further cycles of EtBr/CsCl centrifugation. The EtBr was removed by extraction (3 times) with an equal volume of NaCl-saturated isopropanol (Appendix B). The salt concentration of the solution was reduced by diluting the sample with 2 volumes of water and the DNA was precipitated by the addition of an equal volume of isopropanol. The plasmid DNA was finally resuspended in 0.5 ml TE buffer and the concentration was determined spectrophotometrically as described below (C.2). Plasmid yields were of the order of 2 - 4 mg/l culture. Plasmid identity was confirmed by restriction enzyme digestion and by comparison with its restriction map.

C.2 Determination of DNA concentration. The concentration of DNA solutions was assessed spectrophotometrically by

monitoring the absorbance of the solutions between 220 and 310 nm. The concentration was determined by using the conversion where 1 absorbance unit at 260 nm is equivalent to 50 µg DNA/ml (Maniatis et al., 1982). The 260/280 ratio was in the order of 2.

C.3 Restriction endonuclease digestion. Restriction digests were carried out using one of the four restriction buffers (Appendix B) according to the salt requirements of the particular enzyme (Appendix D). The enzyme SmaI required a unique buffer (Appendix B). Multiple restrictions, requiring different enzymes, could be combined in a single digest provided the salt requirements of the enzymes were compatible (Appendix D). If this was not possible, then the digestion was done sequentially using the enzyme with the lowest salt and highest temperature requirement first and the salt concentrations and incubation temperatures were adjusted before the addition of the next enzyme. The incubation temperature was 37°C for all enzymes except BclI and TaqI (65°C). Digestion volumes varied depending on the amount of DNA being digested but normally 200-300 ng DNA was digested in a 20 µl volume using 1 unit of enzyme for 1 h. Concentrated enzyme stocks were diluted to 1 or 2 units/µl using a universal restriction enzyme dilution buffer (Appendix B). For electrophoretic analysis, digestions were stopped by the addition of DNA sample loading solution (Appendix B). If a sample was to be used for further enzyme reactions (eg. ligation), it was purified by phenol and chloroform extraction.

The DNA solution was extracted with phenol (1/10 volume, TE-saturated), then an equal volume of chloroform/isoamyl alcohol (24:1) was added and the solution emulsified by shaking. The phases were separated by centrifugation and the aqueous phase was extracted twice with water-saturated ether. If the DNA concentration was less than 2 $\mu\text{g}/100 \text{ ul}$, E. coli tRNA was added (2 $\mu\text{g}/100 \text{ ul}$) before the DNA was precipitated by the addition of 0.1 volume of 5 M NaClO_4 and 1 volume of isopropanol. After 10 min on ice the DNA was pelleted by centrifugation for 10 min in a microfuge, washed with 70% ethanol and resuspended in TE buffer.

C.4 Agarose gel electrophoresis. Agarose gel electrophoresis was carried out using a horizontal submerged gel system similar to that described by Maniatis et al. (1982). Tris-acetate buffer (Appendix B) was used routinely. Sigma type II agarose was used in varying concentrations (0.3 - 2.0% w/v in electrophoresis buffer; Maniatis et al., 1982) depending on the sizes of the fragments being examined. A concentration of 1% (w/v) was used routinely. The amount of DNA loaded per lane also varied with the sizes and number of fragments but under normal circumstances about 200 ng of plasmid DNA was used. Gels were electrophoresed at 2 v/cm for 16 h. For quick gels a Hoefer Minnie Submarine agarose gel unit (model HE33; Hoefer Scientific Instruments, San Francisco) or similar apparatus was used at 5 v/cm for 2 - 4 h. Gels were stained in electrophoresis buffer containing 0.5 $\mu\text{g}/\text{ml}$ EtBr for 15 - 30 min. DNA bands were visualised using a 254 nm transilluminator (Chromato-Vue Model TS-15, UV Products Inc., San Gabriel, Calif., U.S.A.). 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 better quality photograph and a negative were required then Polaroid type 665 film (ASA 64) with an exposure of 120 - 140 sec at f4.7 was used.

Fragment sizes were calculated by extrapolation from a standard curve of the mobility of λ DNA fragments, digested with either HindIII, PstI or BstEII, plotted against the \log_{10} of their molecular weights. Approximately 0.8 - 1.2 μg of a λ digest was loaded per gel lane.

C.5 Ligation Reactions. The methods of Maniatis et al. (1982) were generally used. Different vector to insert ratios were used depending on the sizes of the fragments to be cloned. DNA concentrations in the reactions varied depending on the aim of the ligation. Recircularisation reactions for isolating deletion plasmids contained DNA concentrations of 1 pmole/ml or less. A DNA concentration in the order of 5 pmole/ml was used (Zabeau, pers. comm.) for recombination reactions where limited polymerisation with circle closure was required. Ligation reactions containing DNA, ligation buffer (Appendix B) and water to the required volume, were performed in sterile microfuge tubes. Sticky-end ligations were performed at room

temperature for 3 h or at 15°C overnight using 0.1 - 0.25 U of ligase, whereas blunt-end ligations were performed at room temperature for 2 - 20 h using 20 - 100 X more ligase.

C.6 Rapid subcloning technique. The method of Struhl (1985) was used, with minor modifications, in the subcloning of various plasmid DNA fragments. The DNA fragments were separated by electrophoresis through low melting point (LMP) agarose (SeaPlaque^R agarose, Marine Colloids, Rockland, ME) in Tris-Acetate buffer (50 mM. pH 8.2). The gel was stained with EtBr and the DNA visualised using a 310 nm wavelength transilluminator. The fragments were excised from the gel in as small a volume as possible. The gel slices were melted at 70°C for 5 min, equilibrated at 37°C, and combined in appropriate proportions in a total volume of 10 µl. An equal volume of 2 X concentrated ligation buffer containing T4 ligase was added, the ligation mixed quickly, and incubated at room temperature for 2 to 20 h. Before using the solution to transform E. coli cells, the gel was heated to 70°C for 5 min, and then diluted with 10 volumes of 0.1 M CaCl₂.

C.7 Preparation of competent E. coli cells. Competent cells of E. coli were prepared essentially as described by Dagert and Ehrlich (1979). An overnight culture of E. coli was diluted 1/1 000 into 50 ml prewarmed (37°C) LB medium and grown with shaking to OD₆₀₀ 0.2. The culture was cooled (on ice, 5 min), the cells collected by centrifugation (5 000 x g, 5 min, 4°C), washed with 50 ml of ice cold 0.1 M

MgCl₂, and resuspended in 25 ml of ice cold 0.1 M CaCl₂. After 60 min on ice the cells were collected by centrifugation and resuspended in 5 ml 0.1 M CaCl₂. The competent cells were kept on ice for at least 1 h before use or aged overnight at 4°C to improve their competency. For long term storage competent cells were treated with glycerol to a final concentration of 15% (v/v) and the cells frozen at -70°C. Competent cells prepared in this way were suitable for use for more than 1 year.

C.8 Transformation of *E. coli* competent cells. Plasmid DNA (1 - 5 ng) or a fraction of a ligation reaction (20 - 100 ng DNA) was added to 100 µl of competent cells on ice. After 10 min the cells were induced to take up the DNA by heat-shocking the transformation mix at 42°C for 2 min. One ml of LB medium was added and the transformation mix left at 42°C for a further 30 min to allow expression of the plasmid borne antibiotic marker. Controls included: competent cells with no DNA added; unrestricted plasmid to monitor the transformation frequency and linearised and ligated plasmid to monitor ligation efficiency. The transformation frequency was in the order of 3 000 - 8 000 colonies per nanogram of unrestricted plasmid DNA.

C.9 Radioactive labelling of DNA probes. DNA probes were labelled with [α -³²P]dATP to high specific activity by nick-translation (Rigby *et al.*, 1977). The reagents were obtained in kit form from Amersham (kit no. PB.5025, Amersham Int., UK) and used according to the suppliers specifications, with the exception that all volumes were

reduced by half. The progress of the reaction was monitored by Cherenkov counting of trichloro-acetic acid (TCA) precipitable material recovered on glass filters (Whatman GFC) as described in the kit protocol. Contaminating nucleotides were removed from the radioactively labelled probe preparation using a Sephadex G50 spin column as described by Maniatis et al. (1982). Specific activities of approximately 1×10^7 counts/min/ μg of DNA were routinely obtained. 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.

C.10 Transfer of DNA from agarose gels to nitrocellulose membranes. DNA fragments resolved by agarose gel electrophoresis were transferred to nitrocellulose membranes (GeneScreen^R, New England Nuclear Corp., Boston, MA, USA) as described by Smith and Summers (1980). The DNA was acid depurinated by soaking the gel twice with agitation for 15 min in 2 gel volumes of 0.25 M HCl. After rinsing the gel with distilled water, the DNA was denatured by soaking the gel twice for 15 min in 0.5 N NaOH, 1.5 M NaCl (2 gel volumes) with agitation. The gel was neutralised by soaking twice for 30 min in 1 M ammonium acetate, 0.02 M NaOH (2 gel volumes). The gel was placed on a flat surface and overlaid with GeneScreen^R membrane (previously soaked in 1 M ammonium acetate, 0.02 M NaOH solution) being careful not to entrap air bubbles. Three pieces of Whatman 3MM filter paper (soaked in the same solution) were layered on top of the membrane followed by a 5 cm stack of dry paper towel. A 1 kg weight on a glass plate ensured even contact and

transfer was allowed to continue for 2 h. The membrane was air dried and baked at 80°C for 2 h in vacuo.

C.11 DNA hybridisation. The hybridisation conditions used were essentially as described by Maniatis et al. (1982) with minor modifications. Optimum conditions for hybridisation of probes to C. acetobutylicum DNA (28% G + C, Cummins and Johnson, 1971) were established empirically. The baked membrane was soaked in 6 X SSC buffer (Appendix B) for 2 min and then placed in a plastic bag. Prewarmed (60°C) hybridisation solution (6 X SSC, 0.5% SDS, 5 X Denhardt's solution (Appendix B), 10 mM EDTA, 100 µg/ml denatured salmon sperm DNA) (without probe) was added (0.2 ml/cm² of membrane), the bag sealed (ensuring no entrapment of air) and prehybridisation allowed to continue for 2 h at 60°C with constant agitation. Thereafter half the hybridisation fluid was removed, the denatured probe added, the bag resealed and hybridisation allowed to continue for 10 - 16 h at 60°C with constant agitation. Solutions of increasing stringencies were used to wash the filters. The filter was washed twice for 5 min with 2 X SSC, 0.5% SDS at room temperature followed by 2 30 min washes at 60°C with 0.1 X SSC, 0.1% SDS. The wet filter was sealed in a plastic bag which enabled further washes after autoradiography, if required. The bag was taped flat under Kodak XAR-5 autoradiographic film in an X-ray cassette fitted with a Fuji X-ray intensifying screen. Exposure was allowed to continue for 1 to 3 d at -70°C. The film was processed using Kodak GBX X-ray developer and fixer according to the manufacturer's instructions.

C.12 SDS-Polyacrylamide gel electrophoresis of proteins.

SDS-polyacrylamide gels were prepared according to the method of Laemmli (1970) using a Hoefer gel apparatus (SE600) with 1.5 mm spacers assembled according to the manufacturer's specifications. All buffers and a preparation table for resolving (10%) and stacking gels are given in Appendix B. The resolving gel was prepared and degassed before pouring. Butanol was layered on the gel to promote a sharp interface. After the gel had polymerised (about 30 min at room temperature), the butanol was removed by rinsing with stacking gel buffer, and the stacking gel cast. The gel was submerged into the electrophoresis tank before loading the samples.

Samples were prepared in sample treatment buffer (Appendix B) and placed in a boiling waterbath for 2 min before being loaded onto the gel. Electrophoresis was continued at 35 mA (constant current) per gel (10 lanes) until the dye front had migrated to the end of the gel (four to five hours).

After electrophoresis the gels were stained for 3 h in staining solution with gentle agitation, destained and dried. The protein molecular mass markers, with a size range of 14 400 - 94 000 D, were obtained from Pharmacia, Uppsala, Sweden (Electrophoresis calibration kit, Cat. No. 17-0446-01).

C.13 Non-denaturing (native) polyacrylamide gel electrophoresis. Proteins were separated under non-denaturing conditions using conditions identical to SDS-PAGE with the exception that SDS was omitted from all buffers. Samples were combined with an equal volume of sucrose-dye solution (50% (w/v) sucrose, 0.1% (w/v) bromophenol blue) and loaded directly on the gel.

C.14 Determination of protein concentrations. Protein concentrations in solutions were determined by the method of Lowry et al. (1951). Assays were performed in triplicate using new disposable test tubes (13 X 175 mm). The reaction contained protein solution (0.1 ml), 0.1 M NaOH (0.1 ml) and 1 ml Folin-Lowry solution A (Appendix B). After 10 min at room temperature, 0.1 ml of Folin-Lowry solution B (1/2 dilution in water of Folin and Ciocalteu's phenol reagent; BDH No. 19058) was added followed by vigorous shaking. After 30 min at room temperature, the optical density of the reaction was monitored at OD_{660} . Protein concentrations were calculated using a standard curve (BSA Fraction V; 0.05 - 0.7 mg/ml). Protein samples were diluted such that OD_{660} did not exceed 0.8.

