CHARACTERIZATION AND EXPRESSION OF CELLULOMONAS FIMI ENDOGLUCANASE B GENE AND PROPERTIES OF THE GENE PRODUCT FROM ESCHERICHIA COLI

By

JOSHUA BABATUNDE OWOLABI

B.Sc., University of Ife, 1979

M.Sc., University of Ife, 1982

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

(Department of Microbiology)

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

January 1988

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Microbiology

The University of British Columbia
1956 Main Mall
Vancouver, Canada
V6T 1Y3

Date January 27, 1988

ABSTRACT

In Cellulomonas fimi the cenB gene encodes a secreted endoglucanase (EngB) involved in the degradation of cellulose. The cenB gene carried on a 5.6 kb C. fimi DNA fragment encodes a polypeptide of M_r 110,000 in Escherichia coli. The level of expression of the gene was significantly increased by replacing its normal transcriptional and translational regulatory signals with those of the E. coli lac operon. The intact EngB polypeptide is not required for enzymatic activity: active polypeptides of M_r 95,000 and 82,000 also appear in E. coli and a deletion mutant of cenB encodes an active polypeptide of M_r 72,000. The intact and truncated EngB both bind to microcrystalline cellulose. A simple, rapid affinity chromatography procedure on Avicel was developed for the purification of intact EngB and of the 72,000 deletion derivative. Alignment of the amino-terminal amino acid sequence of the purified intact EngB from E. coli with the partial nucleotide sequence of the cloned C. fimi DNA showed that the mature EngB is preceded by a sequence encoding a putative signal polypeptide of 32 amino acids, a translational initiation codon and a sequence resembling an E. coli ribosome binding site 4 nucleotides before the initiation codon. The signal peptide functions and is correctly processed in E. coli, even when its first 15 amino acids are replaced by the first 7 amino acids of

β-galactosidase. The truncation of EngB does not affect its export to the periplasm of *E. coli*. In the intact EngB, 25% of the residues are hydroxyamino acids. It displays features common to endo-β-1,4-glucanases, since it has a high activity on carboxymethylcellulose. The kinetic parameters for carboxymethylcellulose hydrolysis of both intact and truncated EngB are not significantly different. *C. fimi* protease cleaves intact EngB, in a specific manner, to generate two polypeptides of M_r 65,000 and 43,000; the former has the capacity to bind Avicel. A polyclonal antibody raised against the purified intact EngB recognizes a *C. fimi* extracellular protein of M_r 110,000 as well as 5 polypeptides of lower molecular weight.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to thank Drs. R.A.J. Warren, R.C. Miller Jr., and D.G. Kilburn for their guidance and support during this work. Also, I wish to thank Dr. P.M. Townsley for helpful discussions. I would like to thank Ms. Sandy Keilland and Dr. Robert Olafson for the amino-terminal protein sequence, and Dr. D.J. Mckay for the amino acid composition analysis, also Dr. Tom Atkinson for the oligonucleotide synthesis and Dr. Neil R. Gilkes for information on the purification of C. fimi endoglucanase weakly bound to Avicel. Also , I wish to thank Dr. P. Beguin for collaboration on the construction of pCP3cenB expression vectors. I would like to thank all members of the Cellulase Group for their support when I needed it.

This work was supported by strategic grant 67-0941 from the Natural Sciences and Engineering Research Council of Canada to D.G.K., R.C.M., and R.A.J.W. I wish to thank the Canadian Government for the award of a Canadian Commonwealth Scholarship that enabled me to further my graduate studies. This thesis is dedicated to my wife, Aina, and my children, Olayinka and Akinsanya, for their understanding and support.

LIST OF ABBREVIATIONS

aa	Amino acid(s)
Amp	Ampicillin
cenA	Gene encoding *C. fimi* endoglucanase A
cenB	Gene encoding *C. fimi* endoglucanase B
cex	Gene encoding *C. fimi* exoglucanase
CM	Carboxymethyl
dCTP	Deoxycytosine triphosphate
DNSA	Dinitrosalicylic acid
DP	Degree of polymerization
DS	Degree of substitution
EngA	protein encoded by *cenA*
EngB	protein encoded by *cenB*
Exg	protein encoded by *cex*
FPLC	Pharmacia Fast Protein Liquid Chromatography system
IPTG	Isopropyl-β-D-thiogalactoside
Kan	Kanamycin
kb	1000 base pairs
kDa	1000 daltons
LacZ	*E. coli* β-galactosidase gene
LacZ'	The first 78 amino acids of β-galactosidase including the operator and promoter region of the gene.
LB	Luria broth
min	Minute(s)

M_r	Apparent molecular weight
MUC	4-methylumbelliferyl-β-D-cellobioside
ONPG	o-nitrophenyl-β-D-galactoside
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
p-HBAH	p-hydroxybenzoic acid hydrazide
PMSF	Phenylmethylsulfonyl fluoride
pNPC	p-nitrophenyl-β-D-cellobioside
pNPG	p-nitrophenyl-β-D-glucoside
pNPX	p-nitrophenyl-β-D-xyloside
SDS	Sodium dodecyl sulfate
Tet	tetracycline
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside
X-phosphate	5-bromo-4-chloro-3-indolyl-phosphate
/	plasmid carrier state

INTRODUCTION

I Background

Cellulose is the world's most abundant renewable carbon source. It is found as the main structural element and constituent of the walls of higher plants. Some algae, fungi and certain bacteria (for example, Acetobacter xylinum) also produce some cellulose. In recent years, due to energy shortage, food crisis and pollution, the potential for the use of cellulosic biomass for the production of microbial protein, liquid fuels and industrial chemicals has become recognized.

Cellulose can be hydrolysed by strong acids or by a set of enzymes called cellulases. Hydrolysis of cellulosics with acids poses air pollution problems through the emissions of stack gases containing corrosive chlorides and sulfur oxides. It requires extreme conditions of temperature and pH, and does not yield glucose as a sole end-product. Enzymatic hydrolysis, on the other hand, is a very selective process, being specific for the substrate, cellulose. The final product, glucose, is stable under the relatively mild conditions required for enzymatic hydrolysis. The enzymes are reusable, non-polluting and energy-sparing, but very large quantities are required because of their low specific activities and their poor

conversion of cellulose to glucose.

Cellulosic materials can be pretreated to enhance their enzymatic conversion to glucose (Brownell and Saddler, 1987; Fan *et al.*, 1982; Horton *et al.*, 1980; Millet *et al.*,1976). High yielding cellulase producing strains (Choi *et al.*,1978; Hagget *et al.*, 1978; Moloney *et al.*, 1983; Montenecourt, 1983) as well as mutants that are resistant to catabolite repression (Bailey and Oksanen, 1984; Stewart and Leatherwood, 1976) and end-product inhibition have been isolated (Choudhury *et al.*,1980). Strategies have also been devised for optimizing fermentation conditions which enhance cellulase activity (Chahal, 1985; Watson *et al.*, 1984).

Gene cloning technology affords another approach to increasing cellulase production (Joliff *et al.*, 1986b; O'Neill *et al.*, 1986c; Teeri, 1987). From a more basic point of view, gene cloning facilitates separation and identification of various components of the often complex cellulase systems. A given enzyme free of other cellulolytic components allows determination of substrate specificity and catalytic properties. Besides, the nucleotide sequences of cloned cellulase genes can be determined; the deduced amino acid sequences provide information on the structure of the enzymes. This facilitates analysis of structure-function relationships. An understanding of cellulase structural genes and of the structures, substrate specificities and catalytic properties of the enzymes they encode is a necessary prerequisite to a

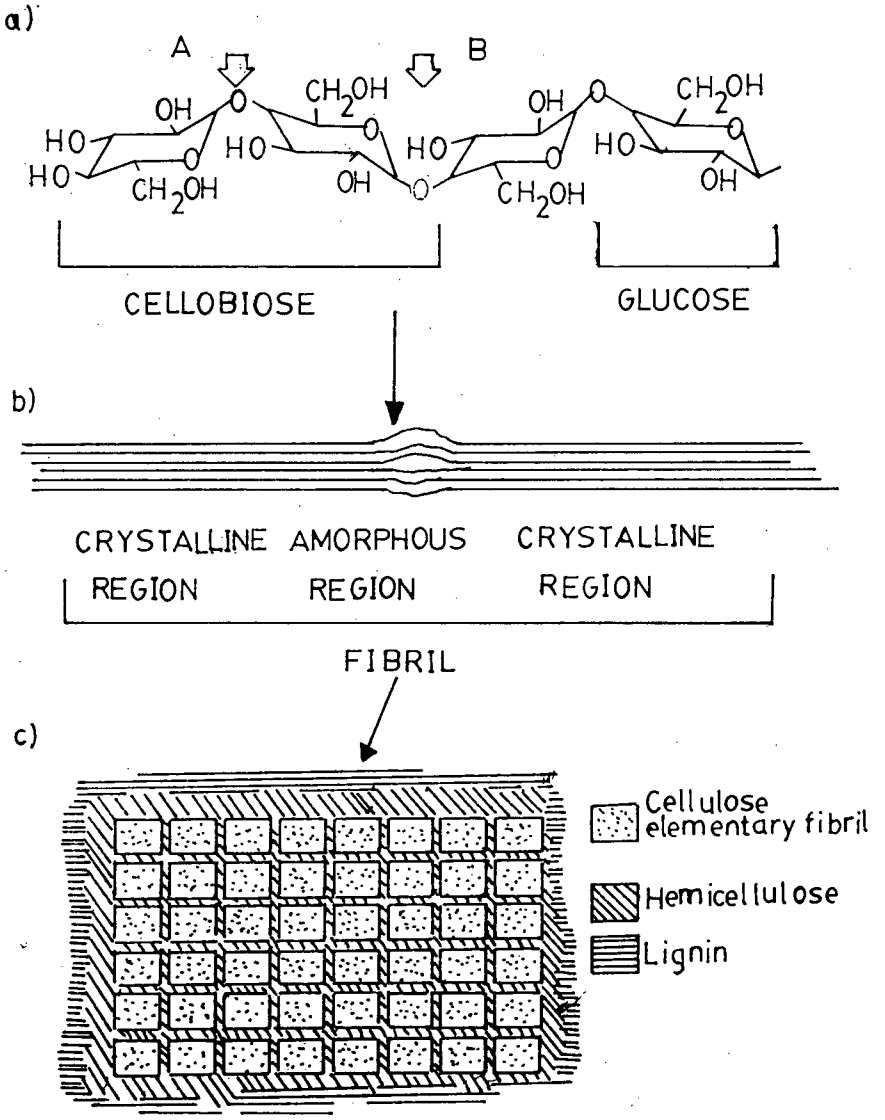
fuller understanding of the relative importance and mechanism of action of each enzyme in the hydrolysis of cellulose, and to the reconstitution of optimized cellulase systems.

II Enzymatic hydrolysis of cellulose

A) Structure of cellulose

Cellulose is a linear homopolymer, composed of glucose residues held together by β-1,4 linkages (Fan *et al.*, 1980a). The glucose units adopt the chair configuration with every other residue rotated 180° around the main axis to give an unstrained linear configuration (Fig. 1a). Thus, the basic functional unit of cellulose is cellobiose. Cellulose is organised into several levels of higher order structure. The smallest structural units are the elementary fibrils, in which a number of polymer chains oriented in parallel are held together by both intra- and inter-chain hydrogen bonds. Within the fibrils there are crystalline, completely ordered regions which alternate with less ordered, amorphous regions (Fig. 1b). Bundles of fibrils form fibers. In the native state, cellulose fibers are associated with other polysaccharides, such as hemicelluloses, and with lignin (Fig. 1c).

Figure 1. Structure of cellulose. (a) Stereochemical representation of a cellulose molecule. Arrows A and B represent β-1,4-linkages lying in different planes within the cellulose fibril. Cleavage at these linkages will generate two different end-group configurations (Wood, 1985). (b) Organization of cellulose molecules in elementary fibrils. Regions of the fibril in which the polymers are highly ordered or are in relative disarray have been termed the crystalline and amorphous regions, respectively. (c) Cross-section of a wood fiber. Cellulose elementary fibrils are embedded in a matrix of hemicellulose and lignin, reducing their accessibility to enzymatic digestion (adapted from Beguin *et al.*, 1987; Fan *et al.*, 1980a).



B) Cellulase-producing organisms

Although a variety of microorganisms can hydrolyse amorphous cellulose, only a few are capable of degrading crystalline cellulose efficiently. Notable in this regard are fungal species such as Trichoderma, Fusarium, Myrothecium, Penicillium, and Sporotrichum; actinomycetes such as Streptomyces and Thermonospora; and bacterial species such as Acetivibrio, Bacteroides, Cellulomonas, Clostridium, and Sporocytophaga (Coughlan, 1985).

In most organisms, the synthesis of cellulases is subject to induction by cellulose and repression by glucose. Neither the nature of the actual inducer nor the mechanism of induction are fully understood. Sophorose (2-O-β-glucopyranosyl-D-glucose) is an excellent inducer of cellulase synthesis in some species, including Trichoderma (Sternberg and Mandels, 1979), Sporotrichum (Eriksson and Hamp, 1978) and Cellulomonas (Stewart and Leatherwood, 1976).

There is some disagreement over the involvement of cAMP in the regulation of cellulase biosynthesis. There is no correlation between induction of cellulases and intracellular cAMP levels in Pseudomonas fluorescens var. cellulosa (Suzuki, 1975) and Trichoderma reesei (Montenecourt, 1983), but cAMP may be involved in regulating cellulase biosynthesis in Thermonospora (Coughlan, 1985).

C) Cellulolytic enzyme systems

Cellulolytic organisms produce complex mixtures of enzymes and other components to effect hydrolysis of native cellulose. The enzymes are mostly hydrolytic, but may be phosphorolytic or oxidative. The major types of hydrolytic enzymes include endoglucanases (endo-1,4-β-glucanase, EC 3.2.1.4) which cause internal cleavage of cellulose chains, exoglucanases (exo-1,4-β-glucanase, EC 3.2.1.91) which remove cellobiose units from the non-reducing ends of the major chain or shorter chains produced by the action of endoglucanases, and cellobiases (β-glucosidase, EC 3.2.1.21) which convert cellobiose to glucose. While fungi usually secrete β-glucosidases bacteria take up cellobiose and hydrolyse it intracellularly.

Cellobiose phosphorylase (EC 2.4.1.20) catalyses the reversible phosphorylation of cellobiose (Alexander, 1968). The oxidative enzymes cellobiose oxidase and cellobiose dehydrogenase oxidise cellobiose and higher cellodextrins to the corresponding lactones in the presence of molecular oxygen or other electron acceptors such as lignin or quinone, respectively (Eriksson, 1978). As some of the organisms producing these non-hydrolytic enzymes do not synthesize β-glucosidases, they provide alternative means of metabolizing cellulose biodegradation products and of diminishing the inhibitory effect of cellobiose on cellulase

activity.

Although, Reese *et al.* (1950) proposed the existence of a non-enzymatic factor, termed C_1, that renders crystalline cellulose amenable to hydrolysis by other components of the cellulase system, the C_1 components of most fungi have been isolated and shown to be exoglucanases (Halliwell and Griffin, 1973; Pettersson, 1975). However, the existence of an essential non-enzymatic component has been recently demonstrated in two separate studies. Griffin *et al.* (1984) isolated from T. reesei cellulase, a factor that generates microfibrils or shorter fibers from filter paper without hydrolysis. In addition to the extracellular enzymes, Ljungdal *et al.* (1983) also isolated a low molecular weight water-insoluble "yellow affinity substance" from culture filtrates of Clostridium thermocellum. This substance binds to the cellulose fibers in the growth medium and promotes the binding of the endoglucanases to the substrates.

Each of the major hydrolytic components of the cellulase complex synthesized by an individual organism may exist in a number of forms. This multiplicity of components could be genetically determined, or be caused by partial proteolysis (Gong and Tsao, 1979; Langsford *et al.*, 1984; Nakayama *et al.*, 1976) or by differential glycosylation of a common polypeptide chain (Moloney *et al.*, 1985). Wood (1985) has suggested that, for stereochemical reasons, at least two types of endoglucanase and exoglucanase should be

required for hydrolysis of crystalline cellulose. Theoretically, he argues, there will be two types of non-reducing end groups in the cellulose crystallite, requiring two different stereospecific enzymes for hydrolysis. Proteolysis has been reported to affect the release of cellulase from the cell walls of some cellulolytic fungi (Kubicek, 1981). Increased specific activity or activation of the endoglucanase from Sporotrichum pulverulentum (Eriksson and Pettersson, 1982) and Penicillium janthinellum (Deshpande *et al.*, 1984b) after treatment with protease(s) from the respective fungus has also been reported.

D) Measurement of cellulase activities

A number of methods are available for the detection and measurement of the activities of all or parts of the cellulase system (Mullings, 1985). A common qualitative assay relies on the incorporation of CMcellulose (a soluble cellulose derivative) into growth media and its interaction with Congo red. Hydrolysis of the CMcellulose by hydrolytic organisms provides zones of clearing (Teather and Wood, 1982). A development of this assay is the Congo red-stained agar replica used to detect cellulase activity in polyacrylamide gels (Beguin, 1983).

Determination of reducing sugar by the dinitrosalicylic acid (DNSA) (Miller, 1959) and the Nelson-Somogyi (Nelson, 1952) procedures are the most popular types of quantitative

assay available. Both methods are subject to significant interference from other reducing substances. Moreover, the response of either reagent varies considerably from sugar to sugar. Although insignificant when the hydrolysis of purified cellulosic substrate is under investigation, these drawbacks can be very serious when impure lignocellulosic materials are used (Rivers *et al.*, 1984). Breuil and Saddler (1985) have also commented on the inadequacies of the reducing sugar assay procedures.

Alternatively, the increase in fluidity accompanying hydrolysis of soluble derivatives of cellulose may be assayed viscometrically (Almin *et al.*,1975). Plots of increase in relative fluidity versus reducing equivalents, indicating the "randomness" of attack on the substrate, have been used to compare individual endoglucanases (Gilkes *et al.*, 1984c). Although the slopes obtained in such curves may help to characterize exoglucanases,it has become more usual to determine exoglucanase activity by measuring the release of p-nitrophenol from p-nitrophenyl-β-cellobioside (pNPC) (Deshpande *et al.*, 1984a) or release of methylumbelliferone from methylumbelliferyl cellobioside (MUC) (van Tilbeurgh *et al.*, 1982). β-glucosidase activity is determined by measuring the release of p-nitrophenol from p-nitrophenyl-β-glucoside or by the release of glucose from cellobiose (Mullings, 1985).

E) Adsorption/desorption and synergism of cellulases

The adsorption of cellulase on the surface of cellulosic material is the first step in hydrolysis. Factors affecting the adsorption of cellulases to cellulose include: the nature of the substrate; its purity; pretreatment and the extent to which it is crystalline or amorphous; enzyme/substrate ratio; the affinity of the multicomponent enzyme used for the substrate; the fact that the topography of the substrate changes as digestion proceeds; inactivation of bound or free enzyme; non-productive binding or immobilization of enzyme; and accumulation of products, especially cellobiose, that inhibit enzyme activity (Castanon and Wilke, 1980; Coughlan, 1985; Moloney and Coughlan, 1983).

On lignocellulosics such as newspaper, the enzymes once bound remain immobilized (Castanon and Wilke, 1980). However, when relatively pure cellulosic materials are used, cellulases are rapidly adsorbed, followed by a slow release of the enzymes to the liquid phase as the hydrolysis proceeds (Moloney and Coughlan, 1983). In the latter, the adsorption behavior obeys Michaelis-Menten kinetics (Moloney and Coughlan, 1983) in that the extent of adsorption increases as cellulose concentration or enzyme concentration increases. Maximum adsorption or desorption of cellulases occurs under the conditions of pH and temperature optimal

for hydrolysis (Moloney and Coughlan, 1983). However, at 0-5°C the extent of adsorption in the initial phase is slow but such adsorption continues until ultimately much more enzyme is bound than at higher temperature (Moloney and Coughlan, 1983; Ryu *et al.*, 1984).

Cellulolytic enzymes adsorb with different degrees of tenacity to cellulosic subtrates. This differential affinity has been exploited for the purposes of enzyme fractionation and purification (Beguin and Eisen, 1978; Boyer *et al.*, 1987; Gilkes *et al.*, 1984c; Halliwell and Griffin, 1978; Nummi *et al.*, 1981; Owolabi *et al.*, 1988; Reese, 1982; Schwarz *et al.*, 1986; van Tilbeurgh *et, al.*, 1984).

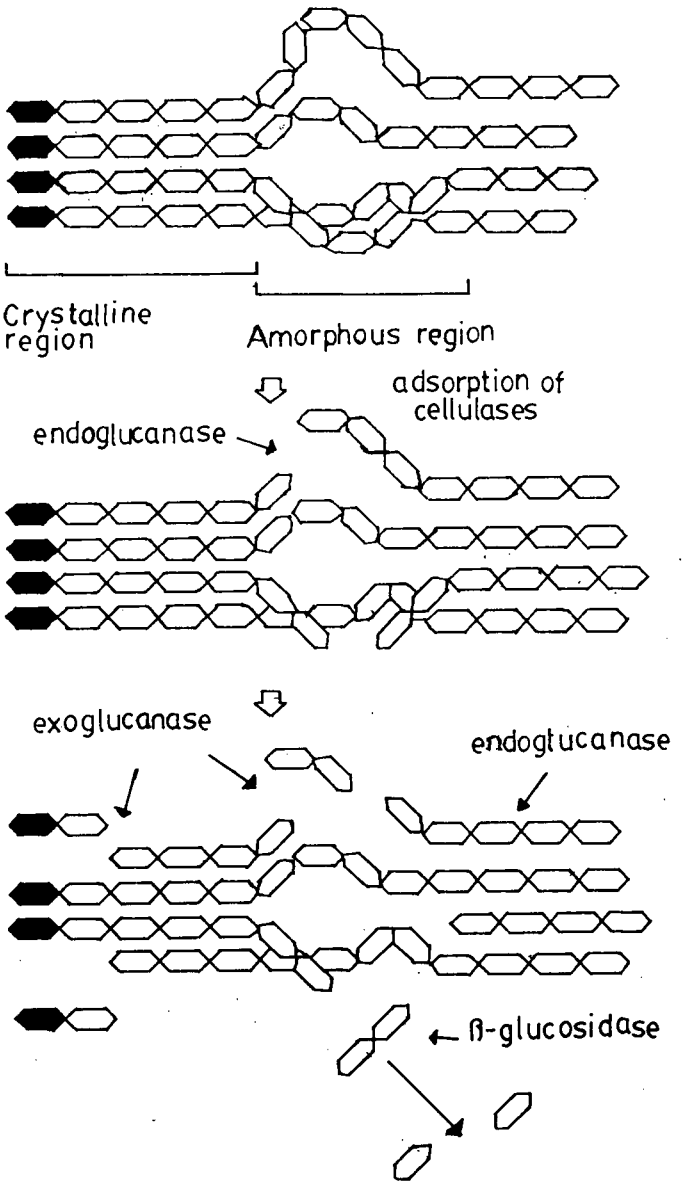
Although the individual components of cellulase systems by themselves have little action on crystalline cellulose, synergism between endo- and exo-glucanases has been shown for the enzymes from a number of fungal species (Eriksson, 1975; Moloney *et al.*,1985; Ryu *et al.*, 1984; Wood, 1975). This synergistic interaction is most marked when highly crystalline substrates are used, is low with amorphous cellulose and is absent with soluble derivatives (Wood and McCrae, 1979). Synergy has been reported to be maximal when the components are in the same ratio as they occur in the crude filtrates (Wood, 1975; Ryu *et al.*, 1984). Cross-synergism between the exoglucanases produced by one organism and the endoglucanase fractions of another has also been demonstrated (Wood, 1975; Moloney *et al.*, 1985).

Not all endoglucanases from a given filtrate are capable of effective synergistic interaction with exoglucanases from the same filtrate (Eriksson, 1975; Wood, 1975). This is best understood in terms of the recent observation that endoglucanases fall into two classes; those that adsorb "strongly" and those that adsorb "weakly" to insoluble cellulose (Ryu *et al.*, 1984). The former predominate in cellulase systems that are highly active against crystalline cellulose, whereas the "weakly" binding forms predominate in filtrates that are relatively inactive against such subtrates. Both forms of enzyme are equally active against soluble substrates (Ryu *et al.*, 1984). A similar study by Klyosov *et al.* (1986) suggests that the tightness of binding of the endoglucanases to the substrates, plays a crucial role in the degradation of crystalline cellulose.

