Characterization of the Mannan-Degrading System of *Cellulomonas fimi*

by

Dominik Stoll

Diplom in Biologie II, Universität Basel, 1992

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

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

(Department of Microbiology and Immunology / Biotechnology Laboratory)

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA
August 1998

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 19/8/98

DE-6 (2/88)

Abstract

In this study the mannan-degrading system of the Gram positive bacterium *Cellulomonas fimi* was characterised. *C. fimi* can degrade different forms of mannan and can use the degradation products as carbon and energy source. This study focuses on the galactomannan-degrading system which was found to be composed of one secreted endo-1,4-β-mannanase (Man26A), one intracellular 1,4-β-mannosidase (Man2A) and one intracellular 1,6-α-galactosidase. The genes encoding Man26A and Man2A have been isolated and sequenced, and the enzyme activities were investigated.

The endo-1,4-β-mannanase (Man26A) has a multidomain structure and comprises a family 26 catalytic domain, a mannan-binding domain (MBD), a S-layer homology domain (SLH domain) and a domain of yet unknown function. Mannanase activity was detected on the cell surface and in the culture supernatant. It is believed that the SLH domain mediates transient binding of Man26A to the cell surface and the MBD mediates binding to the substrate. Strong binding of the MBD to soluble mannan was detected and its potential as an affinity tag for protein purification in aqueous 2-phase systems was tested.

The 1,4-β-mannosidase (Man2A), cleaves β-1,4 mannosidic linkages with net retention of the anomeric configuration. Man2A was transformed into the glycosynthase Man2A E519A by mutating the catalytic nucleophile, E519 to alanine. Glycosynthases are retaining glycosidases without hydrolytic but with synthetic activity. Using α-mannosyl fluoride as donor and *p*-nitrophenyl sugars as acceptors, the glycosynthase Man2A E519A catalyzed the synthesis of β-1,4 and β-1,3 mannosidic linkages.

In this study the biology of mannan-degradation by *C. fimi* was investigated and the biotechnological potential of its components was explored.

Table of Contents

List of Tables

List of Figures

List of Abreviations

2FManβF	2-deoxy-2-fluoro-mannosyl-β-fluoride
A	Absorbance
aa	amino acid
AFGE	Affinity gel electrophoresis
Amp	Ampicillin
Amp^R	Ampicillin resistance
Amu	Atomic mass unit
bp	Base pair
BSA	Bovine serum albumin
CBD	Cellulose-binding domain
CD	Catalytic domain
CIF	Collision induced fragmentation
CM-cellulose	Carboxymethylcellulose
Da	Dalton
DP	Degree of polymerization
EDTA	Ethylenediaminetetraacetic acid
ESMS	Electrospray mass spectrometry
EtOH	Ethanol
FACE	Fluorophore assisted carbohydrate electrophoresis
HCA	Hydrophobic cluster analysis
His_6 tag	Histidine tag
HPLC	High pressure liquid chromatography

I	Inhibitor
INM	Ivory nut mannan
IPTG	Isopropyl-β-D-thiogalactoside
ISV	Ion source voltage
Kan^R	Kanamycin resistance
k_{cat}	enzyme turnover constant
K_d	Dissociation constant
kDa	Kilodalton
k_i	Inactivation rate consant
K_i	Inhibitor equilibrium binding constant
K_m	Michaelis constant
k_{react}	Spontaneous reactivation rate constant
K_{trans}	Dissociation constant
k_{trans}	Transglycosylation rate constant
LB	Luria-Bertani medium
LC	Liquid chromatography
m/z	Mass to charge ratio
Man26A	1,4-β-mannananse
Man2A E519A	Man2A derived glycosynthase
Man2A	1,4-β-mannosidase
MBD	Mannan-binding domain
MCAC	Metal chelating affinity chromatography
M_r	relative molecular mass
MuαGal	Methylumbelliferyl α-galactoside

MUβMan	Methylumbelliferyl β-mannoside
MW	molecular weight
NMR	Nuclear magnetic resonance
PCR	Polymerase chain reaction
PFGE	Pulsed field gel electrophoresis
PNPC	paranitrophenyl β-cellobiose
PNPG	paranitrophenyl β-gentiobiose
PNPGal	paranitrophenyl β-galactose
PNPM	paranitrophenyl β-mannoside
$PNPM_2$	paranitrophenyl β-mannobioside
rpm	Revolutions per minute
RT	Room temperature
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SLH	S-layer homology
TFA	Trifluoroacetic acid
TIC	Total ion chromatogram
TLC	Thin layer chromatography
TYP	Tryptone, yeast extract, phosphate
UV	Ultraviolet
v_{max}	Maximal velocity of enzyme-catalyzed reaction
X-Glu	5-Bromo-4-chloro 3-indolyl-β-D-glucoside

Acknowledgments

I wish to extend my thanks to my supervisor Dr. Tony Warren for giving me the opportunity to pursue my Ph.D. at UBC and for providing continual advice and encouragement. Special thanks go to Dr. Steve Withers for his advice and enthusiasm. I thank Drs. Tom Beatty, Brett Finlay, Neil Gilkes, Doug Kilburn and George Spiegelman for their guidance.

My appreciation also goes to the members of the cellulase lab for their team spirit and friendship. In particular I would like to thank Dr. Andreas Meinke for his mentor- and friendship, Dr. Henrik Stålbrand for helping me to get started on my new project, Al Boraston for many fruitful discussions and for being a very good listener, Brad McLean for his computer tutorials and Dr. Laurent Gal for the transglycosylation discussions. Thanks also go to Lloyd Mackenzie for helping me with the transglycosylation product analysis, Shoumin He for doing all the ESMS work, Renee Mosi, Steve Howard and Dave Vocadlo for substrates and advice. My appreciation also goes to John Smit for his electron microscopy work. I am grateful for the help provided by Emily Kwan and Emily Akow. Many thanks go to Helen Smith for all her work to keep the lab organized and for the many opportunities for cellulase family bonding either at BBQs in her backyard or on the skislopes in Whistler.

I appreciate the financial support received from the Protein Engineering Network of Centres of Excellence (PENCE).

Last but not least, I am especially grateful to my wife, family and friends for their encouragement and support throughout these many years.

1 Introduction

1.1 Gycosidases

Carbohydrates make up most of the organic matter on earth, originating mostly from CO_2 assimilation by oxygenic photosynthesizers, such as plants, algae and cyanobacteria (Raven, 1996). Carbohydrates are found in all forms of life, displaying a wide variety of biological functions, such as energy storage, structural components, fuel and metabolic intermediates (Stryer, 1988).

Most of the carbohydrates participate in glycosidic linkages, such as N-glycosidic linkages, which are found in the base-deoxyribose linkage of DNA nucleotides, or O-glycosidic linkages, which are most commonly found in carbohydrate-carbohydrate linkages. These linkages are the substrates for glycosidases, i.e. N-glycosidases and O-glycosidases, respectively. The abundance of O-glycosidically linked carbohydrates, their biological and medical importance and their biotechnological potential are the driving force for the extensive studies of O-glycosidases (Davies *et al.*, 1998). One class of O-glycosidases is polysaccharidases, which are key components in the recycling of carbohydrates in nature. Polysaccharidases are involved in the microbial degradation of plant cell wall material or in the degradation of nutritional reservoirs, such as starch, glycogen or certain forms of mannan (Warren, 1996).

Glycosidases are also present in lysosomes, where they play an important role in the catabolism of the sugars in sphingolipids, glycoproteins or glycosaminoglycans. The malfunctioning or absence of these enzymes can cause severe diseases, known as lysosomal storage diseases. The loss of β-hexosaminidase activity, for example, results in the accumulation of G_{M2} ganglioside primarily in neuronal lysosomes, causing the

neurodegenerative Tay-Sachs disease (Mahuran, 1995). Other lysosomal storage diseases are Sly syndrome, which is caused by a β-glucuronidase deficiency, and β-mannosidosis, which is caused by insufficient β-mannosidase activity (Neufeld, 1991).

Extensive research on glycosidases not only helps to understand the biological process of glycoside hydrolysis, but has and will continue to reveal many useful applications. Glycosidases, for example, have become useful tools in the field of glycobiology. Structural information about the sugar moiety of glycoproteins can be obtained by the sequential removal of terminal saccharides using exo-acting glycosidases (Edge *et al.*, 1992). Glycosidases are also valuable for crystallography by improving crystallisation by *in-vitro* deglycosylation of glycosylated proteins (Grueninger-Leitch *et al.*, 1996). A novel and very promising application in the development of new therapeutics is the use of glycosidase mutants, i.e. glycosynthases, for the enzymatic *in-vitro* synthesis of glycosidic linkages (Mackenzie *et al.*, 1998; Withers *et al.*, 1998).

The biotechnological potential for plant cell wall degrading polysaccharidases of microbial origin is well established. It includes applications such as the conversion of cellulose to fuel, biobleaching of paperpulp, and 'stone-washing' of denim fabric.

1.2 Catalytic mechanisms of glycosidases

Glycosidases cleave their substrates between a glycon and an aglycon part. They usually have a very high specificity for the atomic nature of the glycon, but fairly low specificity for the aglycon. Aglycons, in nature mostly sugars, can therefore be replaced by chromogenic or fluorogenic substrates such as dinitrophenyl or methylumbelliferyl groups, which make very useful analytic substrates. Two major mechanisms are known for the

cleavage of the glycosidic linkage by glycosidases; cleavage with net retention of the anomeric configuration (retaining enzymes) or with net inversion of the anomeric configuration (inverting enzymes) (Sinnott, 1990). As was proposed by Koshland 45 years ago, the inverting mechanism works via a single chemical step, and the retaining mechanism via a double displacement step (Koshland, 1953).

The double-displacement mechanism of retaining glycosidases is facilitated by acid/base catalysis. The reaction proceeds in two reaction steps, a glycosylation and a deglycosylation step (Figure 1.1). Generally two carboxylic side chains are involved. One acts as an acid and base catalyst, whereas the second carboxylic side group acts as nucleophile. The glycosidic oxygen is protonated by the acid catalyst, and the nucleophile attacks the anomeric center of the glycon, thereby forming a covalent sugar-enzyme intermediate with an inverted anomeric configuration. In the second step, the deglycosylation, water is deprotonated by the acid/base catalyst, increasing its nucleophilicity. Its attack on the anomeric center of the glycosyl-enzyme intermediate, reverses the anomeric configuration again, resulting in an overall retention. The glycosylation and the deglycosylation steps proceed via oxocarbenium ion-like transition states (Davies *et al.*, 1998). A well studied retaining enzyme is the *C. fimi* exoglucanase/xylanase Cex, a member of family 10 (Notenboom *et. al.*, 1998). Kinetic studies have shown that Glu127 acts as the acid/base catalyst and Glu233 as the catalytic nucleophile (MacLeod *et al.*, 1994; Tull *et al.*, 1991). The covalent ester linkage between the enzyme nucleophile and the sugar was demonstrated with mechanism based inhibitors. 3D structure analysis of the retaining enzymes Cex (White *et al.*, 1994), *E. coli* (*lacZ*) β-galactosidase (Jacobson, 1994) and the human β-glucuronidase (Sanjeev *et al.*, 1996) indicated that the carboxylate oxygens from the two catalytic residues are about 5.5 Å apart

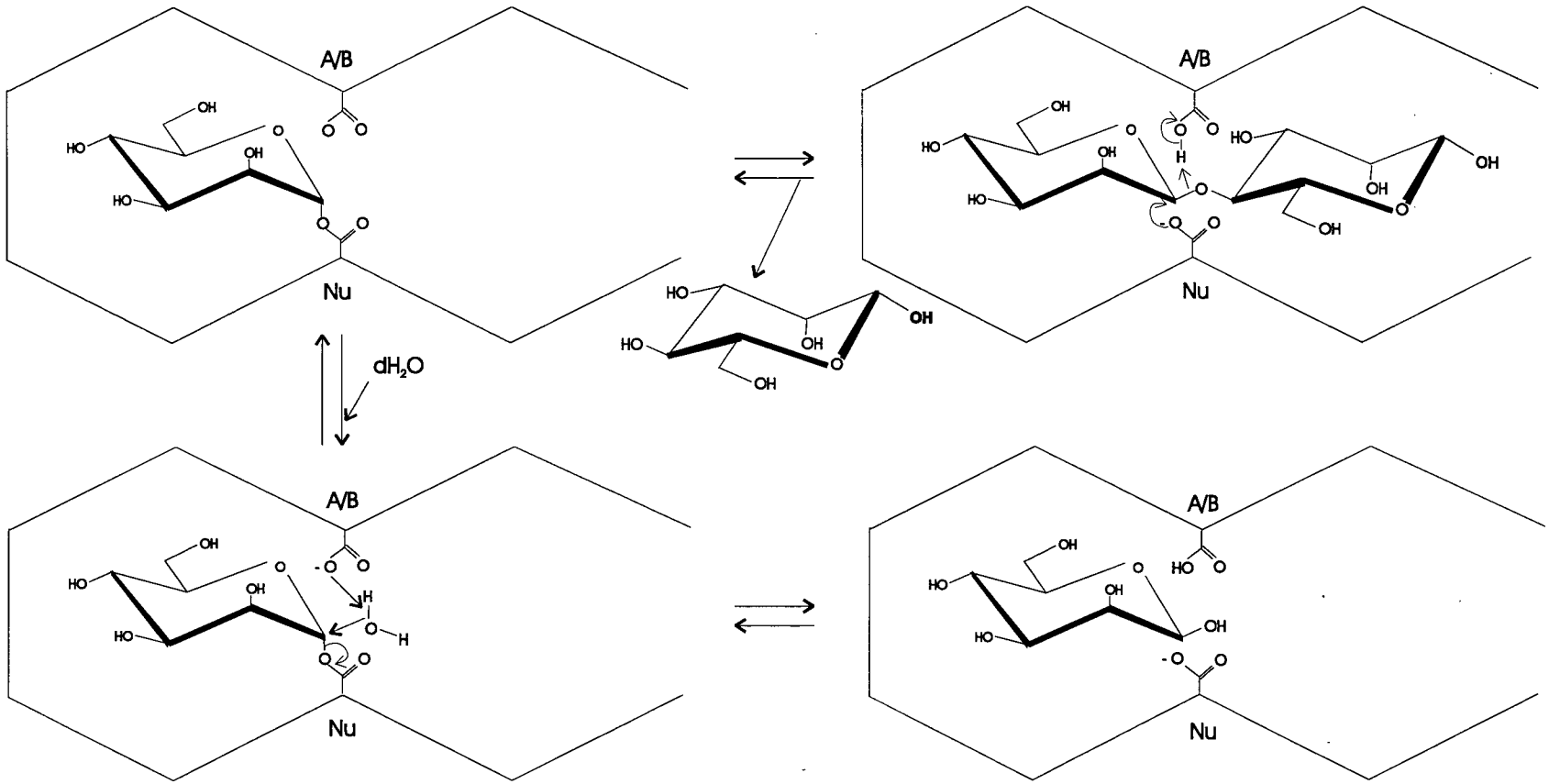


Figure 1.1: Catalytic mechanism of a retaining β-1,4-mannosidase. A/B is the acid/base catalyst and Nu is the catalytic nucleophile. The substrate shown is mannosyl-β-1,4-mannoside. (Adapted from Mackenzie *et al.*, 1998)

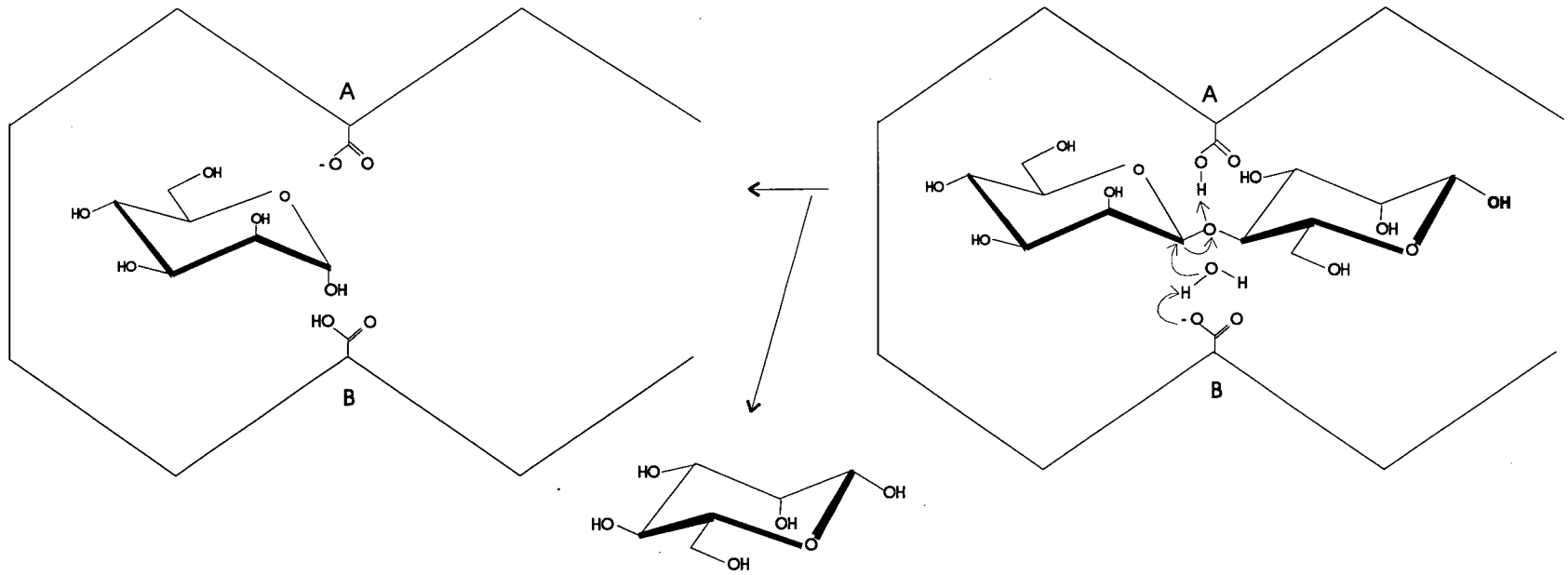


Figure 1.2: Catalytic mechanism of an inverting β-1,4-mannosidase. A is the acid catalyst and B is the base catalyst. The substrate shown is mannosyl-β-1,4-mannoside. (Adapted from Mackenzie *et al.*, 1998)

(McCarter and Withers, 1994). The catalytic nucleophile does not have to be part of the enzyme; it can be part of the substrate. The 2-acetamido side chains, in chitinase or hexosaminidase substrates can attack the anomeric center, forming an oxazoline intermediate. This intermediate corresponds to the glycosyl-enzyme complex. Such a pathway could explain the lack of a catalytic nucleophile in the chitobiase from *Serratia marcescens* (Tews *et al.*, 1996).

In inverting glycosidases, the protonation of the glycosidic oxygen and the deprotonation of the water molecule is catalyzed by carboxylic amino acid side groups, from either aspartates or glutamates. The thereby deprotonated, nucleophilic water can then attack the anomeric center from "the opposite side", resulting in a net inversion of the anomeric configuration (Figure 1.2). As was shown by kinetic studies for the family 6 endoglucanase CenA from *Cellulomonas fimi*, two aspartic acid residues, D252 and D392, act as catalytic acid and catalytic base, respectively. These catalytic residues are conserved in all members of family 6 (Damude *et al.*, 1995). From 3-dimensional structures, e.g. of cellobiohydrolase CbhII from *Trichoderma reseei* (Rouvinen *et al.* 1990) and endo-glucanase E2 form *Thermomonospora fusca* (Spezio *et al.*, 1993), it was found that the two carboxylate oxygens from the catalytic residues are about 9.5 Å apart. This distance is believed to be important to enable simultaneous attack of the glycosidic oxygen by the acid catalyst and the nucleophilic attack at the anomeric center by the deprotonated water (McCarter and Withers, 1994).

1.3 Classification of Glycosidases

A great variety of O-glycosyl hydrolases (EC 3.2.1.-) have been characterized. They are classified into 62 families based on amino acid sequence similarities (Henrissat, 1991; Henrissat, 1993; Warren, personal communication). Since there is a direct relationship between sequence similarities and folding similarities, this classification allows better structural, evolutionary and mechanistic predictions.

Not all members of a family have the same substrate specificity: family 1 comprises β-glucosidases, 6-phospho-β-galactosidases, 6-phospho-β-glucosidases, myrosinases, lactase-phlorizin hydrolases and one β-mannosidase (Bauer *et al.*, 1996). This discrepancy between sequence similarity and different substrate specificity is indicative of evolutionary divergence. The stereochemical outcome of hydrolysis however, is the same for all members within a family (e.g. retaining in family 1). Members within a family also adopt a similar fold. A commonly used tool for predicting the secondary structure of a protein, i.e. the folding, is hydrophobic cluster analysis (HCA) (Gaboriaud *et al.*, 1987). HCA was used to compare the folding of enzymes with known 3D structures from families 1, 2, 5, 10, and 17 with other enzymes from families 1, 2, 5, 10, 17, 26, 30, 35, 39, 42 and 53. Even though they share low levels of sequence identities, HCA analyses revealed that the catalytic domains from members of these families have similar $(\beta/\alpha)_8$ folds. The acid/base catalyst and the nucleophile were found to be located at the C-terminal ends of strands β4 and β7, respectively. These families were placed, based on their structural similarities, into the clan GH-A (Henrissat *et al.*, 1996). Four more of these structural superfamilies were defined; clans GH-B, GH-C, GH-D and GH-E. Based on these classification premises, and the use of HCA, the location of the eight β strands and functional groups were predicted for various

lysosomal enzymes of the clan GH-A, such as the human β-glucosidase (family 30), human β-galactosidase (family 35) and human β-mannosidase (family 2) (Durand *et al.*, 1997).

A further characteristic of glycosidases is their overall topology. 3D structures from more than 22 families have shown, that regardless of protein fold and reaction mechanism variety, the glycosidases have only three different overall topologies. The pocket or the crater topology is optimal for the recognition of saccharide extremities, and is encountered in monosaccharidases, such as β-galactosidases and β-glucosidases. The cleft or groove is an open structure, which allows random binding of several sugar units. This form is commonly found in endo-acting polysaccharidases, such as endo-cellulases and chitinases. The third topology found is the tunnel, which is basically a cleft covered by long loops. This form is commonly found for exo-acting enzymes such as cellobiohydrolases (Davies *et al.*, 1995).

1.4 Strategies for the enzymatic degradation of plant cell wall polysaccharides

The growing plant cell is protected and structurally supported by its primary cell wall. As the plant cell ceases to grow, the secondary cell wall is formed. The secondary cell wall is composed of several lamellae, within which cellulose, a β-1,4-glucose polymer, forms microfibrils that are organized in parallel. Each lamella has a different microfibril orientation. The microfibrils are usually embedded in a matrix of hemicellulose and lignin, forming intimate interactions between the polymers (Tomme *et al.*, 1995). In secondary cell walls, the hemicellulose mainly consists of xylans (β-1,4,-xylose backbone) substituted with a variety of side chains, mannans (β-1,4,-mannose backbone), and glucomannans (mannan with glucose substitutions). Lignin is a highly branched polymer generated by

condensation of aromatic alcohols, coupled to hemicellulose by ester bonds. Only a few organisms are able to degrade lignin by a process involving radical oxidation by peroxidases (Béguin *et al.*, 1994).

The complexity of the natural substrate, i.e. the plant cell wall, led to enzymatic degradation studies using isolated substrates, such as celluloses, xylans and mannans. Since cellulose is the most abundant polysaccharide in the plant cell wall, cellulose degradation has been studied extensively in many microorganisms. These taxonomically and ecologically very diverse cellulolytic microorganisms generally secrete a variety of cellulases and hemicellulases, which can be divided into two broad categories: non-complexed and complexed systems (Gilbert and Hazelwood, 1993). The non-complexed system, also called a non-aggregating system, is characteristic of aerobic fungi and bacteria: cellulases and related polysaccharidases are secreted into the culture medium without forming aggregates. The secreted enzyme in complexed systems, typically produced by anaerobic microorganisms, are aggregated into multienzyme complexes, which can be either cell-bound or released into the culture medium (Béguin and Lemaire, 1996). This distinction between cellulase systems of aerobes and anaerobes applies generally, but is not absolute.

1.4.1 Complexed systems

Secreted cellulases and hemicellulases of some anaerobic microorganisms can associate into large, multienzyme complexes that are very efficient at hydrolyzing crystalline cellulose. These multienzyme complexes are called cellulosomes. One of the best studied organisms producing such cellulosomes is the anaerobic thermophile, *Clostridium*

thermocellum. Up to several hundred cellulosomes form protuberances on the outermost layer of the cell envelope. The cellulosomes have both cellulolytic and xylanolytic activities. *C. thermocellum*, however, is unable to grow on the degradation products of xylanolysis (Bayer *et al.*, 1996; Béguin and Lemaire, 1996).

Cellulosomes are initally cell-bound, but can then detach from the cells without significant loss in enzymatic activity. Their additional ability to bind to cellulose results in adhesion of the cells to the substrate. The binding to cellulose is mediated by a cellulose-binding domain (CBD; type 3), which is part of the scaffoldin protein, CipA. CipA has nine cohesin domains, which are binding sites for the dockerin domains present in most cellulolytic *C. thermocellum* enzymes. Dockerin domains are found in at least nine endo-glucanases, one xylanase, one lichenase and one cellobiohydrolase. The scaffoldin protein CipA itself possesses a C-terminal dockerin domain (type II), which fails to interact with its own cohesin domains (Bayer *et al.*, 1994). It binds to a specific set of cohesin domains (type II), which are found in several cell-bound polypeptides (e.g. OlpB and SdbA). These anchoring proteins have, furthermore, a triply repeated sequence, called the SLH domain. This domain has homology to domains found in S-layer proteins, which mediate non-covalent binding of the proteins to the peptidoglycan, or to other cell envelop components (Lemaire *et al.*, 1998; Egelseer *et al.*, 1998). This binding relay results ultimately in binding of the cellulolytic enzymes to the substrate and to the *C. thermocellum* cell wall.

The fungal cellulolytic system from *Pyromyces* sp. is also aggregated in a cellulose binding complex, which is comparable to the *C. thermocellum* complex. It comprises at least 10 different polypeptides, including cellulase, xylanase and mannanase activities (Ali *et al.*, 1995).

1.4.2 Non-complexed systems:

A well studied, non-complexed system is the plant cell wall-degrading system from *Trichoderma reesei*. This soft-rot, aerobic fungus secretes two major cellobiohydrolases, CBHI and CBHII (Teeri *et al.*, 1987), at least four endoglucanases, EGI, EGII, EGIII and EGV (Saloheimo *et al.*, 1994), two xylanases, XynI and XynII (Törrönen *et al.*, 1995), at least two mannanases (Stålbrand *et al.*, 1995), one β-glucosidase and one β-mannosidase (Stålbrand personal communication).

Synergistic interactions of isolated components from non-complexed systems have been demonstrated in several independent studies. Commonly, the combined activity on crystalline, but not soluble, cellulose was greater than the sum of their individual activities (Tomme *et al.*, 1995). Synergism was found not only between endo- and exo-acting glucanases (endo-exo synergism) but was also reported to occur between two exo-acting cellobiohydrolases, i.e. CBHI and CBHII from *T. reesei* (Nidetzki *et al.*, 1994).

Cellulomonas fimi, a gram positive, mesophilic soil bacterium is another producer of a well studied, non-complexed cellulolytic-hemicellulolytic enzyme system. Since this organism is being investigated in the present study, its system will be described in detail (Section 1.5).

1.5 The cellulolytic-hemicellulolytic enzyme system from *Cellulomonas fimi*

Cellulomonas fimi, a member of the actinomycete family *Cellulomonadaceae* (Stackerbrandt and Prauser, 1991) can use a variety of polysaccharides as carbon sources. *C. fimi* secretes several glucanases and xylanases into the culture medium (Figure 1.3). These

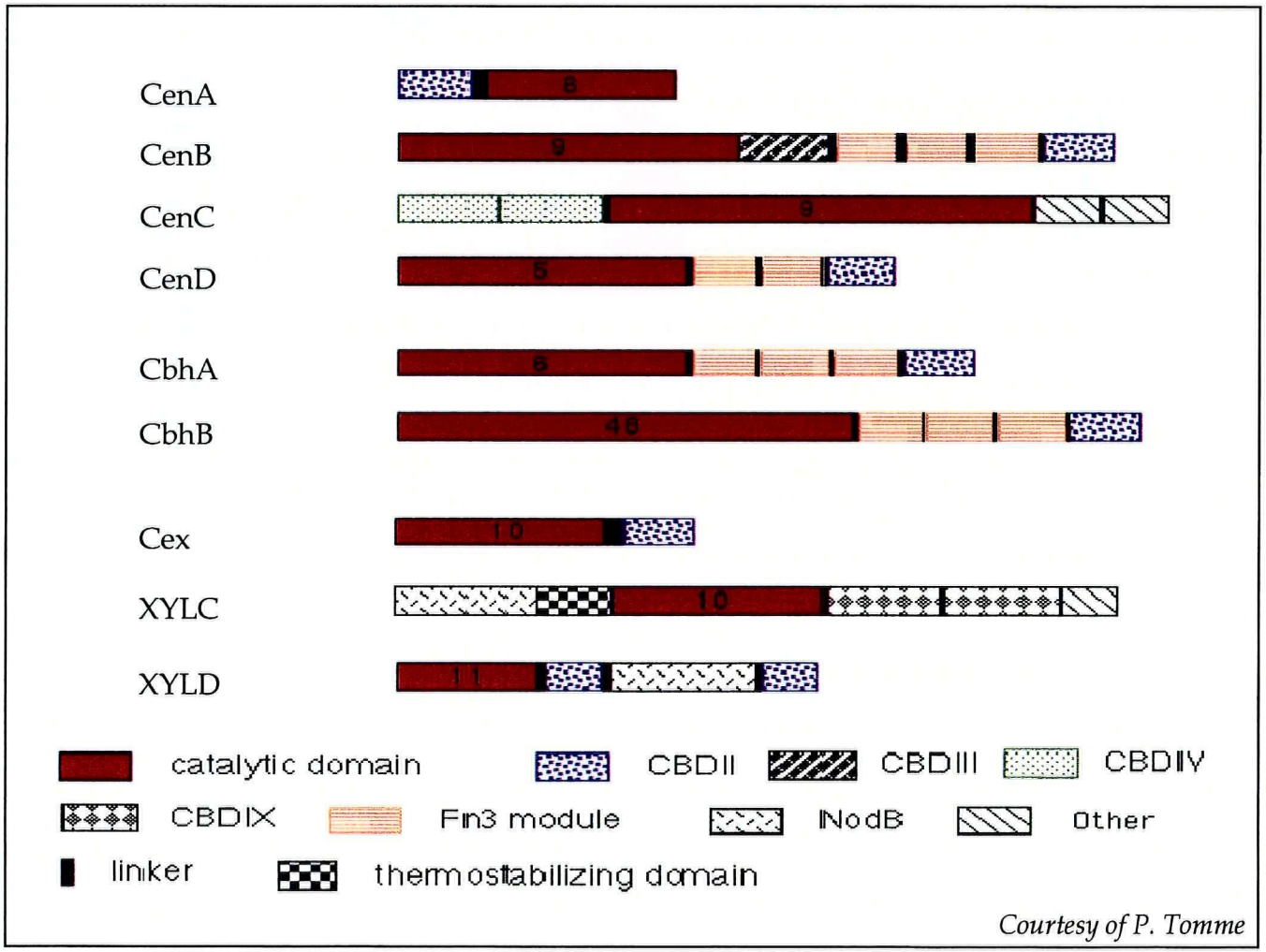


Figure 1.3: Schematic representation of the characterized cellulases and xylanases secreted by *Cellulomonas fimi*.

enzymes all have a modular architecture, composed of two or more domains. The domains are usually separated by linker sequences, such as the proline-threonine rich linker regions found in CenA and Cex (Wong *et al.*, 1986; O'Neill *et al.*, 1986; Gilkes *et al.*, 1991).

Four endoglucanases, CenA, CenB (Owolabi *et al.*, 1988), CenC (Coutinho *et al.*, 1991) and CenD (Meinke *et al.*, 1993), two cellobiohydrolases, CbhA (Meinke *et al.*, 1994) and CbhB (Shen *et al.*, 1995), one mixed function β-1,4-xylanase/β-1,4,-glucanase (Cex) and two xylanases, XYLC (Clarke *et al.*, 1996) and XYLD (Millward-Sadler *et al.*, 1994) have been analyzed to date. Both the cellulases and the xylanases from *C. fimi* have one or more cellulose binding domains (CBD) that bind to soluble and/or insoluble cellulose. Fibronectin type 3 repeats (Fn3) are domains found in CenB, CenD and in the two cellobiohydrolases, CbhA and CbhB. Their function is as yet unknown but it is speculated that they might be involved in protein-protein interactions or might just be used as spatial separation between the catalytic domain and the CBD (Tomme *et al.*, 1995). NodB is a domain which is homologous to the N-acetylglucosamine deacetylating enzyme NodB from *Rhizobium*. The NodB homologues are prevalent among xylanases from different soil bacteria; their function, however, is as yet unknown (Millward-Saddler *et al.*, 1995). The unusual high temperature optimum of the xylanase XYLC (60^0 C), might be explained by the presence of a thermostabilizing domain, adjacent to the catalytic domain (Clarke *et al.*, 1996).

Endo-glucanases and exo-glucanases interact synergistically to optimize cellulose degradation (endo-exo synergy). The endo-glucanases attack the cellulose at disordered, or amorphous regions and thereby provide new sites for the attack of exo-acting enzymes. Differences in initial attack were seen for the four *C. fimi* endo-glucanases (Kleyman-Leyer *et al.*, 1994; Tomme *et al.*, 1996). Like the *T.reesei* cellobiohydrolase CBHI, *C. fimi*

cellobiohydrolase B (CbhB) attacks the β-1,4-glucan chain from the reducing end, whereas CbhA and the corresponding *T.reesei* CBHII attacked the substrate from the non-reducing end (Gilkes *et al.*, 1997; Nidetzky *et al.*, 1994). The major end product from the concerted cellulose degradation by secreted *C. fimi* endo- and exo-cellulases is cellobiose, which once imported into the cells can be further processed into glucose (Wakarchuk *et al.*, 1984, Kim and Pack, 1989).

The biological significance of the secretion of at least four endo-glucanases by *Cellulomonas fimi* is currently unclear. It is possible that the enzymes are more versatile than assumed and that the classification into exo- and endo- acting enzymes is too simplistic, as was shown for CenC. CenC, which was classified as an endo-glucanase, has not only endo- but also exo-activity on CM-cellulose. It is able to hydrolyze linkages within the β-1,4-glucan chains, and then to move processively along the chain, releasing cellobiose. This process was classified as a semiprocessive mechanism (Tomme *et al.*, 1996).

One way to study the role of each cellulase in the biological process of cellulose, or even plant cell wall, degradation would be a genetic approach, i.e. regulation and knockout studies. This approach, however, was obstructed by the inability to manipulate *Cellulomonas fimi* genetically (results not shown).

1.6 Overall objectives

The overall objective of this study was to analyze the ability of *Cellulomonas fimi* to degrade mannan, an important hemicellulose component. It was of scientific interest to analyze the mannan degrading components in order to compare the mannanolytic system to the cellulolytic system and to find out whether the *C. fimi* cellulolytic system is a blueprint for polysaccharide degradation in this organism.

The project was also driven by application-based objectives. The finding of novel carbohydrate binding domains in secreted enzymes might be a very valuable contribution for the study of protein-carbohydrate interactions and the finding of an enzyme with transglycosylation activity might provide a good candidate for the glycosynthase approach to enzymatic synthesis of glycosidic linkages.

2 Materials and Methods

2.1 Buffers, enzymes and chemicals

Buffers and solutions used in this study were generally prepared as described by Sambrook *et al.* (1989). All buffers and solutions were sterilized by autoclaving or filtration. Locust bean gum (LBG) and carboxymethylcellulose were purchased from Sigma, ivory nut mannan, azo-carob galactomannan and mannooligosaccharides were obtained from Megazyme (Australia) and birchwood-xylan from Roth (Germany). Zeocin™ was purchased from Invitrogen (USA), Anti His_6 tag mouse antibodies from Dianova, and anti-mouse IgG-horse radish peroxidase conjugate from Dako Diagnostics (Canada). Potassium was generally used as counterion in phosphate buffer.

2-deoxy-2-fluoro-β-D-mannosyl fluoride (2FManβF) and α-mannosyl fluoride were gifts from Lloyd Mackenzie, Dept. of Chemistry, UBC. CBD_{CenD} was a gift from Al Boraston (Cellulase lab, UBC), and *C. fimi* protease a gift from Emily Kwan (Cellulase lab, UBC).

2.2 Bacterial strains, plasmids and phages

The bacterial strains, plasmids and phages used in this study are described in Tables 2.1-2.5. Bacterial stocks were maintained at -70° C in LB medium containing 15 % glycerol. Plasmid DNA was stored in water at -20° C. Phages were stored in TYP medium at 4° C or -70° C.

Table 2.1.: Bacterial strains

Bacterial strain	*Genotype*	*Reference or source*
Cellulomonas fimi	wild type	ATCC 484
<u>*Escherichia coli* strains:</u>		
BL21(DE3)	F- *ompT hsdSB*(r_B-m_B-)*gal dcm* (DE3)	Grodberg *et al.*, 1988
DH5α	F- *endA1, hsdR17* (*rk, mk*+), *supE44, thi-1, recA1,* (*argF-laczya*)*U169, ϕlacZ M15*	Hanahan, 1983
JM101	*supE thi* Δ(*lac-proAB*)[F′ *traΔ36 proABlacIqZ*ΔM15]	Yanisch-Perrom *et al.*, 1985
XL1-Blue	*endA1, hsdR17, supE44, thi-1, recA1, gyrA96, relA1, lac,* [F′, *proAB, lacIqZΔM15, Tn10*(tet^R)	Jerseth *et al.*, 1992
XLOLR/SOLR	e14-(*mcr*A)Δ(*mcr*CB-*hsd*SMR-*mrr*)171 *sbc*C *rec*B *rec*J *umu*C::Tn5(Kan^R), *uvr*C *lac gyr*A96 *rel*A1*thi*1 *end*A1 λ^R[F′ *pro*AB *lac*I^q Δm15] Su-	Stratagene

Table 2.2: Bacteriophages

Bacteriophages:	*description:*	*Source:*
ExAssist	f1 derived helper phage used for pBluescript excision from λ ZapII	Stratagene
λ ZapII-*C. fimi* genomic library	lambda *C. fimi* expression library	Stratagene

Table 2.3: Plasmids used for cloning and protein expression

Plasmid name:	*Description:*	*Source:*
pZErO™	cloning vector; Zeocin^R	Invitrogen
pBluescript SK and KS	cloning vector; Amp^R	Stratagene
pET27b and pET28a	expression vector; Kan^R	Novagen

Table 2.4: Plasmids used for genome mapping

plasmid name:	*description*	*reference:*
pTAL1	pTZ18R; *cenB*; Amp^R	Meinke *et al.*, 1991
pDAM1-5	pTZ18R; *cenD*; Amp^R	Meinke *et al.*, 1993
pALM1.1	pTZ18R; *cbhA*; Amp^R	Meinke *et al.*, 1994
pTC1n	pTug07K3; *cenC*; Kan^R	Tomme *et al.*, 1996
pTZ18R-cex	pTZ18R; *cex*; Amp^R	MacLeod *et al.*, 1994
pTZ18R-1.6cenA	pTZ18R; *cenA*; Amp^R	Damude *et al.*, 1996
pTugSH3	pTugEO7K3; *cbhB*; Kan^R	Shen *et al.*, 1995
pET28aMad	pET28a; *man2A*; Kan^R	this study
pET27Man26A	pET27b; *man26A*; Kan^R	this study

Table 2.5: Plasmids used for sequencing of *man26A*

plasmid name:	*description:*
pCMan2	pBluescript (SK); 4.3 kbp *man26A* genomic *C. fimi* DNA insert (at *Eco*R I); Amp^R.
pCMan4	pBluescript (SK); 6.3 kbp *man26A* genomic *C. fimi* DNA insert (at *Eco*R I); Amp^R.
pCMan30	pBluescript (SK); 3.0 kbp *man26A* genomic *C. fimi* DNA insert (at *Eco*R I); Amp^R.
pBSManPst0.9	pCMan2 0.9 kbp *Pst* I fragment cloned into pBluescript (KS); Amp^R.
pBSManPst1	pCMan2 1 kbp *Pst* I fragment cloned into pBluescript (KS); Amp^R.
pBSManPst1.9	pCMan2 1.9 kbp *Pst* I fragment cloned into pBluescript (KS); Amp^R.
pBSManXho0.9	pCMan2 0.9 kbp *Xho* I fragment cloned into pBluescript (KS); Amp^R.
pCMan30ΔXho	pCMan30 with deletion of all *Xho* I fragments; Amp^R
pCManBam/Xho0.9	pCMan2 0.9 kbp *Bam*H I-*Xho* I fragment cloned into pBluescript (KS); Amp^R.
pCManΔBam	pCMan2 with deletion of all *Bam*H I fragments; Amp^R
pCManΔKpn	pCMan2 with deletion of all *Kpn* I fragments; Amp^R

Table 2.6: Plasmids used for sequencing of *man2A*

plasmid name:	*description:*
pCMadI	pBluescript (SK); 4.3 kbp *man2A* genomic *C. fimi* DNA insert (at *Eco*R I); Amp^R
pCMadII	pBluescript (SK); 4.3 kbp *man2A* genomic *C. fimi* DNA insert (at *Eco*R I); Amp^R
pBSMadPst/Bam	pCMadI 1.1 kbp *Pst*I-*Xho*I fragment cloned into pBluescript (KS); Amp^R.
pCMadΔBam	pCMadI with deletion of *Bam*H I fragment; Amp^R
pCMadΔPst	pCMan2 with deletion of all *Pst* I fragments; Amp^R
pBSMadXho/Pst	pCMadI 1.1 kbp *Pst* I-*Xho* I fragment cloned into pBluescript (KS); Amp^R.
pCMadΔBamΔXho	pCMadΔBam with deletion of *Xho* I fragment

2.3 Media and growth conditions

E. coli strains were grown in LB medium (10 g tryptone, 5 g yeast extract, 10 g NaCl per liter) at 37° C, or for gene expression at RT in TYP (16 g tryptone, 16 g yeast extract, 5 g NaCl and 2.5 g K_2HPO_4 per liter; pH 7.0) supplemented with either 50 μg kanamycin or 100 μg ampicillin mL^{-1}. pZErO™ clones and subclones thereof were grown in LB low salt (LB with only 0.5% NaCl) and 50 μg/mL Zeocin.

C. fimi was grown in basal salt medium (1 g $NaNO_3$, 1 g K_2HPO_4, 0.5 g KCl, 0.5 g $MgSO_4.7H_2O$, 0.5 g yeast extract per liter; pH 7.2) supplemented with 0.2 % (w/v) carbon source (Stewart *et al.*, 1976). Kanamycin was added at 50 μg/mL. *C. fimi* cultures were grown at 30° C with agitation at 220 rpm. Solid media contained 1.5 % agar.

2.4 Oligodeoxyribonucleotide primers

Primers used for sequencing and PCR reactions are summarized in Tables 2.6-2.8. The primer sequences are indicated in the 5′ end to 3′ end orientation.

Table 2.7: Primers used for *man26A* sequencing and PCR

Primer name:	*Primer Sequence:*
Man1	GTC TTC GGC TGG GAC ACG
Man2	GTGTCGAGGCCGAACACG
Man3	GAT CAA GGC CGA CCC CGT
Man 31	GGA GCT CTA CCG GTT CAC
Man 32	CGT TCT TGA CGT CGG AGT
Man 33	GAC CTG TTC CGT CCT GAC
Man 34	CAA CGC TAC ATG GAG ACG
Man 35	GTG CTG TCG TAG GTG TCG
Man 37	CTC GTA GGC GAC ACC ACC
Man 41	CAA CGT CTA CGT CAA CG
Man 42	GGT ACC TGT ACC GGT TCG AC
Man 43	ACC TCT ACC TCA ACG CAG GC
Man 44	CGC GAG GTG GTA GTC GGA CC
Man 45	AGT ACG CCG AGA CGT GGA T
Man 46	CGC GGT CCG ACT ACG ACC T
Man 47	TAC AAC GTC GTG AGG TCG AG
Man 48	TGA GCA CCC CGA ACG TCG AC
Man 50 (PCR)	GTC CAG GCT AGC GCG GCG CAG CTC GAC GTC GTC (*Nhe* I-site)
Man 51 (PCR)	GGA GGT CAT ATG ACG AAC CGC AGC AGC CGT CCG (*Nde* I-site)
T3	ATT AAC CCT CAC TAA AG
T7	AAT ACG ACT CAC TAT AG

Table 2.8: Primers used for MBD_{1112} PCR

Primer name:	*Primer Sequence:*
MBD 11	AGC GCG CAG CTC GAC AAC ACC ATG GGC ACC GTC ACC GCG ACG GCG (*Nco* I-site)
MBD12	GCC GGG GTG GGC GGC GAT GGC GGC CGC GAC GAG CGA GCC CGA CGC (*Not* I-site)

Table 2.9: Primers used for *man2A* sequencing and PCR

Primer name:	*Primer Sequence:*
Mad1	GGA CCG TCA CGT TCG ACG
Mad2	TCG GTG CGC ACG CAC GTC
Mad3	TCG TTG ACG ACC GCG AGC
Mad4 (PCR)	GCG TCC ATG GTC ACC CAG GAC ATC TAC (*Nco*I-site)
Mad5	CTC GGC GTG CAC GAT CGC GG
Mad10	GAG GTC GAC ACG ACG
Mad11	CCC ACT CGT GGT GCG
Mad12	CCT TGA CGA AGA CGG
Mad13	TCG CCA CCG CTC GAC
Mad21	GAA TGA TCC CGG GTA CCT
Mad22	ATC TTG CGG ACC ATG TTG
Mad23	AAC ATG GTC CGC AAG ATG
Mad24	GTT GAC CGG AAC GTG AAC
Mad25	ATG AAG CCC CAG AGG TTC
Mad26	GTC TGG CAG GAC TTC CTC
Mad27	CCA GCA GTC GTT TGA GCT
Mad28	CGG TTC TGC TCC GAG TTC
Mad29	CGA CCG GTA GTG CTC GAT
Mad51 (PCR)	GCG AAC CGT TCA GCG GCC GCG CGG GAC TGC TGG (*Not*I-site)
MadE519A (used to be MadE622A) (PCR)	CGG CAC GCG TGA GGG TCG ACC ACG TCG TCG GCG GGC CCT GGA AGC CGA AGG CGG AGC AGA ACC
T7 and T3	see table 2.6

2.5 Recombinant DNA techniques

Recombinant DNA work was generally carried out as described by Sambrook *et al.*, 1989. DNA fragments were isolated from agarose gels and purified using the Qiaex II Gel Extraction Kit (Qiagen). Restriction endonucleases were used as recommended by the suppliers (New England Biolabs, GIBCO BRL or Böhringer Mannheim).

Ligations were performed with 100-500 ng of total DNA at insert to vector molar ratios of approximately 10:1 in 20 µL reaction volume. 1 U T4 DNA ligase (GIBCO BRL) was

used per reaction. Ligations were incubated at 23^0 C for approximately 2 h. Reaction mixtures were desalted by butanol precipitation (Thomas *et al.*, 1994). *E. coli* cells were transformed by either electroporation (GenePulser, BioRad) or by heat shock of cells prepared by the quick chemical competent cell protocol in TSS buffer (10 g bacto tryptone, 5 g yeast extract, 5 g NaCl, 100 g PEG 4000, 50 mL DMSO, 50 mL 1M $MgCl_2$ per liter; pH 6.5) (Dower, 1987).

2.5.1 Polymerase chain reaction (PCR)

PCR conditions using *C. fimi* DNA templates were as follows: 100 µL reaction volume with 10-100 ng DNA template, 40 pmol of each primer, recommended polymerase buffer, 10 % DMSO, 0.2 mM 2'-deoxynucleoside 5'triphosphates and 1 unit of either Vent polymerase (New England Biolabs), PWO or HiFi polymerase (both Böhringer Mannheim). DNA was amplified by repeating cycles of denaturation (94^0 C, 30 s), annealing (64-67^0 C, 30 s) and elongation steps (72^0 C, 45-90 s) 30 times.

2.5.2 Primer synthesis and DNA sequencing

Oligodeoxyribonucleotide primers were synthesized by the UBC Nucleic acid and Protein synthesis unit (NAPS) with an Applied Biosystem DNA synthesizer and purified by precipitation with n-butanol (Sawadogo and VanDyke, 1991).

DNA sequencing was performed by NAPS using an Applied Biosystem DNA sequencer Model 377 (Perkin-Elmer). Sequencing reactions were performed according to the AmpliTaq dye termination cycle sequencing protocol with the addition of 5 % DMSO.

2.6 Detection of enzyme acitvity

2.6.1 Plate assays

Mannanase activity was assayed on 2 % agar plates containing 0.5 % LBG in 50 mM buffer (pH 7.0) (Stålbrand *et al.*, 1993). Ten to 20 µL of sample were deposited onto the plate, which then was incubated for 12 to 16 h at 37^0 C. Hydrolysis of the polymer was visualized by flooding the plate with Congo red solution (2 mg/mL) followed by two washing steps with 1 M NaCl (Teather and Wood, 1982).

E. coli clones producing mannanase activity were detected on LB plates (Section 2.2) containing 0.2 % azo-carob galactomannan. Azo-carob galactomannan is a galactomannan dyed with Remazol Brilliant Blue (McCleary, 1978). After 12 to 16 h at 37^0 C zones of mannan hydrolysis (halos) were detectable (Braithwaite *et al.*, 1995).

To detect β-mannosidase or α-galactosidase activity, 0.1 mM to 0.5 mM 4-methylumbelliferyl β-mannoside (MUβMan) or 4-methylumbelliferyl α-galactoside (MUαGal), respectively were added to LB plates. The release of fluorescent methylumbelliferol was detected under UV light (Wakarchuk *et al.*, 1984).

2.6.2 Zymograms

Zymograms for the detection of polypeptides with mannanase activity were prepared as 10 % SDS-polyacrylamide gels (SDS- PAGE) (Laemmli, 1970) supplemented with 0.5 % azo-carob galactomannan (adapted from Braithwaite *et al.*, 1995). Equal volumes of 2 x non-reducing loading buffer were added to samples: 125 mM Tris (pH 6.8), 5 % SDS, 0.3 %

bromophenol blue, 50 % glycerol. Samples (18 μL/lane) were separated at constant voltage (200 V) using a Bio-Rad Mini-PROTEAN™ apparatus. To detect enzyme activity, gels were rinsed and then incubated in 200 mL 50 mM phosphate buffer (pH 7.0) at 37° C until halos were visible (2 to 16 h).

To detect β-mannosidase or α-galactosidase activity, non-reducing SDS-PAGE gels were rinsed in 50 mM phosphate buffer pH 7.0 and incubated in 1 mM MUβMan and 1 mM MUαGal, respectively for approximately 5 min. Hydrolysis was observed under UV (Wakarchuk *et al.*, 1984). To detect β-glucosidase activity, gels were incubated in 0.5 mM 5-bromo-4-chloro-indolyl β-D-glucoside (X-Glu) (Kohchi *et al.*, 1996).

2.7 Library sceening

The *C. fimi* lambda ZapII library prepared by Stratagene was screened as recommended by the supplier with a few changes: 150 mm diameter petri dishes were filled with 30 ml NZY medium (5 g NaCl, 2 g $MgSO_4.7H_2O$, 5 g yeast extract, 15 g agar and 10 g casamino acids per liter). On top of the solidified NZY medium 10 mL of 0.4 % Azo-carob galactomannan containing 1.5 % agar solution was added. The plates were dried for several hours before the overlay of phage and host cells mixed in 6.5 mL NZY top agar (0.7 % agar, 0.1 to 0.5 mM IPTG) was poured on. Mannanase activity was detected as halos, which were visible after incubation periods of 16 to 24 h at 37° C (adapted from Braithwaite *et al.*, 1995). After secondary and tertiary screenings, mannanase positive phagemids (pBluescript SK) containing genomic DNA inserts were excised *in vivo* from lambda DNA, recircularized and transduced into *E. coli* XLOLR host cells (Shen *et al.*, 1995).

The excised form of the entire genomic library was used to screen for β-mannosidase producing *E. coli* clones. Only replica plates containing 0.5 mM MUβMan were induced with 0.3 mM IPTG.

2.8 Production and purification of recombinant proteins

Genes were expressed from pET vectors (Novagen) in *E. coli* BL21(DE3) host cells. Cells were grown at RT and 150 rpm in TYP supplemented with 50 μg/mL kanamycin to mid exponential phase. Protein synthesis was induced with 0.2 to 0.4 mM IPTG at RT for 24 to 36 h. Cultures were centrifuged at 5000 rpm for 10 min. Either the supernatant, as for the purification of secreted Man26A, or the cells for the purification of Man2A WT, Man2A E519A, MBD_{1112}, or Man26A, were further processed. Cells from 1 or 2 liter cultures were resuspended in 1x Binding Buffer (5 mM imidazole, 500 mM NaCl, 20 mM Tris-HCl pH 7.9) and ruptured by passing them 3 times through the French pressure cell. Insoluble cell debris was sedimented for 30 min at 15 k rpm and the cell-free extract was loaded onto a 5-10 mL His-bind® metal chelating affinity column (MCAC) at a flow rate of 0.5 mL/min. Column chromatography was performed as recommended by the supplier (Novagen).

Culture supernatant was concentrated to 50 mL and buffer exchanged with Binding Buffer using an Ultrasette™ tagential flow concentrator with a 1 kDa cut off (Filtron) . The concentrate was loaded onto a 15 mL His-bind column (Novagen) at a flow rate of 0.5 mL/min.

Columns were washed with 3 to 6 column volumes of Binding Buffer at 1 mL/min. Proteins were eluted by step wise increases in the concentration of immidazole from 0 to 0.5 M imidazole elution buffer (0.5 M NaCl, 20 mM Tris-HCl pH 7.9) at a flow rate of 1

mL/min. The desired proteins were eluted with 50 mM to 120 mM immidazole, with the exception of MBD_{1112} which eluted at 20 mM. The appropriate fractions were pooled, buffer exchanged and concentrated by diafiltration through an Amicon PM10 membrane. The MBD_{1112} fraction was concentrated with a Microsep™ Microconcentrator (3 kDa cutoff; Filtron). Protein purity was estimated by Coomassie-stained SDS-PAGE gels.

2.9 Partial purification of Man2A from *C. fimi*

C. fimi cells from a 2 L LBG culture were collected by centrifugation (10 min at 5 k rpm). The resuspended cells were ruptured by passing them 4 times through the French pressure cell. Debris and unbroken cells were removed by centrifugation at 40 k rpm for 1 h. The supernatant was buffer exchanged with 20 mM Bis-Tris HCl pH 5.8. The sample was fractionated by anion-exchange chromatography (EconoQ, Bio-Rad) with a linear gradient of NaCl (0 mM to 700 mM) in 20 mM Bis-Tris HCl pH 5.8 (20 column volumes). Fractions were tested for β-mannosidase activity with 200 μM MUβMan. The fractions with the highest activity were pooled, concentrated and prepared for N-terminal sequence analysis (Section 2.11).

