THE CONSTRUCTION AND CHARACTERIZATION
OF A PRO-THR BOX DELETION OF A CELLULOMONAS FIMI
ENDOGLUCANASE (Cen A)

By

Shen Hua
B.Sc., Nan Kai University, Tianjin, China, 1982

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES
(Department of Microbiology)

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA
February 1990
© Shen Hua, 1990
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
Vancouver, Canada

Date 17 April, 1990
ABSTRACT

The catalytic domain is separated from the cellulose-binding domain in *Cellulomonas fimi* endoglucanase CenA by a proline-threonine rich sequence called the Pro-Thr box. To study the function of the Pro-Thr box region, a deletion mutant, *cenAAPT*, was made from *cenA* by an oligonucleotide directed *in vitro* mutagenesis. The truncated enzyme, CenAAPT, was purified to homogeneity by affinity chromatography on cellulose and characterized. Comparing CenAAPT to CenA, the following characteristics were observed: 1) the Pro-Thr box affected the migration of CenA on SDS-PAGE; 2) the deletion of the Pro-Thr box altered the high affinity interaction with cellulose; 3) the truncated enzyme showed 40-50% reduction in catalytic activity towards both microcrystalline and amorphous cellulose; 4) the truncated enzyme was as sensitive as CenA to a *C. fimi* protease, and both enzymes were cleaved at the same site adjacent to the binding domain. The Pro-Thr box is not essential for the catalytic activity of CenA or its binding to cellulose, but it does contribute to both functions.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS</td>
<td>viii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>ix</td>
</tr>
<tr>
<td><strong>INTRODUCTION</strong></td>
<td>1</td>
</tr>
<tr>
<td>I  Cellulose and cellulases</td>
<td>1</td>
</tr>
<tr>
<td>II CenA and Cex of <em>C. fimi</em></td>
<td>2</td>
</tr>
<tr>
<td>III Structural features of CenA and Cex</td>
<td>2</td>
</tr>
<tr>
<td>IV Objectives</td>
<td>3</td>
</tr>
<tr>
<td><strong>MATERIALS AND METHODS</strong></td>
<td>6</td>
</tr>
<tr>
<td>I  Bacterial strains, phage and plasmid</td>
<td>6</td>
</tr>
<tr>
<td>II Media</td>
<td>6</td>
</tr>
<tr>
<td>III Buffers</td>
<td>7</td>
</tr>
<tr>
<td>IV Enzymes, reagents and DNA recombination techniques</td>
<td>7</td>
</tr>
<tr>
<td>V  Gel electrophoresis of DNA</td>
<td>7</td>
</tr>
<tr>
<td>VI Deletion of the Pro-Thr box-encoding sequence from <em>cenA</em></td>
<td>7</td>
</tr>
<tr>
<td>(A) Construction of an M13mp19<em>cenA</em> clone</td>
<td>8</td>
</tr>
<tr>
<td>(B) Preparation of uracil-containing M13mp19<em>cenA</em> single-stranded DNA</td>
<td>8</td>
</tr>
<tr>
<td>(C) Synthetic oligodeoxyribonucleotides</td>
<td>8</td>
</tr>
<tr>
<td>(D) Primer annealing reaction and T7 DNA polymerase extension</td>
<td>9</td>
</tr>
</tbody>
</table>
TABLE OF CONTENTS (continued)

(E) Transfection and purification of phage 9
(F) DNA sequence determination 9
VII The construction of pUC18-1.5cenAΔPT 9
VIII Screening colonies for endoglucanase activity 10
IX Purification of CenAΔPT 10
   (A) Growth of cells 10
   (B) Preparation of cell extract 10
   (C) Cellulose affinity chromatography 10
   (D) Electrophoresis of proteins 11
X Binding affinity for Avicel 11
XI Sensitivity to C. fimi protease 12
   (A) Preparation of C. fimi protease 12
   (B) Hydrolysis of CenAΔPT with C. fimi protease 12
XII Assay of enzyme activity 12
   (A) Carboxymethylcellulase assay 12
   (B) Activity on Avicel and phosphoric acid swollen Avicel 13
   (C) Cellulose azure assay 13
   (D) C. fimi protease assay 13
XIII Determination of protein concentration 13

RESULTS 14
I Construction of cenAΔPT 14
   (A) Subcloning of cenA gene into M13mp19 14
   (B) in vitro mutagenesis 14
| TABLE OF CONTENTS (continued) |
|-----------------------------|---|
| (C) Subcloning of cen\text{A}\Delta PT into pUC18 | 15 |
| II Purification of Cen\text{A}\Delta PT by affinity chromatography on cellulose | 15 |
| III Characterization of Cen\text{A}\Delta PT | 15 |
| (A) Sensitivity to \textit{C. fimi} protease | 15 |
| (B) Binding to Avicel | 16 |
| (C) Activity | 16 |
| DISCUSSION | 36 |
| REFERENCES | 40 |
LIST OF TABLES

Table 1: Bacterial strains, phage and plasmid  
Table II: Purification of CenAΔPT  
Table III N-terminal amino acid sequence of p30 produced from ngCenAΔPT by *C. fimi* protease  
Table IV: Extinction coefficient of ngCenAΔPT  
Table V: Enzyme activity against a range of cellulosic substrates  
Table VI: *M*ₚ determination for ngCenA and ngCenAΔPT
LIST OF FIGURES

Fig. 1: Structure of CenA and Cex of *C. fimi* 4

Fig. 2: Scheme for the construction of CenA Pro-Thr box deletion plasmid pUC18-\( cenA\Delta PT \) 22

Fig. 3: Screening of M13mp19-\( cenA \) clones by *BamHI* digestion of phage DNA 25

Fig. 4: Screening of potential deletion mutants (M13mp19-\( cenA\Delta PT \)) by *HpaI* digestion of RF DNA 26

Fig. 5: Dideoxy nucleotide sequencing gel showing deletion of the Pro-Thr box coding sequence from the \( cenA \) gene 27

Fig. 6: Screening of potential deletion mutants (pUC18-\( cenA\Delta PT \)) for CMCase activity 28

Fig. 7: Deletion of ngCenA\( \Delta PT \) in cell extracts 29

Fig. 8: Restriction analysis of pUC18-\( cenA\Delta PT \) clones 30

Fig. 9: Purification of ngCenA\( \Delta PT \) by affinity chromatography on cellulose 31

