VTT PUBLICATIONS 277

Structure-function studies of two polysaccharide-degrading enzymes: *Bacillus stearothermophilus* α-amylase and *Trichoderma reesei* cellobiohydrolase II

Anu Koivula

VTT Biotecnology and Food Research

Department of Biosciences
Division of Biochemistry
University of Helsinki
Helsinki, Finland

ACADEMIC DISSERTATION

To be presented with the permission of the Faculty of Science of the University of Helsinki, for public criticism at the Department of Biosciences, Biocenter, Viikinkaari 5, Helsinki, on August 2nd, 1996, at 12 o'clock noon.

TECHNICAL RESEARCH CENTRE OF FINLAND
ESPOO 1996

ISBN 951-38-4935-X
ISSN 1235-0621

JULKAISIJA – UTGIVARE – PUBLISHER

Valtion teknillinen tutkimuskeskus (VTT), Vuorimiehentie 5, PL 42, 02151 ESPOO
puh. vaihde (09) 4561, telekopio 456 4374

Statens tekniska forskningscentral (VTT), Bergsmansvägen 5, PB 42, 02151 ESBO
tel. växel (09) 4561, telefax 456 4374

Technical Research Centre of Finland (VTT), Vuorimiehentie 5, P.O.Box 42, FIN–02151 ESPOO, Finland
phone internat. + 358 9 4561, telefax + 358 9 456 4374

VTT Bio- ja elintarviketekniikka, Geenitekniikka, Biologinkuja 1, PL 1500, 02044 VTT
puh. vaihde (09) 4561, faksi (09) 455 2103

VTT Bio- och livsmedelsteknik, Genteknik, Biologgränden 1, PB 1500, 02044 VTT
tel. växel (09) 4561, fax (09) 455 2103

VTT Biotechnology and Food Research, Genetic Engineering, Biologinkuja 1, P.O.Box 1500, FIN–02044 VTT, Finland
phone internat. + 358 9 4561, fax + 358 9 455 2103

Technical editing Leena Ukskoski

VTT OFFSETPAINO, ESPOO 1996

Koivula, Anu. Structure-function studies of two polysaccharide-degrading enzymes: *Bacillus stearothermophilus* α-amylase and *Trichoderma reesei* cellobiohydrolase II. Espoo 1996, Technical Research Centre of Finland, VTT Publications 277. 97 p. + app. 45 p.

UCD 577.21:577.51.3

Keywords enzymes, carbohydrates, amylase, cellulase, protein structure, cellobiohydrolase, random mutagenesis, site-directed mutagenesis, *Bacillus stearothermophilus*, *Trichoderma reesei*, theses

ABSTRACT

Amylases and cellulases are important enzymes both for the global carbon cycle on earth and for biotechnical applications. They are capable of degrading polysaccharides, which are chemically simple polymers of repeating glucose units that form very complex and water-insoluble macroscopic structures. The enzymatic degradation of starch and cellulose is poorly understood at the molecular level. The cloning and DNA sequence determination of amylase- and cellulase-encoding genes from various organisms allows the use of modern techniques of molecular biology to study and alter their function.

In the present investigation two different approaches were used to study the relationship between enzyme structure and function. In the first of these a random mutagenesis method developed in this study was applied. This method is applicable even without structural knowledge of a protein. The complete *Bacillus stearothermophilus* α-amylase gene was subjected to random mutagenesis, which was optimised to produce a single amino acid change at a time in the corresponding enzyme. Nearly 100 different mutant α-amylases were produced in an *Escherichia coli* host. The mutated enzymes had different activities on substrate plates used in screening. The location of the mutation(s) was analysed by sequencing the mutated gene. A molecular model of *Bacillus stearothermophilus* α-amylase was constructed on the basis of sequence alignment and the known structure of a fungal α-amylase from *Aspergillus oryzae*. The protein is expected to fold into three domains (A, B and C), like many other α-amylases. The data obtained from the analysis of the random mutants was compared to the constructed three-dimensional model. A reasonably good overall correlation was obtained between the mutant data and the model. Two areas on the α-amylase structural model were identified as being important for the activity on polymeric substrate: the open active site cleft situated between domains A and B and containing conserved amino acids known to be important for catalysis and multiple binding of glucose units in other α-amylases; and an interface between the catalytic domain A and domain C about 30 Å away from the active site groove.

The three-dimensional structure of *Trichoderma reesei* cellobiohydrolase II (CBHII) catalytic domain has been determined. The catalytic domain has an α/β barrel fold similar to that of the α-amylase catalytic domain A. Two stable surface loops generate a 20 Å long tunnel for substrate binding and catalysis. The active site tunnel contains four defined binding sites (A-D) for glucosyl units. In the

sites A, C and D, sugar binding is characterized by hydrophobic interactions of the glucosyl units with W135 (A), W269 (C) and W367 (D). At the entrance of the active site tunnel, W272 apparently gives rise to favourable stacking interactions with the sixth glucosyl unit. This site is referred to as the putative binding site F. CBHII, like all glycosidases, cleaves the glycosidic linkage by general acid catalysis. CBHII is an inverting enzyme in which D221, situated between subsites B and C, acts as a proton donor. D175 lying next to D221 either stabilizes hypothetical carbonium ion intermediates or facilitates the protonation of D221, or both. The base has not yet been identified although the role has been attributed to D401, which is located on the opposite side of D221 with respect to the glycosidic linkage.

Site-directed mutagenesis was used to study the role of three amino acids in the active site of CBHII. A tyrosine residue, Y169, located at site B close enough to interact with both D175 and the sugar hydroxyl at site B, was mutated to phenylalanine. The resultant Y169F mutant showed increased binding but reduced catalytic rate on small soluble cello-oligosaccharides. One plausible explanation for these results is that the Y169 residue promotes the distorted sugar conformation at site B, possibly by forming a hydrogen bond, and that this conformation is essential for efficient catalysis by CBHII. The proposed existence of a strained sugar conformation at the subsite preceding the scissile bond gains support from structural studies of other inverting enzymes. In addition, the pH-activity profile of the Y169F mutant declines at low pH, suggesting that the second role of Y169 is to affect the protonation state of the active site carboxylates, D175 and D221.

The tryptophan residues at subsites A and F were also mutated. In both cases removal of the indole ring affected the catalytic rate. On the basis of the experimental data on both the wild type and mutated enzymes the tight binding of an intact glucose ring in binding site A seems to be essential for efficient catalysis and is partly dictated by the W135. It was also shown that the CBHII active site tunnel contains at least one additional binding site (F) at the mouth of the tunnel. Mutations of W272 at subsite F increased the hydrolysis rates of short cello-oligosaccharides, apparently by affecting the non-productive binding modes. It is plausible that site F is important for the breakdown of crystalline cellulose.

PREFACE

This work was carried out in the Biotechnology and Food Research Institute (earlier the Biotehnical Laboratory) of the Technical Research Centre of Finland (VTT). I wish to thank Research Director, Professor Juha Ahvenainen, Professor Tor-Magnus Enari and Professor Matti Linko for excellent working facilities and the encouraging atmosphere towards thesis work at VTT. I also thank Professor Carl Gahmberg, head of the Division of Biochemistry at the University of Helsinki for his flexibility concerning this work. I am grateful to Assistant Professor Reijo Laakso and Docent Mauno Vihinen for critical reading of my thesis.

My sincere thanks are due to Professor Hans Söderlund, who guided me in the world of gene technology during my study years, and who has always had time for helpful conversations concerning the work. I warmly thank Docent Päivi Lehtovaara, who supervised me during the random mutagenesis work and taught me a lot about the basics of scientific work. I also thank Professor Jonathan Knowles for his enthusiasm and for bringing new ideas to the laboratory. Docent Tuula Teeri is gratefully acknowledged for leading the cellulase group with never-ending optimism and helping with the articles and with this manuscript.

I thank my collaborators, Jaana Bamford, Ari Hemminki, Liisa Holm, Alwyn Jones, Tiina Kinnari, Gerard Kleywegt, Arja Lappalainen, Arja Mäntylä, Juha Rouvinen, Pirkko Suominen, Michael Szardenings, Olle Teleman, Anne Valkeajärvi and Susanna Virtanen for sharing their efforts and ideas. Collaboration with Torbjörn Drakenberg, Vesa Harjunpää and Anita Teleman of the VTT Chemical Technology has been very helpful in revealing the secrets of enzyme kinetics. I warmly thank Professor Marc Claeyssens for the pleasant and fruitful time I spent in his laboratory in Ghent. The stimulating discussions of everybody involved in random mutagenesis or cellulase work during the years is also acknowledged. My biggest hugs go to Tapani Reinikainen and Laura Ruohonen, who tackled the challenging world of *Trichoderma* cellulases with me. Their professional skills and friendship have been indispensable.

My special thanks are due to Kariitta Berg, Tuula Kuurila, Ulla Lahtinen and Riitta Suihkonen for their skillful technical assistance and positive attitude. I am grateful to Michael Bailey for revising the language of this thesis and to Oili Lappalainen for excellent secretarial work. I wish to thank all the members of Gene group, both former and present, visiting and permanent, for a pleasant working atmosphere and interesting discussions on various matters of Life. Especially Ansku Hoffrén, Tiina Nakari-Setälä, Eini Nyyssönen, Merja Penttilä, Anu Saloheimo, Markku Saloheimo and Sipo Vanhanen are thanked for being good friends and arranging off-science activities both in and outside the lab.

Financial support from Finnish Academy during the years 1988-1991 is gratefully acknowledged.

Finally, I wish to express my gratitude to my late mother and father for their unquestioning love and support and to my close friends and relatives for their caring attitude.

CONTENTS

Appendices of this publication are not included in the PDF version. Please order the printed version to get the complete publication (http://www.inf.vtt.fi/pdf/publications/1996/)

LIST OF PUBLICATIONS

This thesis is based on the following articles, referred to in the text by the Roman numerals given below. In addition, unpublished data is presented concerning the mutagenesis of the CBHII active site residues W135 and W272. Parts of these data have been previously reviewed in the proceedings publications: Ruohonen *et al*., 1993; Teeri *et al*., 1995; Koivula *et al*., 1996.

I Lehtovaara, P. M., Koivula, A. K., Bamford, J. and Knowles, J. K. C. 1988. A new method for random mutagenesis of complete genes: enzymatic generation of mutant libraries *in vitro*. Protein Eng. 2, pp. 63 - 68.

II Holm, L., Koivula, A. K., Lehtovaara, P. M. , Hemminki, A. and Knowles, J. K. C. 1990. Random mutagenesis used to probe the structure and function of *Bacillus* s*tearothermophilus* alpha-amylase. Protein Eng. 3, pp. 181 - 191.

III Koivula, A., Reinikainen, T., Ruohonen, L., Valkeajärvi, A., Claeyssens, M., Teleman, O., Kleywegt, G., Szardenings, M., Rouvinen, J., Jones, T. A. and Teeri, T. T. 1996. The active site of *Trichoderma reesei* cellobiohydrolase II: The role of tyrosine-169. Accepted for publication in Protein Eng.

IV Koivula, A., Lappalainen, A., Virtanen, S., Mäntylä, A. L., Suominen, P. and Teeri, T. T. Immunoaffinity purification of cellobiohydrolase II mutants from *Trichoderma reesei* strains lacking the major endoglucanases. Submitted for publication.

ABBREVIATIONS

amdS	*Aspergillus nidulans* gene encoding acetamidase
CBH	cellobiohydrolase
CBD	cellulose-binding domain
cbh	*Trichoderma reesei* gene encoding cellobiohydrolase
CGTase	cyclodextrin glycosyltransferase
CMC	carboxymethyl cellulose
DEAE	diethylaminoethyl
DNS	dinitrosalicylic acid
DP	degree of polymerization
EC	Enzyme Commission
EG	endoglucanase
egl	*Trichoderma reesei* gene encoding endoglucanase
ELISA	enzyme-linked-immunosorbent-assay
Fn3	fibronectin type III
FPLC	fast protein liquid chromatography
HCA	hydrophobic cluster analysis
HEC	hydroxyethyl cellulose
HEWL	hen egg-white lysozyme
HPLC	high performance liquid chromatography
MeUmb	4-methylumbelliferyl
$MeUmb(Glc)_1$	4-methylumbelliferyl-β-D-glucoside
$MeUmb(Glc)_2$	4-methylumbelliferyl-β-D-cellobioside
$MeUmb(Glc)_3$	4-methylumbelliferyl-β-D-cellotrioside
NaAc	sodium acetate
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
SBD	starch-binding domain
SDS	sodium dodecyl sulphate
TIM	triose phosphate isomerase
TLC	thin-layer chromatography
3D	three-dimensional
trpC	*Aspergillus nidulans* gene encoding phosphoribosylanthranilate isomerase
wt	wild type
X-gal	bromo-4-chloro-3-indolyl-β-D-galactopyranoside

1 INTRODUCTION

Cellulose and starch are the two most common carbohydrate polymers in nature. They are synthesized by all higher plants, using glucose as a building block. Starch functions as the major source of carbon and energy for most organisms, providing e.g. up to 80 % of the total caloric intake of humans. Cellulose functions as a structural element of plant cell walls but also provides an energy source for many micro-organisms in plant litter and soil.

Due to the complexity and resistance to degradation of both polymers, complete degradation of cellulose and starch usually requires the action of many enzymes. Cellulose and starch require different enzymatic systems for their hydrolysis. Enzymes capable of degrading starch are called amylolytic enzymes, or amylases, and enzymes hydrolysing cellulose are called cellulases. Amylases and cellulases are important enzymes in many biotechnical applications. Various starch hydrolysates are used as sweeteners in the food industry (Kennedy *et al.*, 1988). Amylases are also used to convert cereal starch into fermentable sugars for *Saccharomyces cerevisiae* yeast strains in the brewing industry (Kennedy *et al.*, 1988). The textile industry (e.g. biobleaching of indigo dyed jeans) utilizes cellulases, whereas the detergent and paper industries use both amylases and cellulases (Kennedy *et al.*, 1988; Pommier *et al*., 1990; Lange, 1993; Prasad *et al*., 1993; Béguin and Aubert, 1994; Borchert *et al*., 1995).

Increasing demand for more environmentally friendly processes will lead to increasing use of hydrolytic and other enzymes instead of chemicals. The conditions used in biotechnical processes (e.g. temperature and pH) may be different from the physiological conditions in which enzymes operate optimally. In some processes, combinations of protein properties not present in nature are required. With the aid of gene technology it has become possible to alter or engineer protein properties in a controlled manner. Genes coding for proteins can be isolated, sequenced and expressed in foreign host organisms. The structure of a gene can also be specifically altered, thus affecting the structure of the corresponding protein. Many hydrolytic enzymes are composed of different functional units, called domains. These domains or modules can be grafted from different enzymes to form new protein entities with new combinations of functions. Furthermore, alterations can also be made within domains. If the three-dimensional (3D) structure or structural model of the protein domain is available, mutations can be designed on the basis of the structure. Even without structural data it is possible to make mutations randomly and then screen for altered (improved) properties. All these approaches can also be used to gain more profound knowledge of the structure-function relationships of the protein. The best correlation can be made when the 3D structure of the protein is available.

The enzymatic actions of amylases and cellulases have been studied for many years because of their applicability in industrial processes. Three-dimensional structures of the enzyme proteins have also emerged, allowing protein engineers to study the enzymatic behaviour in detail. Despite this

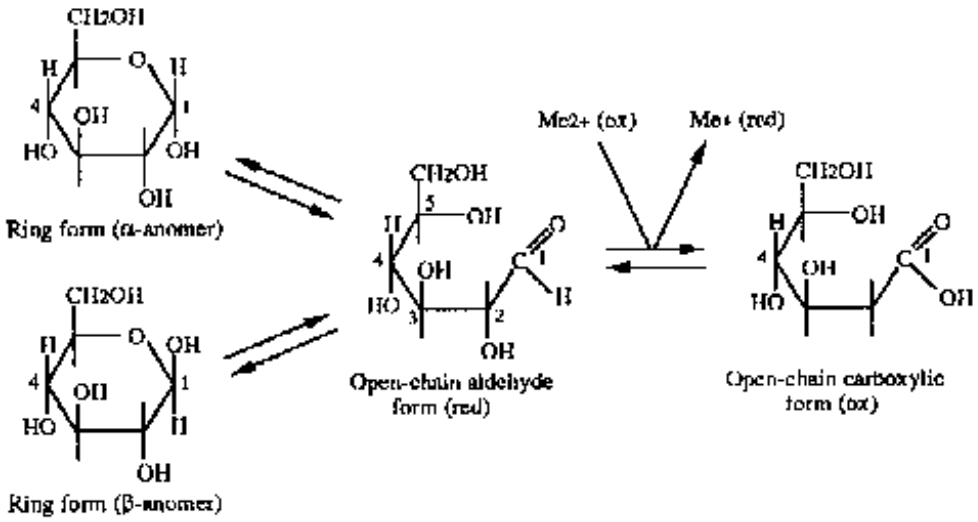

Figure 1. Equilibrium between α- and β-D-glucopyranosyl forms of D-glucose in solution and the oxidation in the presence of metal ions. The α- and β-anomers differ only at the configuration of the anomeric C1 carbon. The open-chain aldehyde form of D-glucose is able to reduce metal ions (designated Me) while it is oxidized to a carboxylic form. The carbon numbering is indicated. Modified from Divne, 1994.

progress, there are still many open and challenging questions, which arise from the complexity of both the substrates and the enzyme systems.

1.1 CELLULOSE AND STARCH

During photosynthesis, carbon dioxide and water are converted to D-glucose in plants. D-Glucose can occur in two anomeric forms, α-D-glucose and β-D-glucose. In water, α- and β-D-glucose (D-glucopyranosyl) are present in an equilibrium mediated by an open-chain aldehyde form (Fig. 1). This aldehyde group on the anomeric carbon C1 has the ability to reduce metal ions, such as silver or copper, while it is itself oxidized to a carboxylic acid. Thus glucose and other carbohydrates having free C1 carbon are called reducing sugars.

D-Glucose is the basic building block of both cellulose and starch. Starch is the storage polymer for energy and is composed of α-glucose. Starch occurs mainly in the seeds, roots and tubers of higher plants. Some algae and bacteria produce a similar reserve polysaccharide called phytoglycogen**.** The majority of cellulose is produced as a component of cell walls of higher plants, where it

functions as a structural element. In addition, some micro-organisms and invertebrates can synthesize cellulose. Cellulose is composed of β-glucose units.

Starch is organized into granules which contain small amounts of non-carbohydrate components, particularly lipids, proteins and phosphate. Cellulose in the cell walls is usually closely associated with hemicellulose and lignin, forming very complex structures. These structures are proposed to involve covalent bonds from lignin to hemicellulose and possibly also to cellulose (Scalbert *et al*., 1985; Higuchi, 1990). Hemicelluloses possess a β-1,4-linked backbone similar to cellulose, but generally have a heterogeneous composition of various sugar units and branches (Wong *et al*., 1988). They are usually named according to the main sugar residue in the backbone, xylans (backbone of D-xylose units) and mannans (D-mannose) being the major groups. Lignin differs from the other structural components of plant cell walls. It is a complex polymer of aromatic alcohols formed by random coupling and containing a diverse range of chemical linkages (Crawford and Crawford, 1984; Higuchi, 1990). It is resistant to microbial degradation and also acts as a physical barrier inhibiting the biodegradation of cellulosic components of wood (Crawford and Crawford, 1984).

1.1.1 The structural organization of cellulose

Cellulose is composed of linear homopolymers of D-glucose units. The D-glucose rings are in a chair conformation ($^4C_1$) and always linked via β-1,4-bonds (Fig. 2). The smallest structural repeating unit in native cellulose is cellobiose, a disaccharide, in which the orientation of glucose rings is related by 2-fold symmetry to each other. This arrangement of glucose units leads to an extended conformation of the cellulose chain, in which successive glycosidic oxygens point in opposite directions (Fig. 2). The number of glucose units, i.e. the degree of polymerization (DP) of cellulose, varies from 100 to 14,000 depending on the source. In cell walls, cellulose exists as highly ordered crystalline fibers, which are formed when individual cellulose chains pack together via inter- and intramolecular hydrogen bonds.

In native cellulose (also called cellulose I) the glucose chains have parallel orientation, all reducing ends of the individual chains being at one end and the non-reducing at the other end of a crystalline microfibril (Fig. 2). Recent structural studies have shown that native cellulose is composed of two different crystal forms, a two-chain monoclinic phase (Iβ) and a single-chain triclinic phase (Iα) (Atalla and VanderHart, 1984; Michell, 1990; Sugiyama *et al*., 1991a; 1991b;). The two crystal structures have similar conformations but different hydrogen bonding patterns and they can co-exist within a single microfibril. Neither of the crystal lattices contains water. The monoclinic phase is more stable, apparently due to more favorable intermolecular hydrogen bonding (Sugiyama *et al*., 1991a; Atalla, 1993; Heiner *et al*., 1995). The relative amounts of Iα and Iβ forms vary depending on the origin of the cellulose. Algal-bacterial type cellulose, e.g. cellulose isolated from *Acetobacter xylinum* or *Valonia ventricosa*, is rich in the Iα component, whereas in higher plant cellulose the dominant form is

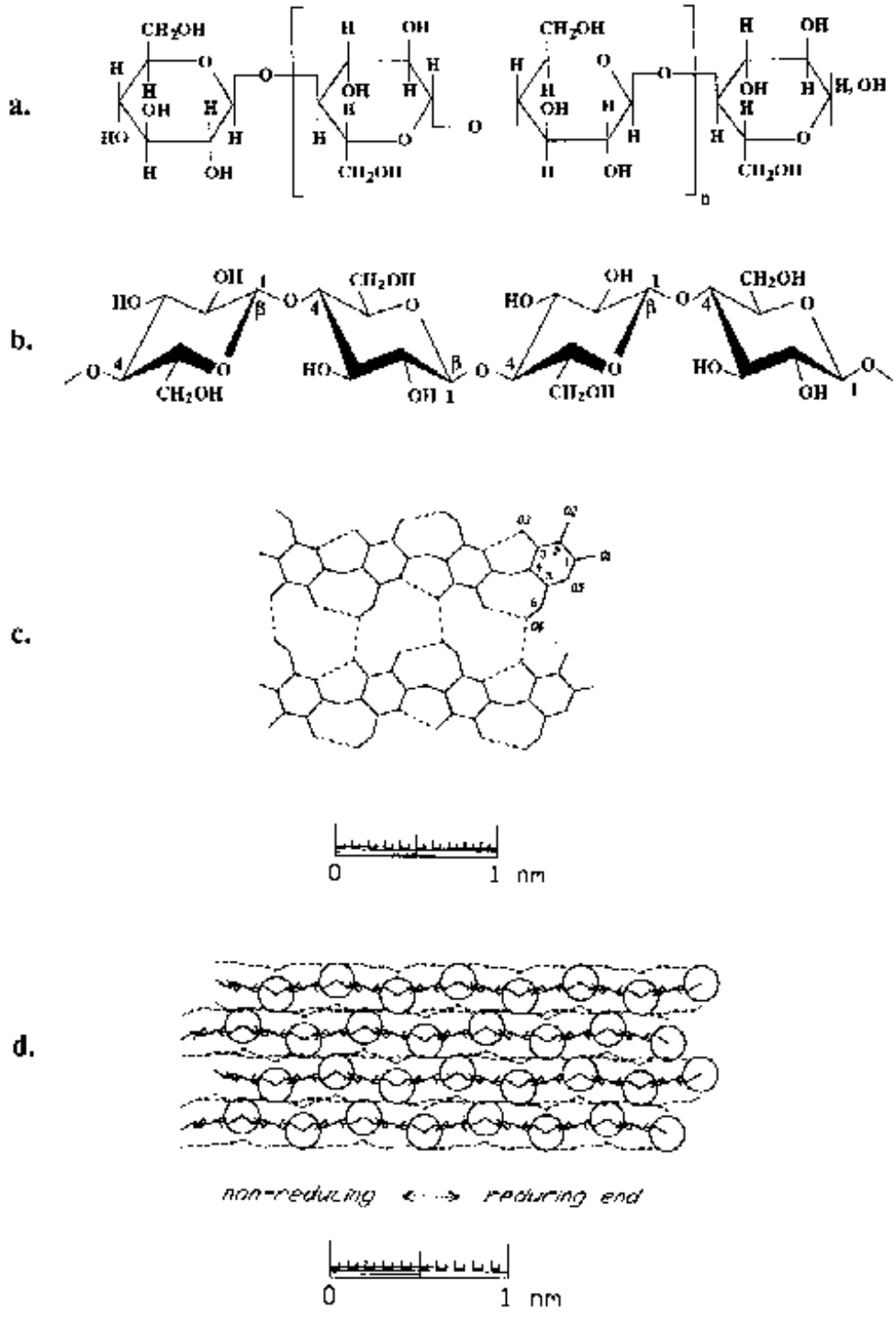

Figure 2. Structure of cellulose. The smallest repeating unit in cellulose is cellobiose (a), in which the glucose units are in chair conformation and are related by 2-fold symmetry to each other (b). From Hon, 1994. The parallel strands form intra- and intermolecular hydrogen bonds and pack into sheets (top view) (c) and then into layers forming microfibrils (side view) (d). From Ståhlberg, 1991.

the Iβ structure. In addition, the crystal dimensions vary depending on the cellulose origin (Atalla, 1993; Sugiyama *et al.*, 1991a; 1993). The cross section of algal *Valonia* cellulose is 20 nm, whereas in the cell walls of higher plants the microfibrils may be only 2 nm wide (Chanzy, 1990).

Cellulose microfibrils are not uniformly crystalline but also contain less ordered, or "amorphous" regions. The crystallinity of isolated celluloses varies depending on the origin and pretreatment. The crystallinity index of cellulose is a relative property, which can be measured by X-ray diffraction or solid state NMR. Cellulose isolated from *Valonia* is usually referred to as 100% crystalline whereas amorphous, acid swollen cellulose is 0% crystalline (Kulshreshta and Dwelz, 1973).

1.1.2 The structure and properties of starch

Starch is usually composed of two different fractions, amylose and amylopectin (Fig. 3). Amylose is mainly a linear polysaccharide formed by α-1,4-linked D-glucose residues and having some α-1,6-linked branching points. Amylopectin has a highly branched structure. The degree of polymerization (DP) of amylose chains ranges from 350 to 1000 glucose units depending on the source (Whistler and Daniel, 1985). The DP of the straight chain units in amylopectin ranges from 10 to 60, but the total DP of amylopectin can be several thousands. The relative content of amylose and amylopectin varies with the source of starch. The amylose content in most common starches, e.g. in barley, corn and potato, is 20 to 30% (Shannon and Garwood, 1984).

Native starch is partly crystalline. According to the currently accepted concept amylopectin forms the crystalline component whereas amylose exists mainly in the amorphous form (Zobel, 1992). Structural studies have shown that native starch has crystalline polymorphism similar to that of native cellulose. In X-ray diffraction, cereal starch typically gives A-type patterns of monoclinic symmetry, and tuber starch B-type patterns of hexagonal symmetry (Imberty *et al*., 1991). The three-dimensional organizations of A- and B-structures in a crystal lattice differ, but their molecular conformations are practically identical. Both structures have left-handed double helices with parallel strands (Fig. 3). Double helices contain six glucose units per turn in each chain and the glucose units are in a chair conformation ($^4C_1$). Within the double helix, there are inter-chain but no intra-chain hydrogen bonds. In addition, parallelly packed double helices are connected through a hydrogen bonding network. The double helix is very compact and there is no room for water in its center (Imberty *et al*., 1988; 1991). The crystal lattice of B-type starch contains more water molecules than the A-structure, which is proposed to be the reason for higher stability of the A-structure. It also appears that amylosic fragments with higher DP yield B-type crystals.