F) Mechanism of cellulose hydrolysis

The mechanism of cellulase action as it relates to solubilization of crystalline cellulose is still controversial (Fan *et al.*, 1980b; Griffin *et al.*, 1984; Mandels, 1982; Reese *et al.*, 1950; Ryu *et al.*, 1984; Wood, 1985). A model accomodating the observations of various investigators has been proposed by Coughlan (1985) and is shown in Fig. 2. In the first step, amorphogenesis, the crystalline substrate is rendered more accessible to the hydrolytic enzymes by non-enzymatic factors. Further

Figure 2. Schematic representation of the action of cellulases on a cellulose fibril. Individual glucose residues of the cellulose chains are represented by hexagons. The non-reducing end of a cellulose polymer is denoted by a filled hexagon (adapted from Beguin *et al.*, 1987).



hydrolysis is then brought about by the combined actions of endoglucanases and exoglucanases, displaying synergism and perhaps competitive adsorption. Lastly, β-glucosidases act on the cellobiose to produce glucose. This model, although based on fungal studies, is thought to hold true for bacterial cellulase systems.

III The application of molecular biology to the study of Cellulomonas fimi cellulases

The first molecular cloning of a cellulase gene was reported in 1982 (Whittle *et al.*, 1982). Since then several research groups have reported the molecular cloning of about 60 cellulase genes from 20 different organisms (Table I). The research focus has been on the characterization and the heterologous expression of the cloned cellulase genes. Notable among the achievements in the field is the determination of nucleotide sequences of 14 of the cellulase genes (Beguin *et al.*, 1985; Fukumori *et al.*, 1986a,b; Grepinet and Beguin, 1986; Joliff *et al.*, 1986a; Kohchi and Toh-e, 1985; O'Neill *et al.*, 1986a; Penttila *et al.*, 1986, Robson and Chambliss, 1987; Shoemaker *et al.*, 1983, 1984; Teeri *et al.*, 1987; Wakarchuk *et al.*, 1988; Wong *et al.*, 1986), and the overexpression in Escherichia coli of a cellulase gene which led to the first crystallization of a cellulase (Joliff *et al.*, 1986b). Currently, the structural organization of recombinant cellulases is being studied in detail and speculations about the active sites and substrate

Table I. Summary of cloned cellulase genes

Organism	Genes cloned no.	type	Screening method	Reference
Agrobacterium ATCC21400	1	β-glucosidase	DNA probe	Wakarchuk *et al.*, 1986
Bacillus subtilis	1	endoglucanase	indicator plates	Koide *et al.*, 1986
B. subtilis DLG	1	endoglucanase	immunological	Robson and Chambliss, 1986
Bacillus sp. strain N-4	2	endoglucanases	indicator plates	Sashihara *et al.*, 1984
Bacillus sp. strain 1139	1	endoglucanase	indicator plates	Fukumori *et al.*, 1986a
Bacteroides succinogenes	1	endoglucanase	indicator plates	Collier *et al.*, 1984
Cellulomonas fimi	2	endoglucanases	immunological	Whittle *et al.*, 1982
	1	endoglucanase	DNA probe	B. Moser, personal communication
	1	exoglucanase	enzyme assays	Gilkes *et al.*, 1984a,b
	1	β-glucosidase	enzyme assays	N. Bates, personal communication
Cellulomonas uda	1	β-glucosidase	indicator plates	Nakamura *et al.*, 1986
Cellulomonas mixtus	1	β-glucosidase	indicator plates	Wayne and Pemberton, 1986
	1	endoglucanase	indicator plates	Wayne and Pemberton, 1986
Clostridium acetobutylicum	1	β-glucosidase	enzyme assays	Zappe *et al.*, 1986
	1	endoglucanase	indicator plates	Zappe *et al.*, 1986
Clostridium thermocellum	7	endoglucanases	indicator plates	Cornet *et al.*, 1983b
	3	exoglucanases	enzyme assays	Millet *et al.*, 1985
	2	β-glucosidases	indicator plates	Schwarz *et al.*, 1985
	13	endoglucanases	indicator plates	Romaniec *et al.*, 1987
Escherichia adecarboxylata	1	β-glucosidase	growth on cellobiose	Armentrout and Brown, 1981
Erwinia caratovora	1	β-glucosidase	complementation	Barras *et al.*, 1984

Table I continued

Organism	Genes cloned no.	type	Screening method	Reference
Erwinia chrysanthemi	1	endoglucanase	indicator plates	van Gijsegem *et al.*, 1985
	1	endoglucanase	indicator plates	Barras *et al.*, 1984
	1	endoglucanase	indicator plates	Kotoujansky *et al.*, 1985
Thermonospora fusca YX	2	endoglucanases	enzyme assays	Collmer and Wilson, 1983
Pseudomonas sp.	1	endoglucanase	indicator plates	Wolff *et al.*, 1986
Pseudomonas fluorescens sub sp. cellulosa	1	endoglucanase	indicator plates	Wolff *et al.*, 1986
Ruminococcus flavefaciens	1	endoglucanase	indicator plates	Barras and Thomson, 1987
Aspergillus niger	1	β-glucosidase	indicator plates	Penttila *et al.*, 1984
Candida pelliculosa	1	β-glucosidase	indicator plates	Kohchi and Toh-e, 1985
Kluyveromyces fragilis	1	β-glucosidase	enzyme assays	Raynal and Guerineau, 1984
Trichoderma reesei	1	exoglucanase	differential hybridization of cDNA probes;	Shoemaker *et al.*, 1983
	1	endoglucanase	hybrid selection of mRNA	Penttila *et al.*, 1986
	1	exoglucanase	differential hybridization of cDNA probes;	Teeri *et al.*, 1983
	1	exoglucanase	hybrid selection of mRNA	Teeri *et al.*, 1987

binding domains are already emerging (Enari and Niku-Paavola, 1987; Knowles *et al.*, 1987; Langsford *et al.*, 1987; Teeri *et al.*, 1987; van Tilbeurgh *et al.*, 1986; Warren *et al.*, 1986, 1988; Yaguchi *et al.*, 1983). This discussion will be limited to the application of molecular biology to the study of Cellulomonas fimi cellulases. An excellent review article by Beguin *et al.* (1987) provides a detailed account of the cloning and characterization of recombinant cellulases.

The genus Cellulomonas comprises a group of coryneform Gram-variable, mesophilic, facultative anaerobic rods, many of which are capable of degrading cellulose (Keddie, 1974). Although not as extensively studied as Trichoderma spp., there is a considerable body of information on their basic physiology and biochemistry. Cellulomonas spp. have the potential for single cell protein production from waste cellulosics (Hitchner and Leatherwood, 1980; Enriques, 1981), and also for direct saccharification of cellulosics (Hagget *et al.*, 1979). Several types of mutant of Cellulomonas strains have been isolated: those resistant to end-product inhibition (Choudhury *et al.*, 1980) or catabolite repression (Stewart and Leatherwood, 1976); and those that produce increased levels of cellulases (Choi *et al.*, 1978).

The most intensively studied Cellulomonas sp. at both the biochemical and genetic levels is Cellulomonas fimi (Beguin and Eisen, 1978; Beguin *et al.*, 1977; Gilkes *et al.*,

1984a,b,c; Greenberg *et al.*, 1987a,b; Guo *et al.*, 1988; Langsford *et al.*, 1984, 1987; O'Neill *et al.*, 1986a,b,c; Owolabi *et al.*, 1988; Wakarchuk *et al.*, 1984; Warren *et al.*, 1986, 1988; Whittle *et al.*, 1982; Wong *et al.*, 1986). This organism produces a complex mixture of enzymes active against CMcellulose during growth on microcrystalline cellulose (Avicel) (Beguin *et al.*, 1977; Beguin and Eisen, 1978; Langsford *et al.*, 1984), at least some of which are glycosylated (Beguin and Eisen, 1978; Langsford *et al.*, 1984). Several of the enzymes bind strongly to the substrate; others bind weakly and are removed by water or dilute buffer (Beguin *et al.*, 1977, Beguin and Eisen, 1978; Langsford *et al.*, 1984; Owolabi *et al.*, 1988). Two of the tightly bound enzymes, both glycoproteins, have been characterized in detail. They are eluted from Avicel with 6 M guanidium hydrochloride. One of them is an exoglucanase (Exg); the other, an endoglucanase (EngA) (Gilkes *et al.*, 1984c; Langsford *et al.*, 1984). Both enzymes hydrolyse CMcellulose, albeit with different kinetics, but only the Exg hydrolyses pNPC and MUC.

The structural genes, cenA for EngA and cex for Exg, were cloned in E. coli on the vector pBR322 (Gilkes *et al.*, 1984a; Whittle *et al.*, 1982). The genes have been sequenced; their coding regions were identified by comparison of the nucleotide sequences with the NH_2-terminal amino acid (aa) sequences of the enzymes purified from C. fimi (O'Neill *et al.*, 1986a; Wong *et al.*, 1986). The cenA

gene (1347 bp) encodes a polypeptide of 449 aa, including a 31 aa leader peptide (Wong *et al.*, 1986). The cex gene(1452 bp) encodes a polypeptide of 484 aa, including a 41 aa long leader peptide (O'Neill *et al.*, 1986a). Both leader peptides function in the processing and export of each enzyme into the periplasm of E. coli (O'Neill *et al.*, 1986b; Wong *et al.*, 1986).

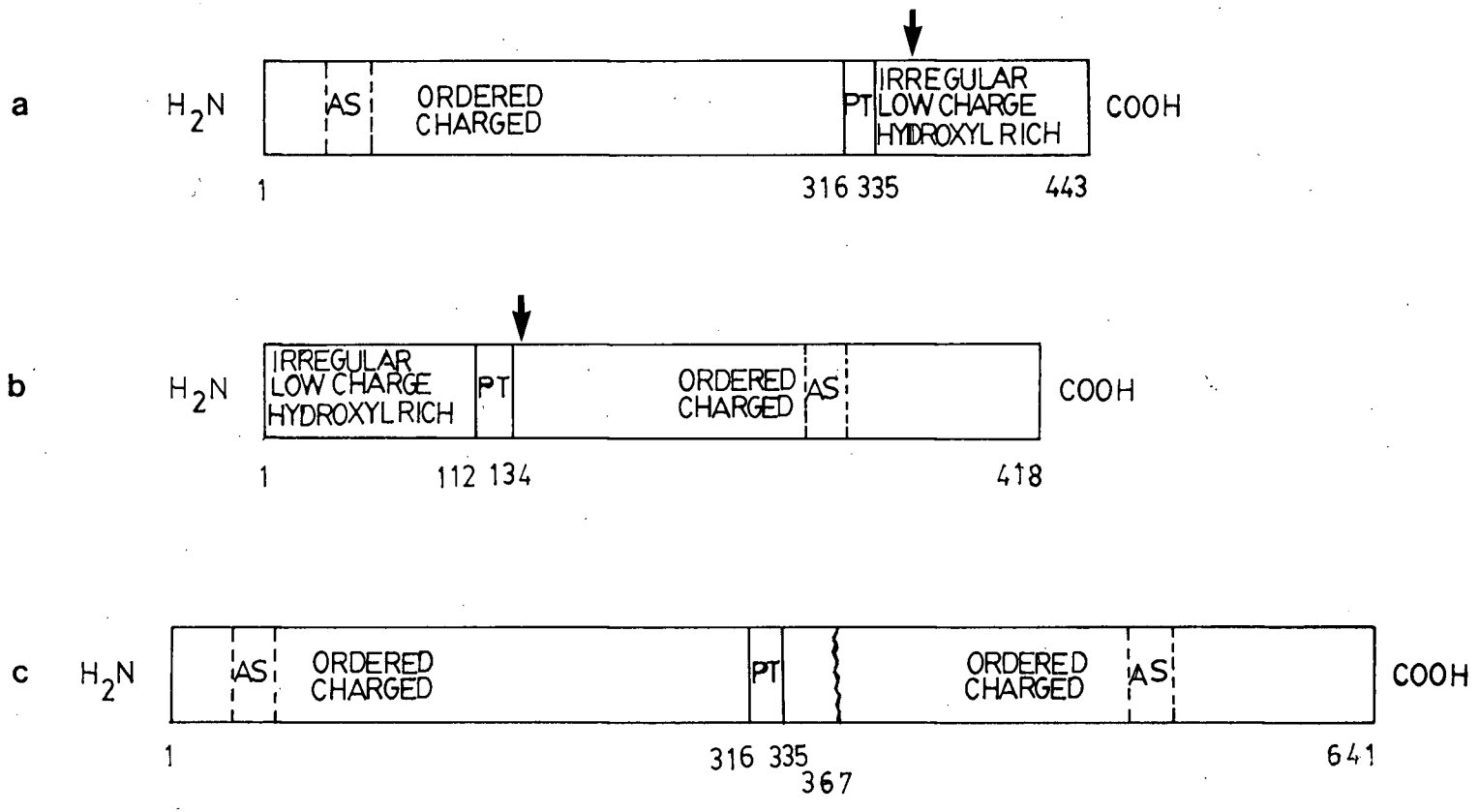
Although the natural cex and cenA promoters (Greenberg *et al.*, 1987a) were present on the cloned fragments, operon fusion experiments in E. coli showed that transcription of these genes in E. coli absolutely required E. coli promoter sequences (O'Neill *et al.*, 1986c; Wong *et al.*, 1986).

Using the above information, both the cenA and cex genes have been engineered for overproduction in E. coli (Guo *et al.*, 1988; O'Neill *et al.*, 1986c). To obtain the overproduction of Exg, the cex coding sequence was fused to a synthetic ribosome-binding site and an initiating ATG , and placed under the control of the leftward promoter of bacteriophage lambda contained on the run-away replication plasmid vector pCP3 (O'Neill *et al.*, 1986c). With the exception of an inserted asparagine adjacent to the initiating ATG, the highly expressed Exg (20 % of total cellular protein) was identical to the native exoglucanase. However, the pCP3-cex-directed overproduction of Exg led to its accumulation in the cytoplasm as aggregates devoid of activity. The aggregates sedimented readily and could be solubilized in either 6 M Urea and 5 M guanidium

hydrochloride. Following solubilization and dialysis, partial recovery of Exg activity was obtained (O'Neill *et al.*, 1986c). To obtain overexpression of the cenA gene, its coding sequence was fused to the translational and transcriptional signals of the E. coli lac operon in pUC18 (Guo *et al.*, 1988). In this construct, the length of the EngA leader peptide was increased by 9 aa, none of which was basic. Nonetheless, the hybrid leader peptide was recognized and processed by E. coli. Moreover, the 800-fold overexpression caused the cells to release EngA in the culture medium by non-specific leakage from the periplasm in a temperature-dependent manner (Guo *et al.*, 1988). It has been suggested that the accumulation of a protein to a high level in the periplasm destabilizes the outer membrane of E. coli resulting in the leakage of periplasmic proteins to the medium (Gatz and Hillen, 1986).

Comparison of the predicted aa sequences of Exg (443 aa) and EngA (418 aa) show that each enzyme has three distinct regions: a short sequence of about 20 aa composed only of proline and threonine (the Pro-Thr box), which is conserved almost perfectly in the two enzymes; a region rich in hydroxyamino acids but of low charge density, which is 50% conserved; and a region, comprising about 70% of the polypeptide, which is poorly conserved but contains a potential active site, Glu - Xaa_7 - Asn - Xaa_6 - Thr. The order of the regions is reversed in the two enzymes (Fig. 3; Langsford *et al.*, 1987; Warren *et al.*, 1986). The Pro-Thr

Figure 3. Overall structures of Exg (a) and EngA (b) of C. fimi, and of the fusion protein Exg-EngA (c) from E. coli. PT denotes the Pro-Thr box; AS denotes the putative active site; arrows indicate C. fimi protease cleavage sites of recombinant enzymes. Numbers refer to amino acids, begining at the mature NH_2-termini. The fusion protein lacks the Pro-Thr box and all of the irregular low charge hydroxyl rich region of EngA, and most of a similar region of Exg (adapted from Langsford *et al.*, 1987; Warren *et al.*, 1986 and 1988).


a
H_2N
AS
ORDERED CHARGED
PT
IRREGULAR LOW CHARGE HYDROXYL RICH
COOH
1
316 335
443
b
H_2N
IRREGULAR LOW CHARGE HYDROXYL RICH
PT
ORDERED CHARGED
AS
COOH
1
112 134
418
c
H_2N
AS
ORDERED CHARGED
PT
ORDERED CHARGED
AS
COOH
1
316 335
367
641

box may be a site for O-linked glycosylation. The sequence Asn-Xaa-Ser/Thr which occurs 6-8 times in both enzymes (4 such sites in the conserved irregular regions) is a potential glycosylation site for N-linked sugars (Warren *et al.*, 1986). Mannose appears to be the sole glycosyl component of the two enzymes from C. fimi (Arfman *et al.*, 1987; M. Langsford, personal communication).

Glycosylated EngA and Exg from C. fimi have been compared with their non-glycosylated counterparts produced in E. coli (Langsford *et al.*, 1987). Glycosylation of the enzymes does not significantly affect their kinetic and substrate binding properties or their stabilities toward heat and pH. However, the glycosylated enzymes are protected from attack by C. fimi protease when bound to Avicel, whereas the non-glycosylated enzymes yield active, truncated products, 30 kDa from EngA (48 kDa) and 39 kDa from Exg (46.5 kDa), with greatly reduced affinity for cellulose. Immunological characterization of the proteolysis products from both recombinant enzymes with the antiserum to synthetic Exg Pro-Thr box showed the presence of 15 kDa and 39 kDa immunoreactive fragments from EngA and Exg, respectively (Langsford *et al.*, 1987). These data are consistent with a bifunctional organization of the enzymes: the C. fimi protease specifically cleaves both enzymes at sites close to the carboxyl-terminus of the Pro-Thr box, generating two fragments, one containing the catalytic region, and the other the cellulose binding region (Fig. 3).

In another study, the sequences of the cex and cenA genes encoding the catalytic domains of Exg and EngA were joined together (Warren *et al*, 1988). The gene fusion encodes a polypeptide with both exoglucanase and endoglucanase activities but unable to bind to Avicel, presumably because it lacks an intact substrate binding region (Fig. 3). These properties further support the bifunctional organization of the enzymes.

IV Objectives of this thesis

The primary goal of the Cellulase Research Group at the University of British Columbia is the characterization of all the components of the C. fimi cellulase system. The successful reconstruction by molecular genetics of the entire cellulase system depends on an understanding of all the structural genes involved. To this end, another endoglucanase gene was also cloned from C. fimi into pBR322 (Gilkes *et al*., 1984a). The resulting plasmid, pEC3, was reported to contain a 5.6 kb BamHI fragment of C. fimi DNA (Gilkes *et al*.,1984a) expressing endoglucanase activity in E. coli, albeit at a very low level (Gilkes *et al*., 1984c). A significant fraction of this activity was found in the periplasm of the heterologous host (Gilkes *et al* ., 1984a). This thesis describes the structural gene (cenB) for C. fimi endoglucanase B (EngB) carried by pEC3, its increased expression in E. coli, and the purification and characterization of recombinant EngB from E. coli.

MATERIALS AND METHODS

I Bacterial strains, plasmids and phages

C. fimi ATCC 484 was the source of the endoglucanase B gene. The E. coli strains C600, JM101, RR1 and BD1854 were described previously (Appleyard, 1954; Jensen *et al.*, 1984; Peacock *et al.*, 1981; Yanisch-Perron *et al.*, 1985). The plasmids pBR322, pcI857, pCP3, pEC3, pUC18, pUC19 and pDR540 were described previously (Bolivar *et al.*, 1977; Gilkes *et al.*, 1984a; Norrander *et al.*, 1983; Remaut *et al*, 1983; Russell and Bennett, 1982; Yanisch-Perron *et al.*, 1985). The phage vectors M13mp11 and M13mp18 were described previously (Messing, 1983; Yanisch-Perron *et al.*, 1985). The characteristics of these bacterial strains, plasmids and phages are given in Table II.

II Cultivation conditions

C. fimi was grown at 30°C in basal salt medium (Hitchner and Leatherwood, 1980) supplemented with 3% Avicel or 0.1% glycerol. Except for E. coli JM101, which was maintained on M9 minimal medium plates (Miller, 1972), all other E. coli strains were grown in Luria broth (LB) (Miller, 1972). Ampicillin (100 μg/ml) was used when growing bacteria containing plasmids. Solid media contained 11 g agar per liter. When screening recombinant clones for endoglucanase activity 0.1% CMcellulose (Low viscosity, DP

Table II. Bacterial strains, plasmids and phages

Bacterial strain	Relevant genotype	Reference
E. coli C600	*thi*-1 *thr*-1 *leuB6 lacY1 tonA21 supE44*	Appleyard, 1954
E. coli BD1854	*minA minB thi rpSL tonA lac*	Jensen *et al.*, 1984
E. coli RR1	*hsdS20*(r_B^--m_B^-) *ara*-14 *proA2 galK2 rpsL20 supE44*	Peacock *et al.*, 1981
E. coli JM101	*supE thi* Δ(*lac-proAB*) [F'*traD36 proAB lacI*^{q}*Z*ΔM15]	Yanisch-Perron *et al.*, 1985
C. fimi ATCC484	cellulose utilization	

Plasmids	Relevant features	Reference
pBR322	Amp^R, Tet^R	Bolivar *et al.*, 1977
pEC3	Amp^R	Gilkes *et al.*, 1984a
pUC18	Amp^R *lacZ'*	Yanisch-Perron *et al.*, 1985
pUC19	Amp^R *lacZ'*	Norrander *et al.*, 1983
pDR540	Amp^R *galK*$^+$	Russell and Bennett, 1982
pCP3	Amp^R, ts runaway replication, $\lambda O_L^+P_L^+$	Remaut *et al.*, 1983
p*c*I857	Kan^R, λ*c*I857^{ts}	Remaut *et al.*, 1983

Phage	Relevant features	Reference
M13mp11	*lacZ'*	Messing, 1983
M13mp18	*lacZ'*	Yanisch-Perron *et al.*, 1985

400) was incorporated into LB-agar plates.

III Biological screening for endoglucanase activity

The interaction of the dye Congo red with intact β-glucans provided a basis for a rapid and sensitive assay system for screening recombinant clones with endoglucanase activity. E. coli clones carrying recombinant plasmids expressing endoglucanase were picked onto both a master LB-agar plate and a LB-agar-CMcellulose plate. The use of CMcellulose for the qualitative assay of endoglucanases has been described previously (Teather and Wood, 1982). After incubation of the plates at 30 to 37°C for 8-12 hours, the colonies were removed and the plates were flooded with Congo red dye solution (2 mg/ml) then rocked gently on a platform for 15 min. The Congo red solution was poured off, and the plates were flooded with 1 M NaCl and rocked as above. The zones of hydrolysis were pale yellow against a red background.

IV Preparation and localization of proteins

A) Elution of Avicel-bound C. fimi proteins with water

5 liters of a 6 day stationary phase culture of C. fimi (grown in 3% Avicel-basal salts medium) were allowed to settle at 23°C for 1 hour. After removing the supernatant by aspiration, the cellulose slurry was centrifuged at 4000 g for 15 min at 4°C. The remaining supernatant and cells

were discarded. The cellulose pellet was resuspended in distilled water to a total volume of 1.5 litres and centrifuged as above. The final pellet was discarded. The water eluates were combined, adjusted to 0.02% sodium azide, clarified by further centrifugation and concentrated 250-fold by pressure filtration through an Amicon PM10 membrane.

B) Extracellular C. fimi protease

A 100ml 2 day exponential phase culture of C. fimi (grown in 0.1% glycerol-basal salts medium) was centrifuged at 7000 g for 20 min at 4°C. The supernatant was filtered through a 0.22μm filter unit (Millipore), adjusted to 0.02% sodium azide and kept at 4°C.

C) Localization of recombinant EngB

Proteins from total cell extracts were prepared by breaking the cells with a French press (Whittle *et al.*, 1982). Periplasmic proteins were isolated by osmotic shock (Nossal and Heppel, 1966). Cytoplasmic proteins were prepared by rupturing the osmotically shocked cells with a French press.

V Purification of intact or truncated recombinant EngB

E. coli RR1 containing pJB301 or pJB303 was grown in 10 litres Luria broth in a stirred fermenter at 30°C. Exponentially growing cells were harvested with a centrifuge

(Sharples) and washed with 2 litres of ice-cold 0.01 M Tris-HCl (pH 7.1) - 0.03 M NaCl. The periplasmic fraction of the cells was prepared as described (Nossal and Heppel, 1966) and kept in buffer A (50 mM potassium phosphate, pH 6.9, 0.02% sodium azide) containing phenylmethylsulfonyl fluoride (20 μg/ml) at 4°C until needed.

Avicel was stirred gently in distilled water (20ml/g), and centrifuged at 4000 g for 5 min to remove "fines". This step was repeated 4 times, and the Avicel was finally resuspended in the same volume of water. After autoclaving for 30 min (121°C, 15 lb/in^2) it was left to cool to room temperature (23°C). The aqueous phase was replaced and discarded twice as above, and finally replaced with buffer A. After overnight equilibration at 4°C, the liquid was removed. The wet settled Avicel was resuspended in ice-cold buffer A (4ml/g) and kept on ice.