Appendix D

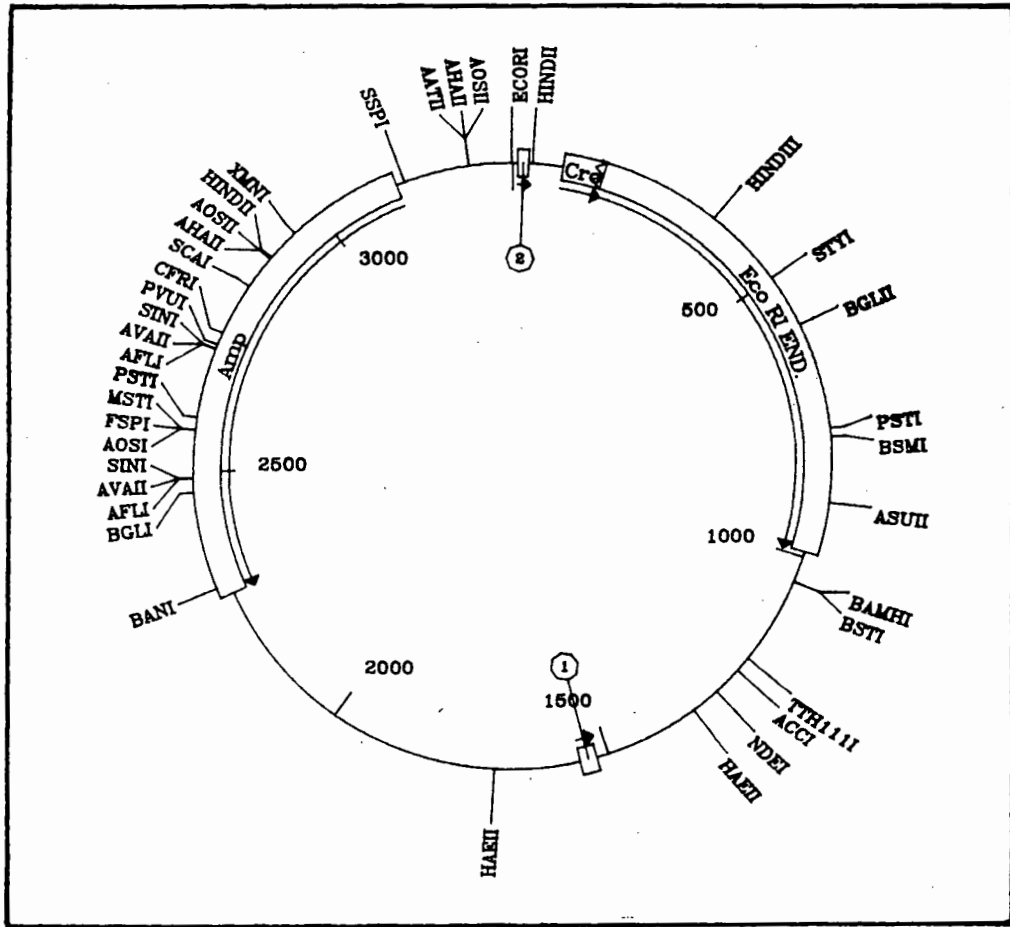
NaCl requirements for restriction enzymes

+++ , 30 - 100% activity compared to recommended conditions;
 ++ , 10 - 30% activity compared to recommended conditions;
 + , < 10% activity compared to recommended conditions;
 * , conditions not recommended due to star activity. The recommended conditions for restriction enzymes are given by the manufacturers (from New England Biolabs Catalog, 1986/87).

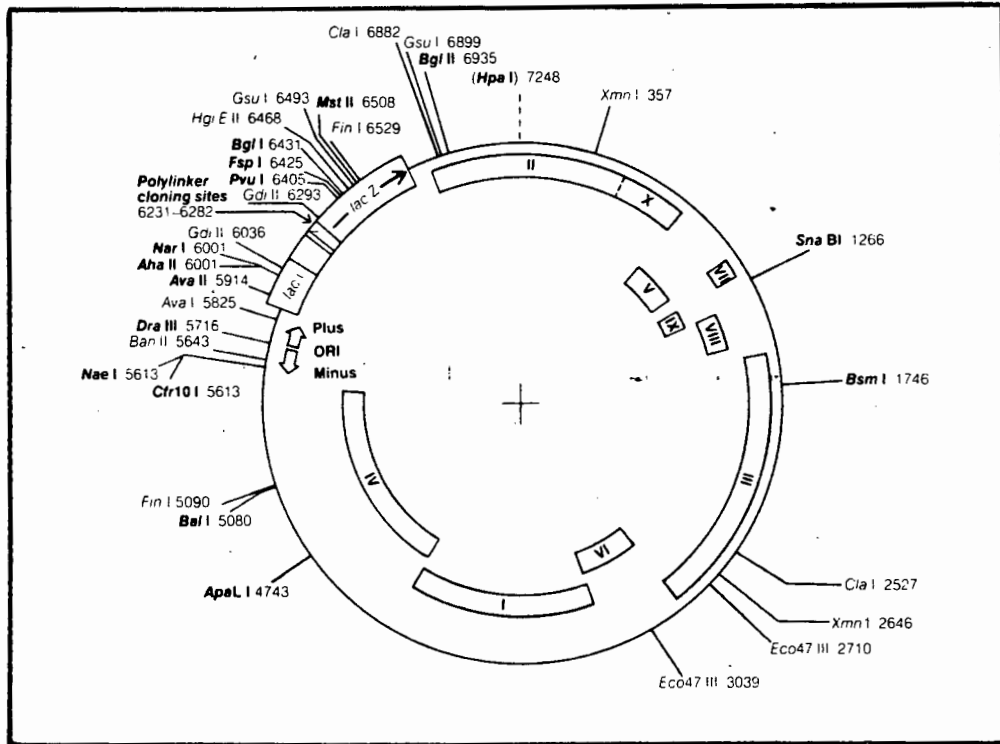
ENZYME	0 mM NaCl	50 mM NaCl	100 mM NaCl	150 mM NaCl	ENZYME	0 mM NaCl	50 mM NaCl	100 mM NaCl	150 mM NaCl	ENZYME	0 mM NaCl	50 mM NaCl	100 mM NaCl	150 mM NaCl
Aat II	+	++	++	+	EcoR I	*	+++	+++	+++	Not I	+	+++	+++	+++
Acc I	+++	+++	+	+	EcoR V	*	+++	+	+++	Nru I	+	+++	+++	+++
Aha II	+	++	+++	+++	FnuD II	+++	++	++	+	Nsi I	++	++	++	+
Alu I	+	+++	+++	+++	Fnu4H I	+++	+++	+++	+	PaeR7 I	+++	+++	+++	+
Apa I	+++	+++	++	+	Fok I	+++	+++	+++	+++	PflM I	++	+++	+++	+
ApaL I	+++	++	+	+	Fsp I	+	+++	+++	+++	PpuM I	+++	+++	+++	+
Ava I	+++	+++	+++	+++	Hae II	+++	+++	+++	+++	Pst I	+++	+++	+++	+++
Ava II	+++	+++	++	+	Hae III	+++	+++	+++	+++	Pvu I	+	+	+++	+++
Avr II	+++	+++	+++	++	Hga I	+++	+++	+	+	Pvu II	+++	+++	+++	+++
Bal I	+++	++	++	+	HgiA I	+	+	++	+++	Rsa I	+++	+++	+++	+++
BamH I	+	++	+++	+++	Hha I	+	+	+	+	Rsr II	+++	+++	+	+
Ban I	+++	+++	++	++	Hinc II	++	+++	+++	+++	Sac I	+++	+++	++	+
Ban II	+++	+++	+++	+++	Hind III	++	+++	+++	+++	Sac II	+++	+++	+++	+
Bbv I	+++	+++	+++	+++	Hinf I	++	+++	+++	+++	Sal I	+	+	++	+++
Bcl I	+	+++	+++	+	HinP I	+++	+++	+++	+++	Sau3A I	+++	+++	+++	+++
Bgl I	+	+++	+++	+++	Hpa I	+	+++	+++	+	Sau96 I	+++	+++	+++	+++
Bgl II	+	+++	+++	+++	Hpa II	+++	+++	+++	+	Sca I	+	+++	+++	+++
Bsm I	+++	+++	+++	+++	Hph I	+++	+++	+++	+	ScrF I	+	+++	+++	+++
Bsp 1286	+++	+++	++	+	Kpn I	+++	+	+	+	SfaN I	+	+	+++	+++
BspM I	++	++	+++	+++	Mbo I	++	++	+++	+++	Sma I	+	+	+	+
BspM II	+	+	+++	+++	Mbo II	+++	+++	+++	+++	SnaB I	+++	+++	++	+
BssH II	+++	+++	+++	+++	Mlu I	++	+++	+++	+++	Spe I	+	+++	+++	+++
BstE II	++	++	+++	+++	Mnl I	+++	+++	+++	+++	Sph I	+	+	+++	+++
BstN I	++	++	+++	+++	Msp I	+++	+++	+++	+++	Ssp I	++	+++	+++	+++
BstX I	+++	+++	+++	+++	Mst II	+	+	+++	+++	Stu I	+++	+++	+++	+++
Cla I	+++	+++	++	++	Nae I	+++	+++	+++	+	Sty I	+	+	+++	+++
Dde I	++	+++	+++	+++	Nar I	+++	+++	+	+	Taq I	+++	+++	+++	+
Dpn I	+	++	+++	+++	Nci I	+++	+++	++	+	Tth111 I	+++	+++	+++	+
Dra I	+	+++	+	+	Nco I	+	++	+++	+++	Xba I	+	+++	+++	+++
Dra III	++	+++	+++	++	Nde I	+	+	++	+++	Xho I	++	+++	+++	+++
Eae I	++	+++	++	+	Nhe I	+++	+++	+++	++	Xho II	+++	+++	++	+
Eag I	++	+++	+++	+++	Nla III	+	+	+	+	Xma I	+++	+++	++	+
EcoO 109	+++	+++	+++	+++	Nla IV	+	+	+	+	Xmn I	+++	+++	+	+

Appendix E

Vector restriction maps



E.1 Restriction map of pEcoR251 (Zabeau, pers. comm.). 1, ori; 2, λ P_I. Restriction sites are shown for those restriction enzymes that cleave the molecule once or twice.



E.3 Restriction map of M13mp18 and M13mp19 (from New England Biolabs catalog, 1986/87). The multiple cloning site polylinker is identical to pUC18/pUC19. The map shows restriction enzyme sites that cleave the molecule once or twice; unique restriction enzyme sites are shown in bold type. The boxes represent the coding sequences.

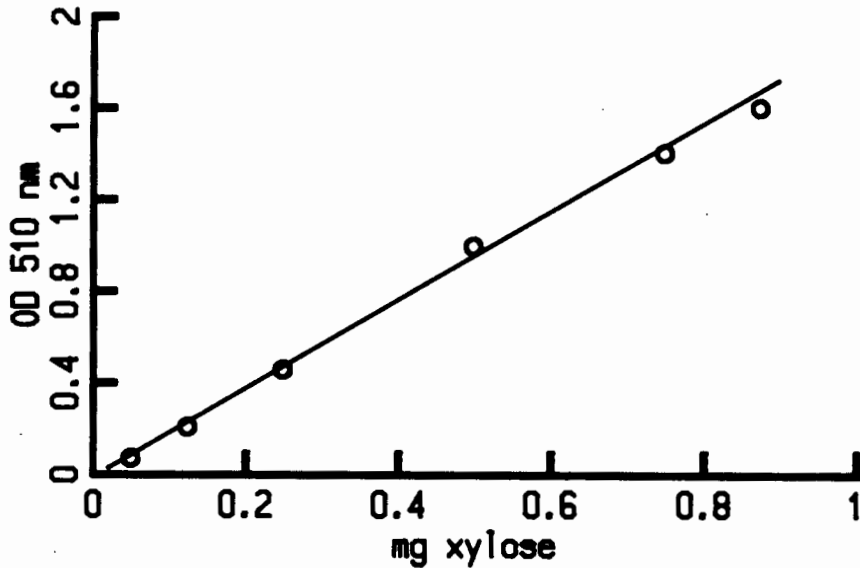
Appendix F

One-and three-letter and codes used for amino acids.