Fig. 10: SDS-PAGE analysis of the purification of ngCenA\( \Delta PT \) 32

Fig. 11: Time course of proteolysis of ngCenA and ngCenA\( \Delta PT \) 33

Fig. 12: Adsorption of ngCenA\( \Delta PT \) and ngCenA to Avicel 34

Fig. 13: Sequence alignments of CenA and CenA\( \Delta PT \) Pro-Thr box and flanking regions and locations of *C. fimi* protease primary cleavage sites in the nonglucosylated protein 35
### ABBREVIATIONS AND SYMBOLS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ap</td>
<td>ampicillin gene</td>
</tr>
<tr>
<td>BSA</td>
<td>bovine serum albumin</td>
</tr>
<tr>
<td>CBD</td>
<td>cellulose binding domain</td>
</tr>
<tr>
<td>CMC(ase)</td>
<td>carboxymethyl cellulose(ase)</td>
</tr>
<tr>
<td>dNTP</td>
<td>deoxynucleoside triphosphate</td>
</tr>
<tr>
<td>Λ</td>
<td>deletion in DNA or protein</td>
</tr>
<tr>
<td>$E_{205}/E_{280}$</td>
<td>extinction coefficient at 205nm or 280nm</td>
</tr>
<tr>
<td>HBAH</td>
<td>hydroxybenzonic acid hydrazine</td>
</tr>
<tr>
<td>GuHCl</td>
<td>guanidine-HCl</td>
</tr>
<tr>
<td>IPTG</td>
<td>isopropyl-β-D-thiogalactoside</td>
</tr>
<tr>
<td>kb</td>
<td>1000 base pairs</td>
</tr>
<tr>
<td>kDa</td>
<td>kilodalton</td>
</tr>
<tr>
<td>lacP/O</td>
<td><em>E.coli</em> β-galactosidase gene promotor and inhibitor</td>
</tr>
<tr>
<td>Mr</td>
<td>relative molecular mass</td>
</tr>
<tr>
<td>PAGE</td>
<td>polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PMSF</td>
<td>phenylmethylsulfonyl fluoride</td>
</tr>
<tr>
<td>RF DNA</td>
<td>replication form DNA</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulfate</td>
</tr>
<tr>
<td>UV</td>
<td>ultra violet</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada. I would like to thank Drs. R.A.J. Warren, D.G. Kilburn, and R. C. Miller Jr. for their expert guidance and for opening up my eyes to the challenges of scientific research. I would like to thank Dr. N.R. Gilkes for helpful discussions. Also, my thanks to all the members of the cellulase group especially my laboratory fellows in Room 206, H. Smith, E. Kwan, J. Coutinho, E. Ong, and A. Macleod for their friendship and kindness which have made my stay in the laboratory an enjoyable and pleasant one.

Special thanks are due to E. Ong for taking the photographs in this thesis, to all my friends in China and elsewhere, and to my family, for their love and encouragement throughout the years.

I dedicate this thesis to my parents.
INTRODUCTION

I. Cellulose and cellulases

Cellulolytic microorganisms produce a complex of enzymes capable of hydrolysing β-1,4-glucosidic bonds, and degrading cellulose to cellobiose or glucose. Many cellulolytic bacteria and fungi have been identified and characterized (Beguin, et al., 1987; Coughlan, 1985). Potential applications of cellulase for the production of fuels from biomass and degradation of agricultural products have intensified research in the enzymatic degradation of cellulose.

Cellulose is the major component of the cell wall of plants. It is a linear polymer of up to 14,000 glucose residues joined by β-1,4-glucosidic bonds. The linear chains of cellulose are organized in bundles of parallel chains to form fibrils. These fibrils comprise both ordered structure, the crystalline regions, and less ordered structure, the amorphous regions.

Cellulase systems are composed of a variety of enzymes with different specificities and modes of action, which act together to hydrolyze cellulose. Cellulose degradation requires the activities of three basic types of enzyme (Coughlan, 1985; Coughlan and Ljungdahl, 1988; Wood, 1985). Endoglucanases attack the amorphous region of cellulose fibers, cleaving internal β-1,4-glucosidic bonds at random and generating new non-reducing ends. Exoglucanases attack the cellulose molecules preferentially from the non-reducing end, liberating cellobiose subunits and then proceeding into the crystalline regions of the fibers. β-glucosidases hydrolyze cellobiose and low molecular weight cellodextrins into glucose.

The rapid development in the study of the molecular biology of cellulase genes has opened new areas of interest: 1) the organization of cellulase genes, 2) their regulation at the molecular level, and 3) the study of structural features required for enzyme activity. This thesis focuses on the influence of the Pro-Thr box on the function of *C. fimi* CenA.
II CenA and Cex of *C. fimi*

The gram-positive cellulolytic bacterium *Cellulomonas fimi* produces a complex array of cellulases when grown on cellulosic material. Of these, CenA and Cex are the two major components in the culture medium; they bind tightly to Avicel. The genes encoding CenA and Cex have been cloned in *E. coli* and their nucleotide sequences determined (Gilkes *et al.*, 1984a; O’Neill *et al.*, 1986; Whittle *et al.*, 1982; Wong *et al.*, 1986). CenA hydrolyzes Cm-cellulose at random while Cex shows preference for the termini (Gilkes *et al.*, 1984b). The native CenA and Cex from *C. fimi* are glycoproteins with M_r of 49.3 and 53 kDa, respectively. The recombinant CenA and Cex from *E. coli* are non-glycosylated, with M_r of 48.7 and 47.3 kDa, respectively (Langsford *et al.*, 1987). The glycosylated CenA and Cex are protected from attack by a *C. fimi* protease when bound to cellulose, but the non-glycosylated recombinant CenA and Cex are both cleaved precisely at the C-terminus of the Pro-Thr box, resulting in the formation of a 30 kDa fragment from CenA and a 39 kDa fragment from Cex. Both of these fragments retain catalytic activity but no longer bind to cellulose (Langsford *et al.*, 1987; Gilkes *et al.*, 1988). Therefore, both CenA and Cex have independently functioning catalytic and cellulose-binding domains.

III Structural features of CenA and Cex

CenA consists of a polypeptide of 449 amino acids including a leader peptide of 31 amino acids which functions to export CenA to the periplasm of *E. coli* (Wong *et al.*, 1986). Cex consists of a polypeptide of 484 amino acids including a 41 amino acid leader peptide which functions to export Cex to the periplasm of *E. coli* (O’Neill *et al.*, 1986). Each enzyme contains three distinct regions (Gilkes *et al.*, 1988; Warren *et al.*, 1986) (Fig. 1a). The binding domain which is a region of low charge, rich in hydroxyamino acids with 112 amino acid residues at the N-terminus of CenA, and 108 amino acid residues at the C-terminus of Cex. These regions are highly conserved, with about 50% sequence homology (Fig. 1b). The more highly charged catalytic domains,
which are 283 amino acid residues at the C-terminus of CenA, and 315 amino acid residues at the N-terminus of Cex, both contain sequences which resemble the sequence at the active site of hen egg white lysozyme (Warren et al., 1986). The third region consists of a highly conserved short sequence of about 20 residues containing only proline and threonine, termed the Pro-Thr box \((PT)_4T(PT)_7\) for CenA and \((PT)_3T(PT)_3T(PT)_3\) for Cex\) (Fig. 1c), which separates the binding domain and the catalytic domain.

A similar domain structure occurs in many fungal and bacterial cellulases and \(\beta\)-glucosidases. In most cases, proline and hydroxyamino acid rich segments separate the functional domains (review by Beguin, in press). The most intensively studied cellulases are \(T.\ reesei\) CBH I and CBH II. CBH I and CBH II contain a conserved element of approximately 30 residues, at the C-terminus of CBH I and at the N-terminus of CBH II, which appears to be a binding domain. The core enzymes, obtained by mild papain treatment, lack the conserved region, retain activity towards soluble substrates but have much reduced Avicelase activity. A Pro-Thr-Ser rich sequence connects the binding domain and the core enzymes (Tomme et al., 1988; van Tilbeurgh et al., 1986). Intact and core CBH I and CBH II were studied by small angle X-ray scattering (Schmuck et al., 1986; Abuja et al., 1988a,b). Both intact enzymes are tadpole-shaped, with an isotropic head and a long tail. The core enzymes lack tails.