A 350 ml volume of periplasmic fluid (from 10 litres of culture) was mixed with 150 ml of the autoclaved Avicel (containing approximately 25 g of dry Avicel), and kept on ice for 1 hour. All subsequent steps were carried out at room temperature. The unadsorbed material was separated from Avicel by filtration through sintered glass (GF/C). The Avicel-enzyme complex was washed once with buffer A by filtration as above, resuspended in half its volume of fresh buffer and packed into a column (2.5 cm x 13 cm). Proteins bound to Avicel were eluted with a concave descending gradient of buffer A (55 ml) and water (800 ml), at a flow

rate of 30 ml/hour. Fractions containing endoglucanase activity were pooled and concentrated by pressure filtration through an Amicon PM10 membrane before loading or loaded directly onto a Pharmacia Mono Q anion exchange column that was equilibrated in buffer B (20 mM Tris-HCl, pH 7.5). The protein was eluted with a 30-40 ml gradient of zero to 0.15 or 0.20 M NaCl in buffer B. Those fractions with the highest specific activity were pooled and desalted by gel filtration using Bio-Gel P-6DG.

VI Enzyme assays and protein determination

A) Reducing sugar assays

CMcellulase activity was determined by colorimetric estimation of sugars using the DNSA (Miller *et al.*, 1960) or p-hydroxybenzoic acid hydrazide (p-HBAH) (Lever, 1973) method. The assay conditions for the DNSA method were as follows: 0.25 ml of appropriately diluted enzyme solution was mixed with 0.5 ml of 4% CMcellulose (Low viscosity, DP 400) in buffer A. After 1 hour incubation at 37°C, the reaction was stopped with 0.8 ml of dinitrosalicylic acid reagent (Miller, 1959); 0.05 ml of 0.1% glucose was added, and the tubes were steamed for 15 min. The absorbance was read at 550 nm against appropriate blanks containing equivalent amounts of enzyme added after mixing substrate with DNSA reagent. One unit of enzyme released 1 nmole glucose equivalents per min by reference to a glucose

standard curve. The assay conditions for the p-HBAH method were as follows: 0.1 ml of appropriately diluted enzyme solution was mixed with 0.4 ml of 0.5% CMcellulose (Low viscosity, DP 400) in buffer C (50 mM sodium citrate, pH 6.8). After incubation at 30°C for 30 min, the reaction was stopped with 1.0 ml p-HBAH reagent (Lever, 1973); and the tubes were steamed for 12 min. The absorbance was read at 420nm against blanks containing equivalent amounts of enzyme added after mixing substrate with p-HBAH reagent. One unit of enzyme activity released 1 nmole glucose equivalents per min by reference to a glucose standard curve.

Other polysaccharides (0.8 to 2.0 mg/ml, final concentration) were incubated with appropriately diluted enzyme in buffer C (total volume was 0.5 ml) at 30°C for 30 min. The reducing sugar was determined with the p-HBAH reagent as described above.

B) Aryl glycosidase assays

Aryl glycosidase activities (p-nitrophenyl cellobiosidase, p-nitrophenyl xylosidase, p-nitrophenyl glucosidase) were determined by incubation of 0.5 ml of 5 mM substrate with 0.5 ml of appropriately diluted enzyme in buffer A at 37°C for 1 hour. The reaction was stopped with 0.5 ml of 1 M Na_2CO_3. The absorbance was read at 410 nm against blanks containing equivalent amounts of enzyme added after mixing substrate with 1 M Na_2CO_3. One unit of aryl glycosidase was that amount of enzyme which liberated 1 nmol

of p-nitrophenol per min at 37°C.

C) β-lactamase assay

β-lactamase activity was assayed spectrophotometrically following the release of nitrocefoic acid from nitrocefin (O'Callaghan *et al.*, 1972). A 1.0 mM stock solution of nitrocefin was prepared by dissolving 5.2 mg of nitrocefin in 0.5 ml of DMSO and then diluting this solution into 9.5 ml of buffer A. The assay contained 0.05 ml of 1.0 mM nitrocefin, 0.9 ml of buffer A, and 0.05 ml of appropriately diluted cell extract. This solution was immediately mixed in a cuvette and the rate of change of absorption at 486 nm was recorded. The results were plotted as change in O.D. at 486 nm versus time; only values in the linear range of the plot were used for calculations. A 0.01 M solution of nitrocefoic acid has an O.D. at 486 nm = 1.55 (O'Callaghan *et al.*, 1972). One unit of enzyme activity released 1 nmole of nitrocefoic acid per min at 23°C.

D) β-galactosidase assay

β-galactosidase activity was measured according to Miller(1972), by the hydrolysis of o-nitrophenyl-β-galactoside (ONPG). The assay conditions were as follows: 0.2 ml of ONPG (4 mg/ml) was added to 1.0 ml of pre-incubated enzyme and Z-buffer (Miller, 1972). After 10 min incubation at 28°C, the reaction was stopped with 0.5 ml of 1 M Na_2CO_3 solution. The absorbance was read at 420 nm

against a Z-buffer blank. One nmole/ml of o-nitrophenol (ONP) has an O.D. at 420nm = 0.0045. β-galactosidase activity was calculated as described (Miller, 1972). One unit of β-galactosidase activity is the amount of enzyme which produces 1 nmole o-nitrophenol per min at 28°C.

E) Protease assay

Protease activity was determined by the hide powder blue assay (Rinderkneckt *et al.*, 1968). 1.5 ml enzyme in buffer D (50 mM Tris-HCl pH 7.8) was mixed with 10 mg hide powder azure. After 1 hour incubation at 37°C, the reaction tubes were cooled on ice and centrifuged. The absorbancies at 595 nm of the supernatant solutions were taken against buffer blanks. Protease activity was standardized with collagenase (Sigma); one unit of activity releases 1 μmol L-leucine in 5 hours at pH 7.8 and 37°C.

F) Protein determination

Protein was determined by the dye binding assay (Bradford, 1976) using bovine plasma albumin as standard.

VII DNA methodology

A) Plasmid DNA isolation and analysis

Plasmid DNA for restriction analysis was isolated by the alkaline lysis procedure (Birnboim and Doly, 1979). Plasmid DNA to be sequenced was purified by banding in

CsCl-ethidium bromide density gradients (Maniatis *et al.*, 1982). The M13RF and viral DNAs were isolated from infected cultures as described (Messing, 1983). DNA restriction fragments were resolved by agarose gel electrophoresis (Maniatis *et al.*, 1982).

B) Oligonucleotide synthesis and purification

The deoxyoligoribonucleotide 5'-CTTTATGCTTCCGGCTCGTA-3' (20-mer) to be used as primer for sequencing junctional regions of fusion plasmids was synthesized by Dr. T. Atkinson at the University of British Columbia, using an Applied Biosystems automated DNA synthesizer model 380A as described (Atkinson and Smith, 1984). The oligonucleotide was purified by polyacrylamide gel electrophoresis (PAGE) on a 20% acrylamide-7M urea sequencing gel, located by UV shadowing on a thin-layer-chromatography (TLC) plate, and extracted from the gel by the crush and soak method (Atkinson and Smith, 1984).

C) DNA sequencing

Fusion regions of deletions at the 5' end of the C. fimi chromosomal DNA subcloned in pUC19, generated with Exonuclease III and S_1 nuclease (Guo and Wu, 1983; Roberts and Lauer, 1979), were sequenced by the double stranded sequencing technique of Hattori and Sakaki (1986), using the above oligonucleotide as primer. The 400 bp BamHI - PstI fragment and some restriction fragments of the deleted DNA

were subcloned into M13mp11 and/or M13mp18, then sequenced by the enzymatic procedure of Sanger *et al*. (1977). The presence and number of the PTIS (portable translation initiation site) sequences inserted in the pCP3 expression vector were confirmed by DNA sequencing using the base-specific chemical degradation method (Maxam and Gilbert, 1980).

VIII Minicells

Minicells were prepared from E. coli BD1854 transformed with appropriate plasmids using the method of Jensen *et al*. (1984), with modifications. The cells were harvested by centrifugation and resuspended in 20 ml of M9 minimal medium (Miller, 1972). The minicells were separated from real cells by sedimentation in a 30 ml 5-20% sucrose gradient for 15 min at 5000rpm in a Beckman SW-27 rotor at 4°C. After tapping the minicells and pelleting them, the sucrose gradient step was repeated. The minicells were again tapped, pelleted, washed once in 5 ml M9 medium, and resuspended in the same medium to an A_{450} of about 2. Samples (0.5 ml) were prepared containing 0.4 ml of the above suspension of minicells, glucose (0.1%), 5 mM each of the protein amino acids except methionine, thiamine (2 μg/ml), biotin (2 μg/ml), histidine (100 μg/ml), and preincubated for 20 min at 37°C. 5 μCi of [^{35}S] methionine were added to each sample and incubated further for 1 h at the same temperature.

IX Electrophoretic analysis of proteins

Proteins were analysed by PAGE in sodium dodecyl sulfate (SDS) containing gels as previously described (Laemmli, 1970). The stacking gels were 3% acrylamide and the separating gels were 7.5% acrylamide (1.5 mm thick). The ratio of acrylamide to bis-acrylamide was 30:0.8. Electrophoresis was performed at a constant voltage of 70 for stacking and 120-140 for separation. Gels were cooled with running tap water.

Endoglucanase activity in the gels was detected with modifications of the Congo red-stained agar replica method (Beguin, 1983). After polyacrylamide gel electrophoresis, SDS was removed by washing the gel four to six times for 30 min in buffer A. The first two washes contained 25% isopropanol. The washed and partially dried gel was laid on top of a thin sheet (0.75 mm thickness) of 2% agarose containing 0.1% CMcellulose (High viscosity grade, DP 3000) in buffer A. After overnight incubation at 30°C, zones of CMcellulose hydrolysis were revealed by staining the agarose replica with Congo red.

Protein was visualized by staining with 0.03% Coomassie brilliant blue dissolved in 10% acetic acid and 25% 2-propanol. Excess stain was removed by soaking the gel in 10% acetic acid until background staining was minimal. Destained gels were dried at room temperature in cellophane

sandwiches (Juang *et al.*, 1984).

Radioactive gels containing [^{35}S] methionine-labelled proteins were dried onto Whatman 3MM paper under vacuum and the labelled proteins were located by autoradiography.

X Immunological detection of EngB

0.1 mg of Avicel affinity purified intact EngB (110 kDa) was mixed with Freund's complete and incomplete adjuvant (1:1), and injected subcutaneously at 3 week intervals into a New Zealand white rabbit. Serum was collected a week after the third injection. The antiserum was tested against the intact EngB, using an enzyme-linked-immunoadsorbent assay (ELISA) as described (Voller *et al.*, 1976).

Polypeptides on polyacrylamide gels were detected by immunoblotting as previously described (Towbin *et al.*, 1979), using the alkaline phosphatase/5-bromo-4-chloro-3-indolyl-phosphate (X-phosphate) detection system (Blake *et al.*, 1984).

XI Determination of NH_2-terminal amino acid sequence and amino acid composition of intact EngB

The NH_2-terminal amino acid sequence of intact recombinant EngB was determined by automated Edman degradation using an Applied Biosystems model 470A gas sequenator utilizing the resident sequencing program. The aa residues were analyzed by reversed phase HPLC

chromatography. These analyses were provided by the University of Victoria protein sequencing facility. The aa composition of intact recombinant EngB was determined at the University of Calgary protein sequencing facility.

XII Enzymes and reagents

Growth media components were from Difco. All restriction endonucleases were purchased from Bethesda Research Laboratories, Burlington, Canada or Pharmacia, Dorval, Canada. Calf intestinal phosphatase was from Boehringer Mannheim, Dorval, Canada. IPTG, X-Gal, CMcellulose (low viscosity, DP 400, DS 0.7; high viscosity, DP 3000, DS 0.7), PMSF, ONPG, collagenase, and hide blue powder were from Sigma, St. Louis, Missouri, U.S.A. Acrylamide, bis-acrylamide, Bio-Gel P-D6G and Coomassie brilliant blue were from Bio-Rad, Mississauga, Ontario, Canada. Avicel (Type PH-101) was from PMC International, Cork,Ireland. Nitrocefin was a gift from Glaxo Group Res. Ltd., Greenford, U.K. SDS was from BDH Biochemicals, Toronto, Canada. Radioactive deoxyribonucleotide-5-triphosphates and [^{35}S] methionine were from NEN Research Products, Boston, U.S.A. and Amersham Canada Ltd, Oakville, Canada, respectively. All solvents used for FPLC were of HPLC grade and were obtained from Fisher Scientific, Vancouver, Canada.

RESULTS

I Genetic characterization and increased expression of the cenB gene

A) Determination of the direction of transcription of the cenB gene

The original library of C. fimi genes was constructed by ligating a BamHI digest of genomic DNA into the BamHI site of pBR322 (Gilkes *et al.*, 1984a). E. coli C600 cells were transformed with the ligation mix. Expression of endoglucanase activity in recombinant clones was determined using an immunological plate screening method (Whittle *et al.*, 1982), CMcellulose-Congo red indicator plate or by assaying CMcellulase activity in cell extracts with the DNSA method (Gilkes *et al.*, 1984a). The plasmid pEC3 encodes an endoglucanase (Fig. 4; Gilkes *et al.*, 1984c) on an insert of about 5.6 kb (Fig. 5; Gilkes *et al.*, 1984a). The cloned C. fimi gene (cenB) is different and distinct from cenA and cex, based on the restriction endonuclease maps of their DNA fragments and the degree of randomness of attack of CMcellulose by the crude cell extracts (Gilkes *et al.*, 1984a,c).

In order to identify the transcriptional regulatory signal for the expression of the cenB gene, derivatives of

Figure 4. Detection of *E. coli* C600/pEC3-encoded endoglucanase on LB-CMcellulose agar plate with Congo red.

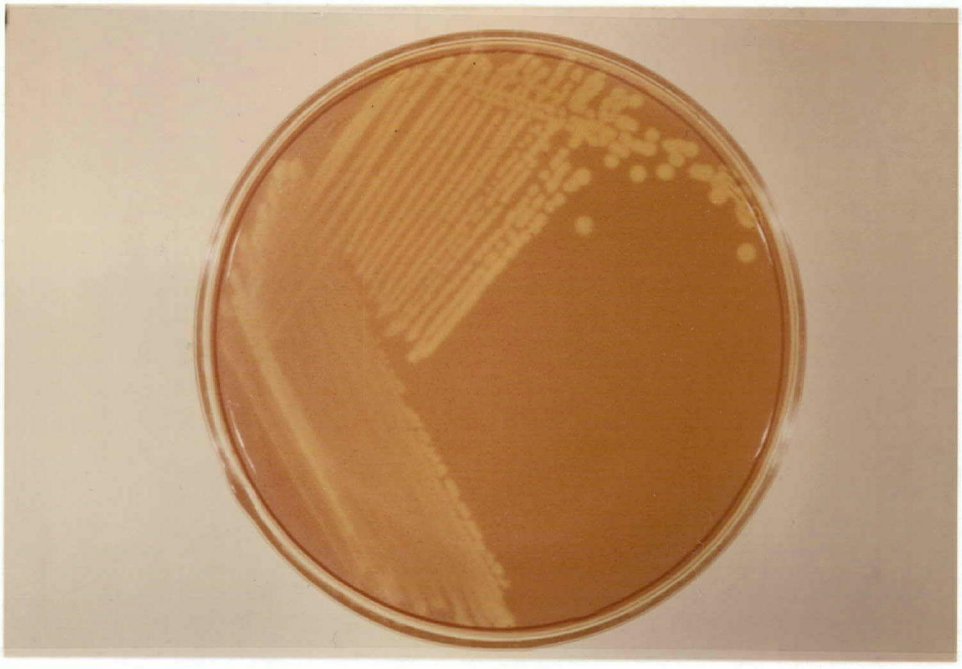
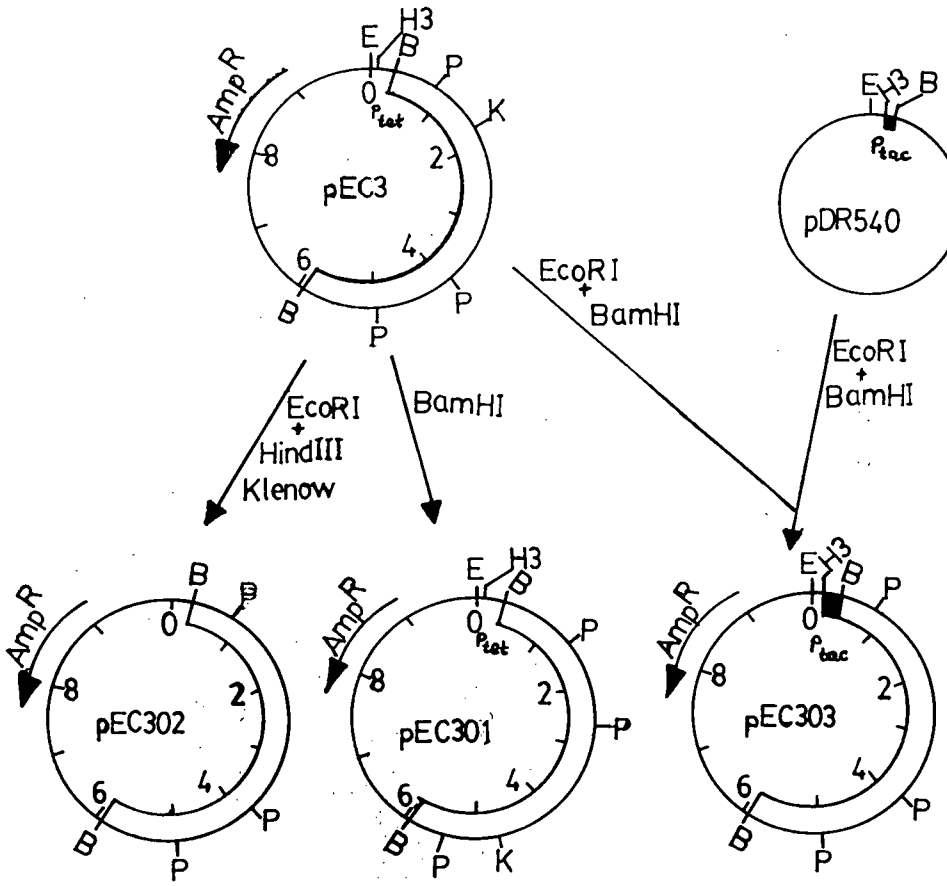


Figure 5. Construction of various pEC3 derivatives. The cenB gene was contained on a 5.6kb BamHI fragment (boxed areas) inserted into the BamHI site of pBR322 to yield pEC3 (Gilkes *et al.*, 1984a; Whittle *et al.*, 1982). To construct pEC301, pEC3 was digested to completion with BamHI and religated; plasmids with inserts in opposite orientation to P_{tet} were confirmed by EcoRI and KpnI double digests which generated 4.7kb and 5.3kb EcoRI-KpnI fragments. pEC3 was restricted with EcoRI and HindIII, the ends were "filled in" in the presence of the Klenow fragment of DNA polymerase I and ligated to obtain pEC302. pEC303 was obtained by substituting the promoter and the first 289 bp of the tetracycline coding region with the tac promoter fragment from the plasmid pDR540 (Russell and Bennett, 1982). pEC3 was restricted with BamHI and EcoRI and the 5.6kb BamHI and 4.0kb EcoRI-BamHI fragments were isolated. The tac promoter was gel-purified as an EcoRI-BamHI fragment from pDR540. Ligation of the tac promoter fragment, the 5.6kb BamHI fragment and the 4.0kb EcoRI-BamHI fragment yielded pEC303. Ptet, tetracycline promoter; P_{tac}, tac promoter. Restriction sites: B, BamHI; E, EcoRI; H3, HindIII; K, KpnI; P, PstI.



pEC3 were constructed (Fig. 5). First, the 5.6 kb BamHI DNA fragment was inverted with respect to the tet promoter in pBR322 (Sutcliffe,1979, Stuber and Bujard, 1981), to give the plasmid pEC301. Secondly, the tet promoter was inactivated by digestion of pEC3 with EcoRI and HindIII, "fill in" of the staggered ends with the Klenow fragment of DNA polymerase I, and circle reclosure. The resulting plasmid was designated pEC302. The plasmid pEC303 (Fig. 5) was constructed by replacing the tet promoter of pEC3 with the tac promoter from pDR540 (Russell and Bennett, 1982). Endoglucanase expression from either pEC301 or pEC302 was drastically reduced (Table III). The specific endoglucanase activity was increased about 7-fold in pEC303 (Table III). These results indicate that the transcription of the cenB gene in pEC3 is dependent upon the tet promoter. The increased expression of the gene in pEC303 further supports this view, and reflects the "strength" of the tac promoter (de Boer and Shepard, 1983).

The electrophoretic pattern of proteins synthesized in minicells of E. coli BD1854 containing pBR322, pEC3 or pEC303 is shown in Fig. 6. In addition to the β-lactamase proteins of pBR322 (laneB) a polypeptide of M_r 110,000 was synthesized in pEC3 and pEC303 (lanes C and D).

B) Delineation of the 5' end of the cenB gene

The strategy for targeting deletions from the 5' end of the 5.6 kb BamHI fragment is shown in Fig. 7. The 5.6 kb

Table III. CMcellulase activities of various *cenB* clones

Host Strain	Plasmid[a]	CMcellulase[b] U/mg
	pEC3	7.61
C600	pEC301	0.28
	pEC302	0.16
	pEC303	51.10
	pJB3	38.90
RR1	pJB301	146.70
	pJB302	167.00
	pJB303	156.70

a) For details of plasmid structure see Figures 5 and 11. pEC3 and pJB3 are the parental clones containing the entire *cenB* gene. pEC301, 302 and 303 are derivatives of pEC3 produced to investigate control of transcription. pJB301 is a gene fusion of *lacZ* and *cenB*. pJB302 and 303 are deletion derivatives of pJB301.

b) CMcellulase activity was determined by assay of reducing sugars with DNSA.

Figure 6. Autoradiogram of polypeptides encoded by pBR322, pEC3 and pEC303. E. coli BD1854 was transformed with plasmids pBR322, pEC3 and pEC303. The proteins encoded by the plasmids were labelled in mini cells (Jensen *et* al., 1984). The labelled proteins were analyzed by SDS-PAGE. Lane A, BD1854; LaneB, BD1854/pBR322; Lane C, BD1854/pEC3; Lane D, BD1854/pEC303. The molecular weight standards were as shown. The positions of the bla and cenB gene products are indicated by arrows. The exposure was too short to visualize the tet gene product.