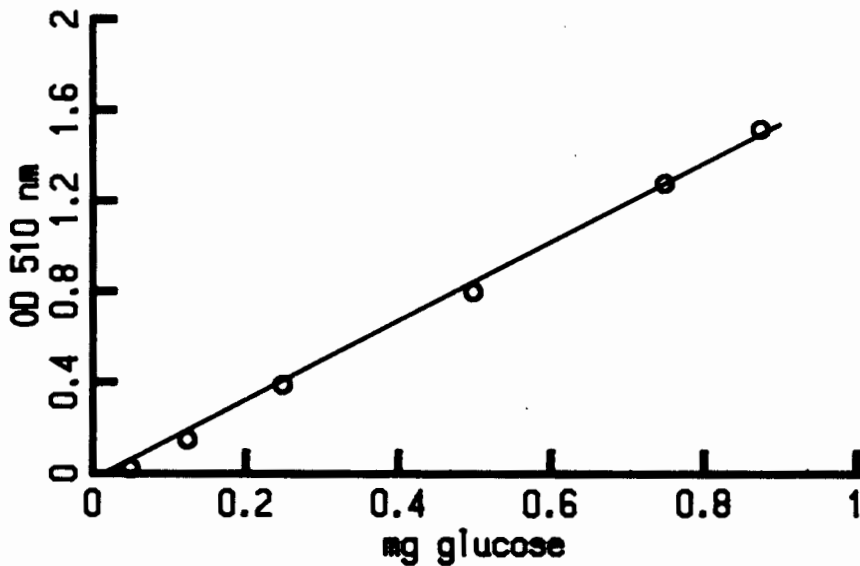
Amino acid	Code	Code	Amino acid
Alanine	Ala A	A	Alanine
Arginine	Arg R	C	Cysteine
Asparagine	Asn N	D	Aspartic acid
Aspartic acid	Asp D	E	Glutamic acid
Cysteine	Cys C	F	Phenylalanine
Glutamine	Gln Q	G	Glycine
Glutamic acid	Glu E	H	Histidine
Glycine	Gly G	I	Isoleucine
Histidine	His H	K	Lysine
Isoleucine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

Appendix G

Standard curves used for the calculation of xylanase and endoglucanase activities.



G.1 Standard curve for the DNS-xylose assay. Assay conditions were as described in Chapter 3 with the exception that the enzyme solution was replaced by 50 μ l of solutions containing varying quantities of xylose. The equation for the least squares regression line was calculated as $y = 1.926X - 0.0086$.



G.2 Standard curve for the DNS-glucose assay. Assay conditions were as described in Chapter 3 with the exception that the enzyme solution was replaced by 50 μ l of solutions containing varying quantities of glucose. The equation for the least squares regression line was calculated as $y = 1.739X - 0.03$.

References

- Ahmed, A. 1984. Plasmid vectors for positive galactose-resistant selection of cloned DNA in Escherichia coli. *Gene* 28:37-43.
- Akin, D. E. 1980. Evaluation by electron microscopy and anaerobic culture of types of rumen bacteria associated with digestion of forage cell walls. *Appl. Environ. Microbiol.* 39:242-252.
- Alberts, B., D. Bray, V. Lewis, M. Raff, K. Roberts, and J. D. Watson. 1983. *Molecular Biology of the Cell*. Garland, New York.
- Allcock, E. R., S. J. Reid, D. T. Jones, and D. R. Woods. 1982. Clostridium acetobutylicum protoplast formation and regeneration. *Appl. Environ. Microbiol.* 43:719-721.
- Allcock, E. R., and D. R. Woods. 1981. Carboxymethyl cellulase and cellobiase production by Clostridium acetobutylicum in an industrial fermentation medium. *Appl. Environ. Microbiol.* 41:539-541.
- Almin, K. E., K-E. Eriksson, and B. Pettersson. 1975. Extracellular enzyme system utilised by the fungus Sporotrichum pulverulentum (Chrysosporium lignorum) for the breakdown of cellulose. *Eur. J. Biochem.* 51:207-211.
- Alurralde, J. L., and G. Ellenrieder. 1984. Effect of attached carbohydrates on the activity of Trichoderma viride cellulases. *Enzyme Microb. Technol.* 6:467-470.
- Amer, G. I., and S. W. Drew. 1980. Microbiology of lignin degradation. *Annu. Rep. Fermentat. Processes* 4:67-103.
- Ames, G. F. L., C. Prody, and S. Kustu. 1984. Simple, rapid, and quantitative release of periplasmic proteins by chloroform. *J. Bacteriol.* 160:1181-1183.
- Armentrout, R. W., R. D. Brown. 1981. Molecular cloning of genes for cellobiose utilization and their expression in Escherichia coli. *Appl. Environ. Microbiol.* 41:1355-1362.
- Aspinall, G. O. 1970. *Polysaccharides*. Pergamon Press, Oxford.
- Avgerinos, G. C., and D. I. C. Wang. 1980. Direct microbial conversion of cellulose to ethanol. *Annu. Rep. Fermentat. Processes* 4:165-191.
- Bahl, H., W. Andersch, and G. Gottschalk. 1982. Continuous production of acetone and butanol by Clostridium acetobutylicum in a two-stage phosphate limited chemostat. *Eur. J. Appl. Microbiol. Biotechnol.* 15:201-205.

- Bailey, M. J. 1981. The effect of β -glucosidase on some assays for cellulolytic enzymes. *Biotechnol. Lett.* 3:695-700.
- Bailey, R. W., and J. C. MacRae. 1970. The hydrolysis by rumen and caecal microbial enzymes of hemicellulose in plant and digesta particles. *J. Agric. Sci.* 75:321-326.
- Barras, F., M. H. Boyer, J. P. Chambost, and M. Chippaux. 1984. Construction of a genomic library of Erwinia chrysanthemi and molecular cloning of cellulase gene. *Mol. Gen. Genet.* 197:513-514.
- Bayer, E. A., and R. Lamed. 1986. Ultrastructure of the cell surface cellulosome of Clostridium thermocellum and its interaction with cellulose. *J. Bacteriol.* 167:828-836.
- Beggs, J. D. 1978. Transformation of yeast by a replicating hybrid plasmid. *Nature* 275:104-108.
- Beguin, P. 1983. Detection of cellulase activity in polyacrylamide gels using Congo red-stained agar replicas. *Anal. Biochem.* 131:333-336.
- Beguin, P., P. Cornet, and J-P. Aubert. 1985. Sequence of a cellulase gene of the thermophilic bacterium Clostridium thermocellum. *J. Bacteriol.* 162:102-105.
- Berenger, J. F., C. Frixon, J. Bigliardi, and N. Creuzet. 1985. Production, purification, and properties of thermostable xylanase from Clostridium stercorarium. *Can. J. Microbiol.* 31:635-643.
- Berghem, L. E. R., L. G. Pettersson, and U. B. Axio-Fredriksson. 1975. The mechanism of enzymatic cellulose degradation. Characterization and enzymatic properties of a β -1,4-glucan cellobiohydrolase from Trichoderma viride. *Eur. J. Biochem.* 53:55-62.
- Bernier, R., and M. Desrochers. 1985. Molecular cloning of a β -xylosidase gene from Bacillus subtilis. *J. Gen. Appl. Microbiol.* 31:513-518.
- Bernier, R. Jr, H. Driguez, and M. Desrochers. 1983. Molecular cloning of a Bacillus subtilis xylanase gene in Escherichia coli. *Gene* 26:59-65.
- Bernier, R. Jr, D. Rho, Y. Arcand, and M. Desrochers. 1985. Partial characterization of an Escherichia coli strain harboring a xylanase encoding plasmid. *Biotechnol. Lett.* 7:797-802.
- Biely, P. 1985. Microbial xylanolytic systems. *Trends Biotechnol.* 3:286-290.

- Biely, P., Z. Kratky, A. Kockova-Kratochvilova, and S. Bauer. 1978. Xylan-degrading activity in yeasts: Growth in xylose, xylan and hemicelluloses. *Folia Microbiol.* 23:366-371.
- Biely, P., C. R. MacKenzie, J. Puls, and H. Schneider. 1986. Cooperativity of esterases and xylanases in the enzymatic degradation of acetyl xylan. *Bio/Technol.* 4:731-733.
- Biely, P., O. Markovic, and D. Mislovicova. 1985a. Sensitive detection of endo-1,4- β -glucanases and endo-1,4- β -xylanases in gels. *Anal. Biochem.* 144:147-151.
- Biely, P., D. Mislovicova, and R. Toman. 1985b. Soluble chromogenic substrates for the assay of endo-1,4- β -xylanases and endo-1,4- β -glucanases. *Anal. Biochem.* 144:142-146.
- Biely, P., and E. Petrakova. 1984. Novel inducers of the xylan-degrading enzyme system of Cryptococcus albidus. *J. Bacteriol.* 160:408-412.
- Biely, P., M. Vrsanska, and Z. Kratky. 1981. Mechanisms of substrate digestion by endo-1,4- β -xylanase of Cryptococcus albidus. Lysozyme-type pattern of action. *Eur. J. Biochem.* 119:565-571.
- Bisaria, V. S., and T. K. Ghose. 1981. Biodegradation of cellulosic materials: substrates, microorganisms, enzymes and products. *Enzyme Microb. Technol.* 3:90-104.
- Blake, C. C. F., L. N. Johnson, G. A. Mair, A. C. T. North, D. C. Phillips, and V. R. Sarma. 1967. Crystallographic studies of the activity of hen egg-white lysozyme. *Proc. Roy. Soc. Lond.* B167:378-388.
- Blaschek, H. P., and M. A. Klacik. 1984. Role of DNase in recovery of plasmid DNA from Clostridium perfringens. *Appl. Environ. Microbiol.* 48:178-181.
- Blotkamp, P. J., M. Takagi, M. S. Pemberton, and G. H. Emert. 1978. Enzymatic hydrolysis of cellulose and simultaneous fermentation to alcohol. *Biochem. Eng.* 74:85-90.
- Bochner, B. R., H. C. Huang, G. L. Schieven, and B. N. Ames. 1980. Positive selection for loss of tetracycline resistance. *J. Bacteriol.* 143:926-932.
- Bolivar, F., R. L. Rodriguez, P. J. Greene, M. C. Betlach, H. L. Heyneker, H. W. Boyer, J. H. Crosa, and S. Falkow. 1977. Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. *Gene* 2:95-113.

- Borriss, R., H. Baumlein, and J. Hofemeister. 1985. Expression in Escherichia coli of a cloned β -glucanase gene from Bacillus amyloliquefaciens. Appl. Microbiol. Biotechnol. 22:63-71.
- Botterman, J., and M. Zabeau. 1985. High-level production of the EcoRI endonuclease under the control of the P_L promoter of bacteriophage lambda. Gene 37:229-239.
- Bowring, S. N., and J. G. Morris. 1985. Mutagenesis of Clostridium acetobutylicum. J. Appl. Bacteriol. 58:577-584.
- Boyer, H. W., and D. Roulland-Dussoix. 1969. A complementation analysis of the restriction and modification of DNA in Escherichia coli. J. Mol. Biol. 41:459-472.
- Boyer, M-H., B. Cami, J. P. Chambost, M. Magnan, and J. Cattaneo. 1987. Characterization of a new endoglucanase from Erwinia chrysanthemi. Eur. J. Biochem. 162:311-316.
- Breuil, C., and J. N. Saddler. 1985. Comparison of the 3,5-dinitrosalicylic acid and Nelson-Somogyi methods of assaying for reducing sugars and determining cellulase activity. Enzyme Microb. Technol. 7:327-332.
- Buchholtz, S. E., M. M. Dooley, and D. E. Eveleigh. 1987. Zymomonas - an alcoholic enigma. Trends Biotechnol. 5:199-204.
- Burns, D. M., and I. R. Beacham. 1984. Positive selection vectors: a small plasmid vector useful for the direct selection of Sau3A-generated overlapping DNA fragments. Gene 27:323-325.
- Canevascini, G. 1985. A cellulase assay coupled to cellobiose dehydrogenase. Anal. Biochem. 147:419-427.
- Cantwell, B. A., and D. J. McConnell. 1983. Molecular cloning and expression of a Bacillus subtilis β -glucanase gene in Escherichia coli. Gene 23:211-219.
- Cauchon, N., and A. LeDuy. 1984. Effect of dilution on carboxymethylcellulase and xylanase assays. Biotechnol. Bioeng. 26:988-991.
- Chanzy, H., and B. Henrissat. 1985. Undirectional degradation of Valonia cellulose microcrystals subjected to cellulase action. FEBS Lett. 184:285-288.
- Chanzy, H., B. Henrissat, R. Vuong, and M. Schulein. 1983. The action of 1,4- β -D-glucan cellobiohydrolase on Valonia cellulose microcrystals. An electron microscope study. FEBS Lett. 153:113-117.
- Chen, C. M., M. Gritzali, and D. W. Stafford. 1987. Nucleotide sequence and deduced primary structure of cellobiohydrolase II from Trichoderma reesei. Bio/Technol. 5:274-278.