Amylose and amylopectin are packed into granules which are insoluble in water below 50^o C. The size and shape vary depending on the plant, the diameter of the granules ranging from 2 to 100 μm (Whistler and Daniel, 1985). The orientation of the starch chains is thought to be perpen-

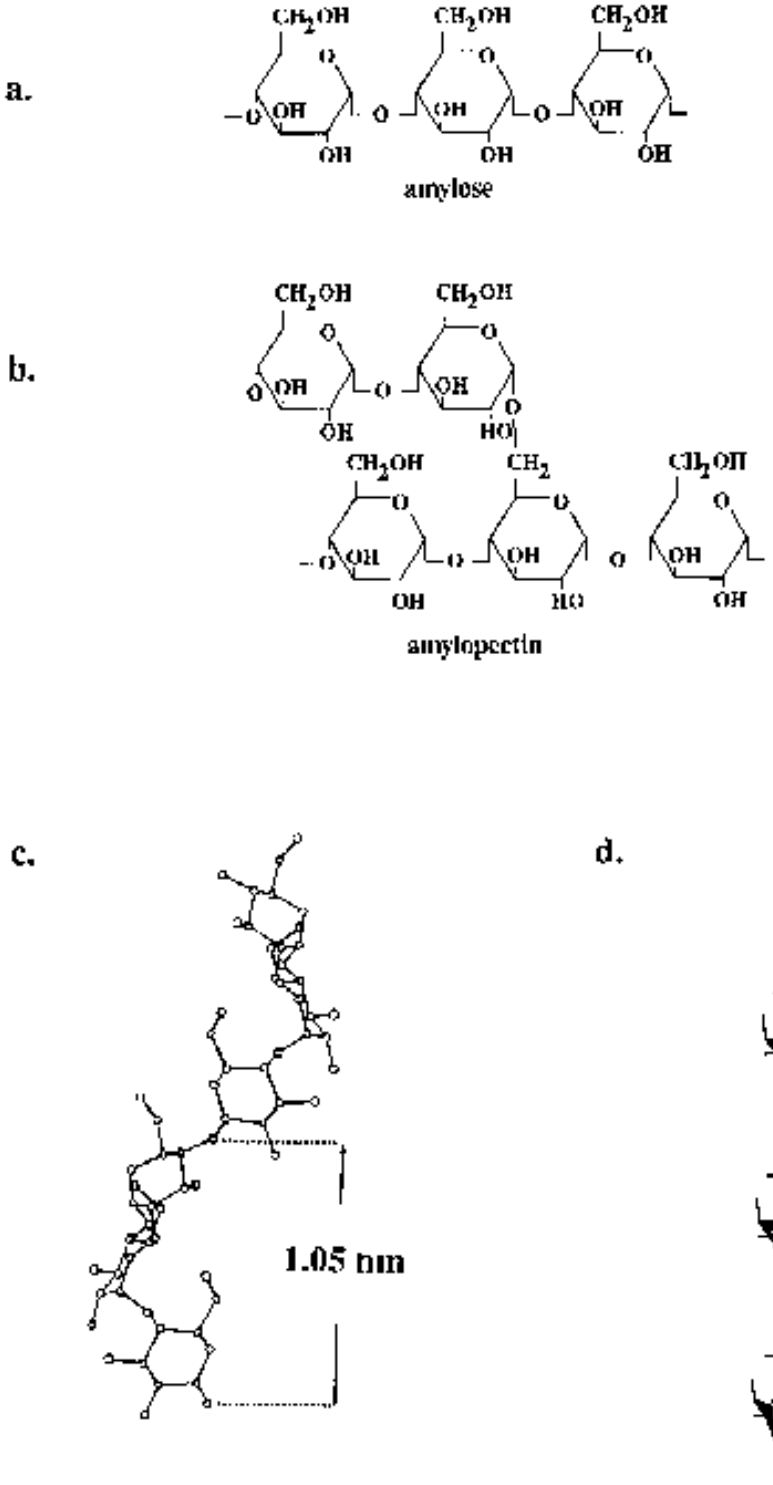

Figure 3. Starch components. Amylose (a) exists mainly in the amorphous fraction of starch granules whereas amylopectin (b) forms the crystalline part. The picture (c) shows one amylosic strand in the left-handed conformation, having six glucose units in a helical repeat in 2.1 nm. The double helix is generated by the association of two parallel single strands (d). Double helices contain inter-chain (not shown) but no intra-chain hydrogen bonds From Imberty et al., 1991.

dicular to the granule surface (French, 1984). It is not known how the amylose and amylopectin molecules are arranged relative to each other, but they are suggested to form separate entities. Amylose probably occurs in close association with fatty acids and might then adopt a single helix conformation (Zobel, 1992). A ”cluster” model has been presented for amylopectin (Fig. 4). As can be seen from Fig. 4 amylopectin has only one functional reducing end group in the molecule. The α-1,6 branch points are apparently located in the amorphous regions between crystallites but they are thought to initiate the double helix formation (Imberty *et al*., 1991). The crystallinity of native starch varies between 15 and 45% and is a relative measure similarly defined as the crystallinity of cellulose (French, 1984).

1.2 ENZYMATIC DEGRADATION OF CELLULOSE AND STARCH

Besides being present in higher plants, amylolytic enzymes are also found in microbes and animals. The degradation of cellulose occurs mainly through the action of micro-organisms in plant litter and soil. Since the native substrate is water-insoluble and cannot penetrate into cells, the biodegradation of cellulose and starch occurs extracellularly. Amylases are mainly secreted into the medium or are found membrane-bound. Some mirobial strains are known to produce intracellular amylases; the reason for this is not known (Vihinen and Mäntsälä, 1989). Cellulases act either in

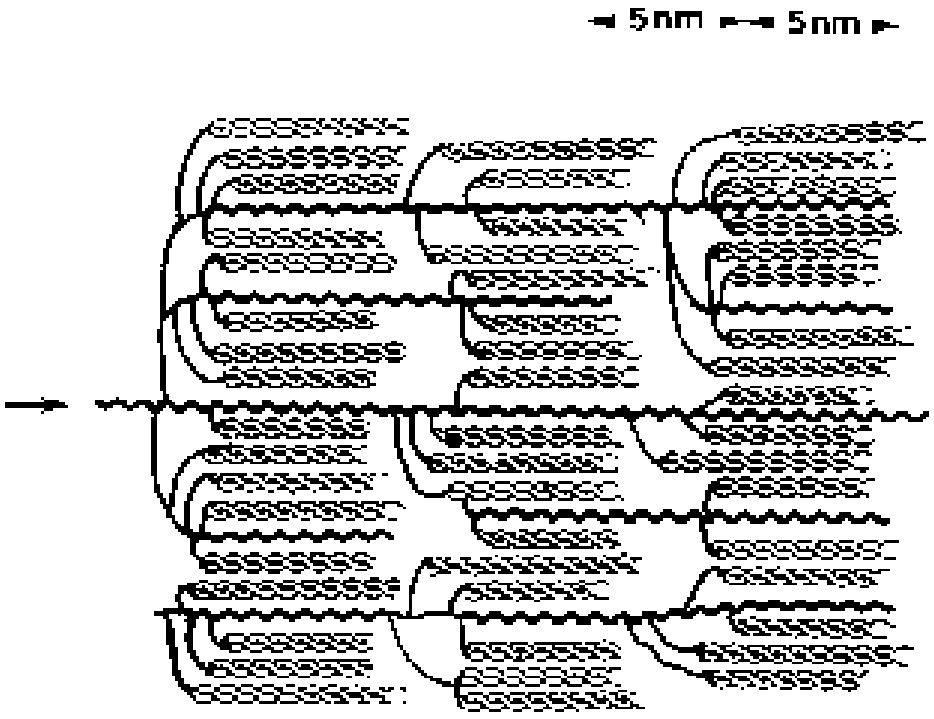

Figure 4. The cluster model of the amylopectin molecule in a starch granule. The only reducing end of the molecule is shown. The α-1,6-linkages of amylopectin apparently promote the double helix formation. From Imberty et al., 1991.

association with the outer cell envelope layer or extracellularly. The effective hydrolysis of both polysaccharides demands the action of many enzymes due to the complexity of the substrate, although a prolonged incubation with one particular enzyme can lead to (almost) complete hydrolysis. Few micro-organisms produce a complete set of enzymes capable of degrading the insoluble substrates efficiently. Complete cellulase systems are often associated with other, related hydrolases, particularly xylanases. This is apparently due to the close association of hemicelluloses with celluloses in plant cell walls.

Amylolytic and cellulolytic enzymes can be classified using various criteria. As more enzymes have been isolated and characterized it has become clear that the classifications hitherto used are somewhat problematic. Many hydrolases have overlapping specificities and the use of poorly defined substrates further hinders proper classification. In addition the nomenclature of various amylases and cellulases is not consistent and causes confusion.

The traditional division of amylases and cellulases is based on their mode of action. Exoenzymes cleave sugar units from the ends of the polysaccharide chain, whereas endoenzymes make more random cuts along the chain. Exo- and endoenzymes act co-operatively, or synergistically, on crystalline substrates; a phenomenon that has been studied mostly on cellulases (Nidetzky *et al*., 1993; 1994a; Irwin *et al*., 1993; Béguin and Aubert, 1994). The current hypothesis to explain endo-exo synergism of cellulases is that endo-acting enzymes provide new chain ends for the exo-acting enzymes. Exo-exo synergism has also been demonstrated with some bacterial and fungal cellulases (Nidetzky *et al*., 1993; Irwin *et al*., 1993; Medve *et al*., 1994). This is postulated to be due to two types of exocellulases produced: one that attacks cellulose from the reducing end and another that attacks from the non-reducing end (Biely *et al*., 1993; Divne *et al.* 1994; Shen *et al*., 1995; Barr *et al*., 1996).

The complete hydrolysis of cellulose to glucose demands the action of exoglucanases (also called cellobiohydrolases), endoglucanases and β-glucosidases (Fig. 5). Cellobiohydrolases (1,4-β-D-glucan cellobio-hydrolase, EC 3.2.1.91) are usually active on crystalline cellulose and are lacking from incomplete cellulase systems. Endoglucanases (1,4-β-D- glucan-4-glucanohydrolase, EC 3.2.1.4) are more active against the amorphous regions of cellulose and they can also hydrolyze substituted celluloses, such as carboxymethylcellulose (CMC) and hydroxyethyl-cellulose (HEC). Cellobiohydrolases cleave disaccharide (cellobiose) units either from non-reducing or reducing ends, whereas endoglucanses hydrolyze the cellulose chains internally. β-Glucosidases (EC 3.2.1.21) are needed to cleave cellobiose and other soluble oligosaccharides to glucose (Béguin, 1990). This last step is important since cellobiose is usually an end-product inhibitor of cellulases.

The variety of amylolytic enzymes is larger than that of cellulases. Amylolytic enzymes can be divided into classes according to which linkages the enzymes are capable of degrading, α-1,4 or α-1,6. The enzymes hydrolyzing α-1,4 linkages of starch are α-amylases, β-amylases, glucoamylases, exo-acting α-amylases and cyclomaltodextrin glycosyl-

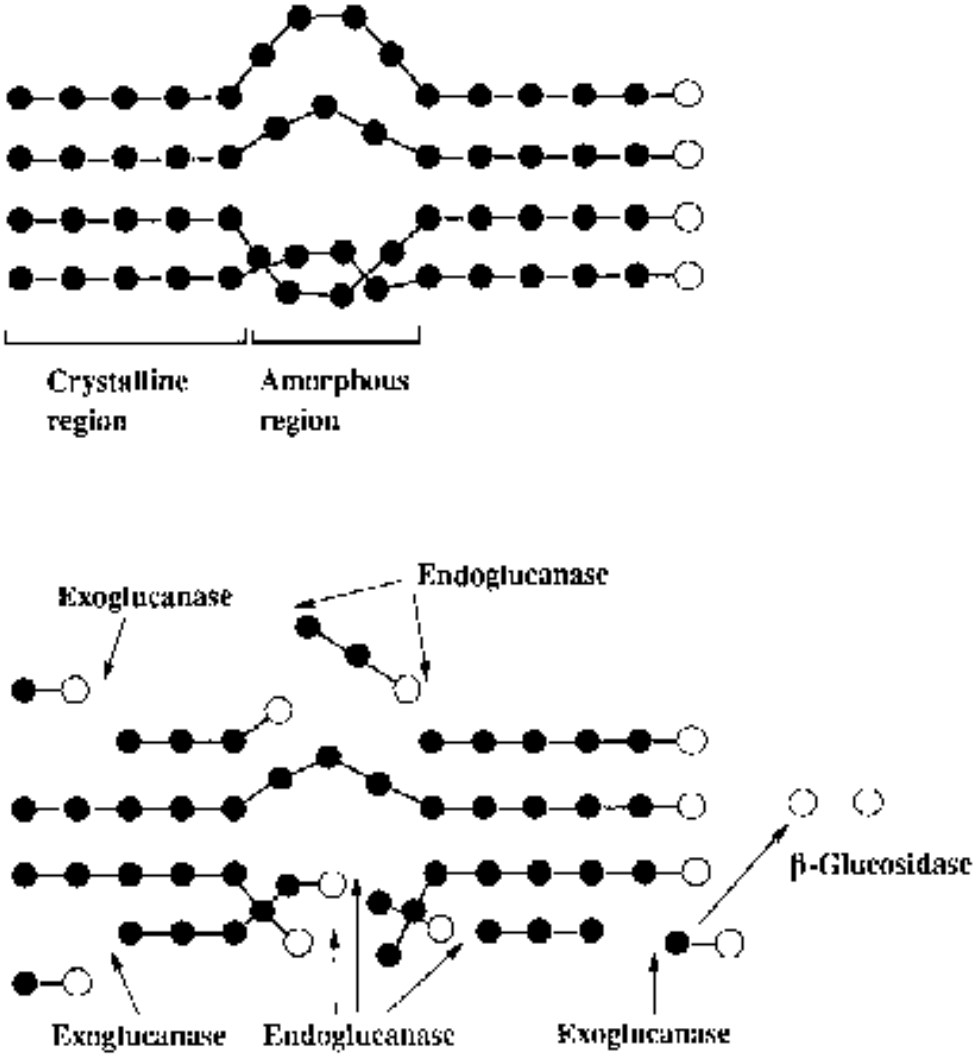

Figure 5. The complete enzymatic degradation of crystalline cellulose requires the action of various enzymes. Exoglucanases (also called cellobiohydrolases) degrade cellobiose units from either the non-reducing or reducing end of the microfibril. Endoglucanases attack the amorphous regions of cellulose and cleave the internal glycosidic linkages. β-glucosidases hydrolyze cellobiose and other small soluble oligosaccharides into glucose. Glucose units are indicated as spheres and reducing ends are shown in white.

transferases (Fig. 6). α-Glucosidases, which are also referred to as maltases, cleave α-1,4-linked glucose units from short oligosaccharides (Fig. 6). The naming of α-amylase and β-amylase is of historical origin andis based on the configuration of the products released. When the hydrolysis products formed are in α-D-configuration, the enzymes are called α-amylases (1,4-α-D-glucan glucanohydrolase, EC 3.2.1.1), whereas β-amylases (1,4-α-D-glucan maltohydrolase, EC 3.2.1.2) release products with β-D-configuration. α-Amylases are endoenzymes, whereas β-amylases attack glucans in an exo fashion from the non-reducing end. α-Amylases thus

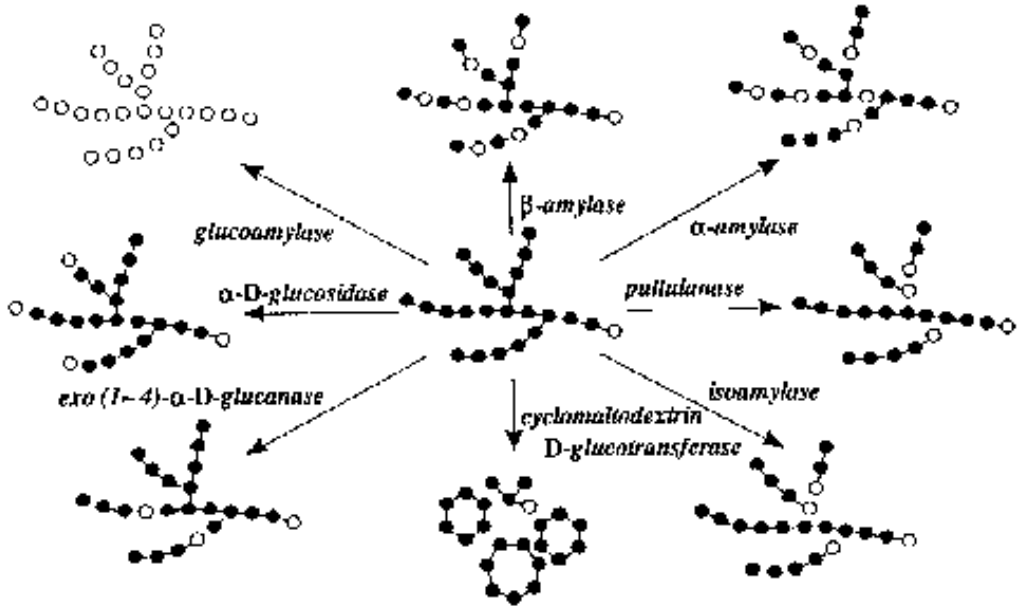

Figure 6. Modes of action of different amylases on starch. See text for details. Glucose units are indicated as spheres and reducing ends are shown in white. Modified from Kennedy et al., 1988.

resemble endocellulases, whereas β-amylases act like some of the exocellulases, e.g. *Trichoderma reesei* cellobiohydrolase II. Another exo-acting amylase is glucoamylase (1,4-α-D-glucan glucohydrolase, EC 3.2.1.3), which releases β-D-glucose from the non-reducing end of starch chains. In addition, α-amylases can sometimes cleave α-1,6 linkages with a low catalytic rate and they can bypass an α-D-1,6 branch linkage, which β-amylases cannot. Glucoamylase and several α-glucosidases also have an inherent α-D-1,6-activity.

Amylases which have primary affinity for α-1,6 linkages and catalyse the hydrolysis of α-1,6-glycosidic bonds in amylopectin and related polymers are called debranching enzymes. Pullulanases (α-dextrin 6-glucanohydrolase, EC 3.2.1.41) and isoamylases (glycogen 6-glucanohydrolase, EC 3.2.1.68) are the most widely studied debranching enzymes found mainly in bacteria (Vihinen and Mäntsälä, 1989).

1.3 MODULAR STRUCTURES OF CELLULASES AND AMYLASES

Amylolytic and cellulolytic enzymes are often multifunctional proteins composed of distinct domains, a catalytic domain and one or more domains involved in substrate binding or multi-enzyme complex formation. They may also have additional domains of unknown function and be composed of more than 10 different domains. The different domains are usually joined together by distinct linker peptides (Gilkes *et al.*, 1991). Most linkers are rich in glycine, proline,

serine and threonine residues, and are often O-glycosylated (Tomme *et al*., 1988b; Williamson *et al*., 1992).

The catalytic domains of all known glycosyl hydrolases, comprising about 500 sequences, are classified into 52 families on the basis of sequence alignment and a secondary structure prediction method called hydrophobic cluster analysis (HCA) (Henrissat, 1991; Henrissat and Bairoch, 1993 and unpublished; Tomme *et al*., 1995a; Shen *et al*., 1995). Different cellulases have been allocated to 11 families (families 5 - 10, 12, 26, 44, 45, 48) and amylases to 3 (families 13 - 15). A total of nine crystal structures of cellulase catalytic domains from six different families and 10 amylase structures from 3 different families have been published (Table 1). As the 3D structures of catalytic domains have become available it has become clear that although the amino acid sequence similarity is usually low, the fold is conserved within one family. The families do not follow the nomenclature of the enzymes (EC classes set by the International Union of Biochemistry), which is mainly based on substrate specificity. Thus, one family may contain both exo- and endoglucanases. Moreover, e.g. family 5 contains mostly cellulases but also a β-mannanase, an exo-1,3-β-glucanase and a cellodextrinase (Henrissat, 1994; Henrissat and Bairoch, 1993). All α-amylases and many other amylolytic enzymes belong to one superfamily (family 13), which contains different activities from over ten enzyme classes.

Some amylases and most cellulases hitherto studied have a second domain which binds to the substrate but does not hydrolyse it. The role of the cellulose-binding domain (CBD) of bacterial and fungal cellulases is well established (Teeri *et al*., 1995; Din *et al*., 1995). Different studies have shown that catalytic domains have a relatively weak affinity for crystalline substrate, whereas isolated CBDs carry most of the binding properties of cellulases. The proteolytic removal of CBD has no effect on the enzyme's capacity to hydrolyse small soluble substrates, whereas the hydrolysis of insoluble cellulose is clearly reduced (van Tilbeurgh *et al*., 1986; Ståhlberg, *et al*., 1991; Tomme *et al*., 1988b; Kleman-Leyer *et al*., 1994). Some bacterial CBDs are capable of disrupting cotton fibres and releasing small particles from the substrate, thus taking a more active role than just anchoring cellulase to their substrates (Din *et al*., 1991; Kilburn *et al*., 1993). CBDs are usually located either in the N- or the C-terminus of the protein, although internal functional CBDs also exist (Nevalainen and Penttilä, 1995; Din *et al*., 1995). The cellulose-binding domains have also been classified into families. At the moment ten different CBD families can be distinguished, of which the CBDs from families I and II are the most extensively studied (see also section 1.3.1. and 1.3.2.) (Tomme *et al*., 1995b). Sequence comparison of the domains identified as CBDs usually shows marked sequence similarity among enzymes from the same organism.

Various amylolytic enzymes, such as glucoamylases, β-amylases, cyclodextrin glycosyltransferases and some α-amylases are reported to have a C-terminal raw starch-binding domain (SBD). All SBDs hitherto studied are about 100 amino acids long, share some conserved amino acids and are

Table 1. Catalogue of amylases and cellulases for which X-ray crystallographic analysis has been reported and mechanisms suggested. The classification is according to the glycosyl hydrolase families (Henrissat and Bairoch, 1993).

Family	Enzyme	Organism	Catalytic domain fold	EC number	Mechanism[1]	Ref.
5	Endoglucanase A	*Clostridium cellulolyticum*	α/β barrel	3.2.1.4	retaining	1
5	Endoglucanase CelC	*Clostridium thermocellum*	α/β barrel	3.2.1.4	retaining	2
6	Cellobiohydrolase II	*Trichoderma reesei*	α/β barrel	3.2.1.91	inverting	3
	Endoglucanase E2	*Thermomonospora fusca*	α/β barrel	3.2.1.4	inverting	4
7	Cellobiohydrolase I	*Trichoderma reesei*	β sandwich	3.2.1.91	retaining	5
	Endoglucanase I	*Humicola insolens*	β sandwich	3.2.1.4	retaining	6
9	Endoglucanase D	*Clostridium thermocellum*	α/α barrel	3.2.1.4	inverting	7
10	Exoglucanase/ Xylanase Cex	*Cellulomonas fimi*	α/β barrel	3.2.1.91/ 3.2.1.8	retaining	8
13	α-amylase	*Aspergillus oryzae*	α/β barrel	3.2.1.1	retaining	9, 10
13	α-amylase	*Aspergllus niger*	α/β barrel	3.2.1.1	retaining	10
13	α-amylase	*Bacillus licheniformis*	α/β barrel	3.2.1.1	retaining	11
13	α-amylase	barley	α/β barrel	3.2.1.1	retaining	12
13	α-amylase	porcine pancreatic	α/β barrel	3.2.1.1	retaining	13
13	α-amylase	human pancreatic	α/β barrel	3.2.1.1	retaining	14
13	Cyclodextrin glycosyltransferase	*Bacillus circulans* (Strain 8)	α/β barrel	2.4.1.19	retaining	15
13	Cyclodextrin glycosyltransferase	*Bacillus circulans* (Strain 251)	α/β barrel	2.4.1.19	retaining	16
14	β-amylase	*Glycine max* (Soybean)	α/β barrel	3.2.1.2	inverting	17
14	β-amylase	Sweet potatoe	α/β barrel	3.2.1.2	inverting	18
15	Glucoamylase	*Aspergillus awamorii*	α/α barrel	3.2.1.3	inverting	19
45	Endoglucanase V	*Humicola insolens*	β-barrel	3.2.1.4	inverting	20

[1] The stereochemical mechanism has not necessarily been determined for all the listed enzymes but is assumed to follow the family division.

References: 1. Ducros *et al*., 1995; 2. Dominguez *et al*., 1996; 3. Rouvinen *et al*., 1990; 4. Spezio *et al*., 1993; 5. Divne *et al*., 1994; 6. Davies and Schülein 1995; 7. Juy *et al.*, 1992; 8. White *et al.*, 1994; 9. Matsuura *et al.*, 1984; 10. Boel *et al.*, 1990; 11. Machius *et al.*, 1995; 12. Kadziola *et al.*, 1994; 13. Buisson *et al.*, 1987; 14. Brayer *et al.*, 1995; 15. Klein and Schulz, 1991; 16. Lawson *et al.*, 1994; 17. Mikami *et al.*, 1993; 18. Cheong *et al.*, 1995; 19. Aleshin *et al.*, 1992; 20. Davies *et al.*, 1993.

predicted to be composed of β-strands (see also section 1.3.4) (Jespersen *et al*., 1991; Svensson *et al*., 1989). The binding domain has been shown to be important for adsorption and degradation of granular starch by glucoamylases (Takahashi *et al.* 1985). Since most α-amylases lack the binding domain but can still attack

granular starch, it is apparently not an absolute requirement for degradation of raw starch (Svensson *et al*., 1989).

One of the additional domains found in various bacterial carbohydrases, for example in four *Cellulomonas fimi* cellulases and in *Clostridium thermohydrosulfuricum* α-amylase-pullulanase, is a fibronectin type III (Fn3) domain approximately 100 amino acids long (Béguin and Aubert, 1994). Fn3 -like sequences occur in various extracellular matrix and adhesion proteins in higher eukaryotes, where they are involved in protein-protein interactions. It has been proposed that Fn3-type domains are capable of unfolding reversibly to provide elasticity in muscle fibres (Erickson, 1994). In carbohydrases Fn3 modules often occur between the catalytic and substrate binding domain, and it has been suggested that they could either serve as linkers between these domains or also provide more flexibility through unfolding (Din *et al*., 1995).

1.3.1 Fungal cellulases

Cellulolytic fungi, especially the aerobic soft-rot species (e.g *Trichoderma, Humicola* and *Penicillium* species) and white-rot fungi (e.g. *Phanerochaete*), are among the most studied cellulolytic micro-organisms. They are capable of producing all the activities needed for complete hydrolysis of native cellulose. They usually produce one or two exoglucanases, several endoglucanases and at least one β-glucosidase, all of which are secreted into the culture medium (Wood *et al*., 1989; Wood and Garcia-Campayo, 1990; Covert *et al*., 1992a; Schülein *et al*., 1993; Nevalainen and Penttilä, 1995). Such a diversity may be needed in part to cope with the physical heterogeneity of the substrate. In addition, the structure of cellulose changes during the degradation process and may require different enzymes at different stages. The exact number of enzymes reported may also vary because the purification of cellulases is difficult (Wood *et al*., 1989; Reinikainen *et al*., 1995a). Post-translational modifications such as glycosylation and proteolysis are known sources of cellulase diversity (Béguin, 1990; Srisodsuk *et al*., 1993), and protein-protein interactions are also reported to cause further problems in purification (Sprey and Bochem, 1993; Wood *et al*., 1989).