2.10 N-terminal amino acid sequencing

Protein samples were subjected to SDS-PAGE (Laemmli, 1970) then electroblotted onto polyvinylidene difluoride (PVDF) membranes (Millipore). Protein bands were visualized by Coomassie staining (0.025 % Coomassie in 40 % MeOH) and the bands of interest were excised and submitted for N-terminal sequence analysis by the NAPS unit at University of British Columbia or by the Protein Microchemistry Facility, University of

Victoria. The N-terminal sequence was determined by automated Edman degradation, using a Perkin Elmer Applied Biosystems 476A gas-phase sequenator (Matsudaira, 1990).

2.11 Determination of protein concentration

Protein concentrations were determined from A_{280} using the Beer-Lambert law (Absorption (A)= extinction coefficient (ε) x path length (l) x molar concentration (c)). The molar extinction coefficient (ε) was calculated according to equation E 2.1 (Pace *et al.*, 1995).

E 2.1: $$\varepsilon(280)(M^{-1}.cm^{-1}) = (\#Trp)(5{,}500) + (\#\ Tyr)(1{,}490) + (\#Cys)(125)$$

2.12 Size exclusion chromatography of Man2A

Fractogel TSK HW-55S from EM-Science (Merck) was used as matrix for size exclusion chromatography of Man2A (100 μg) with a 400 x 10 mm column. A buffer of 50 mM phosphate pH 7.0 with 100 mM NaCl was used (flow rate of 0.5 mL/min) as mobile phase. The void volume was determined with Blue Dextran, and the column was equilibrated with ferritin (440 kDa), catalase (233 kDa), aldolase (158 kDa) and ovalbumin (50 kDa) (Gehrmann *et al.*, 1994). The sample volume was 50 μL.

2.13 Proteolysis of Man26A

Man26A was treated with *C. fimi* protease at 37° C (1 U protease/0.55 nmol Man26A in 20 mM Tris/HCl pH 7.5 (Meinke *et al.*, 1992)). Samples were removed and immediately

added to SDS loading buffer and heated at 100^0 C for 2 min. Samples were analyzed by 12 % SDS-PAGE (Laemmli, 1970) with approximately 6 μg protein loaded per lane.

2.14 Matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOFMS)

Protein samples (0.5 to 3 mg/mL) were desalted by drop dialysis for 12 to 16 h by depositing 5 to 10 μL drops on dialysis membrane disks (MF, pore size 0.025 μm, Millipore) floating on water (500 mL). One μL of dialyzed sample was loaded onto the sample holder and dried for 5 min. The sample was overlaid with 1 μl matrix (supersaturated sinapinic acid solution in 70 % acetonitrile, 0.1 % trifluoroacetate) and dried for 5 min (Kallweit *et al.*, 1996). Mass spectra were recorded on a Mass Phoresis instrument (Ciphergen Inc.).

2.15 Localization of Man26A in *C. fimi* cultures

One hour before analyzing samples from a 6 day old *C. fimi* culture, grown on minimal medium plus 0.2 % LBG, chloramphenicol was added (40 μg/mL). Mannanase activity in the culture, supernatant, and in cells (washed 3 times with 50 mM phosphate buffer pH 7.0) was tested on azo-carob galactomannan (McCleary, 1978). Reactions with 75 μL azo-carob galactomannan (2 %), 5 μL azide(10 %), 6 μL chloramphenicol (20 mg/mL) and 210 μL sample were incubated at 37^0 C for 16 h. Reactions were stopped by adding 750 μL EtOH (95 %) and the precipitated mannan was removed by centifugation for 2 min at 13 k rpm. The release of EtOH-soluble RBB labeled mannooligosaccharides in the supernatant was determined from A_{590}. Datapoints were collected in triplicates.

2.16 Affinity gel electrophoresis (AFGE) and Western blotting

The binding of proteins to soluble substrates was analyzed by affinity gel electrophoresis (Takeo *et al.*, 1972; Takeo, 1984). 7.5 % separating polyacrylamide gels (Mini-Protean II system (Bio-Rad); 0.75 mm spacers) were supplemented with soluble galactomannan, either locust bean gum or azo-carob galactomannan, at final concentrations from 0 to 1 %. Protein samples in native loading buffer (20 % glycerol, 0.1 % Bromopheol blue, 125 mM Tris/HCl pH 8.8) were loaded on the gels. Five μg/lane were used for detection by Coomassie staining and 70 ng/lane for Western blotting. Samples were electrophoresed at 150 V for 2 h at 4^0 C. Protein bands were visualized either by staining with Coomassie blue, by zymograms (for azo-carob galactomannan gels only) or by Western blotting (Towbin *et al.*, 1979). For Western blotting, the protein bands were transferred electrophoretically to PVDF membranes (Millipore) and probed with mouse anti-His_6 antibodies (Dianova; 1/250 dilution) using anti-mouse IgG-horse radish peroxidase conjugate (Dako Diagnostics, Canada) as secondary antibody (1/5000 dilution). ECL™ detection reagents were used as recommended by the supplier (Amersham).

The relative mobility, which is the ratio of migration distance of the protein to that of the reference protein, was measured at various substrate concentrations and the values used to calculate dissociation constants according to equation E 4.1 (Section 4.2.3). Reference

proteins were either Man2A or acetylated BSA (New England Biolabs). CBD_{CenD} was used in Western blots to mark the top of the separating gels.

Binding to monomeric sugars was tested by competition affinity gel electrophoresis (Takeo, 1984) through 1.25×10^{-2} % LBG affinity gels (*vide supra*) containing either mannose or galactose (1.8 %). Gels were electrophoresed as described and proteins with a His_6 tag were detected by Western blotting (*vide supra*).

The concentration of LBG was calculated by measuring total sugar concentration using the phenol-sulfuric acid method (Chaplin, 1986) and reducing sugar concentration with the the *p*-hydroxybenzoic acid hydrazide method (Lever, 1973). Both concentrations were calculated as D-galactose and D-mannose equivalents assuming the presence of 5 mannose residues for each galactose residue, because these sugars differ in their responses in the assay (Lever, 1977).

2.17 Genomic Mapping

2.17.1 Preparation of genomic DNA

The protocol for preparation of *C. fimi* genomic DNA was adapted from Birren and Lai (1993). *C. fimi* cultures were grown in LB lowsalt (0.5 g NaCl per liter) supplemented with 0.5 % glucose and 50 μg kanamycin mL^{-1} to mid log phase (A_{600}= 0.8 to 1.5). Two hours before harvesting, 3 μg Penicillin G $mL^{-1}$was added. Cells were harvested by centrifugation (5 k rpm, 10 min) and washed in Wa1 (10 mM Tris/HCl, pH 7.2, 100 mM EDTA, 200 mM NaCl). The washed cells were resuspended in Wa1 (1/20 of the culture volume) and warmed to 37^0 C. The suspension was mixed with an equal volume of molten 1.5 % low melting preparative grade agarose (Bio-Rad) at 40^0 C. The cell suspension was transferred to

1 cc syringes. After the agarose solidified, the tips of the syringes were cut off and little agarose-embedded cell blocks (2 to 5 mm) were cut with sterile razor blades. Twenty to thirty blocks were incubated in sterile LS1 (10 mM Tris/HCl pH 7.5, 50 mM NaCl, 100 mM EDTA, 0.2 % Na-Deoxycholate, 0.5 % NaSarcosyl, 3 mg lysozyme mL^{-1}) for 16 h at 37^0 C with gentle agitation. The agarose blocks were transfered to 30 mL sterile DB1 (0.5 M EDTA, 1 % NaLaurosylsarkosyl, 0.5 mg/mL Proteinase K) and incubated for 48 h at 53^0 C.

Blocks were extensively dialysed with restriction enzyme buffer prior to DNA digestions with restriction endonucleases. Each block was digested in 100 µL of the appropriate restriction buffer containing 20 to 50 U restriction endonuclease for 16 h at the temperature recommended by the supplier.

2.17.2 Pulsed field gel electrophoresis

Genomic DNA blocks were embedded in 1.2 % agarose gels (high strength analytical ultra pure DNA grade agarose (Bio-Rad), in TAE buffer). DNA size standards were *Saccharomyces cerevisiae* chromosmes (NEB) and *Hin*d III digested lambda DNA (BRL). Pulsed field gel electrophoresis was performed on a 2015 Pulsaphor apparatus (LKB; Pharmacia). The gel was submerged in sterile 1 x TAE buffer and equilibrated to 12^0 C. Switch intervals of 60 s for 9 h, 80 s for 9 h, 100 s for 9 h and 120 s for 9 h were applied with a constant voltage of 160 V. DNA bands were visualized under UV light after ethidium bromide staining (50 µg/mL in H_2O).

2.17.3 Southern blotting of pulsed field gels

DNA bands from PFGE gels were transfered onto Hybond™ N+ membranes (Amersham) by capillary alkaline blotting (Sambrook *et al.*, 1989). Genes were detected by using the Fluorescein Gene Images™ system (Amersham) according to the recommendation of the supplier. DNA probes were labeled either by the random primer method, or by PCR (Feinberg, *et al.*, 1983). PCR conditions were as described in Section 2.5.1, whereby the deoxynucleotides were supplemented with fluorescein-labeled dUTP. The probes were: *man26A*, 1.2 kbp *Sac* II fragment; *man2A*, 650 bp PCR fragment (Mad26, Mad29: PCR primers); *cex*, 511 bp *Nru* I fragment; *cenB*, 811 bp *Nru* I fragment; *cenD*, 711 bp *Pvu* I fragment, *cbhB*: 750 *Sst* I fragment; *xynD*, PCR product 101; the *cenA* and *cenC* specific probes were the entire plasmids (Table 2.4).

The hybridization temperature applied ranged from 63^0 C to 72^0 C and the wash stringency was 1 x SSC, 0.1 % (w/v) SDS and 0.5 x SSC, 0.1 % SDS (20 x SSC: 3 M NaCl, 0.3 M Na_3citrate, pH 7.0), at the same temperatures as the hybridization.

2.18 Enzymology

2.18.1 pH and temperature optima for Man2A

The pH optimum for Man2A was determined by measuring PNPM hydrolysis at A_{400}. 0.3 pmol Man2A was incubated in 500 µL of 200 µM PNPM, 0.1 % BSA, 150 mM NaCl and 100 mM buffer. Potassium phosphate buffers were used at pH 5.7 to 8.1. After 6 min at RT the reactions were stopped and equilibrated by adding 1 mL of 1 M glycine (pH 10.9). Initial PNPM hydrolysis rates (ΔA_{400}/min) were measured to determine the temperature optimum. 0.08 pmol Man2A was added to 1 mL reaction mixture (220 µM PNPM, 0.1 % BSA and 100 mM phosphate buffer pH 7.0). Hydrolysis rates were measured at

temperatures between 23^0 C and 60^0 C. Reaction mixtures were preincubated at the appropriate temperature for a period of 10 min prior to addition of enzyme. ΔA_{400}/min was measured 3 min after the addition of enzyme.

Thermal stability was determined by incubating Man2A at temperatures ranging from 23^0 C to 55^0 C. (0.2 nmol Man2A in 500 µL 50 mM phosphate buffer pH 7.0, 0.1 % BSA). The loss of activity over time was measured by removing samples and determining the rate of hydrolysis of PNPM at 23^0 C (*vide supra*).

2.18.2 Steady state kinetic parameters for hydrolysis of PNPM by Man2A

Steady state kinetic parameters for PNPM hydrolysis were determined by continuous measurement of the release of *p*-nitrophenol using a Hitachi U 2000 spectrophotometer with a temperature-controlled cell holder set at 37^0 C. Reaction mixtures (1 mL) contained 50 mM phosphate buffer (pH 7.0), 4.2 nM Man2A and 0.1 % BSA. Cuvettes and solutions were prewarmed at 37^0 C. The PNPM concentrations ranged from about 1/20 x K_m to about 5 x K_m. The release of *p*-nitrophenyl was measured at A_{400} and the molarity calculated according to the Beer-Lambert law, using $\varepsilon = 7280$ ($M^{-1}x\ cm^{-1}$) (Kempton and Withers, 1992). Values for v_{max}, K_m and k_{cat}/K_m were calculated from a Lineweaver Burk plot (Figure 3.24) for PNPM concentrations of 10 to 150 µM. A curve obeying the Michaelis-Menten equation was fitted to the initial slope using the computer program GraFit 3.0 (Leatherbarrow, 1992).

2.18.3 Kinetics of inactivation of Man2A

The inactivation of Man2A by 2-deoxy-2-fluoro-mannosyl-β-fluoride (2FManβF) was monitored by incubating of enzyme (8.3 µM) with various concentrations (19.5 µM to 520

μM) of 2FManβF and measuring the release of fluoride using an Orion 96-09 combination fluoride electrode. The reaction mixtures (300 μL) were incubated at 37° C in 50 mM phosphate buffer pH 7.0 containing 0.1 % BSA. Pseudo-first order rate constants at each inactivator concentration (k_{obs}) were determined by fitting each inactivation curve to a first order rate equation with GraFit 3.0 (Leatherbarrow, 1992). Values for the inactivation rate constant (k_i) and the dissociation constant for the inactivator (K_i) were determined by non-linear regression according to equation E 6.1 (Section 6.2.1.2) (Mackenzie *et al.*, 1997).

2.18.4 Reactivation of Man2A after inactivation with 2FManβF

Completely inactivated Man2A was freed from excess inactivator by exchanging the buffer in the sample with 50 mM phosphate buffer, pH 7.0, using Microsep™ Microconcentrators (10 kDa cutoff; Filtron) (Mackenzie *et al.*, 1997). Potential reactivators were screened by incubating 21 pmol of the inactivated enzyme with 2 mM of reactivator in a reaction volume of 200 μL at 23° C. The regain of activity was assayed by removing samples and measuring PNPM hydrolysis at A_{400} (45 μL reactivation mix, 500 μM PNPM, 50 mM phosphate buffer pH 7.0, 0.1 % BSA; 1 mL reactions).

The kinetics of reactivation of inactivated Man2A were determined with gentiobiose (glucosyl-β-1,6-glucoside), by incubating 8.3 μM of the inactivated enzyme with 10 mM, 15 mM, 40 mM and 80 mM gentiobiose in 50 mM phosphate buffer, pH 7.0, 0.1 % BSA in a 500 μL reaction volume at 37° C. Reactivation was monitored by removing samples (7 μL) at appropriate time intervals and assaying the regain of activity with 2 mM PNPM, 50 mM phosphate buffer pH 7.0, 0.1 % BSA, in 600 μL. The observed reactivation rate constant, k_{obs}, for each gentiobiose concentration was determined from the slope of the plots of ln (full rate

minus observed rate) *versus* time. The rate constant (k_{trans}) and dissociation constant (K_{trans}) for the reactivation were determined from a reciprocal plot of k_{obs} *versus* reactivator concentration (Mackenzie *et al.*, 1997).

2.18.5 pH and temperature optima for Man26A

Hydrolysis of $PNPM_2$ by Man26A was measured at pH 4.7 to 8.0, by incubating 0.4 nmol Man26A with 1 mM $PNPM_2$, 150 mM NaCl, 0.1 % BSA and 50 mM buffer in a total volume of 100 µL for 5 h at 32^0 C. Citrate buffer was used from pH 4.7 to pH 6.3 and phosphate buffer from pH 5.7 to pH 8.0. The reactions were stopped and the pH adjusted by addition of 500 µL 1M glycine pH 10.9. A_{400} was recorded.

$PNPM_2$ hydrolysis was analyzed at temperatures between 23^0 C and 50^0 C, by incubating 0.4 nmol Man26A with 1 mM $PNPM_2$, 150 mM NaCl, 0.1 % BSA and 50 mM citrate buffer pH 5.5 in a total volume of 100 µL for 5 h. A_{400} was measured after the addition of 500 µL 1 M glycine pH 10.9. All determinations were done in duplicate.

2.18.6 Steady state kinetic parameters for the hydrolysis of LBG by Man26A

Steady state kinetic parameters for the hydrolysis of LBG by Man26A were determined by monitoring the release of reducing sugars, as D-mannose equivalents, using the *p*-hydroxybenzoic acid hydrazide method (Lever, 1973). Reaction mixtures contained 22 nM Man26A, 0.01 to 4.5 mg LBG mL^{-1}, 50 mM citrate buffer pH 5.5 and 0.1 % BSA in a total reaction volume of 10 mL. The reactions were incubated at 37^0 C for 40 min. The

concentration of reducing ends on undigested substrate was subtracted from the concentration of reducing ends measured after hydrolysis to yield the enzymatic activity.

2.18.7 Hydrolysis of mannan and mannooligosaccharides by Man26A and Man2A

Mannooligosaccharides (1 mM), locust bean gum (0.1 %) and ivory nut mannan (0.1 %) were digested with 0.5 nmol Man2A and/or 0.5 nmol Man26A for 1 h at 37 °C in a 100 µL reaction volume. The reaction was terminated by separating the sugars from the enzyme with Microsep™ Microconcentrators (3 kDa cut off; Filtron). The mono- and oligosaccharides released by Man26A and Man2A were analyzed by fluorophore assisted carbohydrate electrophoresis (FACE®) as recommended by the supplier (Glyko). The standards were α-1,4 glucooligosaccharides and β-1,4 mannooligosaccharides. The O-linked oligosaccharide profiling gels were run at 20 mA (constant current) for approximately 60 min and the temperature was controled at 20° C. Fluorescent carbohydrate bands were visualized under UV light (320 nm) and documented with a digital camera (Kodak DC 40).

2.19 Hydrolysis of mannan by intact and proteolytically cleaved Man26A

Man26A (390 µg) was incubated in 100 µL phosphate buffer pH 7.0 with either 1.4 U *C. fimi* protease or without protease for 2 h at RT. A control reaction was set up with protease alone. Six µL samples from these reactions were transferred to 10 mL of 0.2 % locust bean gum or 0.2 % ivory nut mannan in 50 mM citrate, pH 5.5 (0.2 nmol Man26A/10 mL). After incubation for 30 h at 37° C (with agitation) the release of reducing sugars was determined (Section 2.18.6). The concentrations of reducing ends in the control samples

were subtracted from the concentrations in the test samples, to yield the enzymatic release of reducing ends.

2.20 Proteolysis of Man2A

Man2A was inactivated with 2FManβF as described above (Section 2.18.3). Then 1.2 mg each of Man2A and of inactivated Man2A were treated with 120 μg pepsin at RT in 50 mM phosphate pH 2.0 in 1 mL. Digestions were stopped by freezing (-70° C). The products were analyzed immediately after thawing by liquid chromatography/mass spectrometry (LC/MS) (Mackenzie *et al.*, 1997).

2.21 Electrospray mass spectrometry (ESMS)

Mass spectra were recorded on a PE-Sciex API 300 triple quadrupole mass spectrometer (Sciex, Thornhill, Ontario, Canada) equipped with an Ionspray ion source. Errors in the reported masses are 0.1 %. Peptides were separated by reverse phase HPLC on an Ultrafast Microprotein Analyzer (Micchrom BioResources Inc., Pleasanton, CA) directly interfaced with the mass spectrometer. In each of the MS experiments, the proteolytic digest was loaded onto a C18 column (Reliasil, 1 x 150 mm) equilibrated with solvent A (0.05% trifluoroacetic acid (TFA), 2 % acetonitrile in water). Elution of the peptides was accomplished by using a gradient (0 to 60%) of solvent B (0.045 % TFA, 80 % acetonitrile in water) for 60 min followed by 100 % solvent B for 2 min. Solvents were pumped at a constant flow rate of 50 μL/min. Spectra were obtained in the single quadrupole scan mode (LC/MS) or in the tandem MS product ion scan mode (MS/MS).

In the single quadrupole mode (LC/MS), the quadrupole mass analyzer was scanned over a mass to charge ratio (m/z) range of 300 to 2200 amu with a step size of 0.5 amu and a dwell time of 1.5 ms per step. The ion source voltage (ISV) was set at 5.5 kV and the orifice energy (OR) was 45 V.

The labeled and the corresponding unlabeled peptides were identified by comparative mapping, i.e., by comparing the mass spectra of unlabeled *versus* labeled petide mix.

In the tandem MS daughter ion scan mode, the spectrum was obtained by selectively introducing the parent ion (m/z = 1520 for the labeled and m/z = 1036 for the unlabeled sample) from the first quadrupole (q1) into the collision cell (Q2) and observing the product ions in the third quadrupole (Q3). Thus, Q1 was locked at either 1036 or 1520; Q3 scan range was 100-1040 and 100-1530, respectively; the step size was 0.5; dwell time was 1 ms; ion source voltage (ISV) was 5 kV; orifice energy was 45 V; collision gas = N_2. (Mackenzie *et al.*, 1997)

2.22 Thin layer chromatography (TLC)

Thin layer chromatogaphy was performed on 0.2 mm silica gel aluminum plates (#60 F254; Merck). One to two μL of each reaction mixture or standard was applied to the base of the TLC plate. The plate was either air dried for 10 min or by gentle warming with a hair dryer. The solvent was ethyl actetate : methanol : water (7:2:1 v/v). After air drying,

products were detected either under UV and/or by charring; i.e., heating the TLC plates until bands were visible after dipping them in 10 % H_2SO_4 in methanol.

2.23 Transglycosylation product analysis

NMR spectra were recorded on a 200 Hz Bruker AC-200 and are referenced to the solvent peak. Where required, the interpretations were supported by COSY experiments. Mass spectra were obtained by ion-spray mass spectrometry on a PE-Sciex API III triple quadrupole mass spectrometer (Sciex, Thornhill, Ont., Canada). Reactions were monitored by TLC. Separation of the mixtures of regioisomers (β-1,3 and β-1,4 linked mannobiose and mannotriose) was achieved using both flash chromatography and HPLC. Flash chromatography was performed on short columns (10 to 15 cm) of Merck Kieselgel 60 (230 to 400 mesh). HPLC was performed on a Waters HPLC using a Dynamax column and 70:30 acetonitrile-water as the eluant (flow rate 7.0 mL/min).

3 Characterization of the *Cellulomonas fimi* Mannan-degrading System

3.1 Introduction

3.1.1 Mannan

Mannan is a polymer composed of a β-1,4 linked mannose backbone. It is found in nature in a variety of forms with either backbone substitutions and/or side branches. In soft-woods, mannan is the major hemicellulose component, accounting for up to 25% of the wood dry-weight, whereas in hard-woods heteroxylans are predominant. Soft-wood mannan is present as an acetylated galactoglucomannan (O-acetyl-galactoglucomannan), in which the β-1,4 linked mannose backbone is interspersed with β-1,4 linked glucopyranosides, with a mannose to glucose ratio of 3:1 (Timell 1967; Puls *et al.*, 1993). Galactose monomers are α-1,6 linked side branches linked only to mannose residues (Tenkanen *et al.*, 1997), whereas the acetyl groups are esterified to either the 2 or 3 hydroxyls of the backbone mannoses, and to a much lower extent to glucose (Katz *et al.*, 1964). There are two different forms of galactoglucomannan: water soluble and alkali soluble. They differ in their composition; the water soluble form has a Man:Glu:Gal:Ac ratio of 3:1:1:0.24 whereas the alkali soluble form has a ratio of 3:1:0.1:0.24 (Timell, 1967). The distribution of glucose substitutions is not necessarily uniform: there are runs of glucose residues in pinewood O-acetyl galactoglucomannan, that may be important in the proposed intimate interaction of hemicellulose and cellulose (Tenkanen *et al.*, 1997, Atalla, 1993).

In certain green algae mannan replaces cellulose as the crystalline structural component of the wall (Yui *et al.*, 1997). Mannan-based polysaccharides also function as storage carbohydrates in bulbs and the endosperm of carob seeds, ivory nuts beans (Meier,

1958; McCleary *et al.*, 1985). Ivory nut mannan, from *Phytelephas macrocarpa,* and locust bean gum (or carob-) galactomannan, from *Ceratonia siliqua,* are two commonly used substrates for the study of mannan degrading enzymes (Stålbrand *et al.*, 1993 and 1995; Bolam *et al.*, 1996; McCleary *et al.*, 1983). Ivory nut mannan is an insoluble, crystalline material with linear chains of β-1,4 linked mannosyl residues (Hendrixson *et al.*, 1993). Its crystal structure strongly resembles that of cellulose, with two mannose residues repeating every 10.3 Å (Nieduszynski *et al.*, 1972). Locust bean gum is a heteropolymer, with a β-1,4 linked mannosyl backbone and α-1,6 linked galactose side chains (Figure 3.1). The ratio of galactose to mannose was reported as 1:5, with the galactosyl groups distributed irregularly along the β-1,4 mannosyl backbone (McCleary *et al.*, 1985). Locust bean gum is an industrially important polysaccharide, being used as a stabilizer and a thickener in the food industry (Rol, 1973).

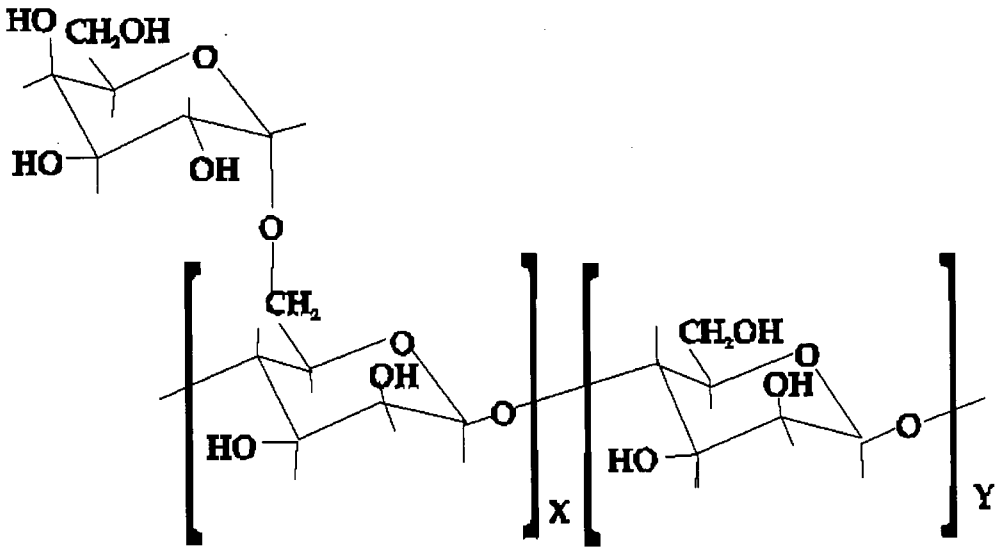


Figure 3.1: Schematic structure of galactomannan. x:y ratio in locust bean gum is 1:4 (McCleary *et al.*, 1983).

3.1.2 Mannan hydrolyzing enzymes

Hydrolysis of acetylgalactoglucomannan into its monomeric components requires the following enzyme activities: endo-β-mannanase, β-mannosidase, α-galactosidase, β-glucosidase and acetyl esterase (Puls and Schuseil, 1993). The study of the mannan-degrading system from *C. fimi* was simplified by studying the degradation of galactomannan (locust bean gum) and mannan (ivory nut mannan), hydrolysis of which requires only endo-β-mannanase, β-mannosidase and, in the case of galactomannan, α-galactosidase.

3.1.2.1 β-mannanase

β-mannanases (mannan endo-1,4-beta-mannosidase; EC 3.2.1.78) are produced by plants, fungi and bacteria. They catalyse the random hydrolysis of β-1,4 mannosidic linkages within the backbone of substrates such as mannan, galactomannan and glucomannan. Among the best characterized β-mannanases are the *Aspergillus niger* mannanase (McCleary *et al.*, 1983), Man1 from *Trichoderma reesei* (Stålbrand *et al.*, 1995; Harjunpää *et al.*, 1995) and ManA from *Pseudomonas fluorescens* ssp. *cellulosa* (Bolam *et al.*, 1996). The sequences of several β-mannanases are known (Braithwaite *et al.*, 1995; Stålbrand *et al.*, 1995; Morris *et al.*, 1995; Arcand *et al.*, 1993; Millward-Sadler *et al.*, 1996). Based on amino acid sequence similarities, β-mannanases can be grouped into two different glycosyl hydrolase families: family 5 and family 26, both members of the superfamily GH-A (Bolam *et al.*, 1996). Mechanistic studies of the family 5 enzyme Man1 from *T. reesei* and family 26 *Pf* ManA have shown that both enzymes cleave their substrate with net retention of the anomeric configuration (Harjunpää *et al.*, 1995; Bolam *et al.*, 1996). Secreted β-mannanases

have a non-modular architecture e.g. *Pf* ManA (Braithwaite *et al.*, 1995), or a modular architecture e.g. *T. reesei* Man1 (Ståhlbrand *et al.*, 1995).

3.1.2.2 β-mannosidase

β-mannosidases (β-1,4-mannoside mannohydrolase EC 3.2.1.25) catalyze the hydrolysis from the non-reducing end of terminal β-D-mannose residues. β-mannosidases from animals, plants, fungi and bacteria have been characterized, but very few β-mannosidase-encoding genes have been sequenced (McCleary *et al.*, 1983; Bauer *et al.*, 1996; Alkhayat *et al.*, 1998). Lysosomal β-mannosidase is involved in degradation of the N-linked oligosaccharide moieties of glycoproteins. A deficiency of lysosomal β-mannosidase in animals and humans is responsible for the neurological disorder β-mannosidosis, which is caused by the accumulation of Man β-1,4 GlcNAc and Man β-1,4 GlcNAc β-1,4 GlcNAc di- and trisaccharide, respectively (Alkhayat *et al.*, 1998; Chen *et al.*, 1995). The genes encoding human (Alkhayat *et al.*, 1998), goat (Leipprandt *et al.*, 1996) and bovine (Chen *et al.*, 1995) β-mannosidase have been sequenced and characterized. Based on their sequence similarities, the eukaryotic β-mannosidases are classified into glycosyl hydrolase family 2, a family with mainly β-galactosidase and β-glucuronidase members (Durand *et al.*, 1997). The only bacterial β-mannosidase sequenced is the intracellular β-mannosidase from the archaeon *Pyrococcus furiosus*, which is a member of glycosyl hydrolase family 1.

Several fungal β-mannosidases and their roles in mannan degradation have been studied. The β-mannosidases from *Aspergillus niger*, *Aspergillus awamori* and *Trichoderma reesei* are secreted into the culture medium and play a role in mannan degradation by releasing mannose preferentially from shorter mannooligosaccharides (Elbein *et al.*, 1977;

Neustroev *et al.*, 1991; Stålbrand personal communication). Furthermore, β-mannosidases from fungal and bacterial sources are useful tools for the analysis of glycoproteins (McCleary *et al.*, 1983; New England Biolabs, 1998).

3.1.2.3 α-galactosidases

α-galactosidases (melibiase, or α-D-galactoside galactohydrolase, EC 3.2.1.22) from plants, fungi and bacteria remove α-1,6 linked galactose groups from galactomannan polymers. The enzymes may be extracellular (Margolles-Clark *et al.*, 1996; McCleary *et al.*, 1983; Talbot *et al.*, 1990) or intracellular (Xiuzhu *et al.*, 1991; Gherardini *et al.*, 1985). α-galactosidases are represented in glycosyl hydrolase families 27 and 36 (Henrissat *et al.*, 1996).

3.1.3 Objectives

It was the aim of this project to study the enzymes of *Cellulomonas fimi* responsible for degrading mannan into its monomers. The genes encoding a β-mannanase and a β-mannosidase were cloned, sequenced and expressed. The products released from mannan and galactomannan by hydrolysis with the recombinant enzymes were identified.

3.2 Results

3.2.1 Identification of galactomannan degrading enzymes produced by *Cellulomonas fimi*

At least one endo- 1,4- β-mannanase, one 1,4- β-mannosidase and one 1,6- α-galactosidase are required to degrade galactomannan into galactose and mannose (Section 3.1.2). To screen for these enzymatic activities, *C. fimi* cultures were grown for six days on minimal medium, suplemented with 0.2 % (w/v) locust bean gum (LBG) as carbon source. The growth rates of LBG cultures were comparable to those of glucose supplemented cultures, indicating that *C. fimi* can degrade LBG and use the degradation products as energy and carbon source. Supernatant and cell-free extract from the LBG culture were assayed for activity by the LBG-Congo red plate assay (Section 2.6.1). Most of the mannanase activity was present in the supernatant, with much lower but detectable activity in the cell-free extract. Culture supernatant was further analyzed by azo-carob galactomannan zymograms (Section 2.6.2). At least three different bands with endo-1,4-β-mannanase activity were detected (Section 3.2.2).

β-mannosidase and α-galactosidase activities were detected only in the cell-free extract. 4-methylumbelliferyl-β-mannoside (MUβMan) and 4-methylumbelliferyl-α-galactoside (MUαGal) were used as substrates. MUβMan and MUαGal zymograms of cell-free extract separated by non-reducing PAGE (Section 2.6.2) revealed only one band of β-mannosidase and one band of α-galactosidase activity. From the electrophoretic mobilities, the M_rs of the α-galactosidase and the β-mannosidase were estimated to be 120 kDa and 60 kDa, respectively. The β-mannosidase comigrated with the *C. fimi* β-glucosidase (Kim and Pack, 1992; Gene bank accession number: M94865).

3.2.2 Carbon source-dependent mannanase synthesis

There were at least three bands with β-mannanase activity (M_rs of 100 kDa, 75 kDa and 30 kDa) on azo-carob galactomannan zymograms of the supernatant of a culture of *C. fimi* grown with LBG. This suggested the presence of more than one β-mannanase encoding gene. Therefore, carbon source-dependent mannanase synthesis was studied.

C. fimi was inoculated into basal salt medium supplemented with either LBG, mannose, glycerol, LBG plus glucose, birch wood xylan, CM-cellulose or galactose as carbon source (0.2 % w/v) or without carbon source. Final cell mass was similar in most cultures. Cultures without carbon source or with CM-cellulose as carbon source grew very slowly. Supernatants from all the cultures were analyzed on zymograms after 2, 6, and 11 days (Figure 3.2). Mannanase gene expression was induced only by LBG or CM-cellulose as carbon source. The LBG culture contained two major (M_rs 75 kDa and 30 kDa), and one minor (M_r 100 kDa) protein with mannanase activity. In the presence of glucose and LBG the synthesis of mannanase activity was repressed. The CM-cellulose culture contained only one protein with β-mannanase activity (M_r 30 kDa). Mannanase activity was not detected in cultures with mannose, galactose, xylan or glycerol as carbon source. These results supported the hypothesis that more than one β-mannananse encoding gene is present in *C. fimi*, and showed that the β-mannanase genes are differentially regulated (this issue will be further addressed in Section 3.7).

β-mannosidase synthesis was induced by LBG and mannose but not by CM-cellulose.

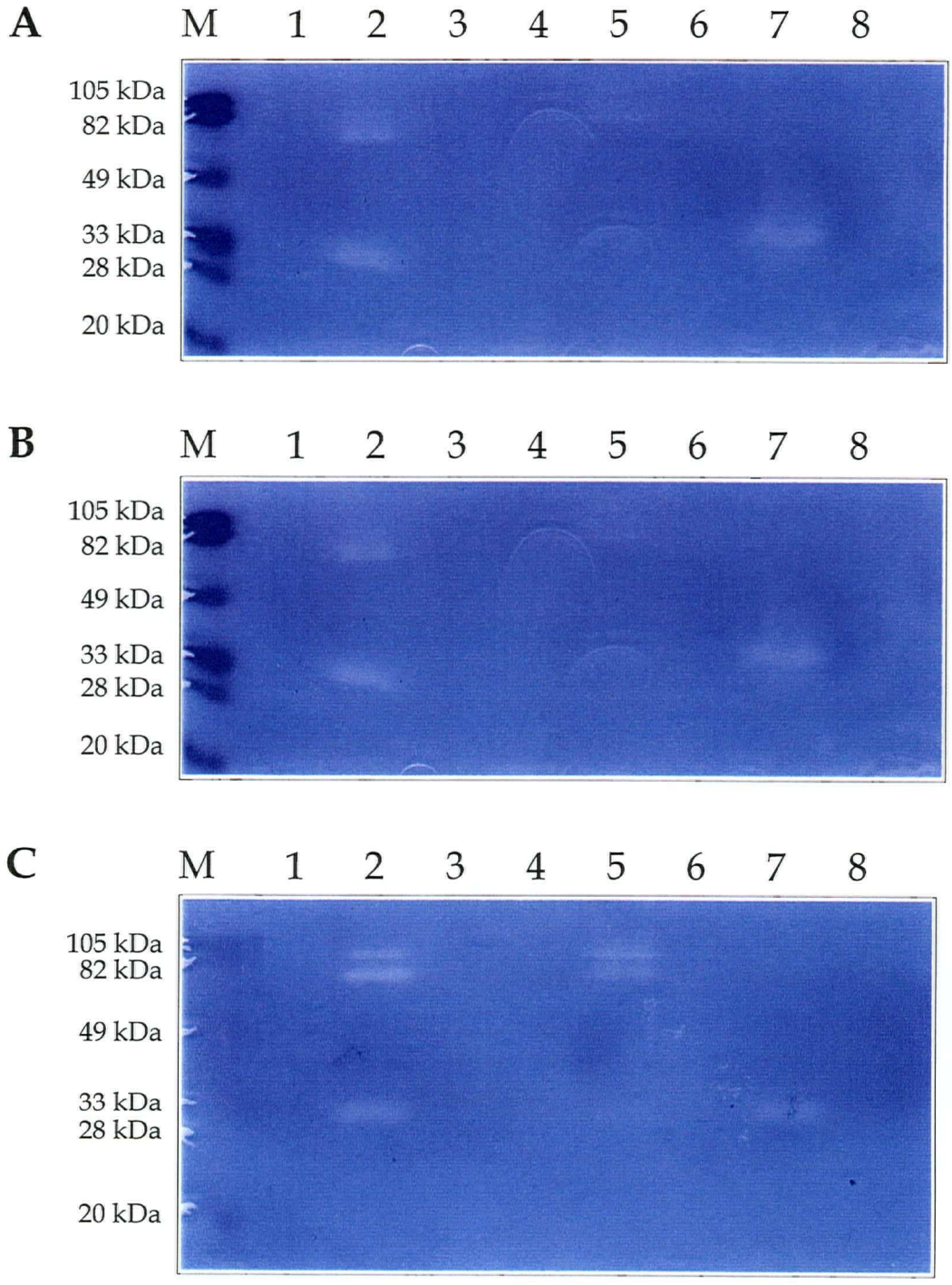


Figure 3.2: Mannanase secretion by *Cellulomonas fimi*. Supernatant samples were analyzed by non-reducing SDS-PAGE-zymograms. *C. fimi* cultures were grown in basal salt medium (Lane 1) or in basal salt medium supplemented with 0.2 % (w/v) locust bean gum (LBG) (Lane 2), mannose (Lane 3), glycerol (Lane 4), glucose and LBG (Lane 5), xylan (Lane 6), CM-cellulose (Lane 7) and galactose (Lane 8); reduced, prestained molecular size standards (Lane M). Panel A: samples from 2 day old cultures; Panel B: samples from 6 day old cultures and Panel C: samples from 11 day old cultures. Zymograms were incubated for 16 h at 37 °C in phosphate buffer.

3.2.3 Screening of the *C. fimi* genomic DNA library for β-mannanase genes

The *C. fimi* genomic library was prepared by inserting genomic DNA fragments (2 to 5 kbp) into the *Eco*R I site of the multiple cloning site of lambda ZAPII (prepared by Stratagene). This created translational fusions of the genomic inserts with the first 36 aa of the *E. coli (lacZ)* β-galactosidase coding sequence transcribed from the *lacZ* promoter (Stratagene; Meinke *et al.*, 1993). Therefore, the *C. fimi* lambda ZapII library could be screened for IPTG-inducible β-mannanase activity on azo-carob galactomannan plates (Section 2.7). Eight plaques with mannanase activity were isolated, their phagemids (pBluescript SK+*C. fimi* DNA insert) excised, and transferred into *E. coli* XLOLR cells. All eight mannanase clones, CMan1-8, secreted mannanase activity into the culture supernatant. In the supernatant of each clone at least two bands with mannanase activity could be detected by zymograms. The only clone that produced higher molecular weight bands with significantly more mannanase activity than the other clones was CMan2. From the other seven clones, slightly more mannanase activity could be detected in the supernatant of CMan4. CMan2 and CMan4 were therefore chosen for further analysis. Restriction maps of their plasmids were established (Figure 3.3). pCMan2 had a 4.3 kbp and pCMan4 a 6.3 kbp insert. The similarity of the restriction pattern from pCMan2 and pCMan4 suggested that their inserts had a DNA fragment in common (Figure 3.3 A). By DNA sequencing it was shown that these two plasmids carried inserts with identical 5′ ends except for 65 extra bp in pCMan4 (Figure 3.4). This suggested that the 6.3 kbp long insert of pCMan4 contained DNA encoding the same mannanase as pCMan2. The insert in pCMan4, however, was 65 bp longer at the 5′ end and 2 kbp longer at the 3′end than

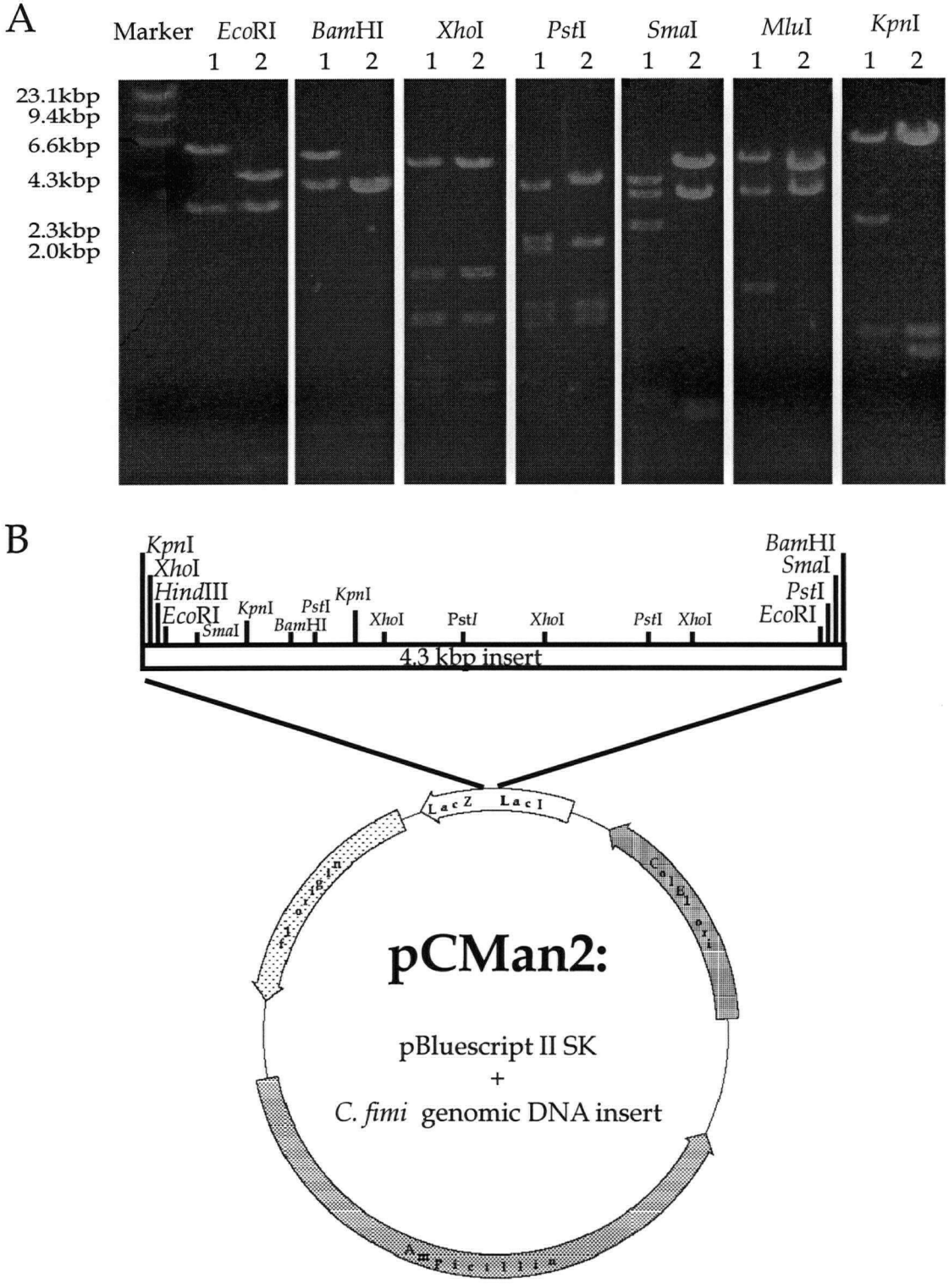


Figure 3.3: Restriction mapping of pCMan2 and pCMan4. Panel A: Restriction digests of pCMan4 (Lane 1) and pCMan2 (Lane 2) separated on a 0.9 % agarose gel. Restriction endonucleases used are indicated. λ *Hin*d III DNA was used as size standard.
Panel B: Plasmid map of pCMan2. The restriction map of the 4.3 kbp *C. fimi* genomic DNA insert of pCMan2 is shown. Not all the fragments could be mapped.

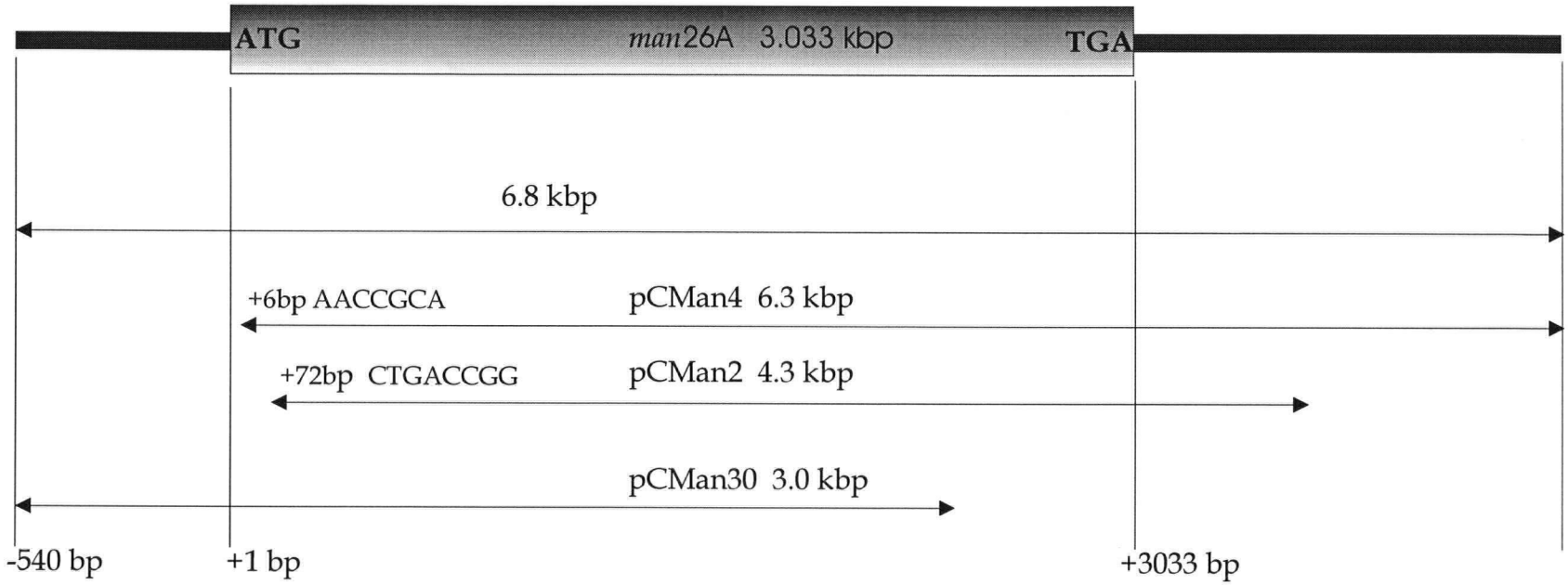


Figure 3.4: A schematic representation of the 6.8 kbp region of *C. fimi* genomic DNA containing the *man26A* open reading frame. The open reading frame is represented as a box, starting with the ATG codon at +1 and ending with the TGA codon at +3033. The genomic DNA inserts of the three plasmids, pCMan4, pCMan2 and pCMan30 and their relative positions are represented by arrows. The arrowheads show the start and the end of the inserts, relative to the *man26A* open reading frame.

pCMan2 (Figure 3.4). Hence, pCMan2 was chosen for DNA sequencing. The subclones and oligodeoxyribonucleotide primers used for sequencing are listed in section 2.2 and 2.3. A putative start codon was found in neither pCMan2 nor in pCMan4.

Because the mannanase activities detected in the supernatants of all the other clones, CMan1, 3, and 5 to 8, were very similar to CMan4 (*vide supra*) the library was rescreened for more mannanase encoding clones. The extracellular mannanase activities of five clones, CMan10, 20, 30, 40 and 50 were analyzed on zymograms. CMan30 secreted two active polypeptides that were very similar in size to those secreted by *C. fimi*. The plasmid insert of CMan30, was 3 kbp in size. Comparisons of restriction digests of the plasmids pCMan30, pCMan2 and pCMan4 showed that all three plasmids had a DNA region in common. An ATG start codon was found in pCMan30, 540 bp downstream from the 5′ end of the insert. This start codon was found to be only 6bp and 72 bp upstream from the 5′ ends of the inserts in pCMan 4 and pCMan2, respectively (Figure 3.4).