- Cheng, S. C., and P. Modrich. 1983. Positive-selection cloning vehicle useful for overproduction of hybrid proteins. *J. Bacteriol.* 154:1005-1008.
- Clark, A. J., and A. D. Margulies. 1965. Isolation and characterization of recombination-deficient mutants of Escherichia coli K12. *Proc. Natl. Acad. Sci. (USA)* 53:451-459.
- Clarke, A. J., and M. Yaguchi. 1985. The role of carboxyl groups in the function of endo- β -1,4-glucanase from Schizophyllum commune. *Eur. J. Biochem.* 149:233-238.
- Clarke, L., and J. Carbon. 1976. A colony bank containing synthetic ColEI hybrid plasmids representative of the entire Escherichia coli genome. *Cell* 9:91-99.
- Collmer, A., and D. B. Wilson. 1983. Cloning and expression of a Thermomonospora YX endocellulase gene in E. coli. *Bio/Technol.* 1:594-601.
- Compere, A. L., and W. L. Griffith. 1979. Evaluation of substrates for butanol production. *Dev. Ind. Microbiol.* 20:509-517.
- Cornet, P., D. Tronik, J. Millet, and J-P. Aubert. 1983. Cloning and expression in Escherichia coli of Clostridium thermocellum genes coding for amino acid synthesis and cellulose hydrolysis. *FEMS Microbiol. Lett.* 16:137-141.
- Costerton, J. W., R. T. Irvin, and K-J. Cheng. 1981. The bacterial glycocalyx in nature and disease. *Annu. Rev. Microbiol.* 35:299-324.
- Coughlan, M. P. 1985a. Cellulose hydrolysis: the potential, the problems and relevant research at Galway. *Biochem. Soc. Trans.* 13:405-406.
- Coughlan, M. P. 1985b. The properties of fungal and bacterial cellulases with comment on their production and application. *Biotechnol. Genetic Engng. Rev.* 3:39-109.
- Coughlan, M. P., K. Hon-Nami, H. Hon-Nami, L. G. Ljungdahl, J. J. Paulin, and W. E. Rigsby. 1985. The cellulytic enzyme complex of Clostridium thermocellum is very large. *Biochem. Biophys. Res. Comm.* 130:904-909.
- Crawford, D. L. 1978. Lignocellulose decomposition by selected Streptomyces strains. *Appl. Environ. Microbiol.* 35:1041-1045.
- Crawford, D. L., and R. L. Crawford. 1980. Microbial degradation of lignin. *Enzyme Microb. Technol.* 2:11-22.

- Crosby, B., B. Collier, D. Y. Thomas, R. M. Teather, and J. D. Erfle. 1984. Cloning and expression in Escherichia coli of cellulase genes from Bacteroides succinogenes. p. 573-576. In: S. Hasnien (ed.), Proc. 5th Can. Bioenergy R and D Seminar. Elsevier Applied Science Publications, Barking, England.
- Cummins, C. S., and J. L. Johnson. 1971. Taxonomy of the clostridia: Wall composition and DNA homologies in Clostridium butyricum and other butyric acid-producing clostridia. J. Gen. Microbiol.:67-33.
- Dagert, M., and S. D. Ehrlich. 1979. Prolonged incubation in calcium chloride improved the competence of Escherichia coli cells. Gene 6:23-28.
- Dale, B. E. 1985. Cellulose pretreatments: Technology and techniques. Annu. Rep. Fermentat. Processes 8:299-323.
- Dale, R. M. K., B. A. McClure, and J. P. Houchins. 1985. A rapid single-stranded cloning strategy for producing a sequential series of overlapping clones for use in sequencing: Application to sequencing the corn mitochondrial 18 S rDNA. Plasmid 13:31-40.
- Dean, D. 1981. A plasmid cloning vector for the direct selection of strains carrying recombinant plasmids. Gene 15:99-102.
- De Boer, H. A., L. J. Comstock, and M. Vasser. 1983. The tac promoter: a functional hybrid derived from the trp and lac promoters. Proc. Natl. Acad. Sci. (USA) 80:21-25.
- Denhardt, D. T. 1966. A membrane filter technique for the detection of complementary DNA. Biochem. Biophys. Res. Comm. 23:641-646.
- Dehority, B. A. 1967. Rate of isolated hemicellulose degradation and utilisation by pure cultures of rumen bacteria. Appl. Microbiol. 15:987-993.
- Dekker, R. F. H. 1985. Biodegradation of the hemicelluloses. p. 505-533. In: Higuchi T. (ed.), Biosynthesis and Biodegradation of Wood Components. Academic Press, New York.
- Dekker, R. F. H., and W. A. Lindner. 1979. Bioutilization of lignocellulosic waste materials: a review. S. Afr. J. Sci. 75:65-71.
- Dekker, R. F. H., and G. N. Richards. 1976. Hemicellulases: their occurrence, purification, properties and mode of action. Adv. Carbohydr. Chem. Biochem. 32:277-352.
- Deshpande, M. V., and K. E. Eriksson. 1984. Reutilisation of enzymes for saccharification of lignocellulosic materials. Enzyme Microb. Technol. 6:338-340.

- Doi, R. H., and L-F. Wang. 1986. Multiple procaryotic ribonucleic acid polymerase sigma factors. *Microbiol. Rev.* 50:227-243.
- Duggleby, C. J., and S. A. Jones. 1983. Cloning and expression of the Staphylococcus aureus protein A gene in Escherichia coli. *Nucleic Acids Res.* 11:3065-3076.
- Dunne, C. P. 1982. Relationship between extracellular proteases and the cellulase complex of Trichoderma reesei. *Enzyme Eng.* 6:355-356.
- Durre, P., A. Kuhn, and G. Gottschalk. 1986. Treatment with allyl alcohol selects specifically for mutants of Clostridium acetobutylicum defective in butanol synthesis. *FEMS Microbiol. Lett.* 36:77-81.
- Ehrlich, S. D., B. Niaudet, and B. Michel. 1982. Use of plasmids from Staphylococcus aureus for cloning of DNA in Bacillus subtilis. *Curr. Top. Microbiol. Immun.* 96:19-30.
- Enari, T-M., and M-L. Niku-Paavola. 1987. Enzymatic hydrolysis of cellulose: is the current theory of the mechanisms of hydrolysis valid?. *CRC Crit. Rev. Biotechnol.* 5:67-87.
- Ennis, B. M., and I. S. Maddox. 1985. Use of Clostridium acetobutylicum P262 for production of solvents from whey permeate. *Biotechnol. Lett.* 7:601-606.
- Erickson, A., and R. H. Deibel. 1973. Turbidometric assay of staphylococcal nuclease. *Appl. Microbiol.* 25:337-341.
- Eriksson, K-E., and B. Pettersson. 1982. Purification and characteristics of two acidic proteases from the white-rot fungus Sporotrichum pulverulentum. *Eur. J. Biochem.* 124:635-642.
- Eveleigh, D. E. 1987. Cellulase: A perspective. *Philos. Trans. R. Soc. Lond.* A321:435-448.
- Fagerstam, L. G., and L. G. Pettersson. 1980. The 1,4- β -glucan cellobiohydrolases of Trichoderma reesei QM 9414. A new type of cellulolytic synergism. *FEBS Lett.* 119:97-100.
- Fagerstam, L. G., L. G. Pettersson, and J. A. Engstrom. 1984. The primary structure of a 1,4-B-glucan cellobiohydrolase from the fungus Trichoderma reesei QM 9414. *FEBS Lett.* 167:309-315.
- Farkas, V., M. Liskova, and P. Biely. 1985. Novel media for the detection of microbial producers of cellulase and xylanase. *FEMS Microbiol. Lett.* 28:137-140.
- Fedorcsak, I., and L. Ehrenberg. 1966. Effects of diethyl pyrocarbonate and methyl methanesulfonate on nucleic acids and nucleases. *Acta Chem. Scand.* 20:107-112.

- Fennington, G., D. Neubauer, and F. Stutzenberger. 1984. Cellulase biosynthesis in a catabolite repression resistant mutant of Thermomonospora curvata. Appl. Environ. Microbiol. 47:201-204.
- Fogarty, W. M., and P. J. Griffin. 1973. Some preliminary observations on the production and properties of a cellulolytic enzyme elaborated by Bacillus polymyxa. Biochem. Soc. Trans. 1:1297-1298.
- Fond, O., E. Petitdemange, H. Petitdemange, and J. M. Engasser. 1983. Cellulose fermentation by a coculture of a mesophilic cellulolytic Clostridium and Clostridium acetobutylicum. Biotechnol. Bioeng. Symp. 13:217-224.
- Forsberg, C. W., T. J. Beveridge, and A. Hellstrom. 1981. Cellulase and xylanase release from Bacteroides succinogenes and its importance in the rumen environment. Appl. Environ. Microbiol. 42:886-896.
- French, A. D. 1978. The crystal structure of native Ramie cellulose. Carbohydr. Res. 61:67-80.
- Fukumori, F., T. Kudo, Y. Narahashi, and K. Horikoshi. 1986a. Molecular cloning and nucleotide sequence of the alkaline cellulase gene from the alkalophilic Bacillus sp. strain 1139. J. Gen. Microbiol. 132:2329-2335.
- Fukumori, F., N. Sashihara, T. Kudo, and K. Horikoshi. 1986b. Nucleotide sequences of two cellulase genes from alkalophilic Bacillus sp. strain N-4 and their strong homology. J. Bacteriol. 168:479-485.
- Fukusaki, E., W. Panbangred, A. Shinmyo, and H. Okada. 1984. The complete nucleotide sequence of the xylanase gene (xynA) of Bacillus pumilus. FEBS Lett. 171:197-201.
- Garen, A., and C. Levinthal. 1960. A fine-structure genetic and chemical study of the enzyme alkaline phosphatase of E. coli. I. Purification and characterization of alkaline phosphatase. Biochim. Biophys. Acta. 38:470-483.
- Gascoigne, J. A., and M. M. Gascoigne. 1960. Biological Degradation of Cellulose. Butterworths, London.
- George, H. A., and J. S. Chen. 1983. Acidic conditions are not obligatory for onset of butanol formation by Clostridium beijerinckii (synonym. C. butylicum). Appl. Environ. Microbiol. 46:321-327.
- Ghose, T. K. 1984. Measurement of cellulase activities. Commission on Biotechnology, International Union of Pure and Applied Chemistry, Indian Institute of Technology, New Delhi.

- Gilbert, I., and G. T. Tsao. 1983. Interaction between solid substrate and cellulase enzymes in cellulose hydrolysis. *Annu. Rev. Ferment. Processes* 6:323-358.
- Gilkes, N. R., D. G. Kilburn, M. L. Langsford, R. C. Miller Jr., W. W. Wakarchuk, R. A. J. Warren, D. J. Whittle, and W. K. R. Wong. 1984a. Isolation and characterisation of Escherichia coli clones expressing cellulase genes from Cellulomonas fimi. *J. Gen. Microbiol.* 130:1377-1384.
- Gilkes, N. R., M. L. Langsford, D. G. Kilburn, R. C. Miller Jr., and R. A. J. Warren. 1984b. Mode of action and substrate specificities of cellulases from cloned bacterial genes. *J. Biol. Chem.* 259:10455-10459.
- Glenn, M., A. Ghosh, and B. K. Ghosh. 1985. Subcellular fractionation of a hypercellulolytic mutant, Trichoderma reesei Rut-C30: Localization of endoglucanase in microsomal fraction. *Appl. Environ. Microbiol.* 50:1137-1143.
- Gong, C-S., M. R. Ladisch, and G. T. Tsao. 1979. Biosynthesis, purification and mode of action of cellulases of Trichoderma reesei. *Adv. Chem. Ser.* 181:261-287.
- Gong, C-S., and G. T. Tsao. 1979. Cellulase and biosynthesis regulation. *Annu. Rep. Fermentat. Processes* 3:111-140.
- Gottwald, M., H. Hippe, and G. Gottschalk. 1984. Formation of n-butanol from D-glucose by strains of "Clostridium tetanomorphum" group. *Appl. Environ. Microbiol.* 48:573-576.
- Graves, M. C., and J. C. Rabinowitz. 1986. In vivo and in vitro transcription of the Clostridium pasteurianum ferredoxin gene. *J. Biol. Chem.* 261:11409-11415.
- Gray, H. B., D. A. Ostrander, J. L. Hodnett, R. J. Legerski, and D. L. Robberson. 1975. Extracellular nucleases of Pseudomonas BAL-31. I. Characterisation of single strand-specific deoxyriboendonuclease and double-strand deoxyriboexonuclease activities. *Nucleic Acids Res.* 2:1459-1492.
- Grepinet, O., and P. Beguin. 1986. Sequence of the cellulase gene of Clostridium thermocellum coding for endoglucanase B. *Nucleic Acids Res.* 14:1791-1799.
- Grethlein, H. E. 1985. The effect of pore size distribution on the rate of enzymatic hydrolysis of cellulosic substrates. *Bio/Technol.* 3:155-160.
- Griffin, H., F. R. Dintzis, L. Krull, and F. L. Baker. 1984. A microfibril generating factor from the enzyme complex of Trichoderma reesei. *Biotechnol. Bioeng.* 26:296-300.