A number of fungal cellulase genes have been cloned (for a recent review see Nevalainen and Penttilä, 1995). This allows sequence comparisons to be carried out among fungal enzymes and comparisons to be made with other glycosyl hydrolases. The overall domain structure of fungal cellulases is generally the same for exoglucanases, or cellobiohydrolases (CBHs), and endoglucanases (EG). All these enzymes consist of a bigger catalytic domain and a small cellulose-binding domain (CBD) separated by an O-glycosylated linker peptide. The CBD is either N- or C-terminal. There are, however, reports of an exoglucanase (Covert *et al*., 1992b) and several endoglucanases (Ward *et al*., 1993; Schülein *et al*., 1993) which have only the catalytic domain. Some anaerobic ruminal fungi (e.g. *Neocallimastix* spp.) show more complex domain architecture and have e.g. multiple catalytic domains (see also section 1.3.5.). In addition, they contain non-catalytic domains, which apparently function as docking domains

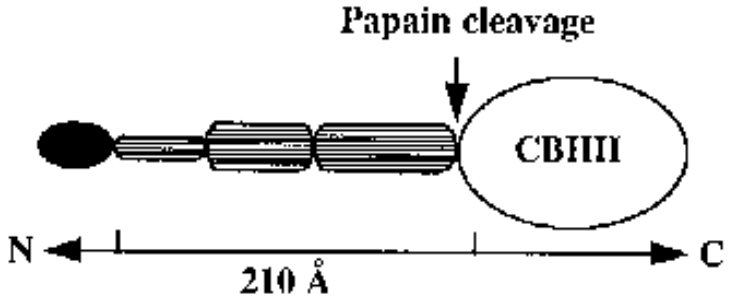

Figure 7. Domain structure of Trichoderma reesei cellobiohydrolase II (CBHII) based on small-angle X-ray scattering measurements. catalytic domain, linker peptide, cellulose-binding domain (CBD).

involved in multienzyme complex formation (Fanutti *et al*., 1995). Some cellulases also exhibit activity against other polymeric substrates; e.g. *Trichoderma reesei* endoglucanase I (EGI) degrades both xylan and cellulose (Biely *et al*., 1991).

The crystallization of entire fungal cellulases has proved to be rather difficult. This is probably due to the linker peptide, which might be flexible and/or heterogenously glycosylated. The two techniques that have been used to evaluate the spatial relationship between the individual domains of fungal cellulases are small-angle X-ray scattering (Abuja *et al*., 1988a, 1988b) and dynamic light scattering (Boisset *et al*., 1995). The results indicate that both *Trichoderma reesei* cellobiohydrolases and *Humicola insolens* endoglucanase have an elongated, tadpole-like shape (Fig. 7). The catalytic domain forms the ellipsoidal head part and the linker adopts extended conformation between the catalytic domain and the CBD. The role of the CBHI linker peptide has been studied by deletion analysis (Srisodsuk *et al*., 1993). Some spatial separation of the two domains is apparently needed for efficient function of CBHI on crystalline cellulose. The linker was suggested to have a proline-glycine rich sequence functioning as a hinge which is followed by a more rigid, extended O-glycosylated part of the linker sequence (Srisodsuk *et al*., 1993). The glycans are apparently not required for the direct enzyme-substrate interactions (Johansson *et al*., 1989).

The catalytic domains produced by one species show no obvious amino acid similarity to each other, with some exceptions such as *T. reesei* CBHI and EGI (Penttilä *et al*., 1986). Of the 11 cellulase families currently classified using the HCA method (Henrissat and Bairoch, 1993; Tomme *et al*., 1995a), fungal cellulases are found in six. The only true exoglucanase structures solved are the two cellobiohydrolases of the filamentous fungus *Trichoderma reesei* (Table 1). The 3D structure of the catalytic domain of *T.reesei* CBHII belonging to family 6

as the first cellulase crystal structure solved at the atomic level (Fig. 8) (Rouvinen *et al*., 1990). The polypeptide folds into an α/β barrel structure similar to triose phosphate isomerase (TIM), with the exception that it contains seven instead of eight β-strands (Fig. 8a). The active site was identified in a tunnel spanning through the whole domain and formed by stable surface loops. This immediately suggested an explanation of the exo-mode of action of CBHII: the cellulose chain could only enter the active site tunnel by one of its ends. On the other hand, the loops enclosing the tunnel were missing in the endoglucanase sequences from the same family. This implied that related endoglucanases should have similar overall folds but the active site should be more open allowing the random binding (and hydrolysis) of the cellulose chain (Rouvinen *et al*., 1990). The 3D structure of *Thermomonospora fusca* endogluganase, E2 (Fig. 8b) (Spezio *et al*., 1993), belonging to the same family 6 confirmed this hypothesis. Furthermore, *T. reesei* CBHI belonging to a different family (family 7) has a completely different fold, but the active site is similarly situated in a tunnel formed by stable loops (Divne *et al*., 1994). It therefore seems that the active site tunnel is a general topological feature of a true cellobiohydrolase. The known exception is the bifunctional exoglucanase/xylanase Cex from the bacterium *Cellulomonas*

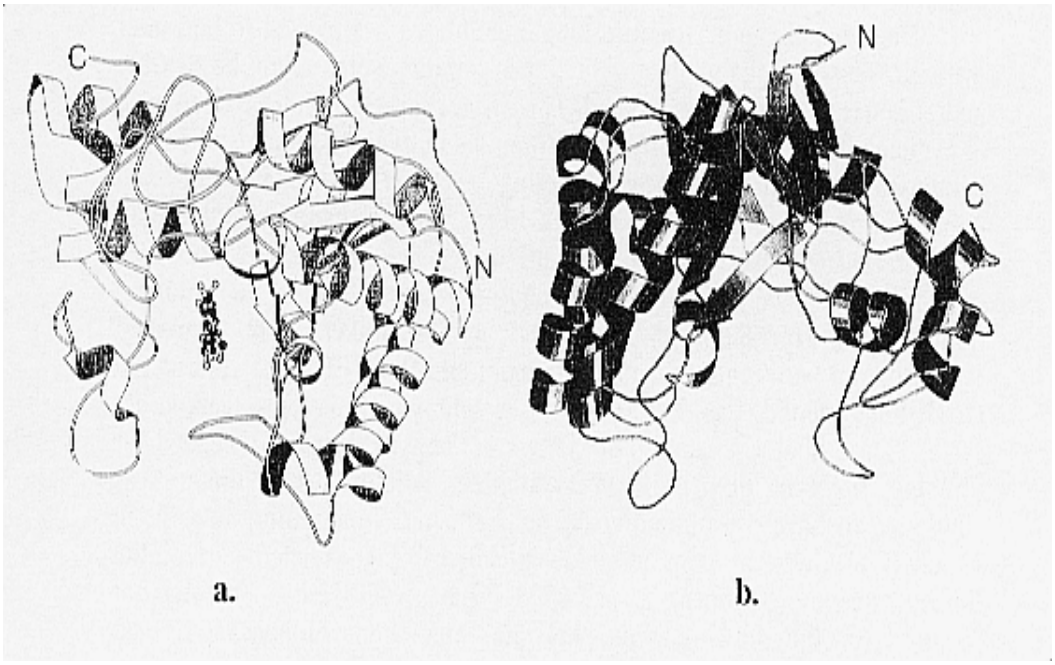

Figure 8. Ribbon presentation of two family 6 catalytic domains. Cellobiohydrolase II (CBHII) from Trichoderma reesei (a) and endoglucanase E2 from Thermomonospora fusca (b) have a TIM barrel central fold (barrel viewed from the side) in which long loops generate the active site at the C-terminal end of the barrel. In the case of CBHII the loops enclose a tunnel for substrate binding and catalysis. In E2 the loops are shorter and consequently a more open active site is generated. CBHII active site contains a bound ligand in the picture. The figure (a) was kindly provided by Alwyn Jones and figure (b) is modified from Wilson et al., 1995.

fimi (Table 1) (see section 1.3.2), which has an open active site architecture (White *et al*., 1994).

The fungal CBDs consist of about 36 amino acid residues and show high sequence identity, which strongly suggests that they have a similar fold (Hoffrén *et al*., 1995). They all belong to the CBD family I (Tomme *et al*., 1995b). The 3D structure of *T.reesei* CBHI CBD has been solved by NMR, revealing a wedge-shaped molecule composed of three antiparallel β-strands and stabilized by two disulfide bridges (Fig. 10a). One side of the wedge is flat, containing three highly conserved aromatic residues (Kraulis *et al*., 1989). Site-directed mutagenesis studies and studies of synthetic CBD peptides have shown that the three aromatic ring structures (either tyrosine ot tryptophan residues) constitute the primary interaction surface with cellulose (Reinikainen *et al*., 1992; 1995b; Linder *et al*., 1995a; 1995b). Residues capable of forming hydrogen bonds on this same surface seem also to be important, whereas direct interaction of the other surface is unlikely (Linder *et al*., 1995a).

1.3.2 Bacterial cellulases

Bacterial cellulases show more diversity in their modular architecture than fungal cellulases. One of the best characterized and most complex multienzyme complexes is produced by the anaerobic bacterium *Clostridium thermocellum*. The cellulase complexes, called cellulosomes,

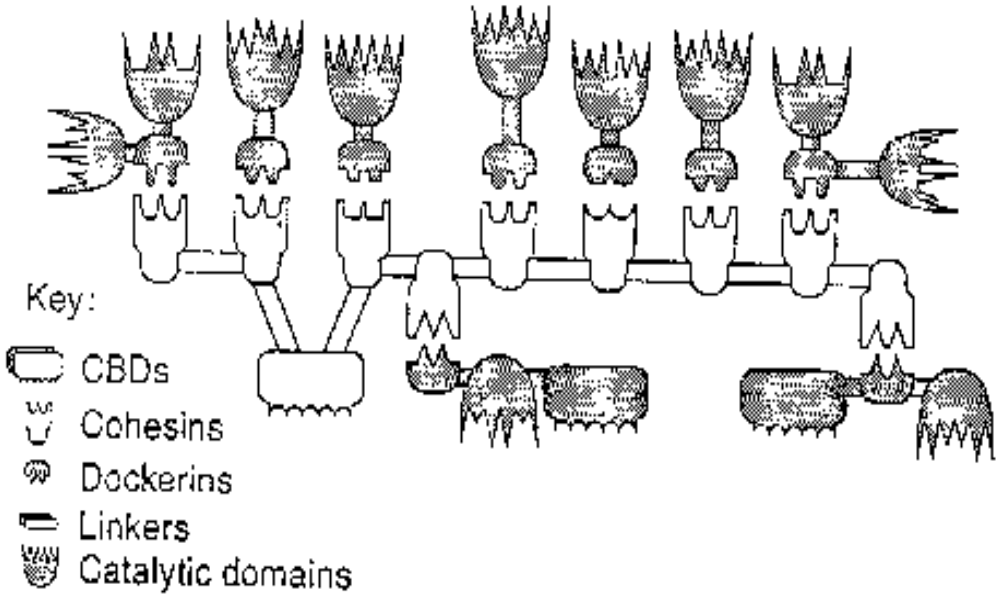

Figure 9. Components of the cellulosome in a simplified model based on the complex from Clostridium thermocellum. All subunits are composed of multiple domains. The scaffoldin (shown as the central structure in white) and some catalytic domains contain other domains (not shown) whose function is still not known. From Bayer et al., 1995.

are very active against crystalline cellulose. The molecular mass of a cellulosome exceeds 2,000 kDa, corresponding to about 20 polypeptide subunits (Béguin and Aubert, 1994). Cellulosomes can pack into polycellulosomal clusters forming protuberance-like organelles on the cell surface (Bayer *et al*., 1994). The largest subunit is a non-catalytic domain called scaffoldin (or S1 or CipA), which forms the core of the cellulosome structure (Fig. 9). It is composed of a single CBD and nine distinct, but closely related domains, cohesins, which bind the catalytic subunits. Catalytic domains comprise the majority of the subunits and exhibit either cellulase or xylanase activity. Catalytic domains contain conserved, duplicated segments called dockerins, which mediate the attachment of the catalytic domains to scaffoldin. The dockerin domains are usually interconnected by linking segments, or linkers, resembling other linkers found in cellulases. Both scaffoldin and some catalytic domains contain other domains, the function of which is currently not known (Bayer *et al*., 1994; 1995).

Cellulomonas fimi cellulases are an example of bacterial cellulases secreted into the medium. Hitherto four endoglucanases (CenA, CenB, CenC and CenD) and two cellobiohydrolases (CbhA and CbhB) have been characterized (for a recent review see Din *et al*., 1995). In addition, one of the two xylanases produced by *C. fimi* is a bifunctioanl hydrolase (Cex) capable of degrading cellulose although being predominantly active on xylan (White *et al*., 1994). The catalytic domain of Cex belongs to family 10, in which all the other members hitherto described are xylanases (Henrissat and Bairoch, 1993). The cellulolytic system of *C. fimi* resembles the corresponding enzyme system of *T. reesei*. All *C.fimi* cellulases are modular proteins comprising of two to six modules or domains, but all sharing two domains in common: a catalytic domain and a cellulose-binding domain (CBD) which can function independently. In four of the enzymes, CenB, CenD, CbhA and CbhB, fibronectin type III (Fn3) repeats separate the N-terminal catalytic domain from the C-terminal CBD. The catalytic domains belong to six different families, whereas all the CBDs except that of CenC belong to family II (Tomme *et al*., 1995b). CenC has tandem CBDs from family IV; CenB has a second, internal CBD from family III.

The *C. fimi* CBDs from family II are the most extensively studied bacterial CBDs. They usually adsorb to both crystalline and amorphous cellulose, and one CBD has also been reported to bind to chitin (Ong *et al*., 1993). Adsorption has been shown to be strongly dependent on aromatic amino acid residues (Din *et al*., 1994). The 3D structure of *C. fimi* Cex CBD belonging to CBD family II was recently solved, revealing a β-barrel fold consisting of 9 antiparallel β-strands and having two faces (Xu *et al*., 1995). One face of the domain contains three tryptophan residues and two of them were shown to interact with cellohexaose when analysed with NMR (Fig. 10b).

1.3.3 α-Amylases and cyclodextrin glycosyltransferases

α-Amylases are probably the most widely distributed of the amylolytic enzymes. They are endo-acting enzymes catalyzing the hydrolysis of α-1,4-

glycosidic linkages of amylose and amylopectin. They liberate glucose and malto-oligosaccharides in the α-anomeric form and rapidly reduce the molecular size of starch. In general, α-amylases are the main enzymes

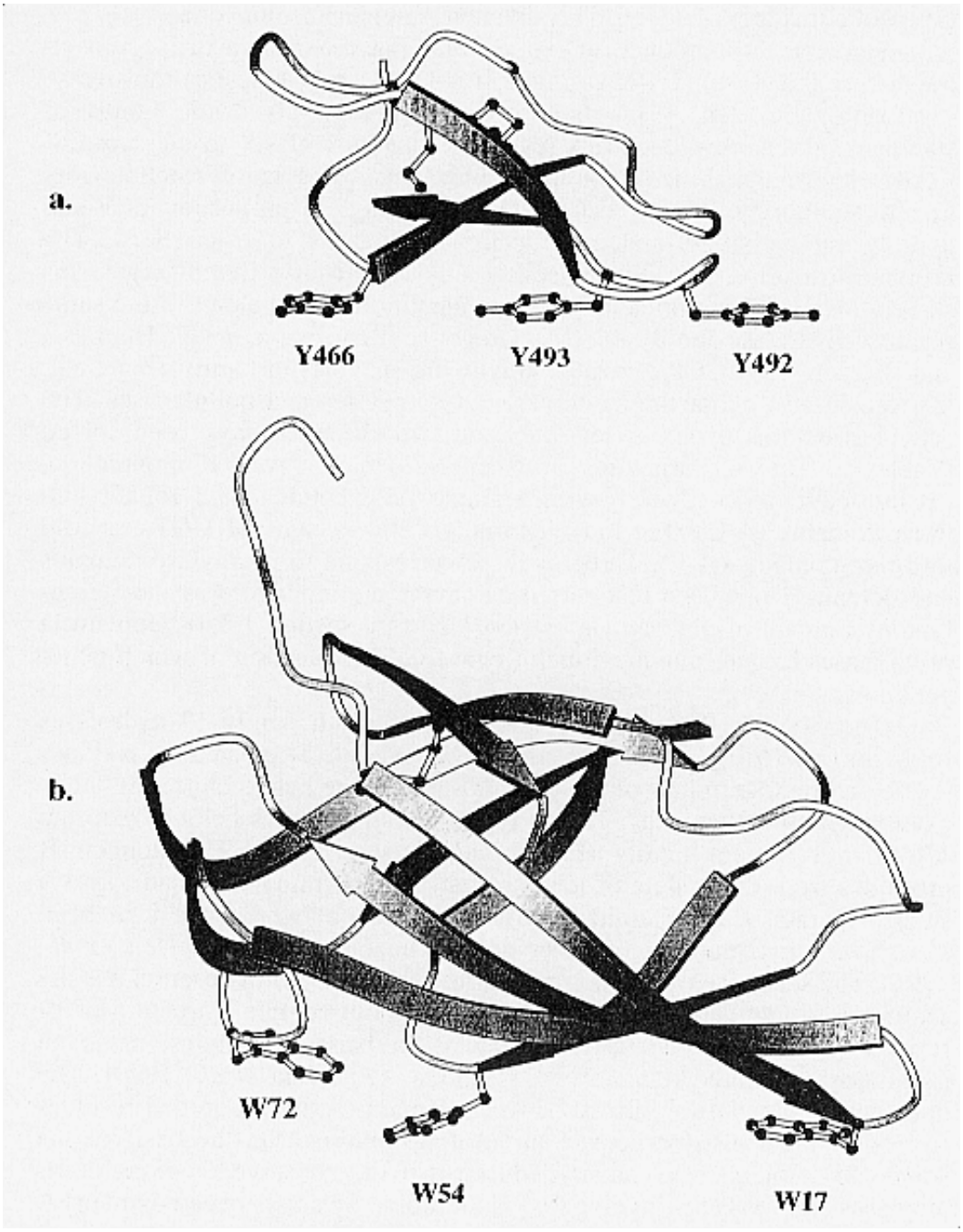

Figure 10. Ribbon presentation of the structures of cellulose-binding domains (CBDs) from T. reesei CBHI and C. fimi Cex, both solved by NMR. The CBD of T. reesei CBHI (a) belongs to CBD family I, whereas the CBD of C. fimi Cex (b) belongs to family II. The side-chains of three tyrosines (Y466, Y493 and Y492) in CBHI CBD and three trytoptophans (W72, W54 and W17) in Cex CBD, implicated in binding to crystalline cellulose, are shown. The conserved disulfide bridges are also shown. From Tomme et al., 1995b.

responsible for degradation of the starch granule. α-Amylases are often divided into two categories according to the degree of hydrolysis of the substrate (Fukumoto and Okada, 1963). Saccharifying α-amylases hydrolyze 50 to 60 % and liquefying α-amylases about 30 to 40 % of the glycosidic linkages of starch. This division is not an absolute one.

Some bacteria produce an extracellular enzyme cyclodextrin glycosyl-transferase (CGTase) (1,4-α-D-glucan 4-α-D-(1,4-α-D-glucano)-transferase (cyclizing), EC 2.4.1.19), which hydrolyses α-1,4-glycosidic bonds of starch to produce cyclodextrins (cyclic compounds of six to eight α-1,4-linked glucose units) via an intramolecular transglycosylation reaction (Fig. 6). In addition, CGTases can display coupling, disproportionation and hydrolysing activities, and are functionally related to α-amylases. The primary structures of CGTases and α-amylases from different organisms share about 30 % amino acid sequence identity and all belong to the same glycosyl hydrolase family 13 (MacGregor and Svensson, 1989; Henrissat and Bairoch, 1993). Other related amylolytic enzymes in family 13 are e.g. α-glucosidases, exo-acting α-amylases, isoamylases and pullulanases. The crystal structures of six α-amylases and two GCTases have been solved (Table 1). The 3D structures of α-amylases have revealed monomeric, calcium-containig enzymes, with a single polypeptide chain folded into three domains (A-C) (Fig.11a). The polypeptide chains of CGTases fold into five domains (A-E). Three of these correspond to α-amylase domains and domain E has been recognised as having a similar fold as the starch-binding domain of glucoamylases (Fig.11b) (see section 1.3.4). Domain D of CGTases has an immunoglobulin-type fold, the function of which is not yet known.

The catalytic domain A of α-amylases and all family 13 hydrolases folds into an $(\alpha/\beta)_8$ -barrel protein. The active site is created by residues located at the C-terminus of the β strands and in the helix-connecting loops extending from these ends. This is typical to all enzymes belonging to the α/β -barrel protein family (Farber and Petsko, 1990). The domain B protrudes from the middle of the barrel at the C-terminal side and forms a small separate, Ca^{2+} -stabilized structural domain, which has also been recognized in a number of other amylolytic enzymes (Jespersen *et al*., 1991). The sequence of this domain varies most; in *Bacillus* α-amylases the sequence is relatively long and folds into a more complex structure of β-strands (Machius *et al*., 1995), whereas in barley α-amylase it is an irregularly structured domain of 64 residues (Kadziola *et al*., 1994). The open active site cleft is formed between domains A and B, so that residues from domain B also participate in substrate binding. On the basis of the solved 3D structures it can be deduced that the conserved regions of α-amylases and CGTases involve the calcium-binding site between domains A and B, and regions belonging to the active site (MacGregor and Svensson, 1989; Klein and Schultz, 1991; Lawson *et al*., 1994).

Both α-amylases and CGTases have a domain C which is relatively conserved and folds into an antiparallel β-barrel. The orientation of domain C relative to domain A varies depending on the type and source of amylase (Brayer *et al*., 1995). The function of this domain is not known. Two structural studies, one from *B. circulans* CGTase (Lawson *et al*., 1994) and the other from pig pancreatic α-amylase (Qian *et al*., 1995), both clearly show, based on electron densities, that maltose units stack against the aromatic side-chain of a tryptophan situated on the surface of domain C (Fig. 11b). The truncation of domain C more

than 30 to 40 amino acids from the C-terminus was shown to be deleterious to the proper folding of the *B. stearothermophilus* α-amylases (Vihinen *et al*., 1994).

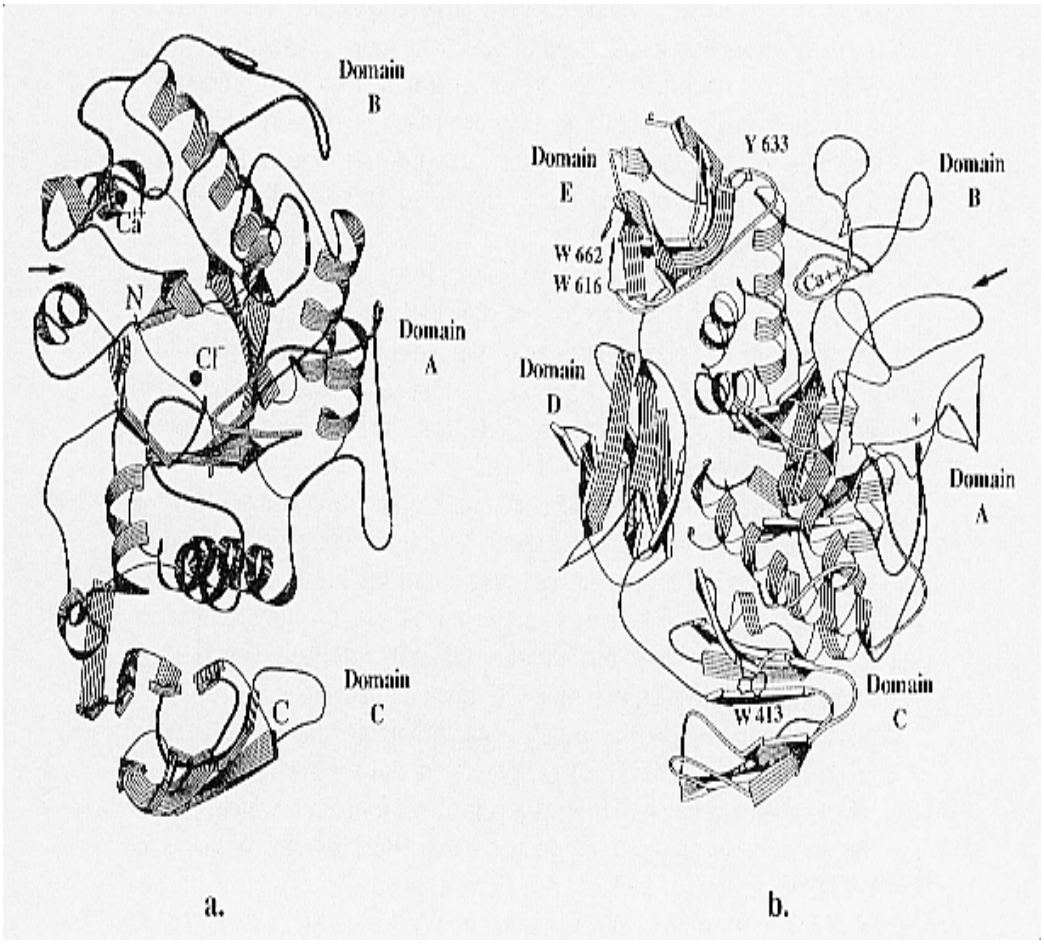

Figure 11. Ribbon presentation of human pancreatic α-amylase and Bacillus circulans CGTase, also showing the domain structure. α-Amylase (a) is composed of three domains (A-C) and CGTase (b) of five domains, of which the first three correspond to the α-amylase domains. Domain A is the catalytic domain having a typical TIM barrel fold. Both enzymes contain a conserved calcium-binding site between domains A and B. Human pancreatic α-amylase contains a chloride binding site in domain A, whereas CGTase contains a second calcium-binding site in domain A (marked with +). Side-chains of W616 and W662 in domain E of CGTase are involved in maltose binding site 1 and Y663 in domain E in maltose binding site 2 of the putative starch-binding domain. A third maltose binding site is found in domain C and involves W413. The active site cleft in domain A is shown with an arrow. From Brayer et al., 1995 (a) and Lawson et al., 1994 (b).

1.3.4 β-Amylases and glucoamylases

β-Amylases cleave β-maltose from the non-reducing ends of α-1,4-linked starch and malto-oligosaccharides in an exo-fashion. The enzyme is produced by higher plants and some micro-organisms. β-Amylases are usually tetrameric proteins composed of identical subunits and belong to family 14 of the glycosyl hydrolases (Henrissat and Bairoch, 1993). The 3D structures of two plant β-amylases have been published (Table 1). The monomeric enzymes are composed of a large core which folds into an $(\alpha/\beta)_8$ -barrel, a smaller lobe formed of three long loops extending from the C-terminal end of the β-barrel, and a long C-terminal loop of 50 - 60 amino acids. The active site pocket in the protein surface involves residues from both the lobe and core parts (Mikami *et al*., 1993; 1994). One of the loops in the lobe region is flexible and moves during substrate binding (see also section 1.4.2) (Mikami *et al*., 1994).

Glucoamylases are produced by bacteria, fungi and yeast. They release β-D-glucose from non-reducing ends of starch and related saccharides. The catalytic domains of glucoamylases belong to family 15, in which the fold is an α/α-barrel composed of 12 α-helices (Aleshin *et al*., 1992) and related to the twisted α-barrel of endoglucanase CelD from the bacterium *Clostridium thermocellum* (Juy *et al*., 1992). Fungal glucoamylases have a similar domain structure to that of fungal cellulases: they are composed of a largecatalytic domain and a smaller starch-binding domain (SBD) joined by an O-glycosylated linker peptide. The function of the linker peptide seems to be similar to that of fungal cellulase linkers, i.e. to provide distance between the catalytic and binding domain, but not itself significantly increasing the binding affinity for the substrate (Williamson *et al*., 1992).

Binding studies on the isolated starch-binding domain (SBD) of *Aspergillus niger* glucoamylase have suggested that for maximum affinity the substrate should be able to form a helical conformation which mimics the conformation of amylose in granular starch (Belshaw and Williamson, 1993). Thermodynamic studies have demonstrated the existence of two ligand binding sites of identical nature in the SBD (Sigurskjold *et al*., 1994b). In addition, two tryptophans have been shown to be important for the binding of starch (Svensson *et al*., 1986; Goto *et al*., 1994). The secondary structure and topology of an SBD from *A. niger* glucoamylase has recently been determined by NMR (Jacks *et al*., 1995). The 108 amino acid long polypeptide folds into two β-sheets composed of eight antiparallel β-strands. The SBD of *A. niger* glucoamylase shares about 37 % amino acid identity to the domain E of cyclodextrin glycosyltransferases (CGTases) (see section 1.3.3) (Svensson *et al*., 1989). The domain E of CGTases contains eight β-strands. The complex structure of a *Bacillus circulans* CGTase revealed two maltose-binding sites on domain E involving stacking interactions with two tryptophan and one tyrosine residues, respectively (Fig. 11b). In addition, both binding interactions involve several hydrogen bonds. The two tryptophans and some of the residues involved in hydrogen bond formation are strictly conserved in all starch-binding domains (Lawson *et al*., 1994; Svensson *et al*., 1989).