3.2.4 Nucleotide- and deduced amino acid sequence of the *C. fimi* mannanase

The nucleotide sequence of its gene and the deduced amino acid sequence of the *C. fimi* mannanase are shown in Figure 3.5. The open reading frame was 3033 bp long, which translated into a 1011 amino acid long protein with a calculated MW of 107,033. The N-terminus of the *C. fimi* mannanase had an amino acid composition rich in positively charged residues, a characteristic of secretion signal peptides (Nielsen *et al.*, 1997). To determine the N-terminus of the secreted and processed *C. fimi* mannanase, concentrated *C. fimi* LBG culture supernatant was separated and analyzed by non-reducing PAGE-

GCCCGACCTGTGGCTGCGCTGTGCGGCGCCAGGGCCCGGCTGACCTCGGGCGACGCCGGTGAAACGCTTCGCGTGCG
AGCCGGGCCCTCGGTAGCGCCTGTCCGACCTGTTCCGTCCTGACTAAGTTGGTAACTTGACGAATCTGTCCGCGTCC
GTCCGGAGCCTCCGGTCGGCGCGGCCCCTGGAGGGCTCGGGCGCCGCTGCCCGGCCCGACCGTCACCT**GGAGG**TTCA

1/1										31/11									
ATG	ACG	AAC	CGC	AGC	CGT	CCG	CGC	GGG	CGC	ACG	GCG	GGC	CAC	GTG	CTC	GCC	GCC	ACC	GCC
M	T	N	R	S	R	P	R	G	R	T	A	G	H	V	L	A	A	T	A
61/21										91/31									
GCG	GCG	CTG	GCC	CTG	ACC	GGC	CTG	TCG	GCG	CTT	CCC	GCC	CAG	TCC	GCA	CCG	GCG	CCC	GCA
A	A	L	A	L	T	G	L	S	A	L	P	A	Q	S	A	P	A	P	A ♣
121/41										151/51									
GCG	CCC	GTC	GCG	GGG	GCG	CTG	CCC	ACC	GCA	CCC	GCC	GAC	GAG	ACC	ATC	GCG	ATC	GTC	GAC
A	P	V	A	G	A	L	P	T ♦	A	P	A	D	E	T	I	A	I	V	D
181/61										211/71									
GCC	GAC	GCG	ACC	GCC	GAG	ACC	CGG	TCG	CTG	CTC	TCC	TAC	CTC	GAC	GGC	GTG	CGC	GGC	GAG
A	D	A	T	A	E	T	R	S	L	L	S	Y	L	D	G	V	R	G	E
241/81										271/91									
GGC	ATC	CTG	TTC	GGC	CAC	CAG	CAC	ACG	ACG	TCG	TTC	GGG	CTC	ACC	ACC	GGA	CCC	ACC	GAC
G	I	L	F	G	H	Q	H	T	T	S	F	G	L	T	T	G	P	T	D
301/101										331/111									
GGC	ACG	ACC	TCC	GAC	GTC	AAG	AAC	GTC	ACG	GGC	GAC	TTT	CCC	GCG	GTC	TTC	GGC	TGG	GAC
G	T	T	S	D	V	K	N	V	T	G	D	F	P	A	V	F	G	W	D
361/121										391/131									
ACG	CTC	ATC	ATC	GAG	GGC	AAC	GAG	CGC	CCC	GGG	CTC	GCC	GAG	AAC	ACG	CGC	GAC	GAG	AAC
T	L	I	I	E	G	N	E	R	P	G	L	A	E	N	T	R	D	E	N
421/141										451/151									
ATC	GCG	CTG	TTC	GCC	GAC	TAC	ATC	CGC	AAG	GCC	GAC	GCG	ATC	GGC	GGC	GTC	AAC	ACC	GTG
I	A	L	F	A	D	Y	I	R	K	A	D	A	I	G	G	V	N	T	V
481/161										511/171									
AGC	GCG	CAC	GTC	GAG	AAC	TTC	GTC	ACC	GGC	GGC	TCG	TTC	TAC	GAC	ACC	TCG	GGC	GAC	ACG
S	A	H	V	E	N	F	V	T	G	G	S	F	Y	D	T	S	G	D	T
541/181										571/191									
CTG	CGC	GCC	GTG	CTG	CCG	GGC	GGC	TCG	CAC	CAC	GCC	GAG	CTC	GTC	GCC	TAC	CTC	GAC	GAC
L	R	A	V	L	P	G	G	S	H	H	A	E	L	V	A	Y	L	D	D
601/201										631/211									
ATC	GCG	GAG	CTC	GCC	GAC	GCG	TCG	CGC	CGC	GAC	GAC	GGC	ACG	CTC	ATC	CCG	ATC	GTC	TTC
I	A	E	L	A	D	A	S	R	R	D	D	G	T	L	I	P	I	V	F
661/221										691/231									
CGG	CCG	TGG	CAC	GAG	AAC	GCC	GGC	TCG	TGG	TTC	TGG	TGG	GGC	GCC	GCG	TAC	GGC	TCA	CCC
R	P	W	H	E	N	A	G	S	W	F	W	W	G	A	A	Y	G	S	P
721/241										751/251									
GGC	GAG	TAC	CAG	GAG	CTC	TAC	CGG	TTC	ACC	GTG	GAG	TAC	CTG	CGC	GAC	GTC	AAG	GGC	GTC
G	E	Y	Q	E	L	Y	R	F	T	V	E	Y	L	R	D	V	K	G	V
781/261										811/271									
TCC	AAC	TTC	CTC	TAC	GCG	TGG	GGT	CCG	GGC	GGC	GGC	TTC	GGC	GGC	AAC	CGC	GAC	GTC	TAC
S	N	F	L	Y	A	W	G	P	G	G	G	F	G	G	N	R	D	V	Y
841/281										871/291									
CTG	CGC	ACC	TAC	CCC	GGC	GAC	GCG	TTC	GTC	GAC	GTG	CTC	GGC	CTC	GAC	ACC	TAC	GAC	AGC
L	R	T	Y	P	G	D	A	F	V	D	V	L	G	L	D	T	Y	D	S
901/301										931/311									
ACC	GGT	TCG	GAC	GCG	TTC	CTC	GCC	GGG	CTC	GTC	GCC	GAC	CTG	CGG	ATG	ATC	GCC	GAG	ATC
T	G	S	D	A	F	L	A	G	L	V	A	D	L	R	M	I	A	E	I
961/321										991/331									
GCC	GAC	GAG	AAG	GGC	AAG	GTG	TCG	GCG	TTC	ACC	GAG	TTC	GGC	GTG	AGC	GGC	GGC	GTG	GGC
A	D	E	K	G	K	V	S	A	F	T	E	F	G	V	S	G	G	V	G
1021/341										1051/351									
ACG	AAC	GGC	TCG	TCG	CCC	GCG	CAG	TGG	TTC	ACC	AAG	GTG	CTC	GCC	GCG	ATC	AAG	GCC	GAC
T	N	G	S	S	P	A	Q	W	F	T	K	V	L	A	A	I	K	A	D
1081/361										1111/371									
CCC	GTC	GCG	AGC	CGC	AAC	GCC	TAC	ATG	GAG	ACG	TGG	GCC	AAC	TTC	GAC	GCC	GGC	CAG	CAC
P	V	A	S	R	N	A	Y	M	E	T	W	A	N	F	D	A	G	Q	H
1141/381										1171/391									
TTC	GTC	CCC	GTG	CCC	GGC	GAC	GCG	CTG	CTC	GAG	GAC	TTC	CAG	GCG	TAC	GCC	GCC	GAC	CCG
F	V	P	V	P	G	D	A	L	L	E	D	F	Q	A	Y	A	A	D	P

1201/401 1231/411
TTC ACG CTG TTC GCG TCC GAG GTC ACG GGC GCG TTC GAC CGG ACC GTC GCC GCA GCG CCC
F T L F A S E V T G A F D R T V A A A P
1261/421 1291/431
GCG CAG CCG GTC GTG CAC ATC GCC TCT CCG GCC GAC GGC GCG CGC GTC GCG TCC GCC CCG
A Q P V V H I A S P A D G A R V A S A P
1321/441 1351/451
ACC ACC GTG CGG GTG CGG GTC GGC GGC ACC GAC GTG CAG TCC GTG ACC GTC GAG GTC GCC
T T V R V R V G G T D V Q S V T V E V A
1381/461 1411/471
CAG GGC GGC ACC GTC GTC GAC ACT CTG GAC CTC GCG TAC GAC GGC GCC CTG TGG TGG ACG
Q G G T V V D T L D L A Y D G A L W W T
1441/481 1471/491 MBD11
GCC CCC TGG TCG CCG ACC AGC GCG CAG CTC GAC AAC AGC ACC TAC ACC GTC ACC GCG ACG
A P W S P T S A Q L D N S T Y T V T A T
1501/501 1531/511
GCG ACG ACC GCC GCC GGG ACG CTC GAC GTC ACG AAC GAG GTC GTC CTC GGG CCG AGG CCG
A T T A A G T L D V T N E V V L G P R P
1561/521 1591/531
ACG TTC CCC GCG GGC GTC GTC GAC GAC TTC GAG GGC TAC GGC GAC GAC ACC GCG CTG CGT
T ♥ F P A G V V D D F E G Y G D D T A L R
1621/541 1651/551
GCC GAG TAC GTG ACC TAC GGC GCC AAC ACG ATC TCG CTC GAG ACG GGG TCC GTC GGG GGC
A E Y V T Y G A N T I S L E T G S V G G
1681/561 1711/571
GGC GCG AAG GCG CTG CGG CTC GAC TAC GAC TTC GCG ACG CAG ACC TAC ACC GGC GTC GGC
G A K A L R L D Y D F A T Q T Y T G V G
1741/581 1771/591
AAG CAG CTG TCC GGC GAC TGG TCC GAC TTC AAC GAG CTC GCG ATC TGG GTC GAC CCC GAC
K Q L S G D W S D F N E L A I W V D P D
1801/601 1831/611
GGC TCG AAC AAC AGG ATG GTC CTG CAG CTC AAC GCC GGT GGT GTC GCC TAC GAG GCG TAC
G S N N R M V L Q L N A G G V A Y E A Y
1861/621 1891/631
CCG TCG CTC GCG GGC GAC GAG CCG CAG CTC GTG ACG ATC CCG TTC GTC GAC TGG CGC CCG
P S L A G D E P Q L V T I P F V D W R P
1921/641 1951/651
GCA CCG TGG GAC ACC GCG CAC GCC GAC CGC CGC ATG TCC GAC GAC GAC CTG CGC GCG CTC
A P W D T A H A D R R M S D D D L R A L
1981/661 2011/671 MBD12
ACG TCG TTC AAC GTC TAC GTC AAC AGC GCC GAG GGC GGT GCC GCG TCG GGC TCG CTC GTC
T S F N V Y ♠ V N S A E G G A A S G S L V
2041/681 2071/691
GTC GAC GAC ATC GCC GCC CAC CCC GGC GTC GAG CCG CCG CCG CTG TTC TCC GAC GTC CCG
V D D I A • A H P G V E P P P L F S D V P
2101/701 2131/711
AAG GGC CAC CCC TAC GAG ACG GAG ATC CTG TGG CTG CAC GCG CAG GGG CTC GAC GAC GGC
K G H P Y E T E I L W L H A Q G L D D G
2161/721 2191/731
TAC GAC GAC GGC ACC TTC CGG CCC GCC AGG CAG GTC AAG CGG CAG GAC GTC GCG CGG CTG
Y D D G T F R P A R Q V K R Q D V A R L
2221/741 2251/751
CTG CAC GCC TAC GAG AAG GCG GTC TTC ACG CCG CCG ACG ACA CCG TCG TTC CTC GAC GTC
L H A Y E K A V F T P P T T P S F L D V
2281/761 2311/771
CGC CGG AGC CAC CCC GCC TAC ACC GCG ATC GAG TGG CTC GTC GCC GAG GGG CTC GTG GAC
R R S H P A Y T A I E W L V A E G L V D
2341/781 2371/791
GAC GGG CGC GTC TTC CTG CCG AGC GCC CCG CTC GAC CGG GCC ACC GCC GCC GAG CTG CTG
D G R V F L P S A P L D R A T A A E L L
2401/801 2431/811
TGG CGG CTC GCC GGG TCC CCC GAG CCC GAG GGG ACG GAG GCG TTC CGC GAC GTC CCC ACG
W R L A G S P E P E G T E A F R D V P T
2461/821 2491/831
TGG CAC CGG TAC CGC ACC GCG ATC ACG TGG GCG ACC GAG GTG GGC GTC GTC GAG CCC GTG
W H R Y R T A I T W A T E V G V V E P V

```
2521/841                                        2551/851
TCG GCG TCG ACG TTC GGG GTG CTC AAG GCC GTG CAG CGG CAG GAG CTC GCG CGG TAC CTG
S   A   S   T   F   G   V   L   K   A   V   Q   R   Q   E   L   A   R   Y   L
2581/861                                        2611/871
TAC CGG TTC GAC GCC CTC CCG TCG CCG CTC GAG CCC GTC GTG CTG TCG GAC TTC GCC GAC
Y   R   F   D   A   L   P   S   P   L   E   P   V   V   L   S   D   F   A   D
2641/881                                        2671/891
GGC GCG CAG GGC TGG GGG CCG CTC GAC GCC GGT CCG GGC TCC GCG ACC GCG TCC GGC GGC
G   A   Q   G   W   G   P   L   D   A   G   P   G   S   A   T   A   S   G   G
2701/901                                        2731/911
ACG CTC ACG ATC GAG GCC GCC GCG CCC GAC GGC GGG TGG TTC TCG TTC ACC CCG TCC GTC
T   L   T   I   E   A   A   A   P   D   G   G   W   F   S   F   T   P   S   V
2761/921                                        2791/931
GCC GAC TGG ACC GGG CGG ACG GCG CTC GCC CTC GAC GTG GTC TCC ACC ACC GGG TTC GAC
A   D   W   T   G   R   T   A   L   A   L   D   V   V   S   T   T   G   F   D
2821/941                                        2851/951
ACC AAG GCG GCC CTG CAG GTC GGC TCG ACC TGG CAG TGG TGC GAG ACC GCG CAG GCC GGG
T   K   A   A   L   Q   V   G   S   T   W   Q   W   C   E   T   A   Q   A   G
2881/961                                        2911/971
TGG GTG TCC GGG CCG CGC ACC GGA GAC GAC GCG CTC GTG CTC GAC CTC ACG ACG TTG TAC
W   V   S   G   P   R   T   G   D   D   A   L   V   L   D   L   T   T   L   Y
2941/981                                        2971/991
GCG GGG TGC GGC GCC CAG CTC GCC GAC GTC AAG CGC GTG AAC CTC TAC CTC AAC GCA GGC
A   G   C   G   A   Q   L   A   D   V   K   R   V   N   L   Y   L   N   A   G
3001/1001                                       3031/1011
ACC CAC GTG ATC GAC GAC GTC GAG CTG CGC TGA  CGCCTGGACCGCACGCGCGGGGGCCGGTCACCC
T   H   V   I   D   D   V   E   L   R   *

ACGTGGTGGCCGGCCCCCGCCACGCGAGGGCGTCGGCTCGCCCGGGGCCCGGGCGGTGCGCCCGCCCGGGGGGCGCT
CGGGTCAGGCGGGGGTGAGGAGGCCGTCGGTGACCAGGCCGCGGACCGAGGGCAGGACGTCGGCCGCGACGGCGTCG
```

Figure 3.5: Nucleotide sequence and deduced amino acid sequence of *man26A*. Putative promoter sequences, identified based on similarities to other *C. fimi* promoters are underlined. The putative ribosome binding site is shown in bold. A possible transcription termination hairpin is overlined. The PCR primers, MBD11 and MBD12, used for sub-cloning of mbd_{1112} are indicated by arrows. ♣: Predicted signal peptide processing site, ♦: experimentally identified N-terminus of Man26A, ♥: predicted C-terminus of Man26A catalytic domain, ♠: N-terminus of 29 kDa SLH domain fragment, •: N-terminus of 21 kDa SLH domain fragment (See text for details).

zymogram and blotted onto a PVDF membrane for N-terminal sequence analysis. The N-terminal sequence of the processed mannanase, corresponding to the 75 kDa active polypeptide (Section 3.2.2), was determined by Edman degradation as $_{50}$APADET$_{55}$, with the starting methionine being position 1. The cleavage between Thr 49 and Ala 50 was not in agreement with the signal peptide cleavage site as predicted by computer analysis using the program SignalP (Nielsen *et al.*, 1997). Cleavage was predicted to occur between Ala 40 and Ala 41 in the sequence $_{37}$PAPA ⇩ APV$_{43}$ (⇩: indicating the cleavage site). This prediction was in agreement with the (-3, -1) rule (Nielsen *et al.*, 1997) and with the consensus cleavage sequence **A/V**X**A** ⇩ **A (** X can be any amino acid) from secreted *C. fimi* glycanases. Secreted *C. fimi* protease preferentially cleaves C-terminal to threonines (Gilkes *et al.*, 1989, Sandercock unpublished results), suggesting that the experimentally determined N-terminus APADET (cleaved after Thr 49) of the secreted mannanase might have resulted from signal peptidase processing followed by proteolysis by the secreted *C. fimi* protease (see also Section 3.2.6).

The amino acid sequence of the *C. fimi* mannanase was compared to those of other proteins. The sequence of the N-terminal half of the *C. fimi* mannanase was similar to those of the catalytic domains of mannanases in glycosyl hydrolase family 26 (EC 3.2.1.78). The highest identity was with mannanase ManA from *Pseudomonas fluorescens* ssp. *cellulosa* (Pf ManA) (Braithwaite *et al.*, 1995). The two proteins were 42 % identical over a sequence of 328 aa. The *C. fimi* mannanase was also similar to mannanases ManB from *Bacillus subtilis* (Mendoza *et al.*, 1995), ManB from *Caldocellulosiruptor saccharolyticus* (Lüthi *et al.*, 1991) and ManA, ManB and ManC from *Piromyces sp.* (Millward-Sadler *et al.*, 1996), all members of family 26 (Figure 3.6).

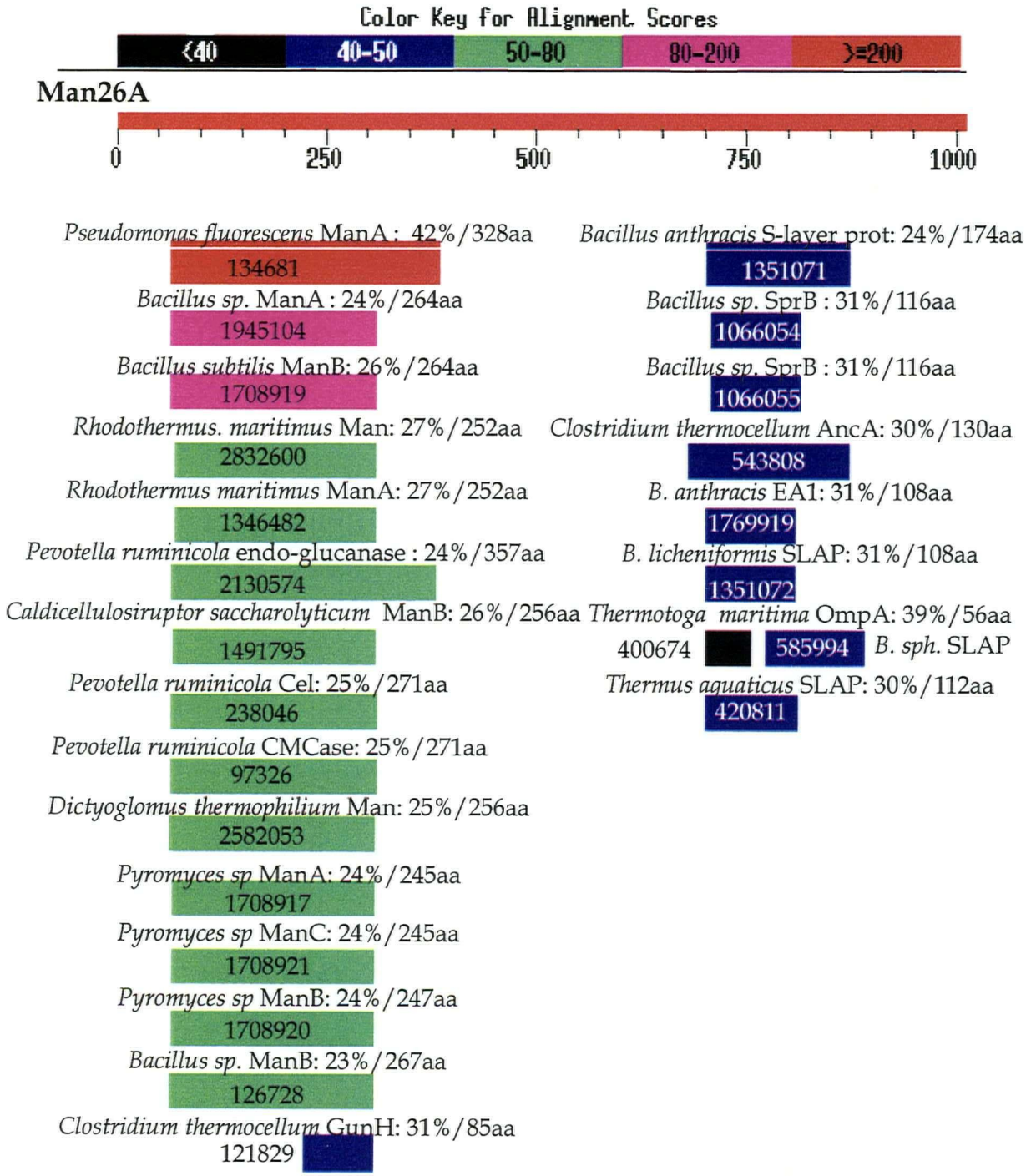


Figure 3.6: Results from protein sequence similarity search with Man26A. Each bar represents an amino acid sequence which is similar to the sequence of Man26A. The position of each bar is aligned for the homology with Man26A. The alignment score is colour-coded and the identities and alignment lengths are indicated for each protein sequence. The numbers in the or beside the bars are the sequence IDs.(Blast search; Altschul *et al.*, 1997).

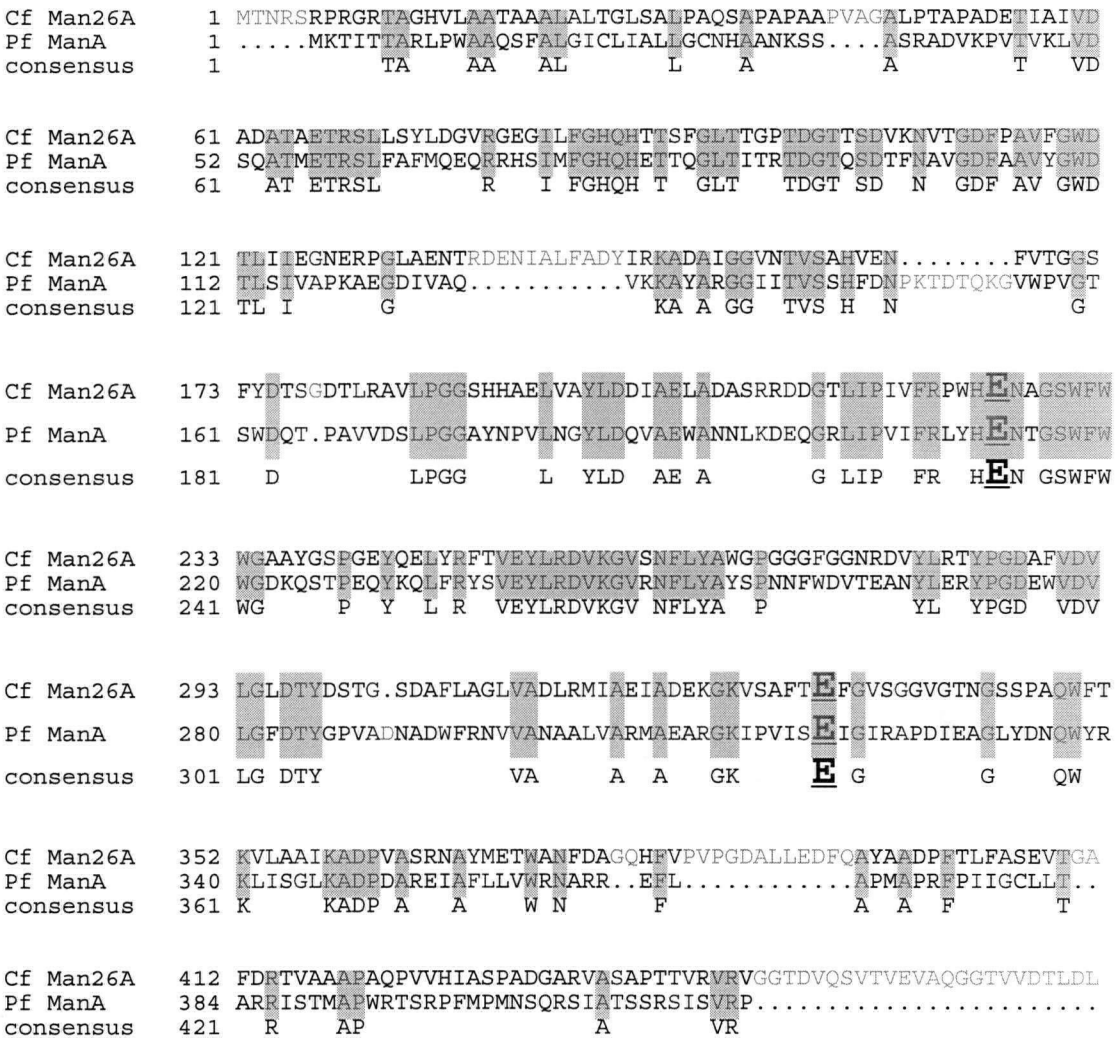


Figure 3.7: Alignment of the first 460 aa from *Cellulomonas fimi* mannanase, Cf Man26A with the entire sequence of the *Pseudomonas fluorescens* mannanase, Pf ManA. Identical residues are highlighted. The two family 26 proteins share 42 % identity.
For Pf ManA, Glu212 and Glu320 were experimentally determined to be the catalytic residues acting as acid/base catalyst and nucleophile, respectively. In Man26A these residues are conserved and correspond to Glu225 and Glu332. The catalytic residues are underlined. Alignment was performed with Clustal W (Tompson *et. al.*, 1994).

The best studied enzyme in family 26 is Pf ManA. This enzyme, representative for all family 26 members, cleaves the substrate *via* a double displacement mechanism, with a net retention of the configuration at the anomeric center. The catalytic residues in Pf ManA were determined by site-directed mutation of conserved family 26 carboxylic residues and kinetic studies of these mutants. Glu212 was identified as the acid-base catalyst, and Glu320 as the catalytic nucleophile (Bolam *et al.*, 1996). Both catalytic residues from Pf ManA are conserved in the *C.fimi* mannanase and correspond to Glu 225 as acid/base catalyst and to Glu 332 as the catalytic nucleophile (Figure 3.7). In accordance with the nomenclature proposed recently the *C. fimi* mannanase, the first family 26 enzyme from *C. fimi* to be described, was named Man26A (or Cf Man26A) (Henrissat *et al.*, 1998).

In Man26A, between residues 680 and 860, another region with homology to other proteins was found. All the proteins sharing identities with Man26A in this region are either S-layer proteins, or proteins with a S-layer-homology (SLH) domain, e.g. *Bacillus anthracis* S-layer protein (24 % identical residues over a sequence of 174 aa), *Bacillus sp* SprB (31 % identity over 116 aa), *Clostridium thermocellum* ORF3p, also reported as ANCA (30 % identity over 130 aa), and the endoglucanase from *Clostridium josui* (23 % identity over 119 aa), to name just a few (Figure 3.6). Typically a SLH domain is composed of three repeated aa sequences (Olabarria *et al.*, 1996). The *C. fimi* Man26A SLH domain homology region was also composed of three repeats, each of which was about 60 aa long. Twenty percent of the amino acids were conserved in all three repeats (Figure 3.8).

Figure 3.8: Alignment of SLH domain repeats (1-3) from Cf Man26A. 696, 757 and 815 are indicate the position of the first amino acid residues of each repeat in the Man26A sequence. The residues that are identical in two or all three repeats are highlighted.

```
1. 696FSDVPKGHPYETEILWLHAQGLDDGYDDGTFRPARQVKRQDVARLLHAYEKAVFTPPTTP
2. 757FLDVRRSHPAYTAIEWLVAEGLVD  DGRVFLPSAPLDRATAAELLWRLAGSPEPEGTEA
3. 815FRDVPTWHRYRTAITWATEVGVVEPVSASTFGVLKAVQRQELARYLYRFDALPSPLEPVV
```

SLH domains are generally involved in anchoring S-layer proteins to the bacterial cell wall. They also occur in other secreted proteins, such as xylanases, pullulanases and cellulosome anchoring proteins (Ries *et al.*, 1997; Lemaire *et al.*, 1993; Fujino *et al.*, 1993a). Cf Man26A is the first mannanase reported to have a SLH domain (Section 3.2.7).

The sequences in Man26A between aa 470 to 680 and 940 to 1011 did not show any significant homologies to other protein sequences.

3.2.5 Sub-cloning of the gene and production of *C. fimi* mannanase, Man26A

A flow sheet illustrating the sub-cloning of *man26A* into the expression vector pET27b is shown in Figure 3.9. The *E.coli* strain BL21(DE3) was the host strain used for production of Man26A fused to a C-terminal His_6 tag. Approximately 10 mg of purified protein were obtained from 1 liter of culture supernatant after purification by MCAC (Section 2.8). This suggested that the leader peptide from Man26A was recognized by *E.coli* cells as a secretion signal. However, this yield was not very good. Better yields were obtained by purifying the enzyme from a cell-free extract: 60 mg of purified and properly folded enzyme were obtained per liter of culture (Section 2.8 and Figure 3.10).

3.2.6 Analysis of the modular organization of *C. fimi* Man26A

The results obtained from aa sequence analysis (Section 3.2.4) suggested the presence of at least two domains. The modular architecture of Man26A was further analyzed. Since folded domains are protected against proteolytic attack by proteases, such as the secreted *C. fimi* protease, whereas the interdomain regions are very susceptible (Gilkes *et al.*, 1989),

Figure 3.9 (following page): Generation of pET27Man26A, encoding the *C. fimi* β-mannanase, Man26A with a C-terminal His_6 tag. A *Nde* I restriction site was introduced at the 5' end and a *Nhe* I restriction site was introduced at the 3' end of *man26A*. Only the restriciton sites that were relevant for the cloning steps are shown.

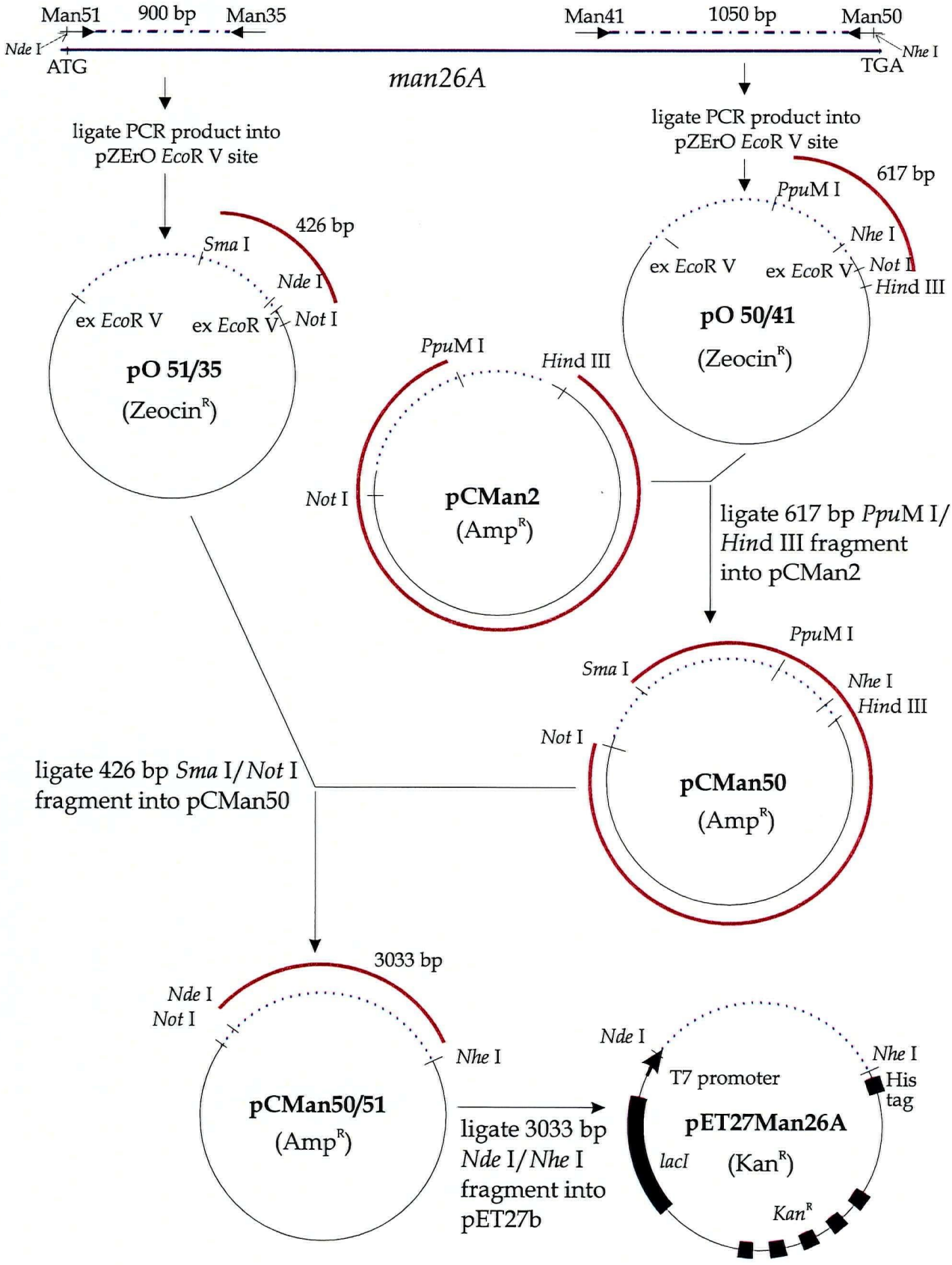
——► : Man51 (*Nde* I), Man35, Man41, Man50 (*Nhe* I) PCR primers

..... : *man26A* or *man26A* derived plasmid insert

– · – · – · : PCR product

——— : *man26A* PCR template

——— : DNA fragments used for the sub cloning steps. The DNA fragments were obtained by restriction endonuclease digestions, separation on agarose gels and purification gel extraction.


Man51
900 bp
Man35
Man41
1050 bp
Man50
Nde I
ATG
man26A
Nhe I
TGA
ligate PCR product into pZErO EcoR V site
ligate PCR product into pZErO EcoR V site
426 bp
Sma I
Nde I
Not I
ex EcoR V
ex EcoR V
pO 51/35
(Zeocin^R)
617 bp
PpuM I
Nhe I
Not I
Hind III
ex EcoR V
ex EcoR V
pO 50/41
(Zeocin^R)
PpuM I
Hind III
Not I
pCMan2
(Amp^R)
ligate 617 bp PpuM I/ Hind III fragment into pCMan2
PpuM I
Sma I
Nhe I
Hind III
Not I
pCMan50
(Amp^R)
ligate 426 bp Sma I/ Not I fragment into pCMan50
3033 bp
Nde I
Not I
Nhe I
pCMan50/51
(Amp^R)
ligate 3033 bp Nde I/ Nhe I fragment into pET27b
Nde I
Nhe I
T7 promoter
His tag
pET27Man26A
lacI
(Kan^R)
Kan^R

Man26A produced by *E. coli* was subjected to proteolysis by the *C. fimi* protease to probe its modular architecture (Section 2.13; Figure 3.10). The preparation of purified Man26A contained degradation products with M_rs of 104, 70, 50, 28 and 21 kDa. Treatment of Man26A with *C. fimi* protease produced similar fragments (Figure 3.10). The first proteolytic event produced the polypeptide with a M_r of 104 kDa, corresponding probably to the loss of the N-terminal 50 aa (Section 3.2.4). This polypetide was further digested into a polypeptide with an estimated M_r of 53 kDa (*via* a 70 kDa intermediate) and into a polypeptide with an estimated M_r of 28 kDa. These two polypeptides were further cleaved to produce fragments with M_rs of 50 kDa and 21 kDa, respectively, which resisted further proteolysis, even after prolonged (48 h) incubation with protease. Analysis of the two final fragments by mass spectrometry (MALDI-TOFMS) gave masses of 50.4 and 21.2 kDa. N-terminal amino acid sequencing showed that the peptide of M_r 50 kDa corresponded to the catalytic domain, the peptide of M_r 28 kDa corresponded to the SLH domain, and the peptide of M_r 21 kDa arose by cleavage at the N- and C-termini of the 28 kDa peptide (Figure 3.11).

The fragmentation pattern was consistent with the hypothesis that Cf Man26A is a modular enzyme, composed of two structural units that resist proteolysis; the catalytic domain (CD 26) and the SLH domain, which are both flanked by exposed protease susceptible sequences.

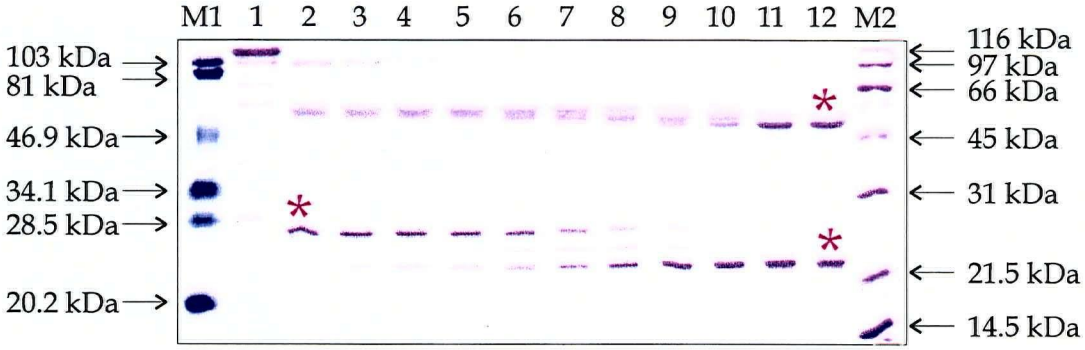


Figure 3.10: SDS-PAGE (12 %) of protease-digested Man26A. Recombinant Cf Man26A was digested with *C. fimi* protease at 37 °C (1 U protease/60 μg Man26A). Samples were removed at different times (Lane 1-12) 1: O min, 2: 15 min, 3: 30 min, 4: 45 min, 5: 60 min, 6: 90 min, 7: 2 h, 8: 3 h, 9: 5 h, 10: 7 h, 11: 16 h, 12: 24 h. Digestion was stopped by boiling samples in SDS loading buffer for 2 min. Molecular weight standards are shown in Lane M1 and M2. The M_rs are indicated. The N-termini of peptides corresponding to bands labeled with an asterisk (*) were sequenced.

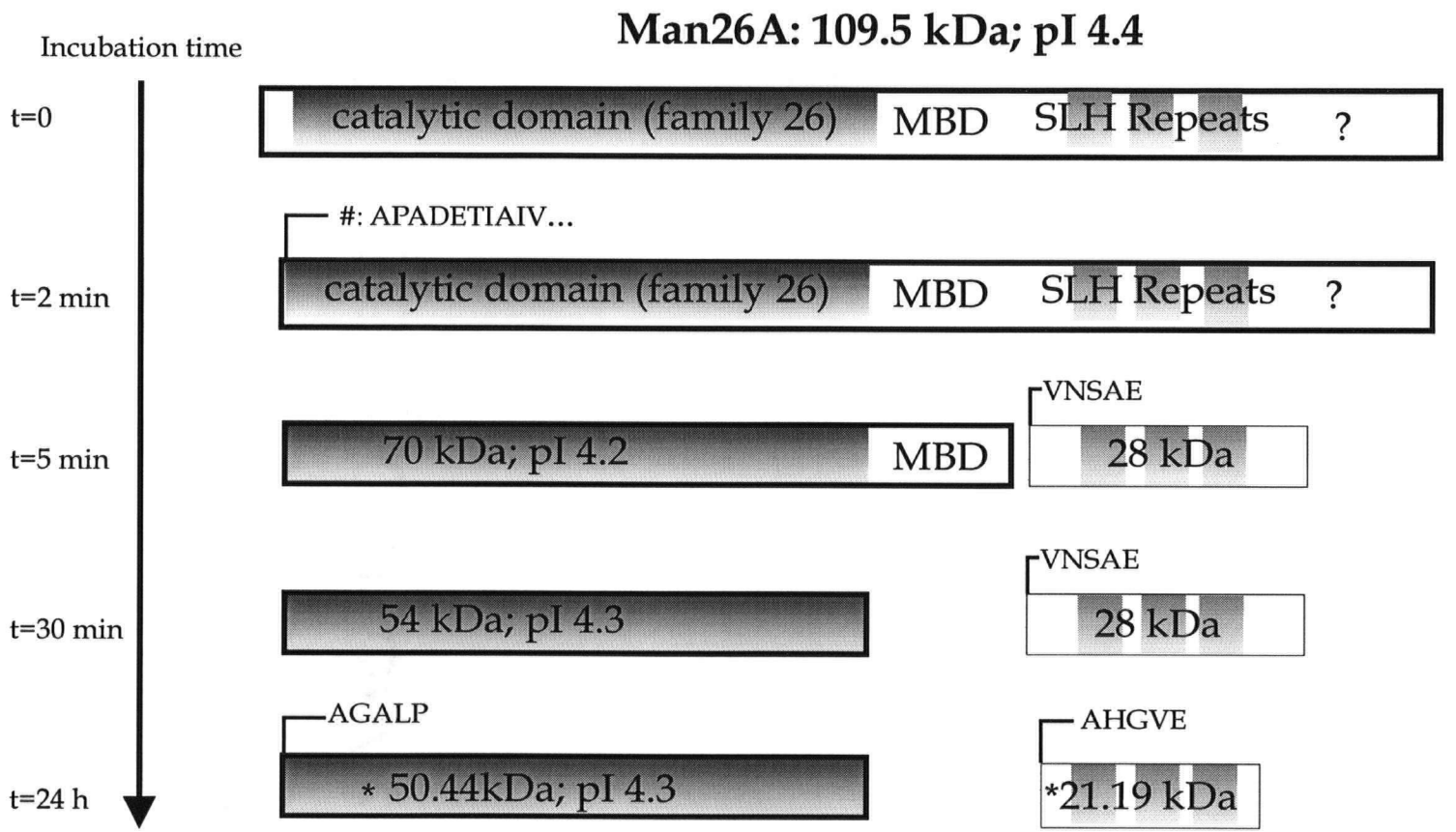


Figure 3.11: Schematic representation of proteolytic degradation of *C. fimi* mannanase Man26A. Recombinant Man26A was treated with *C. fimi* protease. The unprocessed enzyme and the fragments produced by proteolysis are represented as boxes and the digestion time is shown to the left of the boxes. The M_rs and calculated pIs are indicated in the boxes. The masses of fragments marked with * were determined by mass spectrometry. # indicates that the N-terminal sequence was obtained from Man26A, produced by *C. fimi*.

3.2.7 Localization of mannanase activity

The presence of a SLH domain in Man26A suggested that Man26A might be cell-bound. A *C. fimi* culture was grown with LBG as carbon source. After six days protein synthesis was inhibited by adding chloramphenicol to the culture before removing samples. Supernatant, washed cells and total culture samples were analyzed for mannanase activity by measuring azo-carob galactomannan hydrolysis (Figure 3.12; Section 2.15). More mannanase activity was found to be cell-bound, than in the supernatant. The highest activity was detected in the total culture samples.

These results agreed with the idea that the SLH-domain in Man26A attaches the mannanase to the cell surface. This attachment might be only transient, with subsequent release of some Man26A into the culture supernatant (Lemaire *et al.*, 1995). The activity in the supernatant might also be a consequence of proteolytic cleavage between the catalytic domain and the SLH domain, with release of only the catalytic domain into the culture supernatant (Section 3.2.6). To test *in vitro* binding of Man26A to *C. fimi* cell walls, Man26A was added to a peptidoglycan fraction, which was prepared from *C. fimi* cells as described by Lemaire *et al.* (1995). However, binding of Man26A to this cell wall preparation could not be detected.

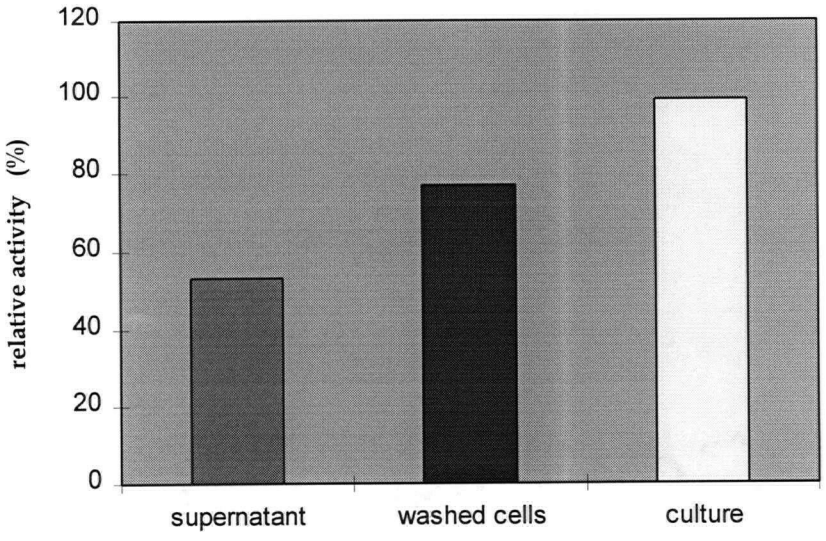


Figure 3.12: Analysis of mannanase activity from a 6 day old *C. fimi* culture. Supernatant, washed cells and total culture were analyzed on azo-carob galactomannan. The release of EtOH soluble oligosaccharides was measured at A_{590nm} after incubation at 37^0 C for 16 h. The activity from total culture was 100 %.

3.2.8 Is Man26A the only mannanase produced by *C. fimi* ?

Since recombinant Man26A was very unstable and susceptible to proteolysis (Figure 3.10), it was possible that all the polypeptides with mannanase activity detected in culture supernatants arose by proteolysis of Man26A. Zymograms of *C. fimi* culture supernatants and of Man26A proteolytic digests were compared. Supernatants from *C. fimi* cultures grown with ivory nut mannan (INM), LBG, CM-cellulose, and a mixture of CM-cellulose and LBG were compared to the patterns produced by recombinant Man26A, either untreated, treated with purified *C. fimi* protease, or treated with supernatant from the *C. fimi* culture with LBG (Figure 3.13). It was possible to obtain bands of mannanase activity similar to those found in *C. fimi* cultures, by treating recombinant Man26A with protease. It

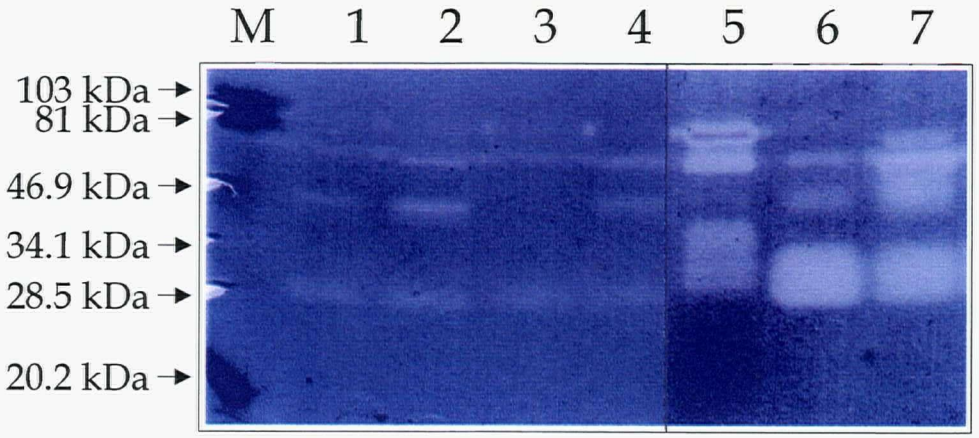


Figure 3.13: Non-reducing SDS-PAGE-zymogram to test the origin of multiple mannanase bands in *C. fimi* culture supernatants. Supernatants from four different *C. fimi* cultures were tested. The cultures were grown for 11 days at 30 °C in minimal medium supplemented with either ivory nut mannan (Lane 1), locust bean gum (Lane 2), carboxymethyl cellulose (Lane 3) or locust bean gum and carboxymethyl cellulose (Lane 4). Lane 5-7 recombinant Man26A (0.05µg), untreated (Lane 5), treated with *C. fimi* protease(Lane 6), or treated with culture supernatant from *C. fimi* culture grown on locust bean gum (Lane7).
The protease digestions were incubated for 1hr at 37 °C. To detect activity the zymogram was incubated in phosphate buffer at 37°C, before staining with Coomassie.

appeared, therefore, that Man26A was the only mannanase secreted by *C. fimi* involved in mannan degradation. The different polypeptides with mannanase activity present in supernatants from CM-cellulose and LBG-grown *C. fimi* cultures were not encoded by two or more independently regulated genes (Section 3.2); Rather they probably reflected different levels of secreted protease in the different cultures. It is believed that the secretion of protease by *C. fimi* and the protease susceptibility of Man26A has biological significance (Section 5)

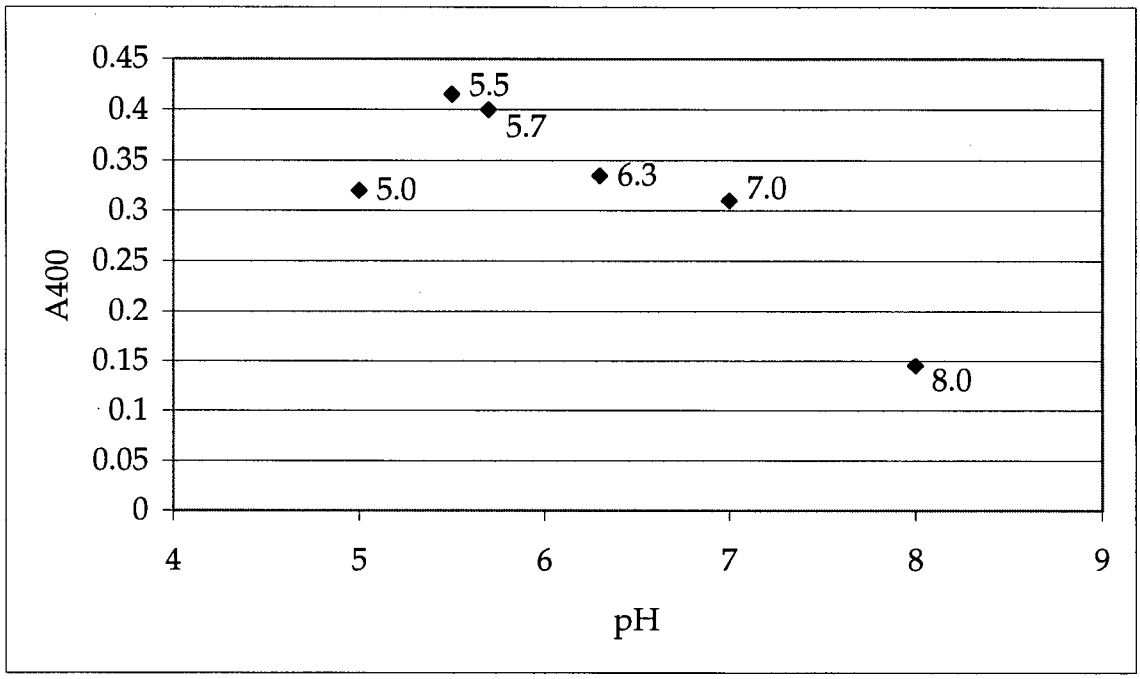
3.2.9 pH- and temperature optima, and kinetic parameters for Man26A

The pH- and temperature optima were determined for Man26A with *p*-nitrophenyl-mannobioside ($PNPM_2$) as substrate. This compound was synthesized by Man2A E519A catalyzed transglycosylation (Section 6.2.4). Optimal hydrolysis of $PNPM_2$ by Man26A occured at pH 5.5 and 42^0 C (Figure 3.14). Due to very slow hydrolysis rates, the reactions had to be incubated for more than 5 h to detect activity. Therefore these results also suggested that Man26A was stable for at least 2 hrs at 42^0 C (Section 2.18.5).

The kinetics of LBG hydrolysis by Man26A were analyzed for concentrations ranging from 0.01 mg/mL to 4.5 mg/mL. No higher concentations were used in order to avoid viscosity problems. The apparent value of k_{cat}/K_m = 1151 mL/(mg.min) was calculated from the initial slope of the plot of v *versus* s (Figure 3.15).

3.2.10 Screening of the *C. fimi* genomic DNA library for β-mannosidase

The *C. fimi* lambda-ZapII library was screened for expression of β-mannosidase activity using MUβMan as substrate. Out of 9×10^5 plaques screened, 2 plaques produced β-



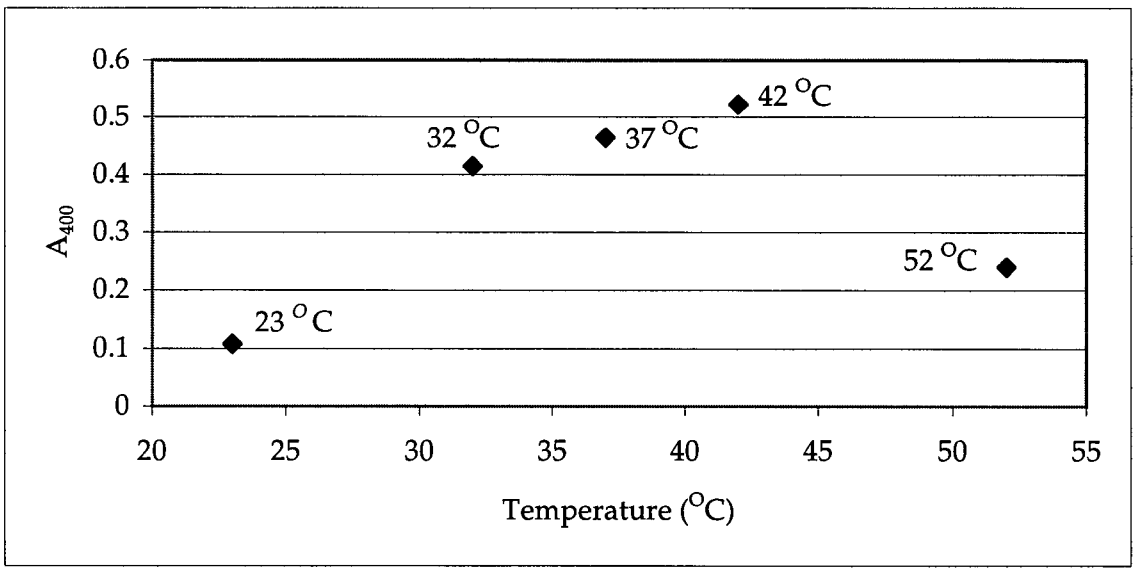


Figure 3.14: Panel A: pH optimum of Man26A. Panel B: Temperature optimum of Man26A. The hydrolysis of 1 mM $PNPM_2$ with 0.36 nmol Man26A was determined measuring A_{400}. 100 µL reactions were incubated for 5 h and stopped by the addition of 500 µL 1 M glycine, pH 10.9.

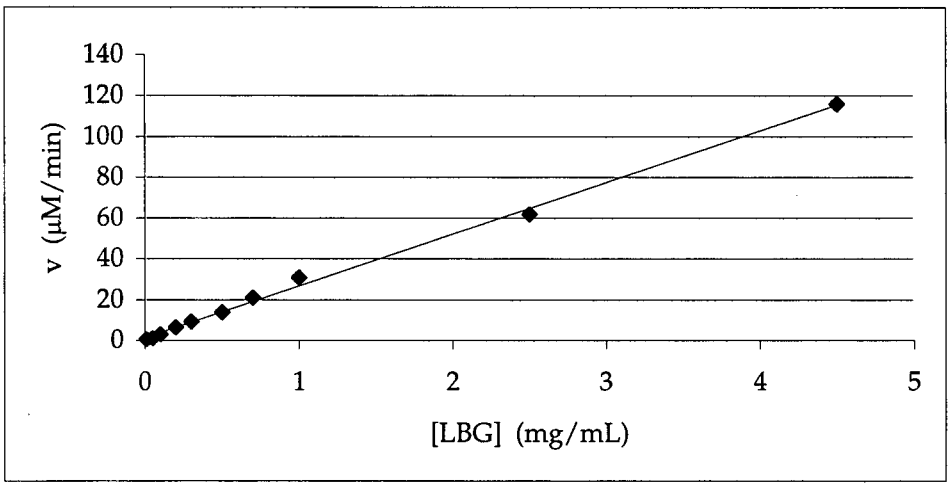


Figure 3.15: Steady-state kinetics for hydrolysis of locust bean gum (LBG) by Man26A. Substrate concentrations from 0.01 mg/mL to 4.5 mg/mL LBG were hydrolyzed with 22 nM Man26A in citrate buffer pH 5.5 at 37° C for 40 min. v is the rate of reducing sugars released per minute. The apparent k_{cat}/K_m = 1151.8 mL/mg.min was calculated from the initial slope of the plot v *versus* [LBG]. To avoid viscosity effects, only concentrations < 4.5 mg/mL were tested.

mannosidase activity. These 2 plaques however, did not produce any positive plaques in the secondary screening procedure, possibly a consequence of exposure to methylumbelliferyl. An alternative approach was used, whereby the *C. fimi* library was rescreened in its excised form, i.e. as *E. coli* colonies containing *C. fimi* genomic DNA fragments inserted into the pBluescript phagemids. The colonies were replica plated on plates containing MUβM to screen for activity. Two clones with β-mannosidase activity were isolated, CMadI and CMadII. The sizes of the genomic DNA inserts were found to be 2.6 kbp and 6.5 kbp for the plasmids pCMadI and pCMadII, respectively. Restriction mapping revealed that both inserts had a DNA region in common, which was confirmed by sequencing the 5′ ends of the inserts. The 5′ end of the pCMadI insert was found only 150 bp downstream from the beginning of the pCMadII insert. The shorter insert of pCMadI was sequenced. The subclones and primers used for sequencing are listed in Section 2.2 and 2.4. pCMadI contained the entire *C. fimi* β-mannosidase open reading frame.

3.2.11 Nucleotide sequence of the *C. fimi* β-mannosidase gene and its deduced amino acid sequence

The nucleotide sequence of its gene and the deduced amino acid sequence of the intracellular *C. fimi* β-mannosidase are shown in Figure 3.16. The open reading frame is 2526 bp long and encodes a protein of 842 amino acids. The calculated molecular weight of this protein is 94,960, with a calculated pI of 5.1.