- Gritzali, M., and R. D. Brown. 1979. The cellulase system of Trichoderma. Relationships between purified extracellular enzymes from induced or cellulose-grown cells. *Adv. Chem. Ser.* 181:237-260.
- Groleau, D., and C. W. Forsberg. 1983. Partial characterization of the extracellular carboxymethylcellulase activity produced by the rumen bacterium Bacteroides succinogenes. *Can. J. Microbiol.* 29:504-517.
- Grutter, M. G., L. H. Weaver, and B. W. Matthews. 1983. Goose lysozyme structure: an evolutionary link between hen and bacteriophage lysozymes? *Nature* 303:828-831.
- Guisseppi, A., B. Cami, J-L. Aymeric, G. Ball, and N. Creuzet. 1988. Homology between endoglucanase Z of Erwinia chrysanthemi and endoglucanases of Bacillus subtilis and alkalophilic Bacillus. *Mol. Microbiol.* In press.
- Hagan, C. E., and G. J. Warren. 1982. Lethality of palindromic DNA and its use in selection of recombinant plasmids. *Gene* 19:147-151.
- Halliwell, G., and M. Griffin. 1973. The nature and mode of action of the cellulolytic component C₁ of Trichoderma koningii on native cellulose. *Biochem. J.* 135:587-594.
- Harwood, C. R., D. M. Williams, and P. S. Lovett. 1983. Nucleotide sequence of a Bacillus pumilus gene specifying chloramphenicol acetyltransferase. *Gene* 24:163-169.
- Hayward, A. C. 1977. Occurrence of glycoside hydrolases in plant pathogenic and related bacteria. *J. Appl. Bacteriol.* 43:407-412.
- Hennecke, H., I. Gunther, and F. Binder. 1982. A novel cloning vector for the direct selection of recombinant DNA in E. coli. *Gene* 19:231-234.
- Henrissat, B., H. Driguez, C. Viet, and M. Schulein. 1985. Synergism of cellulases from Trichoderma reesei in the degradation of cellulose. *Bio/Technol.* 3:722-726.
- Hinchliffe, E. 1984. Cloning and Expression of a Bacillus subtilis endo-1,3-1,4- β -D-glucanase gene in Escherichia coli K12. *J. Gen. Microbiol.* 130:1285-1291.
- Hinnen, A., J. B. Hicks, and G. R. Fink. 1978. Transformation of yeast. *Proc. Natl. Acad. Sci. (USA)* 75:1929-1933.
- Hofemeister, J., A. Kurtz, R. Borriss, and J. Knowles. 1986. The β -glucanase gene from Bacillus amyloliquefaciens shows extensive homology with that of Bacillus subtilis. *Gene* 49:177-187.

- Hohn, B. 1979. In vitro packaging of lambda and cosmid DNA. *Methods Enzymol.* 68:209-309.
- Hohn, B., and J. Collins. 1980. A small cosmid for efficient cloning of large DNA fragments. *Gene* 11:291-298.
- Holt, R. A., G. M. Stephens, and J. G. Morris. 1984. Production of solvents by Clostridium acetobutylicum cultures maintained at neutral pH. *Appl. Environ. Microbiol.* 48:1166-1170.
- Honda, H., T. Kudu, and K. Horikoshi. 1985. Molecular cloning and expression of the xylanase gene of alkalophilic Bacillus sp. strain C-125 in Escherichia coli. *J. Bacteriol.* 161:784-785.
- Honda, S. 1984. High performance liquid chromatography of mono and oligosaccharides. *Anal. Biochem.* 140:1-47.
- Hongo, M., A. Murata, K. Kono, and F. Kato. 1968. Lysogeny and bacteriocinogeny in strains of Clostridium species. *Agric. Biol. Chem.* 32:27-33.
- Horikoshi, K., M. Nakao, Y. Kurono, and N. Sashihara. 1984. Cellulases of an alkalophilic Bacillus strain isolated from soil. *Can. J. Microbiol.* 30:774-779.
- Howard, B. H., G. Jones, and M. R. Purdom. 1960. The pentosanases of some rumen bacteria. *Biochem. J.* 74:173-182.
- Howell, S. H. 1982. Plant molecular vehicles: Potential vectors for introducing foreign DNA into plants. *Ann. Rev. Plant Physiol.* 33:609-650.
- Huang, L., L. N. Gibbins, and C. W. Forsberg. 1985. Transmembrane pH gradient and membrane potential in Clostridium acetobutylicum during growth under acetogenic and solventogenic conditions. *Appl. Environ. Microbiol.* 50:1043-1047.
- Huang, J. S., and J. Tang. 1976. Sensitive assay for cellulase and dextranase. *Anal. Biochem.* 73:369-377.
- Ish-Horowicz, D., and J. F. Burke. 1981. Rapid and efficient cosmid cloning. *Nucleic Acids Res.* 9:2989-2998.
- Ishihara, T. 1980. p. 17-27. T. K. Kirk, T. Higuchi and H. Chang (eds.), *Lignin Biodegradation: Microbiology, Chemistry, and potential applications.* Vol. 2. CRC Press Boca Raton, Florida.
- Iwasaki, A., H. Kishida, and M. Okanishi. 1986. Molecular cloning of a xylanase gene from Streptomyces sp. No. 36a and its expression in Streptomycetes. *J. Antibiot.* 39:985-993.

- Janssen, P. J., W. A. Jones, D. T. Jones, and D. R. Woods. 1988. Molecular analysis and regulation of the glnA gene of the Gram-positive anaerobe Clostridium acetobutylicum. J. Bacteriol. 170:400-408.
- Joglekar, A. V., N. G. Karanth, and M. C. Srinivasan. 1983. Significance of β -D-glucosidase in the measurement of exo- β -D-glucanase activity of cellulolytic fungi.. Enzyme Microb. Technol. 5:25-29.
- Johnson, E. A., M. Sakojo, G. Halliwell, A. Madia, and A. L. Demain. 1982. Saccharification of complex cellulosic substrates by the cellulase system from Clostridium thermocellum. Appl. Environ. Microbiol. 43:1125-1132.
- Joliff, G. J., P. Beguin, and J-P. Aubert. 1986. Nucleotide sequence of the cellulase gene celD encoding endoglucanase D of Clostridium thermocellum. Nucleic Acids Res. 14:8605-8613.
- Jones, D. T., W. A. Jones, and D. R. Woods. 1985. Production of recombinants after protoplast fusion in Clostridium acetobutylicum P262. J. Gen. Microbiol. 131:1213-1216.
- Jones, D. T., A. van der Westhuizen, S. Long, E. R. Allcock, S. J. Reid, and D. R. Woods. 1982. Solvent production and morphological changes in Clostridium acetobutylicum. Appl. Environ. Microbiol. 43:1434-1439.
- Jones, D. T., and D. R. Woods. 1986. Acetone-butanol fermentation revisited. Microbiol. Rev. 50:484-524.
- Kammel, W. G., and C. P. Kubicek. 1985. Absence of postsecretional modification in extracellular proteins of Trichoderma reesei during growth on cellulose. J. Appl. Biochem. 7:138-144.
- Kauri, T., and D. J. Kushner. 1985. Role of contact in bacterial degradation of cellulose. FEMS Microbiol. Ecol. 31:301-306.
- Kemp, D. J., and A. F. Cowman. 1981. Direct immunoassay for detecting Escherichia coli colonies that contain polypeptides encoded by cloned DNA segments. Proc. Natl. Acad. Sci. (USA) 78:4520-4524.
- Khan, A. W., D. Tremblay, and A. LeDuy. 1986. Assay of xylanase and xylosidase activities in bacterial and fungal cultures. Enzyme Microb. Technol. 8:373-377.
- Kiel, J. A. K. W., J. P. M. J. Vossen, and G. Venema. 1987. A general method for the construction of E. coli mutants by homologous recombination and plasmid segregation. Mol. Gen. Genet. 207:294-301.

- Kieser, T., D. A. Hopwood, H. M. Wright, and C. J. Thompson. 1982. pIJ101, a multicopy broad host range Streptomyces plasmid: functional analysis and development of DNA cloning vectors. *Mol. Gen. Genet.* 185:223-238.
- Klyosov, A. A. 1986. Enzymatic conversion of cellulosic materials to sugars and alcohol. *Appl. Biochem. Biotechnol.* 12:249-300.
- Knowles, J., P. Lehtovaara, and T. Teeri. 1987. Cellulase families and their genes. *Trends Biotechnol.* 5:255-261.
- Koenigs, J. W. 1975. Hydrogen peroxide and iron: A microbial cellulolytic system. *Biotechnol. Bioeng. Symp.* 5:151-159.
- Kohchi, C., and A. Toh-e. 1985. Nucleotide sequence of Candida pelliculosa β -glucosidase gene. *Nucleic Acids Res.* 13:6273-6282.
- Koide, Y., A. Nakamura, T. Uozumi, and T. Beppu. 1986. Molecular cloning of a cellulase gene from Bacillus subtilis and its expression in Escherichia coli. *Agric. Biol. Chem.* 50:233-237.
- Kotoujansky, A., A. Diolez, M. Boccara, Y. Bertheau, T. Andro, and A. Coleno. 1985. Molecular cloning of Erwinia chrysanthemi pectinase and cellulase structural genes. *EMBO J.* 4:781-785.
- Kubicek, C. P. 1981. Release of carboxymethyl-cellulase and beta-glucosidase from cell walls of Trichoderma reesei. *Eur. J. Appl. Microbiol. Biotechnol.* 13:226-231.
- Kudo, H., K-J. Cheng, and J. W. Costerton. 1987. Electron microscopic study of the methylcellulose-mediated detachment of cellulolytic rumen bacteria from cellulose fibres. *Can. J. Microbiol.* 33:267-272.
- Kudo, T., A. Ohkoshi, and K. Horikoshi. 1985. Molecular cloning and expression of a xylanase gene of alkalophilic Aeromonas sp. No 212 in Escherichia coli. *J. Gen. Microbiol.* 131:2825-2830.
- Kuhn, I., F. H. Stephenson, H. W. Boyer, and P. J. Greene. 1986. Positive-selection vectors utilizing lethality of the EcoRI endonuclease. *Gene* 44:253-263.
- Kushner, S. R. 1978. An improved method for transformation of Escherichia coli with ColE1-derived plasmids. p. 17. In: H. B. Boyer and S. Nicosia (eds.), *Genetic Engineering*. Elsevier, Amsterdam.

- Labavitch, J. M., L. E. Freeman, and P. Albersheim. 1976. Structure of plant cell walls: Purification and characterisation of a β -1,4-galactanase which degrades a structural component of the primary cell walls of dicots. *J. Biol. Chem.* 251:5904-5910.
- Labudova, I., and V. Farkas. 1983. Multiple enzyme forms in the cellulase system of Trichoderma reesei during its growth on cellulose. *Biochim. Biophys. Acta.* 744:135-140.
- Ladisch, M. R., K. W. Lin, M. Voloch, and G. T. Tsao. 1983. Process considerations in the enzymatic hydrolysis of biomass. *Enzyme Microb. Technol.* 5:82-100.
- Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227:680-685.
- Lamed, R., and E. A. Bayer. 1986. Contact and cellulolysis in Clostridium thermocellum via extensile surface organelles. *Experientia* 42:72-73.
- Lamed, R., R. Kenig, E. Setter, and E. A. Bayer. 1985. Major characteristics of the cellulolytic system of Clostridium thermocellum coincide with those of the purified cellulosome. *Enzyme Microb. Technol.* 7:37-41.
- Lamed, R., J. Naimark, E. Morgenstern, and E. A. Bayer. 1987. Specialised cell surface structures in cellulolytic bacteria. *J. Bacteriol.* 169:3792-3800.
- Lamed, R., E. Setter, and E. A. Bayer. 1983. Characterization of a cellulose-binding, cellulase-containing complex in Clostridium thermocellum. *J. Bacteriol.* 156:828-836.
- Leedle, J. A. Z., and R. B. Hespell. 1980. Differential carbohydrate media and anaerobic plating techniques in delineating carbohydrate-utilising subgroups in rumen bacterial populations. *Appl. Environ. Microbiol.* 39:709-719.
- Lee, S. F., and C. W. Forsberg. 1987. Isolation and some properties of a β -D-xylosidase from Clostridium acetobutylicum ATCC 824. *Appl. Environ. Microbiol.* 53:651-654.
- Lee, S. F., C. W. Forsberg, and L. N. Gibbins. 1985a. Cellulolytic activity of Clostridium acetobutylicum. *Appl. Environ. Microbiol.* 50:220-228.
- Lee, S. F., C. W. Forsberg, and L. N. Gibbins. 1985b. Xylanolytic activity of Clostridium acetobutylicum. *Appl. Environ. Microbiol.* 50:1068-1076.
- Lee, S. F., C. W. Forsberg, and J. B. Rattray. 1987. Purification and properties of two endoxylanases from Clostridium acetobutylicum ATCC 824. *Appl. Environ. Microbiol.* 53:644-650.

- Legler, G., and E. Bause. 1973. Epoxyalkyl oligo-(1,4)- β -D-glucosides as active site-directed inhibitors of cellulases. *Carbohydr. Res.* 28:45-52.
- Lejeune, A., C. Colson, and D. E. Eveleigh. 1986. Cloning of an endoglucanase gene from Pseudomonas fluorescens var. cellulosa into Escherichia coli and Pseudomonas fluorescens. *J. Ind. Microbiol.* 1:79-86.
- Lemmel, S. A., R. Datta, and J. R. Frankiewicz. 1985. Fermentation of xylan by Clostridium acetobutylicum. *Enzyme Microb. Technol.* 8:217-221.
- Lever, M. 1977. Carbohydrate determination with 4-hydroxybenzoic acid hydrazide (PAHBAH): effect of bismuth on the reaction. *Anal. Biochem.* 81:21-27.
- Lin, Y. L., and H. P. Blaschek. 1984. Transformation of heat-treated Clostridium acetobutylicum protoplasts with pUB110 plasmid DNA. *Appl. Environ. Microbiol.* 48:737-742.
- Little, P. F. R., and S. H. Cross. 1985. A cosmid vector that facilitates restriction enzyme mapping. *Proc. Natl. Acad. Sci. (USA)* 82:3159-3163.
- Long, S., D. T. Jones, and D. R. Woods. 1983. Sporulation of Clostridium acetobutylicum P262 in a defined medium. *Appl. Environ. Microbiol.* 45:1389-1393.
- Long, S., D. T. Jones, and D. R. Woods. 1984. The relationship between sporulation and solvent production in Clostridium acetobutylicum P262. *Biotechnol. Lett.* 6:529-534.
- Lorowitz, W., and D. Clark. 1982. Escherichia coli mutants with a temperature-sensitive alcohol dehydrogenase. *J. Bacteriol.* 152:935-938.
- Love, D. R., and M. B. Streiff. 1987. Molecular cloning of a β -glucosidase gene from an extremely thermophilic anaerobe in Escherichia coli and Bacillus subtilis. *Bio/Technol.* 5:384-387.
- Lowry, O. H., N. J. Rosebrough, A. L. Farr, and R. J. Randall. 1951. Protein measurement with the folin phenol reagent. *J. Biol. Chem.* 193:265-275.
- Lutzen, N. W., M. H. Nielson, K. M. Oxenboell, M. Schulein, and B. Stentebjerg-Olesen. 1983. Cellulases and their application in the conversion of lignocellulose to fermentable sugars. *Philos. Trans. R. Soc. Lond.* B300:283-291.
- MacKay, R. M., S. Baird, M. J. Dove, J. A. Erratt, M. Gines, F. Moranelli, A. Nasim, G. E. Willick, M. Yaguchi, and V. L. Seligy. 1985. Glucanase gene diversity in prokaryotic and eukaryotic organisms. *BioSyst.* 18:279-292.