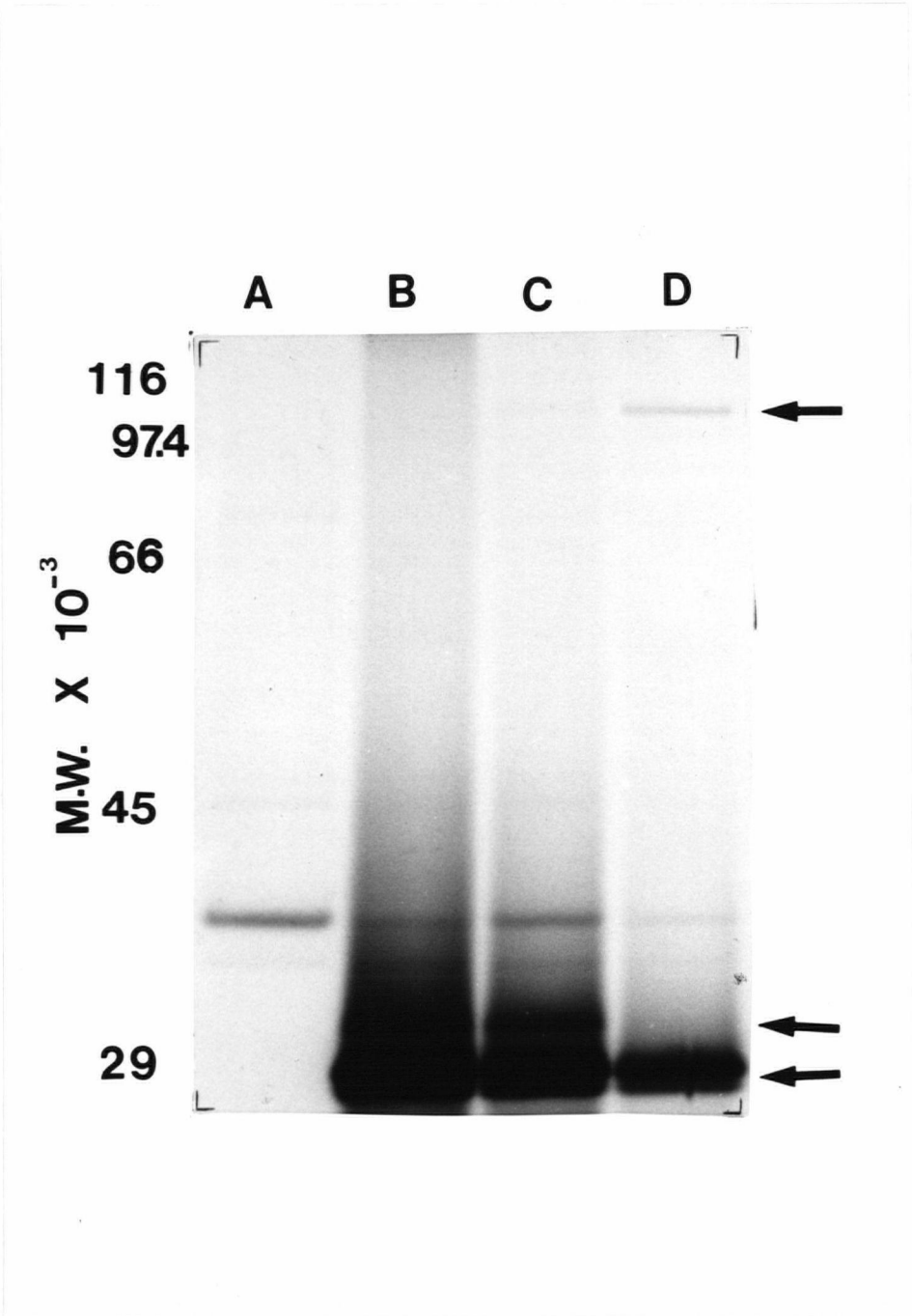
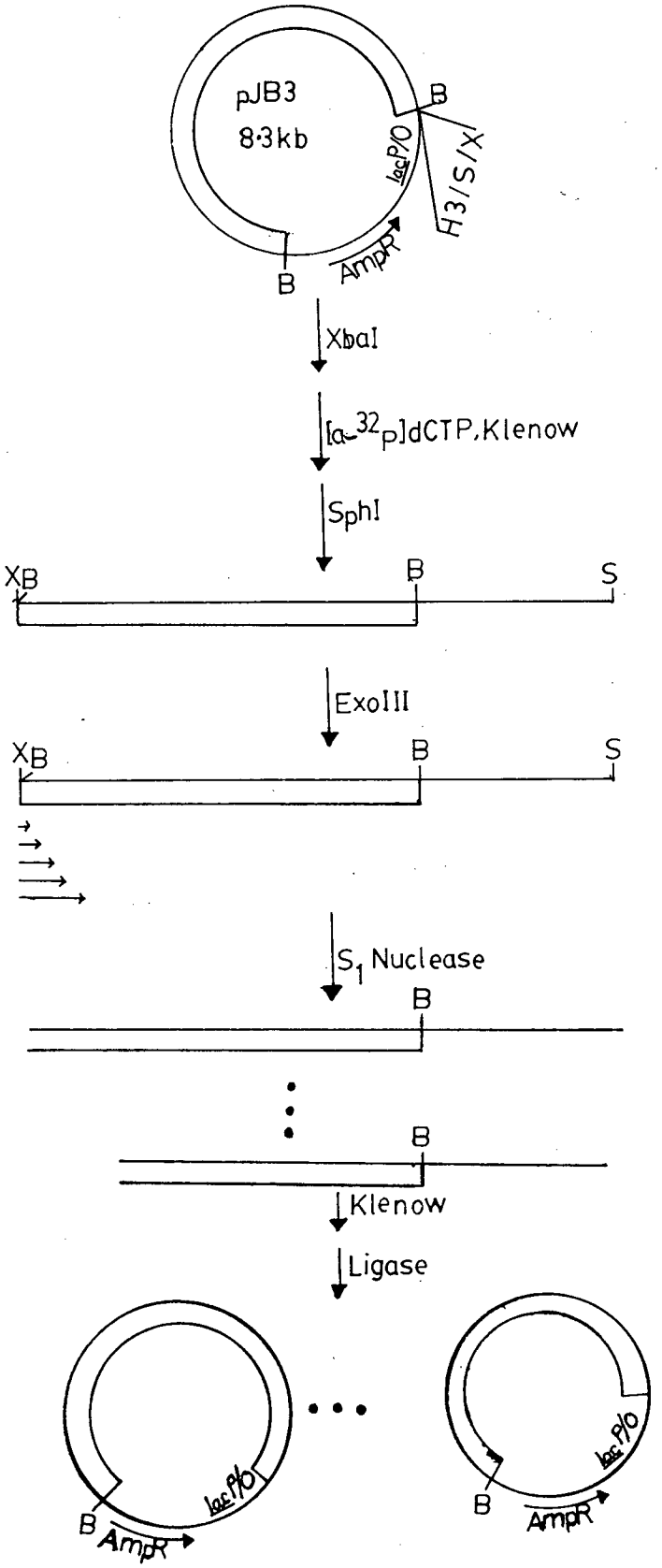


Figure 7. Scheme for targeting deletions from the 5' end of the cenB gene. pJB3 contained the 5.6kb BamHI fragment (boxed area) carrying the cenB gene subcloned into the BamHI site of pUC19 such that the transcription was regulated from the lac promoter. The plasmid was opened at the XbaI site, end-labelled with [α-^{32}P] dCTP in the presence of the Klenow fragment of DNA polymerase I, and digested with SphI. Varying amounts of DNA were removed by Exo III and S_1 nuclease as described (Guo and Wu, 1983; Roberts and Lauer, 1979). Only relevant restriction sites are shown. Restriction sites: B, BamHI; H3, HindIII; S, SphI; X, XbaI.


pJB3
8·3kb
B
lacP/O
A3/S/X
AmpR
B
XbaI
[α-32P]dCTP, Klenow
SphI
XB
B
S
ExoIII
XB
B
S
S1 Nuclease
B
B
Klenow
Ligase
B
AmpR
lacP/O
B
AmpR
lacP/O

BamHI fragment of pEC3 was subcloned into pUC19 such that the 5' end was adjacent to the lac promoter/operator region to give plasmid pJB3. Plasmid pJB3 was linearized at the XbaI site, end-labelled with [α-^{32}P] dCTP and digested to completion with SphI. Then the deletions near the 5'end of the insert were made using Exonuclease III and nuclease S_1 as described (Guo and Wu, 1983; Roberts and Lauer, 1979).

Several clones were examined for endoglucanase activity and insert size (Fig. 8). Deletion of about 100 to 250 bp from the 5' end of the 5.6 kb insert did not affect the level of expression of the cenB gene. However, deletions of more than 385 bp prevented expression of the gene. A deletion mapped at about 325 bp from the 5' end of the insert gave a significantly increased level of endoglucanase activity. The plasmid was designated pJB301; it carries an in-frame fusion between codon 16 for the EngB signal peptide (18 codons before the cleavage site) and codon 7 for the α-fragment of β-galactosidase (Fig. 9).

C) Localization of the 3' end of the cenB gene

The plasmid, pJB301, was cleaved partially with SmaI or PstI and religated. This resulted in the deletion of various lengths of DNA on the 3' side of the cenB gene. Transformants were screened for endoglucanase activity (Fig. 10) and plasmids from active clones were characterized by restriction mapping (Fig. 11). The cells from selected clones were assayed quantitatively for endoglucanase

Figure 8. Extents of deletions into the 5' end of the cenB gene and effects on CMcellulase activity. E. coli RRI Amp^R clones containing various deletion plasmids were screened for CMcellulase activity with Congo red. The plasmids were isolated and the sizes of deletions were determined by restriction enzyme analysis and agarose gel electrophoresis. Only the C. fimi DNA fragments are shown. (a) parental fragment; (b-f) deleted fragments with sizes are shown. Transcription depended on the lac promoter/operator provided by pUC19. CMcellulase phenotypes: "-", inactive; "+", active; "+++", significantly active. Restriction sites: B, BamHI; P, PstI.

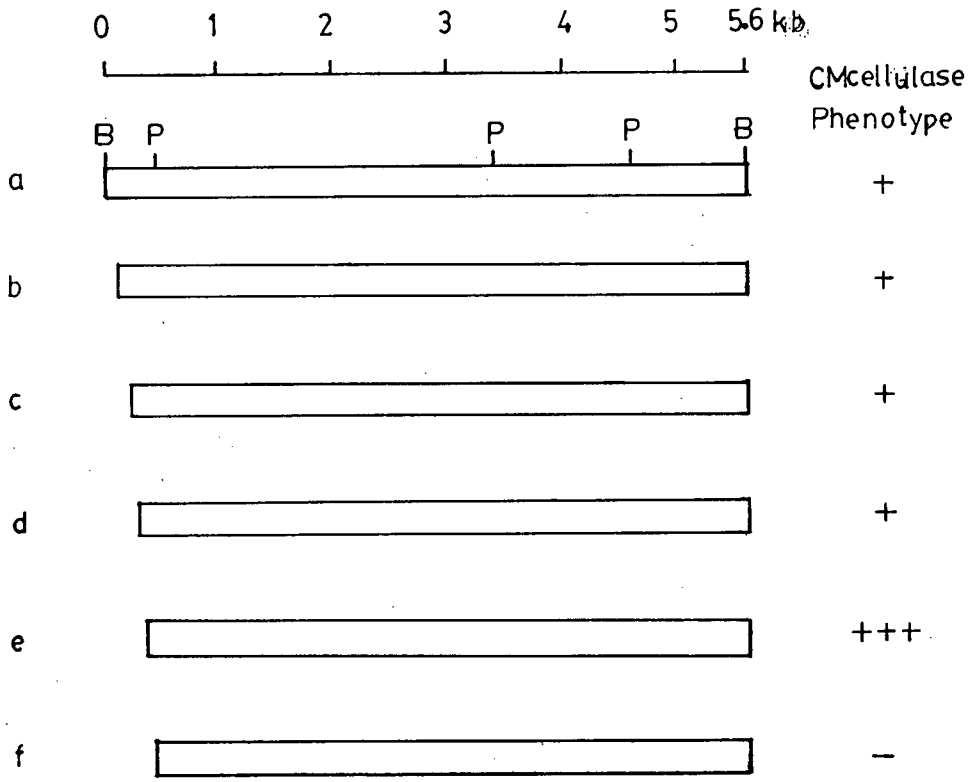


Figure 9. Nucleotide sequence of the RBS, translational initiation site and amino-terminus of the fusion junction of the lacZ'-cenB expression-secretion plasmid, pJB301. The first cenB codon in the fusion plasmid retains its original position number (see Fig. 13). The nucleotides and amino acids derived from lacZ' are underlined.

```
                                                    -18          -15
                          M   T   M   I   T   P   S   L   A   V   A   V   G   V
... ACACAGGAACAGCT ATG ACC ATG ATT ACG CCA AGC CTC GCC GTC GCC GTC GGG GTG

    -10                   -5                -1  +1
 L   V   A   P   L   A   T   G   A   A   A   A   P
CTC GTC GCC CCG CTC GCG ACC GGC GCG GCC GCC GCG CCC ...
```

Figure 10. Screening of various cenB subclones on a CMcellulose-Congo red indicator plate. E. coli RRI cells containing appropriate plasmids were screened for CMcellulase activity as described (Materials and Methods). A, pJB3; B, pJB301; C, pJB302; D, pJB303.

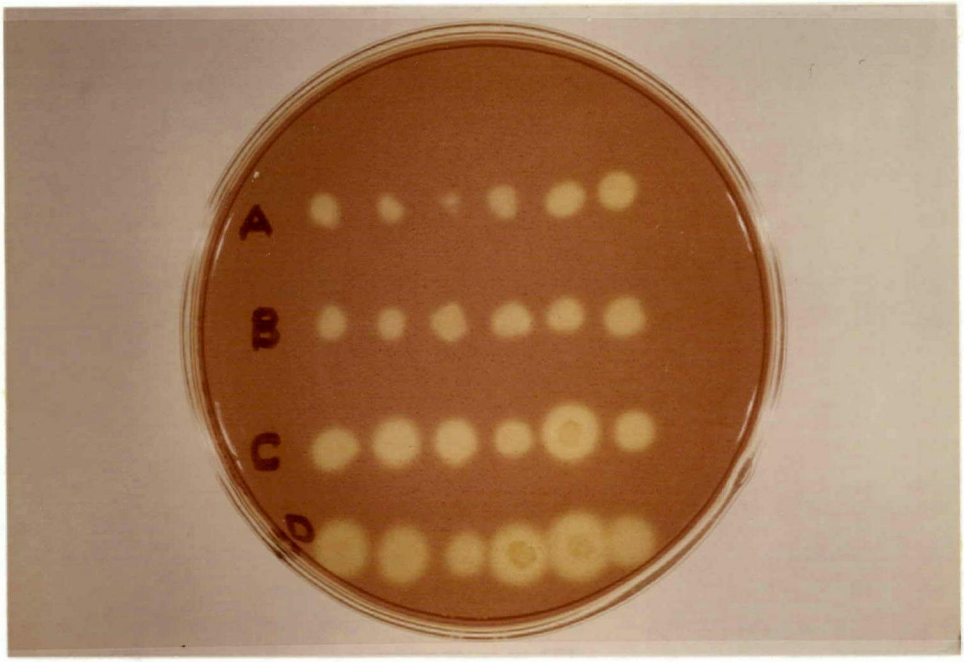
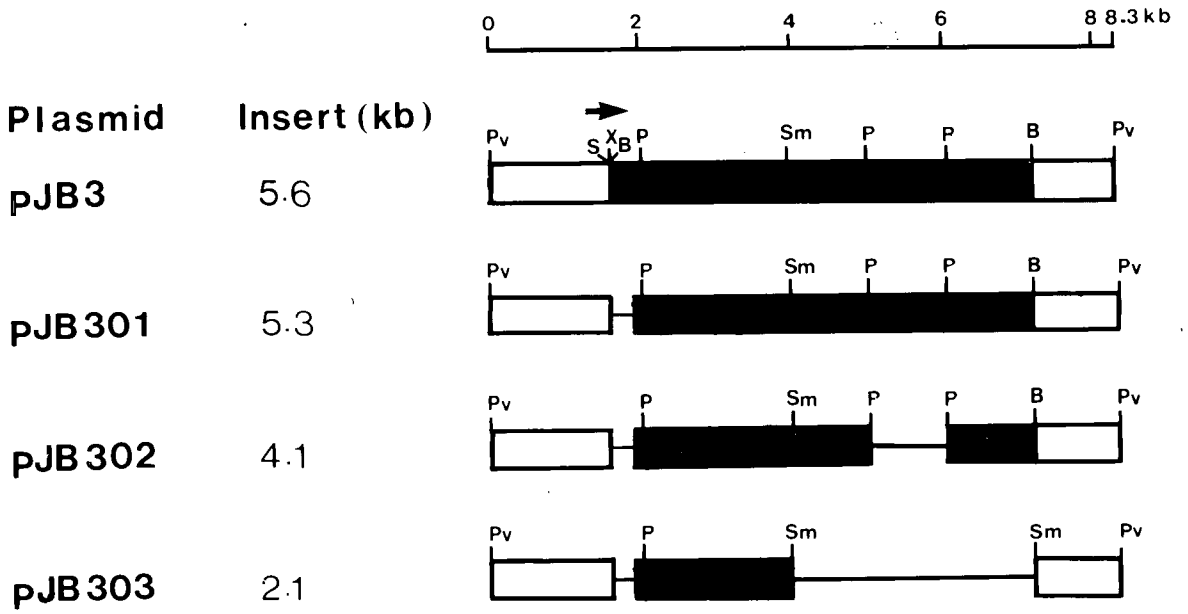


Figure 11. Diagrams of pJB3 and its deletion derivatives. The circular plasmids are shown in a linear fashion for clarity. The open bar represents PUC19 DNA; the solid bar represents C. fimi DNA; the single line represents the regions deleted in each derivative. The arrow at the top indicates the functional orientation for the lac promoter. The length of C. fimi DNA fragment in each plasmid is indicated. The deletion in pJB303 extends to the SmaI site (Sm) of pUC19. Restriction sites: B, BamHI; P, PstI; Pv, PvuII; S, SphI, Sm, SmaI; X, XbaI.



activity (Table III). The shortest, uninterrupted fragment of the 5.3 kb insert which expressed endoglucanase activity equal to that of pJB301 was 2.1 kb long, it was found in pJB303 (Fig. 11, Table III).

Cell extracts of E. coli RRI containing pJB301, pJB302 and pJB303 were subjected to SDS-PAGE analysis. The separated polypeptides were screened for endoglucanase activity (Materials and Methods). This activity was easily detected in SDS gels prepared from samples heated in loading buffer for 2 min at 65°C prior to loading. Control experiments showed no noticeable difference in migration of the polypeptides from gently heated or boiled samples (data not shown). Three active polypeptides of M_r 110,000, 95,000 and 82,000 were observed for pJB301 and pJB302; pJB303 encoded an active polypeptide of M_r 72,000 (Fig. 12)

D) Structure of the 5' terminal region of the cenB gene

The 400 bp BamHI-PstI fragment of pJB3 (Fig. 11) and the deleted inserts generated with Exonuclease III and S_1 nuclease were subcloned into M13mp11 and M13mp18, and the single stranded DNA sequenced using the dideoxy chain termination method (Sanger *et al.*, 1977).

The nucleotide sequence and the deduced amino acid sequence of the C. fimi DNA fragment containing the 5'region of the cenB gene are shown in Fig. 13. The sequence corresponding to the amino terminus of the mature enzyme was located using the amino acid sequence of the protein

Figure 12. SDS-PAGE and zymograms of total cellular proteins from E. coli RRI containing pJB301 or its deletion derivatives. Lanes A,B,C, are total cellular proteins stained with Coomassie brilliant blue. Lane A, RRI/pJB301; Lane B, RRI/pJB302; Lane C, RRI/pJB303. Lanes 1,2, and 3 are zymograms of A,B,C, respectively. The arrows indicate active endoglucanase components. The molecular weight markers were rabbit muscle myosin, 205,000; E. coli β-galactosidase, 116,000; rabbit muscle phosphorylase b, 97,4000; bovine serum albumin, 66,000; ovalbumin, 45,000; carbonic anhydrase, 29,000.

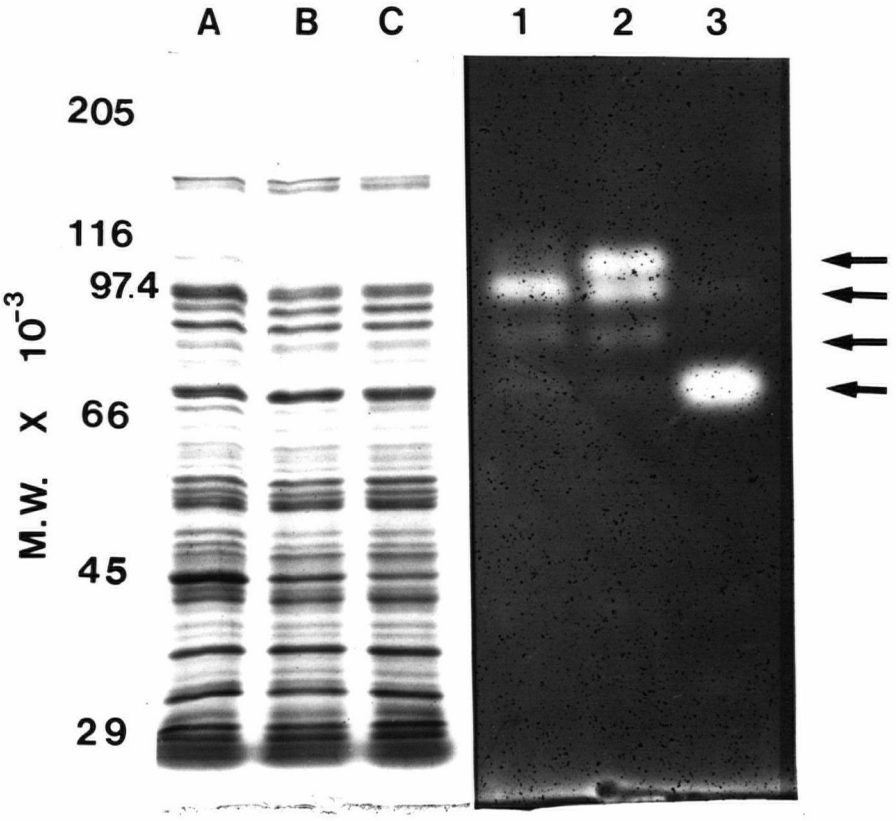
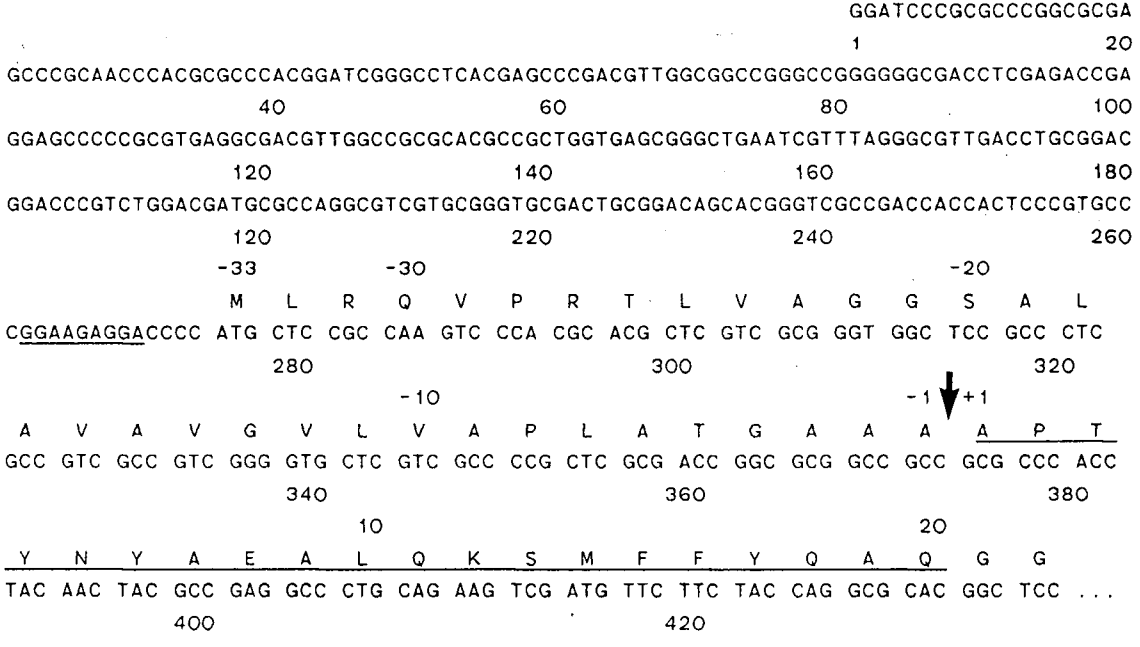


Figure 13. Nucleotide sequence of the 5′ terminus of cenB gene and the deduced NH_2-terminal sequence of EngB. The putative RBS is underlined. The arrow indicates the E. coli signal peptide processing site. The underlined amino acids were determined by automated Edman degradation of the purified intact EngB from E. coli RRI/pJB301.



purified from E. coli (see Results,section III). This sequence is preceded by one encoding a putative signal peptide of 33 amino acids, with a hydrophilic amino-terminus of 7 amino acids, including two arginines, followed by a hydrophobic sequence of 26 amino acids. The translational start codon at nucleotide 275 is preceded by a stretch of nucleotides (GGAAGAGGA) closely resembling other ribosome binding sites (RBSs) (Lofdahl *et al.*, 1983; Gold *et al.*, 1981; McLaughlin *et al.*, 1981; Vasantha *et al.*, 1984). The DNA sequence upstream of the initiation codon contains sequences similar to those of cenA and cex promoter sequences (Fig. 14, Greenberg *et al.*, 1987a). Recently, a cluster of four sites located 24 to 75 bases 5' to the start codon of the cenB gene was identified by S_1 nuclease protection studies (Greenberg *et al.*, 1987b).

II Export of EngB in E. coli

A significant fraction of the endoglucanase encoded by pEC3 was translocated into the E. coli periplasmic space (Gilkes *et al.*, 1984a). This indicates that the EngB signal peptide functions in E. coli. The isolation of the plasmid pJB301 provided an opportunity to study the effect of deleting part of the EngB signal peptide, or forming a hybrid signal peptide, on the export of EngB in this organism. It was also of interest to examine the effect of truncation of EngB on its export.

Figure 14. Comparison of the cenB, cex and cenA 5′ flanking regions. Conserved nucleotides between cenB (this thesis), cex (Greenberg *et al.*, 1987a; O'Neill *et al.*, 1986a), and cenA (Greenberg *et al.*, 1987a; Wong *et al.*, 1986) 5′ flanking regions are denoted by * Mapped mRNA start sites (Greenberg *et al.*, 1987a,b) are underlined.

```
          5'->3'

cenB: GCTG AATCGTTTAGGGCGTTGACCTGCGGACGGACCCGTC TGG ACGATGCG...
      ** * ***** **  *    *  ** ******** ***  * *   * * ***
cex : GCCGAAAT GATTCAGCACCT CCC GCGGACGGGCCCCACGTCACAGGGTGCACC...
        ******  *** *  ** * * ** *  ****    ****   ****   *
cenA: TAGGAAATCC  TCATCCGCT CGC GCCGTGGGGCATT CGTC   GGGTTTCCTCGTCG....
```

Cultures of E. coli RRI containing pJB3, pJB301, pJB302 and pJB303 were harvested in the exponential phase of growth. The periplasmic and cytoplasmic fractions were prepared (Materials and Methods) and assayed for endoglucanase activity. For comparison, β-galactosidase, a cytoplasmic enzyme, and β-lactamase, a periplasmic enzyme, were also measured. About 35% of the endoglucanase activity determined by pJB3 appeared in the periplasm (Table IV). In the fusion plasmid pJB301 and derivatives pJB302 and pJB303, the signal peptide was as described (Fig. 9), and in pJB303, the carboxy terminal region of the mature protein was deleted (Figs. 11 and 12). Nonetheless, endoglucanase activity appeared largely in the periplasms of cells carrying these plasmids (Table IV).