Similarities of the *C. fimi* β-mannosidase to other proteins were analyzed (Figure 3.17). It shared significant sequence identity with mammalian β-mannosidases (EC 3.2.1.25), e.g. 32 % identical residues in a 667 aa long alignment with the caprine β-mannosidase (Figure 3.17; Figure 3.18) (Leipprandt *et al.*, 1996). High alignment scores were also obtained with

ACCGGTCAGTGTGCCACTCGATTGACTTACTTTATTCAGTTGTCTAAGAATGATCCCGGGTACCTGGCAGCCGACTCGG

CACCTCCCCGGAGGCTCCGAGCCGCCCGCGACGACGCGGGACGGACGCGCCGGACCCGCGGACCC**AGGAG**CGCGCGTCG

```
1/1                                             31/11
ATG ATC ACC CAG GAC CTC TAC GAC GGC TGG ACC CTC ACG GCC GTC TCC GGA CCC GTC CCG
M   I   T   Q   D   L   Y   D   G   W   T   L   T   A   V   S   G   P   V   P
61/21                                           91/31
GCC GAG CTG GCC GGC GTG CGC GTG CGG GCC CGC GTC CCC GGG ACG AGC CAC ACC GCG CTG
A   E   L   A   G   V   R   V   R   A   R   V   P   G   T   S   H   T   A   L
121/41                                          151/51
CTC GAC GAG GGC CTG ATC CCC GAC CCG TAC CTC GAC CGC AAC GAG GAC GTC CTC GCG TGG
L   D   E   G   L   I   P   D   P   Y   L   D   R   N   E   D   V   L   A   W
181/61                                          211/71
ATG CGG CGC ACC GAC TGG GCC TAC GAG CGC GAG CTC GTC CTC GAC CCG GCC GCC GCC GAC
M   R   R   T   D   W   A   Y   E   R   E   L   V   L   D   P   A   A   A   D
241/81                                          271/91
GAG CGC GTC GAC CTC GTC TTC GGC GGC ATC GAC ACC GTC GGG ACC GTC ACG TTC GAC GGC
E   R   V   D   L   V   F   G   G   I   D   T   V   G   T   V   T   F   D   G
301/101                                         331/111
CAC GAG CTC GGC CGC ACC GCC AAC CAG CAC CGG TCC TAC CGG TTC GAC GTG CGC GCC CTG
H   E   L   G   R   T   A   N   Q   H   R   S   Y   R   F   D   V   R   A   L
361/121                                         391/131
CTG CGC CCC GAC ACC CAG CGC CTG CGC GTG GAC CTG CGG GCC GCG ATC GTG CAC GCC GAG
L   R   P   D   T   Q   R   C   R   V   D   L   R   A   A   I   V   H   A   E
421/141                                         451/151
GCG GAG CGC GAG CGC CTC GGC CAC CGC CCG CTG GCC TAC CCG CAG CCG TTC AAC ATG GTC
A   E   R   E   R   L   G   H   R   P   L   A   Y   P   Q   P   F   N   M   V
481/161                                         511/171
CGC AAG ATG GCG TGC TCG TTC GGC TGG GAC TGG GGC CCC GAC CTG CAG ACC GCG GGC CTG
R   K   M   A   C   S   F   G   W   D   W   G   P   D   L   Q   T   A   G   L
541/181                                         571/191
TGG AAG CCG GTC CGC GTC GAG CGG TGG CGC ACC GCG CGG CTC GCG TCG GTG CGC ACG CAC
W   K   P   V   R   V   E   R   W   R   T   A   R   L   A   S   V   R   T   H
601/201                                         631/211
GTC ACG GTC GAC CCC GAC GGC ACC GGC CGG GTG CGC GTG CTC GTG GAC CTC GAG CGC TCG
V   T   V   D   P   D   G   T   G   R   V   R   V   L   V   D   L   E   R   S
661/221                                         691/231
GGC CTG CCC GGT GGC GAC GCG CCC GTC ACG CTG CGC GCC CGC GTC CTG TCG GCC GAC GTC
G   L   P   G   G   D   A   P   V   T   L   R   A   R   V   L   S   A   D   V
721/241                                         751/251
GCC GTC ACC GTG CCC GGC ACG GCC ACG TCG GCC GTC GTC GAG CTC GAG GTC CCC CGG GCA
A   V   T   V   P   G   T   A   T   S   A   V   V   E   L   E   V   P   R   A
781/261                                         811/271
CCC CTG TGG TGG CCG GTC GGC CAC GGC CCG CAG CCG CTG TCC GAC CTC ACC GTC ACG CTC
P   L   W   W   P   V   G   H   G   P   Q   P   L   S   D   L   T   V   T   L
841/281                                         871/291
GCG ACG CGC GAC GAC GAG CCG CTC GAC TCC TGG TCC CGG CGC ATC GGG TTC CGC ACG GTC
A   T   R   D   D   E   P   L   D   S   W   S   R   R   I   G   F   R   T   V
901/301                                         931/311
GAG GTC GAC ACG ACG CCC GAC GAG GAC GGG ACG CCG TTC ACG TTC CGC GTC AAC GGG CGG
E   V   D   T   T   P   D   E   D   G   T   P   F   T   F   R   V   N   G   R
```

961/321 991/331
CCG GTC TTC GTC AAG GGC GCC AAC TGG ATC CCC GAC GAC CAC CTG CTC ACC CGC ATC ACG
P V F V K G A N W I P D D H L L T R I T
1021/341 1051/351
CGC GAG CGC CTC GCG CAC CGG CTC GAC CAG GCC GTC GAG GCG AAC CTC AAC CTG CTG CGC
R E R L A H R L D Q A V E A N L N L L R
1081/361 1111/371
GTC TGG GGC GGC GGC ATC TAC GAG TCC GAG GAC TTC TAC GAC CTG TGC GAC GAG CGC GGC
V W G G G I Y E S E D F Y D L C D E R G
1141/381 1171/391
CTG CTC GTC TGG CAG GAC TTC CTC CTC GCG TGC GCG GCC TAC CCC GAG GAG CAG CCC ATC
L L V W Q D F L L A C A A Y P E E Q P I
1201/401 1231/411
TGG GAC GAG CTC GAG GCC GAG GCG CGC GAG AAC GTC GCC CGG CTC ACG CCG CAC GCC TCG
W D E L E A E A R E N V A R L T P H A S
1261/421 1291/431
CTC GTG CTG TGG AAC GGC GGC AAC GAG AAC CTC TGG GGC TTC ATG GAC TGG GGC TGG CCG
L V L W N G G N E N L W G F M D W G W P
1321/441 1351/451
CAG GAG CTC GAG GGC CGC ACG TGG GGC TAC CGG CTG GCG ACC GAG CTG CTC AAG GGC GTC
Q E L E G R T W G Y R L A T E L L K G V
1381/461 1411/471
GTC GCG GAG CTC GAC CCG ACG CGC CCG TAC GCC GAC GGC AGC CCG TAC TCC CCC GGC TTC
V A E L D P T R P Y A D G S P Y S P G F
1441/481 1471/491
GCG CTC GAC GAC GTC CAC CCG AAC GAC CCG GAC CAC GGC ACG CAC CAC GAG TGG GAG GTC
A L D D V H P N D P D H G T H H E W E V
1501/501 1531/511
TGG AAC CGC GTC GAC TAC TCC GCG TAC CGC GAC GAC GTG CCG CGG TTC TGC TCC GAG TTC
W N R V D Y S A Y R D D V P R F C S E F
1561/521 1591/531
GGC TTC CAG GGC CCG CCG ACG TGG TCG ACC CTC ACG CGT GCC GTC CGC GCC GAC GAC GGC
G F Q G P P T W S T L T R A V R A D D G
1621/541 1651/551
GGC CCG CTG ACC AAG GAC GAC CCG ACG TTC CTG CTG CAC CAG AAG GCC GAG GAC GGC AAC
G P L T K D D P T F L L H Q K A E D G N
1681/561 1711/571
GGC AAG CTC GAC CGC GGC CTC GCG CCG CAC CTG GGC GTG CCC GCC GGC TTC GTC GAC TGG
G K L D R G L A P H L G V P A G F V D W
1741/581 1771/591
CAC TGG GCG ACG CAG CTC AAC CAG GCC CGC GCC GTC GCG TTC GCG ATC GAG CAC TAC CGG
H W A T Q L N Q A R A V A F A I E H Y R
1801/601 1831/611
TCG TGG TGG CCC CGG ACC GCG GGC GCG ATC GTG TGG CAG CTC AAC GAC TGC TGG CCG GTG
S W W P R T A G A I V W Q L N D C W P V
1861/621 1891/631
ACG TCG TGG GCC GCG ATC GAC GGC GAC GAG CGG GTC AAG CCG CTG TGG CAC GCC CTG CGC
T S W A A I D G D E R V K P L W H A L R
1921/641 1951/651
CGC GCC TAC GCC CCG CGC CTG CTC ACC GTG CAG CCG CGC GAC GGC CGC GAC GAG CTC GCG
R A Y A P R L L T V Q P R D G R D E L A
1981/661 2011/671
GTC GTC AAC GAC ACG GGT GGC CTC TGG CAG GGC ACG GTC CGG CTC TCG CGG CGC ACG CTC
V V N D T G G L W Q G T V R L S R R T L

```
2041/681                                        2071/691
GAC GGT GCG ACG CTC GCC GAG GTC GAG CTC GGG CTG GCC GTC GGC GCG TGG TCG GTC GGC
D   G   A   T   L   A   E   V   E   L   G   L   A   V   G   A   W   S   V   G
2101/701                                        2131/711
CTG TTC GCG CTG CCC GAC GAG GTC GCC GCG CCC GAC GAC GCG GCC GGC GAG GTC CTG GTC
L   F   A   L   P   D   E   V   A   A   P   D   D   A   A   G   E   V   L   V
2161/721                                        2191/731
GTC GAC CTC GGG GAC GTC CGC ACG GTC CAC ACG TGG GCG CAG GAC GTC GAC CTG CGC CTC
V   D   L   G   D   V   R   T   V   H   T   W   A   Q   D   V   D   L   R   L
2221/741                                        2251/751
GAC CCC GAC CCG GTG TCC GCG ACC GTG TCG CCG CTG CAG GAC GGC TAC CGC GTC GAC GTG
D   P   D   P   V   S   A   T   V   S   P   L   Q   D   G   Y   R   V   D   V
2281/761                                        2311/771
ACC GCC CGG ACG TTC GCC CGC TCG GTC ACG CTG CAC GTG GAC CGC CTC GAC CCC GAC GCG
T   A   R   T   F   A   R   S   V   T   L   H   V   D   R   L   D   P   D   A
2341/781                                        2371/791
ACG GTC GAC GAC GCC CTC GTC GAC GTG CCG GCG GGC GAG ACG TTC TCG TTC CAC GTC CGC
T   V   D   D   A   L   V   D   V   P   A   G   E   T   F   S   F   H   V   R
2401/801                                        2431/811
ACC TCC GCA CGG TTC GAC GCC GCC GCC CTC ACG CGC AGC CCC GTG CTG CGC ACG GCG AAC
T   S   A   R   F   D   A   A   A   L   T   R   S   P   V   L   R   T   A   N
2461/821                                        2491/831
GAC GTG GTC GTG CCG CGC GGC GCG ACG GCA CCG GCG GGC GCC GAG CGT GAT CTC CAG CAG
D   V   V   V   P   R   G   A   T   A   P   A   G   A   E   R   D   L   Q   Q

2521/841      ------------------------>        <------------------------
TCC CGC TGA ACG GTT CGC GCC CGC CAC TAC GCT GGC GGG CGT GAC CCA CGG GAG GAA TTC
S   R   *

GAT ATC AAG CTT ATC GAT ACC GTC GAC CTC GAG GGG GCC CGT ACC AAT
```

Figure 3.16: Nucleotide sequence of *man2A* and deduced amino acid sequence of the protein. The putative -35 and -10 promoter regions, identified based on similarities to other *C. fimi* promoters, are underlined. The putative ribosome binding site is shown in bold. The sequences marked with arrows highlight possible trancription termination sequences.

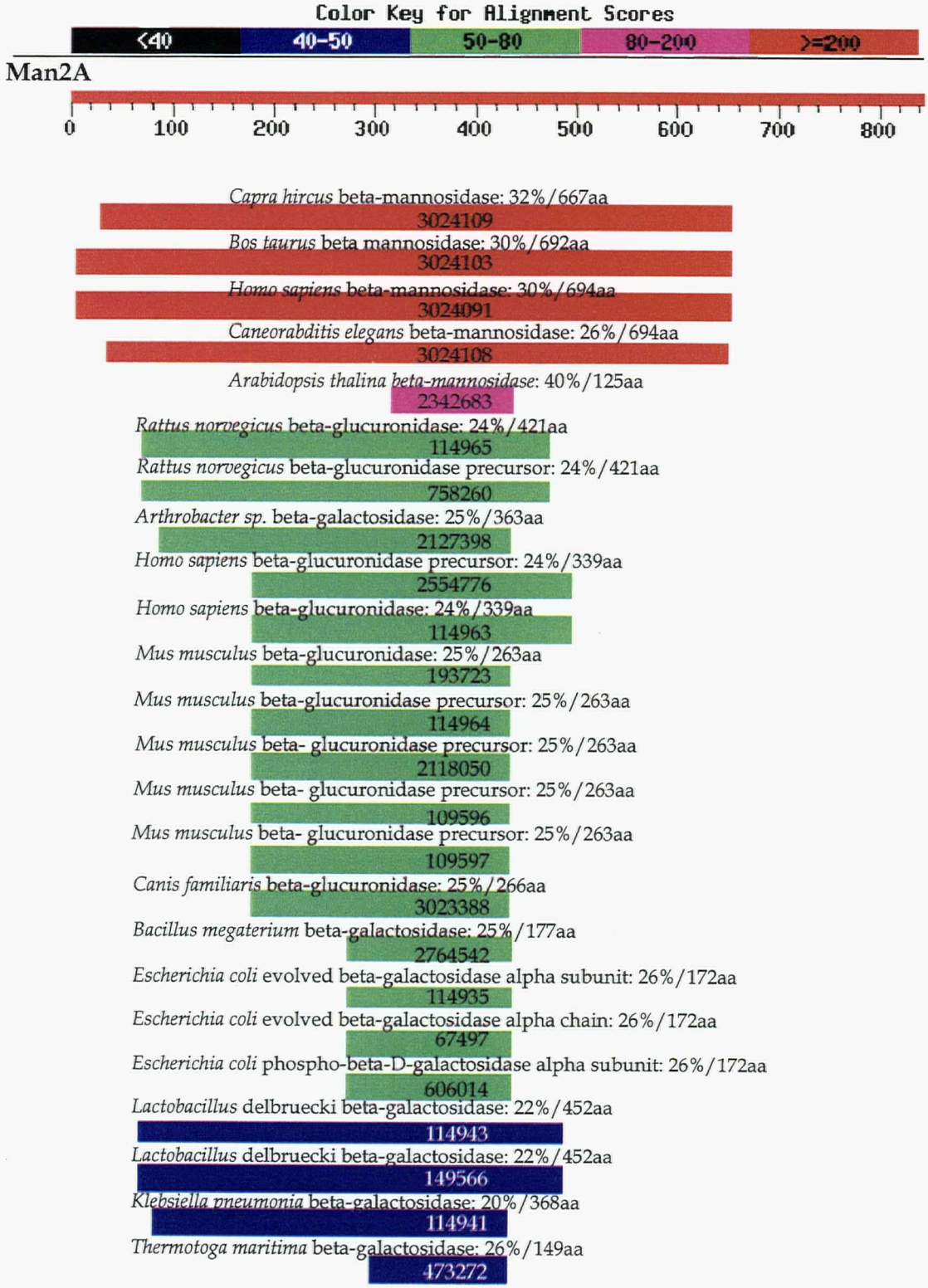


Figure 3.17: Results from protein sequence similarity search with Man2A. Each bar represents an amino acid sequence which is similar to the sequence of Man2A. The position of each bar is aligned for the homology with Man2A. The alignment score is colour-coded and the identities and alignment lengths are indicated for each protein sequence. The numbers in the or beside the bars are the sequence IDs.(Blast search; Altschul *et al.*, 1997).

```
Goat Mad     1...MLLRLLLLLAPCGAGFATEVVSISLRGNWKIHNGNGSLQLPAAVPGCVHSALFNKRI
Bovine Mad   1...MLLRLLLLLAPCGAGFATKVVSISLRGNWKIHSGNGSLQLPATVPGCVHSALFNKRI
Human Mad    1...MRLHLLLLLALCGAGTTAAELSYSLRGNWSICNGNGSLELPGAVPGCVHSALFQQGL
C.fimi Mad   1 ...MITQDLYDGWTLTAVSGP.VPAELAG.........VRVRARVPGTSHTALLDEGL
consensus    1          L                 L G                VPG  H AL

Goat Mad    58 IKDPYYRFNNLDYRWIALDNWTYIKKFKLHSDMSEWNKVNLVFEGIDTVAVVLLNSVPIG
Bovine Mad  58 IKDPYYRFNNLDYRWIALDNWTYIKKFKLHSDMSTWSKVNLVFEGIDTVAVVLLNSVPIG
Human Mad   58 IQDSYYRFNDLNHRWVSLDNWTYSKEFKIPFEISKWQKVNLILEGVDTVSKILFNEVTIG
C.fimi Mad  46 IPDPYLDRNEDVLAWMRRTDWAYERELVL.DPAAADERVDLVFGGIDTVGTVTFDGHELG
consensus   61 I D Y   N     W     W Y              V L   G DT          G

Goat Mad   118 KTDNMFRRYSFDITHMVKA.VNIIEVRFQSPVIYANQRS...ERHTAYWVPPNCPPPVQD
Bovine Mad 118 KTDNMFRRYSFDITHTVKA.VNIIEVRFQSPVVYANQRS...ERHTAYWVPPNCPPPVQD
Human Mad  118 ETDNMFNRYSFDITNVVRD.VNSIELRFQSAVLYAAQQS...KAHTXYQVPPDCPPLVQK
C.fimi Mad 105 RTANQHRSYRFDVRALLRPDTQRVRVDLRAAIVHAEAER.....ER.LGHRPLAYPQP..
consensus  121  T N    Y FD                         A               P   P

Goat Mad   174 GECHVNFIRKMQCSFGWDWGPSFPTQGIWKDVRIEAYNICHLNYFMFTPIYDNYMETWNL
Bovine Mad 174 GECHVNFIRKMQCSFGWDWGPSFPTQGIWKDVRIEAYNVCHLNYFMFTPIYDNYMKTWNL
Human Mad  174 GECHVNFVRKEQCSFSWDWGPSFPTQGIWKDVRIEAYNICHLNYFTFSPIYDKSAQEWNL
C.fimi Mad 157 ....FNMVRKMACSFGWDWGPDLQTAGLWKPVRVERWRTARLASVRTHVTVD...PDGTG
consensus  181      N  RK  CSF WDWGP   T G WK VR E      L         D

Goat Mad   234 KIESSFDVVSSKLVSGEAIVAIPELNIQQRNNIELRHGE....R.TVKLFVKIDKAVIVE
Bovine Mad 234 KIESSFDVVSSKLVSGEAIVAIPELNIQQTNNIELQHGE....R.TVELFVKIDKAIIVE
Human Mad  234 EIESTFDVVSSKPVGGQVIXAIPKLQTQQTYSIELQPGK....R.IVELFVNISKNITVE
C.fimi Mad 210 RVRVLVDLERSGLPGGDAPVTLRARVLS.ADVAVTVPGT......ATSAVVELEVPRAP.
consensus  241       D       G                      G          V

Goat Mad   289 TWWPHGHGNQTGYDMTVTFEL.DGGLRFEKSAKVYFRTVELVEEPIQ.NSP..GLTFYFK
Bovine Mad 289 TWWPHGHGNQTGYNMSVIFEL.DGGLRFEKSAKVYFRTVELVEEPIQ.NSP..GLSFYFK
Human Mad  289 TWWPHGHGNQTGYNMTVLFEL.DGGLNIEKSAKVYFRTVELIEEPIK.GSP..GLSFYFK
C.fimi Mad 262 LWWPVGHGPQPLSDLTVTLATRDDEPLDSWSRRIGFRTVEVDTTPDEDGTP.....FTFR
consensus  301  WWP G G Q      V     D       S    FRTVE    P     P     F F

Goat Mad   345 INGLPIFLKGSNWIPADSFQ.DRVTSDMLRLLLQSVVDANMNALRVWGGGIYEQDEFYEL
Bovine Mad 345 INGLPIFLKGSNWIPADSFQ.DRVTSAMLRLLLQSVVDANMNALRVWGGGVYEQDEFYEL
Human Mad  345 INGFPIFLKGSNWIPADSFQ.DRVTSELLRLLLQSVVDANMNTLRVWGGGIYEQDEFYEL
C.fimi Mad 317 VNGRPVFVKGANWIPDDHLL.TRITRERLAHRLDQAVEANLNLLRVWGGGIYESEDFYDL
consensus  361  N  P F KG NWIP D     R T   L   L   V AN N LRVWGGG YE   FYL

Goat Mad   404CDELGIMIWQDFMFACALYPTDEDFMDSVREEVTHQVRRLKSHPSIITWSGNNENEAALM
Bovine Mad 404CDELGIMIWQDFMFACALYPTDKDFMDSVREEVTHQVRRLKSHPSIITWSGNNENEAALM
Human Mad  404CDELGIMVWQDFMFACALYPTDQGFLDSVTAEVAYQIKRLKSHPSIIIWSGNNENEEALM
C.fimi Mad 376CDERGLLVWQDFLLACAAYPEEQPIWDELEAEARENVARLTPHASLVLWNGGNEN.LWGF
consensus  421CDE G   WQDF  ACA YP      D    E      RL  H S   W G NEN

Goat Mad   464 MGWYDTKPGYLHTYIKDYVTLYVKNIRTIVLEGDQTRPFIISSPTNGAKTTAEGWLSPNP
Bovine Mad 464 MGWYDTKPGYLQTYIKDYVTLYVKNIRTIVLEGDQTRPFITSSPTNGAKTIAEGWLSPNP
Human Mad  464 MNWYHISFTDRPIYIKDYVTLYVKNIRELVLAGDKSRPFITSSPTNGAETVAEAWVSQNP
C.fimi Mad 435 MDWGWPQELEGRTWGYRLATELLK...GVVAELDPTRPYADGSP..YSPGFALDDVHPND
consensus  481 M W               T   K     V   D  RP    SP       A
```

```
Goat Mad    524 YDLNYGDVHFYD.YMSDCWNWRTFPKARFVSEYGYQSWPSFSTLEKVSSEED.WSYES.
Bovine Mad  524 YDLNYGDVHFYD.YVSDCWNWRTFPKARFVSEYGYQSWPSFSTLEKVSSEED.WSYRS.
Human Mad   524 NSNYFGDVHFYD.YISDCWNWKVFPKARFASEYGYQSWPSFSTLEKVSSTED.WSFNS.
C.fimi Mad  490 PD..HGTHHEWEVWNRVDYSAYRDDVPRFCSEFGFQGPPTWSTLTRAVRADDGGPLTKD
consensus   541      G  H                  RF SE G Q  P  STL       D

Goat Mad    580 .SFALHRQHLINGNSEMLQQIELHFKLPNSA.............DQLRRFKDTLYLTQV
Bovine Mad  580 .SFALHRQHLINGNNEMLHQIELHFKLPNST.............DQLRRFKDTLYLTQV
Human Mad   580 .KFSLHRQHHEGGNKQMLYQAGLHFKLPQST.............DPLRTFKDTIYLTQV
C.fimi Mad  548 PTFLLHQKAE.DGNGKLDRGLAPHLGVPAGF.................VDW.HWATQL
consensus   601   F LH      GN         H   P                        TQ

Goat Mad    625 MQAQCVKTETEFYRRSRNEIVDG.KGHTMGALYWQLNDIWQAPSWSSLEYGGKWKMLHYF
Bovine Mad  625 MQAQCVKTETEFYRRSRSEIVNG.KGHTMGALYWQLNDIWQAPSWSSLEYGGKWKMLHYF
Human Mad   625 MQAQCVKTETEFYRRSRSEIVDQ.QGHTMGALYWQLNDIWQAPSWASLEYGGKWKMLHYF
C.fimi Mad  587 NQARAVAFAIEHYRSWWP........RTAGAIVWQLNDCWPVTSWAAIDGDERVKPLWHA
consensus   661  QA  V    E YR             T GA  WQLND W   SW         K L

Goat Mad    684 ARRFFAPLLPVGFEDKD..VLFIYGVSDLPSDHQMMLTVRVHTWSSLELVCSELTNPFVM
Bovine Mad  684 ARHFFAPLLPVGFEDKD..MLFIYGASHLHSDQQMMLTVRVHTWSSLELVCSESTNPFVI
Human Mad   684 AQNFFAPLLPVGFENEN..TFYIYGVSDLHSDYSMTLSVRVHTWSSLEPVCSRVTERFVM
C.fimi Mad  639 LRRAYAPRLLTVQPRDG..RDELAVVNDTG..GLWQGTVRLSRRTLDGATLAEVELGLAV
consensus   72        P L                          VR

Goat Mad    742 KAGESVVLYSKPVPELLKGCPGCTRQSCVVSFYLSTDGELLSPINYHFLSSLKNAKGLHK
Bovine Mad  742 KAGESVLLYTKPVPELLKGCPGCTRQSCVVSFYLSTDGELLSPINYHFLSSLKNAKGLHK
Human Mad   742 KGGEAVCLYEEPVSELLRRCGNCTRESCVVSFYLSADHELLSPTNYHFLSSPKEAVGLCK
C.fimi Mad  695 GAWSVG...LFALPDEVAAPDDAAGEVLVVDLGDVR........TVHTWAQDVDLR.LDP
consensus   781                                VV               H         L

Goat Mad    802 ANITATISQQGNTFVFDLKTSAVAPFVWLDVGS.IPGRFSDNGFLMTEKTRTVFFYPWKP
Bovine Mad  802 ANITATISQQGDTFVFDLKTSAVAPFVWLDVGS.IPGRFSDNGFLMTEKTRTVFFYPWKP
Human Mad   802 AQITAIISQQGDIFVFDLETSAVAPFVWLDVGS.IPGRFSDNGFLMTEKTRTILFYPWEP
C.fimi Mad  743 DPVSATVSPLQDGYRVDVTARTFARSVTLHVDRSTPTRRSTTPSSTCRRARRSRSTSAPP
consensus   841     A  S        D      A  V L V    P R S          R       P

Goat Mad    861 TSKSELEQSFHVTSLADTY.
Bovine Mad  861 TSKSELEQSFHVTSLADTY.
Human Mad   861 TSKNELEQSFHVTSLTDIY.
C.fimi Mad  803 H..................
consensus   901
```

Figure 3.18: Alignment of family 2 β-mannosidases. Goat Mad: *Capra hircus* β-mannosidase, Bovine Mad: *Bos taurus* β-mannosidase, Human Mad: *Homo sapiens* β-mannosidase, C.fimi Mad: *Cellulomonas fimi* β-mannosidase. Identities in all five protein sequences are highlighted and indicated in capital letters in the consensus sequence. The putative catalytic residues are underlined. Alignment was performed with Clustal W (Thompson *et. al.*, 1994).

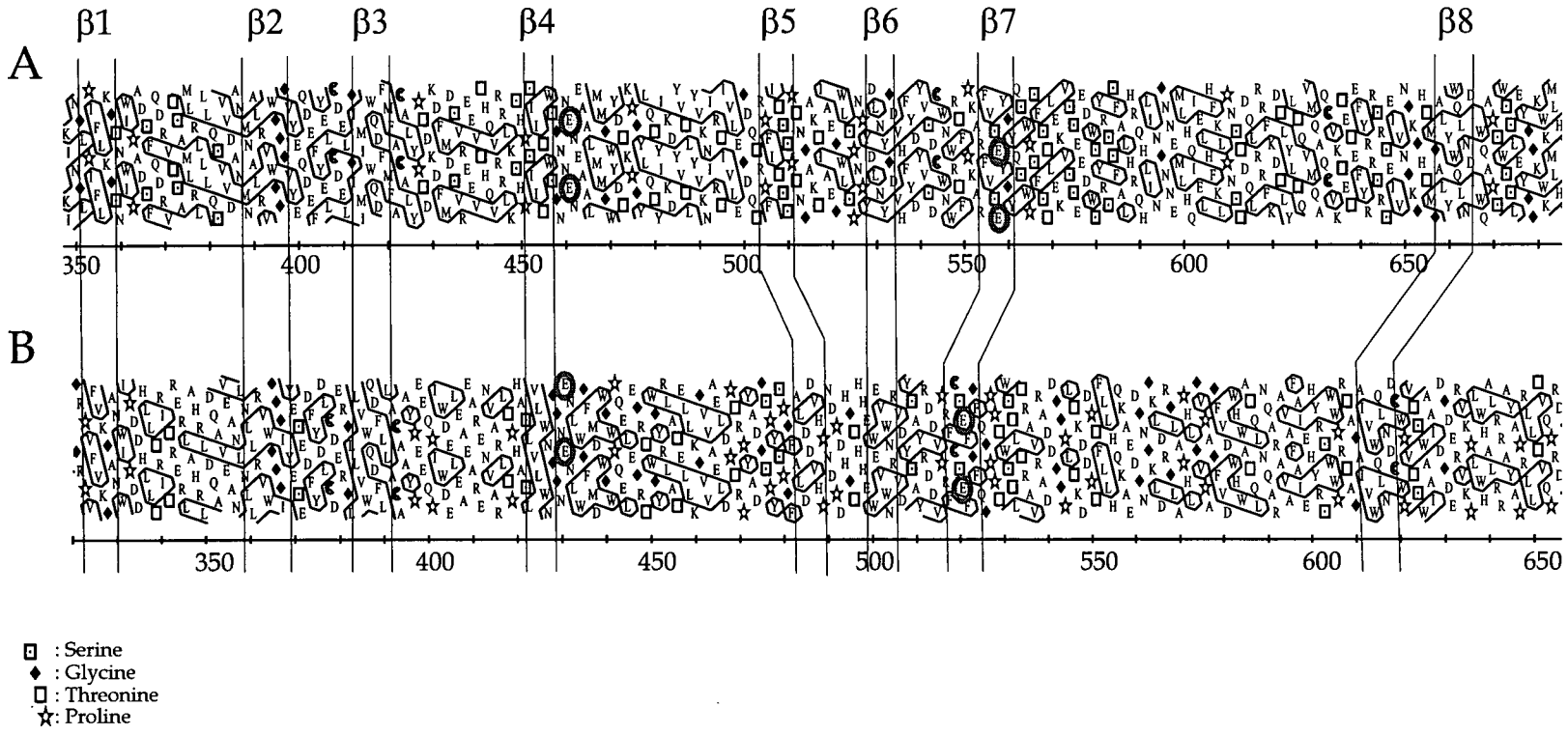


Figure 3.19: Comparison of hydrophobic cluster analysis plots of the catalytic region of *Bos taurus* β-mannosidase; Panel A (Durand *et al.* 1997) and *Cellulomonas fimi* β-mannosidase; Panel B. The 8 β-strands from the $(\beta/\alpha)_8$ fold are indicated. The putative catalytic residues are highlighted.

human and bovine β-mannosidases (Alkhayat *et al.*, 1988; Chen *et al.*, 1995). The *C. fimi* enzyme also shared sequence identity with β-glucuronidases and β-galactosidases, but the alignment scores were much lower (Figure 3.17). All of the enzymes sharing identity with the *C. fimi* β-mannosidase were members of glycosyl hydrolase family 2 (Henrissat 1991). Therefore, in accordance with the newly proposed nomenclature, the *C. fimi* β-mannosidase was named Man2A (or Cf Man2A) (Henrissat *et al.*, 1998). All the family 2 members cleave sugars from the non-reducing end of the substrate with retention of the anomeric configuration (Gebler *et al.*, 1992). Family 2 is a member of clan GH-A, which represents a group of more than 200 distantly related protein sequences from families 1, 2, 5, 10, 17, 26, 30, 35, 39 and 42. Analysis of known 3D structures of GH-A members revealed that their catalytic domains adopt similar $(\beta/\alpha)_8$ folds, with the acid/base catalyst and the catalytic nucleophile at the C-terminal end of strands β4 and β7, respectively (Henrissat *et al.*, 1996). The same folding motif, the $(\beta/\alpha)_8$ barrel, and the position of the catalytic residues were predicted by HCA analysis for the bovine β-mannosidase (Durand *et al.*, 1997). The similarities between the HCA plots of the bovine and the *C. fimi* β-mannosidase were used to predict the positions of the acid/base catalyst as Glu429 and catalytic nucleophile as Glu519 in Man2A (Figure 3.19).

For each glycosyl hydrolase family, signature consensus patterns were defined in order to reduce ambiguity in classification of new enzymes. The mannosidase sequences however, do not agree completely with the consensus patterns as defined for family 2 (Henrissat *et al.*, 1996). One of the conserved regions in family 2 is around the general acid/base catalyst (Gebler *et al.*, 1992). The second conserved region, that was defined as a signature pattern, is located about sixty residues upstream of the acid/base catalyst. These

two signature patterns are shown below for Cf Man2A, with the residues highlighted that do not agree with the family 2 signature sequences (26 %):

consensus sequence 1: N L L R V W G G G I Y E S E D F Y D L C D E R G L L V W

consensus sequence 2: L T P H A S L V L W N G G N E

This suggested that the β-mannosidases might be a subfamily of family 2.

3. 2.12 Sub-cloning and expression of *man2A*

The gene *man2A* was sub-cloned into the expression vector pET28a(+) (Novagen) (Figure 3.20). The resulting clone, pET28Mad, was used to produce recombinant Man2A with a C-terminal H_6 tag in *E. coli* BL21(DE3) cells (Section 2.8). The protein was purified by MCAC. Up to 300 mg of the purified cytoplasmic protein were obtained with a purity estimated to be >96 % (Figure 3.21). The N-terminal amino acid sequence of the purified recombinant protein was identical to that of the partially purified *C. fimi* β-mannosidase (Section 2.9). Therefore, the recombinant, purified enzyme was used for further analyses.

3.2.13 pH optimum of *C. fimi* recombinant Man2A WT

PNPM hydrolysis by Man2A was measured at pHs ranging from 5.5 to 8.2. The pH optimum was 7.0 (Section 2.18.1; Figure 3.22 A).

Figure 3.20 (following page): Generation of pET28aMad, encoding the *C. fimi* β-mannosidase, Man2A with a C-terminal His_6 tag. A *Nco* I restriction site was introduced at the 5' end and a *Not* I restriction site was introduced at the 3' end of *man2A*. Only the restriciton sites that were relevant for the sub-cloning steps are shown.

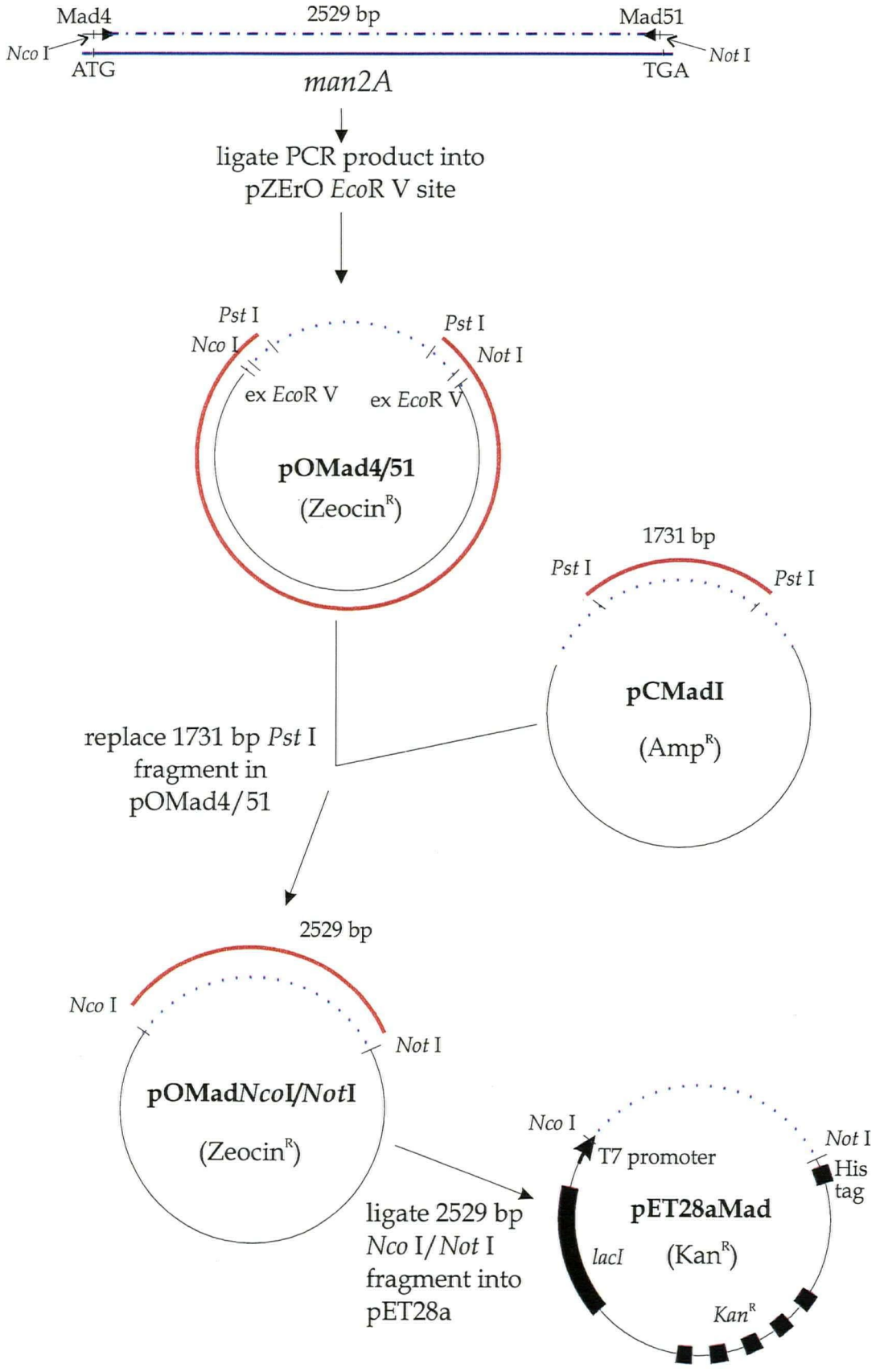
——► : Mad4 (*Nco* I) and Mad51 (*Not* I) PCR primers

..... : *man2A* or *man2A* derived plasmid insert

-·-·-· : PCR product

——— : *man2A* PCR template

——— : DNA fragments used for the sub-cloning steps. The DNA fragments were obtained by restriction endonuclease digestion, separation on agarose gels and purification by gel extraction.


Mad4
2529 bp
Mad51
Nco I
Not I
ATG
TGA
man2A
ligate PCR product into
pZErO EcoR V site
Pst I
Pst I
Nco I
Not I
ex EcoR V
ex EcoR V
pOMad4/51
(Zeocin^R)
1731 bp
Pst I
Pst I
pCMadI
(Amp^R)
replace 1731 bp Pst I
fragment in
pOMad4/51
2529 bp
Nco I
Not I
pOMadNcoI/NotI
(Zeocin^R)
ligate 2529 bp
Nco I/Not I
fragment into
pET28a
Nco I
Not I
T7 promoter
His
tag
pET28aMad
lacI
(Kan^R)
Kan^R

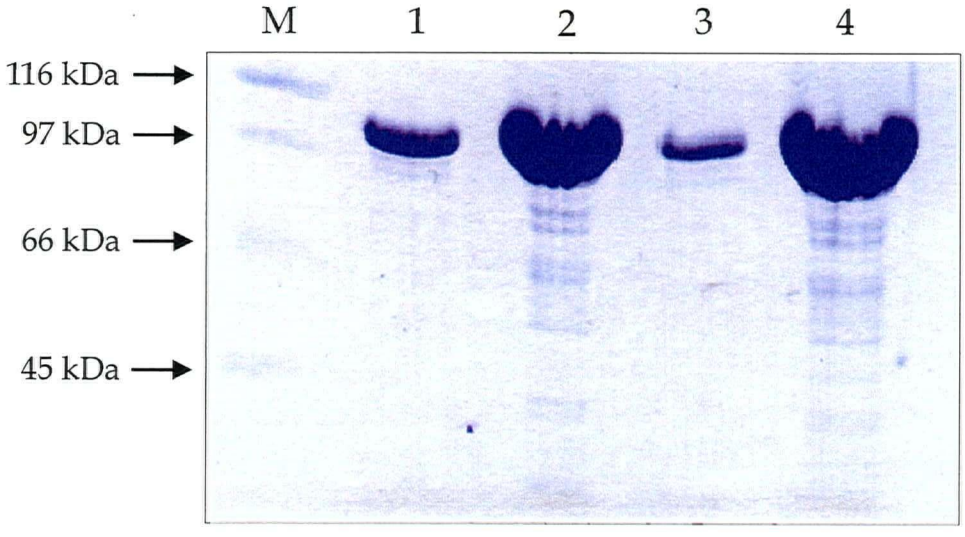


Figure 3.21: *C. fimi* β-mannosidase, Man2A WT and Man2A E519A, purified by MCAC, were separated on a 7.5% SDS-PAG (Phast System, Pharmacia) and stained with Coomassie. Lane M: molecular size standards, Lane 1: 1 μg Man2A E519A, Lane 2: 10 μg Man2A E519A, Lane 3: 1 μg Man2A and Lane 4: 10 μg Man2A.

3.2.14 Temperature optimum and thermostability of recomabinant Man2A WT

PNPM hydrolysis was measured at temperatures ranging from 23^0 C to 65^0 C. The rates were calculated from the ΔA_{400nm}/min at 190 s after the addition of enzyme. The fastest hydrolysis rate was obtained at 55 ^{0}C. This hydrolysis rate was therefore set as 100 %. The activities of Man2A WT at other temperatures were expressed as values relative to the hydrolysis rate obtained at 55^0 C.

The thermostability of Man2A WT was determined by incubating the enzyme at temperatures ranging from 23^0 C to 55^0 C (Section 2.18.5). During the 4 hour incubation samples were removed and PNPM hydrolysis rate was assayed at 23^0 C. A 3D representation of relative hydrolysis rates as a function of time and temperature is shown in Figure 3.22, with hydrolysis rates at 55^0 C and t = 190 s corresponding to 100 %. The optimum temperature was defined as the temperature at which the enzyme displayed highest hydrolysis rates and thermostability for at least 2 hours, which was 37^0 C for Man2A WT. The half-life of Man2A at 37^0 C was calculated to be 27.3 h.

3.2.15 Steady-state kinetic parameters for hydrolysis of PNPM by Man2A

The enzyme (4.2 nM) was inhibited by substrate concentrations >400 μM; therefore values for the apparent k_{cat} = 167 s^{-1}, K_m = 0.3 mM and k_{cat}/K_m = 501 $s^{-1}.mM^{-1}$ were only estimates (Figure 3.24). Substrate inhibition can result if a second substrate molecule binds to the enzyme-substrate complex, ES, to produce an inactive complex, SES (Cornish-Bowden, 1979):

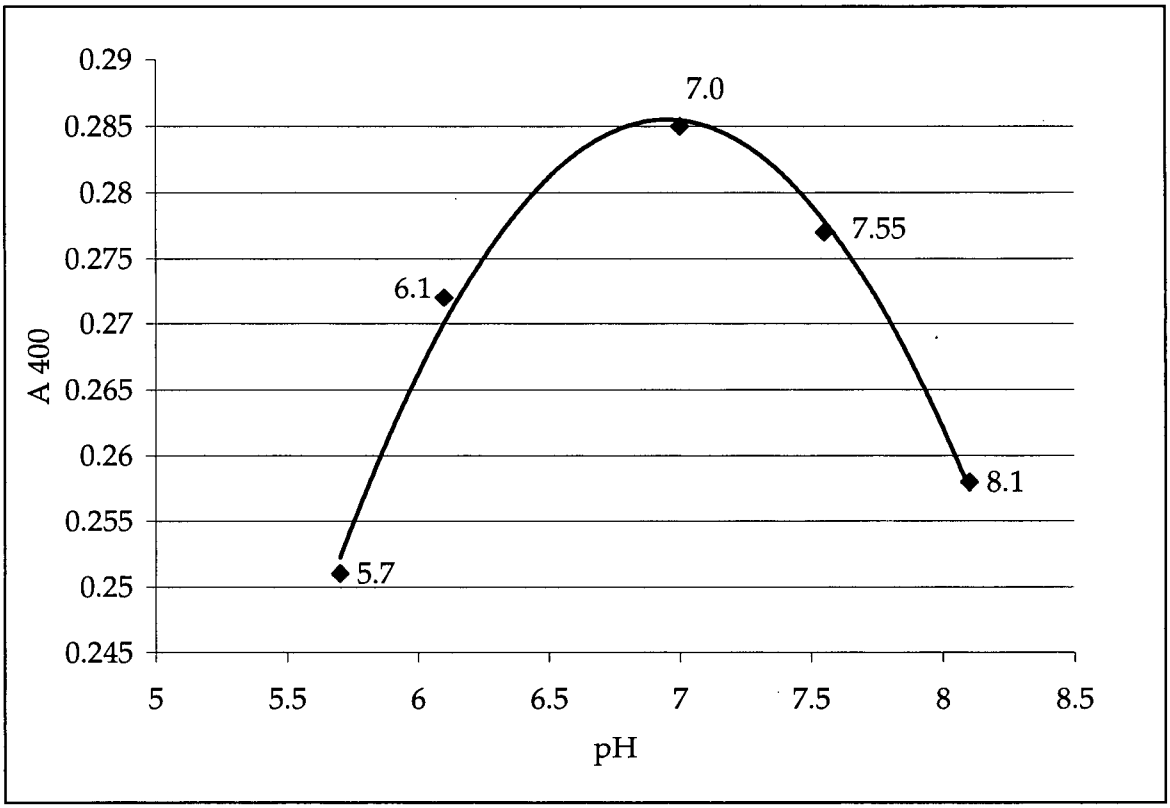


Figure 3.22: pH optimum of Man2A. The release of 4-nitrophenol was measured at A_{400} after stopping the reactions with 1 M glycine pH 10.9.

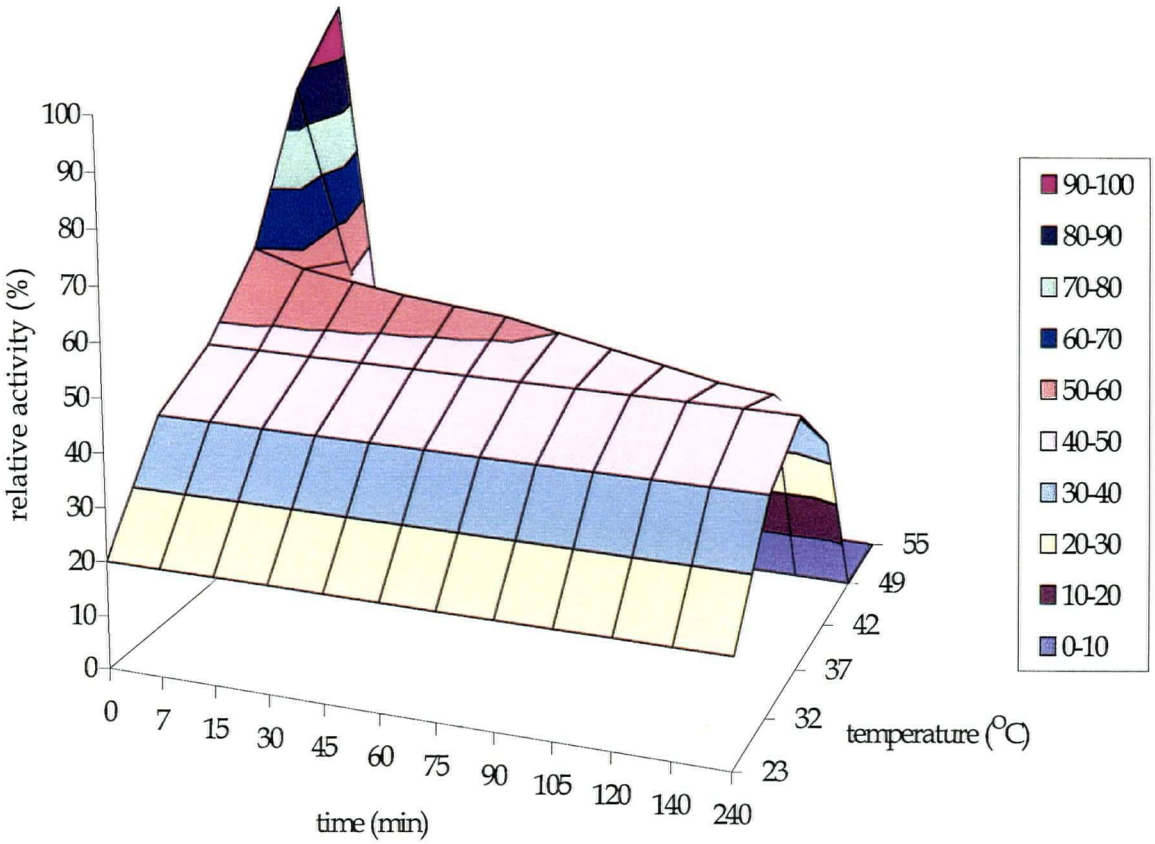


Figure 3.23: 3D plot of Man2A temperature optimum *versus* temperature stability measured as the relative activity on PNPM. The temperature optimum was determined by PNPM hydrolysis at temperatures from 23 °C to 60 °C. Temperature stability was monitored by incubating Man2A for 4 h at temperatures from 23 °C to 55 °C. Stability was examined by removing aliquots and measuring hydrolysis of PNPM at 23 °C. The activity of Man2A at 55 °C was set as 100 %.

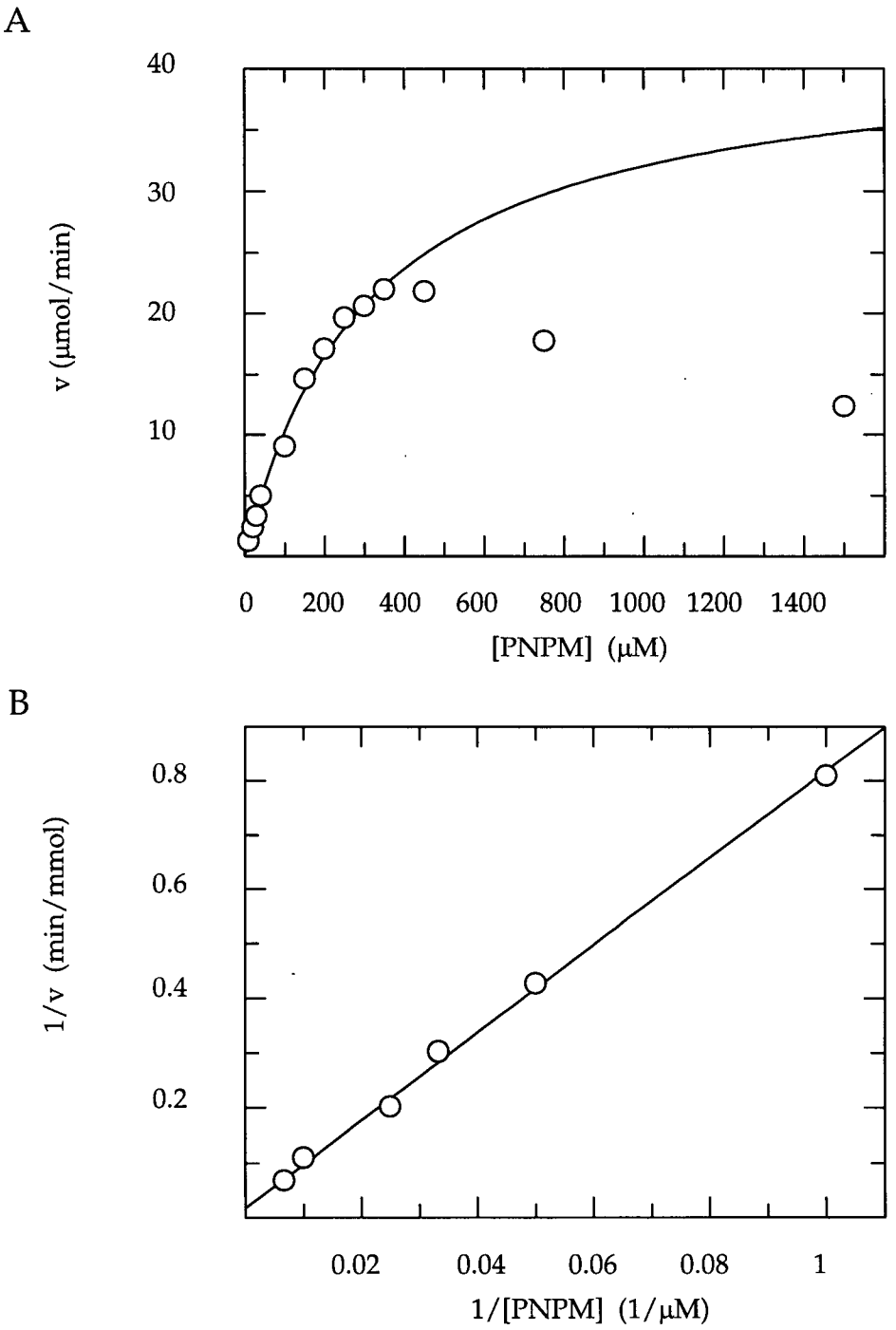
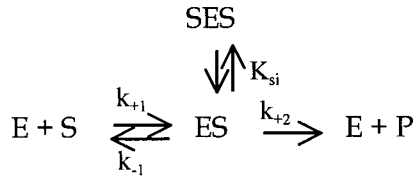


Figure 3.24: Steady-state kinetics of PNPM hydrolysis by Cf Man2A. Panel A: Kinetic data obtained from incubation of 4.2 nM Man2A with PNPM (10 µM to 1.5 mM). The curve shown was fitted to apparent Km = 0.3 mM and vmax = 42 mmol/min, which were calculated from the double reciprocal replot of hydrolysis initial rates with 10 µM to 150 µM PNPM (Panel B).



Man2A WT did not hydrolyze *p*-nitrophenyl-α-mannoside, *p*-nitrophenyl-β-N-acetylglucoside, *p*-nitrophenyl-β-glucoside, *p*-nitrophenyl-β-xyloside, *p*-nitrophenyl-β-cellobioside or *p*-nitrophenyl-β-gentiobioside. Low activity was detected on *p*-nitrophenyl -β-galactoside. Man2A WT was also able to slowly hydrolyze Man-β-1,4-GlucNAc into mannose and N-acetyl glucosamine, as was determined by FACE analysis (data not shown).

3.2.16 Size exclusion chromatography of Man2A

Size exclusion chromatograpy was used to determine the mass of native Man2A. Ferritin (440 kDa), catalase (233 kDa), aldolase (158 kDa) and ovalbumin (50 kDa) were used as molecular mass standards. The log of M_r was plotted against ve/vo (elution volume/void volume) (Figure 3.25). The M_r of Man2A was calculated to be 1.03×10^4 Da, indicating that it was a monomer.

3.2.17 Mannan and mannooligosaccharide hydrolysis by Man26A and Man2A

Mannan, galactomannan and five manno-oligosaccharides were subjected to hydrolysis by *C. fimi* β-mannosidase, Man2A, β-mannanase, Man26A, or by both enzymes (Section 2.18.7). The released products were analyzed by fluorophore assisted carbohydrate electrophoresis (FACE) using glucose (α-1,4 linked) and mannose (β-1,4 linked) oligosaccharides as standards (Figures 3.26 and 3.27). Man26A cleaved randomly within the β-1,4 linked backbone of galactomannan, releasing oligosaccharides of different sizes. This

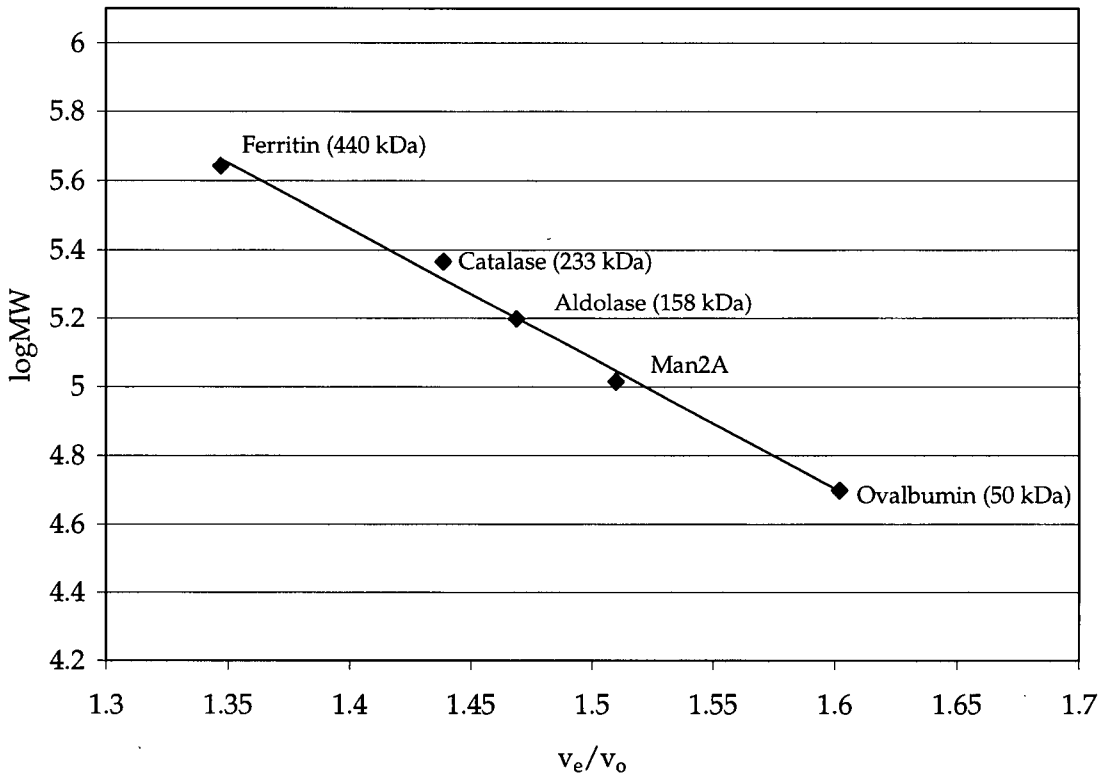


Figure 3.23: Size exclusion chromatography of Man2A. The column was calibrated with ferritin (440 kDa), catalase (233 kDa), aldolase (158 kDa) and ovalbumin (50 kDa). The void volume (v_o) was determined with Dextran Blue. Fractogel® TSK HW-55S was used as matrix.

cleavage pattern was characteristic of an endo-acting enzyme. Man2A, however, was unable to hydrolyze galactomannan into oligosaccharides, but released low quantities of mannose. Man26A in combination with Man2A produced a mix of oligosaccharides from galactomannan, very similar to the products released by Man26A alone, the only difference being the complete hydrolysis of the disaccharide (Figure 3.26 A). Unsubstituted mannan (ivory nut mannan) was hydrolyzed by Man26A into mainly mannotriose, mannobiose and mannose, whereas Man2A produced significant amounts of mannose as the only product. Both enzymes together released mannose as the only product (Figure 3.26 B).