- MacKay, R. M., A. Lo, G. Willick, M. Zuker, S. Baird, M. Dove, F. Moranelli, and V. Seligy. 1986. Structure of a Bacillus subtilis endo- β -1,4-glucanase gene. *Nucleic Acids Res.* 14:9159-9170.
- Mandels, M. 1975. Microbial sources of cellulase. *Biotechnol. Bioeng. Symp.* 5:81-105.
- Mandels, M., R. Andreotti, and C. Roche. 1976. Measurement of saccharifying cellulase. *Biotechnol. Bioeng. Symp.* 6:21-33.
- Mandels, M., G. L. Miller, and R. W. Slater. 1961. Separation of fungal carbohydrates by starch block zone electrophoresis. *Arch. Biochem. Biophys.* 93:115-121.
- Mandelstam, J., K. McQuillen, and I. Dawes. 1982. *Biochemistry of Bacterial Growth.* Blackwell Scientific Publications, Oxford.
- Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. *Molecular cloning: A laboratory manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- Maniatis, T., R. C. Hardison, E. Lacy, J. Lauer, C. O'Connell, D. Quon, D. K. Sim, and A. Efstratiadis. 1978. The isolation of structural genes from libraries of eukaryotic DNA. *Cell* 15:687-701.
- Manley, R. St.-J. 1964. Fine structure of native cellulose microfibrils. *Nature* 204:1155-1157.
- Manoil, C., and J. Beckwith. 1985. Tnp ϕ A: a transposon probe for protein export signals. *Proc. Natl. Acad. Sci. (USA)* 82:8129-8133.
- Marchal, R., D. Blanchet, and J. P. Vandecasteele. 1985. Industrial optimization of acetone-butanol fermentation: a study of the utilization of Jerusalem artichokes. *Appl. Microbiol. Biotechnol.* 23:92-98.
- Marmur, J. 1961. A procedure for the isolation of deoxyribonucleic acid from microorganisms. *J. Mol. Biol.* 3:208-218.
- Marsden, W. L., P. P. Gray, and N. W. Dunn. 1982. Alternative pathway for glucose production from cellulase using Trichoderma reesei QM9414 cellulase. *Biotechnol. Lett.* 4:589-594.
- Matthews, B. W., M. G. Grutter, W. F. Anderson, and S. J. Remington. 1981. Common precursor of lysozyme of hen egg-white and bacteriophage T4. *Nature* 290:334-335.
- McLaughlin, J. R., C. L. Murray, and J. C. Rabinowitz. 1981. Unique features in the ribosome binding site sequence of the Gram-positive Staphylococcus aureus β -lactamase gene. *J. Biol. Chem.* 256:11283-11291.

- Merivouri, H., K. M. Siegler, J. A. Sands, and B. S. Montenecourt. 1985. Regulation of cellulase biosynthesis and secretion in fungi. *Biochem. Soc. Trans.* 13:411-414.
- Mes-Hartree, M., and J. N. Saddler. 1982. Butanol production of Clostridium acetobutylicum grown on sugars found in hemicellulose hydrolysates. *Biotechnol. Lett.* 4:247-252.
- Messing, J. 1983. New M13 vectors for cloning. *Methods Enzymol.* 101:20-28.
- Messing, J., R. Crea, and P. H. Seeburg. 1981. A system for shotgun DNA sequencing. *Nucleic Acids Res.* 9:309-321.
- Miller, G. L. 1959. Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Anal. Chem.* 31:426-428.
- Miller, J. H. 1972. *Experiments in Molecular Genetics.* Cold Spring Harbor laboratory, Cold Spring Harbor, New York.
- Millet, J., D. Petre, P. Beguin, O. Raynaud, and J-P. Aubert. 1985. Cloning of ten distinct DNA fragments of Clostridium thermocellum coding for cellulases. *FEMS Microbiol. Lett.* 29:145-149.
- Mondou, F., F. Shareck, R. Morosoli, and D. Kluepfel. 1986. Cloning of the xylanase gene of Streptomyces lividans. *Gene* 49:323-329.
- Montenecourt, B. S., and D. E. Eveleigh. 1977. Semiquantitative plate assay for determination of cellulase production by Trichoderma viride. *Appl. Environ. Microbiol.* 33:178-183.
- Montenecourt, B. S., D. E. Eveleigh, G. K. Elmund, and J. Parcells. 1978. Antibiotic disks - an improvement in the filter paper assay for cellulase. *Biotechnol. Bioeng.* 20:297-300.
- Morosoli, R., C. Roy, and M. Yaguchi. 1986. Isolation and partial primary sequence of a xylanase from the yeast Cryptococcus albidus. *Biochim. Biophys. Acta.* 870:473-478.
- Mullings, R., J. H. Parish. 1984. New reducing sugar assay for the study of cellulases. *Enzyme Microb. Technol.* 6:491-496.
- Murphy-Holland, K., and D. E. Eveleigh. 1985. Secretion, activity and stability of deglycosylated cellulase of Trichoderma reesei. p. 193. *In: Proc. 85th Am. Soc. Microbiol. K130.* Am. Soc. Microbiol., Washington DC.

- Murphy, N., D. J. McConnell, and B. A. Cantwell. 1984. The DNA sequence of the gene and genetic control sites for the excreted *B. subtilis* enzyme β -glucanase. *Nucleic Acids Res.* 12:5355-5367.
- Muzyczka, N. 1980. Construction of an SV40 derived cloning vector. *Gene* 11:63-67.
- Nakayama, M., Y. Tomita, H. Suzuki, and K. Nisizawa. 1976. Partial proteolysis of some cellulase components from *Trichoderma viride* and the substrate specificity of the modified products. *J. Biochem. Tokyo* 79:955-966.
- Nelson, N. 1952. A photometric adaptation of the Somogyi method for the determination of glucose. *J. Biol. Chem.* 153:376-380.
- Ng, T. K., A. Ben-Bassat, and J. G. Zeikus. 1981. Ethanol production by thermophilic bacteria: fermentation of cellulosic substrates by cocultures of *Clostridium thermocellum* and *Clostridium thermohydrosulfuricum*. *Appl. Environ. Microbiol.* 41:1337-1343.
- Ng, T. K., and J. G. Zeikus. 1981. Purification and characterization of an endoglucanase (1,4- β -D-glucan glucanohydrolase) from *Clostridium thermocellum*. *Biochem. J.* 199:341-350.
- Nikolnikov, S., G. Posfai, and B. Sain. 1984. The construction of a versatile plasmid vector that allows direct selection of fragments cloned into six unique sites of the cI gene of coliphage 434. *Gene* 30:261-265.
- Niku-Paavola, M-L., A. Lappalainen, T-M. Enari, and M. Nummi. 1985. A new appraisal of the endoglucanases of the fungus *Trichoderma reesei*. *Biochem. J.* 231:75-81.
- Nilsson, B., M. Uhlen, S. Josephson, S. Gatenbeck, and L. Philipson. 1983. An improved positive selection plasmid vector constructed by oligonucleotide mediated mutagenesis. *Nucleic Acids Res.* 11:8019-8030.
- Nummi, M., M-L. Niku-Paavola, T-M. Enari, and V. Raunio. 1981. Isolation of cellulases by means of biospecific sorption on amorphous cellulose. *Anal. Biochem.* 116:137-141.
- Nummi, M., M-L. Niku-Paavola, A. Lappalainen, T-M. Enari, and V. Raunio. 1983. Cellobiohydrolase from *Trichoderma reesei*. *Biochem. J.* 215:677-684.
- O'Brien, R. W., and J. G. Morris. 1971. Oxygen and the growth and metabolism of *Clostridium acetobutylicum*. *J. Gen. Microbiol.* 68:307-318.
- O'Connor, C. D., and G. O. Humphreys. 1982. Expression of the *EcoRI* restriction-modification system and the construction of positive-selection cloning vectors. *Gene* 20:219-229.

- Odier, E., and B. Monties. 1978. Biodegradation of wheat lignin by Xanthomonas 23. Ann. Microbiol. (Paris) 129A:361-378.
- Ogata, S., and M. Hongo. 1979. Bacteriophages of the genus Clostridium. Adv. Appl. Microbiol. 25:241-273.
- O'Neill, G., S. H. Goh, R. A. J. Warren, D. G. Kilburn, R. C. Miller Jr. 1986. Structure of the gene encoding the exoglucanase of Cellulomonas fimi. Gene 44:325-330.
- Oultram, J. D., and M. Young. 1985. Conjugal transfer of plasmid pAMP1 from Streptococcus lactis and Bacillus subtilis to Clostridium acetobutylicum. FEMS Microbiol. Lett. 27:129-134.
- Oultram, J. D., and M. Young. 1987. Conjugal transfer of a small plasmid from Bacillus subtilis to Clostridium acetobutylicum by cointegrate formation with plasmid pAMP1. FEMS Microbiol. Lett. 42:113-119.
- Ounine, K., H. Petitdemange, G. Raval, and R. Gay. 1983. Acetone-butanol production from pentoses by Clostridium acetobutylicum. Biotechnol. Lett. 5:605-610.
- Pahel, G., and B. Tyler. 1979. A new glnA-linked regulatory gene for glutamine synthetase in Escherichia coli. Proc. Natl. Acad. Sci. (USA) 76:4544-4548.
- Paice, M. G., R. Bourbonnais, M. Desrochers, L. Jurasek, and M. Yaguchi. 1986. A xylanase gene from Bacillus subtilis: nucleotide sequence and comparison with B. pumilus gene. Arch. Microbiol. 144:201-206.
- Paice, M. G., M. Desrochers, D. Rho, L. Jurasek, C. Roy, C. F. Rollin, E. De Miguel, and M. Yaguchi. 1984. Two forms of endoglucanase from the basidiomycete Schizophyllum commune and their relationship to other β -1,4-glucoside hydrolases. Bio/Technol. 2:535-539.
- Paice, M. G., and L. Jurasek. 1984. Removing hemicellulose from pulps by specific enzymic hydrolysis. J. Wood Chem. Technol. 4:187-198.
- Panbangred, W., T. Kondo, S. Negoro, A. Shinmyo, and H. Okada. 1983a. Molecular cloning of the genes for xylan degradation of Bacillus pumilis and their expression in Escherichia coli. Mol. Gen. Genet. 192:335-341.
- Panbangred, W., A. Shinmyo, S. Kinoshita, and H. Okada. 1983b. Purification and properties of endoxylanase produced by Bacillus pumilus. Agric. Biol. Chem. 47:957-963.

- Pardee, A. B., F. Jacob, and J. Monod. 1959. The genetic control and cytoplasmic expression of "inducibility" in the synthesis of beta-galactosidase by E. coli. J. Mol. Biol. 1:165-178.
- Park, S. H. P., and M. Y. Pack. 1986. Cloning and expression of a Bacillus cellulase gene in Escherichia coli. Enzyme Microb. Technol. 8:725-728.
- Penttila, M., P. Lehtovaara, H. Nevalainen, R. Bhikhabhai, and J. Knowles. 1986. Homology between cellulase genes of Trichoderma reesei: Complete nucleotide sequence of the endoglucanase I gene. Gene 45:253-263.
- Penttila, M., K. M. H. Nevalainen, A. Raynal, and J. K. C. Knowles. 1984. Cloning of Aspergillus niger genes in yeast. Expression of the gene coding Aspergillus β -glucosidase. Mol. Gen. Genet. 194:494-499.
- Pearlman, D., and H. O. Halvorsen. 1983. A putative signal peptidase recognition site and sequence in eukaryotic and prokaryotic signal peptides. J. Mol. Biol. 167:391-409.
- Petitdemange, E., O. Fond, F. Caillet, H. Petitdemange, and R. Gay. 1983. A novel one step process for cellulose fermentation using mesophilic cellulolytic and glycolytic Clostridia. Biotechnol. Lett. 5:119-124.
- Pettersson, L. G., L. G. Fagerstam, R. Bhikhabhai, and K. Leandroer. 1981. The cellulase complex of Trichoderma reesei QM9414. p. 39-42. In: The Ekman Days International Symposium on Wood Pulping Chemistry Vol. 3. SPCI, Stockholm.
- Poncz, M., D. Solowiejczyk, M. Ballantine, E. Schwartz, and S. Surrey. 1982. "Nonrandom" DNA sequence analysis in bacteriophage M13 by the dideoxy chain-termination method. Proc. Natl. Acad. Sci. (USA) 79:4298-4302.
- Pouwels, P. H., B. E. Enger-Valk, and W. J. Brammar. 1985. Cloning vectors. A laboratory manual. Elsevier, Amsterdam.
- Prescott, S. G., and C. G. Dunn. 1959. Industrial Microbiology. 3rd ed. McGraw-Hill, New York.
- Puls, J., K. Poutanen, H-U. Korner, and L. Viikari. 1985. Biotechnical utilisation of wood carbohydrates after steaming pretreatment. Appl. Microbiol. Biotechnol. 22:416-423.
- Queen, C., and L. J. Korn. 1984. A comprehensive sequence analysis program for the IBM personal computer. Nucleic Acids Res. 12:581-599.
- Rando, R. R. 1974. Allyl alcohol-induced irreversible inhibition of yeast alcohol dehydrogenase. Biochem. Pharmacol. 23:2328-2331.