IV Objectives

The striking similarities in the organization of CenA and Cex and the extensive sequence homology between their binding domains make it very interesting to understand the relationship between their structural features and their differential substrate specificities. The role(s) of the Pro-Thr box, almost perfectly conserved and sited between functional domains in both enzymes, is not yet clear. Determining the role of the Pro-Thr box in CenA should contribute to an understanding of the mechanism of action of CenA. It should also contribute to the understanding of other cellulase systems because of the common occurrence of the presence of proline-and hydroxyamino acid-rich
Fig. 1 Structure of CenA and Cex of *C. fim*  (Adapted from Warren *et al.*, 1986.)

a. Overall structures. PT denotes Pro-Thr box; AS denotes putative active site.

b. Conservation of amino acid sequence in the cellulose-binding domains. The numbers refer to amino acid sequences of the intact proteins, with 1 being the N-terminal amino acid. Conserved residues are boxed. * denotes a gap left in the sequence to improve the alignment. + denotes pairs of nonconserved residues which belong to the same structurally related exchange group. clusters of hydroxyamino acids are underlined.

c. Sequence conservation in the Pro-Thr boxes. * denotes a gap left in the sequence to improve the alignment.
segments connecting domains in cellulases from various organisms. This work, concerns the function of the Pro-Thr box in CenA. The strategy was to delete precisely from the cenA gene the sequence encoding the Pro-Thr box, using oligonucleotide directed in vitro mutagenesis. The polypeptide produced by the deletion mutant was purified and its affinity for cellulose, its catalytic activity and its sensitivity to C. fimi protease were examined.
MATERIALS AND METHODS

I  Bacterial strains, phage and plasmid

A list of the bacterial strains, phage and plasmid used in these studies is given in Table I. Stock cultures of bacteria were maintained at -20°C in 40% glycerol LB medium or -80°C in 10% DMSO LB medium.

Table I. Bacterial strains, phage and plasmid

<table>
<thead>
<tr>
<th>Bacterial strain/ phage/plasmid</th>
<th>Genetic characters</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>E. coli JM101</td>
<td>supE thi (lac-proAB) [F traD36 proAB lac^r ZΔ M15]</td>
<td>Yanisch-Perron et al., 1985</td>
</tr>
<tr>
<td>E. coli RZ1032</td>
<td>HfrKL16 PO/45 [lysA(61-62)], dut1, ung1, thi1, relA1 Zbd-279::Tn10, supE44</td>
<td>Kunkel et al., 1987</td>
</tr>
<tr>
<td>M13mp19</td>
<td>lac</td>
<td>Norrander et al., 1983</td>
</tr>
<tr>
<td>pUC18</td>
<td>lac Ap^R</td>
<td>Norrander et al., 1983</td>
</tr>
<tr>
<td>pUC18-1.6cenA</td>
<td>lac Ap^R</td>
<td>Guo et al., 1988</td>
</tr>
</tbody>
</table>

II  Media

LB medium contained per liter; 10 g Bacto-tryptone, 5 g Bacto-yeast extract and 5 g NaCl. 2xYT medium contained per liter; 16 g Bacto-tryptone, 10 g Bacto-yeast extract.
and 5 g NaCl. The pH of all the media was adjusted to 7.2 with NaOH. M9 medium contained per liter; 6 g Na2HPO4, 3 g KH2PO4, 0.5 g NaCl, 1 g NH4Cl, 5 g Casamino acids, 1 mg thiamine with the addition, after autoclaving, of 10 ml of 0.01M CaCl2, 2 ml of 1M MgSO4 and 10 ml of 20% glucose. Agar plates contained 1.5% agar.

III Buffers

The composition and preparation of buffers were described previously (Maniatis et al., 1982; Messing, 1983).

IV Enzymes, reagents and DNA recombination techniques

In general, DNA preparation, restriction enzyme reactions and recombinant DNA techniques were as described by Maniatis et al. (1982). Restriction endonucleases, T7 DNA polymerase, T4 DNA ligase, and deoxyribonucleotides were purchased from Pharmacia Inc. (Dorval, Quebec), New England Biolabs Inc. (Beverly, MA, USA), or BRL Inc. (Burlington, Ontario). All other chemicals and reagents were from BDH Chemicals (Vancouver, Canada) or Sigma Chemical Company (St. Louis, Mo.).

V Gel electrophoresis of DNA

Agarose gels (0.7 %) and polyacrylamide gels (6-12%) were used for the analysis and preparation of large (0.5-12 kb) and small (35-450 bp) DNA fragments, respectively. For preparative purposes, large DNA fragments were separated on 0.7% agarose gels running in TAE buffer. The bands of DNA were located by staining with ethidium bromide, then excised. The DNA was recovered with the Geneclean™ kit (BIO/CAN Scientific Inc. Mississauga, Ontario) as follows: gel slices were dissolved in NaI solution by incubating at 50°C for 5 min; the DNA was bound to glassmilk washed with NEW solution, and eluted with TE buffer.

VI Deletion of the Pro-Thr box-encoding sequence from cenA

(A) Construction of an M13mp19cenA clone

The 1588 bp SstI fragment from pUC18-1.6cenA was cloned into the SstI site of
M13mp19 to give M13mp19cenA.

(B) Preparation of uracil-containing M13mp19cenA single-stranded DNA

M13mp19 phage containing uracil in the DNA was prepared as described by Kunkel (1985) with modifications. E. coli RZ1032 (Kunkel et al., 1987) cells were grown at 37°C in TY medium with vigorous shaking to a density of 10^8 cells per ml; M13mp19-cenA phage were then added at a multiplicity of infection of 20. Incubation at 37°C was continued for another 5 hours, then the culture was centrifuged at 5,000 x g for 5 min at 4°C. The clear supernatant containing the phage was used for a second cycle of growth, identical to the first, to obtain a phage titer on the ung host (RZ1032) 6,000 fold greater than that on the ung^+ host (JM101). The phage suspension was made 3% in polyethylene glycol (PEG 8000) and 0.5 M in NaCl. After centrifugation at 5000 x g for 15 min, the phage pellet was resuspended in TE (pH 7.6), then extracted once with phenol and once with chloroform/isoamyl alcohol (24:1). The DNA was precipitated with ethanol and resuspended in TE. The DNA concentration was determined by absorbance at 260nm. The purity of the DNA was analysed by agarose gel electrophoresis.