1.3.5 Module shuffling

The multiple domains of amylases and cellulases usually function independently. It is likely that during evolution the early genes were transferred as modules encoding single domains, which were subsequently recombined to yield genes coding for multidomain proteins (Béguin and Aubert, 1994). As has been discussed above various combinations of catalytic domains and substrate-binding domains are possible within amylase or cellulase families. In addition, some cellulases contain multiple cellulose-binding domains, e.g. *C. fimi* CenB and CenC (see section 1.3.2). The rumenal anaerobic bacterium *Caldocellum saccharolyticum* produces several hydrolases with multiple catalytic domains and the association usually occurs through one or several CBDs (Saul *et al*., 1990; Gibbs *et al*., 1992; Te'o *et al*., 1995). In addition to amylases and cellulases, chitinases, hemicellulases and lysozymes also hydrolyse insoluble polysaccharides. Many of these enzymes also contain separate domains (Gilkes *et al*., 1991). Since the substrates coexist in nature (e.g. both cellulose and hemicellulose are found in plant cell walls), it is perhaps not surprising that the degrading enzymes sometimes have mixed substrate binding or catalytic specificities due to domain shuffling. There are many bacterial and fungal hemicellulases, e.g. xylanases, mannanases, acetyl esterases and an arabinofuranoside that have been reported to have a cellulose-binding domain. A *Streptomyces* chitinase involved in the degradation of insoluble chitin was also found to contain a CBD (Béguin and Aubert, 1994; Henrissat, 1994; Ståhlbrandt *et al*., 1995; Margolles-Clark *et al*., 1996). Furthermore, fibronectin type III (Fn3) modules are encountered in many carbohydrases, such as cellulases, amylases and chitinases (Béguin and Aubert, 1994). As discussed earlier, the function of this domain is not known but it is proposed to serve as a linker between domains.

The modular structure of various carbohydrases can be utilised in different biotechnical applications. In particular the properties of the cellulose-binding domains have already been used successfully. The fusion is made at the DNA level to a protein of interest. The produced fusion protein bears a binding domain which can be used for purification and immobilization on cheap cellulose-based column matrices. Different types of proteins, such as an alkaline phosphatase, a β-glucosidase, a human interleukin and an antibody fragment have been fused to CBDs and shown to retain their activity and bind to cellulose (Greenwood *et al*., 1992; Ong *et al*., 1991, 1995; Reinikainen *et al*., 1996). A biosensor based on glucose oxidase - CBD conjugate has also been constructed (Kilburn *et al*., 1993) and shown to be regenerable for the detection of glucose.

1.4 CATALYTIC MECHANISMS AND SUBSTRATE BINDING OF GLYCOSYL HYDROLASES

1.4.1 Breaking down the O-glycosidic bond

Cellulases and amylases, like all glycosidases, catalyze the transfer of a glycosyl group to water utilizing general acid catalysis. The enzymatic break-down of a glycosidic bond is also a stereoselective process, in which the configuration about the anomeric centre (C1 carbon) can either be inverted or retained. Different mechanisms are required for such different stereochemical outcomes, as suggested over 40 years ago by Koshland (1953) (Fig. 12). Both mechanisms contain a pair of carboxylic acid residues suitably disposed on either side of the bond to be cleaved. The roles of these two residues are different. Inverting enzymes use a single-displacement mechanism involving a general acid and a base. In this mechanism, a general acid (AH) donates a proton to the leaving glycosidic oxygen and a general base (B-) assists the nucleophilic attack of water (Fig. 12). The mechanism for retaining glycosidases, by contrast, involves a double-displacement reaction in which a covalent glycosyl-enzyme intermediate has been shown to be formed with many but not all retaining enzymes. In the first step the deprotonated carboxylate acting as a nucleophile (Nu) attacks at the anomeric centre and displaces the glycosidic oxygen. This process is assisted by a general acid catalytic residue (AH) which is deprotonated (A-). In the second step water attacks at the anomeric centre with general base-catalytic assistance (A-) displacing the nucleophile (Nu) and releasing the product sugar (Fig. 12) (Koshland, 1953; Sinnott, 1990; Withers and Aebersold, 1995).

The stereochemical course of the hydrolysis is usually determined by proton NMR, in which the α- and β-anomeric protons give different chemical shifts (Withers *et al*., 1986). In all cases studied the same enzymatic mechanism (inverting or retaining) has been conserved within a glycosyl hydrolase family (Gebler *et al*., 1992). Thus both the fold of the protein and the stereochemistry of the hydrolysis reaction appear to be conserved within a given family. Retaining enzymes are also able to bring about transglycosylation, in which in the second step of the reaction (Fig. 12) another oligosaccharide instead of a water molecule attacks at the anomeric centre, leading to elongation of the saccharide chain (Sinnott, 1990). Transglycosylation has been postulated to play a role in the induction of cellulase promoters (Biely *et al*., 1991; Fowler *et al*., 1993; Nevalainen and Penttilä, 1995) and it can be applied in the enzymatic synthesis of oligosaccharides (Gusakov *et al*., 1991; Ogawa *et al*., 1990). Inverting amylases such as glucoamylases have been shown to be able to reverse the hydrolysis by condensation reaction (Sinnott, 1990; Nikolov *et al*., 1989). This ability has also been used in the synthesis of some gluco-oligosaccharides (Nakano *et al*., 1995).

Both of the reaction mechanisms involve oxocarbonium ion-like transition states (McCarter and Withers, 1994; Tanaka *et al*., 1994). With some retaining enzymes, most notably with the hen egg-white lysozyme

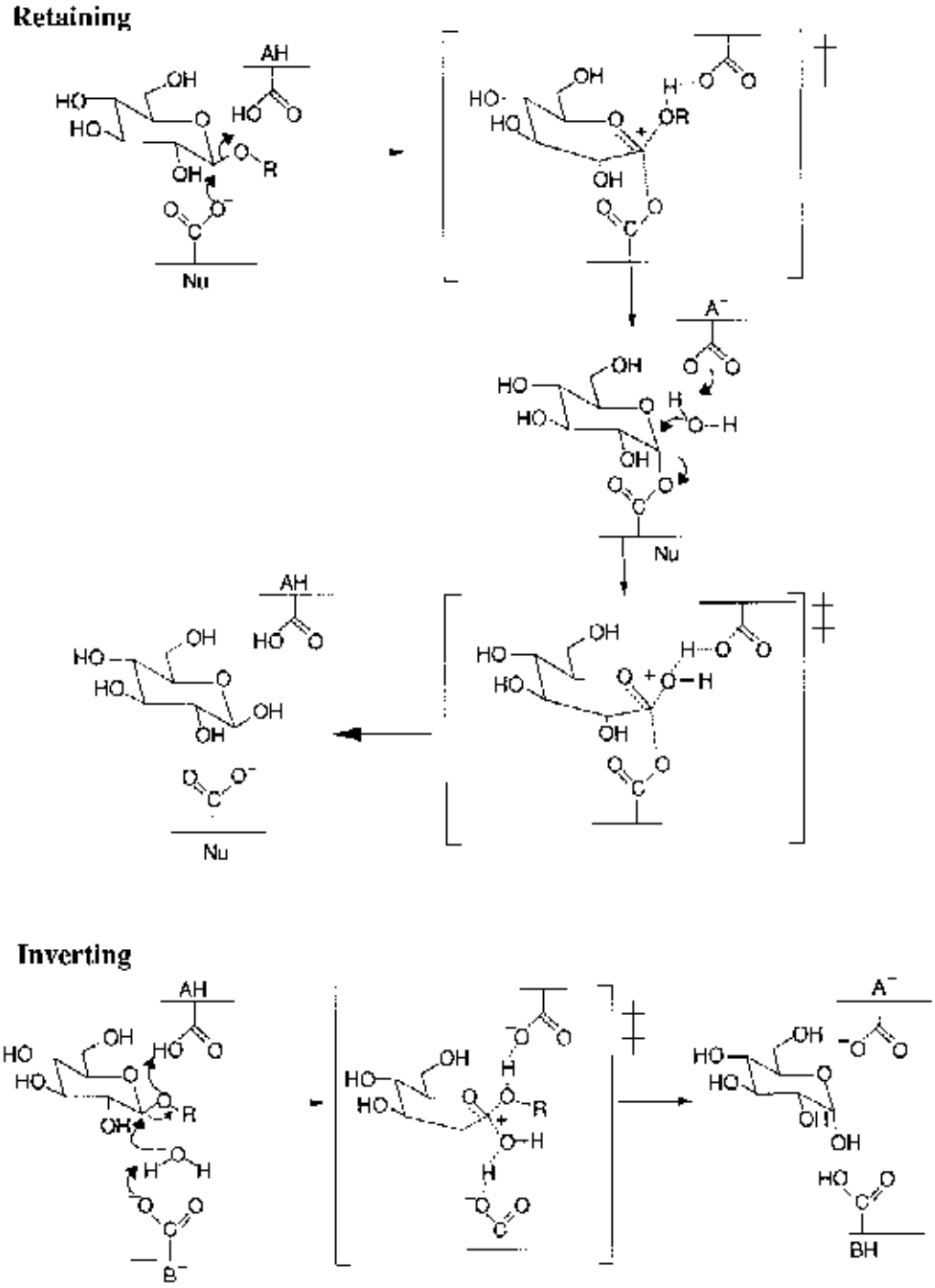

Figure 12. The two different reaction mechanisms of glycosyl hydrolases, shown here with β-1,4-glycanases. See text for details. Modified from Withers and Aebersold, 1995.

(HEWL), attempts to prove the existence of a covalent glycosyl-enzyme intermediate have failed (Sinnott, 1990). Another type of mechanism has been suggested in which the positively charged oxocarbonium is stabilized by the catalytic nucleophile through electrostatic interactions (Sinnott, 1990). In addition, ring distortion at the subsite preceding the scissile bond has been suggested as an element of the catalytic machinery for both inverting and retaining glycosyl hydrolases (Harris *et al*, 1993; Strynadka and James, 1991; Kuroki *et al.*, 1993; Varghese *et al.*, 1992). The identification of active site carboxylic acid residues usually occurs through X-ray crystallographic determination of the 3D structure of protein. Chemical modification and/or site-directed mutagenesis is then used to identify and assign specific functional roles to residues involved in substrate binding and catalysis. In the case of retaining enzymes, the active site carboxylic acids can be identified without the protein structure by the use of affinity labels or mechanism-based inactivators to specifically derivatise the key amino acids. 2-Deoxy-2-fluoro-derivatives of glucose and cellobiose have successfully been used as mechanism-based inhibitors which covalently bind the nucleophile in the retaining enzymes (McCarter and Withers, 1994; Withers and Aebersold, 1995).

A mutagenesis study of an inverting T4 lysozyme showed that the reaction mechanism can also be altered (Kuroki *et al.*, 1995). A single substitution, T26E, in the active site cleft produced an enzyme that cleaved the bond but bound the product covalently to the enzyme (Kuroki *et al.*, 1993). The mutation T26H produced an enzyme which was catalytically active but which retained the configuration at the anomeric centre, contrary to the wt enzyme (Kuroki *et al.*, 1995). The crystal structure of the T26H mutant revealed that histidine-26 acts as a nucleophile in the hydrolysis. The authors suggested a general approach to engineer the mechanism of glycosyl hydrolases. A mutation providing an imidazole or a carboxylate next to the base catalyst in an inverting enzyme (Fig. 12) should block the attack of the solvent and function as a nucleophile, thus retaining the mechanism. The nucleophilic group in retaining enzymes (Fig. 12) should be substituted with a smaller amino acid, which would leave more space for the water to attack and lead to inversion of the anomeric configuration (Kuroki *et al.*, 1995). A related experiment has been described for a retaining β-glucosidase in which mutation E358A in the presence of a small nucleophilic anion led to an inverting mechanism (Withers, 1995).

1.4.2 The active site

Of the 52 glycosyl hydrolase families hitherto classified, the 3D structures are known for 22 families (for review see Davies and Henrissat, 1995). The TIM barrel type of fold is the most common and is found in nine different families. Davies and Henrissat (1995) compared all the solved 3D structures and concluded that the active site topologies can be divided into only three classes regardless of the stereochemistry or the overall fold of the enzyme (Fig. 13). Pocket- or crater-like active sites are encountered in enzymes hydrolysing monosaccharide units from chain ends, such as β-glucosidases, glucoamylases and β-amylases (Fig. 13a). The two latter exoenzymes, which cleave starch granules, have adapted to a substrate with a large number of non-reducing ends exposed at the surface (see section 1.1.2). This kind of active site is not optimal for a substrate like native crystalline cellulose, which has very few chain ends available (see section 1.1.1).

The second class of topologies includes active sites situated in a tunnel which allows the enzyme to release products from chain ends while remaining bound to the polysaccharide chain (Fig. 13b). The only examples existing at the moment are the two *T. reesei* cellobiohydrolases (CBHI and CBHII), which are active especially on crystalline cellulose. It is currently under discussion whether the loops closing the tunnel can open occasionally to allow a random binding or whether the cellulose chain penetrates the tunnel only from one entrance. The third class of topologies is a derivative from the previous one. Here the loops are shorter or turned away from the tunnel, making the active site into a cleft or groove open from both ends (Fig. 13c). The open active site structure allows random binding along the polysaccharide chain and is commonly found in endo-acting enzymes, such as α-amylases, endocellulases, xylanases and lysozymes.

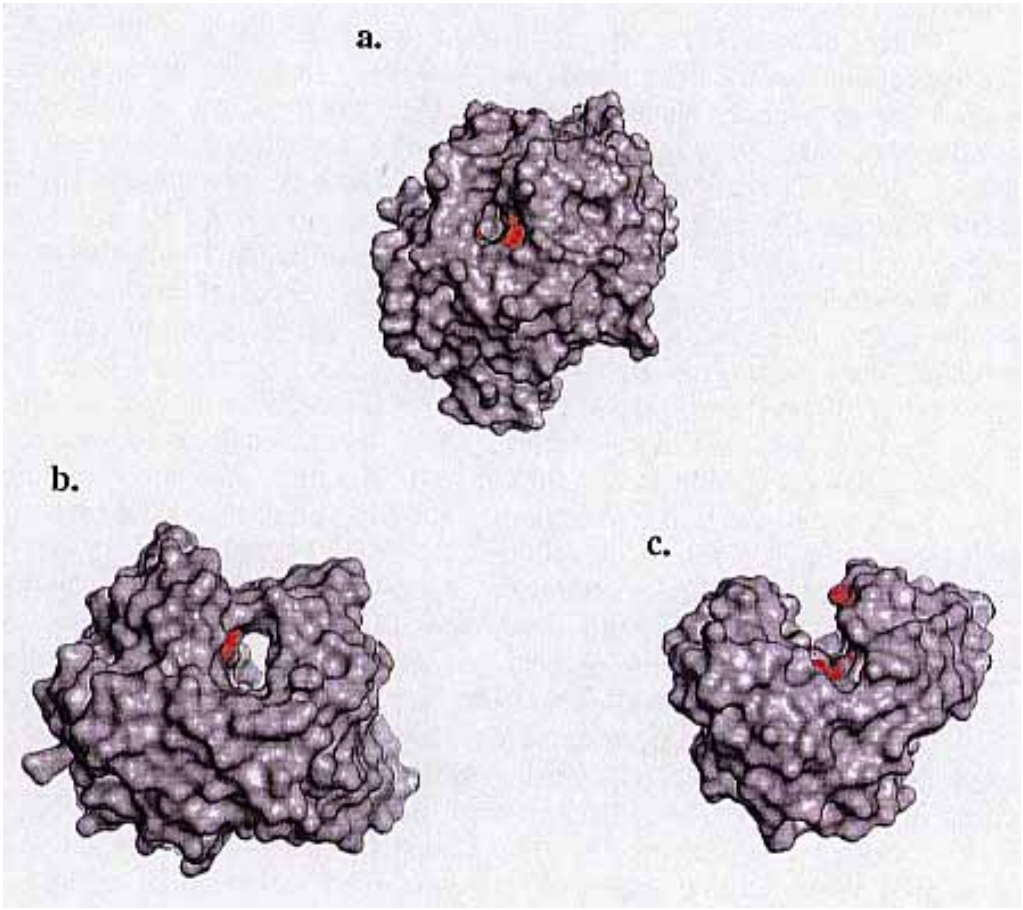

Figure 13. The three active site topologies of glycosyl hydrolases described by Davies and Henrissat (1995). The pocket of A. awamori glucoamylase (a). The tunnel of T. reesei CBHII (b). The cleft of T. fusca E2 (c). E2 belongs to the same family 6 as CBHII. The proposed catalytic amino acids are shaded in red. From Davies and Henrissat (1995).

Structural studies have confirmed that the active sites of glycosyl hydrolases are composed of multiple binding sites, or subsites, for the sugar units of polymeric substrates. The substrate binding sites are commonly lined with aromatic residues, especially tryptophans and tyrosines, which make hydrophobic stacking interactions with the sugar rings. In addition, the active sites contain many residues which form hydrogen bonds to the substrate either directly or via water molecules (Rouvinen *et al*., 1990; Svensson and Søgaard, 1993; Mikami *et al*., 1993; Aleshin *et al*., 1994; Divne *et al*., 1994). The hydrogen bonds are apparently responsible for the specificity of the binding and also dictate the directionality (reducing vs. non-reducing end) of the binding (Fersht *et al*., 1985; Davies *et al*., 1995). The H-bonding networks and planar stacking of aromatic ring structures with sugar rings both seem to be a common theme in many carbohydrate degrading enzymes (e.g. in cellulases and amylases) or carbohydrate binding proteins (e.g. in many lectins and periplasmic binding proteins) (Bourne *et al*., 1993; Quiocho 1993; Toone, 1994). Site-directed mutagenesis studies also imply the importance of these interactions in carbohydrate-degrading enzymes (Bourne *et al*., 1993; Svensson and Søgaard, 1993; Svensson, 1994).

Many amino acid residues involved in substrate binding also appear to be important for the hydrolysis. A conserved loop region containing a tryptophan residue in glucoamylase is suggested to play a critical role in ground-state and in transition state enzyme complexes (Natarajan and Sierks, 1996). The scissile bond in glucoamylases is between the first and second subsite whereas the tryptophan is stacked up against the third sugar ring. The glucoamylase complex structure (Aleshin *et al*., 1994) also reveals that this stacking interaction may exert strain on the glycosidic linkage between the second and third glucose units. Such a strain could be transferred to the scissile glycosidic linkage or it could be used to promote the release of products (Aleshin *et al*., 1994). Interestingly, an equivalent tryptophan (W84) in an α-amylase from the yeast *Saccharomycopsis fibuligera* has been addressed a similar role. Mutation of this tryptophan to leucine enhanced the transglycosylation activity and reduced the hydrolysis activity on small soluble malto-oligosaccharides (Matsui *et al*., 1994). The cyclization (intramolecular transglycosylation) activity of a bacterial CGTase could be reduced while keeping the hydrolysis activity unchanged by removing a tyrosine residue near the cleavage site. α-Amylases which are functionally related to CGTases are known to possess a small residue (Gly, Ser or Val) at the corresponding position (Penninga *et al*., 1995). A condensation reaction of an inverting glucoamylase could be improved by mutating serine to tyrosine in a distant subsite 4 (Svensson *et al*., 1995).

Various types of conformational changes in and around the active site have also been reported. *Humicola insolens* EGV, soybean β-amylase and pig pancreatic α-amylase 3D structures of liganded and unliganded enzymes have revealed a conformational change of a flexible loop. In complex structures the loop moves closer to the ligand, forming part of the active site and shielding the reaction center from solvent (Mikami *et al*., 1994; Davies *et al*., 1995; Qian *et al*., 1994). As a consequence of this lid flipping in EGV, a third aspartate is brought into close

contact with the two catalytic aspartates. Site-directed mutagenesis of the third aspartate reduced the k_{cat} value whereas the K_m value remained almost unchanged. The role of this aspartate is suggested to be either in transition state stabilization or in determining the protonation state of the proton donor (Davies *et al*., 1995).

1.5 ENZYMOLOGY OF AMYLASES AND CELLULASES

1.5.1 Activity measurements

The diversity and heterogeneity of natural substrates coupled with the mixed specificities of individual enzymes presents a problem in the characterization of amylases and cellulases. Furthermore, the enzymatic degradation of native insoluble substrates involves steps and mechanisms which are not yet understood at the molecular level. Therefore biochemical studies always use cellulose and starch in some modified form to enable or simplify analyses. There are basically four different types of substrates used for activity measurements: purified insoluble substrates approximated to a native substrate, modified insoluble substrates, soluble modified polysaccharides and soluble oligosaccharides. Catalytic activity is usually measured by quantifying formed soluble saccharides or chromophoric aglycon. The action of enzymes on insoluble substrates can also be assayed by other means. For example, changes in the physico-chemical properties of the insoluble substrate, such as degree of polymerisation or small particle formation can be followed (Kleman-Leyer *et al*., 1992, 1994; Walker *et al*., 1990; 1992). Carboxymethyl cellulose (CMC) and hydroxyethyl cellulose (HEC) are water-soluble substrates used to assay endoglucanase activities. The activity can be measured as a decrease in viscosity (Hulme, 1988). A similar viscosimetric method has been used to measure α-amylase activity on starch pastes (Marciniak and Kula, 1982).

The measurement of soluble products from insoluble or soluble polymeric substrates often means assaying the formed reducing sugars. One of the simplest and most widely used is the 3,5-dinitrosalisylic acid (DNS) method (Sumner, 1921; Miller, 1959; Ghose, 1987). However, the colour development in the reaction is not strictly proportional to the number of reducing sugars present, but also to the length of the oligosaccharides, leading to higher apparent reducing values with longer sugars (Robyt and Whelan, 1972). DNS itself also breaks down the substrate. In addition, the reaction is rather insensitive, and many metal ions, thiols and phosphate ions disturb the reaction. Several other reducing sugar determination methods have also been developed, e.g. the Somogyi-Nelson method (Somogyi, 1952; Nelson, 1944), alkaline ferricyanide method, (Robyt *et al*., 1972; Halliwell and Lovelady, 1981), the parahydroxybenzoic acid hydrazine (Lever, 1972) and the 2,2'-bichinconinate method (Waffenschmidt and Jaenicke, 1987; Garcia *et al*., 1993) which are reported to be more sensitive and/or to yield more reliable results.

Sensitivity can also be increased by the use of [U-^{14}C]-labelled cellulose (Vardanis and Finkelman, 1981). In some cases dye groups have been attached to the polymeric substrate, e.g. dyed Avicel (Wood and Bhat, 1988), and dyed amylose- and amylopectin (Klein et al., 1970; Babson et al., 1968) and dyed and

cross-linked starch (Ceska *et al*., 1969; Phadebas tablets, Pharmacia). The enzymatic assay is based on color released from the substrate.

The coupled assay presented by Canevascini (1985) is based on the oxidation of cellobiose by a cellobiose oxidase or cellobiose dehydrogenase in the presence of an electron acceptor. There are many possibilities for the electron acceptor (Schou, 1993). The assay may be performed continuously in the reaction mixture or discontinuosly by stopping the cellulase action first. Since the enzymatic oxidation is not very specific, the reducing end of the substrate is usually modified (e.g. methylated or reduced) to prevent background problems (Schou *et al*., 1993). Corresponding methods have been used for amylases, in which the concentration of released glucose is determined either by glucose oxidase/peroxidase (Kunst *et al*., 1984) or by hexokinase/glucose-6-phosphate dehydrogenase method (Rauscher, 1984).

The use of soluble oligosaccharide substrates is well suited for kinetic measurements in mechanistic studies of enzyme action. Besides kinetic constants, cleavage patterns can be determined and subsite mapping performed. Separation and analysis of reaction products is usually done either by HPLC or TLC (Doner, 1988; Dubreucq *et al*., 1989; Bhat *et al*., 1990; Ruohonen *et al*., 1993; Braun *et al*., 1993; Nidetzky *et al*., 1994b). Alternatively, oligosaccharides coupled with a chromophoric aglycon have been used as substrates. The aglycon is usually either methylumbelliferyl or *p*-nitrophenyl. When this group is cleaved off by the enzyme, a measurable increase in fluorescence or absorbance occurs. These substrates are only useful for enzymes capable of hydrolysing the aglycon (van Tilbeurgh *et al*., 1982; van Tilbeurgh and Claeyssens, 1985; Claeyssens, 1988; van Tilbeurgh *et al*., 1988; Bhat *et al*., 1990; Macarron *et al*., 1993). Oligosaccharides carrying a radioactive label (^{3}H or ^{14}C) at the reducing end sugar also allow sensitive detection and can be used with all enzymes. Radioactively labelled cello-oligosaccharides and malto-oligosaccharides have been used particularly to monitor end-product distribution, thus providing information about which end of the substrate is attacked (Allen and Thoma, 1976b; Bhat *et al*., 1990; Vršanská and Biely, 1992; Biely *et al*., 1993; Matsui *et al*., 1995). Proton NMR can be used for the continuous monitoring of the reaction while information concerning the stereochemistry and attacked end is also obtained (Withers *et al*., 1986; Knowles *et al*., 1988; Claeyssens *et al*., 1990; Teleman *et al*., 1995).

1.5.2 Subsite mapping and ligand binding studies

The existence of multiple binding sites in tandem array was predicted for many depolymerases involving carbohydrases before any structural data was available. This subsite model led to a postulation that short oligomeric substrates can have several possible binding modes, of which some are productive and lead to bond cleavage while others are non-productive. Subsite mapping aims to determine the number of subsites, locate the site of cleavage and assign a certain free energy of binding to each subsite of various hydrolases (Hiromi et al., 1973; Allen and Thoma, 1976a; 1976b; Thoma and Crook, 1982). It is based on the use

of a series of oligomeric substrates of increasing length to determine the bond cleavage frequencies together with the chain length dependence of the Michaelis constant (K_m) and the turnover rate (k_{cat}). The general assumption is that the overall binding is a simple sum of binding energies for each subsite. It has been argued by Sigurskjold and coworkers (1994a) that the overall binding is a cooperative process of interactions involving several subsites and cannot be split into distinct subsites, each giving an additive contribution to binding. With the 3D structures now available for some of these enzymes, it is also evident that the assignment of different subsites on the basis of steady-state kinetics does not necessarily correlate with the structural data. For *T. reesei* CBHI four subsites were predicted on the basis of bond cleavage frequencies and relative rates of hydrolysis (Biely *et al.,* 1993). However, the complex structures of various cello-oligosaccharides with CBHI wt and inactive mutants have revealed clear densities for sugar rings at ten different subsites (Alwyn Jones, personal communication).

The roles of aromatic side chains in substrate binding have been monitored with protein difference spectroscopy. Sugar-protein interactions affect the microenvironment of the chromophore and cause spectral shifts. Difference spectroscopy has been applied e.g. to *A. niger* glucoamylase to quantitate the dissociation constants on various inhibitors. Chemical modifications and several mutant glucoamylases were used to determine the roles of individual aromatic residues in these binding events (Svensson and Sierks, 1992). The binding of soluble substrates and ligands to *T. reesei* CBHII wt enzyme has also been studied by difference spectroscopy and by fluorescence titration (van Tilbeurgh *et al*., 1985; 1989). The latter method is based on quenching of the fluorescence of 4-methylumbelliferyl-β-D-glucosides either partially or completely as a result of binding to CBHII. The dissociation constants of CBHII wt for free oligosaccharides (up to cellotriose) have been measured by displacement titration of the MeUmb-glucosides (van Tilbeurgh *et al*., 1989). Thermodynamics of binding have been studied with both of the above mentioned hydrolases (Sigurskjold *et al*., 1994a; van Tilbeurgh *et al*., 1989). Sigurskjold and coworkers used titration microcalorimetry to determine the thermodynamic parameters (ΔG^o, ΔS^o, ΔH^o) and the stoichiometry of binding for the fungal glucoamylase. The binding kinetics and involvement of transient intermediates on a reaction pathway has also been studied with different carbohydrases. Stopped-flow fluorometry was used to investigate the binding kinetics of MeUmb-glucosides to *T. reesei* CBHII wt (van Tilbeurgh *et al*., 1989) and of malto-oligosaccharides to two glucoamylases (Olsen *et al*., 1993; Ohnishi *et al*., 1994).