Mannose was the major product released from mannohexaose, -tetraose, -triose and -biose by Man2A (Figure 3.27), whereas Man26A hydrolyzed mannohexaose, -pentaose and -tetraose into mannobiose and mannose. Hydrolysis of mannotriose into mannobiose and mannose by Man26A was slow, and mannobiose was not hydrolyzed at all (Figure 3.27).

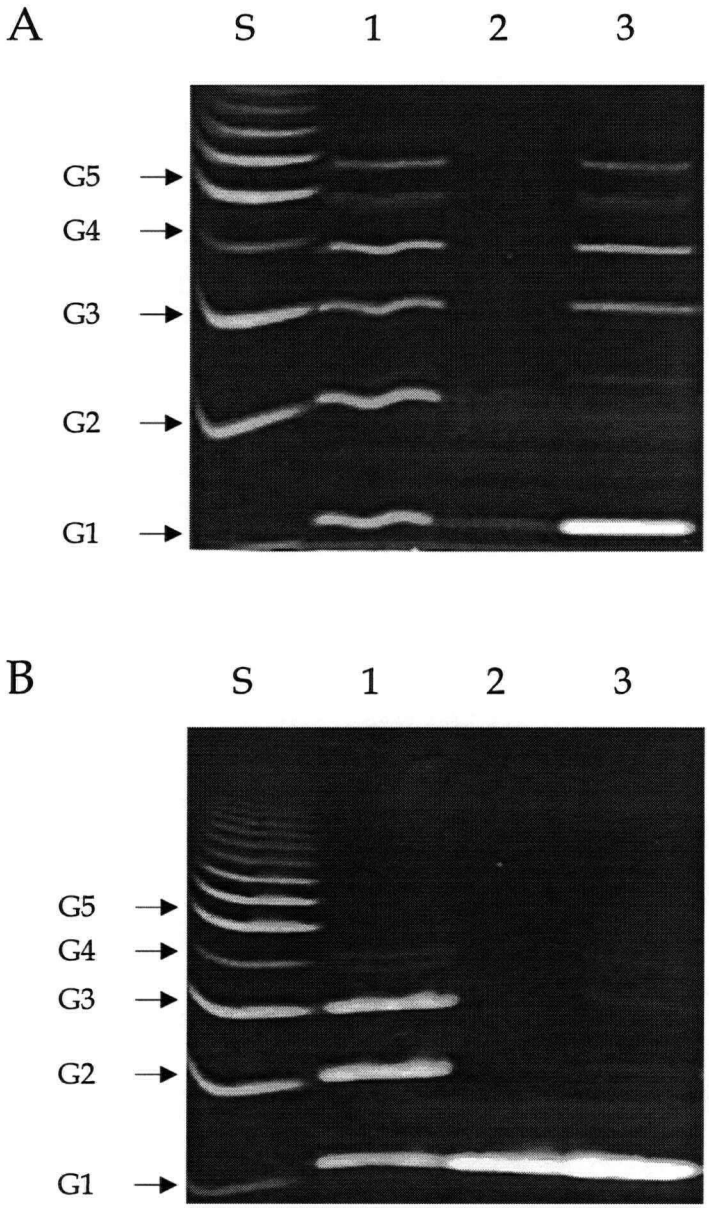


Figure 3.26: Analysis of products formed by hydrolysis of locust bean gum (Panel A) and ivory nut mannan (Panel B) with β-mannanase (Man26A) and β-mannosidase (Man2A). The products were analyzed by fluorophore assisted carbohydrate electrophoresis (FACE) and compared to α-1,4- glucose oligosaccharide standards (Lane S). Lane 1: Man26A, Lane 2: Man2A, Lane 3: Man26A and Man2A.
In a 100 μL reaction volume, 0.1 % (w/v) substrate was digested with 0.5 nmol of either Man26A, Man2A or of both. Incubation time was 1 h at 37° C.

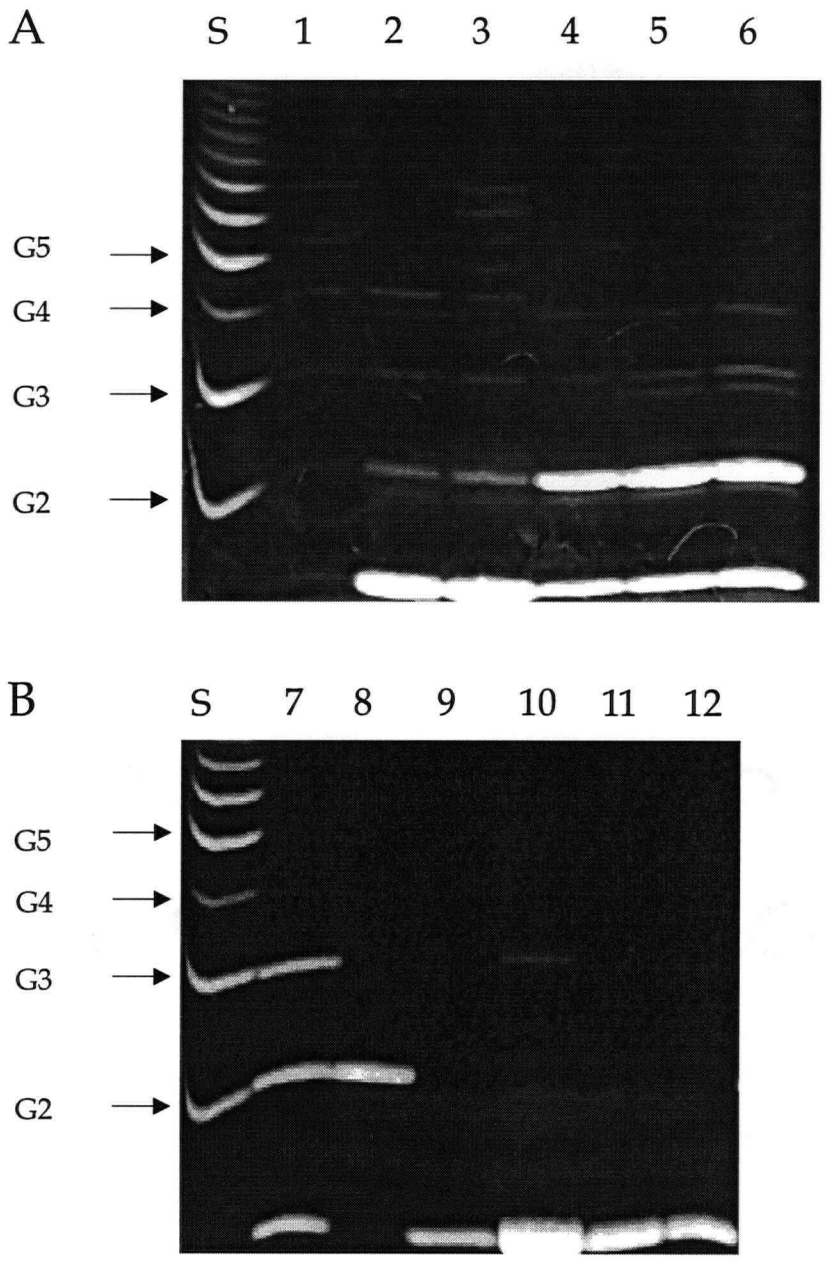


Figure 3.27: Analysis of products formed by hydrolysis of mannooligosaccharides with β-mannanase (Man26A) and β-mannosidase (Man2A). The products were analyzed by fluorophore assisted carbohydrate electrophoresis (FACE) and compared to α-1,4-glucose oligosaccharide standards (Lane S) and β-1,4 mannooligosaccharide standards (M1-M6)(Lane 1). Lane 2: mannotetraose Man2A digested, Lane 3: manno-hexaose Man2A digested. Lanes 4-9 are mannooligosaccharides digested with Man26A, Lane 4: mannotetraose, Lane 5: mannopentaose; Lane 6: mannohexaose; Lane 7: mannotiose; Lane 8: mannobiose; and Lane 9: mannose.Lanes 10-12 are Man2A digests of Lane 10: mannotriose, Lane 11: mannobiose and Lane 12: mannose. 1 mM oligo-saccharide solution was incubated with 0.5 nM enzyme at 37°C for 1 h (100 μL reaction volume)

3.3 Discussion

3.3.1 Molecular architecture of Man26A

A modular architecture is common for β-mannanases. Additional domains such as cellulose-binding domains and cohesins may also be present (Stålbrand *et al.*, 1995; Millward-Sadler *et al.*, 1996; Morris *et al.*, 1995). Some β-mannanases comprise only catalytic domains (Arcand *et al.*, 1993, Braithwaite *et al.*, 1995). The *C. fimi* β-mannanase, Man26A is comprised of a family 26 catalytic domain, a mannan-binding domain (Chapter 4), a SLH domain and a C-terminal domain of unknown function. This study is the first report of a β-mannanase with a SLH domain and the first such domain described for any *C. fimi* protein. SLH domains occur in other enzymes, including xylanases, pullulanases, lichenases and for endo-glucanases (reviewed in Beveridge *et al.*, 1997). In growing *C. thermocellum* cells, a significant fraction of the cellulosomes are cell-bound. This interaction is mediated by surface proteins that bind to the cell envelope *via* SLH domains (Section 1.4). SLH domains from S-layer proteins from *Bacillus stearothermophilus* WT and PV76/p6 strains were shown to non-covalently bind to a secondary cell wall polymer composed of Glc-NAc and Man-NAc, but not to peptidoglycan (Egelseer *et al.*, 1998; Ries *et al.*, 1997). It is possible, based on the sequence diversity of SLH domains, that not all the SLH domains recognize the same cell wall components. SLH domains have also been shown to interact with other SLH domains (Lemaire *et al.*, 1995), which could be an indication that *C. fimi* could form small aggregation of enzyme complexes on the cell surface. The C-terminal region of Man26A might act as an additional anchoring segment by recognizing other cell wall components or perhaps other proteins, supporting the function of the SLH domain.

3.3.2 Localization of mannanase activity in *C. fimi* cultures

Man26A appeared to be transiently cell-bound (Figure 3.12). The activity in the unfractionated culture, however, was lower than the combined activities of the supernatant and washed cells. This discrepancy could be explained by the presence of residual LBG in the culture which interfered with the azo-carob galactomannan activity assay, possibly due to a higher affinity of Man26A for LBG than for the azo-substrate. Man26A has a mannan-binding domain (MBD) (Chapter 4). The centrifugation of the bacterial culture in the preparation of the supernatant fraction removes some of the LBG, bound to the MBD of Man26A and hence to the cells, from the supernatant. Repeated washing of the cells carrying bound LBG could have removed some or all of the LBG. The unfractionated culture sample probably had the highest residual LBG concentration, and therefore showed lower detectable activity on azo-carob galactomannan than expected. Nevertheless, *C. fimi* appeared to contain cell-bound β-mannanase which, depending on the culture conditions, possibly, could be released into the culture supernatant.

More experiments are necessary to study the cell-association of Man26A. One possible experiment would be to synthesize a fluorescently labeled, mechanism-based mannanase inhibitor and detect inactivated enzyme bound to the cell surface by fluorescence microscopy.

3.3.3 The *C. fimi* β-mannosidase, Man2A

The *man2A* gene is only the second bacterial, and the first eubacterial β-mannosidase gene described (Bauer *et al.*, 1996). The Man2A protein is closely related to the mammalian lysosomal β-mannosidases, on the basis of primary amino acid sequence alignments (Figure

3.18). Man2A was inhibited by low concentrations of PNPM. *T. reesei* β-glucosidase (Estrada *et al.*, 1990) and β-galactosidases *from E. coli, Aspergillus* and Jack bean (Pocsi *et al.*, 1993) are also inhibited by substrate. Inhibition of Man2A at low concentrations of substrate suggested that the intracellular mannooligosaccharide levels are controlled, in order to maintain optimal activity of the enzyme. Substrate inhibition will be discussed in Section 6.3.4.3.

The *C. fimi* β-mannosidase migrated considerably faster during SDS-PAGE under non-reducing conditions than under reducing conditions, an indication that native Man2A forms a tight fold, possibly stabilized by disulfide bonds formed between two or more of the six cysteine residues in Man2A.

3.3.4 Degradation of mannooligosaccharides by Man26A and Man2A

Man2A hydrolyzes all of the β-1,4 mannooligosaccharides tested M2 to M6 to mannose as the major product, with some minor products, e.g. mannotetraose cleavage also produced some mannobiose and less mannotriose as minor products. Mannopentaose is hydrolyzed 8 times faster than mannobiose by the exo-β-manannase from guar seeds, whereas the hydrolysis rates from the β-mannosidase from *Helix pomatia* are identical for mannooligosaccharides with DPs from 2 to 5 (McCleary, 1983). The accumulation of more mannobiose than mannotriose during hydrolysis of mannotetraose and mannohexaose by Man2A indicated that it might prefer oligosaccharides longer than mannobiose as substrates (Figure 3.27). The production of mannotriose and mannobiose also suggested that the degradation of mannooligosaccharides by Man2A was not processive, at least not for mannooligosaccharides with DP from 2 to 6. As the optimum DP of the substrate for

Man2A has not been determined, it is unclear if Man2A is an exo-β-mannanase or a β-mannosidase.

The major products released from mannohexaose, -pentaose and -tetraose by Man26A were mannobiose and mannose. Mannotriose was hydrolyzed only slowly by Man26A, to mannobiose and mannose, whereas mannobiose was not hydrolyzed at all. Hydrolysis of mannopentaose by the *T. reesei* β-mannanase, BMANI occurs via a fast initial cleavage into mannotriose and mannobiose, followed by a slower hydrolysis of mannotriose into mannobiose and mannose (Hariupää *et al.*, 1995). In other systems, hydrolysis of mannohexaose and mannopentaose produces mainly mannobiose and mannotriose, because enzymes such as the *Aspergillus niger* β-mannanase are unable to cleave mannotriose (McCleary *et al.*, 1983; Arcand *et al.*, 1993). Man26A cleaves mannotriose into mannobiose and mannose. Mannotetraose is also cleaved into mannobiose and mannose. Presumably mannotetraose was first cleaved into mannose and mannotriose, as with the *A. niger* mannanase (McCleary *et al.*, 1983). The *A. niger* mannanase, and other mannanases, also transglycosylate (Coulombel *et al.*, 1981; McCleary *et al.*, 1983). It is possible that Man26A hydrolyzes mannotetraose into mannobiose, and that the mannobiosyl-Man26A covalent intermediate transfers mannobiose onto mannotetraose, producing mannohexaose; mannohexaose could then be hydrolyzed to mannobiose and mannose as described. Mannotriose is produced from mannohexaose, -pentaose and –tetraose, but it appeared as a doublet in the FACE gels (Figure 3.27). This doublet, however, was not produced from mannotriose. Since the regiochemistry of glycosidic linkages influences the mobility of oligosaccharides in FACE gels, e.g. α-1,3 mannobiose migrates faster than α-1,4 mannobiose (Linda Sandercock, personal communication), it is possible that one of the bands seen for mannotriose is β-1,3 linked mannotriose, whereas the other one, comigrating with the

standard, is β-1,4 linked mannotriose. Only the faster migrating form of mannotetraose, presumably including a linkage other than β-1,4, was detected. Transglycosylation seems to occur with mannohexaose, -pentaose and -tetraose but not with mannotriose and mannobiose. Man2A also transglycosylates (Chapter 6). The cleavage of mannooligosaccharides by Man26A is illustrated in Figure 3.28.

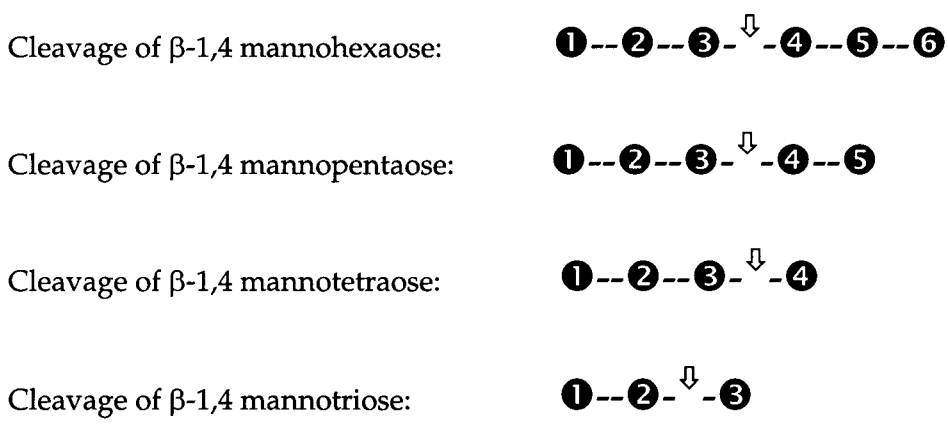


Figure 3.28: Scheme of β-1,4 mannooligosaccharide cleavage by Man26A. Cleavage sites are indicated by arrows and β-1,4 linked mannose molecules by filled circles. The numbers indicate the relative position of the sugars.

3.3.5 Hydrolysis of galactomannan and mannan by Man26A and Man2A

The release of a variety of mannooligosaccharides from LBG, a substituted galactomannan, suggests that the *C. fimi* β-mannanase was an enzyme with endo activity (Figure 3.26). Mannotriose, mannobiose and mannose were released from unsubstitued mannan (INM) indicating a preference of Man26A for mannooligosaccharides of DP >3. Man2A released only insignificant amounts of mannose from LBG, indicating that Man2A attacked the substrate only from the non-reducing end and that α-1,6-galactose side chains block futher hydrolysis. Significantly more mannose is cleaved from INM with Man2A, consistent with a high concentration of free ends (Section 4.2.5). The concerted action of Man26A and Man2A was sufficient to degrade INM into mannose. More mannose was produced in the presence of both enzymes than with Man2A alone. Complete degradation of LBG into monomers could not be obtained with the two enzymes, suggesting a requirement for the intracellular α-galactosidase to remove the galactose side chains (Section 3.2.1). The only detectable difference between products from the hydrolysis of LBG by Man26A and by Man26A and Man2A LBG hydrolysis was the complete cleavage of the disaccharide product in the presence of both enzymes. This disaccharide was therefore mannobiose. The only disaccharide produced from LBG by *A. niger* β-mannanase is mannobiose, but two trisaccharides, mannotriose and 6^1-α-galactosyl-β-D-mannobiose are produced (McCleary *et al.*, 1983). The relative amounts of the products with the mobilities of the G3, G4, G5 and G6 standards were very similar in the Man26A and Man26A + Man2A digestions, indicating that most of these oligosaccharides had an α-galactosyl substitution at their non-reducing ends, which blocked further hydrolysis by Man2A. Thus, a semi-processive hydrolysis mechanism, as was proposed for the *C. fimi* β-1,4 glucanase CenC

(Tomme *et al.*, 1996), is proposed for the β-mannanase Man26A. In this proposal Man26A cleaves the backbone of galactomannan internally, holds on to the substrate with the help of the MBD (Section 4.3.4) and moves along the chain from the non-reducing end towards the reducing end, thereby releasing smaller oligosaccharides. The enzyme moves along the chain until the processive action is stopped by an α-galactosyl side chain. The enzyme has to desorb from the substrate to intiate another endoglycolytic cleavage. This mechanism would explain the production of galactomannanoligosaccharides with galactosyl side chains bound mainly to the mannose at the non-reducing end of the oligomer. Thus the high degree of substitution in LBG, with a mannose to galactose ratio of 5:1 (McCleary *et al.*, 1985), reduces the efficiency of semi-processive cleavage of LBG by Man26A on LBG.

4 Mannan-Binding Domain

4.1 Introduction

4.1.1 Carbohydrate-binding domains of polysaccharidases

Cellulolytic and hemicellulolytic enzymes of non-complexed degradation systems are very often modular enzymes. The most common non-catalytic domains are cellulose-binding domains (CBDs). CBDs are believed to enhance activity by binding the enzyme to the substrate thereby increasing the enzyme concentration on the substrate (Gilkes *et al.*, 1992 and 1993). In complexed cellulolytic systems, e.g in the cellulosome of *C. thermocellum*, the CBD is part of the scaffoldin protein and mediates the anchoring of the entire cellulosome to the substrate. Since cellulosomes are mostly cell bound multienzyme complexes, the scaffoldin CBD is also ultimately responsible for binding the *C. thermocellum* cells to the substrate, allowing improved uptake of the degradation products (Béguin *et al.*, 1996). Substrate binding domains are a common occurence not only in plant cell wall degrading enzymes, but also enzymes involved in starch and chitin degradation that use starch-binding and chitin-binding domains, respectively, to adsorb to the substrate (Williamson *et al.*, 1997, Tomme *et al.*, 1995).

CBDs have been classified into families based on amino acid sequence similarities (Tomme *et al.*, 1995). Currently, more than 170 cellulose-binding domains have been identified and grouped into 17 different families (Creagh *et al.*, 1998; P. Tomme, personal communication). Their sizes range from 33 to over 170 amino acid residues.

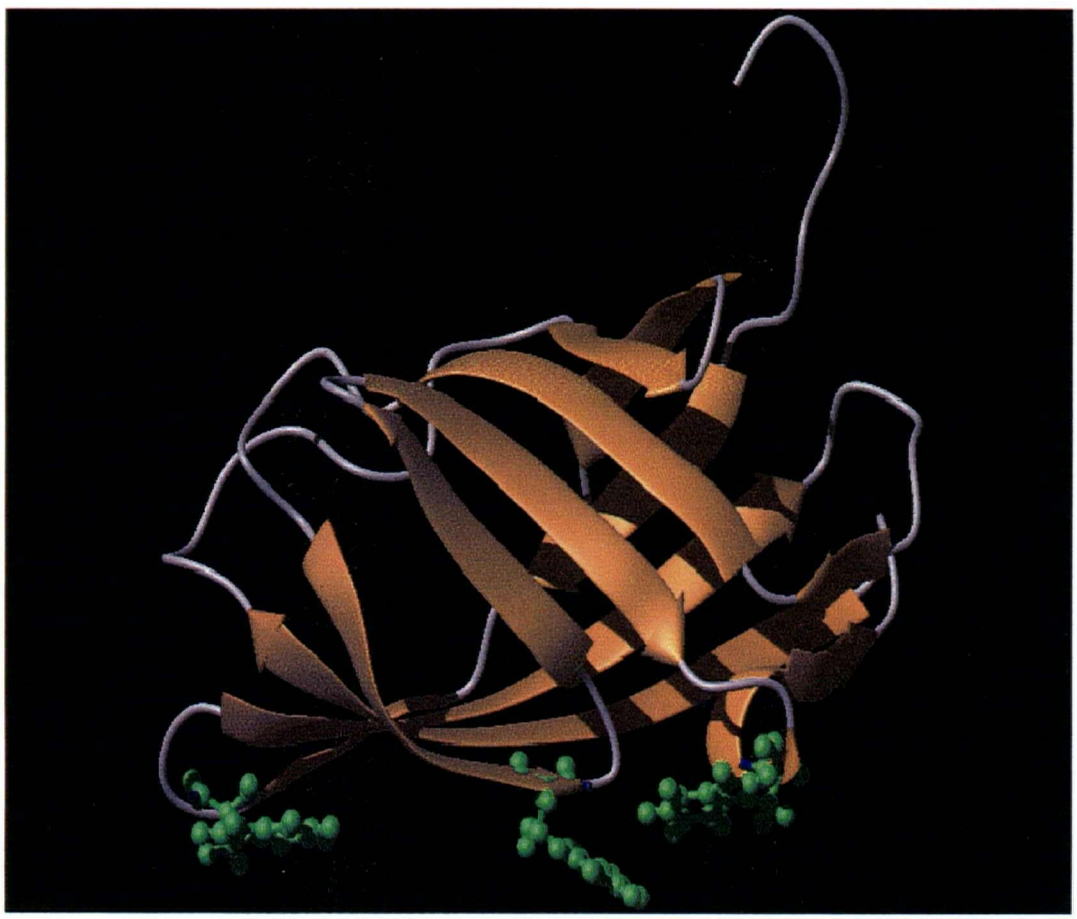
All the cellulases and xylanases secreted by *Cellulomonas fimi* include at least one CBD, including members from CBD families I, II, III, IV and IX. The CBDs differ in their

binding affinities for the different forms of cellulose. The two tandemly repeated CBDs from the *C. fimi* endoglucanase CenC, CBD_{N1} and CBD_{N2}, are unique in their properties. They bind soluble cello-oligosaccharides, and amorphous but not crystalline cellulose (Tomme *et al.*, 1996b, Coutino *et al.*, 1992). The biotechnological potential of CBDs led to detailed structural and functional studies of CBD_{N1} and CBD_{Cex}. The CBD_{cex} from the *C. fimi* xylanase/exoglucanase binds, unlike CBD_{N1} , to crystalline cellulose.

4.1.2 CBD_{cex}

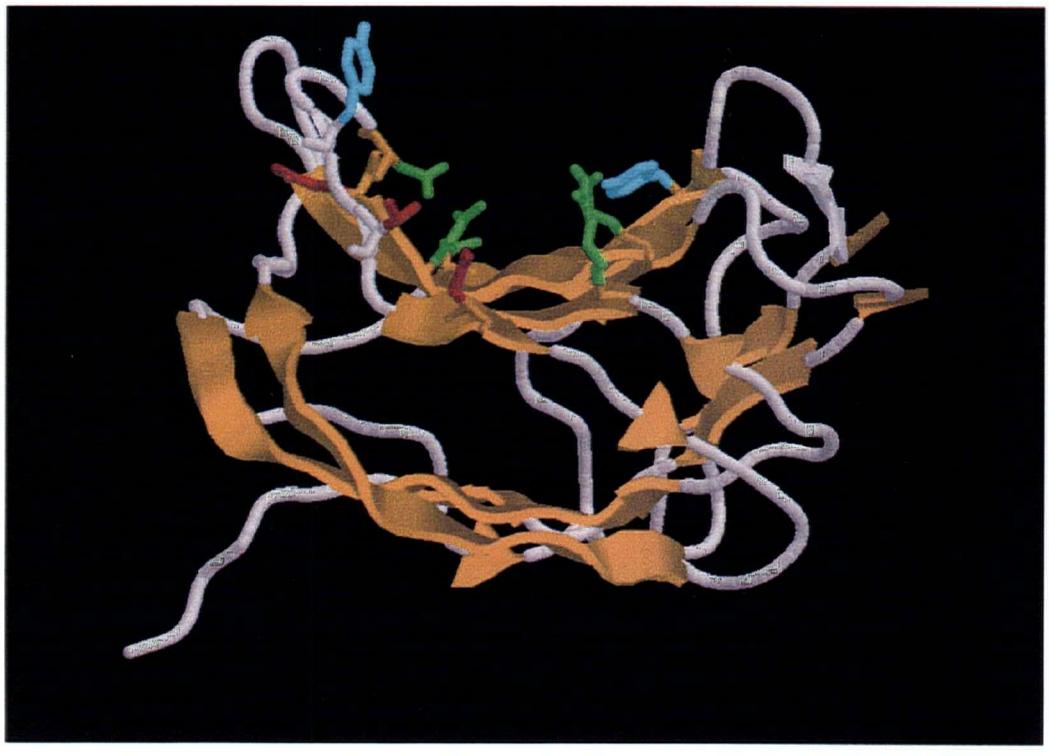
CBD_{cex} is 108 aa long and is at the C-terminal end of Cex (Figure 1.3). Three dimensional structure analysis of this family II CBD revealed a nine-stranded β-barrel fold (Xu *et al.*, 1995). In this structure the three exposed tryptophans, W17, W54 and W72, are aligned on one face as shown in Figure 4.1. The role of these aromatic residues in cellulose binding was demonstrated by chemical modifications (Bray *et al.*, 1996). The arrangement of hydrophobic amino acid side chains on one face is also found in other CBDs. NMR structures of *T. reesei* CBD_{CBHI} (family I) and *Erwinia chrysanthemi* CBD_{EGZ} (family V) revealed either tyrosines or tryptophans on the putative binding faces (Brun *et al.*, 1997). Each of these tyrosines in CBD_{CBHI} was shown by site-directed mutations to be indispensable in cellulose binding (Linder *et al.*, 1995).

Binding studies of CBD_{Cex} on crystalline cellulose showed that binding is entropically driven, indicating significant dehydration of the protein-ligand contact surfaces. Whereas polar interactions, such as hydrogen bonding and van der Waals interactions, provide the major driving force in the enthalpic binding process of CBD_{N1} to soluble sugars (Tomme *et al.*, 1996b, Creagh *et al.*, 1998).



Courtesy of Bradley W. McLean

Figure 4.1: Ribbon diagram of the structure of CBD_{Cex}. The amino acid side chains of three of its five tryptophan residues are shown (W17, W52 and W72). These three tryptophans are aligned along one face of the nine-stranded β-barrel fold. These three tryptophans are conserved in all family II cellulose-binding domains.



Courtesy of J.M. Kormos

Figure 4.2: Ribon diagram of the structure of CBD_{N1}. The ten β-strands are organized in two antiparallel β-sheets, which adopt an overall topology of a jelly-roll β-sandwich. Some of the amino acid side chains involved in substrate binding are shown. Hydrophobic side chains are highlighted in blue, polar residues in green and charged residues in red.

4.1.3 CBD_{N1}

CBD_{N1} is unusual in respect of its binding properties and its structure. It does not bind to crystalline cellulose, but it can bind amorphous cellulose and soluble oligosaccharides (Johnson *et al.*, 1996a; Tomme *et al.*, 1996b). The tertiary fold of this 152 aa long CBD is strikingly different from that of CBD_{Cex}. The jelly-roll β-sandwich lacks a flat binding surface; it forms a binding cleft closely resembling the structure of an endo-1,3-1,4-β-glucanase from *Bacillus macerans* (Johnson *et al.*, 1996b) (Figure 4.2). The binding cleft can accommodate approximately five glucopyranosyl residues. A strip of hydrophobic amino acid residues runs along the center of the binding cleft, whereas the sides are lined by mostly hydrophilic residues. Two tyrosines, Tyr19 which is in the cleft and Tyr85, which is on a loop proximal to the cleft, play critical roles in binding the oligosaccharide chain (Kormos, 1998).

4.1.4 Objectives

It was the aim of this project to study the substrate-binding of the *C. fimi* mannanase, Man26A. A putative mannan-binding domain from this enzyme was sub-cloned and expressed in *E. coli* to study mannan binding.

4.2 Results

4.2.1 Characterization of a novel domain from *C. fimi* mannanase, Man26A

The region between the catalytic domain and the SLH domain in Man26A had no significant sequence similarity to any other known protein sequence (Section 3.2.4). All of the characterized secreted *C. fimi* cellulases and hemicellulases are modular enzymes and have at least one cellulose binding domain. It was therefore hypothesized that Man26A might have either a cellulose or hemicellulose-binding domain. To test this hypothesis, binding of Man26A to the insoluble substrates chitin, mannan, cellulose and xylan was analyzed. However, binding was not detected to any of these polysaccharides (data not shown).

Binding of Man26A to soluble mannan was studied by affinity gel electrophoresis (AFGE). In this technique native protein samples are electrophoretically separated in polyacrylamide gels containing different concentrations of soluble substrate. Reversible binding of the protein to the soluble substrate would reduce its migration distance (mobility). The change in mobility is proportional to the substrate concentration. The mobility of proteins without binding specificity should not change, even in the presence of high concentrations of the polymeric substrate (Nakamura *et al.*, 1992). For the binding study of Man26A and Man2A, azo-carob galactomannan was added to the AFGE gels. The relative mobility, which is used to calculate the dissociation constant, is expressed as the migration distance of the protein sample *versus* the migration distance of a non-binding control. BSA was used as a control in the experiments shown in Figure 4.3. The relative mobility of Man2A was not affected by mannan, not even at a concentration of 1 %,

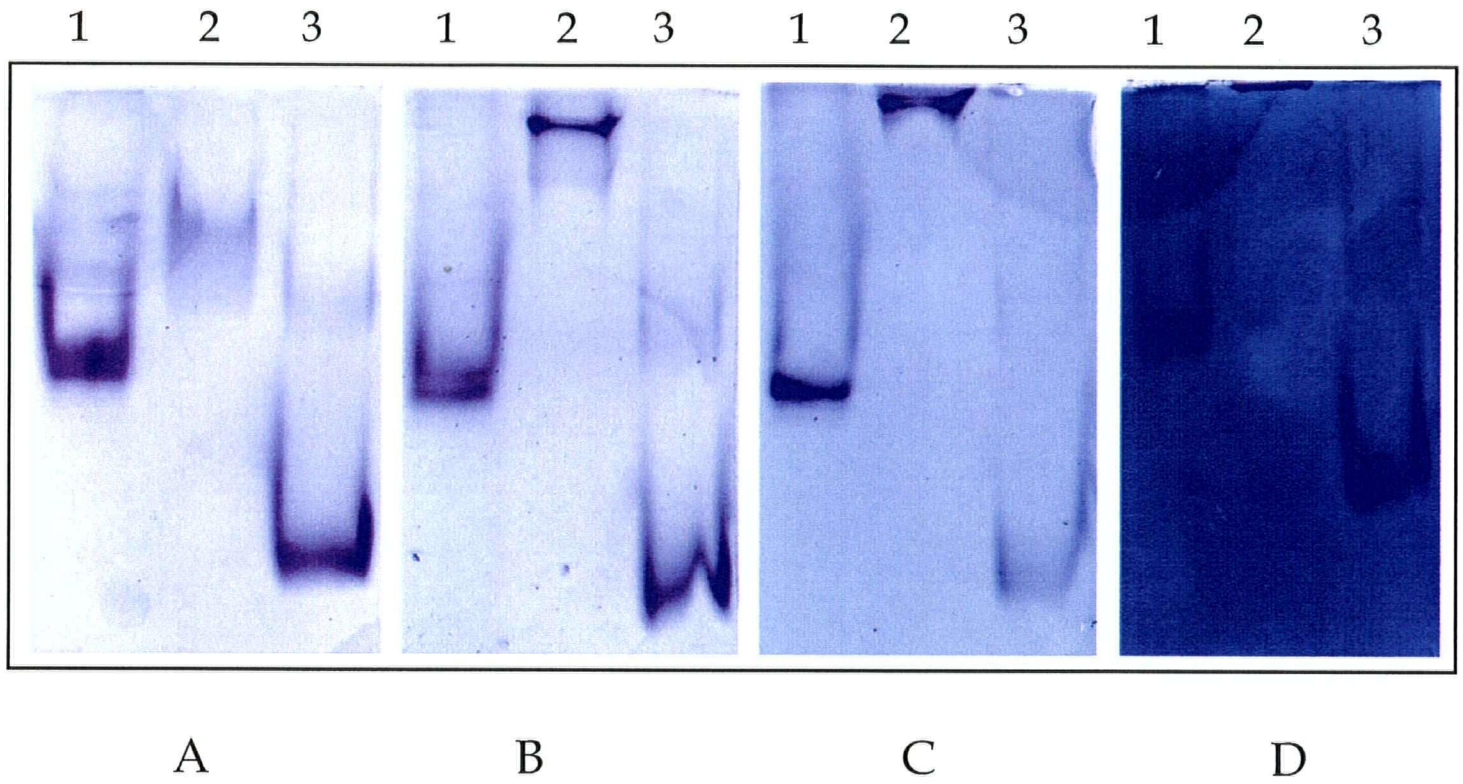


Figure 4.3: Affinity gel electrophoresis of Man2A (Lane 1), Man26A (Lane 2) and BSA (Lane 3) (each 6 µg/lane). Increasing concentrations of macromolecular affinity ligand, i.e. azo-carob galactomannan, were added to the polyacrylamide gels: 0 % (Panel A), 0.05 % (Panel B), 0.2 % (Panel C) and 1 % (Panel D). The gels were run at 150 V for 1.5 h at 4 °C.

indicating that Man2A did not bind to soluble mannan. By contrast, changes in relative mobility of Man26A, were detected at very low substrate concentrations and the presence of 1 % substrate prevented the protein from migrating into the gel. This suggested a relatively strong interaction of Man26A with azo-carob galactomannan.

The AFGE-zymogram method was used to test whether the catalytic domain of Man26A, or the non-catalytic portion of the enzyme was involved in substrate binding (Figure 4.3). The catalytic domain was obtained by protease treatment of Man26A (Section 3.2.6). Changes in relative mobility of undigested Man26A were compared to changes in relative mobility of the Man26A catalytic domain on AF gels with 0 % and 1 % mannan. To confirm that the proteolytic band corresponded to the catalytic domain, the 1 % azo-carob galactomannan gel was incubated in phosphate buffer (1 h at 37^0 C) prior to Coomassie-staining. The presence of mannan did not affect the relative mobility of the catalytic domain, seen as a zone of clearing on the 1 % substrate gel, whereas the intact enzyme did not even migrate into the gel (Figure 4.4). It appeared that mannan binding was due to a domain other than the catalytic domain. It was hypothesized that a mannan-binding domain might be present between the catalytic domain and the SLH domain.

4.2.2 Sub-Cloning of *mbd*$_{1112}$

The DNA fragment encoding the protein portion between the catalytic and SLH domain, the putative mannan-binding domain (MBD), was cloned into the pET28a vector (Novagen) as described: The primers MBD11 (*Nco* I) and MBD12 (*Not* I) were used to amplify the MBD encoding DNA fragment by the polymerase chain reaction (PCR) using pCMan2 as template. The PCR product was cloned into the pZErO™-1.1 vector (Invitrogen)

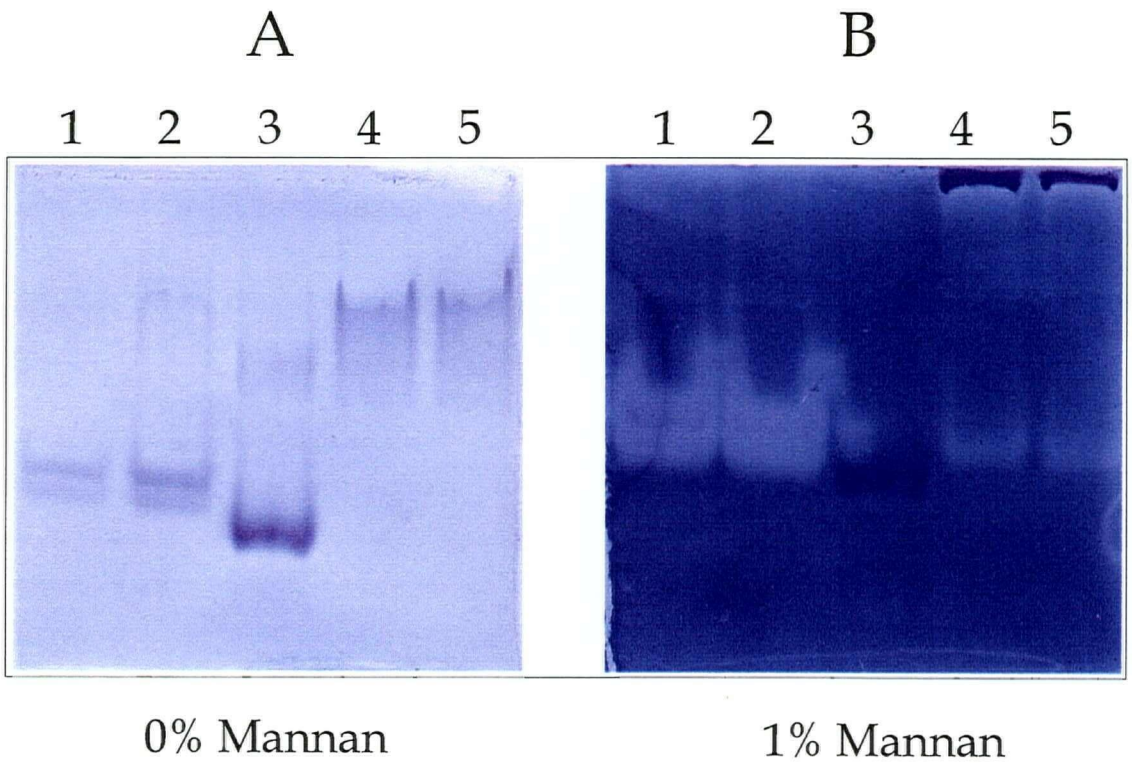


Figure 4.4: Affinity gel electrophoresis (AFGE) of protease treated Man26A (Lane 1 and 2), BSA (Lane 3) and intact Man26A (Lane 4 and 5) without (Panel A) or with 1% (Panel B) macromolecular affinity ligand, i.e. azo-carob galactomannan. To detect mannanase activity the gel was incubated in excess potassium-phosphate buffer (pH 7.0) at 37 °C until clearings were visible. Protein bands were then visualized by Coomassie staining.

at the *Eco*R V restriction site. From this plasmid, pOMBD$_{1112}$, the 570 bp *mbd1112* DNA fragment was excised by *Nco* I and *Not* I restriction endonuclease digestion, and ligated into the expression vector pET28a to obtain pET28MBD$_{1112}$, encoding the putative MBD translationally fused to a C-terminal sequence of six histidines.

The *mbd1112* coding region was expressed in *E.coli* BL21(DE3) cells, producing a protein, MBD$_{1112}$, with a calculated molecular weight of 20,990. Initial expression levels were low and purification by MCAC was unsuccessful, because of poor binding of MBD$_{1112}$ to the affinity column. The experiments presented below were intended to determine the binding characteristics of MBD$_{1112}$ and its potential applications, prior to improving protein expression and purification.

4.2.3 Characteristics of the mannan binding domain (MBD$_{1112}$)

Binding of MBD$_{1112}$ to mannan was analyzed by AFGE. The MBD$_{1112}$ protein bands from the partially purified sample were detected after AFGE separation on Western blots, using oligohistidine specific antibodies (Figure 4.5). Man2A was used as the non-mannan-binding control. To mark the top of the AFGE gels, CBD$_{CenD}$, a protein that does not migrate into the separating gels under native conditions (Boraston, A., and Tomme, P. unpublished results) was applied. The relative mobilities of Man26A and MBD$_{1112}$ were compared in gels including 0 %, 5 x 10^{-4} %, 7.5 x 10^{-4} %, 1.25 x 10^{-3} %, 2.5 x 10^{-3} % and 1.25 x 10^{-2} % locust bean gum and 1 x 10^{-3} %, 1.5x10^{-3} %, 2.5 x 10^{-3} %, 5 x 10^{-3} % and 2.5 x 10^{-2} % azo-carob galactomannan. The relative mobilities of Man26A and MBD$_{1112}$ decreased with increasing substrate concentrations. The decrease was similar for both proteins. However, slightly stronger protein-substrate interactions were detected for Man26A (Figure 4.5). From the

double reciprocal plots of 1/(R-r) *versus* 1/c (see Equation E 4.1) the negative reciprocal of the dissociation constant could be determined as the intercept on the abscissa.

E 4.1:

$$\frac{1}{R - r} = \frac{(1 + K_d/c)}{R - R_c}$$

d: migrating distance of protein (MBD_{1112})

D: migrating distance of reference (Man2A)

r: relative mobility (d/D) of MBD_{1112} in the presence of mannan

R: relative mobility of unbound MBD_{1112}

R_c: relative mobility of MBD_{1112} -mannan complex, or that value obtained in the presence of an excess amount of mannan with all MBD molecules bound.

K_d: dissociation constant of protein for affinity ligand

The dissociation constants for MBD_{1112} were determined to be $K_d = 4.6 \times 10^{-4}$ % ($\pm 4 \times 10^{-5}$) for locust bean gum and $K_d = 5.5 \times 10^{-3}$ % ($\pm 1 \times 10^{-3}$) for azo-carob galactomannan. The 12 fold weaker binding of MBD_{1112} to the azo-substrate could be caused by the Remazol brilliant-Blue R molecules linked to the backbone to an extent of about one dye molecule per 20 sugar residues reducing the accessibility of binding sites. Furthermore, a lower viscosity, as found for the azo-carob galactomannan, could be indicative of molecules with a lower degree of polymerization (DP). Shorter molecules would have fewer potential binding sites, assuming MBD_{1112} binds to the mannan backbone (Section 4.2.4).

In order to calculate the molar dissociation constant, the molarity of the locust bean gum solution was determined by total and reducing sugar analysis, assuming a galactose to

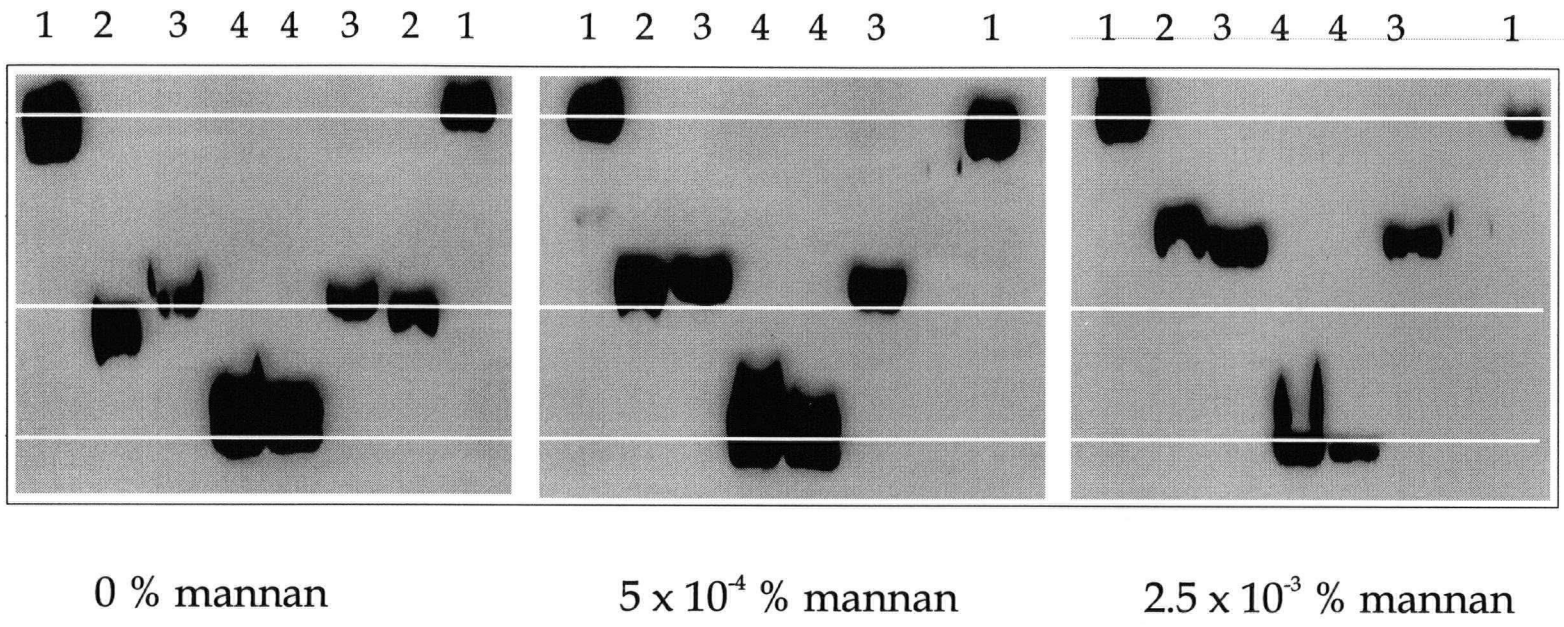


Figure 4.5: Affinity gel electrophoresis of CBD_{CenD} (Lane 1), Man26A (Lane 2), MBD_{1112} (Lane 3), and Man2A (Lane 4). The concentrations of macromolecular affinity ligand, i.e. locust bean gum, are indicated. Gels were run at 150 V for 2 h at 4 °C. The protein bands (~50 ng/lane) were visualized on Western blots using anti-His-tag mouse antibodies (Dianova). Horizontal lines are visual guides for the detection of changes in mobility of Man26A and MBD_{1112}.

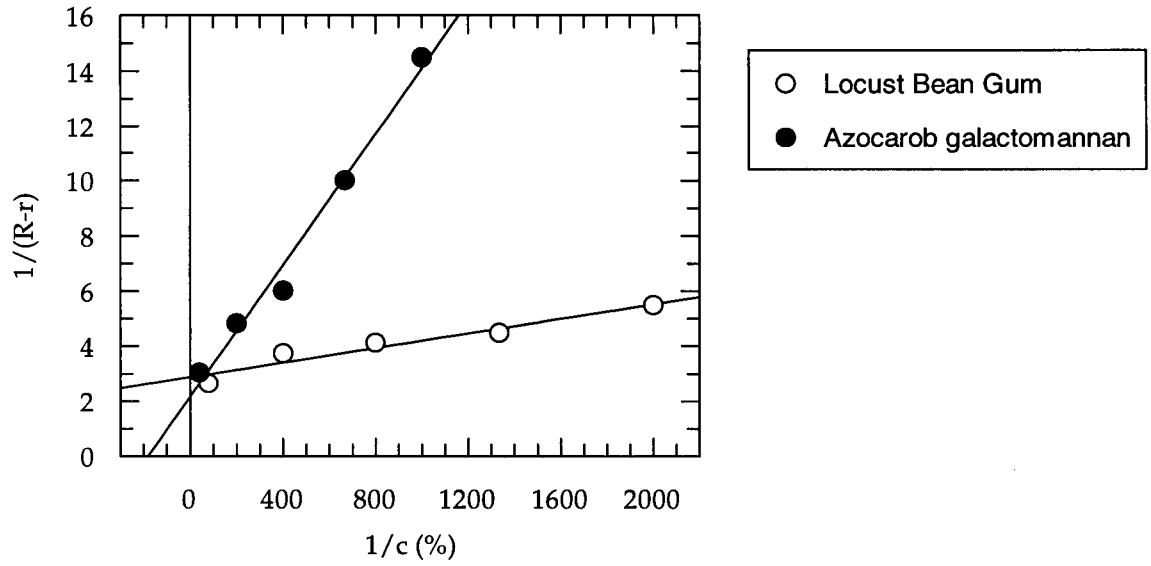


Figure 4.6: Double reciprocal plot of MBD_{1112} AFGE data. Relative mobilities with azo-carob galactomannan and locust bean gum as affinity ligands were plotted as 1/(R-r) *versus* 1/c. Dissociation constants (K_d) were determined from the x-axis intercept (1/-K). K_d for MBD_{1112} on locust bean gum $K_d = 4.6 \times 10^{-4}$ % was and on azo-carob galactomannan $K_d = 5.5 \times 10^{-3}$ %.

mannose ratio of 1:5 (McCleary *et al.*, 1985) (Section 2.16). The soluble galactomannan molecules had an average DP of 130 mannose residues, which meant that a 1 % solution was 0.39 mM. This gave a molar dissociation constant for MBD_{1112} on locust bean gum of $K_d = 1.8 \times 10^{-7}$ M. Although only an approximation, it clearly suggested a high affinity of MBD_{1112} for soluble galactomannan.

4.2.4 Mannan-binding domain MBD1112 is not a lectin

Mannan-binding proteins (MBD) described previously bind to terminal sugars and are, therefore, classified as lectins and were renamed as mannan-binding lectins (MBL)(Mizuno *et al.*, 1981; van Emmerik *et al.*, 1994). It was the goal of the experiments presented in this section to provide evidence that MBD_{1112}, unlike the lectins, binds to the β-1,4-mannose backbone.

The first indication of a non-lectin type of binding for MBD_{1112} came from comparisons of the dissociation constants for locust bean gum and azo-carob galactomannan (Section 4.2.3). If the lower viscosity of the latter corresponded to a lower DP value, its molar concentration would be higher. Therefore, the concentration of terminal sugars would also be greater, presenting more binding sites for mannose-binding lectins. Furthermore, if MBD_{1112} bound only to terminal residues, the presence of Remazol Brilliant Blue should not significantly influence the binding.

Competition affinity gel electrophoresis was used to test the hypothesis that MBD_{1112} is not a lectin-like binding protein. The relative mobility of the mannan-binding domain in a gel with 1.25×10^{-2} % mannan was compared to its relative mobility in 1.25×10^{-2} % mannan gels including 1.8 % mannose or galactose.

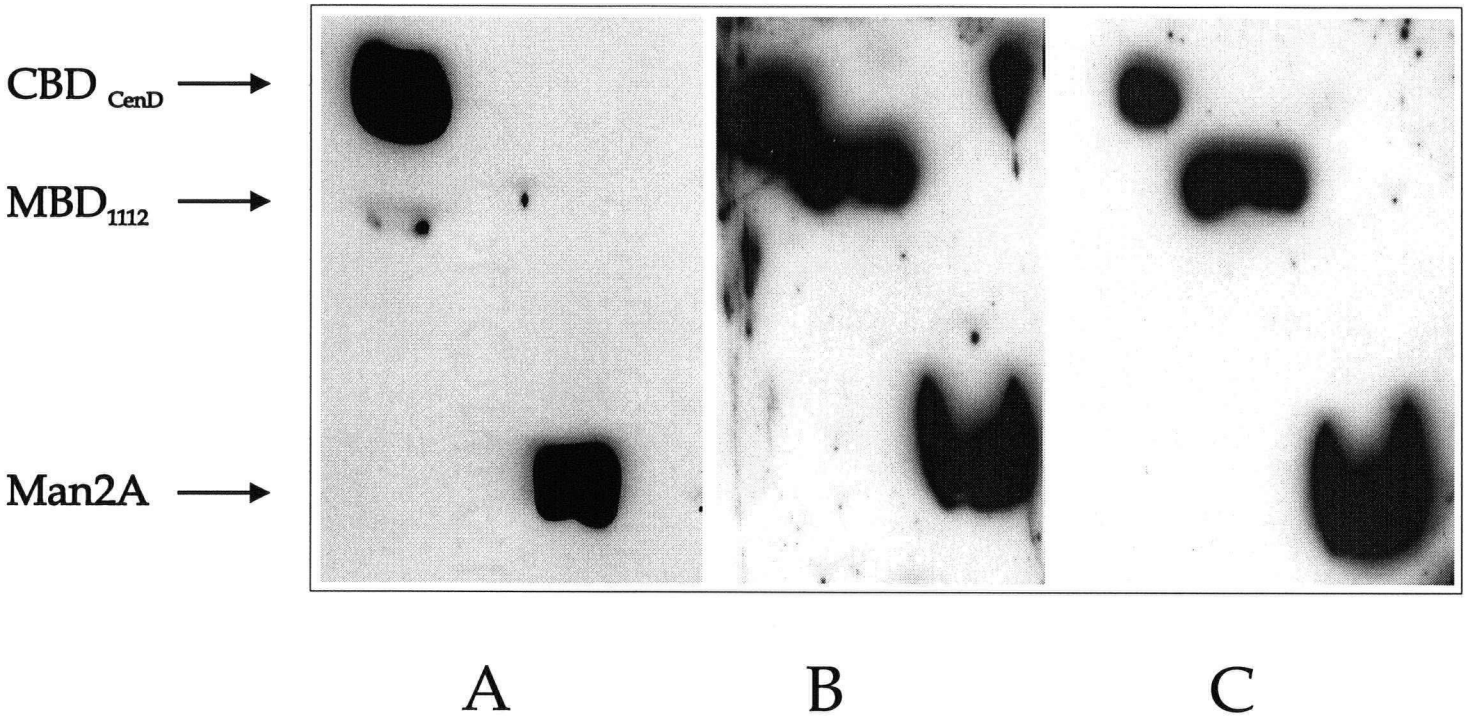


Figure 4.7: Competition affinity gel electrophoresis MBD_{1112}. Controls were CBD_{CenD} and Man2A. Macromolecular affinity ligand, i.e. locust bean gum, was added to all three gels (Panel A, Panel B and Panel C) at a final concentration of 1.25×10^{-2} %. Competitor affinity ligands were added as monosaccharides at a final concentration of 1.8 %: Panel A: no competitor, Panel B: mannose and Panel C: galactose. Protein bands were detected by Western blot.

in the presence of an excess of either galactose or mannose. If MBD_{1112} preferentially bound monomeric sugars, it would have formed a sugar- MBD_{1112} complex and would not have interacted with the mannan. The relative mobility of MBD_{1112} was not changed significantly in the presence of excess galactose or mannose (Figure 4.7), indicating binding preference for polymeric mannan and not for monomeric or terminal mannose or galactose residues, as found in galactomannans. Therefore, the mannan-binding domain from the *C. fimi* mannanase Man26A can be classified as a true mannan-binding domain. This is the first mannan-binding domain to be described.