- Raven, P. H., R. F. Evert, and H. Curtis. 1976. *Biology of Plants*. Worth Publishers, New York.
- Raynal, A., and M. Guerineau. 1984. Cloning and expression of the structural gene for β -glucosidase of Kluyveromyces fragilis in Escherichia coli and Saccharomyces cerevisiae. *Mol. Gen. Genet.* 195:108-115.
- Rees, D. A., E. R. Morris, D. Thom, and J. K. Madden. 1982. Shapes and interactions of carbohydrate chains. p. 195-290. *In: Polysaccharides*. Vol. 1. Academic Press, New York.
- Reese, E. T., A. Maguire, and F. W. Parrish. 1973. Production of β -D-xylopyranosidases by fungi. *Can. J. Microbiol.* 19:1065-1074.
- Reese, E. T., R. G. H. Siu, and H. S. Levinson. 1950. Biological degradation of soluble cellulose derivatives. *J. Bacteriol.* 59:485-491.
- Reid, S. J., E. R. Allcock, D. T. Jones, and D. R. Woods. 1983. Transformation of Clostridium acetobutylicum protoplasts with bacteriophage DNA. *Appl. Environ. Microbiol.* 45:305-307.
- Reilly, P. J. 1981. Xylanases: structure and function. p. 111-129. *In: A. Hollaender, R. Rabson, P. Rogers, A. San Pietro, R. Valentine and R. Wolfe (eds.), Trends in the Biology of Fermentations for Fuels and Chemicals*. Plenum Press, New York.
- Remaut, E., H. Tsao, and W. Fiers. 1983. Improved plasmid vectors with a thermoinducible expression and temperature-regulated runaway replication. *Gene* 22:103-113.
- Reysset, G., J. Hubert, L. Podvin, and M. Sebald. 1987. Protoplast formation and regeneration of Clostridium acetobutylicum strain N1-4080. *J. Gen. Microbiol.* 133:2595-2600.
- Reysset, G., and M. Sebald. 1985. Conjugal transfer of plasmid-mediated antibiotic resistance from streptococci to Clostridium acetobutylicum. *Ann. Inst. Pasteur. Microbiol.* 136B:275-282.
- Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg. 1977. Labelling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. *J. Mol. Biol.* 113:237-251.
- Riske, F. J., D. E. Eveleigh, and J. D. Macmillan. 1986. Purification of cellobiohydrolase I from Trichoderma reesei using monoclonal antibodies in an immunomatrix. *J. Ind. Microbiol.* 1:259-264.

- Roberts, T. M., S. L. Swanberg, A. Poteete, G. Riedel, and K. Backman. 1980. A plasmid cloning vehicle allowing a positive selection for inserted fragments. *Gene*. 12:123-127.
- Robinson, G. C. 1922. A study of the acetone butyl alcohol fermentation of various carbohydrates. *J. Biol. Chem.* 52:125-155.
- Robson, L. M., and G. H. Chambliss. 1986. Cloning of the Bacillus subtilis DLG β -1,4-glucanase gene and its expression in Escherichia coli and B. subtilis. *J. Bacteriol.* 165:612-619.
- Robson, L. M., and G. H. Chambliss. 1987. Endo- β -1,4-glucanase gene of Bacillus subtilis DLG. *J. Bacteriol.* 169:2017-2025.
- Rogers, P. 1986. Genetics and biochemistry of Clostridium relevant to development of fermentation processes. *Adv. Appl. Microbiol.* 31:1-60.
- Roltz, C. 1986a. Ultrasound effect on enzymatic saccharification. *Biotechnol. Lett.* 8:131-136.
- Roltz, C. 1986b. Regulation of adsorption-desorption mechanisms during holocellulosis of biomass. *Trends Biotechnol.* 4:135-136.
- Romaniec, M. P. M., N. G. Clarke, and G. P. Hazlewood. 1987. Molecular cloning of Clostridium thermocellum DNA and the expression of further novel endo- β -1,4-glucanase genes in Escherichia coli. *J. Gen. Microbiol.* 133:1297-1307.
- Rosenberg, S. M., M. M. Stahl, I. Kobayashi, and F. W. Stahl. 1985. Improved in vitro packaging of coliphage lambda DNA: a one-strain system free from endogenous phage. *Gene*. 38:165-175.
- Rybicki, E. P., and M. B. von Wechmar. 1982. Enzyme assisted immune detection of plant virus proteins electroblotted onto nitrocellulose paper. *J. Virol. Meth.* 5:267-278.
- Ryu, D. D. Y., S. B. Lee, T. Tassinari, and C. Macy. 1982. Effect of compression milling on cellulose structure and enzymatic hydrolysis. *Biotechnol. Bioeng.* 24:1047-1067.
- Saddler, J. N. 1986. Factors limiting the efficiency of cellulase enzymes. *Microbiol. Sci.* 3:84-87.
- Saddler, J. N., and M. K-H. Chan. 1982. Optimization of Clostridium thermocellum growth on cellulose and pretreated wood substrates. *Eur. J. Appl. Microbiol. Biotechnol.* 16:99-104.

- Saddler, J. N., and M. K-H. Chan. 1984. Conversion of pretreated lignocellulosic substrates to ethanol by Clostridium thermocellum in mono- and co-culture with Clostridium thermosaccharolyticum and Clostridium thermohydrosulphuricum. *Can. J. Microbiol.* 30:212-220.
- Saddler, J. N., M. K-H. Chan, and G. Louis-Seize. 1981. A one-step process for the conversion of cellulose to ethanol using anaerobic microorganisms in mono- and co-culture. *Biotechnol. Lett.* 3:321-327.
- Sakaguchi, K., 1982. Vectors for gene cloning in Pseudomonas and their applications. *Curr. Topics Microbiol. Immun.* 96:31-46.
- Salyers, A. A., S. E. H. West, J. R. Vercellotti, and T. D. Wilkins. 1977. Fermentation of mucins and plant polysaccharide by anaerobic bacteria from the human colon. *Appl. Environ. Microbiol.* 34:529-533.
- Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. (USA)* 74:5463-5467.
- Sarkanen, K. V., and C. H. Ludwig. 1971. Definition and nomenclature. p. 1. In: K. V. Sarkanen and C. H. Ludwig (eds.), *Lignins: Occurrence, Formation, Structure and Reactions*. Wiley Interscience, New York.
- Sasaki, T., T. Tanaka, N. Nanbu, Y. Sato, and K. Kainuma. 1979. Correlation between X-ray diffraction measurement of cellulose crystalline structure and the susceptibility to microbial cellulase. *Biotechnol. Bioeng.* 21:1031-1042.
- Sashihara, N., T. Kudo, and K. Horikoshi. 1984. Molecular cloning and expression of cellulase genes of alkalophilic Bacillus sp. Strain N-4 in Escherichia coli. *J. Bacteriol.* 158:503-506.
- Savarese, J. J., and S. D. Young. 1978. Combined enzyme hydrolysis of cellulose and yeast fermentation. *Biotechnol. Bioeng.* 20:1291-1293.
- Schmuck, M., I. Pilz, M. Hayn, and H. Esterbauer. 1986. Investigation of cellobiohydrolase from Trichoderma reesei by small angle X-ray scattering. *Biotechnol. Lett.* 8:397-402.
- Schneider, H., G. Mahmoudides, J. L. Labelle, H. Lee, N. Maki, and H. J. McNeill. 1985. Correlation between limitation of growth of Pachysolen tannophilus on D-xylose with the formation of ethanol and other products. *Biotechnol. Lett.* 7:361-364.
- Schumann, W. 1979. Construction of an HpaI and HindIII plasmid vector allowing direct selection of transformants harboring recombinant plasmids. *Mol. Gen. Genet.* 174:221-224.

- Schwarz, W., K. Bronnenmeier, and W. L. Staudenbauer. 1985. Molecular cloning of Clostridium thermocellum genes involved in β -glucan degradation in bacteriophage lambda. *Biotechnol. Lett.* 7:859-864.
- Scopes, R. K. 1982. Protein purification. Principles and Practice. Springer-Verlag, New York.
- Shewale, J. G. 1982. β -glucosidase: its role in cellulase synthesis and hydrolysis of cellulose. *Int. J. Biochem.* 14:435-443.
- Shewale, J. G., and J. C. Sadana. 1981. Purification, characterisation and properties of β -glucosidase enzymes from Sclerotium rolfsii. *Arch. Biochem. Biophys.* 207:185-196.
- Shoemaker, S. P., and R. D. Brown Jr. 1978. Characterisation of endo-1,4- β -D-glucanases purified from Trichoderma viride. *Biochim. Biophys. Acta.* 523:147-161.
- Shoemaker, S., V. Schweickart, M. Ladner, D. Gelfand, S. Kwok, K. Myambo, and M. Innis. 1983. Molecular cloning of exo-cellobiohydrolase I derived from Trichoderma reesei strain L27. *Bio/Technol.* 1:691-696.
- Sipat, A., K. A. Taylor, R. Y. C. Lo, C. W. Forsberg, and P. J. Krell. 1987. Molecular cloning of a xylanase gene from Bacteroides succinogenes and its expression in Escherichia coli. *Appl. Environ. Microbiol.* 53:477-481.
- Smith, G. E., and M. D. Summers. 1980. The bidirectional transfer of DNA and RNA to nitrocellulose or diazobenzoyloxymethyl-paper. *Anal. Biochem.* 109:123-129.
- Sprey, B., and C. Lambert. 1983. Titration curves of cellulases from Trichoderma reesei: demonstration of a cellulase-xylanase- β -glucosidase-containing complex. *FEMS Microbiol. Lett.* 18:217-222.
- Sprey, B., and C. Lambert. 1984. Heterogeneity of cellulase complexes from Trichoderma reesei: a preparative isoelectric focusing study of some extracellular hydrolases. *FEMS Microbiol. Lett.* 23:227-232.
- Stack, R. J., and M. A. Cotta. 1986. Effect of 3-phenylpropanoic acid on growth and cellulose utilisation by cellulolytic ruminal bacteria. *Appl. Environ. Microbiol.* 52:209-210.
- Stack, R. J., and R. E. Hungate. 1984. Effect of 3-phenylpropanoic acid on capsule and cellulases of Ruminococcus albus 8. *Appl. Environ. Microbiol.* 48:218-223.

- Starr, M. P., H. Stolp, H. G. Truper, A. Balows, and H. G. Schlegel. 1981. *The Prokaryotes*. Vol. 1. Springer-Verlag, Berlin.
- Sternberg, D., and G. R. Mandels. 1979. Induction of cellulolytic enzymes in Trichoderma reesei by sophorose. *J. Bacteriol.* 139:761-769.
- Sternberg, N., D. Tiemeier, and L. Enquist. 1977. *In vitro* packaging of a lambda dam vector containing EcoRI DNA fragments of Escherichia coli and phage ϕ 1. *Gene* 1:255-280.
- Stern, M., and G. H. Warrack. 1964. The types of Clostridium perfringens. *J. Pathol. Bacteriol.* 88:279-282.
- Struhl, K. 1985. A rapid method for creating recombinant DNA molecules. *BioTechniques* 3:452-453.
- Swiatek, P. J., S. D. Allen, J. A. Siders, and C. Lee. 1987. DNase production by Clostridium septicum. *J. Clin. Microbiol.* 25:437-438.
- Sykes, R. B., and K. Nordstrom. 1972. Microiodometric determination of β -lactamase activity. *Antimicrob. Agents. Chemother.* 1:94-99.
- Tansey, M. R. 1971. Agar-diffusion assay of cellulolytic ability of thermophilic fungi. *Arch. Microbiol.* 77:1-11.
- Taylor, K. A., B. Crosby, M. McGavin, C. W. Forsberg, and D. Y. Thomas. 1987. Characteristics of the endoglucanase encoded by a cel gene from Bacteroides succinogenes expressed in Escherichia coli. *Appl. Environ. Microbiol.* 53:41-46.
- Teather, R. M., and P. J. Wood. 1982. Use of Congo red-polysaccharide interactions in enumeration and characterization of cellulolytic bacteria from the bovine rumen. *Appl. Environ. Microbiol.* 43:777-780.
- Teeri, T., I. Salovuori, and J. Knowles. 1983. The molecular cloning of the major cellulase gene from Trichoderma reesei. *Bio/Technol.* 1:696-699.
- Teeri, T. T., P. Lehtovaara, S. Kauppinen, I. Salovuori, and J. Knowles. 1987. Homologous domains in Trichoderma reesei cellulolytic enzymes: gene sequence and expression of cellobiohydrolase II. *Gene* 51:43-52.
- Timell, T. E. 1967. Recent progress in the chemistry of wood hemicelluloses. *Wood Sci. Technol.* 1:45-70.
- Timmis, K., and U. Winkler. 1973. Isolation of covalently closed circular deoxyribonucleic acid from bacteria which produce exocellular nuclease. *J. Bacteriol.* 113:508-509.