(C) Synthetic oligodeoxyribonucleotides

A mutagenesis primer 5'-pGC TGG TCG GCT GCG GCG TGA CGT TAA CGC TGG TCG TCG GCA CGG TGC C-3' (48-mer), and a sequencing primer 5'-CGC GGT AGC CCT GCG TCG-3' (17-mer) were synthesized chemically on an Applied Biosystems 380A DNA synthesizer by T. Atkinson using phosphite triester chemistry, essentially as described (Adams et al., 1983; Atkinson and Smith, 1984). The 48-mer was separated from incomplete products by electrophoresis in a 12% polyacrylamide-7M urea sequencing gel, located by UV-shadowing, and extracted from the gel by the crush and soak method (Atkinson and Smith, 1984). The crude17-mer oligonucleotide was dissolved in 1.5 ml of 0.5M ammonium acetate solution. The primers were further purified by binding to Sep-Pak C-18 reverse phase cartridges (Millipore/Waters Assoc., Milford, MA) in water, followed by washing with water, and elution with 20% acetonitrile-80% water. The
concentrations of the oligodeoxynucleotides were determined by absorbance at 260 nm.

(D) Primer annealing reaction and T7 DNA polymerase extension

One pmole of M13mp19-cenA DNA was mixed with one pmole of mutagenesis primer in 18 µl of buffer containing 40mM Tris-Cl (pH 7.5), 20mM MgCl2 and 50mM NaCl. The DNA was denatured by heating the solution in a 1.5 ml microfuge tube at 55°C for 10 min and allowed to anneal by cooling to 23°C over a period of 2 hours and then kept at 4°C overnight. The reaction mixture was then supplemented to give final concentrations of 500µM for each dNTP, 1mM ATP, 3mM DTT, 4 units T7 DNA polymerase, 2 units T4 ligase and 2.4 µl of 5x annealing buffer. (final reaction volume of 31µl). The reaction mixture was incubated at 37°C for 2 hours, then 15mM EDTA pH 8.0 was added to terminate the reaction.

(E) Transfection and purification of phage

The preparation of E. coli JM101 (ung+) competent cells and their transfection with the primer-extended double stranded M13mp19-cenA DNA were performed as described by Messing (1983). The plaques were screened for the deletion mutation by screening RF DNAs for sensitivity to HpaI. The positive plaques were purified by titering the phage on JM101 cells.

(F) DNA sequence determination

The cenAΔPT deletion mutant was sequenced by the dideoxy chain termination method (Sanger et al., 1977). Single stranded templates were prepared from 1.5 ml of phage infected cultures (Messing 1983). Samples from the enzymatic sequencing reactions were analyzed on 6% denaturing (8M urea) polyacrylamide gels.

VII The construction of pUC18-1.5cenAΔPT

The 1525 bp SstI fragment from M13mp19cenAΔPT RF DNA was ligated to pUC18 which had been cut with SstI. E. coli JM101 was transformed with the products.
Transformants were screened for endoglucanase activity (see below) and a positive clone was designated pUC18-1.5cenAΔPT.

VIII Screening colonies for endoglucanase activity

*E. coli* JM101 clones carrying recombinant plasmids were picked onto both a master LB-agar plate containing 100μg ampicillin /ml and a LB-agar plate containing 100μg ampicillin /ml, 0.2 mM IPTG and 1% high-viscosity CMC (carboxymethylcellulose). After incubation at 37°C for 12 hours, the plates were exposed to chloroform vapour for 20 min and incubated at 37°C for 30 min. Then the plates were stained with 0.2% Congo-red for 30 min, and destained with 1M NaCl. Positive colonies were surrounded by a clear halo against a red background.

IX Purification of CenAΔPT

(A) Growth of cells

A 200 ml overnight culture of JM101 pUC18-1.5cenAΔPT was added to 60L LB medium, containing 100μg ampicillin /ml and 0.1mM IPTG, in a 100L fermenter. The culture was grown for 16 h at 37°C; it was stirred at 200 rpm and aerated with 20 L air / min. The cells were harvested with a Sharples centrifuge at 45,000rpm. The cell paste was washed with 50mM phosphate buffer pH 7, 0.02% azide. The cells were resuspended in 1 L 50mM phosphate buffer, 3mM EDTA, 0.02% azide.

(B) Preparation of cell extract

Cells were ruptured by passage twice through a French Press. Immediately after rupture, the extract was made 1mM in PMSF and 0.01mM in pepstatinA, and 1.5% in streptomycin sulfate. The extract was clarified by centrifugation twice at 17,000 rpm in a Beckman JA20 rotor.

(C) Cellulose affinity chromatography

The clarified cell extract was passed through a jacketed CF1 cellulose column (5cm x
30cm) maintained at 4°C at a flow rate of 1ml/min. The column was washed with 1M NaCl, and then 20mM Tris-Cl, pH 7.4 buffer. The enzyme was eluted with a linear gradient of 0-8M GuHCl in 20 mM Tris-Cl, pH 7.4. The absorbance of the eluate was measured at 280nm. Fractions of 10 ml were collected and assayed for CMCase activity. The peak fractions were pooled and the GuHCl was replaced with 50mM phosphate buffer pH 7 by ultrafiltration with an Amicon PM10 membrane in an Amicon pressure cell.

(D) Electrophoresis of proteins

SDS-polyacrylamide gels, 0.75 mm thick, were run according to Laemmli (1970). Relative mass (Mr) was estimated by comparison with Sigma SDS-6H Mr standards run on the same gel. Gels were stained with Coomassie Blue (Zacharius et al., 1969), and preserved by drying between layers of cellophane.

For sequencing, proteins in gels were transferred electrophoretically to polyvinylidene difluoride membrane, and stained with Coomassie Blue. Protein sequences were determined by automated Edman degradation in an Applied Biosystems 470A gas-phase sequenator.

Enzyme activity was detected in gels by a zymogram technique (Beguin, 1983). After electrophoresis, the acrylamide gel was washed twice with 25% isopropanol-50mM citrate pH 6.8 for 30 min, and twice with 50mM citrate buffer pH 6.8, then laid on a 1% agarose gel in citrate buffer and containing 1% CMC (high viscosity), and incubated at 37°C over night. Then the overlays were stained with Congo Red to observe zones of clearing.

X Binding affinity for Avicel

Adsorption was determined at 5°C in 1.5 ml Eppendorf tubes. Tubes contained 20 mg Avicel in 1.15 ml phosphate buffer pH 7.0, and 0.05 - 1.6 mg protein ml⁻¹. Tube contents were mixed continuously by rotation. After equilibration for 3 h, the adsorbent and bound protein were removed by centrifugation (twice at 10,000 xg, 10 min) and the unbound protein concentration ([P], mg protein.ml⁻¹) estimated from the absorption of the
supernatant at 280 nm using an appropriate blank. The bound protein concentration \([P]_{\text{ad}}, \text{mg protein.mL}^{-1}\) was estimated from the difference between the initial protein concentration \([P]_0, \text{mg protein.mL}^{-1}\) and \([P]\). Raw data were graphed as an adsorption isotherm plot and as a Scatchard plot.

XI Sensitivity to \(C.\ fimi\) protease

(A) Preparation of \(C.\ fimi\) protease

Protease was prepared from the culture supernatant after growth of \(C.\ fimi\) on 0.1% glycerol medium (Gilkes, et al., 1988).

(B) Hydrolysis of CenAAPT with \(C.\ fimi\) protease

A 35 \(\mu\)g sample of CenAAPT was incubated with 0.2 unit of \(C.\ fimi\) protease in 200 \(\mu\)l phosphate buffer pH 7 at 37°C. Samples of 20 \(\mu\)l were removed at intervals and the reaction terminated by the addition of PMSF to a final concentration of 3.6 mg/ml. The samples were analyzed by SDS-PAGE. Protein bands were blotted on polyvinylidene difluoride membranes for sequence determinations.