III Purification of intact and truncated EngB from E. coli

Attempts to isolate EngB from E. coli employing an immunoadsorbent scheme described earlier (O'Neill *et al.*, 1986c) were not successful. Therefore a novel affinity chromatography scheme was developed based on the adsorption of the enzyme on Avicel (see Materials and Methods). Osmotic shock fluids of RRI/pJB301 or RRI/pJB303 were adsorbed onto autoclaved Avicel on ice for 1 hour. The Avicel-enzyme complex was washed with buffer A (Materials and Methods), and packed into a short column (2.5cm x 13cm). The adsorbed materials were eluted from the Avicel with a descending, concave gradient of buffer A and water at room

Table IV. Location of EngB, β-lactamase, and β-galactosidase in E. coli RR1 cultures

Plasmid	Fraction	Enzyme activity[a] (and specific activity[b]) EngB[c]	β-lactamase	β-galactosidase
	periplasmic	1.9 (148.3)	770 (53600)	0.85 (67.7)
pJB3	cytoplasmic	3.5 (53.5)	108 (1120)	355 (3669)
	Whole cells	5.4 (38.9)	980 (7210)	427 (4611)
	periplasmic	11.6 (644)	260 (14330)	7.9 (441)
pJB301	cytoplasmic	5.6 (52.3)	12.1 (92)	123 (1156)
	Whole cells	18.1 (146.7)	320 (3400)	142 (1150)
	periplasmic	13.6 (689)	323 (16400)	11.7 (594)
pJB302	cytoplasmic	7.9 (100)	16 (200)	132.2 (1679)
	Whole cells	22.1 (167)	371 (3960)	142 (1515)
	periplasmic	14.1 (726)	420 (21600)	9.1 (467)
pJB303	cytoplasmic	6.9 (68.3)	20 (200)	208 (2037)
	Whole cells	22.6 (156.7)	453 (3700)	265 (1595)

a) Enzyme activity: nmol products released/min/ml culture

b) Numbers in parentheses represent the specific activity: nmol products released/min/mg protein

c) EngB was determined by assay of reducing sugars with DNSA.

temperature. When necessary, pools of endoglucanase-containing fractions were further purified using a Mono Q anion exchange column (Materials and Methods).

A representative elution profile of intact EngB from the Avicel affinity column is shown in Fig. 15. Elution profiles of both intact and truncated EngB from the Mono Q anion exchange column are shown in Figs. 16 and 18, respectively. Intact EngB was purified approximately 50-fold (Fig. 17), with more than 50% recovery of enzymatic activity from the Avicel affinity column (Table V). Truncated EngB was purified approximately 40-fold (Fig. 19), with about 30% recovery of activity from the Avicel column (Table VI). There was no difference in the specific activities of the Mono Q fractions number 22 and 25 of the truncated EngB (Table VI). The specific activity of the truncated EngB was virtually identical with that of the intact enzyme (Tables V and VI). The Mono Q step was unnecessary. It reduced the yield of both intact and truncated EngB 6-fold, and it did not improve their specific activities.

The M_rs of the purified intact and truncated EngB are 110,000 and 72,000, respectively (Figs. 17 and 19), which agree well with the sizes determined by the minicell and zymogram techniques (Figs. 6 and 12).

Figure 15. Affinity chromatography of intact recombinant EngB on Avicel. Shockate obtained from E. coli RRI/pJB301 was mixed with autoclaved Avicel at 0°C for 1 hour. After washing the Avicel-protein complex it was applied to a column (2.5cm x 13cm). Adsorbed proteins were eluted with a descending, concave gradient of buffer A and water at a flow rate of 30ml/hour at 23°C (see Materials and Methods). Fractions of 5 ml were collected. CMcellulase activity was determined using 5µl of each fraction in the DNSA method (Materials and Methods). Only the relevant part of the conductivity gradient is shown.

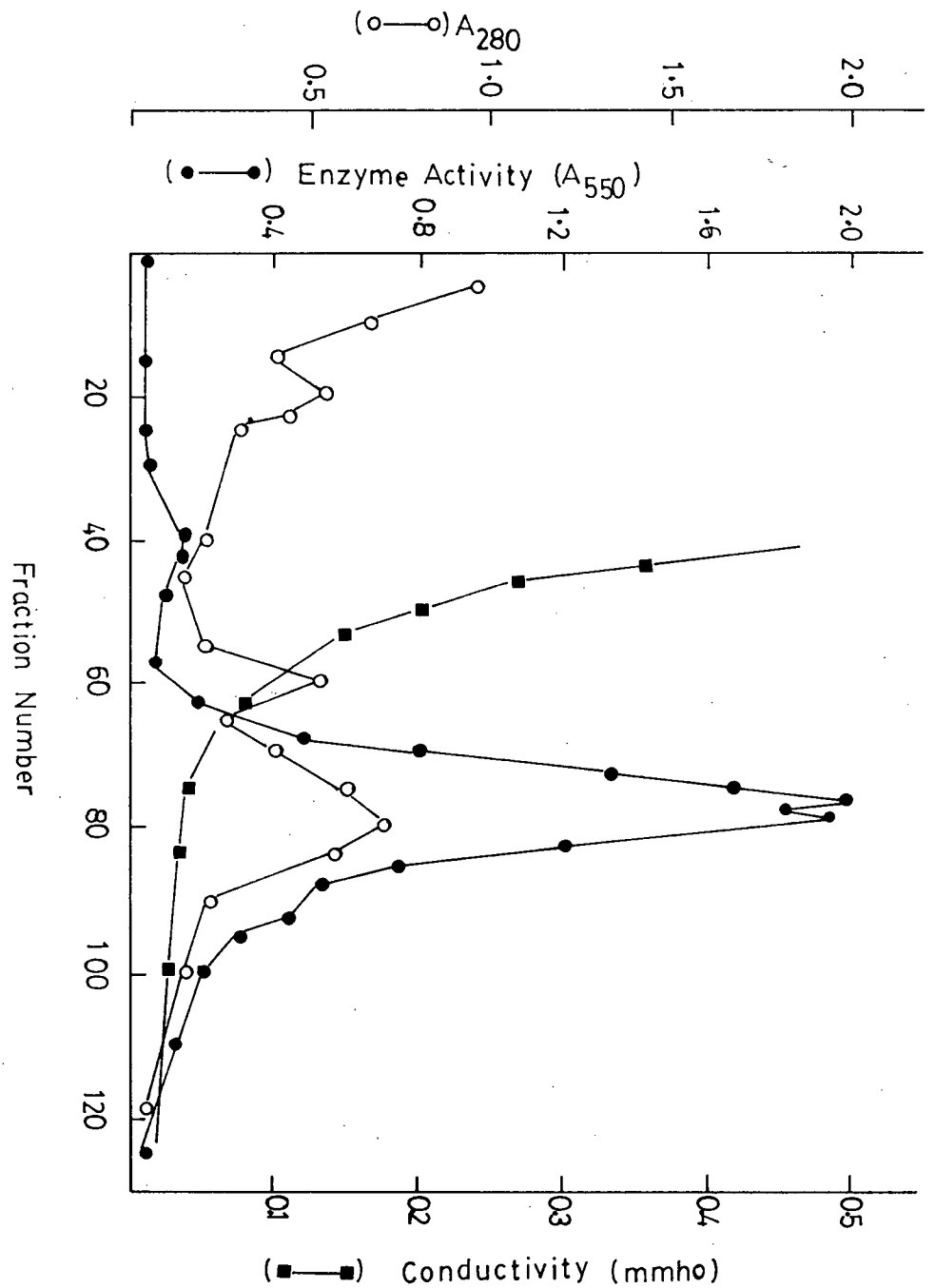

(o—o) A_{280}
0.5
1·0
1·5
2·0
(•—•) Enzyme Activity (A_{550})
0·4
0·8
1·2
1·6
2·0
20
40
60
80
100
120
Fraction Number
0·1
0·2
0·3
0·4
0·5
(■—■) Conductivity (mmho)

Figure 16. Chromatography of intact recombinant EngB on a Mono Q anion-exchange column. The fractions containing CMcellulase activity from the Avicel affinity column were pooled and brought to 20 mM Tris-HCl pH 7.5 (buffer B). The sample (40-50 ml) was pumped onto a Mono Q column equilibrated with buffer B at a flow rate of 1ml/min. The gradient was increased slowly to 15% (0.15 M NaCl in buffer B, 2ml/1% change); then to 25% (0.25 M NaCl in buffer B, 1ml/1% change). Fraction size was 1.0ml. The CMcellulase activity paralleled the major protein peak.

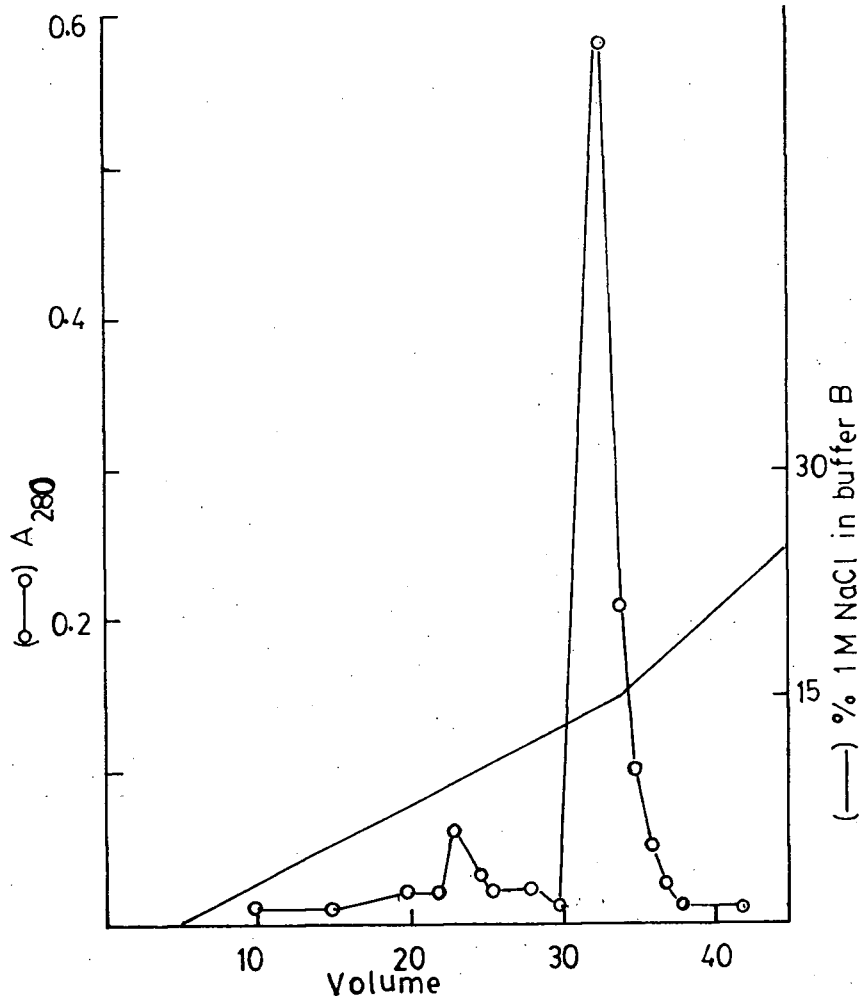


Figure 17. SDS-PAGE analysis of the purification of intact EngB. Samples of various fractions obtained during the purification of intact EngB were analysed on a 7.5% SDS-polyacrylamide gel. Lane A, molecular weight standards ($x10^{-3}$); Lane B, crude cell extract; Lane C, periplasmic fluid; Lane D, pooled active fractions from the Avicel column; Lane E, pooled active fractions from the Mono Q column. The arrow indicates the position of intact EngB.

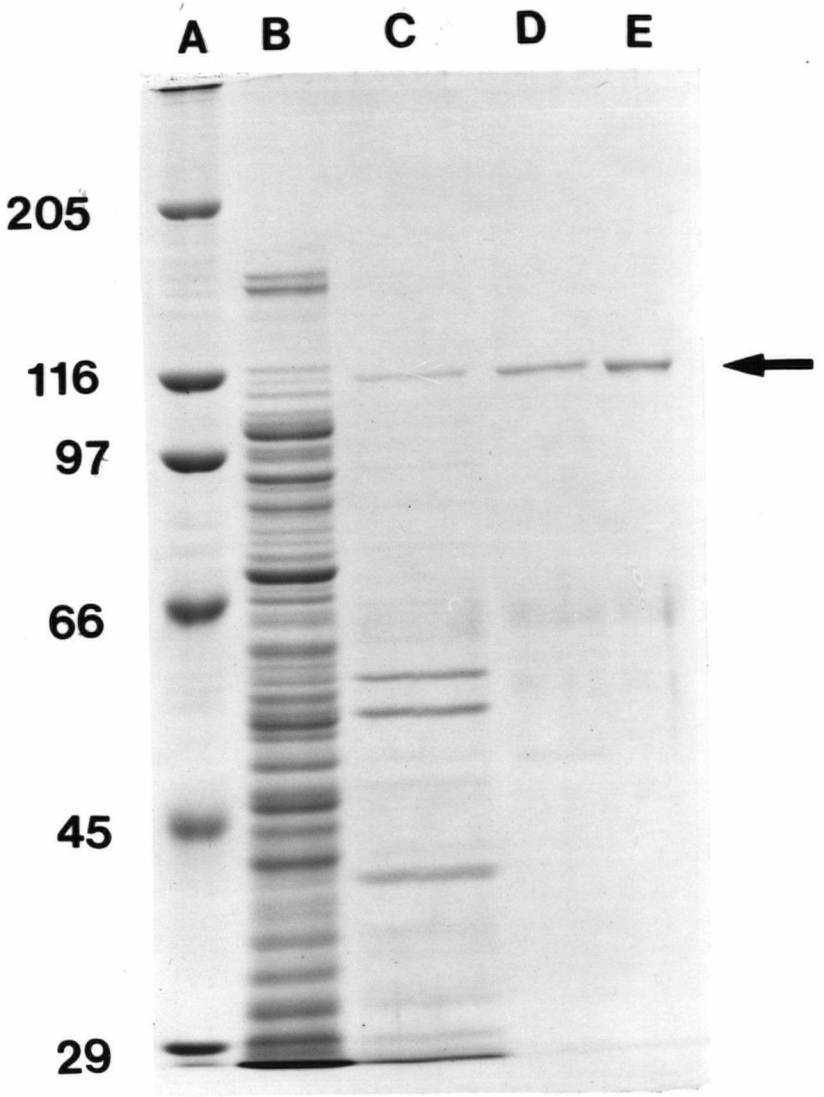


Table V. Purification of intact EngB from E. coli RR1/pJB301

Fraction	Total protein (mg)	Total activity[a] (Units)	Specific activity (Units/mg)	Yield (%)
Periplasmic	38.50	10425	270.7	100
Avicel	0.42	6043	14388.0	58
Mono Q	0.07	950	13571.0	9

a) CMcellulase activity was determined by assay of reducing sugars with DNSA.

Figure 18. Chromatography of truncated recombinant EngB on a Mono Q anion-exchange column. CMcellulase activity pool (100 ml) obtained from Avicel affinity column chromatography of E. coli RRI/pJB303 (data not shown) was concentrated 15-fold by pressure filtration through an Amicon PM 10 membrane. The Amicon retentate (6.5 ml) was brought to 20 mM Tris-HCl, pH 7.5 (buffer B), pumped onto a Mono Q column equilibrated with buffer B at a flow rate of 1ml/min. The gradient was increased slowly to 20% (0.2M NaCl in buffer B, 2 ml/1% change). Peak fraction size was 1.0 ml.

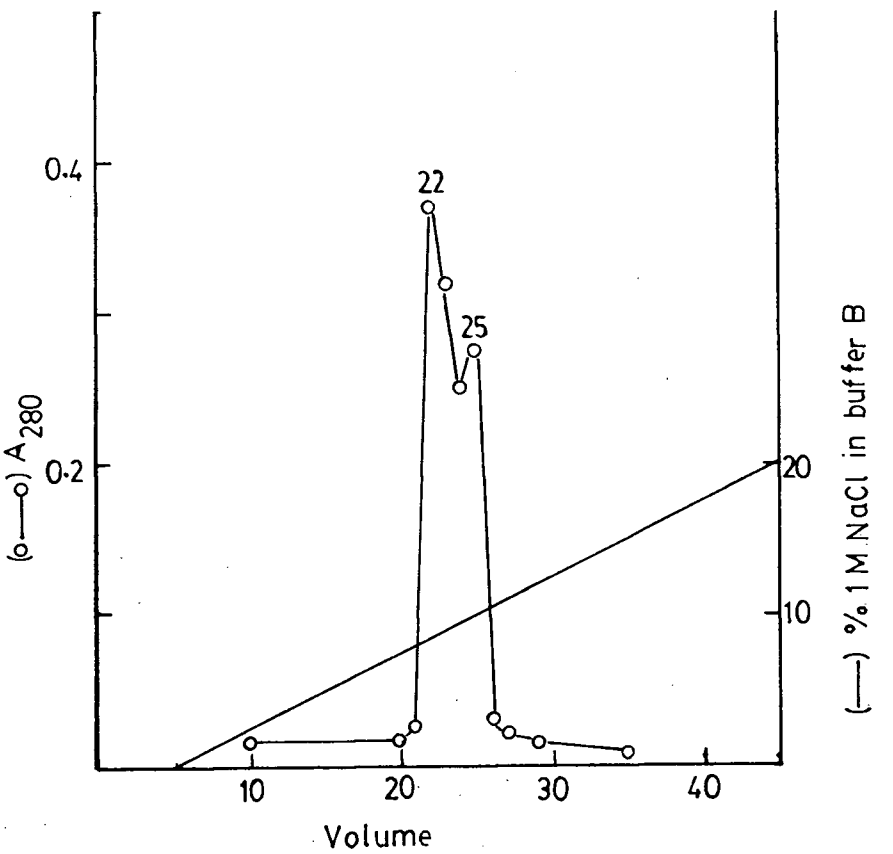


Figure 19. SDS-PAGE analysis of the purification of truncated recombinant EngB. Samples of various fractions obtained during the purification of truncated EngB were analysed on a 7.5% SDS-polyacrylamide gel. Lane A, molecular weight standards ($x10^{-3}$); Lane B, crude cell extract; Lane C, periplasmic fluid; Lane D, pooled active fractions from the Avicel column; Lane E, fraction number 22 from the Mono Q column; Lane F, fraction number 25 from the Mono Q column.

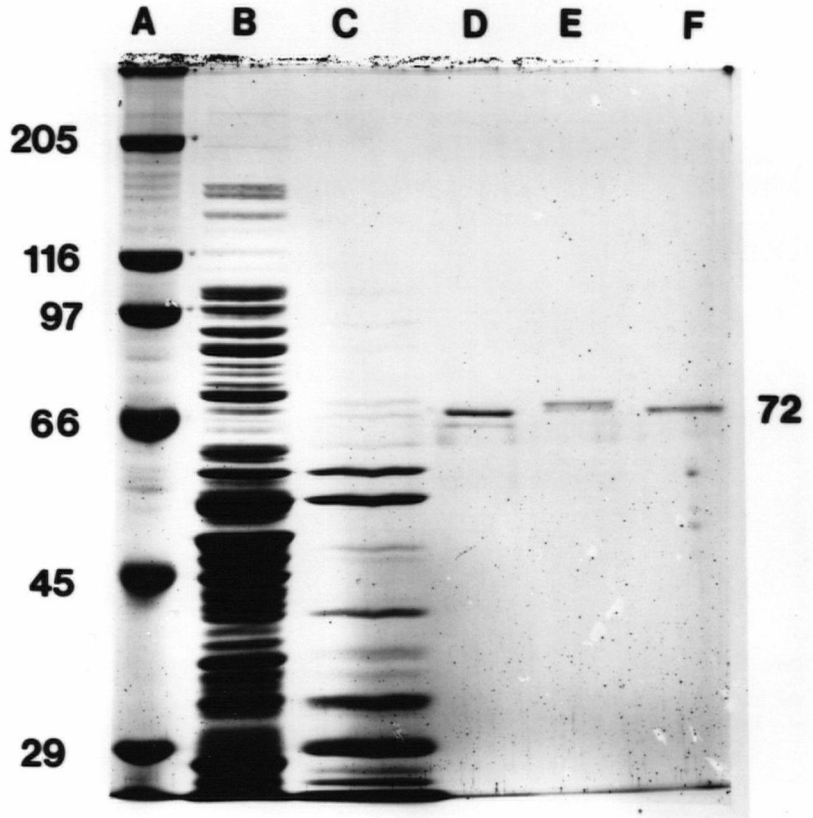


Table VI. Purification of truncated EngB from E. coli RR1/pJB303

Fraction	Total protein (mg)	Total activity[a] (Units)	Specific activity (Units/mg)	Yield (%)
Periplasmic	128.00	40625	317.4	100.0
Avicel	0.94	12188	12925.0	30.0
Mono Q				
Fraction no. 22	0.22	2375	11875.0	5.8
Fraction no. 25	0.12	1417	11808.0	3.5

a) CMcellulase activity was determined by assay of reducing sugars with DNSA.

IV Biochemical characterization of EngB

A) NH_2-terminal amino acid sequence and total amino acid composition of the intact EngB

It was of interest to determine how EngB was processed by E. coli, since both the intact and truncated EngB were exported to the periplasmic space. The first 20 amino acids of the mature intact recombinant EngB are shown underlined in Fig. 13. This reveals the presence of a single NH_2-terminus, the mature product being cleaved at the second Ala-Ala site. The amino acid composition of the purified intact EngB is shown in Table VII.

B) Identification of the C. fimi protein corresponding to EngB

Purified intact recombinant EngB was used to raise antibodies in a rabbit. The antiserum reacted with intact and truncated recombinant EngB (Figs. 20 and 21), and also with a high molecular weight (M_r 110,000) and 4 or 5 lower molecular weight C. fimi proteins. The antibody also reacted with purified native EngA.

C) Action of C. fimi protease on recombinant EngB

C. fimi secretes a serine protease when grown on basal medium plus either CMcellulose, Avicel, or glycerol

Table VII. Amino acid composition of intact EngB polypeptide

Amino acid	Residues[a]/ molecule
Aspartic acid/asparagine	92
Threonine	134
Serine	86
Glutamic acid/glutamine	69
Proline	78
Glycine	109
Alanine	130
Cysteine	8
Valine	85
Methionine	2
Isoleucine	14
Leucine	67
Tyrosine	51
Phenylalanine	30
Histidine	14
Lysine	46
Tryptophan	25
Arginine	25
Total residues	1065
Molecular weight	110,100[b]

a) Average values from three determinations

b) Molecular weight based on the weights of each aa X no. of residues.

Figure 20. Enzyme-linked immunoadsorbent assay of the titre of the antiserum to purified intact recombinant EngB. 100μl aliquots, (0.1μg) of purified intact EngB were transferred onto a multiwell plate. After incubation of the plate for 2 hours at 37°C the wells were washed four times with PBS. Subsequently, the wells were filled to the brim with 2% bovine serum albumin in PBS and the plate was incubated for 1 hour at 37°C to block the remaining protein binding sites on the plate surface. Following this, the wells were washed three times with PBS, filled with 100 μl preimmune (o) or immune (•) antiserum diluted with 0.2% bovine serum albumin in PBS. After incubation of the plate for 16 hours at 4°C, and extensive washing, 100μl of the second antibody, alkaline phosphatase-conjugated goat anti-rabbit IgG, were added at 2000-fold dilution. Subsequently, the plate was incubated for 2 hours at 37°C. The wells were washed, the substrate, p-nitrophenyl phosphate, was added, and the plate was incubated for 1 hour at 37°C. The adsorbance at 405 nm was determined in a Titertek Multiskan.