- Towbin, H., T. Staehelin, and J. Gordon. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. (USA)* 76:4350-4354.
- Tsao, G. T. 1978. Cellulosic material as a renewable resource. *Process Biochem.* 13(10):12-14.
- Tseng, M. C., and J. H. Luong. 1984. Mushroom cultivation - technology for commercial production. *Annu. Rep. Fermentat. Processes* 7:45-79.
- Urano, N., I. Karube, S. Suzuki, T. Yamada, H. Hirochika, and K. Sakaguchi. 1983. Isolation and partial characterization of large plasmids in hydrogen-evolving bacterium *Clostridium butyricum*. *Eur. J. Appl. Microbiol. Biotechnol.* 17:349-354.
- Usdin, K. P., H. Zappe, D. T. Jones, and D. R. Woods. 1985. Cloning and expression of the glutamine synthetase gene from the Gram-positive obligate anaerobe *Clostridium acetobutylicum* in *Escherichia coli*. *Appl. Environ. Microbiol.* 52:413-419.
- van Arsdell, J. N., S. Kwok, V. L. Schweickart, M. B. Ladner, D. H. Gelfand, and M. A. Innis. 1987. Cloning, characterization, and expression in *Saccharomyces cerevisiae* of endoglucanase I from *Trichoderma reesei*. *Bio/Technol.* 5:60-64.
- van Gijsegem, F., A. Toussaint, and E. Schoonejans. 1985. In vivo cloning of the pectate lyase and cellulase genes of *Erwinia chrysanthemi*. *EMBO J.* 4:787-792.
- van Tilbeurgh, H., R. Bhikhabhai, L. G. Pettersson, and M. Claeysens. 1984. Separation of endo- and exo-type cellulases using a new affinity chromatography method. *FEBS Lett.* 169:215-218.
- van Tilbeurgh, H., M. Claeysens, and C. K. De Bruyne. 1982. The use of 4-methylumbelliferyl and other chromophoric glycosides in the study of cellulolytic enzymes. *FEBS Lett.* 149:152-156.
- van Tilbeurgh, H., L. G. Pettersson, R. Bhikhabhai, H. de Boeck, and M. Claeysens. 1985. Studies on the cellulolytic system of *Trichoderma reesei* QM9414. Reaction specificity and thermodynamics of interactions of small substrates and ligands with the 1,4- β -glucan cellobiohydrolase II.. *Eur. J. Biochem.* 148:329-334.
- Veal, D. A., and J. M. Lynch. 1984. Associative cellulolysis and denitrogen fixation by co-cultures of *Trichoderma harzianum* and *Clostridium butyricum*. *Nature* 310:695-697.

- Vernet, T., P. C. K. Lau, S. A. Narang, and L. P. Visentin. 1985. A direct-selection vector derived from pColE3-CA38 and adapted for foreign expression. *Gene* 34:87-93.
- Vernon, C. A. 1967. The mechanisms of hydrolysis of glycosides and their relevance to enzyme catalysed reactions. *Proc. Roy. Soc. Lond.* B167:389-401.
- Vernon, C. A., and B. Banks. 1963. The enzyme hydrolysis of glycosides. *Biochem. J.* 86:7.
- Vogeli, G., E. V. Avedimento, M. Sullivan, J. V. Jr. Maizel, G. Lazano, S. L. Adams, L. Pastan, and B. de Crombrugge. 1980. Isolation and characterisation of genomic DNA coding for $\alpha 2$ type I collagen. *Nucleic Acids Res.* 8:1823-1837.
- Voget, C. E., C. F. Mignone, and R. J. Ertola. 1985. Butanol production from apple pomace. *Biotechnol. Lett.* 7:43-46.
- Wakarchuk, W. W., D. C. Kilburn, R. C. Miller Jr., and R. A. J. Warren. 1986. The molecular cloning and expression of a cellobiase gene from Agrobacterium in Escherichia coli. *Mol. Gen. Genet.* 205:146-152.
- Waldron, C. R. Jr., C. A. Becker-Vallone, and D. E. Eveleigh. 1986. Isolation and characterisation of a cellulolytic actinomycete Microbispora bispora. *Appl. Microbiol. Biotechnol.* 24:477-486.
- Waldron, C. R. Jr., and D. E. Eveleigh. 1986. Saccharification of cellulose by Microbispora bispora. *Appl. Microbiol. Biotechnol.* 24:487-492.
- Waterbury, J. B., C. B. Calloway, and R. D. Turner. 1983. A cellulolytic nitrogen-fixing bacterium cultured from the gland of *Deshayes* in ship worms (Bivalve: *Teredinidae*). *Science* 221:1401-1403.
- Weiss, B., A. Jacquemin-Sablon, T. R. Live, G. C. Fareed, and C. C. Richardson. 1968. Enzymatic breakage and joining of deoxyribonucleic acid. VI. Further purification and properties of polynucleotide ligase from Escherichia coli infected with bacteriophage T4. *J. Biol. Chem.* 243:4543-4555.
- White, A. R., and R. M. Brown Jr. 1981. Enzymatic hydrolysis of cellulose: visual characterisation of the process. *Proc. Natl. Acad. Sci. (USA)* 78:1047-1051.
- Whittle, D. J., D. G. Kilburn, R. A. J. Warren, and R. C. Miller Jr. 1982. Molecular cloning of a Cellulomonas fimi cellulase gene in Escherichia coli. *Gene* 17:139-145.
- Wiegel, J. 1982. Ethanol from cellulose. *Experientia* 38:151-156.

- Wiley, A. S., M. J. Johnson, E. McCoy, and W. H. Peterson. 1941. The fermentation of waste sulfite liquor. *Ind. Eng. Chem.* 35:606-610.
- Wilke, C. R., R. D. Yang, and U. Von Stockar. 1976. Preliminary case analysis for enzymatic hydrolysis of newsprint. *Biotechnol. Bioeng. Symp.* 6:155-176.
- Wilkie, K. C. B. 1979. The hemicelluloses of grasses and cereals. *Adv. Carbohydr. Chem. Biochem.* 36:215-264.
- Wilkie, K. C. B. 1983. Hemicellulose. *Chem. Technol.* 3:306-319.
- Williams, A. G. 1983. Staining reactions for the detection of hemicellulose-degrading bacteria. *FEMS Microbiol. Lett.* 20:253-258.
- Williams, B. G., and F. R. Blattner. 1979. Construction and characterisation of the hybrid bacteriophage lambda Charon vectors for DNA cloning. *J. Virol.* 29:555-575.
- Willick, G. E., and V. L. Seligy. 1985. Multiplicity in cellulases of Schizophyllum commune. Derivation partly from heterogeneity in transcription and glycosylation. *Eur. J. Biochem.* 151:89-96.
- Willis, R. C., R. G. Morris, C. Cirakoglu, G. D. Schellenberg, N. H. Gerber, and C. E. Furlong. 1974. Preparation of the periplasmic binding proteins from Salmonella typhimurium and Escherichia coli. *Arch. Biochem. Biophys.* 161:64-75.
- Wirick, M. G. 1968a. A study of the enzymic degradation of CMC and other cellulose ethers. *J. Polymer Sci.* 6(A-1):1965-1974.
- Wirick, M. G. 1968b. Study of the substitution pattern of hydroxyethylcellulose and its relationship to enzymic degradation. *J. Polymer Sci.* 6(A-1):1705-1718.
- Woese, C. R. 1987. Bacterial evolution. *Microbiol. Rev.* 51:221-271.
- Wolff, B. R., T. A. Mudry, B. R. Glick, and J. J. Pasternak. 1986. Isolation of endoglucanase genes from Pseudomonas fluorescens subsp. cellulosa and a Pseudomonas sp. *Appl. Environ. Microbiol.* 51:1367-1369.
- Wong, W. K., B. Gerhard, Z. M. Guo, D. G. Kilburn, R. Anthony, J. Warren, and R. C. Miller Jr. 1986. Characterization and structure of an endoglucanase gene cenA of Cellulomonas fimi. *Gene* 44:315-324.
- Wong, S-L., C. W. Price, D. S. Goldfarb, and R. H. Doi. 1984. The subtilisin E gene of Bacillus subtilis is transcribed from a σ^{37} promoter in vivo. *Proc. Natl. Acad. Sci. (USA)* 81:1184-1188.

- Wood, P. J. 1980. Specificity in the interaction of direct dyes with polysaccharides. *Carbohydr. Res.* 85:271-287.
- Wood, T. M. 1981. Enzyme interactions involved in fungal degradation of cellulosic materials. p. 31-38. In: The Ekman-Days International Symposium on Wood Pulping Chemistry. Vol. 3. SPCI, Stockholm.
- Wood, T. M. 1985. Properties of cellulolytic enzyme systems. *Biochem. Soc. Trans.* 13:407-410.
- Wood, T. M., and S. I. McCrae. 1972. The purification and properties of the C₁ component of T. koningii cellulase. *Biochem. J.* 128:1183-1192.
- Wood, T. M., and S. I. McCrae. 1979. Synergism between enzymes involved in the solubilisation of native cellulose. p. 181-209. In: R. Brown and L. Jurasek (eds.), *Advances in Chem. Ser.* Vol. 181. American Chemical Society, Washington DC.
- Wood, T. M., and S. I. McCrae. 1982. Purification and some properties of a (1,4)- β -D-glucan glucanohydrolase associated with the cellulase from the fungus Penicillium funiculosum. *Carbohydr. Res.* 110:291-303.
- Wood, W. B. 1966. Host specificity of DNA produced by Escherichia coli: Bacterial mutations affecting the restriction and modification of DNA. *J. Mol. Biol.* 16:118-133.
- Yaguchi, M., C. Roy, C. F. Rollin, M. G. Paice, and L. Jurasek. 1983. A fungal cellulase shows sequence homology with the active site of hen egg-white lysozyme. *Biochem. Biophys. Res. Comm.* 116:408-411.
- Yamazaki, H., K. Ohmura, A. Nakayama, Y. Takeichi, K. Otozai, M. Yamasaki, G. Tamura, and K. Yamane. 1983. α -Amylase genes (amyR2 and amyE) from an α -Amylase-hyperproducing Bacillus subtilis strain: molecular cloning and nucleotide sequences. *J. Bacteriol.* 156:327-337.
- Yang, H. H., M. J. Effland, and T. K. Kirk. 1980. Factors influencing fungal degradation of lignin in a representative lignocellulosic, thermomechanical pulp. *Biotechnol. Bioeng.* 22:65-77.
- Youngleson, J. S., J. D. Santangelo, D. T. Jones, and D. R. Woods. 1987. Cloning and expression of a Clostridium acetobutylicum alcohol dehydrogenase gene in Escherichia coli. *Appl. Environ. Microbiol.* In press.
- Yu, E. K. C., M. K. H. Chan, and J. N. Saddler. 1985a. Butanediol production from lignocellulosic substrates by Klebsiella pneumoniae grown in sequential co-culture with Clostridium thermocellum. *Appl. Microbiol. Biotechnol.* 22:399-404.

- Yu, E. K. C., M. K. H. Chan, and J. N. Saddler. 1985b. Butanol production from cellulosic substrates by sequential co-culture of Clostridium thermocellum and C. acetobutylicum. Biotechnol. Lett. 7:509-514.
- Yu, E. K. C., and J. N. Saddler. 1982. Enhanced production of 2,3-butanediol by Klebsiella pneumoniae grown on high sugar concentrations in the presence of acetic acid. Appl. Environ. Microbiol. 44:777-784.
- Zabeau, M., and K. K. Stanley. 1982. Enhanced expression of cro-beta-galactosidase fusion proteins under the control of the P_R promoter of bacteriophage lambda. EMBO J. 1:1217-1224.
- Zappe, H., D. T. Jones, and D. R. Woods. 1986. Cloning and expression of Clostridium acetobutylicum endoglucanase, cellobiase and amino acid biosynthesis genes in Escherichia coli. J. Gen. Microbiol. 132:1367-1372.
- Zappe, H., D. T. Jones, and D. R. Woods. 1987. Cloning and expression of a xylanase gene from Clostridium acetobutylicum P262 in Escherichia coli. Appl. Microbiol. Biotechnol. 27:57-63.
- Zappe, H., W. A. Jones, D. T. Jones, and D. R. Woods. 1988. Structure of an endo- β -1,4-glucanase gene from Clostridium acetobutylicum P262 showing homology with endoglucanases from Bacillus species. Appl. Environ. Microbiol. In Press.