XII Assay of enzyme activity

(A) Carboxymethylcellulase assay

CMCase activity was determined by two methods. In the purification process, cell extract was incubated with 25 mg substrate/ml in 0.75 ml of phosphate buffer at 37°C for 30 min, then 50 \(\mu\)l of glucose (1mg/ml) and 0.8 ml of DNS reagent (Miller, 1959) were added, and the mixture steamed at 100°C for 15 min, cooled and the absorbance at 550 nm measured. A standard curve was prepared using a 5.56 mM stock solution of glucose.

The purified enzyme was assayed by the hydroxybenzoic acid hydrazide (HBAH) method (Langsford et al., 1987). Enzyme was incubated with 4 mg substrate/ml in 0.5 ml citrate buffer at 30°C for 30 min. Then 1 ml of hydroxybenzoic acid hydrazide reagent was added, the reaction mixture was steamed at 100°C for 12 min, allowed to cool, and the absorbance was measured at 420 nm. A standard curve was determined for each assay.
using a 55.6 nMol stock solution of glucose. Results are expressed as glucose equivalents.

(B) Activity on Avicel and phosphoric acid swollen Avicel

Enzyme was incubated with 15 mg of Avicel or 10 mg of phosphoric acid swollen Avicel in 1.5 ml of 50mM sodium citrate, pH 7.0, 0.2% BSA, 0.02% Na₃N at 37°C for 18 h or 4 h. Reducing sugar release into the reaction supernatant was determined with HBAH reagent.

(C) Cellulose azure assay

Enzyme was incubated with 40 mg of cellulose azure in 2.0 ml of citrate buffer pH 7.0 for 18 h at 37°C. The A₅₈₅ nm of the supernatant was measured after an appropriate time. One unit of activity gave a change in A₅₈₅ nm of 1.0 / h.

(D) C. fimi protease assay

C. fimi protease was incubated with 10 mg of Hide Powder Azure in 1.5 ml of phosphate buffer for 1 h at 37°C. The A₅₈₅ nm of the supernatant was measured after an appropriate time. One unit of activity gave a change in A₅₈₅ nm of 1.0 / h.

XIII Determination of protein concentration

Protein concentration was measured routinely by the dye-binding method (Bradford, 1976) using the Bio-Rad Protein Kit (Bio-Rad Laboratories, Canada).

For the binding assay, protein concentration was based on UV absorption at 280 nm. The extinction coefficient (E¹⁰⁰ mg/ml) for ngCenAΔPT was determined by the far-UV method of Scopes (1974). This value was used to estimate the absolute concentrations of solutions of the protein, which was then used to calculate the extinction coefficient at 280 nm (E¹⁰⁰ mg/ml).
RESULTS

I Construction of cenAΔPT

(A) Subcloning of the cenA gene into M13mp19

The 1588-bp SstI fragment encoding CenA was cloned into the SstI site of M13mp19 (Fig. 2a). The orientation of the insert was determined by restriction analysis of RF DNA (Fig. 3). A clone with the correct orientation was designated M13mp19cenA.

(B) in vitro mutagenesis

Loop-out primer extension was used to delete the sequence encoding the Pro-Thr box of CenA from the cenA gene (Fig. 2b).

If M13 is propagated on an E. coli dut ung strain, the progeny DNA has about 1% of its thymine residues replaced with uracil. Such DNA is degraded rapidly when the phage is plated subsequently on a dut+ ung+ host. This phenomenon can be used to enrich for mutants of M13 strains prepared in vitro (Kunkel, 1985).

M13mp19cenA was prepared by two cycles of infection on E. coli RZ1032 dut ung. The relative efficiencies of plating of the phage obtained were 1.0 and 10^{-6} on the dut ung and dut+ ung+ host, respectively.

The primer for the loop-out was a 48-mer which was complementary to 21 nucleotides on either side of the sequence encoding the Pro-Thr box of CenA, with a hexamer corresponding to a HpaI site replacing the Pro-Thr encoding sequence. There are no HpaI sites in M13mp19, pUC18 (see later) or the SstI fragment encoding CenA.

After primer extension and ligation, the reaction mixture was used to transform a dut+ ung+ E. coli strain. About 30,000 plaques were obtained. Small volume lysates were made from 12 plaques picked at random, and RF DNA prepared from each preparation. The DNA from 5 of the plaques contained a HpaI site (Fig. 4). DNA sequencing confirmed that these phage had the Pro-Thr box encoding sequence replaced by a HpaI recognition site (Fig. 4).
(C) Subcloning of cenAΔPT into pUC18

RF DNA was prepared from two individual M13mp19-cenAΔPT clones. The 1525-bp SstI CenAΔPT-coding fragment was cloned into the SstI site of the pUC18 vector, and transformed into JM101 (Fig. 2 c). The transformants were screened for CMCase activity (Fig. 6). A positive clone was designated pUC18-1.5cenAΔPT.

Extracts prepared from cells carrying pUC18-1.5cenAΔPT contained two polypeptides with CMCase activity (Fig. 7). The polypeptide of Mr 44 kDa corresponded to intact CenAΔPT, that of 30 kDa to an active degradation product.

The restriction pattern of pUC18-1.5cenAΔPT was as expected (Fig. 8).

II Purification of CenAΔPT by affinity chromatography on cellulose

CenAΔPT retained the cellulose-binding domain (CBD) of CenA. It could be purified by affinity chromatography on CF1 cellulose (Table I, and Fig. 9). CenAΔPT was eluted from cellulose with a high concentration of guanidine but could not be eluted with water. The enzyme obtained was virtually homogeneous (Fig. 10). Its Mr was 44.1 kDa, about 4.4 kDa less than that of CenA.

III Characterization of CenAΔPT

CenA from C. fimii is a glycoprotein, and is referred to as gCenA. CenA produced in E. coli is not glycosylated, and is referred to as ngCenA.

(A) Sensitivity to C. fimii protease

Both ngCenA and ngCenAΔPT were converted into a 30 kDa fragment within 24 hours (Fig. 11, A and B, lane 2-9). The enzymes were hydrolysed at the same rate (Fig. 11, A and B, lane 2-9). A transient 36 kDa fragment was produced from ngCenA but not from ngCenAΔPT (Fig. 11 B). The 30 kDa fragment is referred to as p30.

The N-terminal amino acid sequence of p30 from both enzymes was Val-Thr-Pro-Glu-Pro-Thr (Table III). This corresponds to amino acids 135-140 of CenA, and shows that the proteins are cleaved at the same site.
(B) Binding to Avicel

For the binding experiments, protein concentrations were determined by \( A_{280\text{nm}} \). This was considered to be reliable. Concentrations determined by dye binding were 84\% of those determined by absorbance (Table IV).

The adsorption isotherm for ngCenA\( \Delta \)PT with Avicel (Fig. 12 insert) indicated that approximately 1.06 nmol ngCenA\( \Delta \)PT/\( \text{mg}^{-1} \) Avicel was adsorbed at saturation. A Scatchard plot (Fig. 12 main panel) of the same data revealed an almost linear relationship between \([P]_{ad}\) and \([P]_{ad}/[P]\), indicating one class of binding interaction.