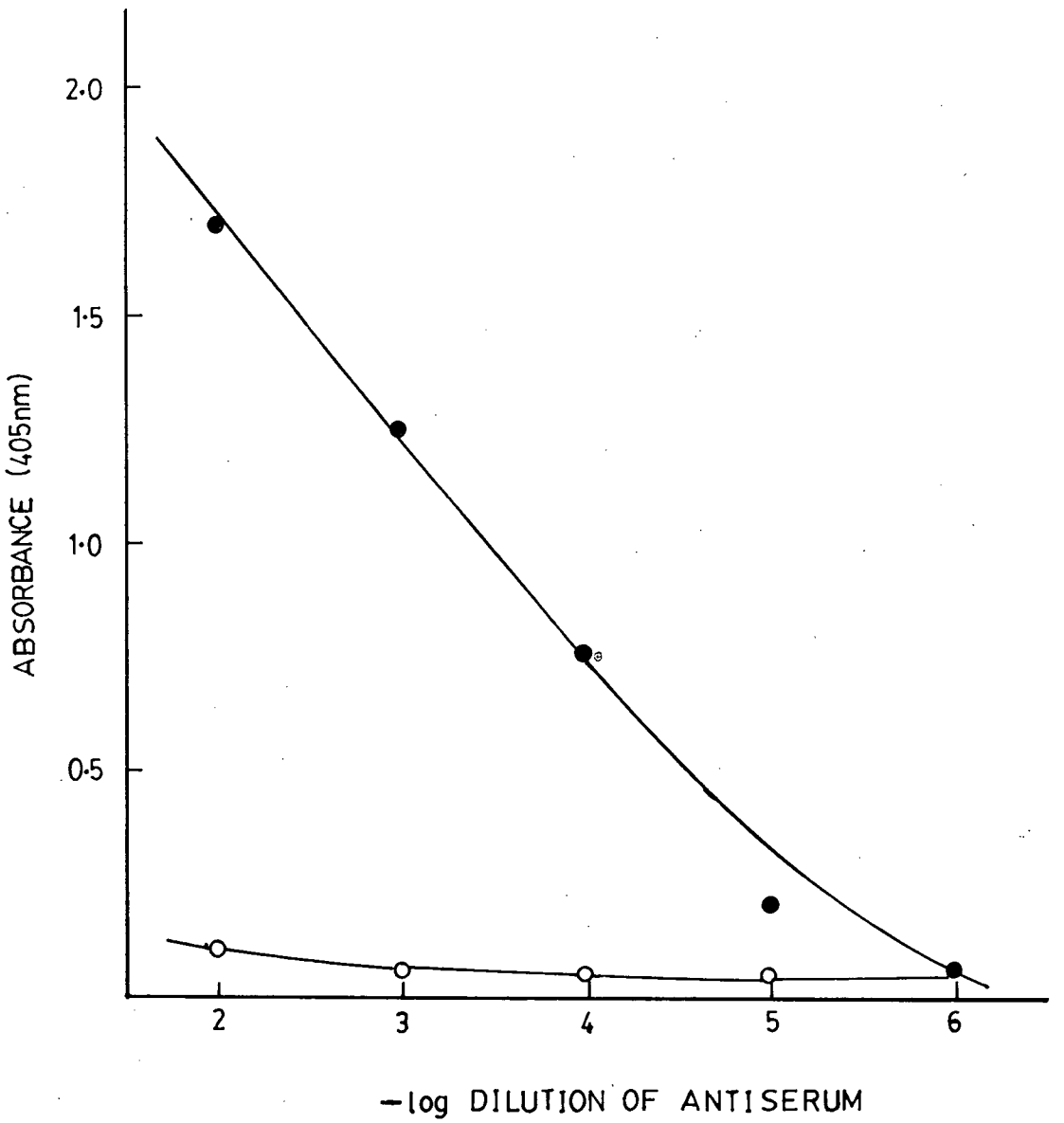
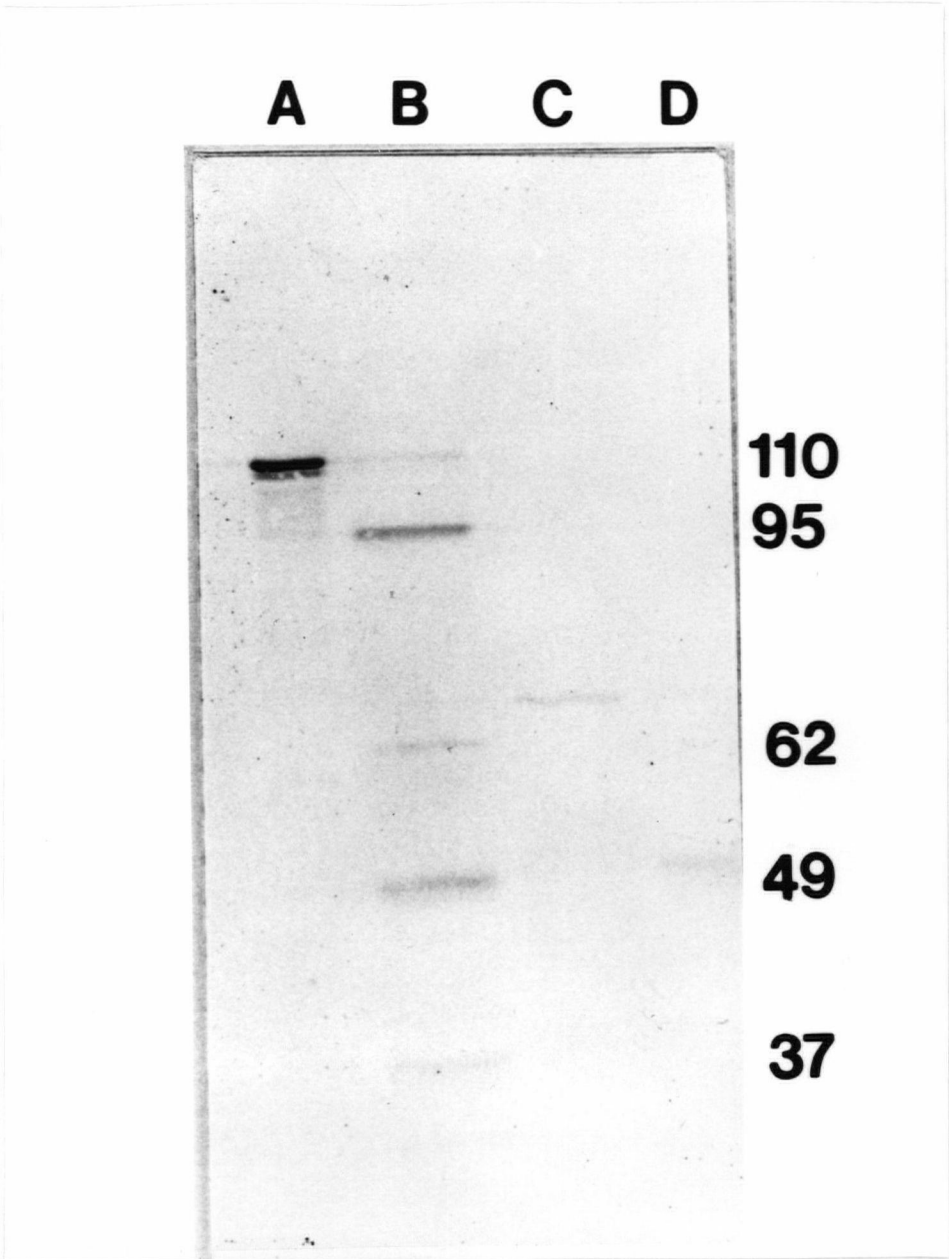

2·0
1·5
1·0
0·5
ABSORBANCE (405nm)
2
3
4
5
6
−log DILUTION OF ANTISERUM

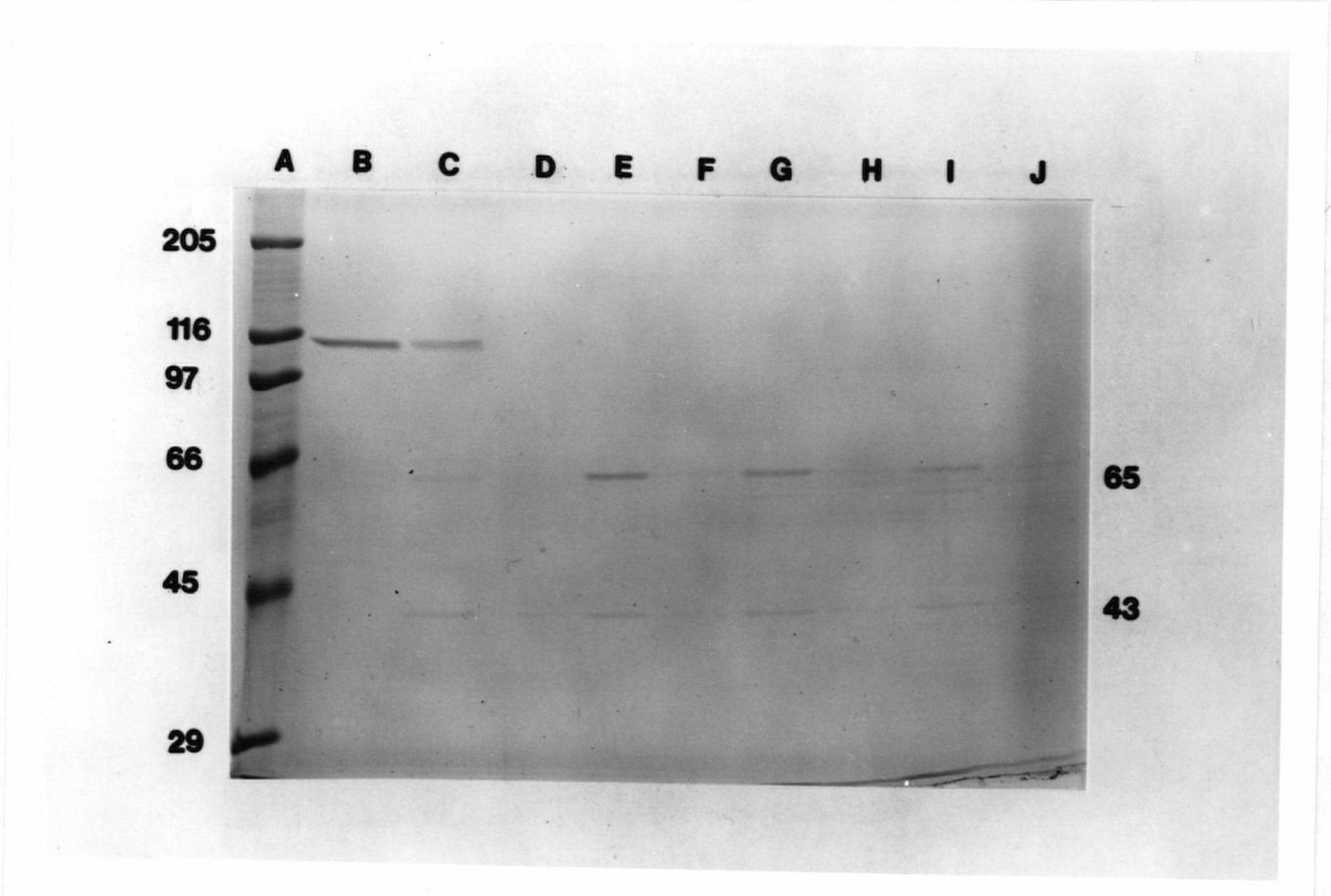
Figure 21. Immunological detection of recombinant EngB and of related polypeptides from C. fimi. The probe was antiserum to intact recombinant EngB. Lane A, intact recombinant EngB; Lane B, C. fimi water-solubilized Avicel bound proteins, 40μg; Lane C, truncated EngB; Lane D, native EngA. The gel was electroblotted onto nitrocellulose. The blot was incubated with rabbit antiserum at a 1:1000 dilution in 1% bovine serum albumin in phosphate buffered saline (1% BSA in PBS) for 16 hours at 4°C. Bound antibody was detected with goat anti-rabbit IgG coupled with alkaline phosphatase. The antibody conjugate was used at 1:2000 in 1% BSA in PBS. The second antibody incubation was for 3 hours at 23°C. The position of bound antibody was detected by the hydrolysis of X-phosphate by alkaline phosphatase. The molecular weights ($x10^{-3}$) are as shown on the right.



(Langsford *et al.*, 1984). It was of interest to examine the susceptibility of the recombinant EngB to crude C. fimi protease. 60 μg of intact EngB powder were dissolved in 2 ml crude extract of C. fimi protease (400 units). Phenylmethylsulfonyl fluoride (100μg/ml, final concentration) was added to 0.5 ml aliquots removed at regular time intervals. Half of each aliquot was examined for the ability of the products to bind Avicel. The other half and the supernatant fractions of centrifuged Avicel-proteolytic product complex were analysed by SDS-PAGE (Fig. 22). Intact EngB was stable in the absence of protease (lane B). During the first hour of incubation with protease, a band corresponding to intact EngB (110 kDa) was present along with two other bands (65 kDa and 43 kDa) (lane C). After 24 hours, only the lower molecular weight bands were seen (lane E). It seemed that longer incubation resulted in further cleavage of the 65 kDa fragment (lanes G,I). The intact EngB and a significant fraction of the 65 kDa fragment were adsorbed by Avicel (D,F,H,J). However, the 43 kDa fragment lacked affinity for the substrate, as shown by its relative abundance in the supernatants after Avicel adsorption (lanes D,F,H,J).

The C. fimi protease has been shown to expose the Pro-Thr box in recombinant Exg and EngA, the proteolytic products being reactive with antiserum to a synthetic Exg Pro-Thr box whereas the intact proteins are unreactive. (Langsford *et al.*, 1987). The antiserum to the synthetic

Figure 22. SDS-PAGE analysis of the effect of the C. fimi protease on intact recombinant EngB. EngB (60μg) was resuspended in 2.0 ml crude C. fimi protease preparation (400 Units) and incubated at 37°C. At given intervals 0.5 ml aliquots (approximately 15μg EngB) were removed and the protease digestion was stopped with PMSF (100μg). Half of each aliquot bound to Avicel and was centrifuged to give supernatant fractions. Samples of EngB before and after protease digestion, and after Avicel adsorption were run on a 7.5% SDS-polyacrylamide gel. Staining was with Coomassie blue. Lane A, molecular weight standards ($x10^{-3}$); intact EngB Lane B, no protease; Lane C, + protease for 1 hour, no Avicel; Lane D, + protease for 1hour, + Avicel; Lane E, + protease for 24 hours, no Avicel; Lane F, + protease for 24 hours, + Avicel; Lane G, + protease for 48 hours, no Avicel; Lane H, + protease for 48hours, + Avicel; Lane I, + protease for 72 hours, no Avicel; Lane J, + protease for 72 hours, + Avicel.



Exg Pro-Thr box was used to probe the intact EngB and its proteolytic products (Fig. 23). The 39 kDa proteolytic product of Exg bound antibody (lane E), whereas a 32 kDa truncated Exg, lacking the Pro-Thr box (Lane F), and the intact Exg (46.8 kDa) (lane D) did not react. Neither intact (110 kDa) (lane A) nor truncated (72 kDa) (lane C) EngB, nor the proteolytic products (65 kDa and 43 kDa) from intact EngB (lane B) reacted with the antiserum. A non-specific band (65 kDa) was seen in all lanes.

D) Substrate specificity of recombinant EngB

Purified intact EngB was tested for its ability to hydrolyse various polysaccharides and aryl-β-glycosides (Table VIII). The results confirmed that EngB is an endoglucanase. It had low but significant activity on lichenan. The activities on Avicel and xylan were 1.4% and 0.5%, respectively, of the activity on CMcellulose. It did not hydrolyse the other glucans nor any of the aryl-β-glycosides.

E) Catalytic properties of intact and truncated EngB

K_m and V_{max} values were determined for the hydrolysis of CMcellulose by purified intact and by truncated EngB (Figs 24 and 25; Table IX).

Figure 23. Western blot analysis of intact and truncated recombinant EngB and proteolytic products of recombinant EngB. The probe was antiserum to synthetic Exg Pro-Thr box. Lane A, intact EngB; Lane B, proteolysis products of intact EngB; Lane C, truncated EngB; Lane D, native Exg; Lane E, 39,000 proteolysis product of recombinant Exg; Lane F, 32,000 truncated product of recombinant Exg. Positions of molecular weight standards ($x10^{-3}$) are as indicated. A non-specific band (65,000) is seen in all lanes.

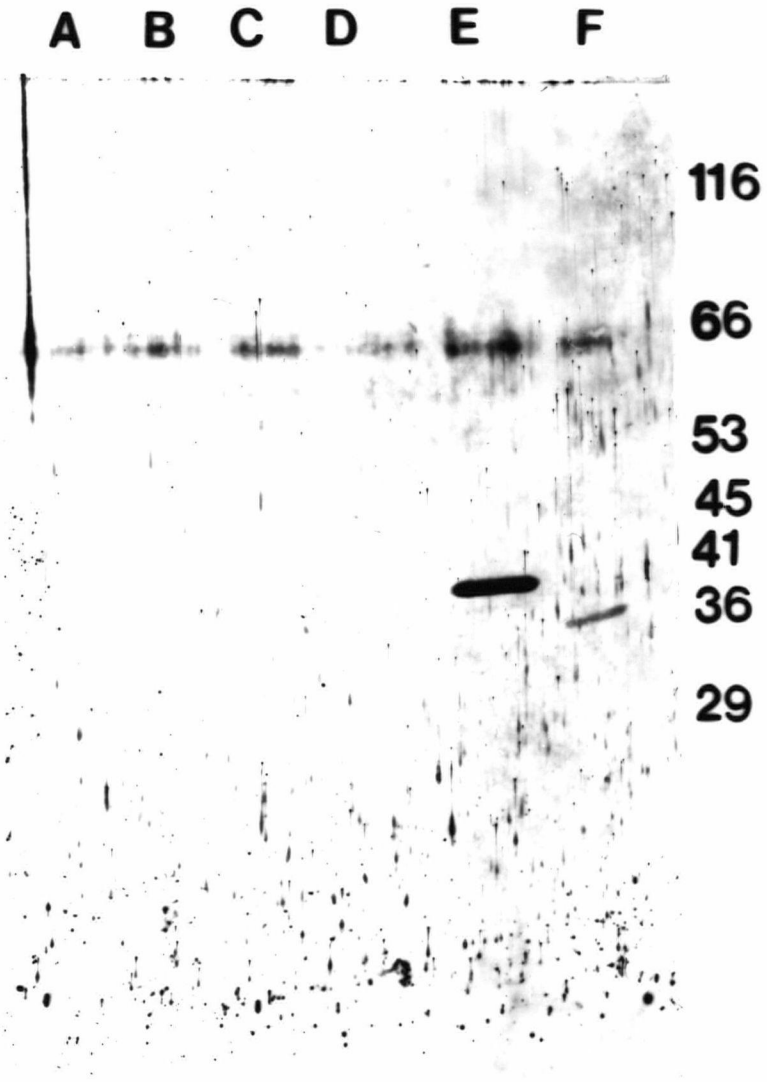


Table VIII. Activity of intact EngB towards various substrates

Substrate	Enzyme specific activity[a]
CMcellulose	17407.8
Lichenan	759.4
Avicel	230.3
Xylan	83.3
Na-polygalacturonate	18.6
Laminarin	3.6
Mannan	<0.01
pNPG	<0.01
pNPC	<0.01
pNPX	<0.01

a) Polymer hydrolysis is expressed as nmol reducing sugar liberated (as glucose equivalents for all polymeric substrates except xylan) /min /mg of protein. Xylanase activity is nmol xylose equivalents/min. Reducing sugars were determined with p-HBAH.

Figure 24. Lineweaver-Burke plot of the kinetics of hydrolysis of CMcellulose by intact recombinant EngB. Initial velocities were determined at 30°C and pH 6.8.

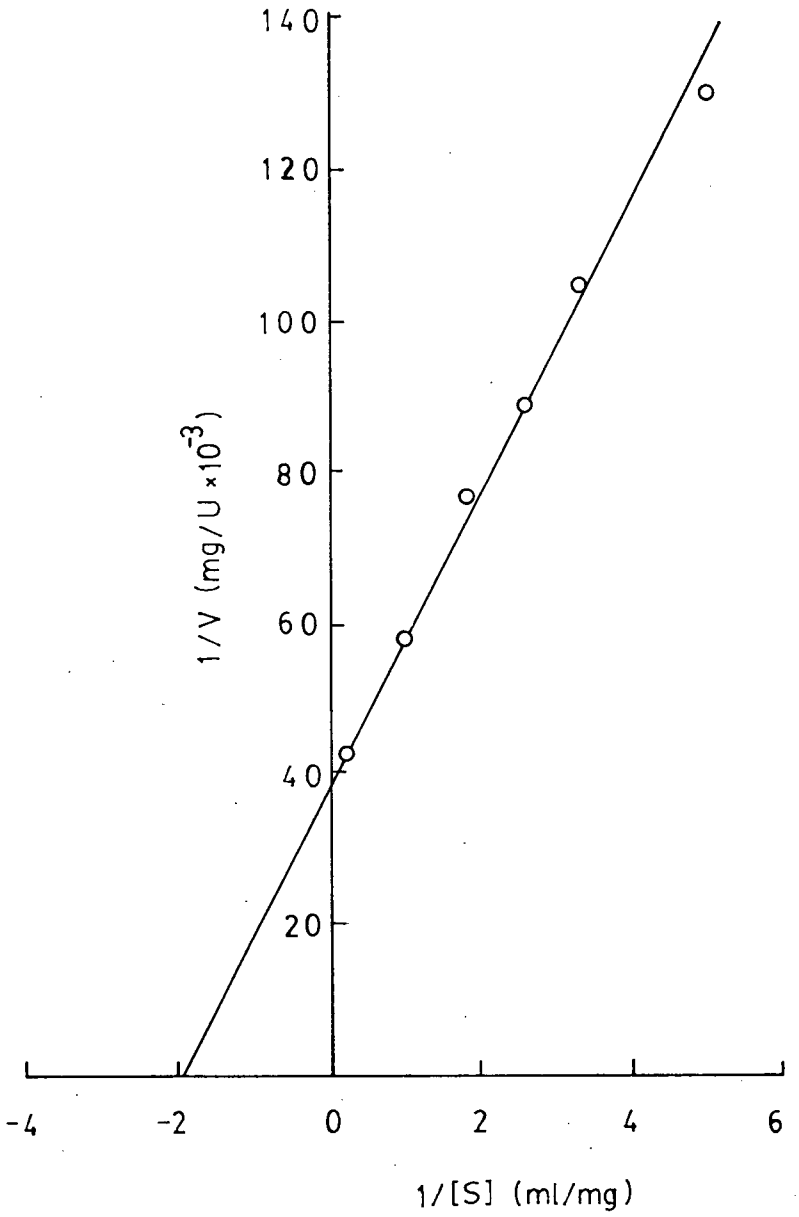


Figure 25. Lineweaver-Burke plot of the kinetics of hydrolysis of CMcellulose by truncated recombinant EngB. Initial velocities were determined at 30°C and pH 6.8.

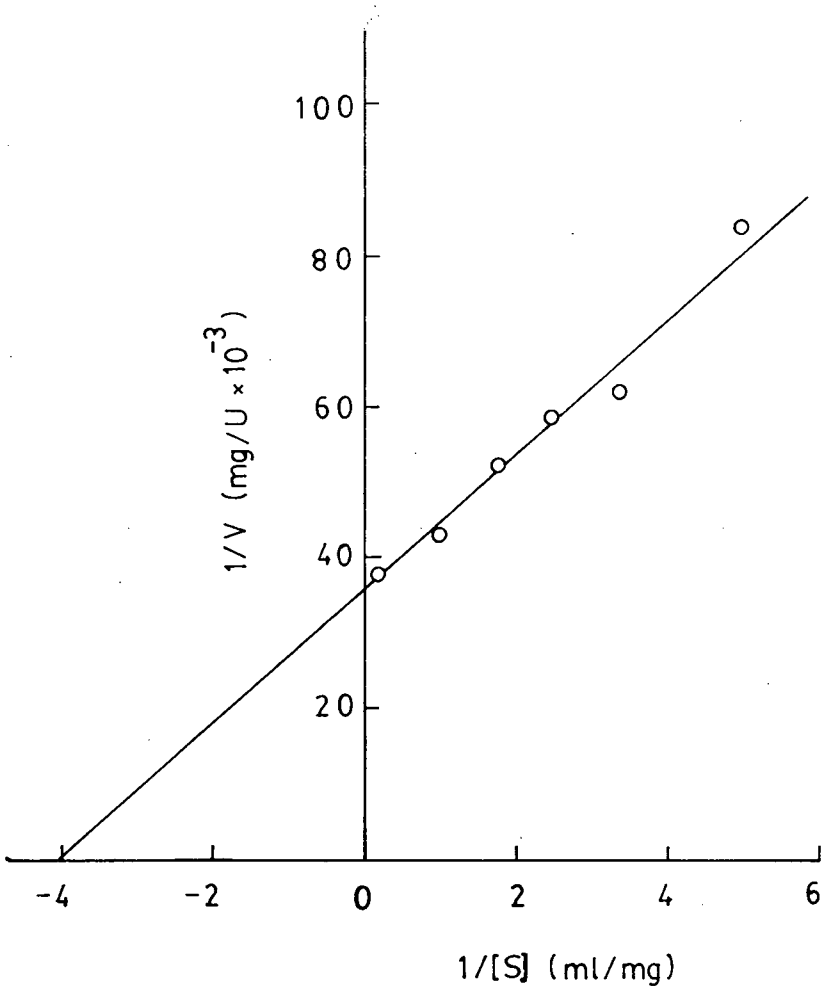


Table IX. Comparison of the kinetic parameters for CMcellulose hydrolysis of intact and truncated EngB, of EngA and of Exg

Enzyme	K_m(mg/ml)	V_{max}(units[c]/mg)
EngA[a]	0.17±0.10	56.60±1.00
Exg[a]	3.04±0.23	35.80±1.90
EngB(110kDa)[b]	0.51±0.05	26.06±1.06
EngB(72kDa)[b]	0.25±0.03	27.77±0.96

a) Values for EngA and Exg are from Langsford *et al.*, 1987.

b) Values were derived from weighted linear regression analysis of the double reciprocal plots, Figs. 24 and 25.

c) Units: μmol glucose equivalents released/min.