4.2.5 The role of MBD in mannan degradation

The role of MBD in Man26A catalyzed mannan degradation was investigated. Locust bean gum and ivory nut mannan were incubated with either intact Man26A or Man26A catalytic domain (prepared by proteolytic cleavage; Section 2.19). Man26A had a lower specific activity on crystalline ivory nut mannan (36.5 nmol/mL.min: released reducing sugars per nmol Man26A) than on locust bean gum (123.6 nmol/mL.min per nmol Man26A). 10 mL of 0.2 % locust bean gum in 50 mM citrate pH 5.5 and 10 mL of 0.2 % ivorynut in 50 mM citrate pH 5.5 were incubated with 0.2 nmol and 1.2 nmol, respectively, of intact Man26A and Man26A catalytic domain. After a 30 h incubation period at 37^0 C, the concentration of released reducing sugars was analyzed.

The results summarized in Table 4.1 indicated that hydrolysis rates on locust bean gum were similar for intact Man26A and for Man26A catalytic domain. However, the hydrolysis rates for Man26A were 1.4 times higher than the rates for Man26A catalytic

domain on crystalline mannan (i.e. ivory nut mannan), indicating a role of MBD in enhancing hydrolysis rates on crystalline substrate.

Table 4.1: Reducing sugars produced by enzymatic mannan hydrolysis (30 h at 37^0 C)

Substrate	Man26A catalytic domain	Man26A
Locust bean gum (0.2 %)	2.6 (±0.18) µMol/mL [a]	2.66 (±0.31) µMol/mL [a]
Ivory nut mannan (0.2 %)	1.69 (±0.23) µMol/mL [b]	2.45 (±0.13) µMol/mL [b]

[a]: enzyme concentration: 0.02 nmol/mL

[b]: enzyme concentration: 0.12 nmol/mL

4.3 Discussion

4.3.1 The mannan-bining domain of *C. fimi* mannanase, Man26A

Unlike all of the *C. fimi* cellulases and xylanases described previously, Man26A does not have a cellulose-binding domain. Binding of Man26A to insoluble substrates, such as mannan, chitin, xylan or cellulose was not detected. Binding of Man26A to soluble mannan however, could be demonstrated by AFGE. The DNA encoding the mannan-binding domain, located between the catalytic domain and the SLH domain, was sub-cloned and expressed. Expression levels were low and the MBD with a C-terminal His-tag (MBD_{1112}) could only be partially purified. In this study, no attempts were undertaken to solve these problems. A possible reason for the purification problems might be the presence of a protease susceptible region in the MBD_{1112} C-terminus (Figure 3.1.4). Cleavage at this site would cause the loss of the affinity tag.

By using the AFGE-Western blot technique, both of the problems could be circumvented and dissociation constants (K_d) could be calculated for MBD_{1112} on soluble mannans. $K_d = 4.6 \times 10^{-4}$ % ($\pm 4 \times 10^{-5}$) on locust bean gum and $K_d = 5.5 \times 10^{-3}$ % ($\pm 1 \times 10^{-3}$) on azo-carob galactomannan. Plotting the mobility data as $1/r$ *versus* c, as is usually done for AFGE data analysis, resulted in convex, instead of linear curves. This non-linear behaviour has been previously reported, e.g. for lower molecular weight dextran-ConA complexes (Takeo, 1984). It is believed that the protein retains some mobility, even when tightly complexed with the substrate. To correct for this behaviour, the data had to be plotted according to Equation 4.1 (Takeo, 1984). The calculations of the molar dissociation contant for MBD_{1112} ($K_d = 1.8 \times 10^{-7}$ M) on locust bean gum were based on an average DP value of 130. This value, however, was obtained by assuming a 1:5 galactose to mannose ratio, as has

been reported in the literature (McCleary, *et al.*, 1985). Therefore, the molar dissociation constant for MBD_{1112} can only be seen as an approximation. The production of larger quantity and purer quality of the mannan-binding domain will allow to do more sensitive assays (i.e. isothermal titration calorimetry or surface plasmon resonance) to obtain more accurate binding affinties (Holmskov *et al.*, 1996).

4.3.2 Comparison of MBD_{1112} to CBD_{N1}

MBD_{1112} and CBD_{N1} are similar in respect to their binding preferences. Both domains bind soluble substrates but do not bind crystalline substrates. Both domains are about 150 amino acid residues long and have relatively low calculated pIs (3.5 for CBD_{N1} and 4.3 for MBD_{1112}). Unlike Man26A, which has only one binding domain, CenC has two similar tandemly repeated binding domains, CBD_{N1} and CBD_{N2} (Coutinho *et al.*, 1992). AFGE experiments were also used to determine a dissociation constant of CBD_{N1} on soluble barley β-glucan (β-1,4 and β-1,3 glucose polymer), $K_d = 7.57 \times 10^{-6}$ M (Kormos, 1998), which indicated 42 fold weaker binding of CBD_{N1} to barley β-glucan compared to binding of MBD_{1112} to locust bean gum.

Comparison of the putative binding mechanisms for CBD_{Cex}, which binds with its hydrophobic binding face to a crystalline substrate (Creagh *et al.*, 1996, Bray *et al.*, 1996), and for CBD_{N1}, which binds flexible, single-stranded soluble substrate in a binding cleft, leads to the speculation that MBD_{1112} adopts a similar overall topology to CBD_{N1}, forming a substrate binding cleft rather than a binding face. For both binding site topologies, hydrophobic amino acid side chains (i.e., tryptophans or tyrosines) were shown to be key components in carbohydrate-protein interaction (Bray *et al.*, 1996; Kormos, 1998; McLean

personal communication). With 8 tyrosines and 4 tryptophans present in MBD_{1112} there are many putative carbohydrate-binding amino acid side chains.

4.3.3 Lectins

Another class of proteins binding to mannan is the collectins. Collectins are C-type (Ca^{2+} dependent) carbohydrate-binding proteins that play an important role in the innate immune defence against microorganisms (Lu, 1997). C-type lectins, such as the mannan binding lectin (MBL, previously known as mannan binding protein, MBP) and collectin-43 (CL-43) bind mainly to terminal mannose and glucose in the presence of Ca^{2+}. By surface plasmon resonance, the dissociation constant for the monomeric CL-43 lectin was determined on *Saccharomyces cerevisiae* mannan as $K_d = 2.68 \times 10^{-8}$ (Holmskov *et al.*, 1996). However, the collectins bind only to terminal sugars. The results from competition AFGE indicated, that MBD_{1112} preferentially bound to polymeric β-1,4 linked mannose rather than to monomeric mannose or galactose. It was therefore proposed that MBD_{1112} did not bind to mannan in a lectin-like fashion but rather as a true β-1,4 mannan-binding protein. This is the first report of such a mannan-binding domain.

4.3.4 The role of MBD_{1112} in Man26A catalyzed mannan degradation

The role of MBD_{1112} in mannan hydrolysis by Man26A was investigated on locust bean gum and on ivory nut mannan. The latter is an unsubstituted, crystalline β-1,4-mannan, whereas locust bean gum is more amorphous and has (α-1,6) galactose substitutions. Intact Man26A and Man26A catalytic domain released similar quantities of reducing sugars from locust bean gum over a 30 h incubation period. However, the intact

Man26A was 1.4 times more efficient than Man26A catalytic domain in producing reducing sugars on ivory nut mannan. Similar results were obtained for the degradation of bacterial micro-crystalline cellulose (BMCC) with CenA. Intact CenA was 3.3 times more efficient in degrading the substrate, than the catalytic domain without the binding domain (Gilkes *et al.*, 1992). For CenA, the results agree with the proposed role of CBDs to increase the local concentration of catalytic domain on the substrate. This, however, can not be the role of MBD_{1112} because binding of Man26A to ivory nut mannan could not be detected. It is possible that Man26A is able to degrade ivory-nut mannan semi-processively, but semi-processivity on locust bean gum is impaired by the galactose substitutions. Semi-processivity, endo-cleavage followed by processive exo-cleavage, has been suggested previously for CenC based on CMC viscosimetric analyses (Tomme *et al.*, 1996). CenC and Man26A both have at least one binding domain that bind soluble oligosaccharides. Therefore, it is hypothesized that the binding domain is involved in the semi-processive cleavage of soluble or amorphous substrates, by "feeding" the oligosaccharide chains to the catalytic domain, and that substitutions, such as the galactose side chains in locust bean gum inhibit processive hydrolysis. Semi-processivity was also proposed for the *Thermomonospora fusca* E4 β-glucanase (Irwin *et al.*, 1998). The crystal structure of E4 β-glucanase catalytic domain linked to the family IIIc E4 CBD has been solved (Sakon et al., 1997). The binding face of the type IIIc CBD and the shallow catalytic cleft of the catalytic domain are aligned in E4 in such a way that a cellulose strand could bind along both domains and enable semi-processive hydrolysis (Irwin *et al.*, 1998).

By analogy it is thought that the catalytic domain of Man26A cleaves crystalline mannan (endo-acting) releasing soluble mannan chains to which the MBD binds. The MBD

would feed the strands to the catalytic domain, which would processively degrade the strands (exo-acting) into mannobiose or mannose.

This model could be tested by viscosimetric assays using substrates with different degrees of substitutions, or viscosimetric assays with solubilized ivory nut mannan, comparing Man26A to a non-modular, only endo-acting mannanase, i.e., *P. fluorescens* ManA (Braithwaite *et al.*, 1995).

4.3.5 Protein purification by aqueous two-phase systems

Separation of two phases in an aqueous two-phase system has a wide varyiety of applications. One of them is large-scale protein purification. Phase separation with solutions of galactomannans and polyethylene glycol (PEG), dextran or citrate has been reported (Franco *et al.*, 1996). The binding characteristics of MBD_{1112} make this binding domain a very good candidate to use as an affinity tag in the purification of fusion proteins. In preliminary experiments with locust bean gum-PEG two-phase systems, purification of Man26A from 1.3kg of *E. coli* supernatant was achieved. Adding 12 g of locust bean gum (600 g of 2 % solution) and 200 g of PEG 20,000 to 1.3 kg of supernatant resulted in a two phase system in which Man26A patitioned into the small volume mannan phase (data not shown). However, detailed binding studies with MBD_{1112} are required to find non-denaturing desorption conditions for removal of the fusion protein from the mannan.

5 Model of galactomannan degradation by *C. fimi*

The cellulose hydrolyzing system from the Gram positive soil bacterium *Cellulomonas fimi* does not represent a paradigm for hydrolysis of plant cell wall polysaccharides by this bacterium. The mannan hydrolyzing system in *C. fimi* is, unlike the cellulolytic system, composed of only one secreted "endo-acting" enzyme, i.e. endo- 1,4-β-mannanase (Man26A), one intracellular 1,4-β-mannosidase (Man2A) and one intracellular 1,6-α-galactosidase. Galactomannan, mannan, mannose and galactose were all good carbon sources for *C. fimi*. The secreted β-mannanase can be either cell-bound or released into the culture supernatant as an intact enzyme or as proteolytically processed enzyme. The interaction of Man26A with the cell envelope, presumably with a polysaccharide thereof (Egelseer *et al.*, 1998), is mediated by the SLH domain (S-layer homology domain). This is the first report of a cell-bound *C. fimi* polysaccharidase. Man26A also comprises a substrate binding domain, i.e., the mannan-binding domain (MBD). This MBD does not bind to crystalline but does bind to soluble mannan, similar to CBD_{N1} from the *C. fimi* endoglucanase CenC (Tomme *et al.*, 1996b). Possibly, the MBD is involved in the semi-processive hydrolysis mechanism of mannan by Man26A, by "feeding" soluble, single strands of substrate to the catalytic domain as was proposed for *Thermomonospora fusca* E4 endoglucanase (Irwin *et al.*, 1998). Man26A hydrolyzes galactomannan into a variety of oligosaccharides, which for further hydrolysis have to be imported into the cells. The cellulase system hydrolyses cellulose to cellobiose outside the cells and presumably imports only cellobiose and glucose. Within the cells the galactomannooligosaccharides are hydrolyzed into monomers by the concerted action of the β-mannosidase, Man2A and the 120 kDa α-galactosidase. The amount of Man26A (or its catalytic domain) released into the

culture medium could be dependent on the concentration of secreted *C. fimi* protease. The proposed hydrolysis mechanism of galactomannan hydrolysis by *C. fimi* cultures is summarized in Figure 5.1:

The presence of a substrate-binding domain (MBD) and a cell adhesion SLH domain in Man26A promotes binding of the cells to the substrate enabling efficient uptake of the galactomannooligosaccharides produced by Man26A into the cells where they are further hydrolyzed into monomeric mannose and galactose by the α-galactosidase and the β-mannosidase.

As the cell density increases in the culture (Figure 5A) the concentration of Man26A increases as well, with a concomitant increase in galactomannooligosaccharide product concentration. With the increased product concentration in the environment the cells no longer have to be very close to the substrate to guarantee sufficient oligosaccharide uptake. Therefore the cells can afford to release Man26A into the culture supernatant, which possibly is regulated by proteolytic processing of Man26A and is therefore dependent on the concentration of secreted *C. fimi* protease. At high cell density the concentration of secreted protease is presumably also increased, thereby processing and hence releasing more Man26A into the culture supernatant. In this proposed mannan-degrading system of *C. fimi* it is possible that the proteolytically processed mannanase takes over the "endo" hydrolysis of the substrate and the cell-bound unprocessed mannanase acts as the "exo"-enzyme.

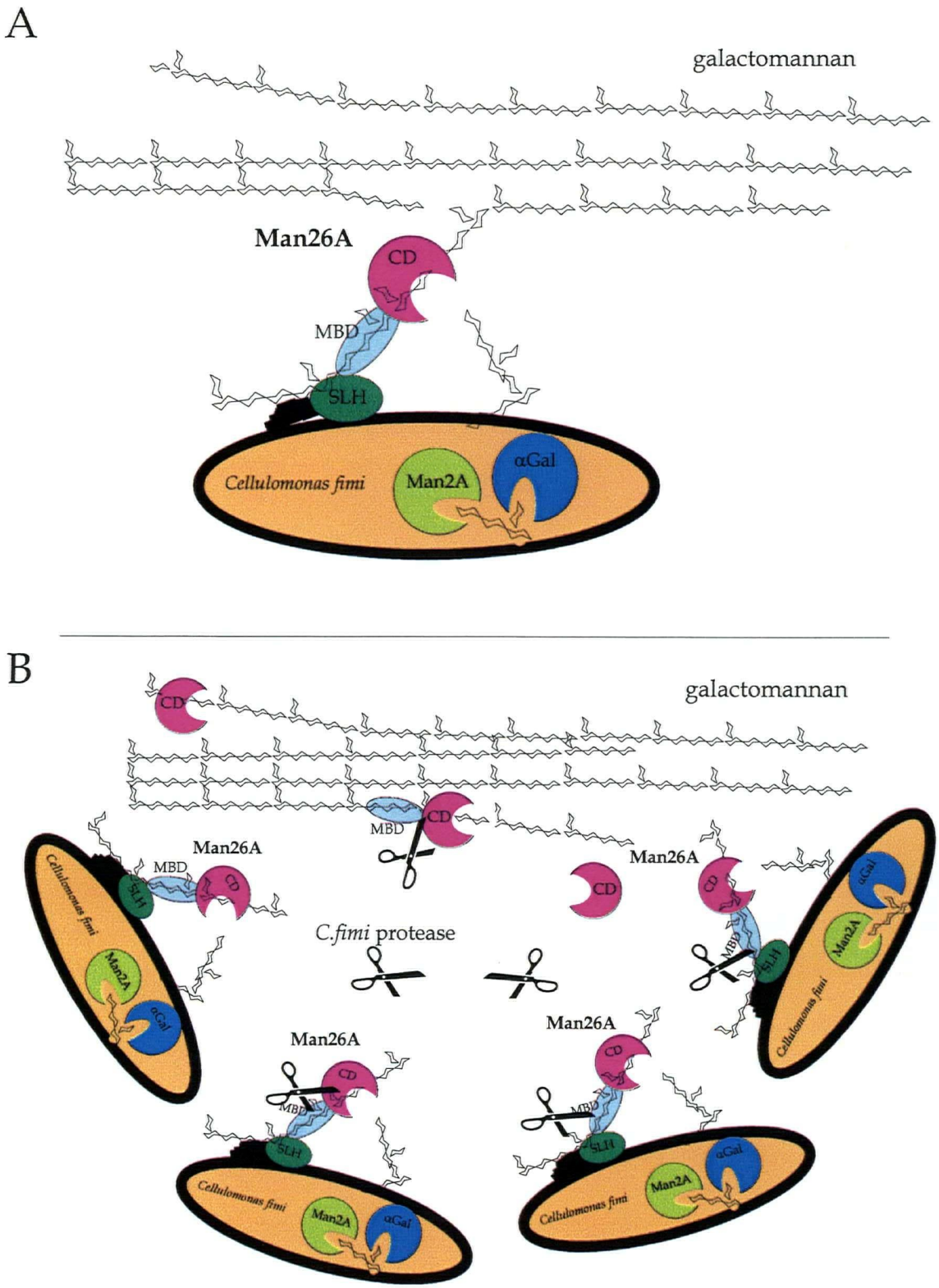


Figure 5.1: Schematic mannan degradation by *Cellulomonas fimi*. CD: catalytic domain, SLH: SLH domain, MBD, mannan-binding domain; αGal: α-galactosidase.

6 Conversion of Man2A into a glycosynthase

6.1 Introduction

6.1.1 Enzymatic synthesis of oligosaccharides

Oligosaccharides play important roles in a range of biological processes such as cell recognition, growth and differentiation (Edelman, 1983) and pathogen adherence (Section 4.3.3). The use of oligosaccharides in studies of these processes and their prospective use as therapeutics, is the driving force for the development of methodologies for the efficient synthesis of oligosaccharides. The synthesis of such hetero-oligosaccharides with regio- and stereospecific interglycosidic linkages by means of classical carbohydrate chemistry is a demanding task (Thiem *et al.*, 1995). Therefore different strategies for the enzymatic formation of oligosaccharides using either glycosyl transferases (EC 2.4.-) glycosyl hydrolases (EC 3.2.-) or glycosynthases are being investigated.

6.1.1.1 Glycosyl transferases

Glycosyl transferases are enzymes that are responsible for the biosynthesis of oligo- and polysaccharides. They catalyze the specific transfer of a monosaccharide from a nucleotide donor sugar to the acceptor substrate (Ichikawa *et al.*, 1992). Several disadvantages limit the use of these enzymes in larger scale *in vitro* processes. A common problem is the cloning, high level expression and stability of specific transferases. Furthermore the high cost of glycosyl nucleotides may be prohibitive, although complex, but effective, recycling schemes have been developed, thereby also reducing the problem of

product inhibition. Another problem with glycosyl transferases is their frequently rigid specificities, limiting their use with unnatural substrates (Ichikawa *et al.*, 1992).

6.1.1.2 Glycosidases

Retaining glycosidases (Section 1.2) are also useful tools in oligosaccharide synthesis. A wide variety of these enzymes is available (Section 1.3), and, unlike glycosyl transferases, they do not require expensive activated sugar donors for oligosaccharide synthesis. Furthermore, their high glycon- and low aglycon specificities allow the synthesis of unnatural substrates (Murata *et al.*, 1997, Yasukochi *et al.*, 1997). The formation of the glycosyl-enzyme intermediate in the synthetic pathway of retaining glycosidases (Figure 6.1, step 1) corresponds to the glycosylation step of the hydrolytic pathway (Figure 1.1). In the deglycosylation step of the synthetic pathway however, the acceptor sugar competes with water for the nucleophilic attack at the anomeric center of the reactive glycosyl-enzyme complex (Figure 6.1). When an acceptor sugar acts as nucleophile the reaction mechanism is referred to as transglycosylation (Toone *et al.*, 1989). A major disadvantage of this method is that newly formed oligosaccharides can be immediately hydrolyzed by the glycosidase. One way to reduce oligosaccharide hydrolysis is by use of readily synthesized, cheap donor sugars, such as nitrophenyl glycosides or glycosyl fluorides. Preferential cleavage of these donor sugars by the enzyme reduces the overall hydrolysis of the synthesized oligosaccharides (Ichikawa *et al.*, 1992). Other means of improving reaction yields are the selective removal of the newly formed products from the reaction or by lowering the activity of water, using organic solvents (Fan *et al.*, 1995). However, even with these improvements reaction yields are typically low.

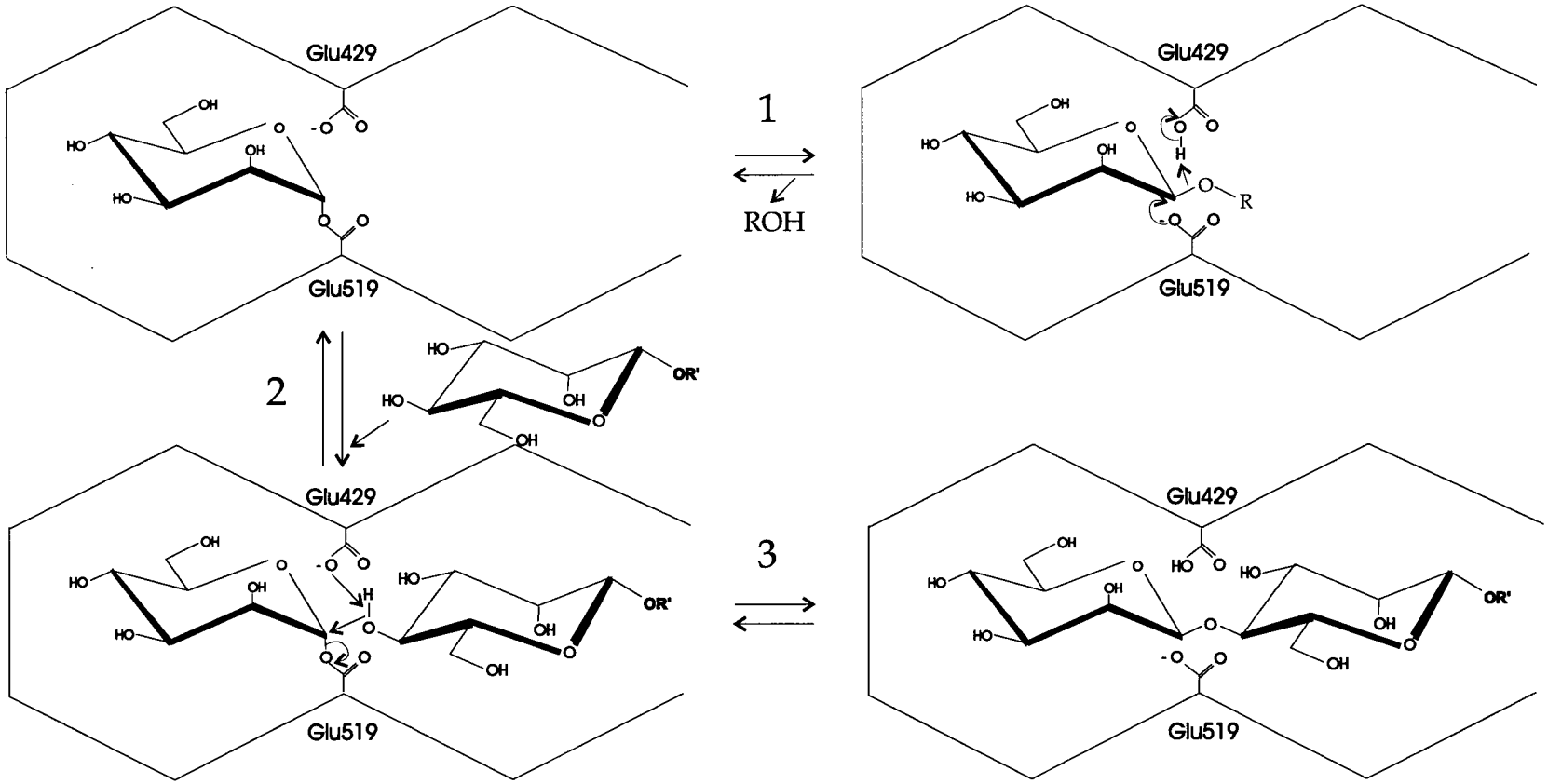


Figure 6.1: Proposed transglycosylation mechanism (reaction step 1-3) by Man2A WT. The acid/base catalyst (Glu429) and the catalytic nucleophile (Glu519) are shown. The leaving group R can be a glycoside or *p*-nitrophenyl group. The transglycosylation reaction is shown with Man βR' as the acceptor molecule. R' can be a hydrogen, a glycoside or a *p*-nitrophenyl group.

6.1.1.3 Glycosynthases

A third and very novel approach to enzymatic oligosaccharide synthesis is the use of glycosynthases. Glycosynthases are mutant forms of glycosidases that are able to synthesize oligosaccharides but can no longer hydrolyze glycosidic linkages (Mackenzie *et al.*, 1998). In the case of retaining glycosidases, the catalytic nucleophile is replaced with a non-nucleophilic amino acid side chain, leaving the rest of the active site, glycon and aglycon site and the acid/base catalyst unaltered. These mutant proteins no longer form covalent enzyme-sugar intermediates, which in wild type glycosidases are formed by the attack of the catalytic nucleophile with inversion of stereochemistry at the anomeric center (Figure 1.1). The catalytic requirement for a glycosyl-enzyme intermediate can be overcome by using activated donor molecules with small aglycon/leaving groups, such as readily synthesized glycosyl fluorides. The anomeric configuration of the activated donor has to be inverted with respect to the natural substrate in order to mimic a glycosyl-enzyme intermediate (Figure 6.2). The binding of an acceptor sugar instead of water in the glycosynthase aglycon site provides the additionally required activation energy to catalyze the release of the leaving group (i.e., fluoride ion) and the formation of the new glycosidic bond.

This approach of using glycosynthases in oligosaccharide synthesis includes all the advantages of using glycosidases for oligosaccharide synthesis, but eliminates most of the disadvantages (Section 6.1.1.2). Prior to this study, the concept of glycosynthase catalyzed oligosaccharide synthesis had been demonstrated only with the glycosynthase derived from the *Agrobacterium* sp. β-glucosidase/galactosidase (Abg) (Mackenzie *et al.*, 1998). This glycosynthase was constructed by mutating the active site nucleophile, glutamate 358, to an

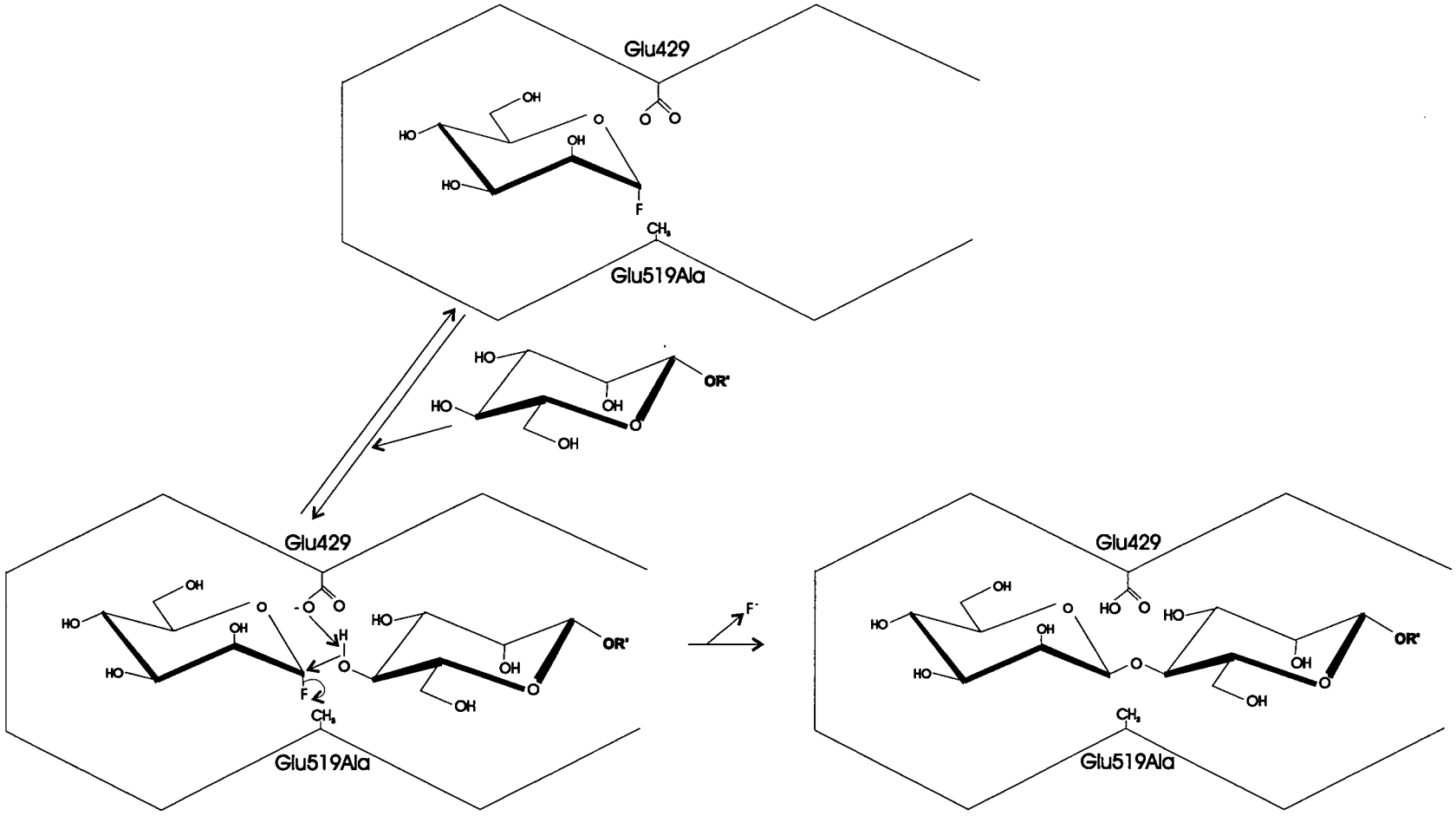


Figure 6.2: Proposed transglycosylation mechanism by glycosynthase Man2A E519A. In this example α–mannosyl fluoride is the activated donor sugar and Man β1R' the acceptor. R' can be a hydrogen, glycoside or a *p*-nitrophenyl group.

alanine (Abg E358A). Abg E358A-catalyzed oligosaccharide synthesis was demonstrated with α-galactosyl fluoride or α-glucosyl fluoride as activated donor sugars and a variety of aryl-glycoside acceptors. Product yields between 64 % and 92 % were obtained (Mackenzie *et al.*, 1998). These product yields demonstrate the great potential for the glycosynthase approach for enzymatic oligosaccharide synthesis. However, with Abg E358A being the first and only report of a glycosynthase, the general applicability of this concept had yet to be proved.

6.1.2 Identification of the catalytic nucleophile

In the absence of structural information the first step towards the conversion of a retaining glycosidase into a glycosynthase is the identification of its catalytic nucleophile. The covalent glycosyl-enzyme intermediate of retaining glycosidases can be trapped with mechanism-based inactivators if the rate of intermediate formation (i.e., glycosylation), is significantly faster than the rate of deglycosylation. Conduritol epoxides, which incorporate an endocyclic epoxide within a cyclitol ring, form one class of such inactivators (Legler 1990). Protonation of the epoxide and the concomitant attack by the enzymic nucleophile leads to epoxide ring opening and covalently links the inactivator to the nucleophile, forming a stable intermediate. The use of conduritol epoxides however, has led to the misassignment of active site residues in several cases, e.g. *E. coli (lacZ)* β-galactosidase (Gebler *et al.*, 1992) and human lysosomal β-glucocerebrosidase (Miao *et al.*, 1994).

Another class of irreversible mechanism-based inactivators is the activated 2-deoxy-2-fluoro glycosides. The substitution of the hydroxyl group at the glycoside 2 position with a fluorine destabilizes the transition state, which slows both the glycosylation and

deglycosylation steps (McCarter *et al.*, 1992, Namchuk *et al.*, 1995). The glycosylation step however, can be accelerated relative to the deglycosylation step by adding a good leaving group, such as 2,4-dinitrophenolate or fluorine, to the sugar acetal center. This causes not only an accumulation of the glycosyl-enzyme intermediate but also an inactivation of the enzyme. Using a combined approach of reverse phase high pressure liquid chromatography (RP-HPLC) and electrospray ionization (ES) tandem mass spectrometry (MS/MS) the labeled catalytic nucleophile can be identified (Withers and Aebersold, 1995). In these techniques, pepsin digested glycosidase, labeled with inactivator or unlabeled, is analayzed by ESMS for peptides with a mass increase, corresponding to the label, by either comparative mapping or by neutral loss (6.2.1.5). The sequence of the labeled peptide can then be determined either by Edman degradation or by MS/MS. The MS/MS sequencing technique is based on collision induced fragmentation of the labeled parent peptide producing daughter ion fragments, whose masses are analyzed. With the deduced glycosidase amino acid sequence available and the mass results from the peptide and the fragments thereof, the sequence of the labeled peptide and the catalytic nucleophile can generally be determined (Withers and Aebersold, 1995).

6.1.3 Objectives

The aim of this project was to transform the *C. fimi* mannosidase, Man2A, into a glycosynthase and to test its ability to synthesize oligosaccharides and to prove that the concept of glycosynthases is generally applicable. The glycosynthase was produced by mutating the carboxylic catalytic nucleophile, which was identified by using the mechanism-based inhibitor MS/MS approach, to an alanine.

6.2 Results

6.2.1 Identification of the catalytic nucleophile in the *C. fimi* mannosidase, Man2A

6.2.1.1 Prediction of the position of catalytic residues

The conversion of a retaining glycosidase into a glycosynthase by mutating the catalytic nucleophile to an alanine requires knowledge of the exact position of the nucleophile. In many cases the nucleophile can be accurately predicted by alignments of the amino acid sequences of members of the same glycosidase family (Figure 3.7). If the catalytic residues were determined for one member of the family, from either 3D structures, labeling studies or from kinetic studies of catalytic residue mutants, the catalytic residues for new members of the family could be inferred if the sequences were sufficiently similar. This method could be used to predict the catalytic residues for the *C. fimi* mannanase Man26A (Figure 3.7). In case of insufficient sequence similarity, secondary structure predictions, e.g. hydrophobic cluster analysis (HCA), can be used to look for common secondary structure motives (Section 1.3). This method however, is not always accurate and has lead to misassignments of catalytic nucleophiles (Henrissat *et al.*, 1995 and Vocadlo, personal communication).

The *C. fimi* mannosidase Man2A is a member of family 2, a retaining glycosyl hydrolase family (Section 3.2.9). Three homology groups can be defined within family 2: the β-mannosidase-, β-glucuronidase- and β-galactosidase subfamilies. Multiple sequence alignment of members of the β-mannosidase subfamily revealed at least two conserved carboxylic amino acids that might function as the catalytic nucleophile (Figure 3.18). Multiple alignment of glycosidase sequences from all three subfamilies, however, showed

insufficient sequence identity C-terminal of the acid/base catalyst to accurately predict the catalytic nucleophile position in Man2A. In Figure 6.3 an alignment of the sequence around the catalytic residues of five family 2 members is shown. The acid/base catalyst, which is part of the family 2 consensus pattern W-[GS]-x (2,3)-N-E, is conserved in all members shown (Section 3.2.11). In *E. coli (lacZ)* β-galactosidase, the glutamate in $_{536}$CEY$_{538}$, was experimentally identified as catalytic nucleophile and, based on kinetic data Glu461 was identified as the acid/base catalyst (Gebler *et al.*, 1992). From comparisons of the 3D structure of *E. coli (lacZ)* β-galactosidase (Jacobsen *et al.*, 1994) and the 3D structure of human lysosomal β-glucuronidase, the acid/base catalyst and the nucleophile were predicted for the β-glucuronidase to be Glu451 and Glu519, respectively (Jain *et al.*, 1996).

The HCA plots of several lysosomal glycosidases, members of the superfamily GH-A, were compared in order to find common motif and to predict the catalytic residues. For the bovine β-mannosidase, Glu457 was predicted to act as acid/base catalyst and Glu554 as nucleophile. These residues are located in the $(\beta/\alpha)_8$ barrel at the C-terminal ends of the β-4 and β-7 strands, respectively (Figure 3.19; Durand *et al.*, 1997). Although these residues are conserved in Man2A and correspond to Glu429 and Glu519, it was considered to be important to accurately identify the nucleophile in Man2A and to examine its role in catalysis.

6.2.1.2 Inactivation of Man2A using 2-deoxy-2-fluoro-β-mannosyl fluoride

Initial studies indicated that 2-deoxy-2-fluoro-β mannosyl fluoride (2FManβF) was a potent inhibitor of the *C. fimi* β-mannosidase, Man2A (data not shown). Complete inactivation of Man2A, measured as loss of PNPM hydrolytic activity, could be obtained by

Figure 6.3: Multiple sequence alignment of glycosidases from family 2. The region around the acid/base catalyst (A/B) and the catalytic nucleophile (Nu) is shown. The roles of the these residues (boxed) were either determined experimentally, or obtained from 3D structures or from HCA predictions. ChMad: Goat beta-mannosidse; CfMad: *C. fimi* beta-annosidase; HGus: human beta-glucuronidase; EcLacZ: *E. coli (lacZ)* beta-galactosidase. Conserved amino acids are highlighted.

incubating the enzyme with excess inactivator. 2FManβF was known to inhibit β-mannosidase activity in rats (*in vivo*) and in rat tissue homogenate (*in vitro*), but not the purified enzyme (McCarter *et al.*, 1994). According to the proposed inactivation mechanism, one fluoride ion is released per inactivated enzyme molecule (Figure 6.4 A and B; McCarter *et al.*, 1992; Street *et al.*, 1992). Therefore, Man2A inactivation was determined by measuring the changes in fluoride concentration with a fluoride electrode. Inactivator (2FManβF) concentrations ranging from 19.5 μM to 520 μM were tested. The rate constant (k_{obs}) for each inactivator concentration was obtained by non-linear regression analysis of the experimental values. An example of time-dependent inactivation is shown in Figure 6.5 A. 8.3 μM Man2A were inactivated in the presence of 520 μM 2FManβF. From the plot of k_{obs} versus inhibitor concentration, the inhibition rate constant k_I = 0.556 ($\pm$ 0.024) min^{-1} and the dissociation constant K_I = 0.412 ($\pm$0.033) mM were obtained by non-linear regression analysis (according to equation E 6.1: $k_{obs} = (k_i \times [I])/ (K_i + [I])$). For the purpose of illustration, k_{obs} *versus* inhibitor concentration was replotted in a double reciprocal plot (Figure 6.5 B). Complete inactivation of 8.3 μM Man2A, at inactivator concentrations of 260 μM, 390 μM and 520 μM, produced an average final concentration of released fluoride of 8.03 μM. Consistent with the inactivation scheme (Figure 6.4 A and B), inactivation of Man2A by 2FManβF followed a 1:1 stoichiometry, was mechanism-based, most likely active-site directed, and the inactivation rate was dependent on 2FManβF concentration in a saturable manner.

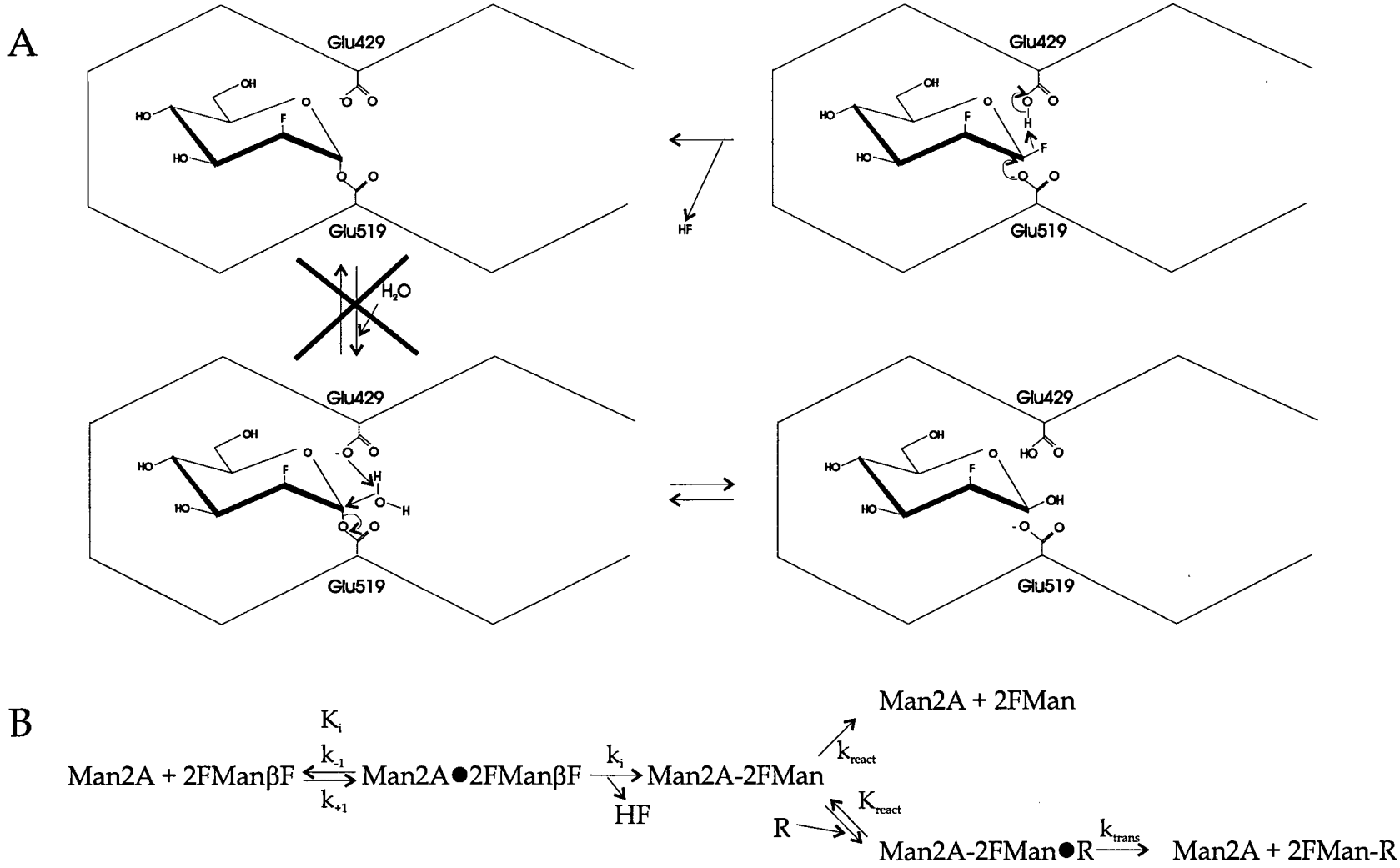


Figure 6.4: Panel A: Proposed inactivation mechanism of Man2A with 2FMan βF. Panel B: Inactivation scheme for the reaction pathway illustrated in Panel A including spontaneous reactivation and reactivation by transglycosylation. ● indicates a Michaelis-Menten complex. R can be e.g. PNPM.

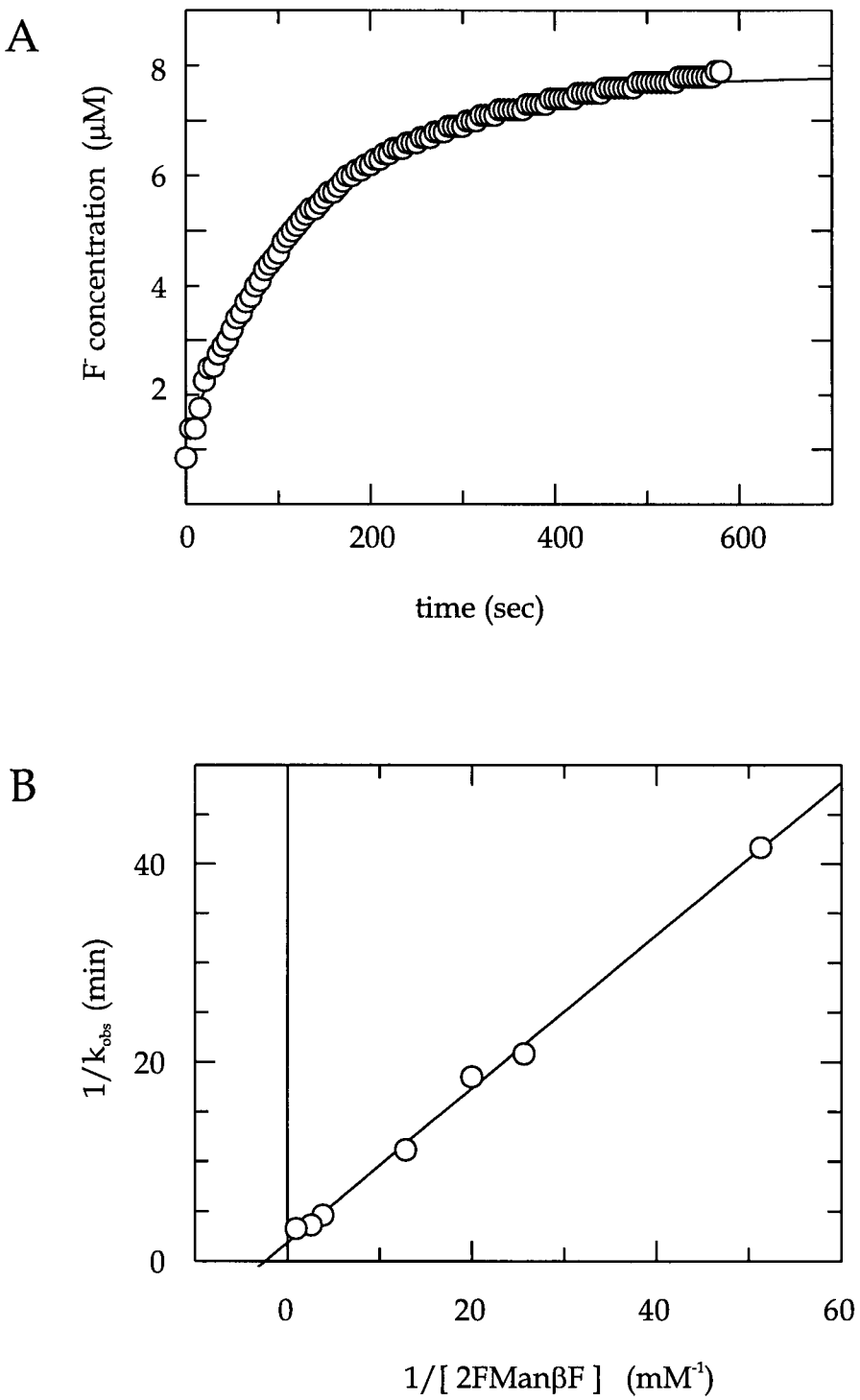
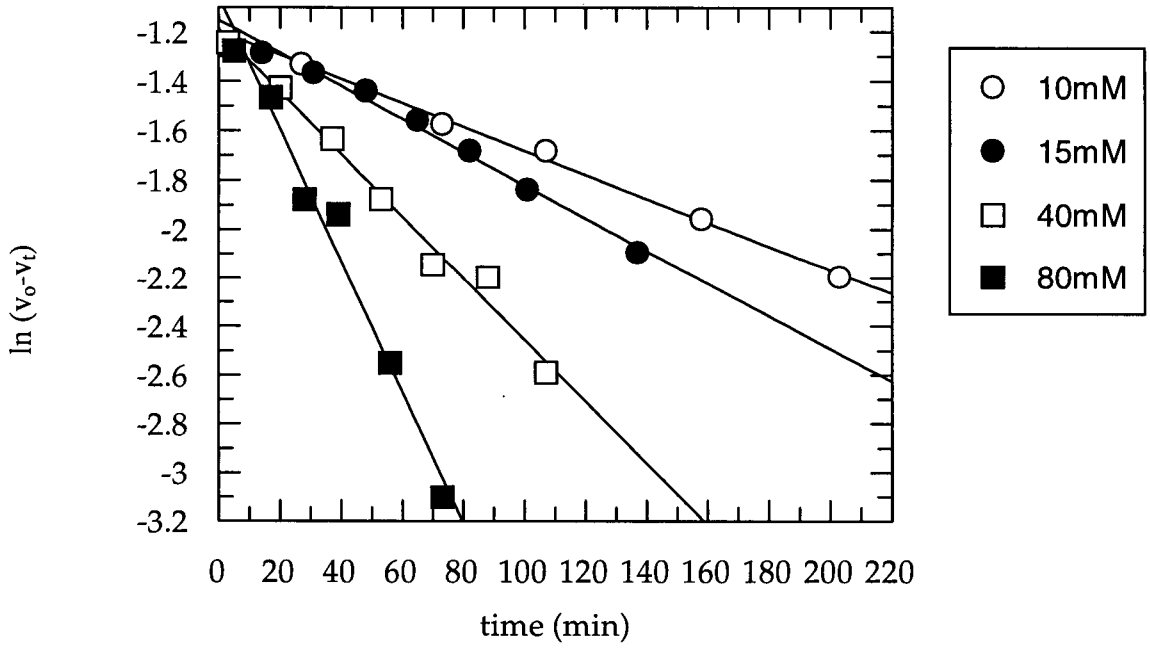


Figure 6.5: Man2A inactivation with 2FManβF. Panel A: example of time-dependent inactivation of 8.3 μM Man2A with 520 μM 2FManβF. k_{obs} were calculated for inactivation reactions with 2FManβF concentrations ranging from 19.5 μM to 520 μM.
Panel B: Double reciprocal plot of first order rate constants (k_{obs}) *versus* 2FManβF concentration.

6.2.1.3 Reactivation of inactivated Man2A

Catalytic competence of the 2FMan-enzyme intermediate was demonstrated by incubation of inactivated enzyme (freed of excess inactivator) in either buffer only or in buffer plus a reactivator, such as the uncleavable glucose disaccharide gentiobiose (Glu β-1,6 Glu). From the reactivation reactions (incubated at 37^0 C) aliquots were removed to measure the regain of PNPM hydrolytic activity due to regeneration of free enzyme. The reactivation in buffer followed a first order process with an apparent rate constant of k_{react} = 0.002 min^{-1}, which corresponded to a half-life of the Man2A-2FMan intermediate of 344 min. The reactivation process was significantly accelerated in the presence of gentiobiose. Reactivation with gentiobiose was gentiobiose concentration-dependent and followed pseudo-first order kinetics. Apparent rate constants (k_{obs}) were determined for gentiobiose concentrations of 10 mM, 15 mM, 40 mM and 80 mM (Figure 6.6 A). From the double reciprocal replot of k_{obs} versus gentiobiose concentration, the reactivation rate constant, k_{trans} = 0.043±0.014 min^{-1} and the dissociation constant K_{trans} = 78±25 mM was estimated (Figure 6.6 B). The increased reactivation rate with gentiobiose suggested that this sugar facilitates turnover of the intermediate, probably, as was shown in other systems, *via* transglycosylation (Street *et al.*, 1992). This sugar-dependent reactivation rate enhancement was used to screen glycoside libraries to find substrates that can act as good acceptor molecules for the glycosynthase catalyzed transglycosylation (Section 6.2.5).

A



B

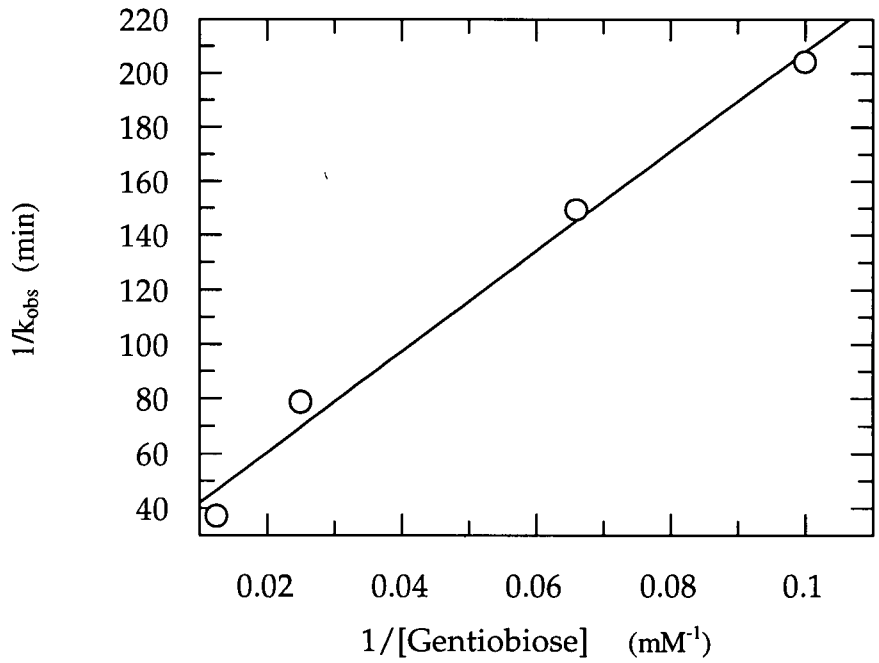


Figure 6.6: 2FMan-Man2A reactivation. Panel A: Time-dependent reactivation of 2FMan-Man2A with 10 mM, 15 mM, 40 mM and 80 mM gentiobiose. Activity *versus* time is shown on a semilogaritmic plot.
Panel B: Pseudo-first order rate constants (k_{obs}) are reploted on a double reciprocal plot *versus* gentiobiose concentration.

6.2.1.4 Analysis of the enzyme-inactivator complex by mass spectrometry

Native and inactivated Man2A were analyzed by ESMS. The native enzyme had a MW of 94,980 (94,960 predicted), whereas the inactivated enzyme, "labeled" with 2FManβF had a MW of 95,139 (95124 predicted). The MW difference between the unlabeled and labeled enzyme of 159 was consistent, within error, with the predicted difference of 164, corresponding to the mass of 2FMan. These results confirmed the 1:1 stoichiometry of the inactivation (Section 6.2.1.2), and further indicated that the inactivator was covalently linked to the enzyme. These findings are in good agreement with the proposed inactivation mechanism (Figure 6.4).