1.6 AIMS OF THE PRESENT STUDY

The general aim was to study the structure-function relationships of two carbohydrate-degrading enzymes, *B. stearothermophilus* α-amylase and *T. reesei* cellobiohydrolase II, using the techniques of protein engineering. The 3D structure of CBHII catalytic domain had been solved and the design of single amino acid

changes was based on the known structure. For *B. stearothermophilus* α-amylase no 3D structure exists, but a molecular model was constructed on the basis of sequence alignment and the known structure of *Aspergillus oryzae* α-amylase. In both cases amino acid changes were made by *in vitro* mutagenesis methods.

The specific aims were:

1. To develop an efficient *in vitro* mutagenesis method for the generation of random mutant libraries of complete genes.
2. To use the random mutagenesis method for the *B. stearothermophilus* α-amylase gene.
3. To improve the purification of *T. reesei* CBHII active site mutants.
4. To study the role of specific amino acid residues in catalysis and substrate binding in the active site of *T. reesei* CBHII.

2 MATERIALS AND METHODS

2.1 STRAINS AND VECTORS

Escherichia coli strain BW313 lacking dUTPase (*dut*) and uracil-DNA glycosylase (*ung*) activities was used as a host to produce the uracil template for all the random mutagenesis work (Kunkel, 1985). Strains JM109 and TG2 (both *dut*+ *ung*+) were used as expression hosts for randomly mutagenised genes. *E. coli* strain DH5α (Promega) was used as the bacterial cloning host for all the *cbh2* mutagenesis work. A UV-induced, tryptophan auxotrophic mutant strain, ALKO2319 (unpublished, Primalco Ltd Biotec) derived from the *T. reesei* strain VTT-D-79125 (Bailey and Nevalainen, 1981) was used as the initial host for constructing new expression strains for CBHII mutants. Two double replacement strains ALKO2877 (CBHII- EGI-) and ALKO3133 (CBHII- EGII-) were constructed as described in detail in IV. In *T. reesei* ALKO2877 (CBHII- EGI-) the genes encoding *cbh2* and *egl1* have been replaced by *trpC* and *amdS* genes, respectively and in strain ALKO3133 (CBHII- EGII-) *trpC* and *amdS* markers replace the *cbh2* and *egl1* genes.

M13mp18 and M13mp19 vectors containing the N-terminal fragment of the *E. coli* β-galactosidase gene were used in the development of the random mutagenesis method. The α-amylase gene from the strain *B. stearothermophilus* ATCC 12980 was also cloned into these phage vectors (Tapani Reinikainen, unpublished). M13mp18 contains the coding strand and M13mp19 the non-coding strand of α-amylase as a 1.9 kb *Hind*III-*Sma*I fragment from the plasmid pCSS1 (Suominen *et al*., 1987). Plasmid pTTc9 (Teeri *et al.*, 1987) contains the *cbh2* cDNA as an *Eco*RI-*Bam*HI restriction fragment in pUC8. The fungal expression vector pUJ10 is a derivative of pBR322 containing the *lac*Z gene linked to the *cbh1* promoter and terminator sequences (Ulla Airaksinen and Merja Penttilä, unpublished). A phleomycin selection plasmid pAN8-1 (Mattern *et al.*, 1987) was used in selection of the *Trichoderma* transformants.

2.2 DNA TECHNIQUES

DNA manipulations were carried out according to standard protocols (Sambrook *et al*., 1989).

2.2.1 Random mutagenesis

The development of the mutagenesis method based on the use of a DNA polymerase lacking proofreading activity is described in detail in I. A single-stranded uracil-containing template from phage vectors M13mp18 and M13mp19 is first prepared in *E. coli* BW313 (Kunkel, 1985). Thereafter a base-limited elongation of the second strand is performed with Klenow polymerase. The elongation is performed in four separate reactions (termed A-, C-, G- and T-) in

which one nucleotide at a time is limiting. Each limiting nucleotide mixture contains some radioactively labelled (either α-^{32}P or α-^{35}S) nucleotide, and the elongation can be checked on a sequencing gel. The limiting nucleotide concentration can be varied depending on how long primer elongation is desired. After removal of all nucleotides the misincorporation step with reverse transcriptase is carried out. Each of the four reaction mixtures contains three misincorporating nucleotides (e.g. the A- reaction contains dCTP, dGTP and dTTP) in optimized proportions to give, in principle, a 1:1:1 ratio of incorrect bases. Finally the elongation to completely double-stranded molecules is performed in the presence of Klenow polymerase and T_4 ligase.

The random mutagenesis of *B. stearothermophilus* α-amylase gene was performed in M13mp18 (coding strand) and M13mp19 (non-coding strand) phage vectors. The 1.9 kb insert contained the gene encoding α-amylase with the signal sequence (1650 bp) under the α-amylase promoter (234 bp). The random mutagenesis was performed as described above. An area of 200 - 300 nucleotides at one time was mutagenised under similar conditions as described in I. Five primers were used to mutagenise the A nucleotides in the whole coding region of α-amylase in the M13mp19 construction. With one of these primers C and G nucleotides were also mutagenised. T and G nucleotides were mutagenised with one primer in the M13mp18 construction.

2.2.2 Site-specific mutagenesis of CBHII

Mutations W135F (TGG→TTC), W135L (TGG→TTG), W272A (TGG→GCG)**,** W272D (TGG→GAC) and Y169F (TAT→TTT) were introduced into the *cbh2* cDNA by a polymerase chain reaction (PCR) (Ho *et al.*, 1989). The nucleotide sequence of the whole fragment subjected to PCR was confirmed. The mutated *cbh2* cDNA was cloned under the *cbh1* promoter of the fungal expression vector by changing the *Ksp*I-*Sma*I restriction fragment (containing the *lacZ* gene) of pUJ10 into the mutated *Eco*RI-*Bam*HI fragment.

E. coli competent cells were transformed with standard protocols (Sambrook *et al.*, 1989). *Trichoderma* transformation was carried out according to a standard method using co-transformation with the selection plasmid pAN8-1 (Penttilä, *et al.*, 1987; Nyyssönen *et al.*, 1993). Screening of the best producing transformants was performed as described earlier (Srisodsuk *et al.*, 1993).

2.3 CHARACTERIZATION OF α-AMYLASE MUTANTS

The mutagenised double-stranded DNA was transfected to *E. coli* JM109 or TG2 and the cells were plated on glucose minimal medium/H-top amylopectin azure plates, pH 7, which were incubated at 37^{o}C overnight. Expression of

functional α-amylase was detected as halos around the plaques. A three-grade scale was used to assay the α-amylase activity as follows: ++ for wt size halos, + for smaller but visible halos and - for non-detectable halos. Mutations were identified by DNA sequencing.

2.4 PREPARATION OF THE IMMUNOAFFINITY CHROMATOGRAPHY COLUMN

Three different monoclonal antibodies (mAbs) were tested for immunoaffinity chromatography by coupling them on CNBr-activated Sepharose 4B (Pharmacia, Sweden) as described by the manufacturer. All the antibodies belong to the IgG1 subclass, and recognize either the catalytic domain (CII-8) or the cellulose-binding domain and the interdomain linker (CII-82 and CII-30) of CBHII (Aho *et al*., 1991). Of the three monoclonals, CII-8 did not bind CBHII in the conditions tested whereas CII-82 bound so tightly that only 7 M urea could be used for elution. The third monoclonal, CII-30, bound CBHII in 30 mM sodium phophate, 250 mM sodium chloride, pH 7.4 and elution was achieved with 0.1 M citrate buffer, pH 2.6-3.0. Further coupling experiments were carried out with this mAb.

Crosslinking the purified mAb CII-30 (against the linker+CBD part of CBHII) (Aho *et al*., 1991) to Protein G Sepharose 4 Fast Flow matrix (Pharmacia, Sweden) was performed as described in IV and was slightly modified from Stern and Podlanski (1993). This column was used to purify the CBHII mutants W135L, W135F, Y169F, W272A and W272D as a final step (see section 2.5). The mutants were bound to the column in 50 mM Na citrate, 150 mM NaCl, pH 6.4 and eluted with 0.1 M citrate, pH 2.7. The fractions were neutralized with 1 M Tris/HCl, pH 8.5.

2.5 EXPRESSION AND PURIFICATION OF CBHII ENZYMES

CBHII wt protein was purified from the *T. reesei* Rut C-30 strain as described by Reinikainen *et al.* (1995a). CBHII mutant proteins W135F, W135L, W272A, W272D were produced in the *Trichoderma* strain ALKO2877 (CBHII- EGI-) and mutant Y169F in strain ALKO3133 (CBHII- EGII-) basically as described in IV. Desalted culture supernatant was run through a DEAE-Sepharose fast flow column (Pharmacia, Sweden) equilibrated in 50 mM NaAc buffer, pH 5.6. The flow-through fraction containing CBHII was further purified with thiocellobioside-based affinity chromatography (Tomme *et al.*, 1988a) and finally with immunoaffinity chromatography (see section 2.4).

W135F, W135L and Y169F catalytic domains for structure determination were produced by papain (Sigma) digestion (Tomme *et al.*, 1988b). Catalytic domains were purified with a MonoQ column (Pharmacia, Sweden) equilibriated in 20 mM bis-Tris pH 7.0 and were eluted with a linear salt gradient (0 - 0.1 M NaCl) using the FPLC System.

2.6 PROTEIN BIOHEMISTRY

2.6.1 Estimation of protein concentration

Double antibody sandwich ELISA was used to quantitate CBHII during the course of purification. It was carried out essentially as described by Bühler (1991). The monoclonal antibody against CBHII linker+CBD (CII-89) (Aho *et al*., 1991) was used as the capture antibody and the polyclonal antibody (Niku-Paavola *et al*., 1986) as the secondary antibody.

Total protein was measured by the Bradford method (Bradford, 1976) using purified CBHII wt as a standard. CBHII wt concentration in a purified preparation was determined from the UV absorbance at 280 nm using a molar extinction coefficient, ε=80,500 $M^{-1}cm^{-1}$ (Tomme, 1991). For CBHII W135F, W135L, W272A and W272D mutants, a molar absorptivity value for tryptophan (5,550 $M^{-1}cm^{-1}$) (Wetlaufer, 1962) was subtracted resulting in ε=74,950 $M^{-1}cm^{-1}$. Similarly, for CBHII Y169F mutant a molar absorptivity value for tyrosine (1,340 $M^{-1}cm^{-1}$) (Wetlaufer, 1962) was subtracted resulting in ε=79,160 $M^{-1}cm^{-1}$. A value of 73,000 $M^{-1}cm^{-1}$ was used for CBHI and EGI (Tomme, 1991). For EGII, the molar absorptivity, ε=81,610 $M^{-1}cm^{-1}$, was calculated from the amino acid sequence (Wetlaufer, 1962; Saloheimo *et al*., 1988). The molecular mass of CBHII is 58 kDa and of IgG 150 kDa.

2.6.2 SDS-PAGE and immunoblotting

The purity of the CBHII wt and mutant preparations was checked and verified by SDS-PAGE (Laemmli, 1970) and Western blotting. Proteins were separated on 10% gels and visualized by staining with BioRad silver staining kit or Coomassie Blue (Merril, 1990). For immunoblotting, samples were run on SDS-PAGE gel and proteins were transferred onto a nitrocellulose filter (Towbin *et al.*, 1979). Bands containing CBHII were identified with monoclonal antibody raised against CBHII catalytic domain [CII-8, (Aho *et al.*, 1991)] as described in IV.

2.7 MEASUREMENT OF CELLULOLYTIC ACTIVITIES

2.7.1 Detection of contaminating cellulase activities

Since neither the wild type nor the mutated CBHII enzymes can cleave the heterosidic linkage of small chromophoric oligosaccharides (van Tilbeurgh *et al*., 1988, and our results), these substrates were used to check the purity of CBHII enzyme preparations. The contaminating activities were measured at saturating substrate concentration in 50mM NaAc buffer, pH 5.0 at 50^{o} C basically as described earlier (van Tilbeurgh *et al*., 1988; Claeyssens, 1988; Tomme, 1991). β-Glucosidase activity was determined using 0.9 mM 4-methylumbelliferyl-β-D-

glucopyranoside (MeUmb$(Glc)_1$) (Sigma, USA) as substrate. CBHI and EGI activities were measured in 0.9 mM MeUmb$(Glc)_2$ (Sigma; USA) using 13 mM D-gluconic acid lactone (Sigma; USA) as an inhibitor of β-glucosidases, and EGII activity was measured in 0.4 mM MeUmb$(Glc)_3$ (Lambda, Austria) using 13 mM D-gluconic acid lactone + 86 mM D-glucose (BDH, England) to inhibit β-glucosidases. The activities were calculated from the initial slopes of the reaction velocity curves. Total endoglucanase activity was measured on 0.9 % (w/v) hydroxyethyl cellulose (HEC) (Fluka, Switzerland) in 50 mM NaAc buffer, pH 5.0 at 50^o C. Samples were taken at different time points and the reducing sugars released were determined by the DNS method (Miller, 1959) using D-glucose as a standard. The activities obtained on the different substrates were divided by the estimated CBHII concentration to obtain "specific activities". These values were compared to the specific activities measured with purified control proteins (CBHI, CBHII , EGI and EGII) in order to calculate the contaminating protein amounts as percentages. Since purified β-glucosidase was not available, estimation of its presence in the CBHII preparations was based on comparison with the wt CBHII preparation.

2.7.2 Enzyme kinetics of CBHII

Turnover numbers and Michaelis constants for the hydrolysis of cello-oligisaccharides (Glc_3 - Glc_6) (Merck) in 10mM NaAc buffer, pH 5.0 at 27° C were determined by HPLC (Waters Millipore, USA) equipped with an RI detector basically as described by Teleman *et al*., 1995. The separation was achieved either on an HC-40 cation-exchange column (Hamilton, USA) or on Aminex HPX-42A (Bio-Rad). The conditions used with CBHII wt and Y169F mutant are described in detail in III. The pH dependence of both enzymes was determined at 27°C with cellotetraose as substrate, as described in III. With the mutants W135F, W135L, W272A and W272D, usually one substrate concentration (200 - 300 μM), assumed to be saturating, was used. In all measurements, samples were taken at least at 8-10 different time points in order to produce reliable values for the initial rates. Kinetic constants were obtained from the initial velocities of the reaction curves by a non-linear regression data analysis program [Enzfitter, (Leatherbarrow, 1987)]. For CBHII wt and the W272A and W272D mutants the hydrolysis was followed longer in order to obtain the whole curve. Progress-curve analysis was then applied to the data basically as described by Teleman *et al*. (1995).

2.8 LIGAND BINDING STUDIES

Association constants for CBHII wt and W135L, W135F and Y169F mutants were measured with small soluble cello-oligosaccharides and their derivatives. Binding of MeUmb$(Glc)_1$ and MeUmb$(Glc)_2$ in 50mM NaAc buffer, pH 5.0 at 16^o C was studied by direct fluorescence quenching titrations as described by van Tilbeurgh *et al.* (1985). Fluorescence was measured with an Aminco SPF-500 ratio spectrofluorometer equipped with a thermostated cuvette holder. Excitation was invariably at 318 nm (spectral bandwith 2 nm) and the

emission was measured at 360 nm with a bandwidth of 15 nm, 20 nm or 40 nm. Glucose, cellobiose and cellotriose (Merck) were used as competitive ligands in the displacement titrations. The dissociation constants for MeUmb-oligosaccharides were obtained by non-linear least-squares fitting to the fluorescence data following the mathematical treatment described by de Boeck *et al.* (1983). The fitting was implemented as a Macintosh program, which also allowed manual fitting.

2.9 CRYSTALLOGRAPHY AND MODEL BUILDING

2.9.1 Structural model of *B. stearothermophilus* α-amylase

The 3D model for α-amylase was constructed using the structure of *Aspergillus oryzae* α-amylase as a template and is described in detail in II. The model building was performed by Liisa Holm.

2.9.2 X-ray crystallography of CBHII Y169F mutant protein

The catalytic domain of CBHII Y169F mutant crystallized in the same space group ($P2_1$, a=49.1Å, b=75.8Å, c=92.9Å and $\beta=103.2^{o}$) as the wild type enzyme (Rouvinen *et al*., 1990). The refinement of the structure was performed as described in III. All the X-ray crystallographic work was performed in Uppsala University, Sweden.

3 RESULTS

3.1 STUDY OF THE STRUCTURE-FUNCTION RELATIOSHIPS OF *B. STEAROTHERMOPHILUS* α-AMYLASE

3.1.1 Random mutagenesis of β-galactosidase and α-amylase genes (Papers I and II)

The principle of the random mutagenesis method is illustrated in Figure 1 (I). The α-fragment of *E. coli* β-galactosidase was used as a target in the method development (I). This region is included in the DNA of the single-stranded phage vector M13mp19 and provides a simple screening criterion for the enzymatic activity. The activity of functional α-fragments is indicated as blue plaques on plates containing chromophoric substrate (X-gal), whereas light blue and white plaques indicate disruption of the function. The precise yield, distribution and equilibrium of various mutants were obtained by isolating and sequencing a large number of plaques of different phenotypes. The mutations were distributed essentially in random fashion (I, Fig. 4). The equilibrium between the different point mutations could be altered by changing the proportions of different misincorporating nucleotides in the mixtures (I, Table I). It was also noted that forcing the misincorporation towards completion increased the proportion of multiple mutations (I, Fig.3). Therefore a 40 - 60 % yield of mutants each containing 1.4 - 1.8 mutations/template appeared to give rise to the maximum efficiency, if the aim was single amino acid changes in the corresponding protein.

The mutagenesis of a thermostable α-amylase from *Bacillus stearothermophilus* was performed in phage vectors M13mp18 and mp19. In these constructions the α-amylase gene is under its own promoter and signal sequence. It is expressed in *E.coli* and secreted into the periplasm, from where it leaks out after prolonged growth (Suominen *et al.*, 1987 and our own results). The pH optimum for *B. stearothermophilus* α-amylase wt enzyme is pH 5 - 6 and it is stable in the pH range 6.5 - 7.5 at 80^{o}C (Vihinen and Mäntsälä, 1990). When transfected *E.coli* cells were plated on amylopectin azure plates (pH 7), halos, i.e. zones of clearance, could be detected around the plaques after overnight growth at 37^{o}C. Random mutagenesis of the α-amylase gene was performed under similar conditions as the mutagensis of the *E.coli* β-galactosidase gene (I). A three-grade scale (++, + and –) was applied to the different sized haloes produced by different α-amylase mutants on substrate plates. The proportion of halos showing reduced or no activity (+ or –) was typically less than 10 %. Transformants from each three activity classes were picked randomly and analysed by sequencing. This led to the identification of 98 mutant enzymes, of which 31 showed no activity, 28 showed decreased activity and 39 had wt-like activity on amylopectin azure plates (II, Table II). 75 residues out of 515 were hit by at least one mutation and they were distributed over most of the protein sequence. Since the mutagenesis was

performed mainly to the A nucleotides of the α-amylase gene in the M13mp19 construction, most of the mutants contained codon changes in A. Some codons also contained mutations in C and G, whereas none of the T nucleotides of the coding strand were mutagenised.

3.1.2 The structural model of *B. stearothermophilus* α-amylase and rationalization of the mutant data (Paper II)

The information of the amino acid change(s) combined to the assayed activity of almost 100 different mutants allowed a comparison to be made to the structural model of α-amylase, which had been constructed independently of the mutant data. The model building was done by first aligning the sequence of *B. stearothermophilus* α-amylase gene with nine other known α-amylase sequences (II, Table I and Fig.1). The solved 3D structure of *Aspergillus oryzae* α-amylase (also called Taka-amylase A) (Matsuura *et al*., 1984) was used as a template (II, Fig. 2). The α-amylase model is composed of three domains: A, B and C. Domain A is the large catalytic domain which folds into an $(\alpha/\beta)_8$ barrel, or TIM barrel, protein. It contains most of the conserved residues in α-amylases, e.g. three carboxylic acids proposed to be involved in catalysis (D234, E264 and D331) and aromatic amino acids involved in substrate binding to the active site. A maltohexaose chain was docked into the substrate binding groove winding across the C-terminal end of the β-barrel in domain A. The six glucose unit long substrate could only be brought into contact with the proposed catalytic carboxylates by introducing a sharp bend into the chain. The glucose units make extensive contacts with a number of residues equivalent to those proposed for Taka-amylase A (Matsuura *et al*., 1984) and some of them being highly conserved in all α-amylases.

B. stearothermophilus α-amylase, like other α-amylases, contains a calcium-binding site between domains A and B near the active site cleft. Some of the amino acids involved in coordinating the Ca^{2+} ion are also conserved among α-amylase sequences. Domain B, which forms a lid above the barrel structure, shows generally very little amino acid similarity to Taka-amylase. Furthermore it contains about 45 amino acid long insert not existing in Taka-amylase and this segment was omitted from the model. The proposed structure for domain B should therefore be treated with considerable caution. Domain C is about 100 amino acids long and also lacks invariant residues. It is presumed to fold into an antiparallel β-sheet structure similar to Taka-amylase A on the basis of preservation of alternating hydrophobic and hydrophilic residues in β-strands. The C-terminal extension of 19 amino acids in domain C was omitted from the model because it does not exist in Taka-amylase A structure.

Figure 3 (II) shows the distribution of the three types of mutations in the structural model of α-amylase. There are two clear areas in which mutations destroy the activity on polymeric substrate: around the active site cleft in domain A where the proposed catalytic residues and tyrosines involved in substrate binding are situated, and more surprisingly, at the interface between domains A and C about 30 Å away from the active site

groove (II, Table II and Fig. 3a). The majority of the mutations that did not show any effect on activity were situated at solvent-exposed surfaces in the 3D model.

The proposed catalytic amino acids in Taka-amylase A correspond to residues D234, E264 and D331 in *B. stearothermophilus* α-amylase. D331 is relatively accessible to solvent but D234 and E264, situated close to each other, appear almost buried at the bottom of a pocket. All three amino acid residues were hit by a mutation (D234G, E264D and D331A), which inactivated the enzyme (II, Table II). This suggests that all of them play a crucial role in catalysis. In addition, a conserved arginine, R232, situated in the active site seems to be important for the functionality of α-amylase. The role of another conserved residue, H238, in the Ca^{2+} ion binding was supported by the data from two mutations H238N and H238Y. H238N had reduced activity whereas mutation H238Y completely abolished the activity.

The second major cluster of inactivating mutations was found between domains A and C (II, Table II and Fig. 3). In the α-amylase model both salt bridges and hydrophobic interactions are found across this interface. Single amino acid substitutions I428F and I354L, both affecting the hydrophobic interactions, severely or totally disintegrated the functionality of the enzyme. In addition, a premature stop codon at residue 442, which removes two-thirds of the domain C, led to inactivation of the enzyme.

3.2 THE CATALYTIC DOMAIN OF *T. REESEI* CBHII

3.2.1 The active site tunnel of CBHII

The overall fold of *T. reesei* CBHII is a variant of the classical TIM barrel with only seven parallel β-strands and five α-helices (Fig. 8a) (Rouvinen *et al*., 1990). At the C-terminal end of the barrel, two long loops stabilised by disulfide bonds form an enclosed tunnel approximately 20 Å long. CBHII protein complex structures with different ligands have revealed four glucosyl binding sites (A-D) within the tunnel (Rouvinen *et al*., 1990; Rouvinen, 1990). Although each subsite is somewhat flattened, they have different cross sections (III, Fig. 4a-d). In the subsites A, C and D tryptophan side chains W135, W367 and W269, respectively, make significant contributions to the formation of the sugar binding sites. In site A, in particular, this results in a narrower tunnel cross section. In addition, the ligand structures have revealed a 20^{o} twist in the chain between subsites C and D. Modelling studies with longer oligomeric substrates also suggest favourable van der Waals interactions between W272 and the sixth glucosyl in the putative binding site F just at the entrance of the tunnel (Fig. 14) (Rouvinen, 1990). This, however, requires about a 110^{o} twist in the cellulose chain between the fourth and sixth glucosyl units. There are no clear hydrogen bonds or van der Waals contacts for the putative subsite E (Rouvinen, 1990).

Besides tryptohan residues the active site tunnel contains many residues which form hydrogen bonds to the substrate (Rouvinen *et al*.,

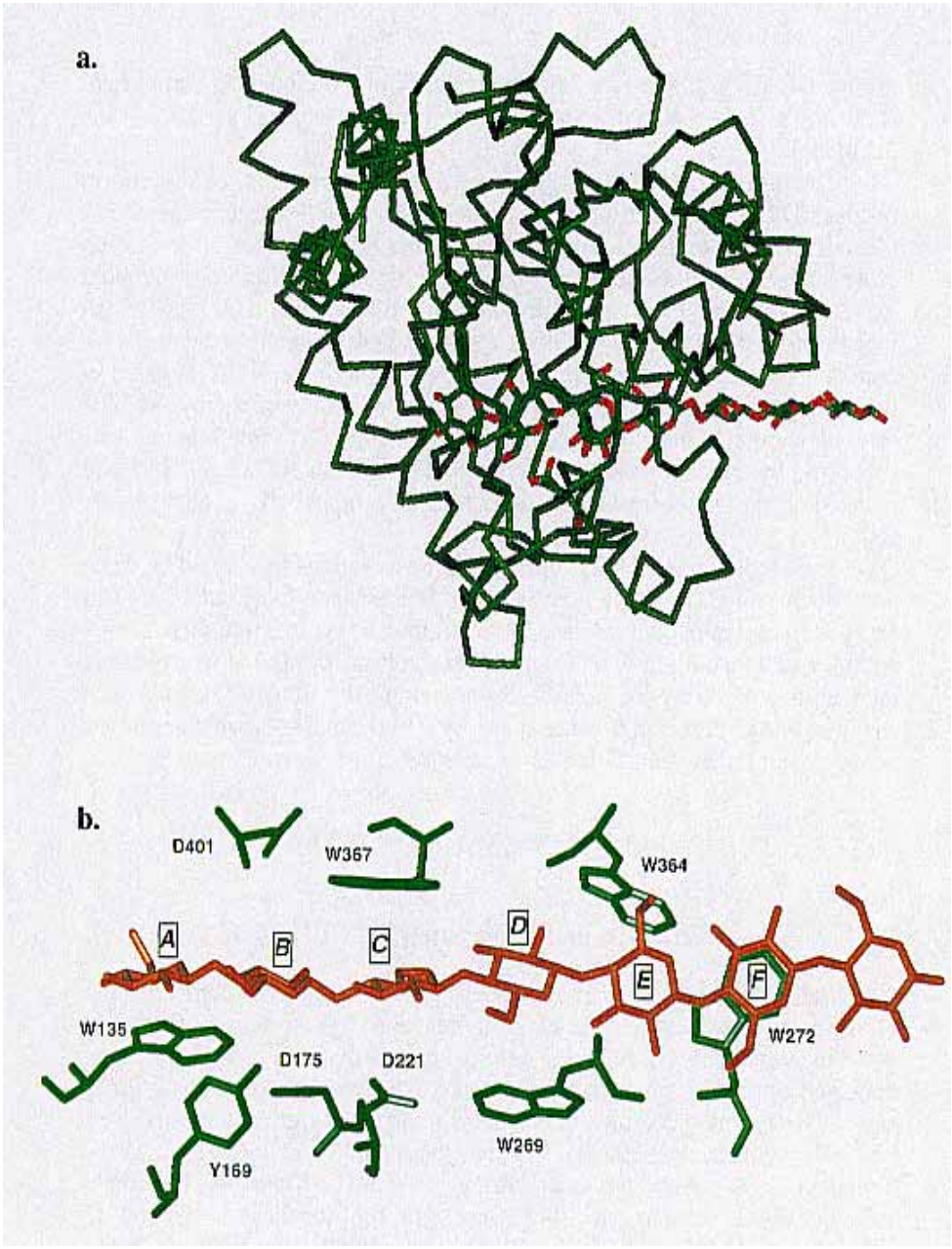

Figure 14. A side-view of the catalytic domain (a) and the active site tunnel (b) of CBHII containing a celloheptaose substrate. The non-reducing end of the substrate binds to subsite A and the bond cleavage occurs between subsite B and C. D221 has been shown to act as a proton donor in the reaction, whereas D401 is suggested to act as a base catalyst. D175 is in hydrogen bond distance from both D221 and Y169. Y169 is also within hydrogen bond distance from the sugar unit at subsite B. The hydrophobic stacking interaction of a tryptophan side-chain and the sugar ring is characteristic in the subsites A, C and D, and in the hypothetical subsite F. The cellulose chain is twisted about 20^{o} between subsites C and D, and 110^{o} between subsites D and F. The figures (a) and (b) were kindly provided by Anna-Marja Hoffrén and Olle Teleman, respectively.