DISCUSSION

The cenB gene encodes a polypeptide of M_r 110,000 in E. coli. A polypeptide of similar size from C. fimi, which binds weakly to the substrate in cultures grown with Avicel, has been purified but not sequenced (N.R. Gilkes, personal communication). An intragenic cenB DNA fragment hybridized very strongly to a species of C. fimi RNA approximately 3,200 bases long isolated from CMcellulose-grown cells (Greenberg *et al.*, 1987b). Furthermore, a polyclonal antibody against the EngB purified from E. coli recognizes a C. fimi extracellular protein of M_r 110,000 as well as 4 or 5 polypeptides of lower molecular weight (Fig. 21). One of these immunologically related polypeptides is EngA. Several of the CMcellulase activities in C. fimi culture supernatants were shown earlier to be immunologically related (Langsford *et al.*, 1984).

The sequence upstream of the translational start site of cenB does not contain an E. coli-like promoter sequence (Rosenberg and Court, 1979), but it does contain sequences similar to those of two other C. fimi promoters (Greenberg *et al.*,1987a). Transcription of cenB in C. fimi is directed from two tandem promoters; the distal cenBp1 regulated promoter and the proximal cenBp2 constitutive promoter (Greenberg *et al.*, 1987b). These regulatory elements are weak or non functional in E. coli, like those of the cenA

and cex genes of C. fimi (O' Neill *et al.*, 1986c; Wong *et al.*, 1986).

When the cenB coding sequence is fused in-frame to the E. coli lac RBS and the sequence coding the first 7aa of β-galactosidase, and placed under the control of the lac promoter on the plasmid vector pUC19, the level of EngB obtained in E. coli is about 0.2% of the total cellular protein in this organism. This agrees with an an earlier observation on the expression of the cex gene in E. coli (O'Neill *et al.*1986c). A hybrid lacZ'-cex gene causes the production of Exg constituting about 0.2% of the total cellular protein (O'Neill *et al.*, 1986c). When the cenB coding sequence is fused to a synthetic RBS and an initiating ATG, and placed under the control of the leftward promoter of bacteriophage lambda contained on the runaway replication plasmid vector pCP3, the level of EngB expression is very low (See Appendix). This contrasts the report of the cex gene expression in a similar construct (O'Neill *et al.*, 1986c). The level of Exg expression obtained in E. coli from a pCP3-cex plasmid is more than 20% of total cellular protein. The reasons for this difference in the expression of the cenB and cex genes from the same high-level expression vector are not clear.

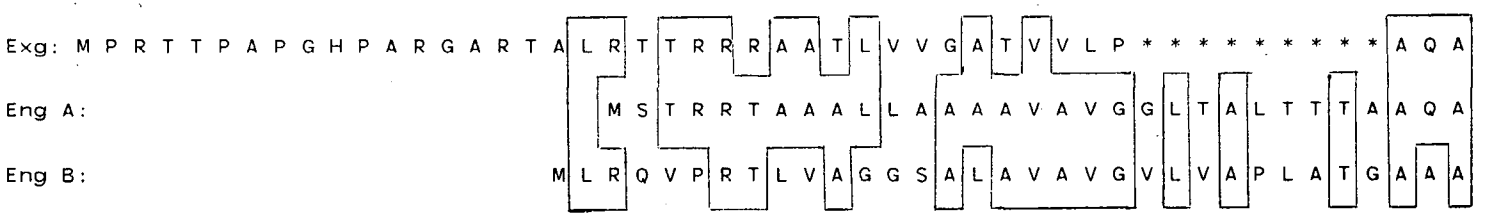
Like many other genes coding for a secreted product, the cenB gene codes for a signal sequence at the amino-terminal end of the protein. Such sequences are believed necessary for efficient export of a protein in most

Gram-negative organisms and are usually composed of a short stretch of charged amino acids (2 to 11 residues) at the amino terminus followed by a longer stretch (14 to 20) residues of strongly hydrophobic amino acids (Inouye and Halegoua, 1980; Silhavy *et al.*, 1983). Signal sequences from Gram-positive organisms have the same structure, but are often slightly longer than those of other systems (Lampen *et al.*1984; Lofdahl *et al.*, 1983; Murphy *et al.*, 1984; Robson and Chambliss, 1987).

Despite its unusual length the EngB signal peptide allows export of EngB in E. coli. Furthermore, replacement of its basic amino-terminal section with the amino-terminal amino acids of β-galactosidase does not block processing and export of EngB. Deletion of the basic amino acids in the signal peptide of E. coli lipoprotein, or their replacement with neutral amino acids, has little effect on lipoprotein export (Inuoye *et al.*, 1982; Vlasuk *et al.*, 1983). However, their replacement with negatively charged amino acids reduced lipoprotein export drastically.

The significance of the hydrophobic region of the EngB signal peptide remains to be determined. The EngA, EngB, and Exg signal peptides exhibit extensive homology in their hydrophobic carboxy terminal sequences (Fig. 26; O'Neill *et al.*, 1986a,b; Wong *et al.*, 1986). The conservation of these sequences, especially in EngA, EngB, suggests an essential role for the structure in protein export and processing. Some mutations in the hydrophobic regions of

Figure 26. Comparison of the Exg, EngA and EngB signal peptides. The amino acid sequences of Exg, EngA and EngB were deduced from the nucleotide sequences (O'Neill *et al.*, 1986a; Wong *et al.*, 1986; this thesis). Conserved residues are boxed. * denotes a gap left in the sequence.



the leader peptides of E. coli lamB and maltose binding proteins prevent export of these proteins (Bedouelle *et al.*, 1980; Emr *et al.*, 1980).

The adsorption of intact and truncated recombinant EngB on Avicel allowed the development of an affinity chromatography for their purification. The procedure is relatively simple, fast, inexpensive and efficient. Autoclaving the Avicel resulted in better adsorption of the recombinant cellulases. The procedure yields EngB of a sufficient purity for amino acid sequence determination, amino acid composition analysis and the raising of polyclonal antibody. Moreover, it is applicable to the purification of Exg and EngA from E. coli, and it also facilitates substrate binding studies (Langsford *et al.*, 1987).

It is noteworthy that 25% of the residues in EngB are hydroxyamino acids (Table VII). Both EngA and Exg contain 20% of hydroxyamino acids, which tend to occur in clusters (Warren *et al.*, 1986). Other glucanases are also rich in hydroxyamino acids: 20% in an α-amylase from Bacillus subtilis (Yang *et al.*, 1983); 22% in a β-glucanase from the same organism (Murphy *et al.*, 1984); 27 and 28%, respectively, in CBHI and endoglucanase from T. reesei (Enari and Niku-Paavola, 1987; Shoemaker *et al.*, 1983, Penttila *et al.*, 1986); 21 and 17% in two endoglucanases from C. thermocellum (Beguin *et al.*, 1985; Grepinet and Beguin, 1986); 30% in an α-amylase from Aspergillus niger

(Boel *et al.*, 1984). All of these enzymes contain clusters of hydroxyamino acids, and it has been suggested that they function as binding domains for insoluble cellulose (Langsford *et al.*, 1987; Warren *et al.*, 1988; van Tilbeurgh *et al.*, 1986).

Preliminary experiments suggested that EngB is an endoglucanase (Gilkes *et al.*, 1984c). cenB clones expressing enzymatic activity form halos on Congo red-stained CMcellulose plates, a feature commonly shared by endoglucanases (Bartley *et al.*, 1984; Teather and Wood, 1982). Purified EngB hydrolyses CMcellulose more than other polysaccharides. Lichenan is a mixed glucan containing β-1,4 and β-1,3 linkages. Presumably, only the β-1,4 linkages are labile since the purified EngB does not attack laminarin, a predominantly β-1,3-linked glucan. The low activity on Avicel suggests that it hydrolyses the amorphous but not the crystalline regions of this substrate. The enzyme does not hydrolyse cellobiose, pNPG or PNPC and is therefore not a cellobiase or exoglucanase. The specific activity of purified EngB on CMcellulose is in the same range as that of EngA and Exg (Table IX; Langsford *et al.*, 1987).

Proteins of M_r 110,000, 95,000 and 82,000 are enzymatically active. It is likely that all three polypeptides are encoded by the cenB gene, whose product is partially hydrolysed in the E. coli cells. Furthermore, a truncated EngB of M_r 72,000 also

possesses enzymatic activity similar to those of the original protein. Taken together, these results provide evidence that the intact EngB is not required for enzymatic activity and that the multiple activities encoded by the complete gene may result from proteolytic processing of the intact product from the carboxyl terminal region. An active truncated product of the alkaline cellulase gene from an alkalophilic Bacillus sp. has been reported (Fukumori *et al.*, 1987). Multiplicity of products of other cellulase genes cloned in E. coli has also been reported: the celA and celB genes from Clostridium thermocellum (Cornet *et al.*, 1983a, Beguin *et al.*, 1983); the endoglucanase gene from Bacteroides succinogenes (Taylor *et al.*, 1987), the endoglucanase gene from Bacillus subtilis (Robson and Chambliss, 1986); and the cex and cenA genes from C. fimi (Z. Guo and N. Arfman, personal communication).

C. fimi protease cleaves intact EngB specifically to generate a truncated polypeptide with both enzymatic activity and substrate binding capacity. It is not presumptuous to expect a similar effect in vivo. Proteolysis has been proposed to account for the multiplicity of C. fimi cellulases (Langsford *et al.*, 1984). Although the polypeptides of M_r 95,000, 62,000 and 37,000 reactive with the antiserum to EngB may be proteolytic products of the intact native EngB, it is equally possible that they are as yet uncharacterized components of the C. fimi cellulase system. What ever their nature, the

antiserum to EngB could be useful for the purification of these proteins by immunoadsorption chromtography.

EngB hydrolyses CMcellulose and binds to Avicel, like EngA and Exg (Gilkes *et al.*, 1984a; Langsford *et al.*, 1987). However, while the intact EngA and Exg are necessary for these two functions (Langsford *et al*, 1987; Warren *et al.*, 1988), the carboxyl terminal one-third of EngB is required for neither. The role of this apparently dispensable region of EngB in cellulose hydrolysis remains to be determined.

It is not certain that the 110,000 C. fimi protein weakly bound to Avicel, purified earlier (N.R. Gilkes, personal communication) or immunologically characterized in this thesis, has the same enzymological properties as those described for the E. coli-made enzyme. However, EngA and Exg, unglycosylated in E. coli, both retain the specificities of the native enzymes from C. fimi (Langsford *et al.*, 1987).

LITERATURE CITED

Alexander, J.K. (1968) Purification and specificity of cellobiose phosphorylase from Clostridium thermocellum. J. Biol. Chem. 243:2899-2904.

Almin, K.E., Eriksson, K.E., and Pettersson, L.G. (1975) Extracellular enzyme system utilized by the fungus Sporotrichum pulverulentum (Chryosporium lignorum) for the breakdown of cellulose. 2. Activities of the five endo-β-1,4-glucanases towards carboxymethylcellulose. Eur. J. Biochem. 51: 207-211.

Appleyard, R.K. (1954) Segregation of new lysogenic types during growth of a doubly lysogenic strain derived from Escherichia coli K-12. Genetics 39:440-452.

Arfman, N., Gilkes, N.R., Guo, Z., Kilburn, D.G., Langsford, M., Miller, R.C., Jr., Owolabi, J., and Warren, R.A.J. (1987) Characterization of Cellulomonas fimi cellulases. In Proceedings of the Sixth Canadian Bioenergy R & D Seminar, in press.

Armentrout, R.W., and Brown, R.D. (1981) Molecular cloning of genes for cellobiose utilization and their expression in Escherichia coli. Appl. Environ. Microbiol. 41: 1355-1362.

Atkinson, T., and Smith, M. (1984) In Oligonucleotide Synthesis: a practical approach. Gait, N.J. ed. IRL press, Washington.

Bailey, M.J., and Oksanen, J. (1984) Cellulase production by mutant strains of Trichoderma reesei on non-cellulosic media. In Proceedings of the Third European Congress of Biotechnology, volume 2:157-162 Verlag Chemie Weinheim.

Barras, F., Boyer, M.H., Chambost, J.P., and Chippaux, M. (1984) Construction of a genomic library of Erwinia chrysanthemi and molecular cloning of cellulase genes. Mol. Gen. Genet. 197:513-514.

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

Bartley, T.D., Murphy-Holland, K., and Eveleigh, D.E. (1984) A method for the detection and differentiation of cellulase components in polyacrylamide gels. Anal Biochem. 140:157-161.

Bedouelle, H., Bassford, P.J., Fowler, A.V., Zabin, I., Beckwith, J., and Hofnung, M. (1980) Mutations which alter the function of the signal sequence of the maltose binding protein of Escherichia coli. Nature 285:78-81.

Beguin, P. (1983) Detection of cellulase activity in polyacrylamide gels using Congo red-stained agar replicas. Anal. Biochem. 131:333-336.

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

Beguin, P., Cornet, P., and Millet, J. (1983) Identification of the endoglucanase of the celB gene of Clostridium thermocellum. Biochimie 65:495-500.

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

Beguin, P., Eisen, H., and Roupas, A. (1977) Free and cellulose bound cellulases from Cellulomonas sp. J. Gen. Microbiol. 101:191-196.

Beguin, P., Gilkes, N.R., Kilburn, D.G., Miller, R.C. Jr., O'Neill, G.P., and Warren, R.A.J. (1987) Cloning of cellulase genes. CRC Critical Rev. Biotechnol. 6:129-162.

Birnboim, H.L., and Doly, J. (1979) A rapid alkaline extraction procedure for screening of recombinant plasmid DNA. Nucl. Acids Res. 7:1513-1523.

Blake, M.S., Johnston, K.H., Russel-Jones, G.J., and Gotschlich, E.C. (1984) A rapid and sensitive method for detection of alkaline phosphatase conjugated anti-antibody on Western blots. Anal. Biochem. 136:175-179.

Boel, E., Hjort, I., Svensson, B., Norris, F., Norris, K.E., and Fiil, N.P. (1984) Glucoamylases G1 and G2 from Aspergillus niger are synthesized from two different but closely related mRNAs. EMBO J. 3:1097-1102.

Bolivar, F., Rodreiquez, R.L., Bethlach, M.C., Heyneker, H.L. Boyer, H.W., Crosa, J.H., and Falkow, S. (1977) Construction and characterization of new cloning vehicles, II. A multipurpose cloning system. Gene 2:95-113.

Boyer, R.F., Allen, T.L., and Dykema, P.A. (1987) Fractionation of Aspergillus niger cellulases by combined ion exchange affinity chromatography. Biotechnol. Bioeng. 29:176-179.

Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72:248-254.

Breuil, C., and Saddler, J.N. (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.

Brownell, H.H., and Saddler, J.N. (1987) Steam pretreatment of lignocellulosic material for enzymatic hydrolysis. Biotechnol. Bioeng. 29:228-235.

Castanon, M., and Wilke, C.R. (1980) Adsorption and recovery of cellulases during hydrolysis of newspaper. Biotechnol. Bioeng. 22:1037-1053.

Chahal, D.S. (1985) Solid-state fermentation with Trichoderma reesei for cellulase production. Appl. Environ. Microbiol. 49:205-210.

Choi, W.Y., Haggett, K.D., and Dunn, N.W. (1978) Isolation of a cotton wool degrading strain of Cellulomonas: mutants with altered ability to degrade cotton wool. Aust. J. Biol. Sci. 31:553-564.

Choudhury, N., Gary, P.P., and Dunn, N.W. (1980) Saccharification of sugar cane bagasse by an enzyme preparation from Cellulomonas: resistance to product inhibition. Biotechnol. Lett. 2:427-428.

Collier, B., Thomas, D.Y., Teather, R.M., Erfle, J.D., and Crosby, W.L. (1984) Cloning and expression in Escherichia coli of cellulase genes from Bacteroides succinogenes. DNA 3:184-188.

Collmer, A., and Wilson, D.B. (1983) Cloning and expression of a Thermonospora YX endocellulase gene in E. coli. Bio/Technology 1:589-594.

Cornet, P., Millet, J., Beguin, P., and Aubert, J.-P. (1983a) Characterization of two cel (cellulose degradation) genes of Clostridium thermocellum coding for endoglucanases. Bio/Technology 1:589-594.

Cornet, P., Tronik, D., Millet, J., and Aubert, J.-P. (1983b) Cloning and expression in Escherichia coli of Clostridium thermocellum genes coding for amino acid synthesis and cellulose hydrolysis. FEMS Microbiol. Lett. 16:137-141.

Coughlan, M.P. (1985) The properties of fungal and bacterial cellulases with comments on their production and application. Biotech. Genetic Engineering Rev. 3:39-109.

de Boer, H. A., and Shepard, H. M. (1983) Strategies for optimizing foreign gene expression in Escherichia coli. Horizons Biochem. Biophys. 7:205-248.

Deshpande, M.V., Eriksson, K.E., and Pettersson, B. (1984a) An assay for selective determination of exo-β-1,4-glucanases in a mixture of cellulolytic enzymes. Anal. Biochem. 138:481-487.

Deshpande, V., Rao, M., Keskar, S., and Mishra, C. (1984b) Occurence of a procellulase in the culture filtrates of Penicillium janthinellum. Enzyme Microb. Technol. 6:371-374.

Emr, S.D., Hedgpeth, J., Clement, J.-M., Silhavy, T.J. and Hofnung, M. (1980) Sequence analysis of mutations that prevent export of λ receptor, an Escherichia coli outer membrane protein. Nature 285:82-85.

Enari, T.-M., and Niku-Paavola, M.-L. (1987) Enzymatic hydrolysis of cellulose: is the current theory of mechanisms of hydrolysis valid? CRC Critical Rev. Biotechnol. 5:67-87.

Enriques, A. (1981) Growth of cellulolytic bacteria on sugarcane bagasse. Biotechnol. Bioeng. 23:1423-1429

Eriksson, K.E. (1975) Enzyme mechanisms involved in the degradation of wood components, p. 263-280. In M. Bailey, T.M. Enari and M. Linko (eds.), Symposium on Enzymatic Hydrolysis of Cellulose. SITRA, Helsinki.

Eriksson, K.E. (1978) Enzyme mechanisms involved in cellulose hydrolysis by the rot fungus Sporotrichum pulverulentum. Biotechnol. Bioeng. 20:317-332.

Eriksson, K.E., and Hamp, S.G. (1978) Regulation of endo -β-D-glucanase production in Sporotrichum pulverulentum. Eur. J. Biochem. 90:183-190.

Eriksson, K.E., and Pettersson, B. (1982) Purification and partial characterization of two acidic proteases from the white-rot fungus Sporotrichum pulverlentum. Eur. J. Biochem. 124:635-642.

Fan, L.T., Lee, Y.-H., and Beardmore, D.H. (1980a) Major chemical and physical features of cellulosic materials as substrates for enzymatic hydrolysis. Adv. Biochem. Engineering 14:101-118.

Fan, L.T., Lee, Y.-H., and Beardmore, D.H. (1980b) Mechanism of enzymatic hydrolysis of cellulose: Effects of the major structural features of cellulose on enzymatic hydrolysis. Biotechnol. Bioeng. 22:177-199.

Fan, L.T., Lee, Y.-H., and Gharpuray, M.M. (1982) The nature of lignocellulosics and their pretreatments for enzymatic hydrolysis. Adv. Biochem. Engineering 23:157-183.

Fukumori, F., Kudo, T., and Horikoshi, K. (1987) Truncation analysis of an alkaline cellulase from an alkalophilic Bacillus species. FEMS Microbiol. Lett. 40:311-314.

Fukumori, F., Kudo, T., Narahashi, Y., and Horikoshi, K. (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., Sashihara, N., Kudo, T., and Horikoshi, K. (1986b) Nucleotide sequences of two cellulase genes from alkalophilic Bacillus sp. strain N-4 and their strong homology. J. Bacteriol. 186:494-504.

Gatz, C., and Hillen, W. (1986) Acinetobacter calcoaceticus encoded mutarotase: nucleotide sequence analysis of the gene and characterization of its secretion in Escherichia coli. Nucl. Acids Res. 14:4309-4323.

Gilkes, N.R., Kilburn, D.G., Langsford, M.L., Miller, R.C.Jr., Wakarchuk, W.W., Warren, R.A.J., Whittle, D.J., and Wong, W.K.R. (1984a) Isolation and characterization of Escherichia coli clones expressing cellulase genes from Cellulomonas fimi. J. Gen. Microbiol. 130: 1377-1384.

Gilkes, N.R., Kilburn, D.G., Miller, R.C.Jr., and Warren, R.A.J. (1984b) A mutant of Escherichia coli that leaks cellulase activity encoded by cloned cellulase genes from Cellulomonas fimi. Bio/Technology 2: 259-263.

Gilkes, N.R., Langsford, M.L., Kilburn, D.G., Miller, R.C.Jr., and Warren, R.A.J. (1984c) Mode of action and substrate specificities of cellulases from cloned bacterial genes. J. Biol. Chem. 259: 10455-10459.

Gold, L., Pribnow, D., Schneider, T., Shinedling, S., Swebelius-Singer, B., and Stormo, G. (1981) Translational initiation in prokaryotes. Ann. Rev. Microbiol. 35:365-403.

Gong, C.S., and Tsao, G.T. (1979) Cellulase and biosynthesis regulation. Ann. Reports on Fermentation Processes 3:111-139.

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

Greenberg, N.M., Warren, R.A.J., Kilburn, D.G., and Miller, R.C.Jr., (1987b) Regulation and initiation of cenB transcripts of Cellulomonas fimi. J. Bacteriol. 169:4674-4677.

Grepinet, O., and Beguin, P. (1986) Sequence of the cellulase gene of Clostridium thermocellum coding for endoglucanase B. Nucl. Acids Res. 14:1791-1799.

Griffin, H., Dintzis, F.R., Krull, L., and Baker, F.L. (1984) A microfibril generating factor from the enzyme complex of Trichoderma reesei. Biotechnol. Bioeng. 26:296-300.

Guo, Z., Arfman, N., Ong, E., Gilkes, N.R., Kilburn, D.G., Warren, R.A.J., and Miller, R.C.Jr. (1988) Leakage of Cellulomonas fimi cellulases from Escherichia coli. FEMS Microbiol. Lett., in press.

Guo, L.-H., and Wu R. (1983) Exonuclease III : Use for DNA sequence analysis and in specific deletions of nucleotides. Methods Enzymol. 100:60-96.

Hagget, K.D., Choi, W.Y., and Dunn, N.W. (1978) Mutants of Cellulomonas which produce increased levels of β-glucosidase. Eur. J. Appl. Microbiol. Biotechnol. 6:189-191.

Hagget, K.D., Gray, P.P., and Dunn, N.W. (1979) Crystalline cellulose degradation by a strain of Cellulomonas and its mutants derivatives. Eur. J. Appl. Microbiol. Biotechnol. 8: 183-190.

Halliwell, G., and Griffin, M. (1973) The nature and mode of action of the cellulolytic component C_1 of Trichoderma koningii on native cellulose. Biochem. J. 135:587-594.

Halliwell. G., and Griffin, M. (1978) Affinity chromatography of the cellulase system of Trichoderma koningii. Biochem. J. 169:713-715.

Hattori, M., and Sakaki, Y. (1986) Dideoxy sequencing method using denatured plasmid templates. Anal. Biochem. 152:232-238.

Hitchner, E.V., and Leatherwood, J.M. (1980) Use of a cellulase derepressed mutant of Cellulomonas in the production of a single-cell protein product from cellulose. Appl. Environ. Microbiol. 39:382-386.

Horton, G.L., Rivers, D.G., and Emert, G.H. (1980) Preparation of cellulosics for enzymatic conversion. Ind. Eng. Chem. Prod. Res. Dev. 19:422-429.

Inouye, M., and Halegoua, S. (1980) Secretion and membrane localization of proteins in Escherichia coli. CRC Critical Rev. Biochem. 7:339-371.

Inouye, S., Soberon, X., Franceschini, T., Nakamura, K., Itakura, K., and Inouye, M. (1982) Role of positive charge on the amino-terminal region of the signal peptide in protein secretion across the membrane. Proc. Natl. Acad. Sci. USA 79:3438-3441.