6.2.1.5 Identification of the labeled active site nucleophile by ESMS

Peptide mixtures from pepsin digests of native and 2FMan-labeled Man2A were separated by RP-HPLC and analyzed by ESMS in LC/MS mode (Section 6.1.2 and 2.21). The total ion chromatograms (TIC) showed a large number of peaks, each of which corresponded to one or more peptides (Figure 6.7 A and D). One method to identify the labeled peptide in the peptide mix is neutral loss tandem mass spectrometry. Collision-induced fragmentation of the labeled peptide results in the cleavage of only the label-carboxyl ester bond, which generally represents the weakest linkage within the peptide. Peptides losing a predetermined mass (i.e., the mass of the label), can then be selected for by tandem mass spectrometry (MS/MS). An attempt to identify the 2FMan-labeled peptide from Man2A by MS/MS in neutral loss mode was unsuccessful. Therefore the mass spectra obtained from the labeled and from the unlabeled peptic digests were analyzed by

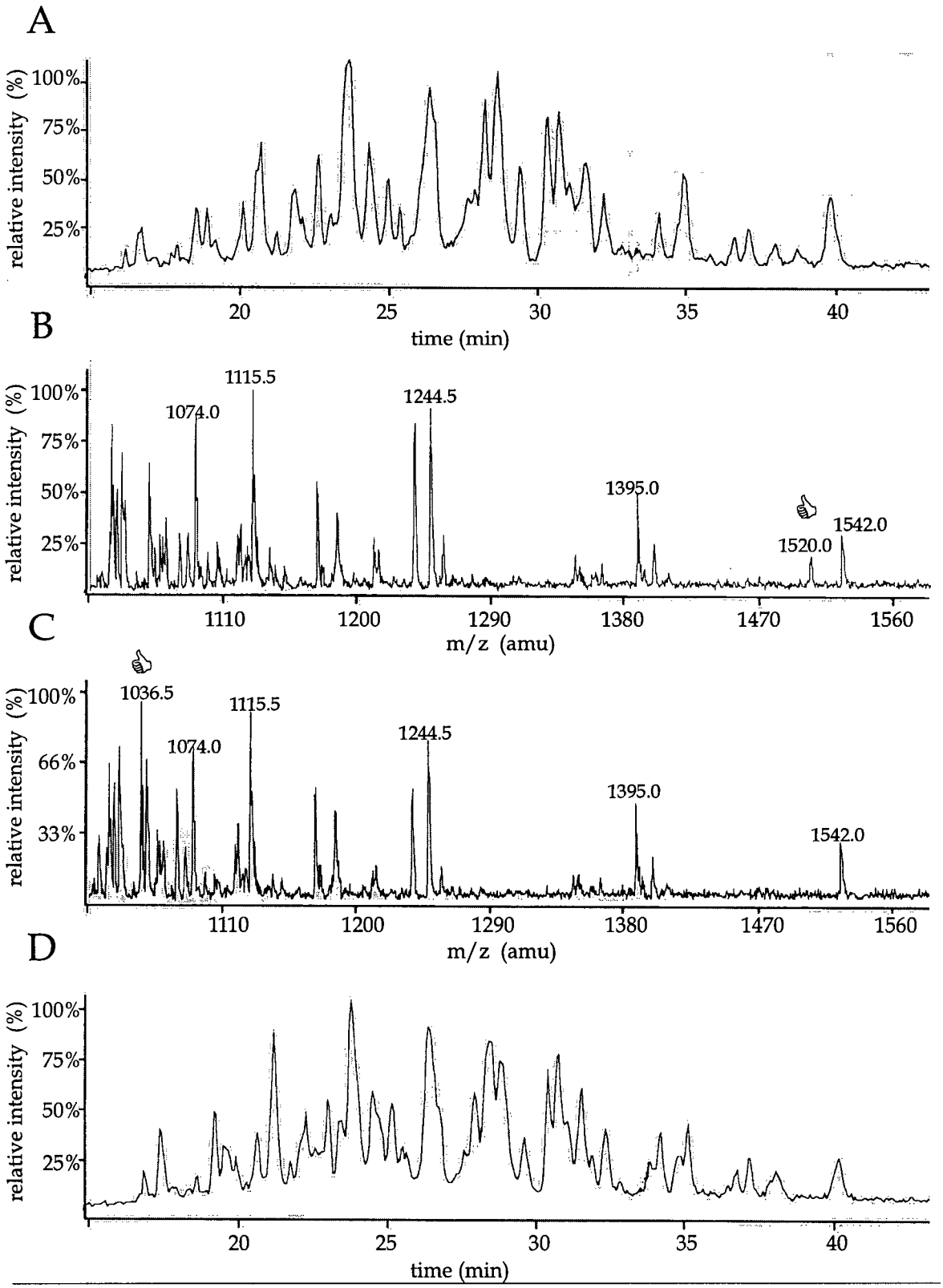


Figure 6.7: Comparison mapping of labeled and unlabeled Man2A pepsin digests. Total ion chromatograms (TIC) from labeled and unlabeled pepsin digests are shown in Panel A and Panel D, respectively. Panel B: Mass spectrum of labeled digest taken at 28.6 min. Panel C: Mass spectrum of unlabeled digest taken at 28.8 min. ☜: indicates the labeled and the corresponding unlabeled peptides.

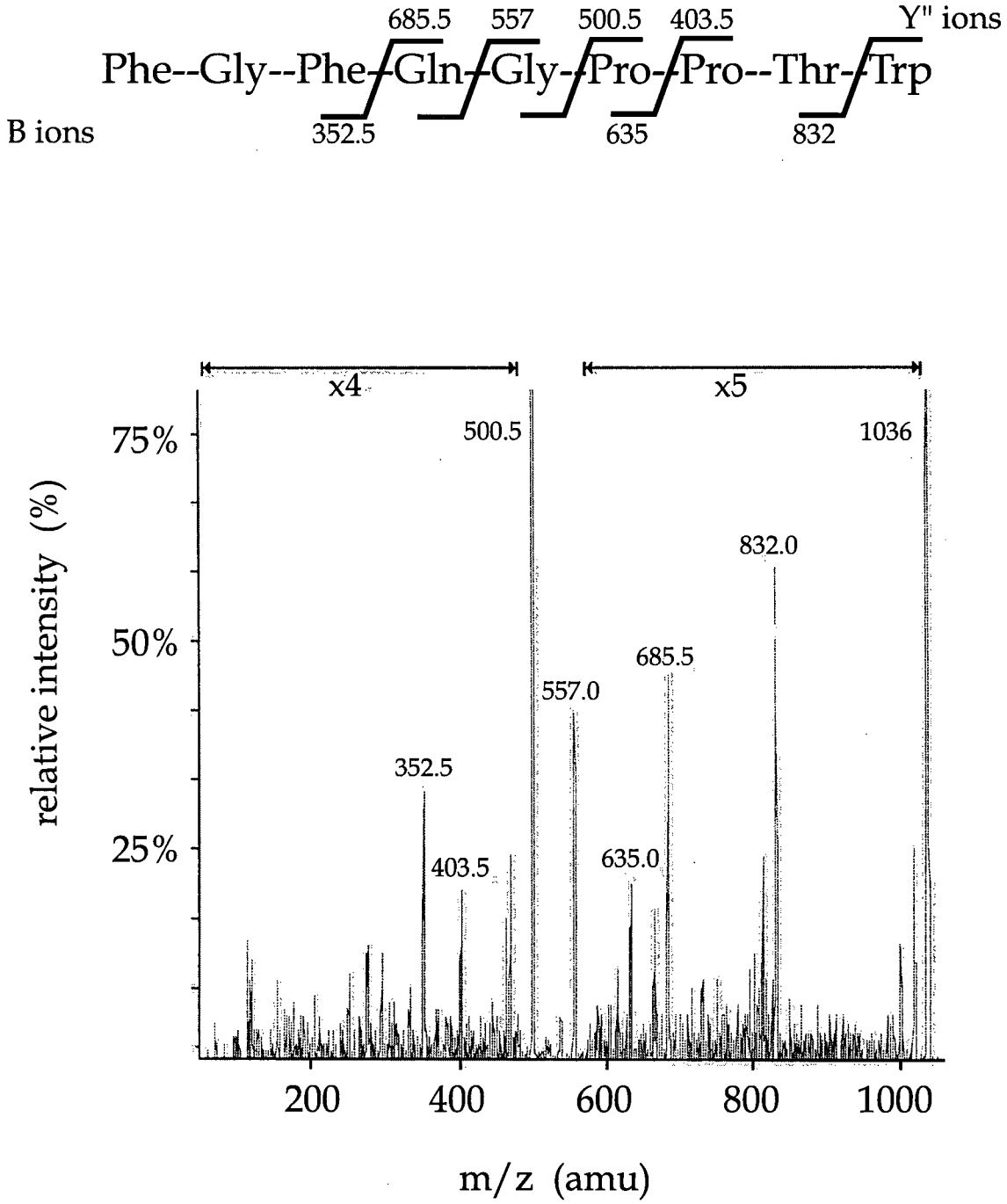


Figure 6.8: ESMS/MS daughter ion spectrum of the unlabeled peptide (MW 1036). The peptide sequence is indicated with the fragment sizes of the B and the Y " ions shown below and above the sequence, respectively.

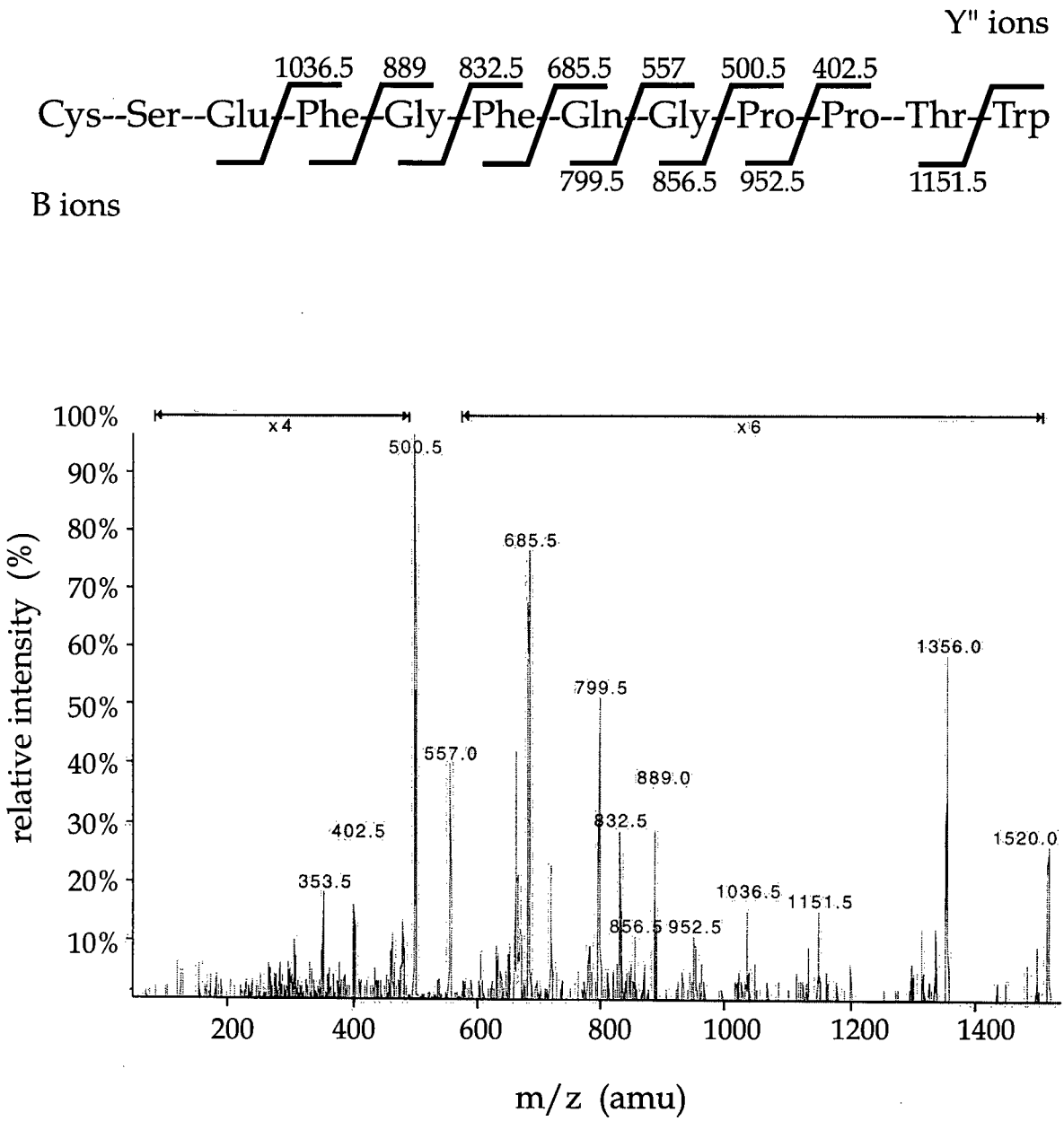


Figure 6.9: ESMS/MS daughter ion spectrum of the labeled peptide (MW 1520). The peptide sequence is indicated with the fragment sizes of the unlabeled B and the Y" ions shown below and above the sequence, respectively.

comparative mapping (Figure 6.7 B and C). A peptide with a MW of 1520 (peptide 1520) was found in the labeled, but not in the unlabeled sample (Figure 6.7 B), whereas a peptide with a MW of 1036.5 (peptide 1036.5) was found in the unlabeled, but not in the labeled peptic digest (Figure 6.7 C). The difference in MW of the two peptides was 483.5, which did not correspond to the expected difference from the label of 164 (*vide infra*).

Since only small quantities of peptide 1520 were purified, the more abundant peptide, peptide 1036.5 from the unlabeled peptide mix was analyzed. Amino acid sequence information was obtained by collision-induced fragmentation (CIF), and mass data were collected in daughter ion scan mode (Figure 6.8). The sequence for peptide 1036.5 was identified as $_{520}$F G F Q G P P T W$_{528}$. This sequence did not contain any carboxylic amino acid that might act as catalytic nucleophile. However, the N-terminus of this peptide was immediately C-terminal of the predicted nucleophile E519. By purifying more of peptide 1520 its sequence was identified by CIF as $_{517}$C S E F G F Q G P P T W$_{528}$ (Figure 6.8). The MW difference between peptide 1520 and peptide 1036.5 corresponded to the MW of the three additional amino acids found in peptide 1520 plus the MW of the label. Since only carboxylic amino acid residues can act as catalytic nucleophiles (Davies *et al.*, 1998), and only one carboxylic amino acid residue was present in the labeled peptide 1520, glutamate E519 was identified as the catalytic nucleophile.

6.2.2 Conversion of Man2A into a Glycosynthase: Mutation E519A

To convert the *C. fimi* mannosidase Man2A into a glycosynthase, the catalytic nucleophile, E519, was mutated to an alanine. This mutation was not only designed to completely abolish hydrolytic activity, but also to create a pocket within the active site, to

allow the use of activated α-sugar donors in transglycosylation reactions (Section 6.1.1.3; Mackenzie *et al.*, 1998).

The scheme for *in vitro* mutagenesis and generation of plasmid pET28aMadE519A, which encodes the Man2A glycosynthase, is illustrated in Figure 6.10. Expression and purification procedures were the same as used for the wild type mannosidase (Section 2.8). The yields (approx. 300 mg/L) and purity (>96 %) of the glycosynthase were comparable to Man2A WT (Figure 3.24). Extreme caution had to be taken not to contaminate the Man2A E519A mutant with Man2A during purification. Purified Man2A E519A did not hydrolyze PNPM, even after a 5 day incubation period with high enzyme concentration (data not shown). The mass of Man2A E519A, with a predicted MW of 94902, was identified by ESMS as 94918.

6.2.3 Transglycosylation reaction by Man2A WT

As has been reviewed in Section 6.1.1.2, retaining glycosidases are not only able to hydrolyze glycosidic linkages, but they are also able to form new glycosidic linkages. Transglycosylation products, hydrolysis substrates themselves, might never accumulate and be detected because hydrolysis is generally much faster than transglycosylation (Harjunpää *et al.*, 1995). One way to reduce the rate of transglycosylation product hydrolysis is to use a substrate with a good leaving group at high concentration, and low enzyme concentrations (Gusakov, *et al.*, 1991; Namchuk *et al.*, 1995). PNPM was used as the substrate in transglycosylation reactions with Man2A WT. Man2A WT 0.25 nmol/mL was added to a saturated PNPM solution (7.5 mg/mL), which was buffered with 100 mM

Figure 6.10 (following page): Generation of pETMadE519A, encoding the glycosynthase Man2A E519A with a C-terminal H_6 tag. PCR mutagenesis of the catalytic nucleophile encoding codon and the three subcloning steps, that were required for the generation of pETMadE519A are illustrated. Only the restriciton sites that were relevant for cloning are shown.

——→ : Mad10 and MadE519A PCR primers

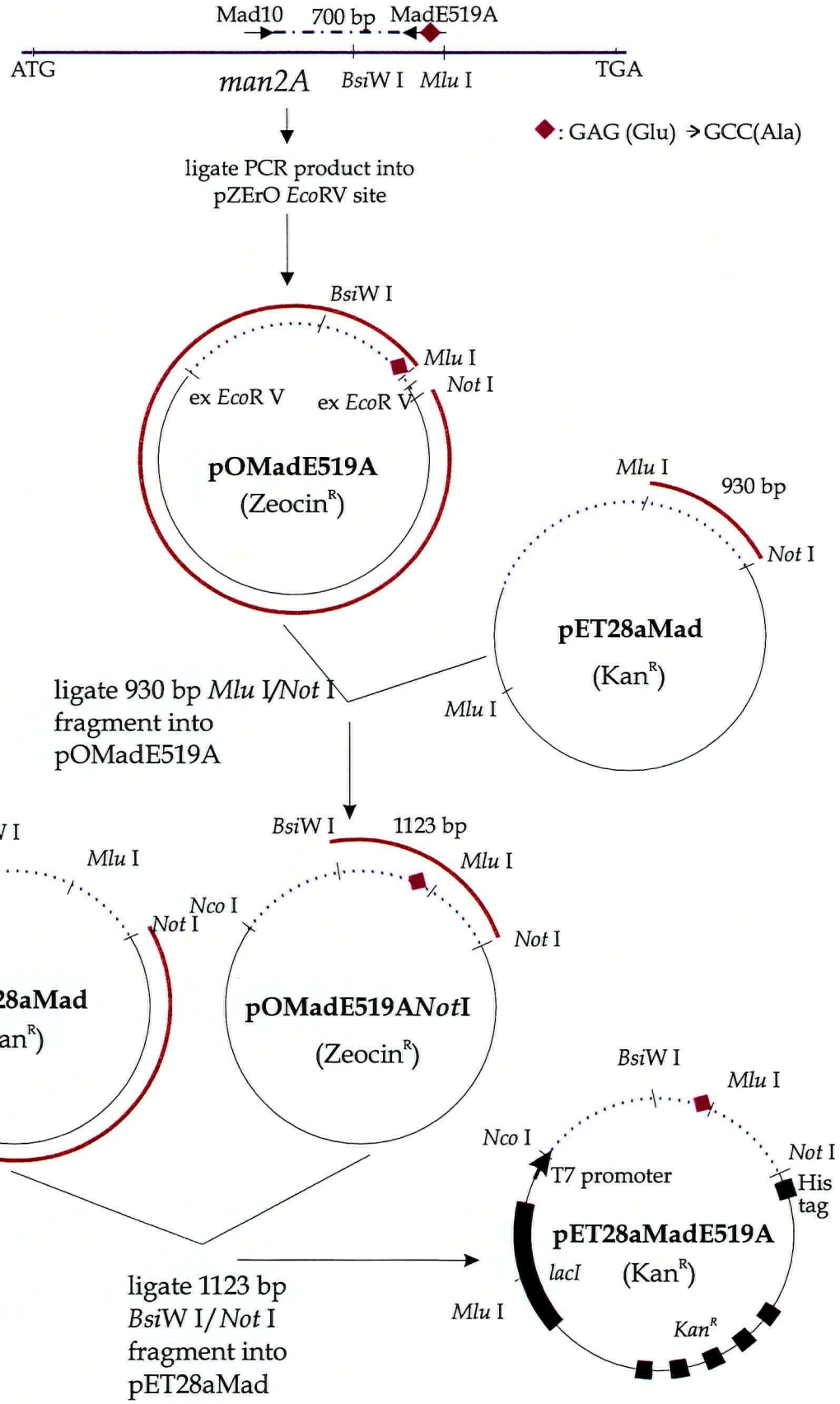
◆ : Mutation introduced by PCR mutagenesis

····· : *man2A* or *man2A* derived plasmid insert

-·-·- : PCR product

——— : *man2A* PCR template

═══ : DNA fragments used for the cloning steps shown. The DNA fragments were obtained by restriction endonuclease digestions, separation on agarose gels and purification by gel extraction.


Mad10
700 bp
MadE519A
ATG
man2A
BsiW I
Mlu I
TGA
◆: GAG (Glu) →GCC(Ala)
ligate PCR product into
pZErO EcoRV site
BsiW I
Mlu I
Not I
ex EcoR V
ex EcoR V
pOMadE519A
(Zeocin^R)
Mlu I
930 bp
Not I
pET28aMad
(Kan^R)
Mlu I
ligate 930 bp Mlu I/Not I
fragment into
pOMadE519A
BsiW I
1123 bp
Mlu I
Nco I
Not I
Not I
pOMadE519ANotI
(Zeocin^R)
BsiW I
Mlu I
Not I
pET28aMad
(Kan^R)
BsiW I
Mlu I
Nco I
Not I
T7 promoter
His
tag
pET28aMadE519A
lacI
(Kan^R)
Mlu I
Kan^R
ligate 1123 bp
BsiW I/Not I
fragment into
pET28aMad

potassium phosphate at pH 7.0,. The reaction was incubated at 33^0 C and the formation of transglycosylation products was monitored by thin layer chromatography (TLC). Mono- and oligosaccharides linked to the para-nitrophenyl group were readily detected under UV light on TLC plates, after separation in ethyl acetate : methanol : water (7:2:1). The major product formed, as identified by TLC, was PNPmannobioside ($PNPM_2$). PNPmannotrioside ($PNPM_3$) was synthesized as well, but in quantities too small for further analysis. The reaction (20 mL) was stopped after 105 min and the products were purified on a silica gel column and separated by RP-HPLC. Two regioisomers of $PNPM_2$ were separated and subsequently analyzed. The results from NMR and ESMS analyses are shown (analysis performed by Lloyd Mackenzie; Section2.23):

4-Nitrophenyl 2,3,4,6-tetra-O-acetyl-β-D-mannopyranosyl-(1,4)-2,3,6-tri-O-acetyl-β-D-manno-pyranoside. ^{1}H NMR (200 MHZ, $CDCl_3$): d 8.20 (d, 2H, $J_{3''',2'''}$ 9.3 Hz, H-3''' Ph), 7.05 (d, 2H,$J_{2''',3'''}$ 9.3 Hz, H-2''' Ph), 5.64 (dd, 1 H, $J_{2,3}$ 3.2 Hz, H2), 5.43 (dd, H, $J_{2',3'}$ 3.4 Hz, H-2'), 5.29 (d, 1 H, $J_{1,2}$ 1.0 Hz, H-1), 5.25 (dd, 1 H, $J_{3,4}$ 9.3 Hz, H-3), 5.22 (dd, 1 H, $J_{4',5'}$ 9.6 Hz, H-4'), 5.03 (dd, 1 H, $J_{3',4'}$ 9.8 Hz, H-3') 4.74 (d, 1 H, $J_{1',2'}$ 1.0 Hz, H-1') 4.40 (dd, 1 H, $J_{6a,6b}$ 12.0 Hz, H-6a), 4.36-4.23 (m, 2 H, H-6a', H-6b'), 4.13 (dd, 1 H, $J_{6b,5}$ 2.7 Hz, H-6b), 4.04 (dd, 1 H, $J_{4,5}$ 9.3 Hz, H-4), 3.88 (m, 1 H, $J_{5,6a}$ 2.7 Hz, H-5), 3.65 (m, 1 H, $J_{55',6'a}$ 5.6 Hz, H-5'), 2.14-1.97 (7 s, 21 H, Ac); Ms (ionspray) 464 (M + 1)

4-Nitrophenyl 2,3,4,6-tetra-O-acetyl-β-D-mannopyranosyl-(1,3)-2,3,6-tri-O-acetyl-β-D-manno-pyranoside. ^{1}H NMR (200 MHZ, $CDCl_3$): d 8.19 (d, 2H, $J_{3''',2'''}$ 9.3 Hz, H-3''' Ph), 7.06 (d, 2H,$J_{2''',3'''}$ 9.3 Hz, H-2''' Ph), 5.62 (dd, 1 H, $J_{2,3}$ 3.1 Hz, H2), 5.32-5.26 (m, 2 H, H-1, H-2'), (dd, 1 H, $J_{4',5''}$ 9.8 Hz, H-4'), 5.14 (dd, 1 H, $J_{4,5}$ 9.0 Hz, H-4), 5.08 (dd, 1 H, $J_{3',2'}$ 3.4 Hz, $J_{3',4'}$ 9.9 Hz, H-3'), 4.75 (d, 1 H, $J_{1',2'}$ 1.0 Hz, H-1'), 4.36-4.22 (m, 3 H, H-6a, H-6a', H-6b'), 4.12 (dd, 1 H, $J_{6b,5}$ 2.7 Hz, $J_{6b,6a}$ 12.3 Hz, H-6b) 4.40 (dd, 1 H, $J_{6a,6b}$ 12.0 Hz, H-6a), 4.36-4.23 (m, 2 H, H-6a', H-

6b'), 4.13 (dd, 1 H, $J_{6b,5}$ 2.7 Hz, H-6b), 4.05 (dd, 1 H, $J_{3,4}$ 9.0 Hz, H-3), 3.88 (m, 1 H, H-5), 3.63 (m, 1 H, H-5'), 2.16-1.97 (7 s, 21 H, Ac); Ms (ionspray) 464 (M + 1)

A molar product yield of 6.7 % and 3.5 % was obtained for β-1,4 linked PNPM2 and β-1,3 linked $PNPM_2$, respectively. By using the glycosynthase Man2A E519A, increased yields in transglycosylation were expected, since no product hydrolysis could occur.

6.2.4 Transglycosylation by glycosynthase Man2A E519A

The glycosynthase, Man2A E519A, was not able to form a covalent glycosyl-enzyme intermediate at position 519 because the catalytic nucleophile (E519) had been mutated to an alanine. Therefore, α-mannosyl fluoride, which imitates the glycosyl-enzyme intermediate, was used as the donor sugar in transglycosylation reactions (Figure 6.2). The aryl glycoside, PNPM, was used as the acceptor molecule. Product formation in the glycosynthase-catalyzed transglycosylation reaction was followed over time by removing aliquots and analysing them by TLC. Transglycosylation with Man2A E519A was observed; however, the reaction proceeded very slowly. From the different conditions tested, it was found that a donor acceptor ratio of 10:1 and a high enzyme concentration gave the most transglycosylation. The pH had a considerable effect on product formation (Figure 6.11). For pH values ranging from 4.0 to 8.0 the pH 5.4 produced the highest levels of transglycosylation products ($PNPM_2$ and $PNPM_3$); pH 5.1 and lower caused precipitation of the enzyme.

To analyze the products formed in the glycosynthase-catalyzed transglycosylation reaction, a 6.3 mL reaction was set up with the following conditions: 41 μM Man2A E519A, 140 mM K-P pH 5.7, 42 mM PNPM and 63 mM α-mannosyl fluoride. The reaction was

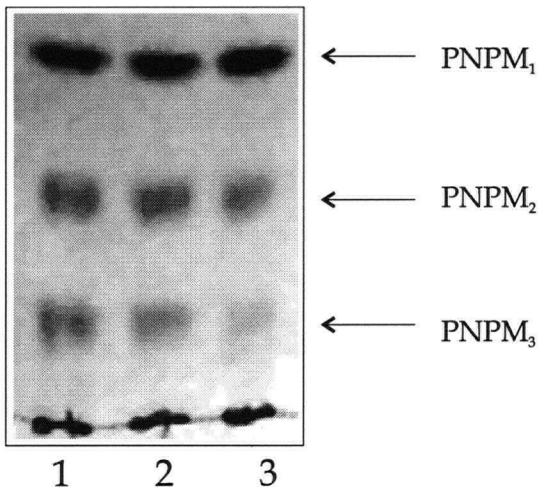


Figure 6.11: pH effect on transglycosylation catalyzed by glycosynthase Man2A E519A. Products were analyzed by thin layer chromatography (TLC) and visualized by UV. The reaction conditions were 150 mM α-mannosyl fluoride, 6 mM PNPmannose, 40 μM Man2A E519A and 150 mM phosphate buffer at Lane 1: pH 5.4, Lane 2: pH 6.0 and Lane 3: pH 7.3. The reactions were incubated for 45 h at RT.

incubated at room temperature and product formation was monitored by TLC. After an incubation period of 5 days the reaction, although not gone to completion, was stopped and the products, $PNPM_2$ and $PNPM_3$ were purified and separated by HPLC. β-1,4 and β–1,3 linkages were identified by NMR and ESMS (Section 6.2.3). β-1,4 linked PNPmannobiose was the major product formed in this transglycosylation reaction with a molar yield of 6.5 % (8.4 mg). 1.4 % (1.8 mg) of the substrate were found to be converted into β-1,3 linked PNPmannobiose. The PNPmannotriose products (0.7 %) could also be separated into two different products by HPLC. Due to the low yields however (0.8 mg and 0.3 mg), they could not be analyzed by NMR.

These results clearly showed that Man2A E519A is a glycosynthase that can be used for enzymatic synthesis of β-1,4 and β-1,3 mannosidic linkages, linkages that are very difficult to synthesize chemically. The optimization of the glycosynthase reaction was addressed in the experiments presented below.

6.2.5 Screening for good transglycosylation acceptor molecules

One way to optimize a glycosynthase catalyzed transglycosylation is to choose good acceptor molecules. Reactivation of inactivated 2FMan-Man2A WT was accelerated with gentiobiose, presumably by transglycosylation of the inactivator to gentiobiose (Section 6.2.1.4). The more efficient a glycoside reactivates the 2FMan-Man2A WT intermediate, the better its qualities to act as an acceptor molecule in transglycosylation by WT enzyme and presumably by the glycosynthase. Nine PNPglycosyl substrates were tested for their ability to reactivate 2FMan-Man2A WT. Inactivated Man2A WT was incubated either in phosphate buffer at pH 7.0 alone or in the presence of 2 mM PNPglycoside. After 90 min and 150 min

incubations at RT, aliquots were removed and the regained activity was assayed on PNPM. The fastest reactivation, from the nine substrates tested, was obtained with PNPgentiobiose. PNPcellobiose was also a good reactivator; however, reactivation was slower than with PNPgentiobiose. Significantly slower reactivation was observed with PNP-β-mannoside (PNPM) and PNP-β-galactoside. PNP-N-acteylglucosamine, PNP-α–mannoside, PNP-β-arabinoside, PNP-β-glucoside and mannose did not show increased reactivation rates compared to reactivation in buffer alone.

Four of these aryl-glycosides were tested as acceptor molecules in transglycosylation experiments with the glycosynthase 2Man E519A (Figure 6.12 A). The conditions used for these reactions were: 90 mM α-mannosyl fluoride, 7.5 mM PNPglycoside, 10 nmol Man2A E519A, 180 mM phosphate buffer at pH 5.7. The reaction volume was 55 μL. Transglycosylation yields with PNPgentiobiose as an acceptor molecule were about 10 times higher, and with PNPcellobioside about 4 times higher than with PNPM (estimated from TLC). PNPgalactose was less efficient than PNPM.

These results indicated that Man2A preferentially bound disaccharides, or perhaps even longer oligosaccharides, in its aglycon site. Therefore, PNPgentiobiose and PNPmannobiose were compared in their qualities as acceptor in glycosynthase catalyzed transglycosylation reactions (Figure 6.12 B). The glycosynthase Man2A E519A had a stronger preference for PNPgentiobiose than for mannose or derivatives thereof as acceptor.

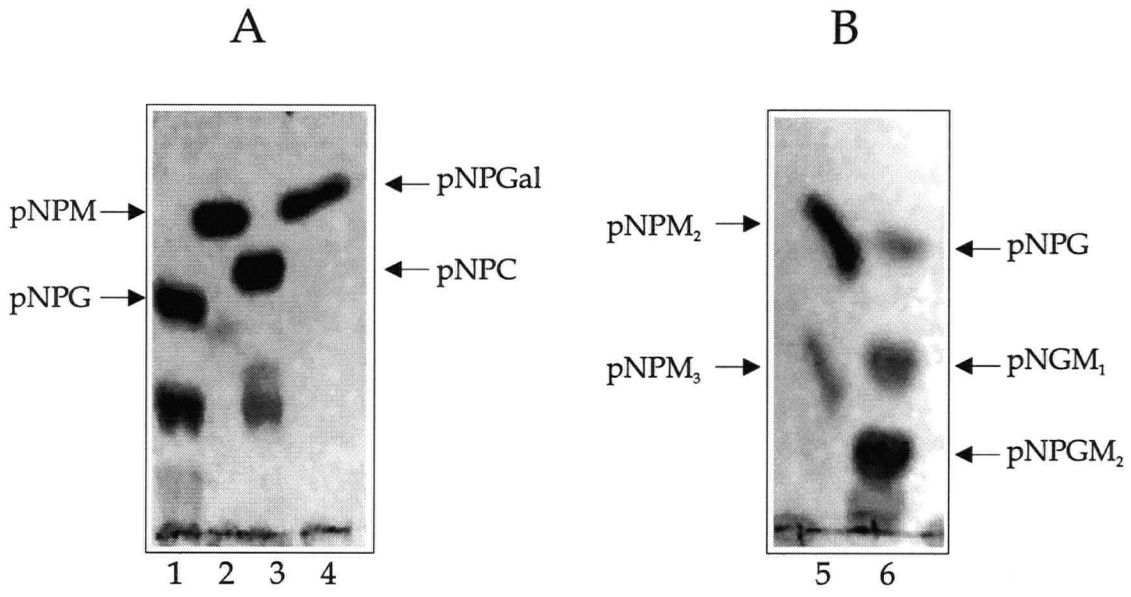


Figure 6.12: Acceptor preference of glycosynthase Man2A E519A in transglycosylation reactions. Comparison of PNP- acceptors. Products were analyzed by thin layer chromatography (TLC).
The substrates tested were Lane 1: PNPgentiobiose (PNPG); Lane 2: PNPmannose; (PNPM) Lane 3: PNPcellobiose (PNPC); Lane 4: PNPgalactose (PNPG).
Panel B: Lane 5: PNPmannobiose ($PNPM_2$) and Lane 6: PNPgentiobiose(PNPG).
Substrates and products are indicated by arrows and labeled. The reactions in Panel A were analysed after 2 days, the reactions in Panel B after 6 days incubation at RT.

6.3 Discussion

6.3.1 Man2A inactivation

The proposed mechanism for Man2A inactivation by 2-deoxy-2-fluoro-β-D-mannosyl fluoride (2FManβF) is shown in Figure 6.4. In order to trap the 2FMan-enzyme intermediate, the rate of formation (inactivation rate; k_i) must be faster than the rate of reactivation (k_{react}). Under the conditions of k_{react} approaching zero, k_i is equal to the glycosylation rate and the dissociation constant Ki, is equal to k_{-1}/k_{+1} (McCarter *et al.*, 1992). The fluorine substitution at the C-2 position in 2FManβF, reduces the glycosylation and deglycosylation rate, whereas the good leaving group at the C-1 position (fluorine) specifically increases the glycosylation rate (McCarter *et al.*, 1994). Such a situation was clearly obtained in the inactivation of Man2A with 2FManβF. Inactivation, which was time-dependent, followed pseudo-first order kinetics and showed a 1:1 stoichiometry. The detection of a glycosyl-enzyme intermediate by ESMS confirmed not only the 1:1 stoichiometry, but was also a very strong indication of the covalent nature of the glycosyl-enzyme linkage.

The catalytic competence of the 2FMan-Man2A intermediate was demonstrated by the analysis of the reactivation rates, which could be accelerated by transglycosylation to acceptor sugars (i.e., gentiobiose). The rate of reactivation of the 2FMan-Man2A intermediate was 21 fold faster upon addition of gentiobiose (k_{trans}) compared to the spontaneous reactivation rate (k_{react}). A 12.6 fold rate increase was observed for the reactivation of the *Candida albicans* 2FGlc-exo-β-1,3-glucanase intermediate with benzyl-thio-β-D-glucopyranoside (Mackenzie *et al.*, 1997). Glucosyl benzene was used to accelerate

the reactivation rate of the *Agrobacterium* sp 2FGlu-β-glucosidase intermediate 530 fold (Street *et al.*, 1992) and even higher turnover rates were found for this system with *p*-nitrophenyl β-glucoside. For all three systems the dissociation constants for the reactivator were very similar (78±25 mM, 59±3.1 mM and 56±9.5 mM, respectively). These differences in reactivation *versus* transglycosylation rates illustrate the affinities of the aglycon site for the reactivator sugars, and the importance of reactivator binding in reducing the activation energy of transglycosylation (Street *et al.*, 1992), which can be expressed as $\Delta\Delta G = RT\ln(k_{react}/k_{trans})$. Reactivation of 2FGlu-Abg with glucosyl benzene caused a larger change ($\Delta\Delta G$ = - 16.1 kJ/mol) than could be obtained from reactivation of 2FMan-Man2A with gentiobiose ($\Delta\Delta G$ = - 7.9 kJ/mol).

6.3.2 Protection of pepsin cleavage site by 2FMan glycosylation

To identify the catalytic nucleophile in Man2A, labeled and unlabeled enzyme samples were digested with pepsin. The peptide mixes were analyzed by ESMS and by comparison mapping (6.2.1.5). Peptide 1520 (MW) was unique for the labeled sample, and peptide 1036.5 (MW) was unique for the unlabeled peptide mix. Peptide 1520 differed from peptide 1036.5 by having three additional amino acids at the N-terminus ($_{517}$CSE$_{519}$), with the glutamate at position 519 identified as the catalytic nucleophile. This cleavage pattern indicated that glycosylation of E519 with 2FMan prevented peptic attack between the two amino acid residues E519 and F520, which in the unlabeled sample was cleaved to produce peptide 1036.5. Glycosylation has been shown to protect proteins against proteolysis, for example in secreted *C. fimi* cellulases (Gilkes *et al.*, 1988).

6.3.3 Prediction of catalytic residues by hydrophobic cluster analysis (HCA)

The catalytic nucleophile in bovine β-mannosidase was predicted to be E554 by comparison of HCA plots from several lysosomal glycosidases of the clan GH-A (Durand *et al.*, 1997). As seen from multiple sequence alignments, the glutamate corresponding to the bovine catalytic nucleophile is conserved in all family 2 mannosidases and corresponds to residue E519 in the *C. fimi* mannosidase, Man2A (Figure 3.19). The amino acid E519 was experimentally identified as the catalytic nucleophile (Section 6.2.1), which supports the validity of HCA predictions.

6.3.4 Transglycosylation by glycosynthase Man2A E519A

6.3.4.1 Stereo- and regiospecificity of transglycosylation

Detection of transglycosylation by Man2A WT was a good indicatior for a successful approach to transglycosylation by the glycosynthase Man2A E519A. This study is the second report of glycosynthase-catalyzed oligosaccharide synthesis using α-glycosyl fluoride as donor sugars. The *Agrobacterium* sp. β-glucosidase/galactosidase (Abg) E358A mutant, the first glycosynthase reported, uses α-glucosyl fluoride and α-galactosyl fluoride as donor sugars, whereas a wide range of aryl- and alkyl glycosides can be used as acceptor molecules. The products formed by Abg glycosynthase are almost exclusively β-1,4 linked, with the exception of glycosyltransfer to β-xylosides, which results in β-1,3 linkages. The stereo- and regiospecificity of Man2A E519A was analyzed for the $PNPM_2$ products. Most of the products were β-1,4 linked although, about 20 % of the products were found to be β-

1,3 linked. This ratio of β-1,4 to β-1,3 linked product (4.6:1) is considerably higher than the 2:1 ratio found in transglycosylation products formed by Man2A WT.

In this study only α-mannosyl fluoride was used as the donor sugar. Since Man2A can hydrolyze PNPgalactoside, α-galactosyl fluoride could possibly be used as an alternative donor sugar in Man2A catalyzed transglycosylation reactions (Section 3.13).

The first application of the glycosynthase Man2A E519A was the synthesis of a substrate ($PNPM_2$) for kinetic analysis of the *C. fimi* mannanase Man26A (Section 3.2. 9).

6.3.4.2 The effect of pH on Man2A E519A transglycosylation

The replacement of the nucleophilic carboxylate with a neutral side chain, or the formation of the glycosyl-enzyme intermediate both reduce the pKa of the acid/base catalyst in a β-glycosidase, thereby setting it up to function optimally as the general base catalyst for the deglycosylation step (McIntosh *et al.*, 1996).

The pH optimum for the Man2A E519A-catalyzed transglycosylation was pH 5.4, as estimated by product yields from TLC, whereas the optimum for PNPM hydrolysis by the WT was pH 7.0 (Section 3.11). These results were unexpected, and could not be explained by the reduction of the pK_a of the acid/base catalyst caused by the mutation of the catalytic nucleophile to alanine (McIntosh *et al.*, 1996). Other factors resulting from changes in enzyme structure that effect substrate binding might be responsible. Alternatively the lower pH might reduce substrate inhibition by reducing the formation of the inhibitory SES complex (Section 3.2.15).

6.3.4.3 Acceptor preference of Man2A E519A

The reactivation of the 2FMan-Man2A intermediate, as demonstrated with gentiobiose (Section 6.2.1.3), can be used as a fast and easy way to screen for suitable transglycosylation acceptors. In this study, nine glycosides were tested. PNPgentiobiose was the best and PNPcellobiose the second best reactivator. Four of these reactivators were also tested as acceptor molecules for glycosynthase catalyzed-transglycosylation. The glycosynthase results agreed with the reactivation results, demonstrating the usefulness of the screening procedure. The best transglycosylation product yields were obtained with PNPgentiobiose, and somewhat lower yields with PNPcellobiose. Both substrates were significantly better acceptors than PNPmannose. In the case of PNPgentiobiose, the reaction went to almost completion after an incubation period of 6 days (Figure 6.12 A). The comparison of PNPmannobiose and PNPgentiobiose as acceptors in Man2A E519A catalyzed transglycosylation reactions demonstrated that PNPgentiobiose is also a better acceptor than PNPmannobiose, which indicated that PNPgentiobiose is not better solely because it is an aryl-disaccharide (Figure 6.12 B). Gentiobiose has previously been reported to be a good aceptor in tranglycosylation reactions with *Fusarium oxysporium* β-glucosidase (Christakopoulos *et al.*, 1994). In light of Man2A being an exo-mannanase, preferential binding of mannose or manno-oligosaccharide in its aglycon site would have been expected. A 3D structure of Man2A with a substrate bound in its aglycon site (e.g. gentiobiose) could give some information about the interactions that are involved in binding of the substrate in the aglycon site. For now, the reasons for preferential

PNPgentiobiose binding are only speculative. One speculation is that the enzyme has a relatively large aglycon site, which can accommodate non-linear 1,6 linked disaccharides, such as galactose α-1,6 mannose disaccharides, as found in galactomannan, or glucose β-1,6 glucose as in gentiobiose.

Another hypothesis is based on substrate inhibition as was demonstrated for Man2A WT with PNPM (Section 3.2.15). It is assumed that Man2A E519A is also able to form a substrate inhibition complex (SES) with PNPM, but not with PNPgentiobiose and PNPcellobiose. Therefore the latter two substrates appear to be better acceptor molecules. From these transglycosylation studies it was also apparent, that α-mannosyl fluoride did not cause strong substrate inhibition in Man2A E519A.

7 Genomic Map of *Cellulomonas fimi*

7.1 Introduction

7.1.1 Genetic organization of cellulase and hemicellulase systems

In many cellulolytic microorganisms, the genes encoding cellulases and hemicellulases are scattered on the genomes with little or no linkage between individual genes (Tomme *et al.*, 1996). Most of the genes encoding components of the *Clostridium thermocellum* cellulosome are scattered on the genome and are transcribed as monocistronic mRNAs (Béguin *et al.*, 1996; Guglielmi and Béguin, 1998). A notable exception in *C. thermocellum* is the cluster of structural cellulosome component encoding genes *cipA-olpB*-ORF2-*olpA*. In this cluster *cipA* and *olpB* are in one operon and ORF2 and *olpA* in another (Fujino *et al.*, 1993a). Several other gene pairs were found to be closely spaced; e.g., *licA* and *celC* were separated by 4 kbp. In the fungus *T. reesei*, mapping of the major cellulase and xylanase encoding genes *cbh1, cbh2, egl1, egl2, bgl1, xyl1* and *xyl2* revealed only two genes in proximity to each other. The genes *cbh2* and *egl2* were located on the same 47 kbp *Not* I fragment of chromosome I (Carter *et al.*, 1992).

Most of the genes encoding enzymes and structural components of the cellulosome from the anaerobe *Clostridium cellulolyticum* are organized in an approximately 20 kbp long cluster. This is the biggest cluster reported for cellulolytic systems. It is comprised of the genes *cipC, celF, celC, celG, celE, cipX, celH* and *celJ* (Gal L., 1997; Bélaich *et al.*, 1997). Two genes encoding additional cellulases, CelA and CelD, however, are not closely linked to this cluster. A cluster, almost identical to *the C. cellulolyticum* cluster, was identified in *Clostridium josui*, suggesting that this organism might be closely related to *C. cellulolyticum*

(Fujino *et al.*, 1993b). A similar but smaller cluster was detected in the bacterium *Caldocellulosiruptor saccarolyticus*. This cluster, on a 12 kbp DNA fragment, contains at least the *celA, celB, manA* and *celC* genes, all of which encode bifunctional and multidomain enzymes (Borges *et al.*, 1993).

7.1.2 Gene cluster in *Cellulomonas fimi*

In *C. fimi,* two genes are closely linked: *cbhA*, the gene encoding cellobiohydrolase A, is upstream of *cenD*, which encodes endoglucanase CenD. The two genes are separated by only 129 bp. Putative transcriptional termination sequences and putative promoter sequences were found between the two open reading frames, indicating that they are not part of an operon (Meinke *et al.*, 1994). This close linkage suggested a possible linkage of more than only two cellulase/hemicellulase encoding genes in *C. fimi.*

7.1.3 Objectives

The aim of this project was to study the distribution of genes encoding cellulases and hemicellulases in *C. fimi,* focusing on a possible linkage between *man26A* and *man2A*. A physical and genetic map was established for the *C. fimi* genome using the techniques of pulsed field gel electrophoresis and Southern blotting.

7.2 Results

7.2.1 Mapping of the genes *man26A* and *man2A* on the *C. fimi* genome

To determine the location of the two genes involved in mannan degradation, *man26A* and *man2A*, a physical and genetic map of the *C. fimi* genome was constructed. Genomic DNA was embedded in agarose blocks and digested with restriction endonucleases (Section 2.17.1), that cleave the genome producing DNA fragments ranging from 100 kbp to 1400 kbp in size. The large DNA fragments were separated in an agarose gel by pulsed field gel electrophoresis (PFGE) (Birren *et al.*, 1993; Section 2.17.2)

Restriction endonucleases with AT-rich recognition sequences were tested for cleavage of the GC-rich *C. fimi* DNA (71.5% G+C (Yamada *et al.*, 1970)). Of the more than twenty enzymes tested, *Mun* I, *Xba* I, *Nde* I, *Hpa* I, *Eco*R I, *Hin*d III and *Nsi*I proved to be good candidates. These enzymes produced 7 to 15 DNA fragments, 20 kbp to 1400 kbp in size that were separable on a 1.2 % agarose gel by PFGE. To yield optimal separation of all fragments within the range of 20 kbp to 1400 kbp, the gels were run for 36 h switching the orientation of the electric field every 60 s for the first 9 h, then every 80 s, 100 s and 120 s for 9 h each. The voltage was 160 V and the temperature was controled at 14^0 C. The DNA bands were visualized after separation with ethidium bromide under UV light (Figure 7.1 A). As molecular size standards, chromosomes from the yeast *Saccharomyces cerevisiae* and *Hin*d III digested λ DNA were used. The sizes of DNA fragments were estimated by comparing their mobilities to the mobilities of the size standards. Fragments released by *Hpa* I, *Mun* I, *Hin*d III and *Nsi* I are shown schematically in Figures 7.2 and 7.3. Addition of fragment lengths from *Hpa* I, *Mun* I and *Hin*d III digests resulted in genome sizes of 4120 kbp,

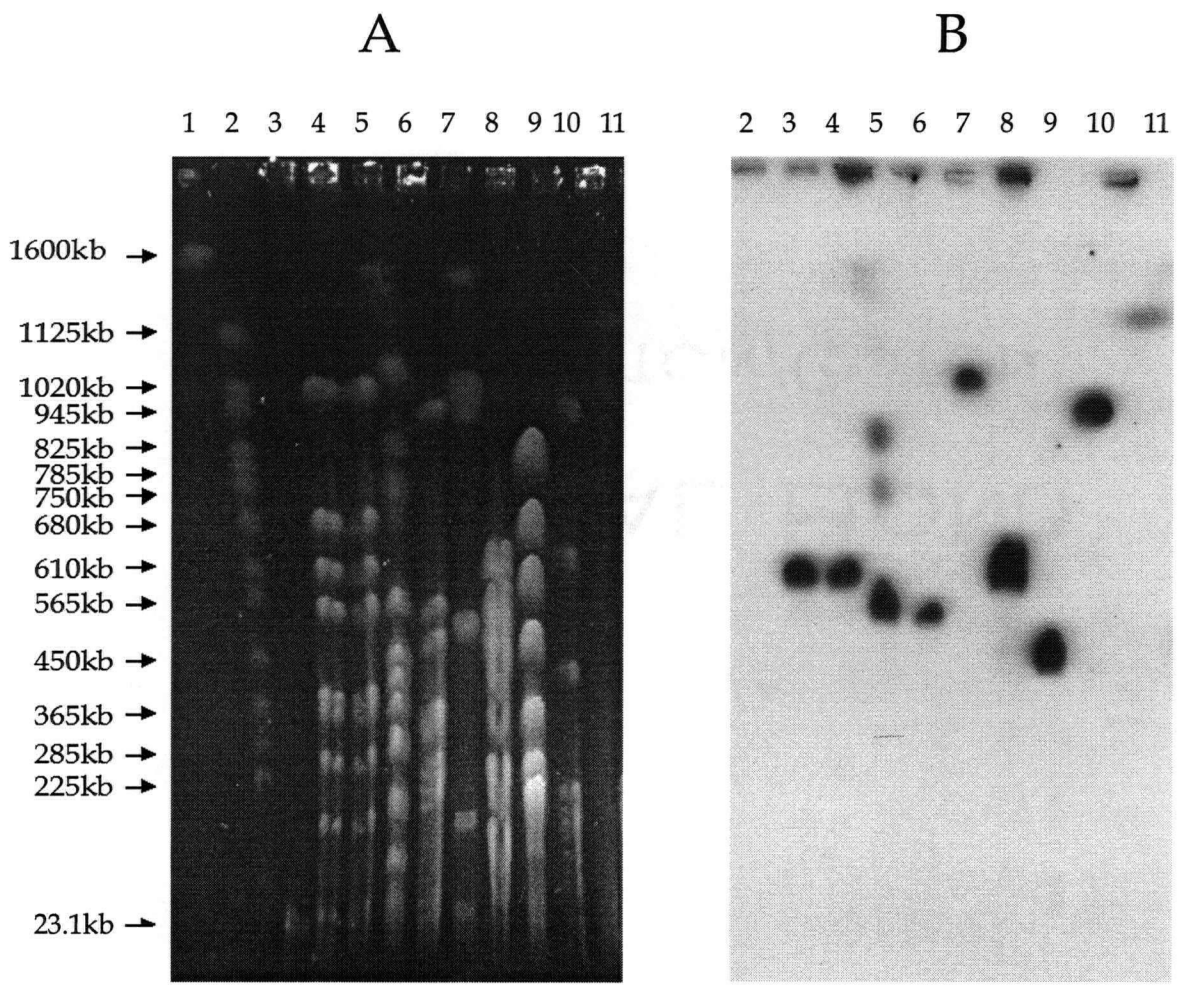


Figure 7.1: Panel A: Separation of restriction fragments of *C. fimi* genomic DNA by pulsed field gel electrophoresis (PFGE). The following parameters were used for PFGE: switch intervals of 60 s, 80 s, 100 s and 120 s for 9 h each at 14°C, 1.2 % agarose and 1 x TAE running buffer.
The sizes of the size standards are indicated: *Saccharomyces cerevisiae* cromosomes (Lane 1) and λ-*Hin*d III (Lane 2).
Panel B: Southern blot of gel shown in Panel A probed with fluorescently labeled, 711 bp *Pvu* I *cenD* fragment.
The *C. fimi* genomic DNA was digested with the following restriction endonucleases: *Mun* I (Lane 3 and Lane 4), *Xba* I (Lane 5), *Nde* I (Lane 6), *Hpa* I (Lane 7), *Eco*R I (Lane 8), *Hin*d III (Lane 9) and *Nsi* I (Lane 10). Undigested genomic *C. fimi* DNA is shown in Lane 11.

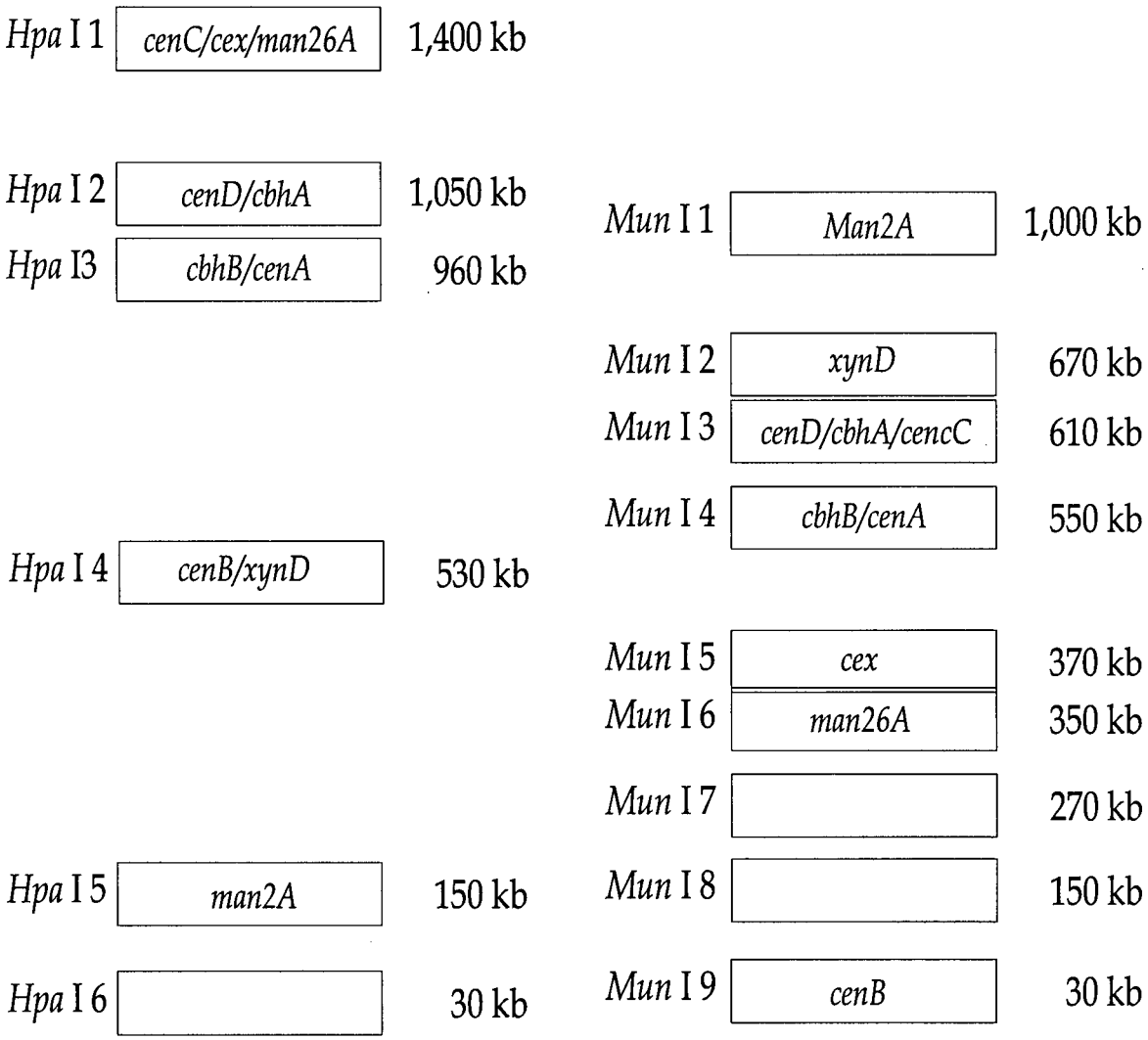


Figure 7.2: Schematic representation of the DNA fragments produced by *Hpa* I and *Mun* I restriction endonuclease digestion of *C. fimi* genomic DNA. The fragments are represented as boxes with their names and sizes indicated. The cellulase and hemicellulase encoding genes were mapped to these fragments. The genome size calculated from the sum of *Hpa* I fragments is 4,120 kbp and 4,000 kbp for the *Mun* I fragments, respectively.