(C) Activity

The molar activities of ngCenA\( \Delta \)PT on a range of substrates were compared to those of ngCenA and p30 (Table V). The molar activity of p30 against CMC was 20\% higher than that of ngCenA. Higher activities for p30 were also observed against phosphoric acid-treated cellulose and cellulose azure, with 100\% and 300\% increases, respectively. However, the molar activity of p30 against Avicel was 22\% lower than that of ngCenA. These results agreed with those obtained previously (Gilkes et al., 1988).

The molar activities of ngCenA\( \Delta \)PT against all these substrates were lower than those of ngCenA: 50\% lower against CMC and phosphoric acid-treated cellulose, 62\% lower against cellulose azure and 40\% lower against Avicel.
<table>
<thead>
<tr>
<th>Material</th>
<th>Total protein (mg)</th>
<th>Total activity(^a) (U)</th>
<th>Specific activity (U/mg)</th>
<th>Recovery (%)</th>
<th>Fold purification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell ext.</td>
<td>19,800</td>
<td>11,090</td>
<td>0.56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strep. sulf.</td>
<td>15,444</td>
<td>11,189</td>
<td>0.725</td>
<td>100.8(^b)</td>
<td>1.3</td>
</tr>
<tr>
<td>Eluate pool</td>
<td>43.32</td>
<td>5,465</td>
<td>126.15</td>
<td>48.84(^c)</td>
<td>225</td>
</tr>
</tbody>
</table>

\(^a\) CMCase activity is expressed as micromoles of glucose equivalents released per min at 37°C.

\(^b\) This more than 100% recovery is due to degradation of CenAAΔPT during purification. The degradation product, p30, has a higher specific activity than CenAAΔPT.

\(^c\) This recovery is underestimated due to the activity contributed by p30 in the crude cell extract.
TABLE III  N-terminal amino acid sequence of p30 produced from ngCenAAPT by C. fimi protease

<table>
<thead>
<tr>
<th>Cycle</th>
<th>Complete digestion (66)(^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>V (21.3)(^b)</td>
</tr>
<tr>
<td>2</td>
<td>T (13.8)</td>
</tr>
<tr>
<td>3</td>
<td>P (13.71)</td>
</tr>
<tr>
<td>4</td>
<td>Q (13.39)</td>
</tr>
<tr>
<td>5</td>
<td>P (12.40)</td>
</tr>
<tr>
<td>6</td>
<td>T (10.18)</td>
</tr>
<tr>
<td>7</td>
<td>S (5.53)</td>
</tr>
<tr>
<td>8</td>
<td>G (10.33)</td>
</tr>
<tr>
<td>9</td>
<td>F (7.08)</td>
</tr>
<tr>
<td>10</td>
<td>Y (7.79)</td>
</tr>
<tr>
<td>11</td>
<td>V (11.19)</td>
</tr>
<tr>
<td>12</td>
<td>D (7.69)</td>
</tr>
<tr>
<td>13</td>
<td>P (6.31)</td>
</tr>
<tr>
<td>14</td>
<td>T (4.08)</td>
</tr>
<tr>
<td>15</td>
<td>T (7.51)</td>
</tr>
<tr>
<td>16</td>
<td>Q (6.30)</td>
</tr>
<tr>
<td>17</td>
<td>G (9.02)</td>
</tr>
<tr>
<td>18</td>
<td>Y (5.36)</td>
</tr>
<tr>
<td>19</td>
<td>R (18.49)</td>
</tr>
<tr>
<td>20</td>
<td>A (14.88)</td>
</tr>
<tr>
<td>21</td>
<td>W (0.51)</td>
</tr>
<tr>
<td>22</td>
<td>Q (5.17)</td>
</tr>
<tr>
<td>23</td>
<td>A (15.48)</td>
</tr>
</tbody>
</table>

a. Sample amount (picomoles)
b. Yield (picomoles).
Table IV  Extinction coefficient of ngCenAΔPT

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>1mg/ml E&lt;sub&gt;205&lt;/sub&gt;</th>
<th>1mg/ml E&lt;sub&gt;280&lt;/sub&gt;</th>
<th>1mg/ml E&lt;sub&gt;280&lt;/sub&gt;</th>
<th>Protein concentration determined by Bradford assay relative to UV assay (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ngCenA</td>
<td>32.78</td>
<td>2.64</td>
<td>2.47</td>
<td>81</td>
</tr>
<tr>
<td>ngCenAΔPT</td>
<td>35.785</td>
<td>2.62</td>
<td>2.59</td>
<td>84</td>
</tr>
<tr>
<td>p30</td>
<td>35.12</td>
<td>2.38</td>
<td>2.42</td>
<td>89</td>
</tr>
</tbody>
</table>

<sup>a</sup> Determined as described by Scopes. (1974).
<sup>b</sup> The absolute concentration of a solution was determined using E<sub>205</sub> and then used to estimate E<sub>280</sub>.
<sup>c</sup> Predicted from Tyr and Trp content, accordingly to Cantor and Shimmel (1980).
<sup>d</sup> The concentration of a solution determined by the Bradford dye-binding assay (relative to BSA) (Bradford, 1976) is expressed as a percentage of that determined from absorbance at 280 nm.
TABLE V. Enzyme activity against a range of cellulosic substrates

<table>
<thead>
<tr>
<th>Sample</th>
<th>Enzyme Activity</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Avicel (mkatal/mol)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>PASC&lt;sup&gt;b&lt;/sup&gt; (katal/mol)</td>
<td>CMC (katal/mol)</td>
<td>Cel. Azure (units/mmol)&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>ngCenA</td>
<td>25.3±0.7</td>
<td>3.6±0.2</td>
<td>20.28±0.3</td>
<td>215±41</td>
</tr>
<tr>
<td>ngCenAΔPT</td>
<td>15.03±0.3</td>
<td>1.88±0.2</td>
<td>11.17±0.1</td>
<td>79.4±6</td>
</tr>
<tr>
<td>p30&lt;sup&gt;c&lt;/sup&gt;</td>
<td>19.77±0.2</td>
<td>7.36±0.6</td>
<td>24.54±0.2</td>
<td>785.8±30</td>
</tr>
</tbody>
</table>

<sup>a</sup> katal/mol: activity of 1 mol product/second/mol enzyme.
<sup>b</sup> PASC, phosphoric acid-swollen cellulose.
<sup>c</sup> units/mmol: production of OD595 = 1/hour/mmol enzyme.
<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Total amino acid residues&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Observed&lt;sup&gt;b&lt;/sup&gt; $M_r$ (kDa)</th>
<th>Predicted&lt;sup&gt;c&lt;/sup&gt; $M_r$ (kDa)</th>
<th>Observed less predicted (kDa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ngCenA</td>
<td>418</td>
<td>48.7</td>
<td>43.8</td>
<td>4.9</td>
</tr>
<tr>
<td>ngCenAΔPT</td>
<td>397</td>
<td>44.1</td>
<td>41.8</td>
<td>2.3</td>
</tr>
<tr>
<td>p30</td>
<td>284</td>
<td>30</td>
<td>30.2</td>
<td>0.2</td>
</tr>
</tbody>
</table>

<sup>a</sup> From the deduced amino acid sequence of the mature enzyme (Wong et al., 1986).