Jensen, K.F., Larsen, J.N., Schack, L., and Sivertsen, A. (1984) Studies on the structure and expression of Escherichia coli pyrC, pyrD, and pyrF using the cloned genes. Eur. J. Biochem. 140:343-352.

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

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

Juang, R.-H., Chang, Y.-D., Sung, H.-Y., and Su, J.-C. (1984) Oven-drying method for polyacrylamide gel slab packed in cellophane sandwich. Anal. Biochem. 141:348-350.

Keddie, R.M. (1984) Genus III. Cellulomonas, p. 629-631. In R.E. Buchanon and N.E. Gibbons (eds), Bergey's manual of determinative bacteriology, 8th ed. The Williams and Wilkins Co., Baltimore.

Klyosov, A., Mitkevich, O.V., and Sinitsyn, A.P. (1986) Role of the activity and adsorption of cellulases in the efficiency of the enzymatic hydrolysis of amorphous and crystalline cellulose. Biochem. 25:540-542.

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

Kohchi, C., and Toh-e, A. (1985) Nucleotide sequence of Candida pelliculosa β-glucosidase gene. Nucl. Acids Res. 13:6273-6282.

Koide, Y., Kakamura, A., Uozumi, T., and Beppu, T. (1986) Molecular cloning of a cellulase gene from Bacillus subtilis and its expression in Escherichia coli. Agric. Biol. Chem. 50:233-237.

Kotoujansky, A., Diolez, A., Boccara, M., Bertheau, Y., Andro, T., and Coleno, A. (1985) Molecular cloning of Erwinia chrysanthemi pectinase and cellulase structural genes. EMBO J. 4:781-785.

Kubicek, C.P. (1981) Release of carboxymethylcellulase and β-glucosidase from walls of Trichoderma reesei. Eur. J. Appl. Microbiol. Biotechnol. 13: 226-231.

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

Lampen, J.O., Wang, W., Mezes, P.S.F., and Yang, Y.-J. (1984) β-lactamases of Bacilli: Nature and processing, p. 129-140. In J. Hoch and A.T. Genesan (eds.), Genetics and Biotechnology of Bacilli. Academic Press, New York.

Langsford, M.L., Gilkes, N.R. Wakarchuk, W.W., Kilburn, D.G., Miller, R.C.Jr., and Warren, R.A.J. (1984) The cellulase system of Cellulomonas fimi. J. Gen. Microbiol. 130:13671376.

Langsford, M.L., Gilkes, N.R., Singh, B., Moser, B., Miller, R.C.Jr., Warren, R.A.J., and Kilburn, D.G. (1987) Glycosylation of bacterial cellulases prevents proteolytic cleavage between functional domains. FEBS Lett. 225:163-167.

Lever, M. (1973) Colorimetric and fluorometric carbohydrate determination with p-hydroxybenzoic acid hyrazide. Biochem. Med. 7:274-281.

Ljungdahl, L., Pettersson, B., Eriksson, K.E., and Wiegel, J. (1983) A yellow affinity substance involved in the cellulolytic system of Clostridium thermocellum. Current Microbiol. 9:195-200.

Lofdahl, S., Guss, B., Uhlen, M., Philipson, L., and Lindberg, M. (1983) Gene for staphylococcal protein A. Proc. Natl. Acad. Sci. USA 80:697-701.

Mandels, M. (1982) Cellulases, p. 35-78. In G.T. Tsao (ed.), Annual reports of fermentation processes, vol. 5. Academic Press Inc., New York.

Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) Molecular cloning: A laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Maxam, A.M., and Gilbert, W. (1980) Sequencing end-labelled DNA with base-specific chemical cleavages. Method Enzymol. 65:499-560.

McLaughlin, J.R., Murray, C.L., and Rabinowitz, J.C. (1981) Unique features in the ribosome binding site sequence of the Gram-positive Staphylococcus aureus β-lactamase gene. J. Biol. Chem. 256:11283-11291.

Messing, J. (1983) New M13 vectors for cloning. Method Enzymol. 101C:20-78.

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

Miller, G.L., Blum, R., Glennon, W.E., and Burton, A.L. (1960) Measurement of carboxymethylcellulase activity. Anal. Biochem. 2:127-132.

Miller, J.H. (1972) Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

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

Millett, M.A., Baker, A.J., and Satter, L.D. (1976) Physical and chemical pretreatments for enhancing cellulose saccharification. Biotechnol. Bioeng. Symp. 6:125-153.

Moloney, A.P., and Coughlan, M.P. (1983) Sorption of Talaromyces emersonii cellulase on cellulosic substrates. Biotechnol. Bioeng. 25:271-280.

Moloney, A.P., Hackett, T.J., Considine, P.J., and Coughlan, M.P. (1983) Isolation of mutants of Talaromyces emersonii CBS 814.70 with enhanced cellulase activity. Enzy. Microb. Technol. 5:260-264.

Moloney, A.P., McCrae, S.I., Wood, T.M., and Coughlan, M.P. (1985) Isolation and characterization of the 1,4-β-D-glucan glucanohydrolases of Talaromyces emersonii. Biochem. J. 225:365-374.

Montenecourt, B.S. (1983) Trichoderma reesei cellulases. Trends Biotechnol. 1:156-161.

Mullings, R. (1985) Measurement of saccharification by cellulases. Enzy. Microb. Technol. 7:586-591.

Murphy, N., McConnell, D.J., and Cantwell, B.A. (1984) The DNA sequence of the gene and genetic control sites for the excreted B. subtilis enzyme β-glucanase. Nucl. Acids Res. 12:5355-5367.

Nakamura, K., Misawa, N., and Kitamura, K. (1986) Cellulase genes of Cellulomonas uda CB4. I. Cloning and expression of β-glucosidase genes in Escherichia coli. J. Biotechnol. 3:239-246.

Nakayama, M., Tomita, Y., Suzuki, H., and Nisizawa, K. (1976) Partial proteolysis of some cellulase components from Trichoderma viride and the substrate specificity of modified products. J. Biochem. 79:955-966.

Nelson, N. (1952) A photometric adaptation of the Somogyi method for the determination of glucose. J. Biol. Chem. 153:376-380.

Norrander, J., Kempe, T., and Messing, J. (1983) Construction of improved M13 vectors using oligodeoxyribonucleotide-directed mutagenesis. Gene 26:101-106.

Nossal, N.G., and Heppel, L.A. (1966) The release of enzymes by osmotic shock from Escherichia coli in exponential phase. J. Biol. Chem. 241: 3055-3062.

Nummi, M, Niku-Paavola, M.-L., Enari, T.-M., and Raunio, V. (1981) Isolation of cellulases by means of biospecific sorption on amorphous cellulose. Anal. Biochem. 116:137-141.

O'Callaghan, C.H., Morris, A., Kirby, S.M., and Shingler, A.H. (1972) Novel method for detection of β -lactamase by using a chromogenic cephalosporin substrate. Antimicrob. Ag. Chemother. 1:283-288.

O'Neill, G., Goh, S.H., Warren, R.A.J. Kilburn, D.G., and Miller, R.C.Jr. (1986a) Structure of the gene encoding the exoglucanase of Cellulomonas fimi. Gene 44:325-330.

O'Neill, G., Warren, R.A.J., Kilburn, D.G., and Miller, R.C.Jr. (1986b) Secretion of a Cellulomonas fimi exoglucanase by Escherichia coli. Gene 44:331-336.

O'Neill, G., Warren, R.A.J., Kilburn, D.G., Miller, R.C.Jr. (1986c) Overproduction of a cellulase gene with a high guanosine plus cytosine content in Escherichia coli. Appl. Environ. Microbiol. 52:737-743.

Owolabi, J.B., Beguin, P., Kilburn, D.G., Miller, R.C.Jr., and Warren, R.A.J. (1988) Expression in Escherichia coli of the structural gene for endoglucanase B of Cellulomonas fimi. Appl. Environ. Microbiol., in press.

Peacock, S.L., McIver, C.M., and Monohan, J.J. (1981) Transformation of E. coli using homopolymer-linked plasmid chimeras. Biochim. Biophys. Acta 655:243-250.

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

Penttila, M.E., Nevalainen, K.M.H., Raynal, A., and Knowles, J.K.C. (1984) Cloning of Aspergillus niger genes in yeast. Expression of the gene coding Aspergillus β-glucosidase. Mol. Gen. Genet. 194:494-499.

Pettersson, L.G. (1975) The mechanism of enzymatic hydrolysis of cellulose by Trichoderma viride, p. 255-280. In M. Bailey, T.M. Enari, and M. Linko, (eds.), Symposium on Enzymatic Hydrolysis of Cellulose. SITRA, Helsinki.

Raynal, A., and Guerineau, M. (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.

Reese, E.T. (1982) Elution of cellulase from cellulose. Process Biochem. 17:2-6.

Reese, E.T., Siu, R.G.H., and Levinson, H.S. (1950) Biological degradation of soluble cellulose derivatives. J. Bacteriol. 59:485-487.

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

Rinderknecht, H., Geokas, M.C., Silverman, P., and Haverback, B.J. (1968) A new ultrasensitive method for the determination of proteolytic activity. Clin. Chim. Acta 21:197-203.

Rivers, D.B., Gracheck, S.J., Woodford, L.C., and Emert, G.H. (1984) Limitations of the DNS assay for reducing sugars from saccharified lignocellulosics. Biotechnol. Bioeng. 26:800-802.

Roberts, T.M., and Lauer, G.D. (1979) Maximizing gene expression on a plasmid using recombination in vitro. Method Enzymol. 68:473-482.

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

Robson, L.M., and Chambliss, G.H. (1987) Endo-β -1,4-glucanase gene of Bacillus subtilis DLG. J. Bacteriol. 169:2017-2025.

Romaniec, M.P.M., Clarke, N.G., and Hazlewood, G.P. (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, M., and Court, D. (1979) Regulatory sequences involved in the promotion and termination of RNA transcription. Ann. Rev. Genet. 13:319-353.

Russell, D.R., and Bennett, G.N. (1982) Construction and analysis of in vivo activity of E. coli promoter hybrids and promoter mutants that alter the -35 to -10 spacing. Gene 20:231-243.

Ryu, D.D.Y., Kim, C., and Mandels, M. (1984) Competitive adsorption of cellulase components and its significance in a synergistic mechanism. Biotechnol. Bioeng. 26:488-496.

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

Sashihara, N., Kudo, T., and Horikoshi, K. (1984) Cloning and expression of cellulase genes of alkalophilic Bacillus sp. strain N-4 in Escherichia coli. J. Bacteriol. 158: 503-506.

Schwarz, W., Bronnenmeier, K., and Staudenbauer, W.L. (1985) Molecular cloning of Clostridium thermocellum genes involved in β-glucan degradation in bacteriophage lambda. Biotechnol. Lett. 7:859-864.

Schwarz, W.H., Grabnitz, F., and Staudenbauer, W.L. (1986) Properties of a Clostridium thermocellum endoglucanase produced in Escherichia coli. Appl. Environ. Microbiol. 51:1293-1299.

Shoemaker, S.P., Gelfand, D.H., Innis, M.A. Kwok, S.Y., Ladner, M.B., and Schweickart, V. (1984) Recombinant fungal cellobiohyrolases. European patent application EP 137,280. European Patent Office.

Shoemaker, S.P., Schweilkart, V., Ladner, M., Gelfand, D., Kwok, S., Myambo, K., and Innis, M. (1983) Molecular cloning of the exocellobiohydrolase derived from Trichoderma reesei strain L27. Bio/Technology 1:691-696.

Silhavy, T.J., Benson, S.A., and Emr, S.D. (1983) Mechanism of protein localization. Microbiol. Rev. 47: 313-344.

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

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

Stuber, D., and Bujard, H. (1981) Organization of transcriptional signals in plasmids pBR322 and pACYC184. Proc. Natl. Acad. Sci. USA 78:167-171.

Sutcliff, J.G. (1979) Complete nucleotide sequence of the Escherichia coli plasmid pBR322. Cold Spring Harbor Symp. Quant. Biol. 43:77-90.

Suzuki, H. (1975) Cellulase formation in Pseudomonas fluorescens var. cellulosa, p. 115-169. In M. Bailey, T.M. Enari, and M. Linko (eds.), Symposium on Enzymatic Hydrolysis of Cellulose. SITRA, Helsinki.

Taylor, K.A., Crosby, B., McGavin, M., Forsberg, C.W., Thomas, D.Y. (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 Wood, P.J. (1982) Use of Congo-red-poly saccharide interactions in enumeration and characterization of cellulolytic bacteria from bovine rumen. Appl. Environ. Microbiol. 43:777-780.

Teeri, T. (1987) The cellulolytic enzyme system of Trichoderma reesei. Molecular cloning, characterization and expression of the cellobiohydrolase genes. Technical Research Centre of Finland. Publications 38. Espoo, Finland.

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

Teeri, T., Salovuori, I., and Knowles, J. (1983) The molecular cloning of the major cellulase gene from Trichoderma reesei. Bio/Technology 1:696-699.

Towbin, H., Staehlin, T., and Gordon, J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. USA 76:4350-4354.

van Gijsegem, F., Toussaint, A., and Schoonejans, E. (1985) In vivo cloning of the pectate lyase and cellulase genes of Erwinia chrysanthemi. EMBO J. 4:787-792.

van Tilbeurgh, H., Bhikhabhai, R., Petterson, L.G., and Claeyssens, M. (1984) Separation of endo- and exo-type cellulases using a new affinity method. FEBS Lett. 169:215-218.

van Tilbeurgh, H., Claeyssens, M., and de Bruyne, C.K. (1982) The use of 4-methylumbelliferyl and other chromophoric glycosides in the study of cellulolytic enzymes. FEBS Lett. 149: 152-156.

van Tilbeurgh, H., Tomme, P., Claeyssens, M., Bhikhabhai, R., and Petterson, G. (1986) Limited proteolysis of the cellobiohydrolase I from Trichoderma reesei. Separation of functional domains. FEBS Lett. 204:223-227.

Vasantha, N., Thompson, L.D., Rhodes, C., Banner, C., Nagle, J., and Filpula, D. (1984) Genes for alkaline protease and neutral protease from Bacillus amyloliquefaciens contain a large open reading frame between the regions coding for the signal sequence and mature protein. J. Bacteriol. 159:811-819.

Vlasuk, G.P., Inouye, S., Ito, H., Itakura, K., and Inouye, M. (1983) Effects of the complete removal of basic amino acid residues from the signal peptide on secretion of lipoprotein in Escherichia coli. J. Biol. Chem. 258:7141-7148.

Voller, A., Bidwell, D., and Barlett, A. (1976) Microplate immunoassays for the immunodiagnosis of virus infections, p. 506-512. In N. Rose, and H. Freedman (eds.), Manual of Clinical Immunology, 1st ed. American Society for Microbiology, Washington, D.C.

Wakarchuk, W.W., Greenberg, N.M., Kilburn, D.G., Miller, R.C.Jr., and Warren, R.A.J. (1988) Structure and transcription analysis of the gene encoding a cellobiase from Agrobacterium ATCC21400. J. Bacteriol. 170:301-307.

Wakarchuk, W.W., Kilburn, D.G., Miller, R.C. Jr., and Warren, R.A.J. (1984) The preliminary characterization of β-glucosidases of Cellulomonas fimi. J. Gen. Microbiol. 130:1385-1389.

Wakarchuk, W.W., Kilburn, D.G., Miller, R.C.Jr., and Warren, R.A.J. (1986) The molecular cloning and expression of a cellobiase gene from an Agrobacterium in Escherichia coli. Mol. Gen. Genet. 205:146-152.

Warren, R.A.J., Beck, C.F., Gilkes, N.R., Kilburn, D.G., Langdford, M.L., Miller, R.C.Jr., O'Neill, G.P., Scheufens, H., and Wong, W.K.R. (1986) Sequence conservation and region shuffling in an endoglucanase and an exoglucanase from Cellulomonas fimi. Proteins: Structure, Function and Genetics 1:335-341.

Warren, R.A.J., Gerhard, B., Gilkes, N.R., Owolabi, J.B., Kilburn, D.G., and Miller, R.C.Jr. (1988) A bifunctional exoglucanase-endoglucanase fusion protein. Gene, in press.

Watson, T.G., Nelligan, I., and Lessing, L. (1984) Cellulase production by Trichoderma reesei (RUT-C30) in fed-batch culture. Biotechnol. Lett. 6:667-672.

Whittle, D.J., Kilburn, D.G., Warren, R.A.J., and Miller, R.C.Jr., (1982) Molecular cloning of a Cellulomonas fimi cellulase gene in Escherichia coli. Gene 17:139-145.

Wolff, B.R., Mudry, T.A., Glick, B.R., and Pasternak, J.J. (1986) Isolation of endoglucanase genes from Pseudomonas fluorescens subsp. cellulosa and a Pseudomonas sp. Appl. Environ. Microbiol. 51:1367-1369.

Wong, W.K.R., Gerhard, B., Gou, Z.M., Kilburn, D.G., Warren, R.A.J., and Miller, R.C.Jr. (1986) Characterization and structure of an endoglucanase gene cenA of Cellulomonas fimi. Gene 44:315-324.

Wood, T.M. (1975) Properties and mode of action of cellulases. Biotechnol. Bioeng. Symp. 5:111-137.

Wood, T.M. (1985) Properties of cellulolytic enzyme systems. Biochem. Soc. Trans. 13:407-410.

Wood, T.M., and McCrae, S.I. (1979) Synergism between enzymes involved in the solubilization of native cellulose, p. 181-209. In R.D. Brown, and L. Jurasek, (eds.), Advances in Chemical Series vol. 181. American Chemical Society, Washington, D.C.

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

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

Yang, M., Galizzi, A., and Henner, D. (1983) Nucleotide sequence of the amylase gene from Bacillus subtilis. Nucl. Acids Res. 11:237-249.

Yanisch-Perron, C., Vieira, J., and Messing, J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of M13mp and pUC vectors. Gene 33:103-119.

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

APPENDIX

I Expression of *cenB* on a thermoinducible runaway replication plasmid

A) Targeted 5' deletions of the *cenB* gene

In order to further engineer the *cenB* gene for overexpression in *E. coli*, another set of targeted deletions of the gene was generated. The plasmid pJB3 was linearized with *Hind*III and digested bi-directionally with Exonuclease III, followed by removal of single-stranded overhanging regions with S_1 nuclease (Materials and Methods). The resulting products were cleaved with *Eco*RI, and ligated into *Sma*I and *Eco*RI cleaved pUC19. *E. coli* C600 cells transformed with the ligation mix were selected for Amp^R and screened for CMcellulase. Plasmids isolated from clones were analysed by restriction enzyme digestion and agarose gel electrophoresis. The fusion junctions of the *lacZ*'-*cenB*Δ plasmids were also sequenced. Representative 5' deletions of *cenB* obtained are as shown (Fig. A1).

B) Construction of pCP3*cenB* expression vectors

The plasmid pCP3 is a derivative of the runaway replication vector pKN402 which contains the lambda pL promoter adjacent to a multiple cloning site (see Materials and Methods). In order to provide the necessary translational initiation signals for the overexpression of *cenB*Δs, a derivative of pCP3 was constructed as shown (Fig. A2). pCP3 deleted for the *Eco*RI-*Bam*HI region in the multiple cloning site was ligated to an *Eco*RI - *Bam*HI portable translation initiation sequence (PTIS, purchased from Pharmacia, Canada). A single copy of the PTIS was observed on the resulting plasmid pCP3::PTIS, by restriction enzyme analysis and sequencing by the base-specific chemical degradation method (data not shown).

*cenB*Δ5 was subcloned into pCP3::PTIS as shown (Fig. A2). The 2.1kb fragment generated from *Bam*HI-*Sma*I digestion of pJB3Δ5 was gel-purified and ligated into the *Bam*HI and *Hinc*II sites of pUC18. *E. coli* cells transformed with the ligation mix were selected for Amp^R and screened for CMcellulase activity. The resulting plasmid, pUC18*cenB*Δ5, was digested with *Bam*HI and *Hind*III; the desired fragment was purified and ligated into *Bam*HI and *Hind*III sites of the pCP3::PTIS expression vector.

In the plasmid, designated pCP3cenBΔ5, the lambda leftward promoter initiates transcription across the cenBΔ5, which is fused in-frame to the initiating codon of the PTIS.

C) EngB synthesis in E. coli/pCP3cenBΔ5

E. coli C600/pcI857 cells carrying pCP3cenBΔ5 were grown in Luria broth supplemented with ampicillin (100μg/ml) and kanamycin (50μg/ml) at 30°C to an O.D. at 600 nm of 0.3. Subsequently, the culture was divided, and parallel samples were grown further at 30°C (noninduced) and at 41°C (induced) for 2 hours. The plasmid pcI857 codes for kanamycin resistance and for the thermolabile cI857 gene product (Materials and Methods). Although the cell extract of the C600/pcI857/pCP3cenBΔ5 induced culture had endoglucanase activity on CMcellulose Congo red indicator plate, the level of expression could not be accurately measured in the reducing sugar assay with the DNSA reagent because the absorbance was not significantly different from that of the reagent blank. Moreover, no distinct polypeptide corresponding to the pCP3cenBΔ5 encoded endoglucanase could be detected by SDS-PAGE analysis.

Figure A1. Sizes of deletions and effects on CMcellulase expression of the *cenB* gene. *E. coli* C600 Amp^R clones containing various deletion plasmids were screened for CMcellulase activity with Congo red. The plasmids were isolated and the sizes of deletions were determined by restriction enzyme analysis and agarose gel electrophoresis. Only the *C. fimi* DNA fragments are shown. a)parental fragment; b) Δ160; c) Δ75; d) Δ5; e) Δ78; f)Δ7. Transcription depended on the *lac* promoter/operator provided by pUC19. CMcellulase phenotypes : " - " , inactive; " +/- " , weakly active; " + " , active. Restriction sites: B, *Bam*HI; P, *Pst*I.

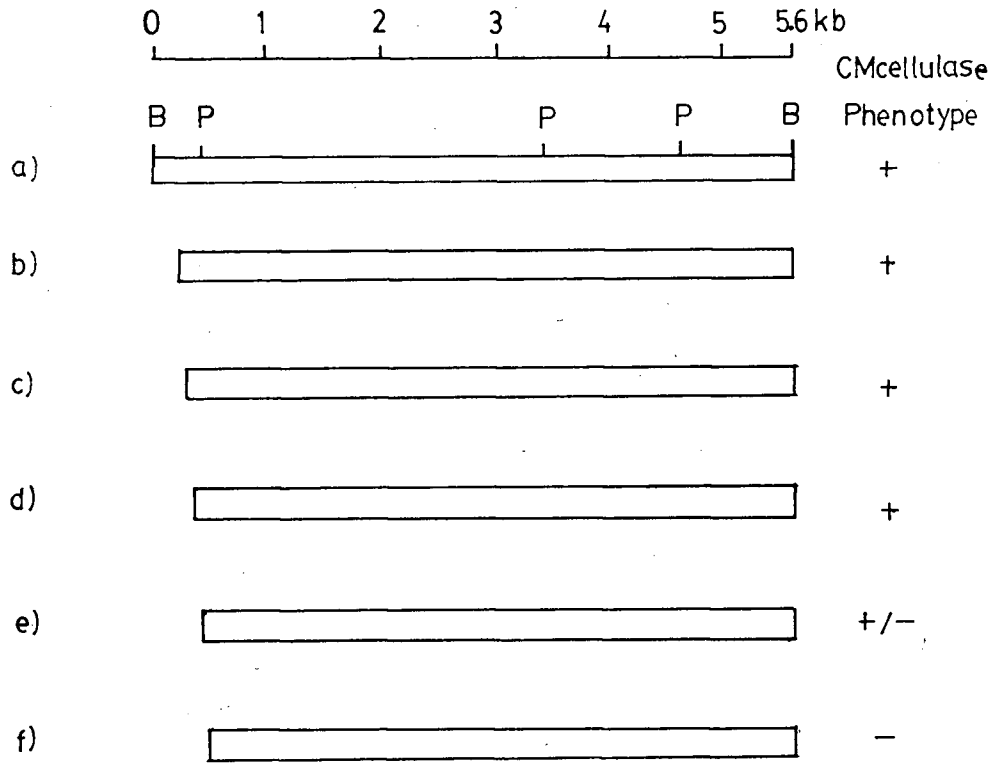


Figure A2. Scheme for the construction of pCP3cenB expression vector. See the text for details. The DNA sequence of the cenBΔ5 coding region immediately 3' to the lambda leftward promoter in pCP3 is shown. The functional orientations of the gene coding for β-lactamase (Amp ^{R}), EngB, and the lac and P_L promoters are indicated by arrows. PTIS; coding sequence of the EngB. Restriction sites: B, BamHI; E, EcoRI; Hc, HincII; H3, HindIII; Sm, SmaI. The deduced amino acid sequence for translated codons is indicated in one-letter code.