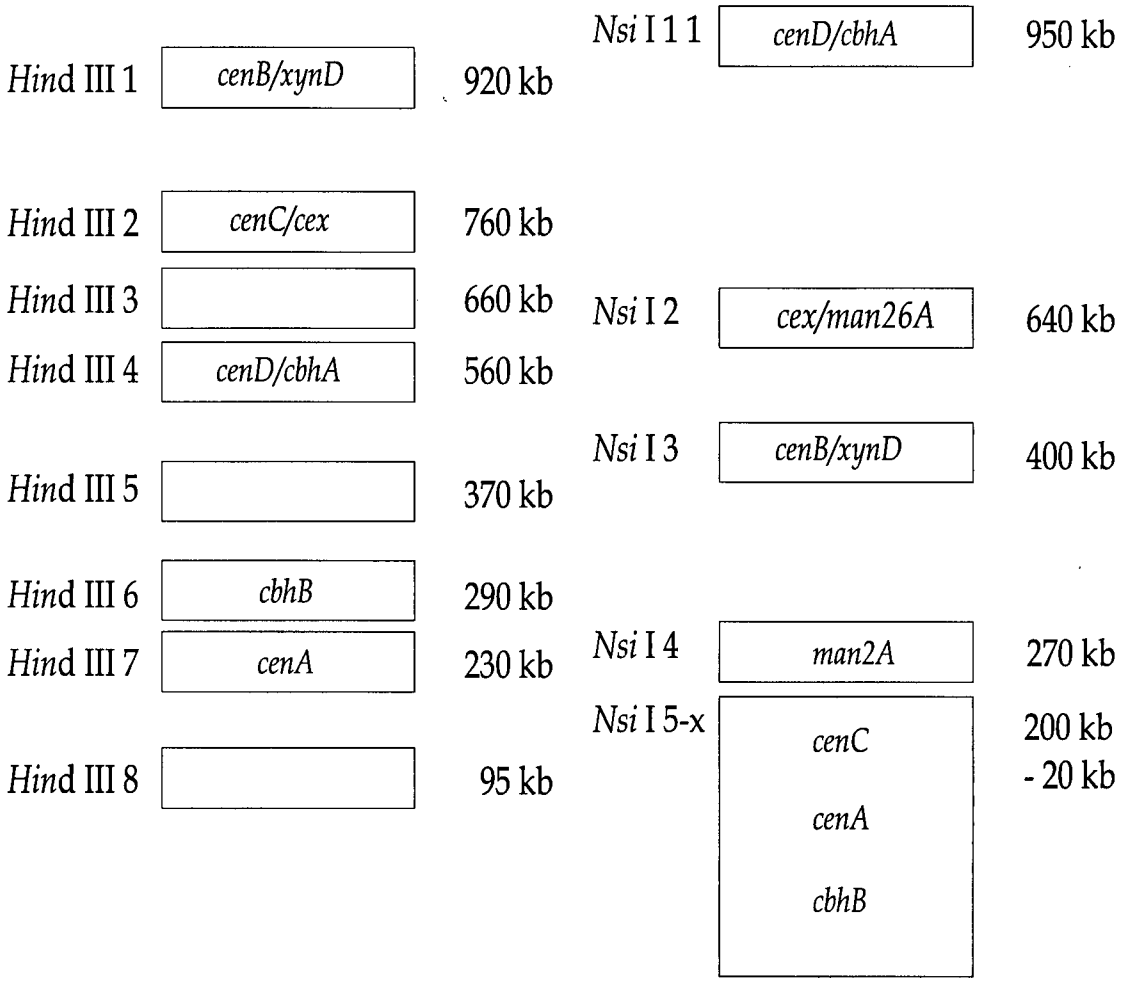


Figure 7.3: Schematic representation of DNA fragments produced by *Hind* III and *Nsi* I restriction endonuclease digestion of *C. fimi* genomic DNA. The fragments are represented as boxes, with their names and sizes indicated. The cellulase and hemicellulase encoding genes were mapped. From the sum of *Hind* III fragments a genome size of 3,885 kbp was calculated .
The DNA fragments produced by *Nsi* I digestion, 200-20 kbp in size, could not be separated under the conditions used, and are therefore represented as one large box. The relative positions of the genes within this region are indicated.

4000 kbp and 3885 kbp, respectively. The average *C. fimi* genome size was therefore estimated to be approximately 4,000 kbp. *C. fimi* DNA treated with *Xba* I resulted only in partial digestion. Therefore, the results obtained from *Xba* I digests were not considered for further analysis.

Undigested *C. fimi* DNA comigrated as a single band with the 1,125 kbp DNA standard on PFG. This suggested that *C. fimi* has a single, circular chromosome, with the electrophoretic mobility of supercoiled DNA.

The PFGE gel shown in Figure 7.1 A was subjected to Southern blot analysis. Fluorescently labeled probes, i.e., the 711bp *Pvu* I *cenD* fragment, were used to locate the cellulase and hemicellulase genes in each set of restriction fragments (Figure 7.1 B) (Section 2.17.3).

Good PFGE and Southern blot results were obtained with the *C. fimi* genes *man26A* , *man2A*, *cenA*, *cenB*, *cenC*, *cenD*, *cbhB*, *cex* and *xynD* as genetic markers from *Mun* I, *Hpa* I and *Hin*d III digests. These results are summarized in the schematic restriction patterns shown in Figure 7.2 and 7.3 and were used to create a genetic map of the *C. fimi* chromosome (Figure 7.4). Results obtained by PFGE and Southern blot analyses of the large DNA fragments from *Nsi* I and *Nde* I digests were included in the map as well. Several DNA fragments produced by *Nsi* I that ranged from 200 kb to 20 kb could not be separated clearly by PFGE. In the schematic representation in Figure 7.3, all of these bands are represented by only one large box.

The map of the *C. fimi* chromosome, shown in Figure 7.4, summarizes all the results obtained by PFGE and Southern blot. Each circle on the map represents the full length of the chromosome and on each circle the results from one restriction endonucleolytic digestion are summarized. Each box represents one restriction fragment and is drawn in its

relative size, and gaps between boxes indicate boundaries. For *Nsi* I and *Nde* I digests only the fragments that contained one of the used a genetic markers could be mapped. The physical linkages of the three largest *Mun* I fragments were analyzed. After separation of *Mun* I digested *C. fimi* DNA by PFGE, the three DNA fragments *Mun* I1, *Mun* I2 and *Mun* I3 (Figure 7.2) were each excised from the gel and purified (QiaexII). After fluorescent labeling they were used as probes for Southern blots. The *Mun* I 1 fragment hybridized to the *Hpa* I 2, *Hpa* I3 and *Hpa* I 5 fragments and to the *Nde* I 1 fragment. Hybridization to fragments from other digests resulted in ambigous signals. The *Mun* I 2 fragment could only be shown to hybridize to the *Hin*d III1 fragment and the *Mun* I 3 fragment was shown to hybridize to *Hin*d III 2, *Hin*d III 4, *Hpa* I 1, *Hpa* I 2 and to *Nde* I 2 fragments. These results were in good agreement with the genomic map as outlined in Figure 7.4.

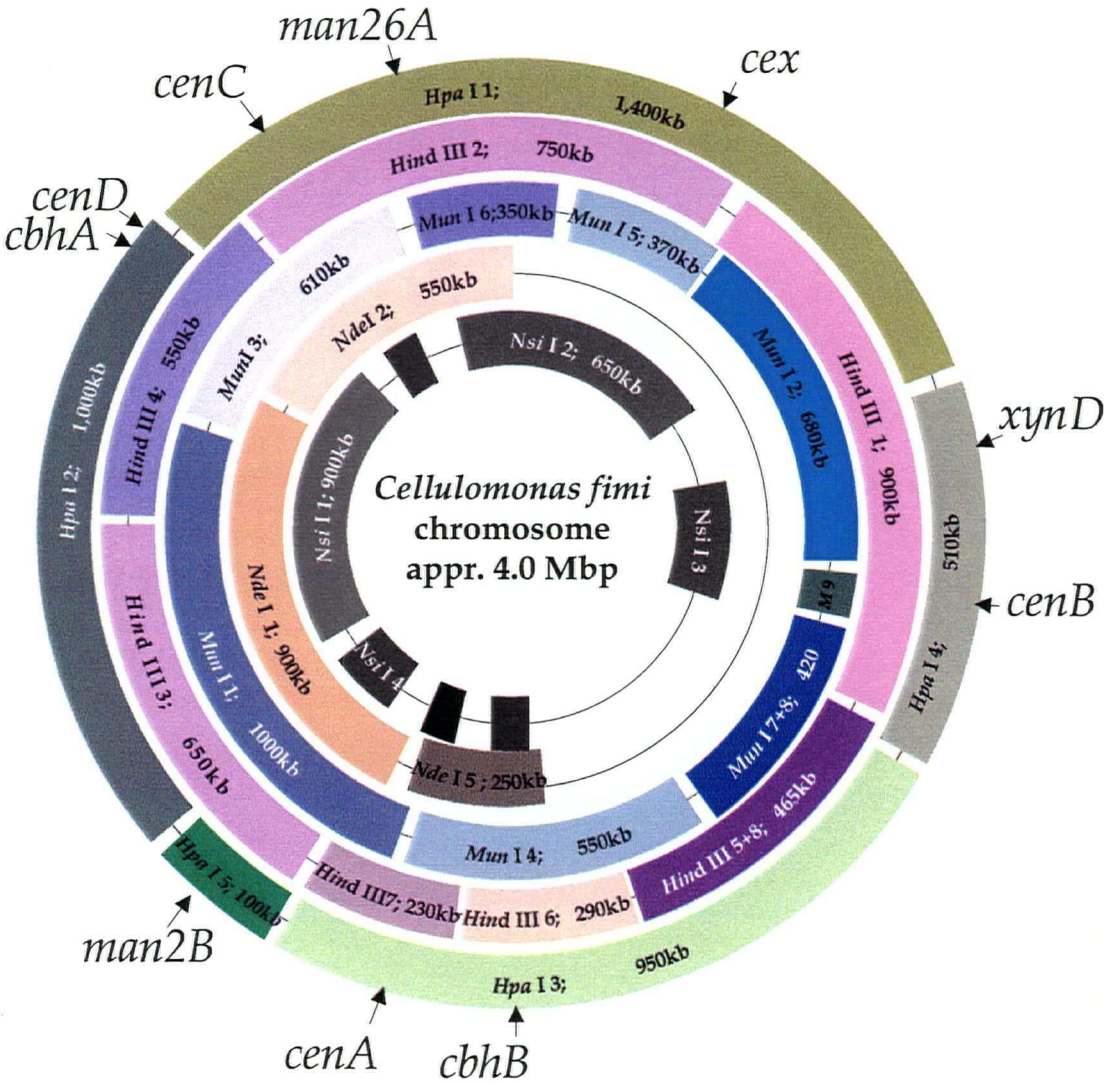


Figure 7.4: Physical and genetic map of the *Cellulomonas fimi* genome. Each circle represents DNA fragments produced by one restriction endonuclease.
One DNA fragment is represented by one box, with the length of the box being proportional to the length of the DNA fragment. The majority of the boxes are labeled with the name and the length of the fragments. Only the fragments that could be physically or genetically mapped are included. The relative positions of the *C. fimi* cellulase and hemicellulase encoding genes are indicated by arrows.

7.3 Discussion

7.3.1 Genome size and geometry

The *C. fimi* genome was determined to be approximately 4,000 kbp in size, comparable to the *E. coli* genome, which is 4,639 kbp (Blattner *et al.*, 1997). Even though sizes of bacterial genomes vary from 600 kbp for *Mycoplasma genitalium* to 12,800 kbp for *Calothrix* strains, the *C. fimi* genome is similar in size to the average Gram positive genome. The average size of Gram positive genomes was calculated to be 3,115 kbp (Trevor, 1996).

Undigested genomic *C. fimi* DNA comigrated with the 1,125 kbp size standard. This suggested that *C. fimi* has only one circular chromosome. The majority of the bacterial genomes studied were found to be circular. However, linear chromosomes have been described, e.g. for *Borelia burgdorferi*, which has a linear chromosome of approximately 1,000 kbp and a number of linear and circular plasmids (Davidson *et al.*, 1992), and for *Streptomyces lividans* and *Streptomyces griseus* both of which have one linear chromosome about 8,000 kbp in size (Lezhava *et al.*, 1995). Plasmids, which might have been useful tools for genetic manipulations, were not detected in *C. fimi* DNA preparations.

7.3.2 Genetic and physical mapping of the *C. fimi* genome

PFGE and Southern blots were used to establish a genetic and physical map of the *C. fimi* genome. The genome map clearly shows that neither the genes encoding the mannan degrading system, *man26A* and *man2A*, nor the genes encoding the cellulose or xylan degrading systems are organized in clusters. The only exception is the close linkage of *cbhA* and *cenD* (Meinke *et al.*, 1994). As in *C. fimi* , most of the genes known to encode cellulases

and xylanases are scattered on the genome of *P. fluorescens* subsp. *cellulosa*. However, the *xynB* and *xynC* genes, encoding a xylanase and an arabinofuranosidase, respectively, are in the same orientation and only 148 base pairs apart (Kellet *et al.*, 1990). These two examples of small clusters might have arisen from gene duplications.

Even though the *C. fimi* genes encoding hemicellulases and cellulases are scattered all over the chromosome some of them might still be part of one or more regulons, and controlled by the same regulatory element(s).

8 References

Ali, B.R.S., Zhou, L., Graves, F.M., Freedman, R.B., and Black, G.W. (1995). Cellulases and hemicellulases of the anaerobic fungus *Pyromyces* constitute a multiprotein cellulose-binding complex and are encoded by multigene families. FEMS Microbiol. Lett. 125:15-22.

Alkhayat AH, Kraemer SA, Leipprandt JR, Macek M, Kleijer WJ, and Friderici KH (1998). Human beta-mannosidase cDNA characterization and first identification of a mutation associated with human beta-mannosidosis. Hum. Mol. Genet. 7:75-83.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25:3389-3402.

Arcand, N., Kluepfl, D., Paradis, F.W., Morosoli, R., and Sharek, F. (1993). β-mannanase of Streptomyces lividans 66: cloning and DNA sequenceing of the *manA* gene and charaterization of the enzyme. Biochem. J. 290:857-863.

Atalla, R.H., Hackney, J.M., Uhlin, I., and Thompson, N.S. (1993) Hemicelluloses as structure regulators in the aggregation of native cellulose. Int J Biol Macromol 15(2):109-112

Atkinson, T. and Smith, M. (1984). Solid phase synthesis of oligodeoxyribonucleotides by the phosphite triester method. In: Oligonucleotide synthesis: a practical approach. N.J. Gait (ed.) IRL Press, Oxford. pp. 35-81.

Bauer, M.W., Bylina, E.J., Swanson, R.V., and Kelly, R.M. (1996). Comparison of a β–glucosidase and a β-mannosidase from the hyperthermophilic archaeon *Pyrococcus furiosus*. J. Biol. Chem. 271(39):23749-23755.

Bayer, E.A., Morag, E., and Lamed, R. (1994). The cellulosome - A treasure trove for biotechnology. Trends Biotechnol. 12:378-386.

Bayer, E.A., Morag, E., Shoham, Y., Tormo, J., and Lamed, R. (1996). In: Bacterial Adhesion: Molecular and Ecological Diversity. pp 155-182. Wiley-Liss, Inc.

Béguin, P., and Aubert, J.-P. (1994). The bioloical degradation of cellulose. FEMS Microbiol. Rev. 13:25-58.

Béguin, P., and Lemaire, M. (1996). The cellulosome: An exocellular, multiprotein complex specialized in cellulose degradation. Crit. Rev. Biochem. Mol. Biol. 31(3):201-236.

Bélaich, J.-P., Tardif, C., Bélaich, A., and Gaudin, C. (1997). The cellulolytic system of *Clostridium cellulolyticum*. J. Biotech. 57:3-14.

Beveridge, T.J., Pouwels, P.H., Sara, M., Kotiranta, A., Lounatma, K., Kari, K., Keruso, E., Haapasolo, M., Egelseer, E.M., Schocher, I., Sleytr, U.B., Morelli, L., Callegari, M.-L., Nomelli, J.F., Bingle, W.H., Smit, J., Leibovitz, E., Lemaire, M., Miras, I., Salamitou, S., Beguin, P., Ohayon, H., Gounon, P., Matuschek, M., Sahm, K., Bahl, H., Grogono-Thomas,

R., Dworkin, J., Blaser, M.J., Woodland, R.M., Newell, D.G., Kessel, M., and Koval, S.F. (1997). Function of S-layers. FEMS Microbiol. Rev. 20:99-149.

Birren, B., and Lai, E. (1993). Pulsed field gel electrophoresis. A practical guide.Academic Press, Inc., San Diego.

Black, G.W., Hazlewood, G.P., Millward-Saddler, S.J., Laurie, J.I., and Gilbert, H.J. (1995). A modular xylanase containing a novel non-catalytic xylan-specific binding domain. Biochem. J. 307:191-195.

Blattner, F.R., Plunkett, G., Bloch C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden M.A., Rose, D.J., Mau, B., and Shao, Y . (1997) The complete genome sequence of **Escherichia coli** K-12. Science. 277(5331):1453-74.

Bolam, D.N., Hughes, N., Virden, R., Lakey, J.H., Hazlewood, G.P., Henrissat, B., Braithwaite, K.L., and Gilbert, H.J. (1996). Mannanase A from *Pseudomonas fluorescens* ssp. *cellulosa* is a retaining glycosyl hydrolase in which E212 and E320 are the putative catalytic residues. Biochemistry 35:16195-16204.

Borges, K.M., and Berquist, P.L. (1993). Pulsed-Field gel electrophoresis study of the genome of *Caldocellum saccharolyticum*. Curr. Microbiol. 27:15-19.

Braithwaite, K.L., Black, G.W., Hazlewood, G.P., Ali, B.R., and Gilbert, H.J. (1995). A non-modular endo-beta-1,4-mannanase from *Pseudomonas fluorescens* subspecies *cellulosa*. Biochem.J. 305:1005-1010.

Bray, M.R., Johnson, P.E., Gilkes, N.R., McIntosh, L.P., Kilburn, D.G., and Warren, R.A.J (1996). Probing the role of tryptophan residues in a cellulose-binding domain by chemical modification. Prot. Sci. 5:2311-2318.

Brun, E., Moriaud, F., Gnas, P., Blackledge, M.J., Barras, F., and Marion, D. (1997). Solution structure of the cellulose-binding domain of the endoglucanase Z secreted by *Erwinia chrysanthemi*. Biochemistry 36:16074-16086.

Carter, G.L., Allison, D., Rey, M.W., and Dunn-Coleman, N.S. (1992). Chromosomal and genetic analysis of the electrophoretic karyotype of *Trichoderma reesei*: mapping of the cellulase and xylanase genes. Mol. Microbiol. 6(15):2167-2174.

Chaplin, M.F. (1986). Chaplin, M.F., Kennedy, J.F. (eds), In: Carbohydrate analysis: a practical approach. IRL Press, Oxford, pp 1-36.

Chen, H., Leipprandt, J.R., Traviss, C.E., Sopher, B.L., Jones, M.Z., Cavanagh, K.T., and Friderici, K.H. (1995). Molecular cloning and characterization of bovine beta-mannosidase. J.Biol.Chem. 270:3841-3848.

Clarke, J.H., Davidson, K., Gilbert, H.J., Fontes, C.M.G.A., and Hazlewood, G.P. (1996). A modular xylanase from mesophilic *Cellulomomas fimi* contains the same cellulose-bining and

thermostabilizing domains as xylanases from thermophilic bacteria. FEMS Micobiol. Lett. 139:27-35.

Cornish-Bowden, A. (1979). Fundamentals of enzyme kinetics. Butterworth & Co Ltd.

Coulombel, C., Clermont, S., Foglietti, M.-J., Percheron, F. (1981). Transglycosylation reactions catalyzed by two β-mannanases. Biochem. J. 195:333-335.

Coutino, J.B., Gilkes, N.R., Warren, R.A.J., Kilburn, D.G., and Miller, R.C. (1992). The binding of *Cellulomonas fimi* endoglucanase C (CenC) to cellulose and Sephadex is mediated by the N-terminal repeats. Mol. Microbiol. 6(9):1243-1252.

Creagh, A.L., Koska, J., Johnson, P.E., Tomme, P., Joshi, M.D., McIntosh, L.P., Kilburn, D.G., Haynes, C.A. (1998). Stability and oligosaccharide binding of the N1 cellulose-binding domain of *Cellulomonas fimi* endoglucanase CenC. Biochemistry. 37(10): 3529-3537

Creagh, A.L., Ong, E., Jervis, E., Kilburn, D.G., and Haynes, C.A. (1996). Binding of the cellulose-binding domain of exoglucanase Cex from Cellulomonas fimi to insoluble microcrystalline cellulose is entropically driven. Proc. Natl. Acad. Sci. USA 93:12229-12234.

Damude, H.G., Withers, S.G., Kilburn, D.G., Miller, R.C. and Warren, R.A.J. (1995). Site-directed mutation of the putative catalytic residues of endoglucanase CenA from *Cellulomonas fimi*. Biochemistry. 34(7): 2220-2224.

Davies, G., and Henrissat, B. (1995). Structure and mechanisms of glycosyl hydrolases. Structure 3(9): 853-859.

Davies, G., Sinnott, M.L. and Withers, S.G. (1998) Glycosyl transfer. Comprehensive biological catalysis. A Mechanistic Reference. Volume I. Academic Press. San Diego.

Del Sal, G, Manfioletti, G, and Schneider, C. (1989). The CTAB-DNA precipitation method: a common mini-scale preparation of template DNA from phagemids, phages or plasmids suitable for sequencing. Biotechniques. 7(5): 514-520

Dower, W.J. (1987). Mol. Biol. Rep. (Bio-Rad Laboratories)1, 5.

Durand, P., Lehn, P., Callebaut, I., Fabrega, S., Henrissat, B, and Mornon, J.-P. (1997). Active-site motifs of lysosomal acid hydrolases: invariant features of clan GH-A glycosyl hydrolases deduced from hydrophobic cluster analysis. Glycobiol. 7(2):277-284.

Edelman, G.M. (1983). Cell adhesion molecules. Science 219:450.

Edge, C.J., Rademacher, T.W., Wormald, M.R., Parekh, R.B., Butters, T.D., Wing, D.R. and Dwerk R.A. (1992). Fast sequencing of oligosaccharides: the reagent-array analysis method. Proc. Natl. Acad. Sci. USA 89:6338-6342.

Egelseer, E.M., Leitner, K., Jarosch, M., Hotzy, C., Zayni, S., Sleytr, U.B., and Sara, M. (1998). The S-layer proteins of two *Bacillus stearothermophilus* wild-type strains are bound via their

N-terminal region to a secondary cell wall polymer of identical chemical composition. J. Bacteriol. 180(6):1488-1495.

Estrada, P., Mata, I., Dominguez, J.M., Castillon, M.P., and Acebal, C. (1990). Kinetic mechanism of β-glucosidase from *Trichoderma reesei* QM 9414. Biochim. Biophys. Acta 1033:298-304.

Fan, J.-Q., Takegawa, K., Iwahara, S., Kondo, A., Kato, I., Abeygunawardana, C., and Lee, Y.C. (1995). Enhanced transglycosylation activity of *Arthrobacter protophormiae* endo-β-N-acetylglucosaminidase in media containing organic solvents. J. Biol. Chem. 270:17723-17729.

Feinberg, A.P., and Vogelstein, B. (1983). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal Biochem. 132: 6-13.

Franco. T.T., Rodrigues, N.R., Serra, G.E., Panegassi, V.R., and Buckeridge, M.S. (1996). Characterization of storage cell wall polysaccharides from Brazilian legume seeds and the formation or aqueous two-phase systems. J. Chromatogr. B. Biomed. Appl. 680(1-2):255-261.

Fujino, T., Béguin, P., and Aubert, J.-P. (1993a). Organisation of a *Clostridium thermocellum* gene cluster encoding the cellulosomal scaffolding protein CipA and a protein possibly involved in the attachement of the cellulosome to the cell surface. J. Bacteriol. 175:1891-1899.

Fujino, T., Karita, S., and Ohmiya, K. (1993b). Nucleotide sequence of the celB gene encoding endo-1,4,-beta-glucanase-2, ORF1 and ORF2 forming a putative cellulase gene cluster of *Clostridium josui*. J. Ferment. Bioeng. 76:243-250.

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

Gal, L. (1997). Etude du cellulosome de *Clostridium cellulolyticum* et de l'un de ses composants: la cellulase CelG. PhD thesis. Universite de Provence Aix-Marseille I.

Gebler, J.C., Aebersold, R., and Withers, S.G. (1992). Glu-537, not Glu-461, is the nucleophile in the active site of *(lacZ)* beta-galactosidase from *Escherichia coli*. J. Biol. Chem. 267:11126-11130.

Gehrmann, M.C., Opper, M., Sedlacek, H.H., Bosslet, K., and Czech, J. (1994). Biochemical properties of recombinant human β-glucuronidase synthesized in baby hamster kidney cells. Biochem. J. 301:821-828.

Gherardini, F., Babcock, M., and Salyers, A. (1985). Purification and characterization of two α-galactosidases associated with the catabolism of guar gum and other α-galactosides by *Bacteroides ovatus*. J. Bacteriol. 161:500-506.

Gilbert, H.J. and Hazelwood, G.P. (1993). Bacterial cellulases and xylanases. J. Gen. Microbiol. 139:187-194.

Gilkes, N.R., Chanzy, H., Kilburn, D.G., Miller, R.C. Jr., Sugiyama, J., Warren, R.A.J., and Henrissat, B. (1993). Visualization of the adsorption of a bacterial endo-beta-1,4-glucanase and its isolated cellulose-binding domain to crystalline cellulose. Internat. J. Biol. Macromol. 15:347

Gilkes, N.R., Henrissat, B., Kilburn, D.G., Miller, R.C., Warren, R.A.J. (1991). Domains in microbial β-1,4-glycanases: Sequence conservation, function, and enzyme families. Microbiol. Rev. 55:303-315.

Gilkes, N.R., Jervis, E., Henrissat, B., Tekant, B., Miller, R.C. Jr., Warrren, R.A.J., and Kilburn, D.G. (1992). The Adsorption of a bacterial cellulase and its two isolated domains to crystaliline cellulose. J. Biol. Chem. 267:6743-6749.

Gilkes, N.R., Kilburn, D.G., Miller, R.C. Jr., and Warren, R.A.J. (1989) Structural and functional analysis of a bacterial cellulase by proteolysis. J. Biol. Chem. 264:17802-17808.

Gilkes, N.R., Kwan, E., Kilburn, D.G., Miller, R.C., and Warren, R.A.J. (1997). Attack of carboxymethyl cellulose at opposite ends by two cellobiohydrolases from *Cellulomonas fimi*. J. Biotechnol. 57:83-90.

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

Grodber, J., and Dunn, J.J. (1988). *ompT* encodes the *Escherichia coli* outer membrane protease that cleaves T7 RNA polymerase during purification. J. Bacteriol. 170:1245-1253

Grueninger-Leitch, F., D'Arcy, A., D'Arcy, B,. and Chene, C. (1996). Deglycosylation of proteins for crystallization using recombinant fusion protein glycosidases. Protein Sci. 5:2617-2622.

Guglielmi, G., and Béguin, P. (1998). Cellulase and hemicellulase genes of *Clostridium thermocellum* from five independent collections contain few overlaps and are widely scattered across the chromosome. FEMS Microbiol. Lett. 161:209-215.

Gusakov, A.V., Protas, O.V., Chernoglazov, V.M., Sinitsyn, A.P., KovalyshevaG.V., Shpanchenko, O.V., and Ermolova, O.V. (1991). Transglycosylation activity of cellobiohydrolase I from *Trichoderma longibrachiatum* on synthetic and natural substrates. Biochim. Biophys. Acta 1073:481-485.

Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. 166:557-580.

Harjunpää, V., Teleman, A., Siika-Aho, M., and Drakenberg, T. (1995). Kinetic and stereochemical studies of mannooligosaccharide hydrolysis catalysed by β-mannanases from *Trichoderma reesei*. Eur. J. Biochem. 234:278-283.

Hauzer, K., Ticha, M., Horejsi, V., and Kocourec, J. (1979) Studies on lectins. XLIV. The pH dependence of lectin interactions with sugars as determined by affinity electrophoresis. Biochim. Biophys. Acta 583:103-109.

Hendrixon, T.L., and Millane, R.P. (1993). In: Cellulosics:chemical, biochemical and material aspects. Kennedy, J.F., Phillips, G.O., Williams, P.A. (Eds.), Ellis Horwood Limited.

Henrissat , B. (1990). A classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem. J. 280:304-316.

Henrissat, B. and Bairoch, A. (1993). New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem. J. 293:781-788.

Henrissat, B., and Bairoch, A. (1996) Updating the sequence based classification of glycosyl hydrolases. Biochem. J. 316:695-696.

Henrissat, B., Callebaut, I., Fabrega, S., Lehn, P., Mornon, J.-P., Davies, G. (1995). Conserved catalytic machinery and the prediction of a common fold for several families of glycosyl hydrolases. Proc. Natl. Acad. Sci. USA 92:7090-7094.

Henrissat, B., Teeri, T.T., and Warren, R.A.J. (1998). A scheme for designating enzymes that hydrolyse the polysaccharides in the cell walls of plants. FEBS Lett. 425:352-354.

Higgins D., Thompson, J., Gibson, T.Thompson, J.D., Higgins, D.G., Gibson T.J. (1994). "CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting,position-specific gap penalties and weight matrix choice." Nucleic Acids Res. 22:4673-4680.

Holmskov, U., Fischer, P.B., Rothman, A., and Horjup, P. (1996). Affinity and kinetic analysis of bovine plasma C-type lectin collectin-43 (CL-43) interacting with mannan. FEBS Lett. 393:314-316.

Hui, P.A., Neukom, H. (1964). Properties of galactomannans. Tappi. 47:39-42.

Ichikawa, Y., Look, G.C., wong, C.-H. (1992). Enzyme-catalyzed oligosaccharide synthesis. Anal. Biochem. 202:215-238.

Irwing, D., Shin, D.H., Barr, B.K., Sakon, J., Karplus, P.A., and Wilson, D.B. (1998). Roles of the catalytic domain and the two cellulose-binding domains of *Thermomonospora fusca* E4 in cellulose hydrolysis. J. Bacteriol. 180:1709-1714.

Jacobson, R.H., Zhang, X.-J., DuBose, R.F., and Matthews, B.W. (1994). Three-dimensional structure of beta-galactosidase from *E. coli*. Nature 369:761-766.

Jain, S., Drendel, W.B., Chen, Z.-W., Mathews, S., Sly, W.S., and Grubb, J.H. (1996). Structure of human beta-glucuronidase reveals canditate lysosomal targeting and active-site motifs. Nature Struc. Biol. 3:375-380.

Jerseth, J. (1992). Strategies 5(1).

Johnson, E. P., Tomme, P., Joshi, M. D., and McIntosh, L. P. (1996a). Interaction of soluble cellooligosaccharides with the N-terminal cellulose-binding domain of *Cellulomonas fimi* CenC. 2. NMR and ultraviolet absorption spectroscopy. Biochem. 35:13895-13906.

Johnson, P.E., Joshi, M.D., Tomme, P., Kilburn, D.G., and McIntosh, L.P. (1996b). Structure of the N-terminal cellulose-binding domain of *Cellulomonas fimi* CenC determined by nuclear magnetic resonance spectroscopy. Biochem. 35:14381-14394.

Kallweit, U., Boernsen, K.O., Kresbach, G.M., and Widmer, H.M. (1996). Matrix compatible buffers for analysis of proteins with matrix-assisted laser desorption/ionization mass spectrometry. Rapid Com. Mass Spec. 10:845-849.

Katz, G. (1965). The location and the significance of the O-acetyl groups in a glucomannan from Parana pine. TAPPI, 48, 34-41.

Kellet, L.E., Poole, D.M., Ferreira, L.M.A., Durrant, A.J., Hazlewood, G.P., Gilbert, H.J. (1990). Xylanase B and an arabinofuranosidase from *Pseudomonas fluorescens* subsp. *cellulosa* contain identical cellulose-binding domains and are encoded by adjacent genes. Biochem J. 272(2): 369-376.

Kempton, J.B., and Withers, S.G. (1992). Mechansm of *Agrobacterium* β-glucosidase: Kinetic studies. Biochemistry 31:9961-9969.

Kim, H.K., and Pack, M.Y. (1989). Cloning and expression of *Cellulomonas fimi* β-glucosidase genes in *Escherichia coli*. Enzyme Microb. Technol. 11:313-316.

Kleman-Leyer, K.M., Glikes, N.R., Miller, R.C. Jr., and Kirk, K.T (1994). Changes in the molecular-size distribution of insoluble celulloses by the action of recombinant *Cellulomonas fimi* cellulases. Biochem. J. 32:463-469.

Kohchi C, and Toh-e A. Cloning of *Candida pelliculosa* beta-glucosidase gene and its expression in *Saccharomyces cerevisiae*. Mol Gen Genet. 203: 89-94

Kormos, J. M. (1998). Master thesis. Functional and mutational analysis of the cellulose-binding domain CBDN1 from *Cellulomonas fimi* β-glucanase C (CenC). University of British Columbia.

Koshland, D.E. (1953). Stereochemistry and the mechanism of enzymatic reactions. Biol. Rev. 28:416-436.

Krawiec, S., and Riley, M. (1990). Organisation of the bacterial chromosome. Microbiol. Rev. 54:502-539.

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

Leatherbarrow, R.J. (1992). GraFit version 3.0. Erithacus Software Ltd., Staines, United Kingdom.

Legler, G. (1990). Glycoside hydrolases: Mechanistic information from studies with reversible and irreversible inhibitors. Adv. Carbohydr. Chem. Biochem. 48:319-385.

Leipprandt, J.R., Kraemer, S.A., Haithcock, B.E., Chen, H., Dyme, J.L., Cavanagh, K.T., Friderici, K.H., and Jones, M.Z. (1996) Caprine beta-mannosidase: sequencing and characterization of the cDNA and identification of the molecular defect of caprine beta-mannosidosis. Genomics 37:51-56.

Lemaire, M., Miras, I., Gounon, P., and Béguin, P. (1998). Identification of a region responsible for binding to the cell wall within the S-layer protein of *Clostridium thermocellum*. Microbiol. 144:211-217.

Lemaire, M., Ohayon, H., Gounon, P., Fujino, T., and Béguin, P. (1993). OlpB, a new outer layer protein form *Clostridium thermocellum*, and binding of its S-layer-like domains to components of the cell envelope. J. Bacteriol. 177:2451-2459.

Letzhava, A., Mizukami, T., Kajitani, T., Kameoka, D., Redenbach, M., Shinkawa, H., Nimi, O., and Kinashi, H. (1995). Physical map of a linear chromosome of *Streptomyces griseus*. J. Bacteriol. 177:6492-9498.

Lever, M. (1973). Colorimetric and fluorometric carbohydrate dtermination with *p*-hydroxybenzoic acid hydrazide. Biochem. Med. &:274-281.

Lever, M. (1977). Carbohydrate determination with 4-hydroxybenzoic acid hydrazide (PAHBAH): effect of bismuth on the reaction. Anal. Biochem. 81:21-27.

Linder, M., Mattinen, M.-L., Kontteli, M., Lindeberg, G., Stahlberg, J., Drakenberg, T., Reinikainen, T., Petterson, G., and Annila, A. (1995). Identification of functionally important amino acids in the cellulose-binding domain of *Trichoderma reesei* cellobiohydrolase I. Protein Sci. 4:1056-1064.

Lu, J. (1997). Collectins: collectors of microorganisms for the innate immune system. BioEssays 19(6):509-518.

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

Mackenzie, L.F., Brooke, G.S., Cutfield, J.F., Sullivan, P.A., and Withers, S.G. (1997). Identification of Glu-330 as the catalytic nucleophile of *Candida albicans* exo-β-1,3-glucanase. J. Biol. Chem. 272:3161-3167.

Mackenzie, L.F., Wang, Q., Warren, R.A.J., Withers, S.G. (1998). Glycosynthases: Mutant glycosidases for oligosaccharide synthesis. J. Am. Chem. Soc. 120:5583-5584.

MacLeod, A.M., Lindhorst, T., Withers, S.G. and Warren, R.A.J. (1994). The acid/base catalyst in the exoglucanase/xylanase (Cex) from *Cellulomonas fimi*. Gene 121:143-147.

Mahuran, D.J. (1995). β-hexosaminidase: Biosynthesis and processing of the normal enzyme, and identification of mutations causing Jewish Tay-Sachs Disease. Clinical Biochemistry 28:101-106.

Margolles-Clark, E., Tenkanen, M., Luonteri, E., and Penttilae (1996) Three alpha-galactosidase genes of *Trichoderma reesei* cloned by expression in yeast. Eur. J. Biochem. 240: 104-111.

Matsudaira, P. (1990). Limited N-terminal sequence analysis. Methods Enzymol. 182:602-613.

McCarter, J.D., Adam, M.J., and Withers, S.G. (1992). Binding energy and catalysis. Biochem. J. 286:721-727.

McCarter, J.D., Adam, M.J., Hartman, N.G., and Withers, S.G. (1994). *In vivo* inhibition of β-glucosidase and β-mannosidase activity in rats by 2-deoxy-2-fluoro-β-glycosyl fluorides and recovery of activity *in vivo* and *in vitro*. Biochem. J. 301:343-348.

McCarter, J.D., and Withers, S.G. (1994). Mechanism of enzymatic glycoside hydrolysis. Curr. Opin. Struct. Biol. 4:885-892.

McCleary, B. (1983). β-D-mannosidase from Helix pomatia. Carb. Res. 111:297-310.

McCleary, B.V. (1978). Soluble, dye-labeled polysaccharides for the assay of endohydrolases. Methods in Enzymol. 160:74-86.

McCleary, B.V., Clark, A.H., Dea, I.C.M., and Rees, D.A. (1985). The fine structure of carob and guar galactomannans. Carb. Res. 139:237-260.

McCleary, B.V., Taravel, F.R., and Joseleau, J.-P. (1983a). Characterisation of the oligosaccharides produced on hydrolysis of galactomannan with β-D-mannanase. Carb. Res. 118:91-109.

McIntosh, L.P., Hand, G., Johnson, P.E., Joshi, M.D., Koerner, M., Plesniak, L.A., Ziser, L., Wakarchuk, W.W., and Withers, S.G. (1996). The pKa of the general acid/base carboxyl group of a glycosidase cycles during catalysis: A 13C-NMR study of *Bacillus circulans* xylanase. Biochem. 35:9958-9966.

Meier, H. (1958). On the structure of cell walls and cell wall mannans from ivory nuts and from dates. Biochim. Biophys. Acta 28:229-240.

Meinke, A., Gilkes, N.R., Kilburn, D.G., Miller, R.C. Jr., and Warren, R.A.J. (1993). Cellulose-binding polypeptides from *Cellulomonas fimi*: Endoglucanase D (CenD), a family A β-1,4-glucanase. J. Bacteriol. 175:1910-1918.

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

Meinke, A., Gilkes, N.r., Kwan, E., Kilburn, D.G., Warren R.A.J., and Miller, R.C. Jr. (1994). Cellobiohydrolase A (CbhA) from the cellulolytic bacterium *Cellulomonas fimi* is a β-1,4,-exocellobiohydrolase analogous to *Trichoderma reesei* CBHII. Mol Microbiol. 12:413-422.

Meinke, A., Schmuck, M., Gilkes, N.R., Kilburn, D.G., Miller, R.C., and Warren, R.A.J. (1992). The tertiary structure of endo-β-1,4-glucanase B (CenB), a multidomain cellulase from the bacterium *Cellulomonas fimi*. Glycobiol. 2:321-326.

Mendoza NS, Arai M, Sugimoto K, Ueda M, Kawaguchi T, Joson LM (1995) Cloning and sequencing of beta-mannanase gene from *Bacillus subtilis* NM- 39. Biochim.Biophys.Acta 1243:552-554.

Miao, S., McCarter, J.D., Grace, M., Grabowski, G., Aebersold, R., and Withers, S.G. (1994). Identification of Glu 340 as the active site nucleophile in human glucocerebrosidase by use of electrospray tandem mass spectrometry. J. Biol. Chem. 269:10975-10978.

Millward-Saddler, S.J., Davidson, K., Hazlewood, G.P., Black, G.W., Gilbert, H.J., and Clarke, J.H. (1995). Novel cellulose-binding domains, NodB homologes and conserved modular architecture in xylanases from the aerobic soil bacteria *Pseudomonas fluorescens* subs. *cellulosa* and *Cellvibio mixtus*. Biochem. J. 312:39-48.

Millward-Sadler, S.J., Hall, J., Black, G.W., Hazlewood, G.P., and Gilbert, H.J. (1996) Evidence that the *Piromyces* gene family encoding endo-1,4-mannanases arose through gene duplication. FEMS Microbiol. Lett. 141:183-188.

Milward-Sadler, S.J., Poole, D.M., Hernissat, B., Hazlewood, G.P., Clarke, J.H., and Gilbert, H.J. (1994). Evidence for a general role for high-affinity non-catalytic cellulose-binding domains in microbial plant cell wall hydrolases. Mol. Microbiol. 11:375-382.

Mizuno, Y., Kozutsumi, Y., Kawasaki, T., and Yamashina, I. (1981) Isolation and characterization of a mannan-binding protein from rat liver. J. Biol. Chem. 256:4247-4252.

Morris, D.D., Reeves, R.A., Gibbs, M.D., Saul, D.J., and Berquist, P.L. (1995). Correction of the β-mannanase domain of the *celC* pseudogene from *Caldocellosiruptor saccharolyticus* and activity of the gene product on kraft pulp. Appl. Environ. Microbiol. 61:2262-2269.

Murata, T., Akimoto, S., Horimoto, M., Kusui, T. (1997). Galactosyl transfer onto *p*-nitrophenyl β-D-glucoside using β-D-galactosidase from *Bacillus cirulans*. Bio. Biotech. Biochemisty 61:1118-1120.

Nakamura, K, Kashiwagi, S., and Takeo, K. (1992) Characterization of the interaction between human plasma fibronectin and collagen by means of affinity electrophoresis. J. Chrom. 597:351-356.

Namchuk, M.N., and Withers, S.G. (1995). Mechanism of *Agrobacterium* β-glucosidase: Kinetic analysis of the role of noncovalent enzyme/substrate interactions. Biochemistry 34:16194-16202.

Neufeld, E.F. (1991). Lysosomal storage diseases. Annu. Rev. Biochem. 60:257-280.

Neustroev, K.N., Krylov, A.S., Firsov, L.M., Abroskina, O.N., and Khorlin, A.Y. (1991). Isolation and properties of β-mannosidase from *Aspergillus awamori*. Biokihimiya 56:1406-1412.

Nidetzki, B., Steiner, W., Hayn, M., and Claeyssens, M. (1994). Cellulose hydrolysis by the cellulases from *Trichoderma reesei*: a new model for synergistic interaction. Biochem. J. 298:705-710.

Nieduszinski, I., and Marchessault, R.H. (1972). the crystalline structure of poly-β,D(1,4)mannose: mannan I. Can. J. Chem. 50:2130-2138.

Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. (1997). Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Prot. Eng. 10:1-6.

Notenboom, V., Birsan, C., Warren, R.A.J., Withers, S.G., and Rose, D.R. (1998). Exploring the cellulose/xylan specificity of the beta-1,4-glycanase cex from *Cellulomonas fimi* through crystallography and mutation. Biochemistry. 37: 4751-4758

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

Olabarria, G., Carrascosa, J., Pedro, M.A., and Berenguer, J. (1996) A conserved motif in S-layer proteins is involved in peptidoglycan binding in *Thermus thermophilus*. J. Bacteriol. 178:4765-4772.

Pace, N.C., Vajdos, F., Fee, L., Grimsely, G., and Gray, T. (1995). How to measure and predict the molar absorption coefficient of a protein. Protein Sci. 4:2411-2423.

Posci, I., Taylor, S.A., Richardson, A.C., Smith, B.V., and Price, R.G. (1993). Comparison of several new chromogenic galactosidases as substrates for various β-D-galactosidases. Biochim. Biophys. Acta 1163:54-60.

Puls, J., and Schuseil, J. (1993). Chemistry of hemicelluloses: relationship between hemicellulose structure and enzymes required for hydrolysis. pp1-28, In: Coughlan, M.P. and Hazlewood, G.P. (Eds.), Hemicellulose and hemicellulases, Portland press, London.

Raven, J.A. (1996). The role of autotrophs in global CO2 cycling. In: Microbial growth on C1 Compounds (Lidstrom, M.E. and Tabita, F.R., eds.), pp. 351-358. Kluwer, Dordrecht.

Ries, W., Hotzy, C., Schocher, I., Sleytr, U.B., and Sagra, M. (1997). Evidence that the N-terminal part of the S-layer protein from *Bacillus stearothermophilus* PV72/p2 recognizes a secondary cell wall polymer. J. Bacteriol. 179:3892-3898.

Rol, F. (1973). Locust bean gum, pp 323-337. Whistler, R.L., Miller, J.N. (Eds.), Industrial gums, Academic press, New York.

Rouvinen, J., Bergfors, T., Teeri, T., Knowles, J.K.C., and Jones, T.A. (1990) Three-dimensional structure of cellobiohydrolase II from *Trichoderma reesei*. Science. 249(4967): 380-386

Sakon, J., Irwin, D., Wilson, D.B., and Karplus, P.A. (1997). Structure and mechanism of endo/exocellulase E4 from *Thermomonospora fusca*. Nature Struct. Biol. 4:810-817.

Saloheimo, A., Henrissat, B., Hoffren, A.M., Teleman, O., and Penttila, M. (1994). A novel, small endoglucanase gene, egl5, from *Trichoderma reesei* isolated by expression in yeast. Mol. Microbiol. 13:219-228.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual. 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Sanjeev, J., Drendel, W.B., Chen, Z.W., Mathews, F.S., Sly, W.S. and and Grubb, J.H. (1996). Structure of human β-glucuronidase reveals candidate lysosomal targeting and active-site motifs. Nature Struct. Biol. 3:375-380.

Sawadogo, M., and Van Dyke, M.W. (1991). A rapid method for the purification of deprotected oligodeoxynucleotides.NAR 19:674.

Shen, H., Gilkes N.R., Kilburn, D.G., Miller, R.C., and Warren, R.A.J. (1995). Cellobiohydrolase B, a second exocellobiohydrolase from the cellulolytic bacterium *Cellulomonas fimi*. Biochem. J. 311:67-74.

Sinnott, M.L. (1990). Catalytic mechanisms of enzymic glycosyl transfer. Chem. Rev. 90:1171-1202.

Spezio, M., Wison, D.B., and Karplus, P.A. (1993) Crystal structure of the catalytic domain of a thermophilic cellulase. Biochemistry 32:9906-9912

Stackerbrandt, E., and Prauser, H. (1991). Assignment of the genera *Cellulomonas, Oerskovia, Promicromonospora* and *Jonesia* to *Cellulomonadacaea* fam. nov. system. Appl. Microbiol. 14:261-265.

Stålbrand, H., Saloheimo, A., Vehmaanpera, J., Henrissat, B., and Pentilla, M. (1995). Cloning and expression in *Saccharomyces cerevisiae* of a *Trichoderma reesei* beta-mannanase gene containing a cellulose-binding domain. Appl. Environ. Microbiol. 61(3):1090-1097.

Stålbrand, H., Siika-aho, M., Tenkanen, M., and Viikari, L. (1993). Purification and characterization of two beta-mannanases from *Trichaderma reesei*. J. Biotech. 29:229-242.

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

Street, I.P., Kempton, J.B., and Withers, S.G. (1992). Inactivation of a β-glucosidase through the accumulation of a stable 2-deoxy-2-fluoro-α-D-glucopyranosyl-enzyme intermediate: a detailed investigation. Biochem. 31: 9970-9978.

Stryer, L. (1988). Carbohydrates. In: Biochemistry pp.331-348. W.H. Freeman and Company, New York.

Takeo, K. (1984). Affinity electrophoresis: Principles and applications. Electrophoresis 5:187-195.

Takeo, K., and Nakamura, S. (1972). Dissociation constant of glucan phosphorylase of rabbit tissues studied by polyacrylamide gel disc electrophoresis. Arch. Biochem. Biophys. 10:1684.

Talbot, G., and Sygusch, J. (1990). Purification and charaterization of thermostable beta-mannanase and alpha-glactosidase from *Bacillus stearothermophilus*. Appl. Environ. Microbiol. 56:3505-3510.

Teather, R.M., and Wood, P.J. (1982). Use of Congo Red - polysaccharide interactions in enumeration and characterization of cellulolytic bacteria from the bovine rumen. Appl. Environ. Microbiol. 43:777-780.

Teeri, T.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.

Tenkanen, M., Makkonen, M., Perttula, M., Viikari, L., and Teleman, A. (1997). Action of *Trichoderma reesei* on galactoglucomannan in pine kraft pulp. J. Biotech. 57:191-204.

Tews, I., Perrakis, A., Oppenheim, A., Dauter, Z., Wilson, K.S. and Vorgias, C.E. (1996). Bacterial chitobiase structure provides insight into catalytic mechanism and the basis of Tay-Sachs disease. Nature Struct. Biol. 3:638-648.

Thiem, J. (1995). Application of enzymes in synthetic carbohydrate chemistry. FEMS Microbiol. Rev. 16:193-211.

Thomas, M.G. (1994). A procedure for second-round differential screening of cDNA libraries. Biotechniques. 16:988-989.

Timell, T.E. (1967). Recent progress in the chemistry of wood hemicelluloses. Wood Sci. Technol. 1:45-70.

Tomme, P, Kwan, E., Gilkes, N.R., Kilburn, D.G., and Warren, R.A.J. (1996a). Characterization of CenC, an enzyme from *Cellulomonas fimi* with both endo- and exoglucanase activities. J. Bacteriol. 178:4216-4223.

Tomme, P., Creagh, A.L., Kilburn, D.G., and Haynes, C.A. (1996b). Interaction of polysaccharides with the N-terminal cellulose-binding domain of *Cellulomonas fimi* CenC. 1. Binding specificity and calorimetric analysis. Biochemistry 35:13885-13894.

Tomme, P., Warren, R.A.J., Gilkes, N.R. (1995). Cellulose hydrolysis by bacteria and fungi. Adv. Microb. Physiol. 37:1-81.

Toone, E.J., Simon, E.S., Bednarski. M.D., and Whiteside, G.M. (1989). Enzyme catalyzed synthesis of carbohydrates. Tetrahedron. 45(259):5365-5422.

Törrönen, A. and Rouvinen, J. (1995). Structural comparison of two major endo-1,4-xylanases from *Trichoderma reesei*. Biochemistry 34:847-856.

Towbin, H., Sthaehelin, T., and Gordon, J. (1970). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. USA 76:4350-4354.

Trevor, J.T. (1996). Genome size in bacteria. Antonie van Leuwenhoek 69:293-303.

Tull, D., Withers, S.G., Gilkes, N.R., Kilburn, D.G., Warren, R.A.J. and Aebersold, R. (1991). Glutamic acid 274 is the nucleophile in the active site of a "retaining" exoglucanase from *Cellulomonas fimi*. J. Biol. Chem. 266(24):15621-15625.

van Emmerik LC, Kuijper EJ, Fijen CA, Dankert J, and Thiel S (1994) Binding of mannan-binding protein to various bacterial pathogens of meningitis. Clin.Exp.Immunol. 97:411-416.

Wakarchuk, W.W., Kilburn, D.G., Miller, R.C. Jr., and Warren, R.A.J. (1984). J. Gen. Microbiol. 130:1385-1389.

Warren, R.A.J. (1996). Microbial hydrolysis of polysaccharides. Annu. Rev. Microbiol. 50:183-212.

White, A., Withers, S.G., Gilkes, N.R. and Rose, D.R. (1994). Crystal structure of the catalytic domain of the b-1,4,-glycanase Cex from *Cellulomonas fimi*. Biochem. 33:12456-12552.

Williamson, M.P., LeGal-Coeffet, M.-F., Sorimachi, K., Furniss, C.S.M., Archer, D.B., Williamson, G. (1997) Function of conserved tryptophans in the Aspergillus niger glucoamylase 1 starch binding domain. Biochemistry. 36(24): 7535-7539

Withers, S.G., and Aebersold, R. (1995). Approaches to labeling and identification of active site residues in glycosidases. Protein Sci. 4:361-372.

Withers, S.G., Mackenzie, L., and Wang, Q. (1998). Methods and compositions for the synthesis of oligosaccharides, and the products formed thereby. United States Patent. Patent Number:5,716,812.

Withers, S.G., Rupitz, K., Trimbur, D., and Warren, R.A.J. (1992). Mechanistic consequences of mutation of the active site nucleophile Glu 358 in *Agrobacterium* β-glucosidase. Biochemistry 31:9979-9985.

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

Wong-Madden, S.T., and Londry, D. (1998). New England Biolabs.

Xiuzhu, D., Schyns, P.J.Y.M.J., and Stams, A.J.M. (1991). Degradation of galactomannan by a *Clostridium butyricum* strain. Antonie van Leuwenhoek 60:109-114.

Xu, G.Y., Ong, E., Gilkes, N.R., Kilburn, D.G., Muhandrian, D.R., Harris-Brandts, M., Carver, J.P., Kay, L.E., Harvey, T.S. (1995). Solution structure of a cellulose binding domain from *Cellulomonas fimi* by nuclear magnetic resonace spectroscopy. Biochemistry 34:6993-7009.

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

Yasukochi, T., Fukase, K., Suda, Y., Takagaki, K., Endo, M., and Kusumoto, S. (1997). Enzymatic synthesis of 4-methylumbelliferyl glycosides of trisaccharide and core tetrasaccharide, Gal(β1-3) Gal(β1-4) Xyl and GlcA(β1-3) Gal(β1-3) Gal(β1-4) Xyl, corresponding to the linkage region of proteoglycans. Bull. Chem. Soc. Jpn., 70:2719-2725.

Yui, T., Miyawaki, K., Yada, M., and Ogawa, K. (1997). An evaluation of crystal structure of mannan I by X-ray powder diffraction and molecular mechanics studies. Int. J. Biol. Macromol. 21(3): 243-250.