<sup>b</sup> Determined by SDS-PAGE.

<sup>c</sup> Calculated from the deduced amino acid sequence.
Fig. 2 Scheme for the construction of CenA Pro-Thr box deletion plasmid pUC18-cenAΔPT

a) subcloning of cenA to M13mp19  
b) oligonucleotide directed mutagenesis  
c) subcloning of cenAΔPT into pUC18

mcs denotes multiple cloning site;  
BD denotes binding domain sequence  
CD denotes catalytic domain of CenA  
PT denotes Pro-Thr box region  
S, B, H and E denote restriction sites for SstI, BamHI, HpaI and EcoRI.  
--- denotes single stranded DNA
prepare ssDNA from ung dug host cell

5'-pGC TGG TCG GCT GCG GCG TGA
CGT TAA GGC TGG TCG GCA GCG TGC C-3'

annealing
T7 DNA polymerase
- d NTP
- ATP
- T4 DNA ligase

transfection to ung⁺ host,
select HpaI digestion positive clones

C.

T4 DNA ligase
Figure 3. Screening of M13mp19-*cenA* clones by *BamHI* digestion of phage DNA

Mini-prep DNA from infected cells was digested with *BamHI* and electrophoresed through a 0.7% agarose gel

- lane 1 and 14: λ DNA digested with *HindIII/EcoRI* (standard)
- lane 2: M13mp19 DNA digested with *BamHI* (control)
- lane 3-13: recombinant phage DNA digested with *BamHI*
  - positive clones: 8.6 kb + 0.23 kb (lane 3, 6, 8, 11, 12)
  - negative clones: 7.47 kb + 1.37 kb (lane 4, 5, 7, 9, 10, 13)
Figure 4. Screening of potential deletion mutants (M13mp19-cenAΔPT) by HpaI digestion of RF DNA. Mini-prep DNA from infected cells was digested with HpaI and electrophoresed through a 0.7% agarose gel.

- lane 1: λ DNA digested with HindIII (standard)
- lane 2: λ DNA digested with HpaI (control)
- lane 3-5: M13mp19 DNA: uncut; cut with HpaI; cut with EcoRI (control)
- lane 6-17: recombinant phages, H: HpaI digestion;
  - -: uncut
  - deletion mutant: HpaI digestion positive (phage clones 3,8,10,11,12)
- lane 18: single-stranded M13mp19-cenAΔPT DNA
- o.c. denotes open circular form DNA
- linear denotes linear form DNA
- c.c.c. denotes covalently closed circular form DNA
N--- Val Pro Thr Thr Ser Val (Asn Val) Thr Pro Gln Pro Thr Ser---C  CenAΔPT replace

N--- Val Pro Thr Thr Ser Val (Pro_{11}Thr_{12}) Thr Pro Gln Pro Thr Ser---C  CenA

Figure 5. Dideoxy nucleotide sequencing gel showing deletion of the Pro-Thr box coding sequence from the cenA gene. The sequence of recombinant phage M13mp19-cenA/M13mp19cenAΔPT was obtained using a chemically synthesized 17-mer oligonucleotide.
Figure 6. Screening of potential deletion mutants (pUC18-\textit{cenA}\Delta PT) for CMCase activity. Transformed clones were transferred to pairs of gridded LB+ampicillin+IPTG+CMC plates. After incubation overnight at 37°C, one plate from each pair was stained with Congo-red and destained with 1M NaCl. Positive clones were surrounded by haloes. pUC18 was the negative control (bottom left), pUC18-\textit{cenA} was the positive control (bottom right).
Figure 7. Detection of ngCenAΔPT in cell extracts  Zymogram of polypeptides from JM101/pUC18-1.5cenAΔPT separated by SDS-PAGE (see Materials and Methods). Lane 1 and 2 denote cenAΔPT clone 1 and 2.
Figure 8. Restriction analysis of pUC18-cenAΔPT clones.

pUC18: control
pUCEC2: control
1 and 2: deletion mutant clones 1 and 2
H, S and B denote HpaI, SstI and BamHI digestion respectively; - denotes uncut DNA sample.
Figure 9. Purification of ngCenAΔPT by affinity chromatography on cellulose. See Materials and Methods for details. After adsorption of the enzyme in 50mM phosphate buffer, pH 7.0, the column was washed with 1M NaCl, then 50mM phosphate buffer. The enzyme was eluted with a 0 to 8 M linear gradient of guanidine-HCl. Fractions were 10ml. ○ A$_{280}$nm, • CMCase activity; △ Guanidine.HCl concentration.
Figure 10. SDS-PAGE analysis of the purification of ngCenAΔPT

- lane 1 cell extract
- lane 2 streptomycin sulfate treated cell extract
- lane 3 purified ngCenA (control)
- lane 4 Mr standard
- lane 5 pool of active fractions from the cellulose column
Figure 11. Time course of proteolysis of ngCenA and ngCenAΔPT. Reaction mixture contained 35 μg of ngCenA (A) or ngCenA (B), 200 μl of phosphate buffer and 0.2 unit of crude C. fimi protease. They were incubated at 37°C. Reactions were sampled at 0, 30min, 1h, 2h, 4h, 6h, 11h, 24h, (A and B, lane 2-9, respectively), treated with PMSF and analyzed by SDS-PAGE (10% acrylamide). Control samples were incubated in the absence of protease, for 24h (A and B, lane 10). All lanes were loaded with sample equivalent to 3.4 μg of initial protein. Lane 1, Mr marker.
Figure 12. Adsorption of ngCenAΔPT and ngCenA to Avicel The main panel is a Scatchard plot of the adsorption data (see Materials and Methods for details). The inset shows the data plotted as adsorption isotherm. • ngCenAΔPT; ○ ngCenA.
Fig. 13 Sequence alignments of CenA and CenAΔPT Pro-Thr box and flanking regions, and locations of *C. fimi* protease primary cleavage sites in the nonglycosylated proteins. Amino acid sequences were deduced from the nucleotide sequence of the *cenA* gene. They are numbered with reference to the amino terminus of mature CenA. Underlined sequence was determined by amino acid sequencing (Table III). ---denotes a gap to allow alignment of the flanking regions. Boxed amino acids are encoded by the nucleotides constituting the *HpaI* site of CenAΔPT.
DISCUSSION

The CenAΔPT polypeptide appeared to be stable in *E. coli* and could be recovered from cell extracts.

The observed and predicted Mₘₚs for the various polypeptides are: ngCenA 48.7 kDa and 43.8 kDa; p30: 30.4 kDa and 29.6 kDa (Gilkes *et al.*, 1988); ngCenAΔPT 44.1 kDa and 41.8 kDa. This shows that the discrepancy between the observed and predicted Mₘₚs for ngCenA is caused largely by the conformation of the Pro-Thr box sequence of the protein, but the conformation of CBD may also be a factor.