1990). Most of these interactions involve charged residues, which are conserved within the glycosyl hydrolase family 6. Hydrogen bonds involving ionizable residues generally cause strengthening of the bonds by 3 kcal/mol as compared to hydrogen bonds involving uncharged residues (Fersht *et al*., 1985).

3.2.2 Catalytic mechanism of CBHII

Kinetic and structural studies of CBHII have shown that the bond cleavage takes place between the subsites B and C with an inversion of the β-1,4-linkage (Rouvinen *et al*, 1990; Ruohonen *et al*., 1993; Knowles *et al*., 1988; Claeyssens *et al*., 1990). Since CBHII works with inversion of configuration the hydrolysis is postulated to proceed through a single-displacement reaction involving a general acid to donate a proton, and a base to assist the nucleophilic attack of water (Koshland, 1953). Two neighbouring aspartic acids 221 and 175 were identified at the centre of the tunnel between subsites B and C close the O-glycosidic linkage and in hydrogen bond distance from each other (Fig. 14) (Rouvinen *et al*., 1990). D221 is in an environment in which it is expected to be protonated, whereas D175 is more likely to be charged. The mutation D221A abolishes practically all of the catalytic activity of CBHII but does not affect the binding properties on small ligands. Mutation D175A severely impairs the catalytic activity of CBHII but also slightly alters its binding behaviour (Ruohonen *et al*., 1993 and Tables 2 and 3). In addition, the D175A mutant was shown to have about 2 % residual activity on soluble polymeric substrate whereas D221A was completely inactive (Laura Ruohonen, personal communication). It thus seems that D221 acts as the proton donor and that D175 has a role either in ensuring the protonation of D221, stabilizing reaction intermediates or both. The catalytic base has not yet been identified experimentally although D401 has been addressed to this role (Fig. 14).

3.2.3 The role of Y169 in the catalytic action of CBHII (Paper III)

A tyrosine residue, Y169, strictly conserved in the CBHII family 6 and located at site B near the scissile glycosidic bond, may also be involved in the catalytic mechanism (Fig. 14). Based on the CBHII structure this residue is in hydrogen bond distance from the sugar ring and the carboxylate D175. Both of these interactions were expected to be due to the hydroxyl group in the tyrosine side chain and therefore a mutation Y169F was designed.

The crystal structure of the Y169F mutant enzyme was determined to a resolution of 2.0 Å (III, Fig. 1). Figure 2b (III) overlays the native and Y169F mutant structures in the active site tunnel and demonstrates that the mutation causes essentially no change in the structure. However, Y169F mutant enzyme shows similar or increased binding (Table 3 and III, Table II) but lowered catalytic rate on small soluble cellooligosaccharides as compared to the CBHII wt enzyme (Table 2 and III, Table I). The catalytic constants for cellotriose and cellotetraose degradation are shown in Table I

Table 2. Catalytic constants for the CBHII wild type and mutant enzymes. Hydrolysis experiments were performed in 10 mM sodium acetate buffer, pH 5.0 at 27°C. Samples were taken at different time points and analyzed with HPLC as described earlier (Ruohonen et al., 1993; Teleman et al., 1995). Kinetic constants were calculated by a non-linear regression analysis (Enzfitter) or by analyzing whole progress-curves as described by Teleman et al. (1995). The error is estimated to be 10 % on the basis of four repeated experiments.

	k_{cat} (min^{-1})				
Protein	Glc_3	Glc_4	Glc_5	Glc_6	Ref.
CBHII wt	3.7	220	60	840	1, 2
D221A	0.01	≤0.2	nd	nd	3
D175A	0.03	≤0.2	nd	nd	3
Y169F	0.9	57	nd	nd	1
W135F	0.1	75	nd	nd	3, 4
W135L	0.02	1	nd	nd	3
W272A	20	300	480	1500	5
W272D	20	240	480	1800	5

nd = not determined

References: 1. Paper III ; 2. Vesa Harjunpää *et al*., submitted; 3. Ruohonen *et al*., 1993; 4. Koivula *et al*., 1996; 5. Anu Koivula, unpublished.

Table 3. Association constants for the CBHII wild type and mutants. Binding constants on MeUmb-glycosides were determined in 50 mM NaAc buffer, pH 5.0 at 16°C or 8°C. The glucose concentration in the Glc +MeUmb(Glc)$_2$ experiment was in all cases 0.3 M. The data was analysed as described in section 2.8. The uncertainty in K_{ass} values is estimated to be 10-15 % on the basis of four repeated experiments.

		K_{ass} (M^{-1})		
Protein	Temp.	MeUmb$(Glc)_2$	Glc +MeUmb$(Glc)_2$	Ref.
CBHII wt	16°C	3×10^5	6×10^5	1, 2
D221A	16°C	2×10^5	nd	3
D175A	16°C	9×10^5	nd	3
Y169F	16°C	$\geq 200 \times 10^5$	17×10^5	1, 2
CBHII wt	8°C	5×10^5	10×10^5	2
W135F	8°C	0.5×10^5	0.9×10^5	2, 4
W135L	8°C	0.2×10^5	0.1×10^5	2, 4

nd = not determined

References: 1. Paper III; 2. Anu Koivula, unpublished; 3. Ruohonen *et al*., 1993; 4. Koivula *et al*., 1996.

(III). The turnover number (k_{cat}) for Y169F mutant has decreased to 25 % of that of the wild-type enzyme at pH 5.0. The specificity constant (k_{cat}/K_m) has also decreased both on cellotriose and on cellotetraose. The reaction rates for cellotetraose degradation were determined as a function of pH. Over the whole pH range studied the activity of the mutant enzyme was lower; in the pH range 4.5-10.0 the activity was invariably 20 - 25 % of the CBHII wt activity, but at more acidic pH values the activity differential increased linearly towards pH 2.0. The pH-activity profiles are presented in Fig. 5 (III).

The association constants of MeUmb-derivatized oligosaccharides (from direct titration) and of free oligosaccharides (from displacement titration) for CBHII wt and Y169F are shown in Table 2 (III). CBHII wt and Y169F mutant enzymes showed almost identical binding affinity towards both glucose and Meumb$(Glc)_1$. On the other hand, the mutation Y169F was found to increase the binding affinity of CBHII for MeUmb$(Glc)_2$ over 50-fold (Table 3) and the affinity for cellobiose or cellotriose 3-fold. The changed binding behaviour was also evidenced by two different titration experiments. In a displacement titration experiment D-glucose displaced MeUmb$(Glc)_2$ from its complex with the mutant enzyme (data not shown). This is in sharp contrast to the CBHII wt enzyme, for which glucose enhances the affinity of MeUmb$(Glc)_2$ in the CBHII complex (van Tilbeurgh *et al*., 1985; 1989). A direct titration experiment with MeUmb$(Glc)_2$ in the presence of a high glucose concentration (0.3 M) gave a K_{ass} value which was lower than the K_{ass} for MeUmb$(Glc)_2$. This is also opposite to the behaviour of CBHII wt enzyme (Table 3).

3.2.4 Characterization of the kinetic behaviour of CBHII wt

Kinetic constants for CBHII wt enzyme were measured on soluble oligosaccharides using either HPLC or proton NMR (III; Ruohonen *et al*., 1993; Teleman *et al*., 1995; Harjunpää *et al*., 1996, submitted). The data analysis was performed either from the initial slopes or from complete hydrolysis curves by a progress-curve analysis method described by Teleman *et al.* (1995). The rate constants are shown in Table 2 and the specificity constants (k_{cat}/K_m) for cellotriose and cellotetraose are presented in Table I (III). Kinetic experiments with cellotriose have showed that this substrate can bind in more than one way in the active site tunnel of CBHII. The productive binding mode is to subsites A-B-C, whereas the non-productive binding was suggested to occur in subsites B-C-D (Teleman *et al*., 1995). The productive binding mode leads to cellotriose degradation into α-cellobiose and glucose (Fig. 15). Cellotetraose is degraded into two cellobioses (Fig. 15). The specificity constant is 700-fold higher for cellotetraose than for cellotriose hydrolysis, due to both increased binding and reaction rate.

Cellopentaose and cellohexaose are degraded in two different ways and can consequently have two different productive binding modes: the substrate is either bound from the non-reducing end to subsite A or to one over A (Fig. 15). The two binding modes have different orientations, in which the overall zig-zag patterns of

glycosidic linkages are related by 2-fold rotation to each other. In this way either α-cellobiose or α-cellotriose is cleaved off from the non-reducing end of the chain. The overall rate constant for cellopentaose is four times lower than for cellotetraose and 14 times lower than for cellohexaose (Table 2). With cellohexaose it was also demonstrated that practically no cellotetraose is released from the enzyme before a second cleavage occurs to produce cellobiose (Fig. 15).

3.2.5 The role of tryptophans W135 and W272 in the active site of CBHII

CBHII crystal structures with different ligands have shown that subsite A is the tightest binding site (Rouvinen *et al*., 1990; Rouvinen, 1990; III). This subsite must be occupied for hydrolysis to occur, as demonstrated by NMR studies failing to detect α-glucose among the hydrolysis products of CBHII (Ruohonen *et al*., 1993; Teleman *et al*., 1995). Earlier binding studies have revealed that subsite A is very specific for an intact D-glucopyranose configuration (van Tilbeurgh *et al*., 1989). Two mutations, W135F and W135L, were made to study the role of a tryptophan residue at subsite A. Substitution of an indole ring to a phenolic ring in W135F mutant causes decreased binding on small ligands (Table 3) and subsequent reduction in the catalytic efficiency of CBHII (Table 2). D-Glucose enhances the affinity of MeUmb$(Glc)_2$ in the enzyme complex similarly to the behaviour of the CBHII wt (van Tilbeurgh *et al*., 1985; 1989). The reaction rate of W135F on cellotriose is about 3 % of the wt value, whereas the reaction rate on cellotetraose is over 30 % of the wt value.

Removal of the aromatic ring structure completely with mutation W135L leads to even more pronounced effects on catalysis and binding (Tables 2 and 3). A high glucose concentration (0.3M) could not be used to increase the association constant for MeUmb$(Glc)_2$, contrary to the behaviour of CBHII wt enzyme (Table 3). The reaction rates on both cellotriose and cellotetraose were only 0.5 % of the corresponding values with CBHII wt. The crystal structures of unliganded mutant proteins W135F and W135L have been solved, revealing basically no change in the 3D structures (Alwyn Jones, personal communication).

The participation of the putative site F in binding and catalysis was studied by making the mutations W272A and W272D. Here the idea was to remove the aromatic ring structure completely (W272A) and to introduce an extra charge (W272D) at the mouth of the tunnel. The kinetic data shows that the two mutants behave similarly on cello-oligosaccharides (Table 2). The hydrolysis rates are in all cases either increased or kept the same as compared to the CBHII wt. The decrease in the k_{cat} value with CBHII wt on cellopentaose over cellotetraose is not detected with either of the mutants (Table 2).

3.3 PURIFICATION OF THE CBHII ACTIVE SITE MUTANTS (PAPER IV)

Accurate assessment of the activities of CBHII mutants requires protein preparations of the highest purity. In particular, endoglucanases with high specific activities on soluble substrates (Tomme, 1991; Macarron *et al*., 1993) need to be removed completely. To improve the purification, recombinant *T. reesei* production strains lacking the genes coding for both CBHII and one of the two major endoglucanases, EGI (strain ALKO2877) or EGII (strain ALKO3133) were constructed (Fig. 1, IV). The second approach was to use an immunoaffinity chromatography as the final purification step.

Of the three different monoclonal antibodies tested CII-30, which is specific for the cellulose-binding domain (CBD) of CBHII, proved to be useful in purification. Two procedures for preparation of the antibody matrix were tested. Crosslinking of the monoclonal antibody to Protein G matrix instead of the conventional immobilization via cyanogen bromide increased the binding efficiency (IV, Table 1). All the tested active site mutants of CBHII (W135F, W135L, Y169F, W272A and W272D) bound to the immunoaffinity column in neutral pH and were eluted at pH 2.7. The purity of the CBHII mutant preparations was tested using small chromophoric substrates and hydroxyethyl cellulose, which are hydrolysed by the other cellulases but not by CBHII. A summary of the purification of CBHII W135L mutant (from strain ALKO2877) is shown in Table 2 (IV). The immunoaffinity column could be used to purify W135L mutant over 800-fold in a single step and resulted in a homogeneous protein preparation free of proteolytically cleaved forms of CBHII (Fig. 2, IV). The use of the double replacement *T. reesei* production strains, especially that lacking the genes coding for both the endogeneous CBHII and EGII, helped to reduce the total endoglucanase activity in the mutant protein preparations as shown in Table 3 (IV).

4 DISCUSSION

4.1 *IN VITRO* MUTAGENESIS (PAPERS I, II, III)

Mutational analysis is a potent tool in structure-function relationship investigations of carbohydrate-degrading or any other enzymes. Mutagenesis can also be used for various applications. There are different ways of making mutations *in vitro*: domain grafting, gene fragment grafting (chimeric genes) and point mutations. One example of domain grafting is the use of cellulose-binding domains as affinity tags to improve purification (discussed in section 1.3.5). Chimeric genes have been constructed e.g. between various α-amylases to study and improve the thermostability needed in industrial applications (Suzuki *et al*., 1989). In the present work point mutations were made using two different approaches which were complementary to each other.

There are still relatively few proteins for which a detailed 3D structure is available. Furthermore, it is not self-evident which mutations should be made in order to obtain a desired effect in function. This has led to the development of various random mutagenesis methods, which all aim to create a library of all possible point mutations in the target DNA. The DNA can be a regulatory region of a gene, a fragment or the whole gene coding for a protein. Earlier *in vitro* random mutagenesis methods were based on the use of UV or various chemical agents to mutagenize DNA. These methods lack a level of control and work usually only with certain nucleotides and/or with certain mutational hot spot areas in DNA (Botstein and Shortle, 1985). Random mutagenesis based on the use of enzymatic misincorporation of nucleotides was introduced by Shortle *et al*. (1982) and Zakour and Loeb (1982) but was used only in a very limited area of a DNA fragment. In this work the enzymatic misincorporation was applied to a longer DNA fragment, first to the β-galacosidase fragment (≈170 bp) and then to the whole coding sequence of the α-amylase gene (1600 bp) by the use of limited elongation of primers. A reverse transcriptase lacking the proofreading activity was used in the misincorporation step and only the three wrong nucleotides were supplied in the reaction. Enrichment of the mutant strand in *E. coli* was achieved by the uracil-template method described by Kunkel (1985). Theoretically, it is possible to change any amino acid codon to any other by single point mutations in 40 % of the cases. This is sufficient to identify many functionally important amino acids, as shown in this study. Alternatively, remutagenesis and selection can be applied for several rounds. Random mutagenesis can also be performed using degenerated, or spiked, oligonucleotides (Oliphant *et al*., 1986; Hermes *et al*., 1989). The method is based on the controlled cross-contamination of the nucleotide reservoirs during oligonucleotide synthesis by a DNA synthesizer. This method is probably more random than the enzymatic methods, which sometimes suffer from some sequence selectivity leading to uneven distribution of the mutations (Hermes *et al*., 1989).

The polymerase chain reaction (PCR) (Saiki *et al*., 1985) has become one of the most powerful tools in molecular biology. It is based on the use of heat-stable DNA polymerase, synthetic oligonucleotides and thermal cycling devices. It was first used to amplify DNA before cloning but soon found applications in all kinds of gene cloning and construction work, in DNA sequencing and in generating different types of mutations and synthetic genes (for a review see Rashtchian, 1995, and Taylor and Logan, 1995). All the point mutations of CBHII in this work were made using PCR by the overlap extension method described by Ho *et al.* (1989).

PCR has also been applied in random mutagenesis. The two different methods used are based either on the use enzymatic misincorporation, also called error-prone PCR (Gram *et al*., 1992), or on synthetic degenerated oligonucleotides (Barbas *et al*., 1992; Siderovski et al., 1992). These have also been used in combination as described by Deng *et al*. (1994). Much of the development work has been done in generating random antibody mutant libraries, where the targets are the six complementary-determining regions (CDRs) needed in the recognition and binding of an antigen (Gram *et al*., 1992; Barbas *et al*., 1992; Deng *et al*. 1994). The power of the random approach when combined to an efficient expression system (phage display) and a rapid screening method (usually on microtiter plates) was very clearly demonstrated in these investigations. The studies have led both to improvement and to the creation of new binding specificities for antibodies useful in many diagnostic and therapeutic applications.

4.2 *B. STEAROTHERMOPHILUS* α-AMYLASE (PAPER II)

At the time of the random mutagenesis work on *B. stearothermophilus* α-amylase, only two 3D structures of α-amylases had been solved (Buisson *et al*., 1987; Matsuura *et al*., 1984). Furthermore, the resolution of Taka-amylase A 3D structure, which was used as a template in the modelling, was only 3.0 Å (Matsuura *et al*., 1984). The low resolution affects the accurate assignment of main chain fold and bond lengths (Vihinen *et al*., 1994). During the past five years many high resolution 3D structures of α-amylases and other amylolytic enzymes have been published (Table 1) and site-directed mutagenesis studies have been performed (Svensson and Søgaard, 1993). This now facilitates a better elucidation of the data presented in paper II.

The model building of *B. stearothermophilus* α-amylase was based on the assumption that all α-amylases share a common structural design, since the structures of porcine pancreatic α-amylase and a fungal Taka-amylase A had very similar overall folds despite the low amino acid identity (23%, see II, Table I). The random mutagenesis data on *B. stearothermophilus* α-amylase also supported the model. The assumption has now been shown to be correct. All known α-amylase sequences belong to the same glycosyl hydrolase family 13 (Henrissat and Bairoch, 1993). Furthermore, the solved crystal structures from family 13 comprising both α-amylase and CGTases have revealed at least three common domains A, B and C showing high structural equivalence (Fig. 11) (Machius *et al*.,

1995; Brayer *et al*., 1995). The refined 3D structures also include a homologous α-amylase from *Bacillus licheniformis* (Machius *et al*., 1995).

The precision of the structural model of *B. stearothermophilus* α-amylase was assumed to be most accurate in the domains A and C, where the best conserved regions are situated. In addition, the structure of the calcium binding site between domains A and B was expected to be similar and to involve conserved amino acids. The solved 3D structures have now revealed that the active site is always situated in a cleft between domains A and B, domain A invariably having the TIM barrel fold as was also predicted for *B. stearothermophilus* α-amylase. Furthermore, both the active site and the calcium binding site show exceptionally pronounced structural similarity so that the highly conserved amino acids are superimposable between various crystal structures (Machius *et al*., 1995; Brayer *et al*., 1995). Domain C in all amylases from family 13 forms a β-sheet structure including a Greek key motif and the overall topology is very similar (Machius *et al*., 1995).

The alignment of domain B sequence was difficult because of the low overall sequence homology. Domain B is a protrusion between β-strand 3 and α-helix 3 of the α/β -barrel and the sequence varies depending on the enzyme. Three homologous thermostable α-amylases from *Bacillus amyloliquefaciens, B. licheniformis* and *B. stearothermophilus* share a similar B domain sequence about 100 amino acids long (II, Table I). The sequence is longer than those of the other aligned α-amylases, containing an insert of about 45 amino acids. This insert was omitted from the structural model of *B. stearothermophilus* α-amylase since no template for modelling existed (II, Fig. 2). It is therfore not surprising that the topological alignment based on four existing crystal structures done by Machius and coworkers (1995) varies most in domain B from the alignment presented in paper II. The 3D structure of *B. licheniformis* α-amylase revealed a domain B with a twisted β-barrel having a large hole in the interior (Machius *et al*., 1995). However, crystallization of the *B. licheniformis* α-amylase could be accomplished only in the calcium-free form. This may result in some artefacts in domain orientation or structural integrity.

Most of the site-directed mutagenesis studies hitherto conducted have concentrated on the active site residues involved in catalysis and/or substrate binding. In addition, *in vitro* mutagenesis has been used to study the thermostability of α-amylases (for review see Svensson and Søgaard, 1993; Svensson, 1994). The catalytic mechanism of starch hydrolysis by α-amylases is still not very well understood. It has been assumed that the catalytic mechanism bears some resemblance to the mechanism of a retaining lysozyme from hen egg-white (HEWL). All amylases (including CGTases) so far studied possess one glutamic acid and two aspartic acid residues in the active site. In the literature all pairs of these carboxylates have been proposed to be the catalytic pair involved in hydrolysis. Site-directed mutagenesis of all three residues has been carried out with many enzymes, such as *B. subtilis* α-amylase (Takase *et al*., 1992), barley α-amylase (Søgaard *et al*., 1993) and Taka-amylase A (Nagashima *et al*., 1992) and has always resulted in an inactive enzyme. This is also the case with the *B.*

stearothermophilus α-amylase studied here (II, Table II) and by others (Vihinen *et al*., 1991). Based on the examination of the complex structure of a mammalian α-amylase Qian *et al*. (1994) suggested that the glutamate (corresponding to E264 in *B. stearothermophilus* α-amylase) would be the proton donor in the reaction.

The active site of α-amylases contain multiple binding sites for the glucose units of starch. Complex structures with ligands are still mostly lacking for different α-amylases, but kinetic studies have suggested e.g. eight subsites for Taka-amylase A (Allen and Thoma, 1976b), 10 subsites for barley α-amylase (MacGregor *et al*., 1994) and 11 subsites for α-amylase of the yeast *Saccharomycopsis fibuligera* (Matsui *et al*., 1995). A complex structure of a pig pancreatic α-amylase shows a V-shaped active site cleft of five subsites in which the substrate binds in a kinked conformation (Qian *et al*., 1994). The modelling studies with *B. stearothermophilus* α-amylase also suggested the involvement of a bent binding groove (II). Several of the α-amylase subsites are characterized by aromatic amino acids which are involved either in the stacking interactions or which make hydrogen bonds to the sugar rings. Site-directed mutagenesis has been used to study the roles of these residues (Svensson and Søgaard, 1993; Matsui *et al*., 1994; Svensson, 1994). In *B. stearothermophilus* α-amylase single or multiple mutations of Y15, Y57, Y60, Y63, Y369 and Y370 led to decreased activity and were situated in the active site cleft in the model. The topological alignment involving the homologous *B. licheniformis* α-amylase (Machius *et al*., 1995) differs slightly in this area so that e.g. the strictly conserved tyrosine near the cleavage site is apparently Y57 instead of Y63, as suggested by our results (numbering according to *B. stearothermophilus* α-amylase).

The calcium binding site in domain B of α-amylases involves four amino acid residues, two of which are strictly conserved (Machius *et al*., 1995; Brayer *et al*., 1995). The calcium binding site in the *B. stearothermophilus* α-amylase model included D105, D203, D205 and H238, where D105, H238 and D203 are correctly aligned with the known calcium binding residues based on α-amylase 3D structures (Machius *et al*., 1995; Brayer *et al*., 1995). Two different point mutations in H238 had a severe effect on the activity (II, Table II).

The roles of domain B and especially of domain C are still unclear. Besides being important for calcium -binding, domain B has been shown to be important for the substrate specificity of barley α-amylase isozymes (Rodenburg *et al*., 1994; Svensson *et al*., 1995). *B. stearothermophilus* α-amylase as well as related α-amylases from *B. amyloliquefaciens* and *B. licheniformis* are all thermostable liquefying amylases with a relatively large domain B as seen from the sequence alignment in Fig. 1 (II). One plausible explanation is that the structure of domain B affects the end-product spectrum of starch hydrolysis. Furthermore, the loops in domain A are shorter as compared to the 3D structure of Taka-amylase, thus possibly also contributing to the substrate binding site of *B. stearothermophilus* α-amylase.

No distinct function can be ascribed with certainty to domain C, although in *Bacillus* CGTase crystal structures domain C was shown to involve a maltose binding site (Fig. 8b) (Lawson *et al*., 1994). In pig pancreatic α-amylase a clear

electron density for maltose was also detected, this time between domains A and C and similarly involving hydrogen bonding and a stacking interaction against a tryptophan ring (Qian *et al*., 1995). Furthermore, the maltose binds in approximate helical conformation (Qian *et al*., 1995). In both cases, it is speculated that domain C may have an anchoring role for natural polysaccharide substrates. C-terminal deletions of *B. stearothermophilus* α-amylase domain C (Vihinen *et al*., 1994) led either to instability problems or slightly decreased K_m values and increased thermal stability in starch hydrolysis. The results from the random mutagenesis support the idea that the domain C is important for the activity on polymeric substrate and may have a role in orienting the amylose chain with respect to the active site cleft in domain A (II, Table II). A concervative amino acid change I354L severely reduced, while mutation I428F totally destroyed the activity; both mutations were situated at the interface between domain A and C in the model. In addition, a stop codon at residue 442 (II, Table II) removing two-thirds of domain C led to inactivation of the enzyme, supporting the results of Vihinen and coworkers (1994).

This work on *B. stearothermophilus* α-amylase demonstrated how the biochemical data obtained from the 98 random mutants could be used to provide evidence supporting the overall fold of the structural model. The strength of the approach was that a sufficiently large pool of mutats could be produced and the activities assayed. On the other hand, the structural model could be used to rationalize the mutant data, thus allowing general conclusions to be drawn from the functionally important areas in the enzyme structure. The individual mutations in the conserved residues confirmed their importance for the function.

4.3 PRODUCTION OF PURE CBHII ACTIVE SITE MUTANTS (PAPER IV)

For the biochemical and structural characterization of CBHII mutants, amounts of 20 - 50 mg of pure enzyme protein are typically needed. Basically this is not a problem since *Trichoderma reesei* is an efficient producer of extracellular proteins, and is capable of secreting about 40 g/l of cellulases in optimised conditions (Durand *et al*., 1988). The strong inducible c*bh1* promoter has been applied successfully in the production of both homologous and heterologous proteins in *T. reesei* (Nyyssönen *et al*., 1993; Nevalainen and Penttilä, 1995). In inducing conditions on cellulose-based culture media all four major cellulases (CBHI, CBHII, EGI and EGII) are produced simultaneously (Nevalainen and Penttilä, 1995). The purification of a single cellulase from *Trichoderma* culture supernatant is not a trivial task. Heterogeneity in protein preparations has been shown to occur because of post-translational modifications, e.g. N- and O-glycosylation (Salovuori *et al*., 1987; Srisodsuk *et al*., 1993; Béguin, 1990), proteolysis (Hagspiel *et al*., 1989; Haab *et al*., 1990) and protein-protein interactions (Sprey and Lambert, 1983; Wood *et al*., 1989). Some cellulases also bind nonspecifically to chromatographic matrices such as Sephadex.

Attempts to produce fully active wild type or mutated cellobiohydrolases in heterologous hosts, such as *E. coli* and *Saccharomyces cerevisiae* (Teeri, 1987; Reinikainen *et al*., 1992; Michael Szardenings, personal communication), have so far not been very successful. Production in *E. coli* caused instability problems, whereas in *Saccharomyces cerevisiae* overglycosylation decreased the specific activity and also affected the binding on crystalline cellulose. This led to the use of a recombinant *T. reesei* strain lacking the endogenous *cbh1* gene but containing all the other cellulase genes to produce CBHI mutants (Srisodsuk *et al*., 1993; Reinikainen *et al*., 1995b). The CBHI mutations were designed to alter the linker or CBD part and the normal purification scheme with an affinity column could be applied (see below).