Proteolysis studies of CenA with different proteases under denaturing conditions show that the resistance of the core protein to proteolysis is not due to an absence of cleavage sites within the primary structure of the core peptide, p30 in the case of ngCenA, but to its conformation (N. Gilkes, unpublished result). Proteolysis of ngCenAΔPT also results in a core peptide of Mₘ 30 kDa which has the same N-terminal amino acid sequence as the p30 released from CenA. This suggests that the sequence before the cleavage site is not critical for recognition by *C. fimii* protease. However, the Pro-Thr box in CenA lies within an extended sequence rich in proline and hydroxyamino acids (N-TTCTGTVPITTS¹⁰¹-¹¹¹[Pro-Thr box]¹¹²-¹³⁴VTPQQPYSG¹³⁵-¹⁴³-C) so the portion left in ngCenAΔPT may mimic the intact Pro-Thr box and give a protease sensitive bridge.

Proteolytic processing of the terminal cellulose-binding domains has been proposed as a means of regulating and adjusting the substrate specificities of cellulases during the hydrolysis of complex carbohydrates (Knowles *et al.*, 1987). While there is insoluble substrate to hydrolyse, cellulases bind to it and the glycosylated enzymes are protected from proteolysis. Once the substrate has been solubilized, the enzymes become free in solution and become susceptible to limited proteolysis. Removal of the cellulose-binding regions produces enzymes with improved affinities for shorter, soluble cellodextrins.

From the adsorption isotherms, 1.06 and 1.14 nmole. mg⁻¹ of ngCenAΔPT and ngCenA, respectively, are required to saturate Avicel PH101 under the stated conditions.
This suggests that the overall binding to cellulose is not influenced by the Pro-Thr box. Adsorption of ngCenA to Avicel involves two types of binding: high affinity and low affinity (Fig. 12, main panel, open circles). Scatchard analysis of the adsorption of ngCenAAPT to Avicel (Fig. 12, main panel closed circles) reveals a low affinity interaction but hardly any high affinity interaction. There are two possible explanations for this (N.R. Gilkes, personal communication). One is that there are two types of site for binding on the cellulose: the crystalline and amorphous regions. The second is that there are two binding sites on the CBD with a single class of binding site on Avicel. The loss of the high affinity component but retention of the low affinity component for binding by ngCenAAPT supports the latter proposal. The removal of the Pro-Thr box may cause some conformational change which eliminates the high affinity site selectively. However, the following evidence favours the first proposal. Avicel has a heterogeneous structure comprising both crystalline and disordered or amorphous regions (Kulshreshtha and Dwetz, 1973). Treatment with H$_3$PO$_4$ converts Avicel to fully amorphous cellulose. The crystalline cellulose structure is changed when it is swollen in concentrated H$_3$PO$_4$, resulting in loss of the parallel chain structure of the original Avicel microcrystals and also a significant reduction in particle size and an increase in surface area. (Blackwell, 1981; Lee et al., 1982; Ooshima et al., 1983; Sarko, 1986). The binding of ngCenA to such amorphous cellulose exhibits only the low affinity component.

ngCenA but not ngCenAAPT can be eluted from cellulose with water. This suggests that the Pro-Thr box may also affect the binding interaction between the enzyme and the crystalline regions of cellulose. This further suggests that some conformational changes may result from the deletion. The relative affinities of CenA and CenAAPT for cellulose need to be examined by adsorbing them to cellulose, then determining the conditions for eluting them.

ngCenAAPT has significantly lower activity than ngCenA on a range of cellulosics, including microcrystalline (Avicel), amorphous (phosphoric acid-treated cellulose, and cellulose azure) and soluble (CM-cellulose) substrates (Table V). This indicates that the Pro-Thr box has a crucial role in the normal behavior of this enzyme. The Pro-Thr box may
give the enzyme a conformation that makes the catalytic domain more accessible to the substrate.

The catalytic activity of p30 against the same range of cellulosic substrates (Table V) agreed with a previous study (Gilkes, et al., 1988). Its decreased activity on microcrystalline cellulose suggests that the binding domain is required for efficient degradation of microcrystalline cellulose. Recent experiments have shown that ngCenA can effect the fragmentation of cotton fibres into small particles, whereas p30 can not (Kilburn et al., 1989). This also implies an active role for the CBD of CenA in cellulose degradation. The activity of p30 against soluble and amorphous substrates was significantly increased suggesting that the binding domain somehow inhibits the function of the catalytic domain when the function of the binding domain is not needed. Perhaps the conformation of the entire enzyme makes it less accessible to the substrate than the core protein alone.

The Proline and hydroxyamino acid rich segments occur in more than 20 cellulases from various organisms (personal communication with R.A.J. Warren, ). Of these, T. reesei cellulases display a bifunctional organization that closely resembles that of CenA and Cex (Tomme et al., 1988; van Tilbeurgh et al., 1986). The binding domain of CBH I is at the N-terminus, that of CBH II is at the C-terminus. A Pro-Thr-Ser rich sequence occurs between the two functional domains of CBH I and a repeated Pro-Thr-Ser sequence occurs between the two functional domains of CBH II. CBH II resembles CenA closely. Studies of the truncated proteins obtained by papain hydrolysis of T. reesei CBH I and CBH II showed that their binding to and hydrolysis of Avicel are markedly reduced. The truncated core protein of CBH II has significantly reduced catalytic activity against amorphous cellulose. This reduced activity in the absence of the binding domain is reminiscent of p30 (Tomme et al., 1988; van Tilbeurgh et al., 1988). However, papain hydrolysis removes the Pro-Thr-Ser region as well as the binding domain, so the effect on binding and hydrolysis of cellulose could be all or partly due to the influence of the Pro-Thr-Ser sequence.

The Pro-Thr box of C. fimi is very similar in sequence to the hinge region of IgA1
The properties of ngCenAΔPT suggest that the Pro-Thr box also functions as a hinge. The elimination of high affinity binding and the significant reduction in catalytic activity could reflect a lack of conformational flexibility. The differential responses to C. fimi protease of native CenA free in solution or bound to Avicel (Langsford, et al., 1987) suggest that there is conformational change when CenA binds to Avicel. It would be interesting to compare the sensitivities of ngCenA and ngCenAΔPT to C. fimi protease when they are bound to cellulose.

Small angle X-ray scattering studies of T. reesei CBH I and CBH II (Abuja et al., 1988; Abuja et al., 1988; Schmuck et al., 1986) indicate that both CBH I and CBH II have a tadpole-like structure. The core proteins produced by partial papain proteolysis of CBH I and CBH II are the heads in both cases, and the binding domains together with the Pro-Thr-Ser rich regions are the tails. Similar analysis of CenA (N. Gilkes, unpublished result) also reveals a tadpole-like structure. Small angle X-ray scattering studies of CenAΔPT and comparison to CenA should offer insights into how the Pro-Thr box is linked to the binding domain.

It has been shown that neither CenA nor Cex are cleaved by the Neisseria gonorrhoeae IgA₁ protease (N. Gilkes, unpublished observations). Replacement of the Pro-Thr box with the IgA hinge region and examination of its sensitivity to the Neisseria gonorrhoeae IgA₁ protease will offer useful information about possible changes in the conformation of CenA caused by a different hinge region.
REFERENCES


Beguin, P., (in press) Molecular biology of cellulose degradation


Coughlan, M.P. 1985. The properties of fungal and bacterial cellulases with comment on