Hitherto the best purification scheme for both cellobiohydrolases from *Trichoderma* culture filtrates includes anion exchange followed by affinity chromatography based on a derivatized cellobioside ligand (Tomme *et al*., 1988a). As shown, this purification method works rather efficiently on wt CBHI and CBHII enzymes, producing homogenous enzyme preparations (IV, Table 3). However, some CBHII active site mutants (shown with W135L mutant in IV, Fig. 2a and Table 2) no longer bind properly to the affinity ligand, which is a substrate analogue. Furthermore, it has been shown that trace amounts of endo-type activity can be found in affinity purified CBHII preparations and lead to unexpected results. One of these endoglucanases has been identified as EGII on the basis of substrate specificity and its very similar protein properties to those of CBHII (Reinikainen *et al*., 1995a). Endoglucanase contaminations cause problems especially when studying either completely or partially inactive CBHII mutants, since the major endoglucanases EGI and EGII have higher catalytic rates (k_{cat} values) on short oligosaccharides as compared to those of CBHII (Tomme, 1991).

In order to be able to purify CBHII active site mutants free of endoglucanase contaminations recombinant *T. reesei* production strains lacking either the genes coding for CBHII and EGI or for CBHII and EGII were constructed. To further improve the purification, an immunoaffinity column based on monoclonal antibody against CBHII CBD was made. The purification was monitored by assaying the activities of contaminating cellulases. This was necessary because no specific substrates are available for assaying CBHII activity. Furthermore, specific substrates alone would not have been adequate with CBHII active site mutants. Especially the two mutations in the catalytic residues (D221A and D175A) causing almost total inactivation demonstrated the importance of the use of sensitive chromophoric substrates for the detection of contaminating activities. The catalytic rates given in Table 2 for these mutants are upper limits due to the contamination problem (Laura Ruohonen, personal communication). The D221A and D175A mutants also demonstrated the limitations of the *Trichoderma* expression system, since 100 % pure enzyme cannot be obtained. With the other active site mutants this was not such a severe problem as they remained quite active and CBHII mutants essentially free of endoglucanase contamination and proteolytic degradation products could be purified from the culture media (IV, Table 3).

For similar structure-function studies in future it would be advantageous to consider other host organisms. Heterologous hosts like insect cells or the yeast *Pichia pastoris* which does not hyperglycosylate as easily as *S. cerevisiae* (Romanos, 1995) could be considered. The powerful secretory capacity of *Trichoderma* could also be applied if the expression could be performed under a promoter active on glucose, on which endogenous cellulase genes are known to be repressed (Nevalainen and Penttilä, 1995). Some studies of this kind of approach already exist (Nakari-Setälä and Penttilä, 1995).

4.4 VARIATIONS IN THE ACTIVE SITE STRUCTURES OF FAMILY 6 CELLULASES

The glycosyl hydrolase family 6 currently contains 15 proteins of bacterial and fungal origin (Tomme *et al*., 1995a). The 3D structures of two catalytic domains have been solved: an exoglucanase (CBHII) from *T. reesei (*Rouvinen *et al*., 1990*)* and an endoglucanase (E2) from *Thermomonospora fusca* (Spezio *et al*., 1993). The amino acid identity of these two catalytic domains is rather low, 26 %, but still the overall fold is an α/β barrel in both cases. There are also some differences in the structures; the most obvious is the the active site topology. This is apparently a general difference among the family 6 endo- and exoglucanases deduced from the sequence alignment: the spatially restricted active site tunnel of CBHII is not formed in the endoglucanase E2. One of the loops forming the active site in E2 is much shorter and the other is turned away and binds a SO_4^{2-} group in the crystal structure (Fig. 8 and Fig. 13). Furthermore, it has been shown with the *Cellulomonas fimi* cellobiohydrolase A from the same family that deletion of a fragment corresponding to the carboxy-proximal loop covering the active site tunnel in *T. reesei* CBHII leads to an enzyme having enhanced endoglucanase activity (Meinke *et al*., 1995). The second difference is in the number and character of subsites of family 6 exo- and endoglucanases. There is apparently space for four subsites in the active site cleft of E2 (Taylor *et al*., 1995). The CBHII catalytic domain is larger and there is an additional lid formed above the entrance of the tunnel, which apparently contains a binding site for the sixth glucose unit as discussed in section 4.5.4 (Fig. 14). In addition, cellobiose binds to subsites A-B in E2 crystals (Spezio *et al*., 1993) whereas two molecules are found (either in X-A + C-D or A-B + C-D as discussed in paper III) in CBHII crystals.

The active sites of both cellulases contain many of the conserved amino acid residues of family 6. The proposed general acid and general base catalytic residues in both enzymes are in equivalent positions about 10 - 11 Å away from each other and on opposite sides of the glycosidic linkage of the substrate (see section 4.5.1). However, a third conserved aspartic acid residue has a different conformation in CBHII and E2. One of the loops containing the D79, corresponding to D175 in CBHII, is no longer part of the active site in E2. In CBHII D175 is within a hydrogen bonding distance from D221 and Y169 and has been shown to be involved in catalysis (Ruohonen *et al*., 1993) (Tables 2 and 3). On the other hand, site-directed mutagenesis of four conserved aspartates in the *C. fimi* endoglucanase CenA from the same family 6 suggested that D252 and D392 (corresponding to D221 and D401, respectively, in CBHII) are the acid and base catalysts (Damude

et al., 1995). Interestingly, *C. fimi* CenA can cleave the heterocidic linkage of chromogenic oligosaccharides, whereas *T. fusca* E2 and *T.reesei* CBHII cannot (Damude *et al*., 1995; David Wilson, personal communication; van Tilbeurgh *et al*., 1985 and our results). The differences in the active structures may be important for catalytic mechanisms and substrate binding, allowing the different functions of these enzymes.

4.5 HOW DOES THE CBHII CATALYTIC DOMAIN FUNCTION

4.5.1 Catalytic amino acids

The critical amino acid in the cleavage of glycosidic linkages seems to be the acid catalyst. Mutations of the proton donor usually lead to a very great loss of catalytic activity but do not affect the K_m on unmodified substrates, whereas the mutagenesis of the base catalyst or nucleophile can result in an enzyme with considerable residual activity (Malcolm *et al*., 1989; Chavaux *et al*., 1992; Svensson and Søgaard, 1993; Frandsen *et al*., 1994; Svensson *et al*., 1995). CBHII cleaves the β-1,4 glycosidic bond of cellulose by acid catalysis with an aspartic acid, D221, as the proton donor, and another aspartate, D175, probably ensures its protonation and stabilises charged reaction intermediates (Ruohonen *et al*., 1993). Due to the inverting mechanism the nucleophilic water should approach the anomeric carbon (C1) from the direction opposite to the proton donor, D221. The base in inverting hydrolases is usually situated so that a water molecule can be accommodated between the base and the sugar anomeric carbon (McCarter and Withers, 1994) (see also Fig. 12). In the case of *T. fusca* endoglucanase E2 from the same family 6 as CBHII, it has been suggested that D265 corresponding to D401 in CBHII acts as the base (Spezio *et al*., 1993). The structure of CBHII reveals, however, that although correctly oriented, D401 is salt-linked to R353 and K395. These interactions would be expected to reduce its capacity to extract a proton from the attacking water molecule.

Kinetic studies by Sinnott and coworkers (Sinnott, 1990; Konstantinidis *et al*., 1993) have cast doubt on the classical single-diplacement mechanism for CBHII, suggesting that a base may not be necessary. Experimental support is obtained from the observation that the hydrolysis of α-cellobiosyl fluoride does not seem to proceed with the Hehre-type resynthesis-hydrolysis mechanism (Kasumi *et al*., 1987). It is thus possible that D401 functions as a weak base, or that the polarization of the water molecule is accomplished by other residue(s). Furthermore, the water molecule needed for the reaction could enter the active site through a narrow tube reaching to the external solvent and filled with water as oberved in the CBHII structure (Rouvinen *et al*., 1990). In this case

endoglucanases like E2 would use a different type of mechanism, since no water tunnel is found in E2 (David Wilson, personal communication).

4.5.2 Ring distortion as an element of the catalytic mechanism (Paper III)

Both inverting and retaining enzymes are thought to operate through transition states with substantial oxocarbonium ion character. In some hydrolases the distortion of the sugar ring towards the transition state has been shown to occur already when the substrate binds to the enzyme (McCarter and Withers, 1994). Kinetic studies of Konstantinidis *et al.* (1993) and the high association constant of cellobionolactonoxime (III, Fig. 5) support the idea that the mechanism of CBHII involves oxocarbonium ion-like transition states (van Tilbeurgh *et al.*, 1989). On the other hand, no conformational changes have been denoted in the structure after ligand binding (Rouvinen *et al*., 1990; Rouvinen, 1990). There is some structural evidence that a sugar ring might adopt a distorted configuration in site B upon binding of MeUmb$(Glc)_2$ (III). In addition, subsite B in CBHII is different from all the other binding sites in that it has no tryptophan and has a protrusion (III, Fig. 4b) which may allow alternative sugar conformations.

The experimental data support the idea that Y169 is involved in catalysis by interacting with the glucose ring at the second binding site. The catalytic rate has decreased, whereas the association constants on small soluble ligands have increased (III, Tables I and II). The most pronounced effect on K_{ass} was observed after binding of methylumbelliferyl-cellobioside (MeUmb$(Glc)_2$), when the Y169F mutant enzyme showed over 50-fold higher affinity as compared to the wild type CBHII (III, Table II). Earlier a significant increase in affinity was observed after binding of a MeUmbGlcXyl ligand to the CBHII wild type (van Tilbeurgh *et al.*, 1989). This ligand is missing a methylhydroxyl group on the sugar, while the Y169F mutant lacks a hydroxyl group in the corresponding position on the enzyme. It is tempting to speculate that this is a ring distortion effect. With more space at site B the sugar ring could adopt a more relaxed conformation whereas in the tighter space (in the wild type enzyme) the ring may be forced to adopt a more strained conformation.

The specificity constant (k_{cat}/K_m) of the wt CBHII is 700-fold higher for cellotetraose than for cellotriose hydrolysis (III, Table I), due to both increased binding and reaction rate. This shows that the sugar binding at subsite D is an important factor in transition state stabilization. Y169F mutant shows transition state destabilization, more pronounced on cellotriose than on cellotetraose (III, Table I), further stressing the importance of subsite D. Since no ligand-induced conformational changes have been observed with CBHII wt structures, catalysis would be facilitated if all chemical species along the reaction coordinate resembled each other, and if on binding at subsite B, the glucose unit were forced into a conformation similar to the transition state. Furthermore, it is possible although speculative that binding of the fourth sugar ring in site D causes strain, which helps the ring distortion at site B. Since the catalysis rate of Y169F mutant enzyme was reduced only to 25 % of the wt activity it is likely that other residues

are also involved in the distortion of the sugar ring. Y169 residue is expected to promote the distorted sugar conformation at site B by forming a hydrogen bond.

Reports of two inverting endoglucanases, namely E2 from *T. fusca* and EGV from *Humicola insolens*, suggest involvement of conformational distortion of the scissile bond (Taylor *et al*., 1995; Davies *et al*., 1995). The complex structure between substrate and an inactive mutant of EGV shows excellent electron density for an oligosaccharide in six of the enzyme's seven subsites but no sugar unit at the subsite preceding the cleavage site (Davies *et al*., 1995). The geometry of the cleavage site suggests that EGV would favour the binding of a sugar with an elongated glycosidic bond resembling the proposed transition state of the scissile bond (Tanaka *et al*., 1994). It has also been proposed that the binding of a glucose ring at the subsite preceding the cleavage site is energetically unfavourable for the substrate and therefore no glucose unit is seen binding to it (Davies *et al*., 1995). In a molecular dynamics simulation study of *T. fusca* E2 (from family 6) a tilt in the second sugar unit at subsite B was observed which also led to a conformational distortion of the scissile bond (Taylor *et al*., 1995). The modelling data also suggested that the conserved tyrosine residue (corresponding to Y169 in CBHII) might have an active role in the catalysis in the proton transfer process. The experimental data on *T. reesei* CBHII Y169F mutant does not support quite such an active role for Y169 in the catalytic action.

In addition, a change in the pH activity profile was observed. Since the altered pH activity profile of Y169F is not caused by a loss in pH stability (III), it is likely that the hydrogen bonding network around D175 has changed and affects the hydrolysis rate of the mutant. It is postulated that one of the functions of Y169 is to modulate the protonation states of the interacting carboxylates of D175 and D221. Maintaining the negative charge of D175 stabilizes the charged reaction intermediates and assures the protonation of D221. The structure reveals that the pK_a of Y169 may in turn be lowered due to its planar stacking interaction with the guanidino group of a conserved R174.

4.5.3 The importance of subsite A

The earlier binding studies carried out before the CBHII structure became available predicted that CBHII contains four subsites A-D and that the bond cleavage occurs between sites C and D. The subsite A was proposed to be an allosteric site where an intact D-glucose binds, increasing the binding affinity of small ligands in subsites B, C and D (van Tilbeurgh *et al*., 1985; 1989). The kinetic and structural data have confirmed the existence of at least four subsites (III, Table I; Rouvinen *et al*., 1990). Kinetic studies with CBHII wt have demonstrated that high glucose concentration can inhibit cellotriose hydrolysis in a non-competitive manner, whereas cellobiose inhibition is competitive (Teleman *et al*., 1995). On the other hand, the mutagenesis data has demonstrated that the bond cleavage occurs between subsites B and C. Furthermore, there has hitherto been no structural evidence of a separate glucose binding site outside the tunnel. The complex structure of CBHII wt with glucose+cellobiose revealed a glucose

unit in subsite A and cellobiose in subsites C and D (Rouvinen *et al.*, 1990), whereas no structural changes could be detected in the enzyme structure. The non-competitive inhibition of cellotriose was explained by assuming that cellotriose can bind in two different ways in the active site tunnel of CBHII: productively (A-B-C) and non-productively (B-C-D) (see section 4.5.4 for a suggestion of additional non-productive binding modes). In the presence of high glucose concentration glucose binds in subsite A and displaces cellotriose. The glucose inhibition appears to be non-competitive and allosteric, since cellotriose is not expelled from the active site (Teleman *et al*., 1995).

The importance of subsite A was studied by making two mutations, W135F and W135L, both removing the indole ring of W135 at site A. The kinetic data on cellotriose and cellotetraose (Table 2) shows that both mutations have a deteriorating effect on the catalytic rates. In addition, the binding affinities on MeUmb-cellobioside have decreased 10-fold or more (Table 3). Since the unliganded mutant enzyme structures of W135F or W135L do not reveal any detectable changes (Alwyn Jones, personal communication), the binding site A thus has a role both in substrate binding and in catalysis. Kinetic studies with CBHII wt enzyme have also demonstrated that subsite A must be occupied by the substrate in order for the cleavage to occur, since no α-glucose has been detected among the hydrolysis products (Fig. 15) (Ruohonen *et al*., 1993).

The "glucose effect", i.e. an increased K_{ass} value for MeUmb-cellobioside in the presence of high glucose concentration (0.3 M), was studied with W135F, W135L and Y169F mutants (Table 3). Both the W135L and Y169F mutants showed altered behaviour: high glucose concentration decreased the K_{ass} value for MeUmb-cellobioside. In addition, D-glucose could be used to displace MeUmb$(Glc)_2$ from its complex with the W135L and Y169F enzyme (data not shown). W135F mutant behaved like the wt enzyme: high glucose concentration doubled the K_{ass} value of MeUmb-cellobioside (Table 3). This data is interpreted as indicating that glucose binds in the active site tunnel of CBHII wt to the subsite A. On the basis of crystallographic data MeUmb$(Glc)_2$ is known to bind to subsites A-B-C, as discussed in the previous section (Rouvinen *et al*., 1990). In the presence of a high glucose concentration, glucose displaces MeUmb$(Glc)_2$ out of the A-B-C binding mode to another binding mode, e.g. to B-C-D and/or C-D-E. Since the binding affinity of MeUmb$(Glc)_2$ to Y169F is so high, glucose is expected to displace the ligand only partially and therefore the measured K_{ass} value reflects the combination of two (or more) binding modes. In the case of the W135L mutant, glucose either does not bind to the subsite A or cannot induce the higher affinity binding mode for MeUmb$(Glc)_2$.

These studies further demonstrate that the high affinity glucose binding site is the subsite A of the active site tunnel of CBHII, in which the affinity is caused by the stacking interaction with the indole ring of W135. The binding of glucose to subsite A may induce some subtle conformational changes, e.g. in subsite D, not detectable in crystallographic experiments and causing the increase in K_{ass} value. As shown by van Tilbeurgh *et al.*

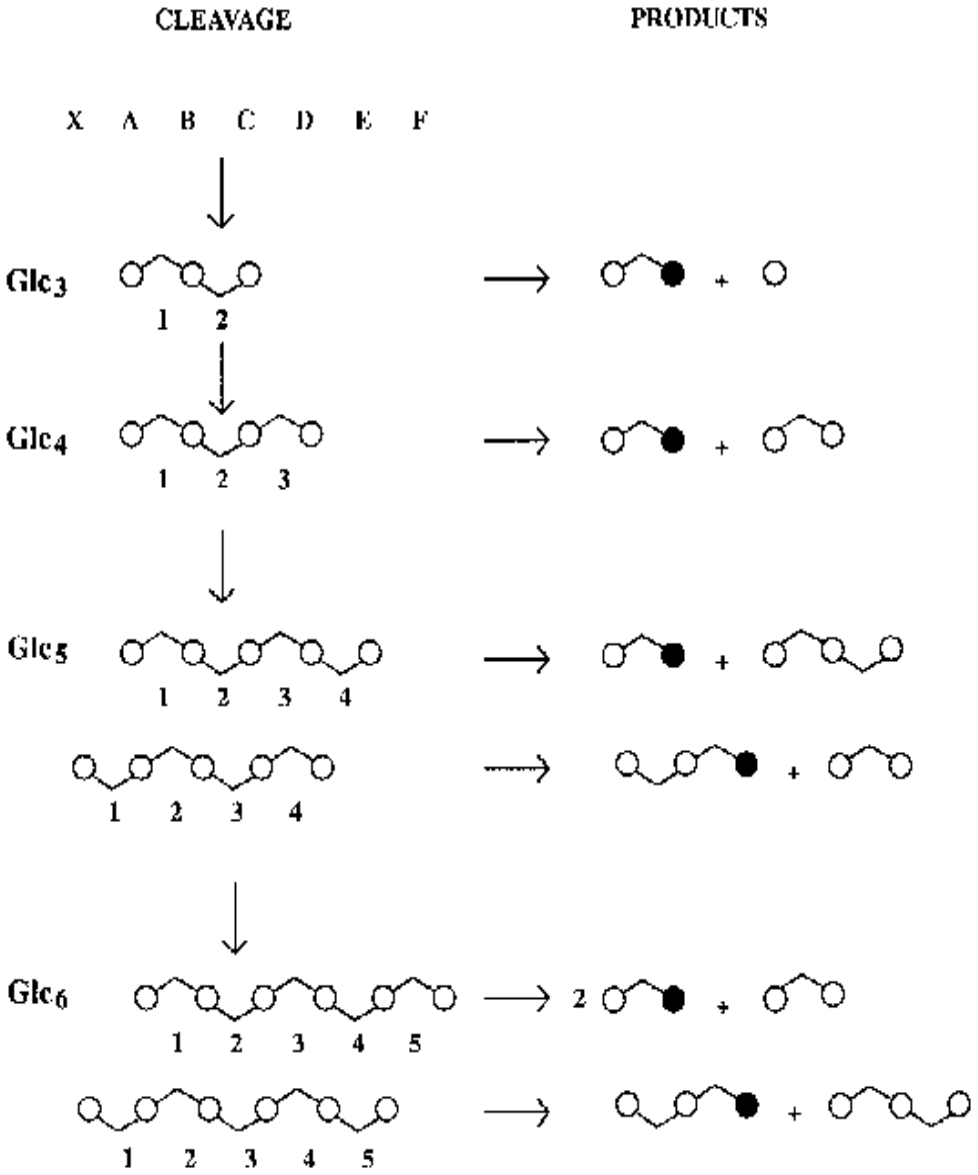

Figure 15. The action pattern of CBHII wt on soluble substrates. Cellobiose is not degraded by CBHII and longer cello-oligosaccharides are cleaved only when the subsite A is occupied. Cello-oligosaccharides can enter the active site tunnel in two different conformations in which the patterns of the alternating glycosidic linkages differ as shown in the picture. The chain cleavage occurs when the orientation of the glycosidic linkage between subsites B and C is "down" towards the catalytic acid residue. Only the productive binding modes are shown. The black sphere denotes the new reducing end of α-anomeric configuration. The picture was kindly provided by Vesa Harjunpää.

(1989), glucose decreases both the the on- and off-rate of MeUmb$(Glc)_2$ binding. Theoretically, a high glucose concentration in the active site tunnel could affect the ligand association and dissociation rates without causing any structural changes in the enzyme. Nevertheless, this does not mean that the subsite coupling upon oligomeric substrate binding in general, suggested by van Tilbeurgh *et al*. (1989), could not hold for CBHII active site tunnel. Experimental confirmation, however, may not be an easy task.

As discussed in the previous section (4.5.2) subsite D is important for the transition state stabilization seen from the 700-fold higher specificity constant (k_{cat}/K_m) for cellotetraose compared with cellotriose hydrolysis (III, Table I). The specificity constant for W135F mutant on cellotetraose was decreased about 10 fold (Teeri *et al*., 1995). This shows that binding to both subsites A and D is needed to reduce the activation energy of the reaction. Furthermore, the hydrophobic interaction of the glucose ring at subsite D with W269 is likely to play a similar role as in subsite A with W135.

4.5.4 Evidence for the new subsite F

The kinetic data on small soluble cello-oligosaccharides with CBHII wt can be rationalized if it is assumed that there are more than four subsites and that non-productive binding is possible with substrates shorter than the number of subsites. In addition, the substrate is expected to enter the active site tunnel in two different orientations with equal probabilities as discussed in section 3.2.4 (Fig. 15). It is also possible that twisting of the oligosaccharide chain between subsites D and F (Fig. 14) does not allow all possible non-productive binding modes e.g. for cellotetraose, as discussed below. The earlier hydrolysis studies on *T. reesei* CBHII (Nidetzky *et al*., 1994b) demonstrated that the progress-curves of oligomeric substrates bend in a manner not explainable by any type of product inhibition. The phenomenon was explained by a non-steady-state model based on the formation of productive and nonproductive enzyme-substrate complexes. However, the hydrolysis curves with CBHII wt enzyme did not show similar behaviour in our experiments. It is possible that the discrepancy between the results reported by us and by Nidetzky and coworkers (1994b) is caused by non-specific binding of CBHII to the surface of glass tubes. This could be detected only in very dilute enzyme concentrations, which were invariably used by Nidetzky and coworkers (1994b) due to a sensitive detector used with HPLC.

The increased turnover numbers for F-site mutants W272A and W272D are assumed to be due to decreased binding affinity for subsite F, which leads to removal of some of the non-productive binding modes. The data obtained with cellopentaose fits these assumptions. For CBHII wt enzyme the catalytic rate on cellopentaose is four times lower than for cellotetraose (Table 2). This is expected to be due to non-productive binding to subsites including a sixth binding site F (e.g. to B-F). The reaction rates of W272 mutants (Table 2) on cellopentaose are 8-fold higher as compared to the wt enzyme. In addition, the hydrolysis rate on cellopentaose is now 2-fold higher than on cellotetraose. Cellotetraose binding to CBHII wt enzyme would not seem to involve a non-productive binding to C-F, as

$k_{cat}(Glc)_4$ is higher than $k_{cat}(Glc)_5$. In addition, no change was observed in the hydrolysis rates with W272 mutants. One plausible explanation is that the twisting of the oligosaccharide chain between subsites D and F does not allow (strong enough) binding of cellotetraose to C-F in the wt or mutant enzymes (see Fig. 14).

Since the catalytic rates of W272 mutants also increased on cellotriose it is speculated that one of the non-productive binding modes of cellotriose involves binding to D-F in CBHII wt enzyme, in addition to B-D as suggested earlier (see section 4.5.3). The slight increase in the hydrolysis rate on cellohexaose is interpreted as non-productive binding mode to B-F and one glucose ring protruding outside the binding area. There might be other plausible explanations for this result. It also remains to be seen whether the binding site F is important in binding and/or breakdown of crystalline cellulose besides binding the cellulose chain tighter in the active site. One of the remaining questions in the enzymatic hydrolysis of crystalline cellulose is how the cellulose chain is lifted from the solid surface to the active site tunnel of cellobiohydrolases. As discussed earlier (see section 4.4), endoglucanase E2 from the CBHII family 6 does not have the corresponding subsite. Other endoglucanases of family 6 also seem to lack the equivalent to W272, whereas it is found in all the cellobiohydrolase sequences.

4.6 PROCESSIVITY OF POLYSACCHARIDE CHAIN CLEAVAGE

Many carbohydrases are assumed to degrade the substrate in a processive manner, catalyzing several bond cleavages before dissociation from the polysaccharide chain. This is also referred as a multiple attack mechanism. Both endo- and exo-types of enzymes have been suggested to have this property (Robyt and French, 1967; Mazur and Nakatani, 1993; Divne *et al*., 1994; Davies and Henrissat, 1995). The processive action mode of a particular carbohydrase is apparently dictated by the active site topology, number and character of subsites, and position of the catalytic amino acids with respect to the number of subsites. Exoenzymes, such as glucoamylases having a funnel-shaped active site with only one entrance (Fig. 13a), are not predicted to make multiple cuts along the substrate before dissociation. On the other hand, the endoenzyme porcine pancreatic α-amylase is known to proceed processively along the substrate (Robyt and French, 1967; Mazur and Nakatani, 1993). This ability can be partially attributed to the open active site cleft (Fig. 13c). The degree of processivity is presumably also a relative term, endoenzymes having a cleft being less processive than the cellobiohydrolases having a tunnel-shaped active site. Some endoglucanases are known to decrease the average DP or viscosity of the polymeric substrate rapidly and can be assumed to favour a multichain attack action pattern, thus behaving like classical endoenzymes (Kleman-Leyer *et al*., 1994; Srisodsuk, 1994).

Processive enzymatic action assumes that after the cleavage the soluble product leaves the active site relatively rapidly, while the polymeric part of the

product remains long enough in the active site for the next cleavage to occur. The latter time can be affected by increasing the affinity or number of subsites for the polymeric part of the chain. In general, the multiple subsites are probably needed to prevent the chain from organizing back to a crystalline phase. As has been shown with some hydrolases, such as α-amylases (Qian *et al*., 1994), the substrate binding groove may contain a bend or a kink. This topological feature might be important for the hydrolysis of native starch having double-helical structures, which need to be opened up before proper binding and cleavage can occur. Some of the additional domains of amylases are probably also needed in this event (II, Lawson *et al*., 1994; Qian *et al*., 1995).

Trichoderma cellobiohydrolases are proposed to act processively because of the tunnel-shaped active site combined to the multiple binding sites (Divne *et al*., 1994; Davies and Henrissat, 1995). In addition, processivity is considered as a key factor for the efficient enzymatic degradation of crystalline cellulose. However, it has proved difficult to demonstrate processivity experimentally on crystalline cellulose. Hydrolysis studies on soluble cello-oligosaccharides by Nidetzky and coworkers (1994b) suggested that *T. reesei* CBHI and CBHII both act processively. The data presented here from the hydrolysis of soluble substrates further demonstrated that both CBHII wt and W272 mutants degrade cellohexaose in a processive manner: no cellotetraose could be detected as an intermediate product (Fig. 15). This is postulated to be due to the hydrolysis rate of cellotetraose being faster than the off-rate which leads to processive action of CBHII. In addition, since no cellotetraose from the non-reducing end of the chain was produced, it is concluded that loop opening must be very rare during cellohexaose hydrolysis by CBHII (Fig. 15) (Harjunpää *et al*., submitted). It is theoretically possible that the active site loops of CBHII open to allow random binding as the first step of crystalline cellulose hydrolysis. However, this event must be very rare, as cellobiohydrolases do not significantly decrease the average DP of crystalline substrates even after prolonged hydrolysis times and the only soluble products encountered are cellobiose and glucose (Srisodsuk, 1994).

5 CONCLUSIONS AND FUTURE PERSPECTIVES

In the present investigation two different approaches were used to study the relationship between enzyme structure and function. This also led to different levels of information. Since the three-dimensional structure of *Bacillus stearothermophilus* α-amylase is not known, the data from random mutants could not be used to draw detailed conclusions concerning e.g. the enzymatic mechanism. On the other hand, important regions in the enzyme structure distant from the active site cleft could be identified as having an unexpected effect on enzyme function. Furthermore, the work demonstrated the power of random mutagenesis to alter protein properties. This kind of approach is applicable for many different purposes as long as an easy expression system and sensitive screening method can be developed. The phage display method used to engineer e.g. monoclonal antibodies and antibody fragments is probably the best example of this. The use of polymerase chain reaction (PCR) has rationalized the field of molecular biology. Besides making all the cloning work much easier, PCR can also be used for mutagenesis, e.g. for random mutagenesis of complete genes or gene fragments. The basic concept of enzymatic misincorporation of wrong nucleotides to create mutations can also be applied in PCR-driven random mutagenesis. The main difference from the random mutagenesis method developed here is the speed with which the mutations can be produced by PCR.

In the second approach, site-specific mutations were made on the basis of the known three-dimensional structure of *T. reesei* CBHII catalytic domain. Structural data of various enzyme-ligand complexes combined to the detailed kinetic and binding experiments with both wt and different mutant enzymes have provided valuable information about how the catalytic domain of CBHII functions. Together with other similar structure-function studies this provides us with a better picture of carbohydrate-degrading enzymes in general. The catalytic mechanism of glycosyl hydrolases can follow two general routes depending on the spatial arrangement of the catalytic carboxylates and lead to retention or inversion of anomeric configuration. The details of the mechanisms are apparently dictated by the overall active site topology and the amino acid residues involved in the substrate binding. Some enzymes are capable of going through conformational changes upon substrate binding. Polymeric substrates can also bind in high energy conformations to promote the catalysis. CBHII has been shown to degrade soluble cello-oligosaccharides processively, suggesting that the loops forming the active site tunnel open very rarely if ever. Evidence for the ring distortion as an element of the catalytic mechansim of CBHII has been obtained. The results also demonstrate that alterations in the substrate binding sites of CBHII affect the catalytic efficiency. In addition, co-operativity may be involved in substrate binding. The hydrophobic stacking interactions of tryptophans or tyrosines both in catalytic and substrate binding domains of various carbohydrases seem to be a

general way of providing high binding affinity towards the polysaccharide substrate.

The catalytic mechanism and substrate binding to the active site or to the substrate binding domain have been the most studied functions of various carbohydrate-degrading enzymes. There are still many open questions waiting to be solved in other fields of enzymatic functions. We do not understand e.g. how breaking of the insoluble substrate is achieved by an individual enzyme, how the catalytic domain and substrate binding domain interact or what is the function of all the additional domains of carbohydrases. Even without fully understanding all the details, applications are already possible but many more will evolve as we learn more of the basic rules and learn to ask the right questions.

REFERENCES

Abuja, P. M., Pilz, I., Claeyssens, M., Tomme, P. 1988a. Domain structure of cellobiohydrolase II as studied by small X-ray scattering, pp. close resemblance to cellobiohydrolase I. Biochem. Biophys. Res. Comm. 156, pp. 180 - 185.

Abuja, P. M., Schmuck, M., Pilz, I., Tomme, P., Claeyssens, M., Esterbauer, H. 1988b. Structural and functional domains of cellobiohydrolase I from *Trichoderma reesei*. Eur. Biophys. J. 15, pp. 339 - 342.

Aho, S., Olkkonen, V., Jalava, T., Paloheimo, M., Bühler, R., Niku-Paavola, M.-L., Bamford, D. H. and Korhola, M. 1991. Monoclonal antibodies against core and cellulose-binding domains of *Trichoderma reesei* cellobiohydrolase I and II and endoglucanase I. Eur. J. Biochem. 200, pp. 643 - 649.

Aleshin, A., Gobulev, A., Firsov, L. M. and Honzatko, R. B. 1992. Crystal structure of glucoamylase from *Aspergillus awamori* var. *X100* to 2.2-Å resolotion. J. Biol. Chem. 267, pp. 19291 - 19298.

Aleshin, A., Firsov, L. M. and Honzatko, R.B. 1994. Refined structure for the complex of acarbose with glucoamylase from *Aspergillus awamori* var. *X100* to 2.4-Å resolotion. J. Biol. Chem. 269, pp. 15631 - 15639.

Allen, J. D. and Thoma, J. A. 1976a. Subsite mapping of enzymes. Depolymerase computer modelling. Biochem. J. 159, pp. 105 - 120.

Allen, J. D. and Thoma, J. A. 1976b. Subsite mapping of enzymes. application of the depolymerase computer model to two α-amylases. Biochem. J. 159, pp. 121 - 132.

Atalla, R. H. and VanderHart, D. L. 1984. Native cellulose: a composite of two distinct crystalline forms. Science 223, pp.283 - 285.

Atalla, R. H. 1993. The structures of native celluloses.In: Suominen, P. and Reinikainen, T. (Eds.). *Trichoderma reesei* cellulases and other hydrolases. Foundation for Biotechnical and Industrial Fermentation Research. Helsinki. Vol. 8, pp. 25 - 39.

Babson, A. L., Kleinman, N. and Megraw, R. E. 1968. A new substrate for serum amylase determination. Clin. Chem. 14, p. 802.

Bailey, M. J. and Nevalainen, H. 1981. Induction, isolation and testing of stable *Trichoderma reesei* mutants with improved production of solubilizing cellulase. Enzyme Microb. Technol. 3, pp. 153 - 157.

Barbas, C. F. III, Bain, J. D., Hoekstra, D. and Lerner, R. 1992. Semisynthetic combinatorial antibody libraries: A chemical solution to the diversity problem. Proc. Natl. Acad. Sci. U.S.A.. 89, pp. 4457 - 4461.

Barr, B. K., Hsieh, Y.-L., Ganem, B. and Wilson, D. B. 1996. Identification of two functionally different classes of exocellulases. Biochemistry 35, pp. 589 - 592.

Bayer, E., Morag, E. and Lamed, R. 1994. The cellulosome - a treasure-trove for biotechnology. TIBTECH 12, pp. 379 - 386.

Bayer, E., Morag, E. Wilchek, M., Lamed, R., Yaron, S. and Shoham, Y 1995. Cellulosome domains for novel biotechnological aplication. In: Petersen, S. B., Svensson, B. and Pedersen, S. (Eds.) Carbohydrate bioengineering. Progress in biotechnology. Elsevier, The Netherlands. Vol 10, pp. 251 - 259.

Béguin, P. 1990. Molecular biology of cellulose degradation. Ann Rev. Microbiol. 44, pp. 219 - 248.

Béguin, P. and Aubert, J.-P. 1994. The biological degradation of cellulose. FEMS Microbiology Reviews 13, pp. 25 - 58.

Belshaw, N. J. and Williamson, G. 1993. Specificity of the binding domain of glucoamylase I. Eur. J. Biochem. 211, pp. 717 - 724.

Bhat, K. M., Hay, A. J., Wood, T. and Claeyssens, M. 1990. Study of the mode of action and site-specifity of the endo-1K4-b-D-glucanases of thr fungus *Penicillium pinophilum* with normal, 1-^{3}H-labeled, reduced and chromogenic cello-oligosaccharides. Biochem. J. 266, pp. 176 - 185.

Biely, P., Vršanská, M. and Claeyssens, M. 1991. The endo-1,4-β-glucanase I from *Trichoderma reesei*. Action on β-1,4-oligomers and polymers derived from D-glucose and D-xylose. Eur. J. Biochem. 200, pp. 157 - 163.

Biely, P., Vršanská, M. and Claeyssens, M. 1993. Mode of action of *Trichoderma reesei* β-1,4-glucanases on cellooligosaccharides. In: Suominen, P. and Reinikainen, T. (Eds.). *Trichoderma reesei* cellulases and other hydrolases Foundation for biotechnical and industrial fermentation research. Helsinki. Vol 8, pp. 99 - 108.

De Boeck, H., Matta, K. L., Claeyssens, M., Sharon, N. and Loontiens, F. G. 1983. Binding of 4-methylumbelliferyl β-D-galactosyl-(1→3)-N-acetyl-β-D-galactosaminide to peanut agglutinin. Eur. J. Biochem. 131, pp. 453 - 460.

Boel, E., Brady, L., Brzozowski, A. M., Derewenda, Z., Dodson, G. G., Jensen, V. J., Petersen, S. B., Swift, H., Thim, L. and Woldike, H. F. 1990. Calcium binding in α-amylases: An X-ray diffraction study at 2.1-Å resolution of two enzymes from *Aspergillus*. Biochemistry 29, pp. 6244 - 6249.

Boisset, C., Borsali, R., Schülein, M. and Henrissat, B. 1995. Dynamic light scattering of the two-domain structure of *Humicola insolens* endoglucanase V. FEBS Letters 376, pp. 49 - 52.

Borchert, T. V., Lassen, S. F., Svedsen, A. and Frantzen, H. B. 1995. Oxidation stable amylases for detergents In: Petersen, S. B., Svensson, B. and Pedersen, S. (Eds.). Carbohydrate bioengineering. Progress in biotechnology. Elsevier, The Netherlands. Vol 10., pp. 175 - 179.

Botstein, D. and Shortle, D. 1985. Strategies and applications of in vitro mutagenesis. Science 229, pp. 1193 - 1201.

Bourne, Y., van Tilbeurgh, H. and Cambillau, C. 1993. Protein-carbohydrate interactions. Curr. Opinion Struct. Biol. 3, pp. 681 - 686.

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

Braun, C., Meinke, A., Ziser, L. and Withers, S. 1993. Simultaneous high-performance liquid chromatographic detrmination of both the cleavage pattern and the stereochemical outcome of the hydrolysis reactions catalyzed by various glycosidases. Anal. Biochem. 212, pp. 259 - 262.

Brayer, G. D., Luo, Y. and Withers, S. 1995. The structure of human pancreatic α-amylase at 1.8 Å resolution and comparisons with related enzymes. Protein Sci. 4, pp. 1730 - 1742.

Buisson, G., Duée, E., Haser, R., and Payan, F. 1987. Three dimensional structure of porcine pancreatic α-amylase at 2.9 Å resolution. Role of calcium in structure and activity. EMBO J. 6, pp. 3909 - 3916.

Bühler, R. 1991. Double-Antibody Sandwich Enzyme-Linked Immunosorbent Assay for Quantitation of Endoglucanase I of *Trichoderma reesei*. Appl. Environ. Microbiol. 57, pp. 3317 - 3321.

Canevascini, G. 1985. A cellulase assay coupled to cellobiose dehydrogenase. Anal. Biochem. 147, pp. 419 - 427.

Ceska, M., Brown, B. and Birath, K. 1969. Ranges of α-amylase activities in human serum and urine and correlations with some other α-amylase methods. Clin. Chim. Acta 26, pp. 445 - 453.

Chanzy, H. 1990. Aspects of cellulose structure. In: Kennedy, J.F., Phillips, G. O. and Williams, P. A. (Eds.). Cellulose sources and exploitation: industrial utilization, biotechnology and physico-chemical properties. Ellis Horwood, New York. pp. 3 - 12.

Chauvaux, S., Béguin, P. and Aubert, J.-P. 1992. Site-directed mutagenesis of essential carboxylic residues in *Clostridium thermocellum* endoglucanase CelD. J. Biol. Chem. 267, pp. 4472 - 4478.

Cheong, C. G., Eom, S. H., Chang, C., Shin, D. H., Song, H. K., Min, K., Moon, J. H., Kim, K. K., Hwang, K. Y. and Suh, S. W. 1995. Crystallization, molecular replacement solution, and refinement of tetrameric beta-amylase from sweet potato. Proteins 21, pp. 105 - 117.

Claeyssens, M. 1988. The use of chromophoric substrates and specific assays in the study of structure-activity relationships of cellulolytic enzymes. In: Aubert, J.-P., Beguin, P. and Millet, J. (Eds.). Biochemistry and Genetics of cellulose Degradation. Academic Press, London. Pp.393 - 397.

Claeyssens, M., Tomme, P., Brewer, C. F. and Hehre, E. J. 1990. Stereochemical course of hydrolysis and hydration reactions catalyzed by cellobiohydrolases I and II from *Trichoderma reesei*. FEBS Lett. 263, pp. 89 - 92.

Covert, S. F., Bolduc, J. and Cullen, D. 1992a. Genomic organization of a cellulase gene family in *Phanerochaete chrysosporium*. Curr. Genet. 22, pp. 407 - 413.

Covert, S. F., Vanden Wymelenberg, A. and Cullen, D. 1992b. Structure, organization, and transcription of a cellobiohydrolase gene cluster from *Phanerochaete chrysosporium*. Appl. Environ. Microbiol. 58, pp. 2168 - 2175.

Crawford, R. L. and Crawford, D. L. 1984. Recent advances in studies of the mechanism of microbial degradation of lignins. Enz. Microb. Technol. 6, pp. 434 - 442.

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

Davies, G. J., Dodson, G. G., Hubbard, R. E., Tolley, S. P., Dauter, Z., Wilson, K. S., Hjort, C., Mikkelsen, J. M., Rasmussen, G. and Schülein, M. 1993. An endoglucanase structure has a novel protein fold but familiar catalytic geometry. Nature 365, pp. 362 - 364.

Davies, G. and Henrissat, B. 1995. Structures and mechanisms of glycosyl hydrolases. Structure 3, pp. 853 - 859.

Davies, G. J., Tolley, S. P., Henrissat, B., Hjort, C. and Schülein M. 1995. Structure of oligosaccharide-bound forms of the endoglucanase V from *Humicola insolens* at 1.9 Å resolution. Biochemistry 34, pp. 16210 - 16220.

Davies, G. J., and Schülein M. 1995. Structural studies of fungal endoglucanases from *Humicola insolens.* In: Petersen, S. B., Svensson, B. and Pedersen, S. (Eds.). Carbohydrate bioengineering. Progress in biotechnology. Elsevier, The Netherlands. Vol 10, pp. 225 - 237.

Deng, S.-J., MacCenzie, C. R., Sadowska, J., Michniewicz, J., Young, N. M., Bundle, D. R. and Narang, S. A. 1994. Selection of antibody single-chain variable fragments with improved carbohydrate-binding by phage display. J. Biol. Chem. 269, pp. 9533 - 9538.

Din, N., Gilkes, N. R., Tekant, B., Miller, R. C. Jr., Warren, R. A. J. and Kilburn, D. G. 1991. Non-hydrolytic disruption of cellulose fibres by the binding domain of a bacterial cellulase. Bio/Technology 9, pp. 1096 - 1099.

Din, N., Forsythe, I. J., Burtnick, L. D., Gilkes, N. R., Miller, R. C. Jr., Warren, R. A. J. and Kilburn, D. G. 1994. The cellulose-binding domain of endoglucanase A CenA from *Cellulomonas fimi* : Evidence for involvement of tryptophan residues in binding. Mol. Microbiol. 11, pp. 747 - 755.

Din, N., Coutinho, J. B., Gilkes, N. R., Jervis, E., Kilburn, D. G., Miller, R. C. Jr., Ong, E., Tomme, P. and Warren, R. A. J. 1995. Interactions of cellulases from *Cellulomonas fimi* with cellulose. In: Petersen, S.B., Svensson, B. and Pedersen, S. (Eds.) Carbohydrate bioengineering. Progress in biotechnology. Elsevier, The Netherlands. Vol 10, pp. 261 - 270.

Divne, C. 1994. The three-dimensional structure of cellobiohydrolase I from *Trichoderma reesei*. Ph.D. Thesis. Uppsala University. Acta Universitatis Upsaliensis.

Divne, C., Ståhlberg, J., Reinikainen, T., Ruohonen, L., Pettersson, G., Knowles, J. K. C., Teeri, T. T. and Jones, T. A. 1994. The three-dimensional structure of cellobiohydrolase I from *Trichoderma reesei* reveals a lysozyme-likeactive site on a lectin-like framework. Science 265, pp. 524 - 528.

Dominguez, R., Souchon, H., Spinelli, S., Dauter, Z., Wilson, K. S., Chavaux, S., Béguin, P. and Alzari, P. 1996. A common protein fold and similar active site in two distinct families of beta-glycanases. Nat. Struct. Biol. 2, pp. 569 - 576.

Doner, L. W. 1988. High-performance thin-layer chromatography of starch, cellulose, xylan, and chitin hydrolyzates. Methods Enzymol. 160, pp. 176 - 180.

Dubreucq, E., Boze, H., Moulin, G. and Galzy, P. 1989. Kinetics of the α-amylase of *Schwanniomyces castellii*. Biotech. Bioeng. 33, pp. 369 - 373.

Ducros, V., Czjzek, M., Belaich, A., Gaudin, C., Fierobe, H.-P., Belaich, J.-P., Davies, G. and Haser, R. 1995. Crystal structure of the catalytic domain of a bacterial cellulase belonging to family 5. Structure 3, pp. 939 - 949.

Durand, H., Clanet, M. and Tiraby, G. 1988. Genetic improvement of *Trichoderma reesei* for large scale cellulase production. Enz. Microb. Technol. 10, pp. 341 - 345.

Erickson, H. P. 1994. Reversible unfolding of fibronectin type III and immunoglobulin domains provides the structural basis for stretch and elasticity of titin and fibronectin. Proc.Natl.Acad.Sci.U.S.A. 91, pp. 10114 - 10118.

Fanutti, C., Ponyi, T., Black, G. W., Hazlewood, G. P. and Gilbert, H. 1995. The conserved noncatalytic 40-residue sequence in cellulases and hemicellulases from anaerobic fungi function as protein docking domain. J. Biol. Chem. 270, pp. 29314 - 29322.

Farber, G. K. and Petsko, G. A. 1990. The evolution of α/β barrel enzymes. TIBS 15, pp. 228 - 234.

Fersht, A. R., Shi, J.-P., Knill-Jones, J., Lowe, D. M, Wilkinson, A. J., Blow, D. M., Brick, P., Carter, P., Waye, M. M. Y. and Winter, G. 1985. Hydrogen bonding and biological specificity analysed by protein engineering. Nature 314, pp. 235 - 238.

Fowler, T., Grizali, M. and Brown, R. D. Jr. 1993. Regulation of the cellulase genes of Trichoderma reesei. In: Suominen, P. and Reinikainen, T. (Eds.). *Trichoderma reesei* cellulases and other hydrolases Foundation for biotechnical and industrial fermentation research, Helsinki. Vol 8, pp. 199 - 210.

Frandsen, T. P., Dupont, C., Lehmbeck, J., Stoffer, B., Sierks, M. R., Honzatko, R. B. and Svensson, B. 1994. Site-directed mutagenesis of the catalytic base glutamic acid 400 in glucoamylase from *Aspergillus niger* and of tyrosine 48 and glutamine 401, both hydrogen-bonded to the γ-carboxylate group of glutamic acid 400. Biochemistry 33, pp. 13808 - 13816.

French, D. 1984. Organization of starch granules. In: Whistler, R. L., Bemiller, J. N. and Paschall, E. F. (Eds.) Starch: chemistry and technology. Academic Press, London, pp. 183 - 245.

Fukumoto, J. and Okada, S. 1963. Studies on bacterial amylase, Amylase types of *Bacillus subtilis* species. J. Ferment. Technol. 41, pp. 427 - 434.

Garcia, E., Johnston, D., Whitaker, J. R. and Shoemaker, S. P. 1993. Assessment of endo-1,4-beta-D-glucanase activity by a rapid colorimetric assay using disodium 2,2' bicinchoninate. J. Food. Biochem. 17, pp. 135 - 145.

Gebler, J., Gilkes, N. R., Claeyssens, M., Wilson, D. B., Béguin, P., Wakarchuk, W. W., Kilburn, D. G., Miller, R. C. Jr., Warren, R. A. J., Withers, S. G. 1992. Stereoselective hydrolysis catalyzed by related b-1,4-glucanases and b-1,4-xylanases. J. Biol. Chem. 267, pp. 12559 - 12561.

Ghose, T. K. 1987. Measurement of cellulase activities. Pure and Appl. Chem. 59, pp. 257 - 268.

Gibbs, M. D., Saul, D. J., Lüthi, E. and Bergquist, P. L., 1992. The β-mannanase from *Caldocellum saccharolyticum* is part of a multidomain enzyme. Appl. Environ. Microbiol. 58, pp. 3864 - 3867.

Gilkes, N. R., Henrissat, B., Kilburn, D. G., Miller, R. C. Jr. and Warren, R. A. J. 1991. Domains in microbial β-1,4-glycanases: sequence conservation. Microbiol. Rev. 55, pp. 303 - 315.

Goto, M., Semimaru, T., Furukawa, K. and Hayashida, S. 1994. Analysis of the raw starch-binding domain by mutation of a glucoamylase from Aspergillus awamori var. kawachi expressed in *Saccharomyces cerevisiae*. Appl. Environ. Microbiol. 60, pp. 3926 - 3930.

Gram, H., Marconi, L.-A., Barbas, C. F. III, Collet, T. A., Lerner, R. A. and Kang, A. S. 1992. *In vitro* selection and affinity maturation of antibodies from naive combinatorial immunoglobulin library. Proc. Natl. Acad. Sci. U.S.A.. 89, pp. 3576 - 3580.

Greenwood, J. M., Ong, E., Gilkes, N. R., Warren R. A. J., Miller, R. C. Jr. and Kilburn, D. G. 1992. Cellulose-binding domains: potential for purification of complex proteins. Protein Eng. 5, pp. 361 - 365.

Gusakov, A. V., Protas, O. V., Chernoglazov, V. M., Sinitsyn, A. P., Kovalysheva, G. V., Shpancenko, O. V. and Ermolova, O. V. 1991. Transglycosylation activity of cellobiohydrolase I from *Trichoderma longibrachiatum* on synthetic and natural substrates. Biochim. Biophys. Acta 1070, pp. 481 - 485.

Haab, D., Hagspiel, K., Szakmary, K. and Kubicek, C. P. 1990. Formation of the extracellular proteases from *Trichoderma reesei* QM 9414 involved in cellulase degradation. *J. Biotechnol.,* 16, pp. 187-198.

Hagspiel, K., Haab, D., Kubicek, C. P. 1989. Protease activity and proteolytic modification of cellulases from a *Trichoderma reesei* QM 9414 selectant. Appl. Microbiol. Biotechnol. 32, pp. 61 - 67.

Halliwell, G. and Lovelady, J. 1981. Utilization of carboxymethylcellulose and enzyme synthesis by *Trichoderma koningii*. J. Gen. Microbiol. 126, pp. 211 - 217.

Harjunpää, V, Teleman, A., Koivula, A., Ruohonen, L., Teeri, T. T., Teleman, O. and Drakenberg, T. 1996. Cello-oligosaccharide hydrolysis by cellobiohydrolase II from *Trichoderma reesei*; association and rate constants derived from an analysis of progress curves. -Submitted.

Harris, E. M. S., Aleshin, A. E., Firsov, L. M. and Honzatko, R. B. 1993. Refined structure for the complex of 1-deoxynojirimycin with glucoamylase from *Aspergillus awamori* var. *X100* to 2.4-Å reslotion. Biochemistry 32, pp. 1618 - 1626.

Heiner, A. P., Sugiyama, J. and Teleman, O. 1995. Crystalline cellulose Iα and Iβ studied by molecular dynamic simulation. Carbohydr. Res. 273, pp. 207 - 223.

Henrissat, B. 1991. A classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem. J. 280, pp. 309 - 316.

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

Henrissat, B 1994. Cellulases and their interaction with cellulose. Cellulose 1, pp. 169 - 196.

Hermes, J.D., Parekh, S.M., Blacklow, S.C., Köster, H. and Knowles, J. 1989. A reliable method for random mutagenesis: the generation of mutant libraries using spiked oligodeoxyribonucleotide primers. Gene, 84, pp. 143 - 151.

Higuchi, T. 1990. Lignin biochemistry. Wood Sci. Technol. 24, pp. 23 - 63.

Hiromi, K., Nitta, Y., Numata, C. and Ono, S. 1973. Subsite affinities of glucoamylase: the examination of the validity of the subsite theory. Biochim. Biophys. Acta 302, pp. 362 - 375.

Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K. and Pease, L. R. 1989. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene 77, pp. 51 - 59.

Hoffrén A.-M., Teeri, T. T. and Teleman, O. 1995. Molecular dynamic simulation of fungal cellulose-binding domains: differences in molecular rigidity but a preserved cellulose binding surface. Protein Eng. 8, pp. 443 - 450.

Hon, D. N.-S. 1994. Cellulose: a random walk along its historical path. Cellulose 1, pp. 1 - 25.

Hulme, M. 1988. Viscosimetric determination of carboxymethylcellulase activity. Methods Enzymol. 166, pp. 130 - 135.

Imberty, A., Chanzy, H., Pérez, S., Buléon, A. and Tran, V. 1988. The double-helical nature of the crystalline part of A-starch. J. Mol. Biol. 201, pp. 365 - 378.

Imberty, A., Buléon, A, Tran, V. and Pérez, S. 1991. Recent advances in knowledge of starch structure. Starch/Stärke 43, pp. 375 - 384.

Irwin, D. C., Spezio, M., Walker, L. P. and Wilson, D. B. 1993. Activity studies of eight purified cellulases: specificity, synergism and binding domain effects. Biotechnol. Bioeng. 42, pp. 1002 - 1013.

Jacks, A. J., Sorimachi, K., Le Gal-Coeffet, M.-F., Williamson, G., Archer, D. anf Williamson, M. P. 1995. ^{1}H and ^{15}N assignments and secondary structure of the starch-binding domain of glucoamylase from *Aspergillus niger*. Eur. J. Biochem. 233, pp. 568 - 578.

Jespersen, H. M., McGregor, E. A., Sierks, M. R. and Svensson, B. 1991. Comparison of the domain-level organization of starch hydrolases and related enzymes. Biochem. J. 280, pp. 51 - 55.

Johansson, G., Ståhlberg, G., Lindeberg, G., Engström, Å. and Pettersson, G. 1989. Isolated fungal cellulase terminal domains and a synthetic minimum analogue bind to cellulose. FEBS Lett. 243, pp. 389 - 393.

Juy, M., Amit, A. G., Alzari, P. M., Poljak, R. J., Claeyssens, M., Béguin, P. and Aubert, J.-P. 1992. Three-dimensional structure of a thermostable bacterial cellulase. Nature 357, pp. 89 - 91.

Kasumi, T., Tsumuraya, Y., Brewer, C. F., Kersters-Hilderson, H., Claeyssens, M. and Hehre, E. J. 1987. Catalytic versatility of *Bacillus pumilis* β-xylosidase: glycosyl transfer and hydrolysis promoted with α- and β-D-xylosyl fluoride. Biochemistry 26, pp. 3010 - 3016.

Kadziola, A., Abe, J., Svensson, B. and Haser, R. 1994. Crystal and molecular structure of barley α-amylase. J. Mol. Biol. 239, pp. 104 - 121.

Kennedy, J. F., Cabalda, V. M. and White, C. A. 1988. Enzymic starch utilization and genetic engineering. TIBTECH 6, pp. 184 - 189.

Kilburn, D. G., Assouline, Z., Din, N., Gilkes, N. R., Ong, E., Tomme, P. and Warren, R. A. J. 1993. Cellulose binding domains: properties and applications. In: Suominen, P. and Reinikainen, T. (Eds.). *Trichoderma reesei* cellulases and other hydrolases. Foundation for Biotechnical and Industrial Fermentation Research. Helsinki. Vol. 8, pp. 281 - 290.

Klein, B., Foreman, J. A. and Searcy, R. L. 1970. New chromogenic substrate for determination of serum amylase activity. Clin. Chem. 16, pp. 32 - 38.

Klein, C. and Schultz, G. E. 1991. Structure of cyclodextrin glycosyltransferase refined at 2.0 Å resolution. J. Mol. Biol. 217, pp. 737 - 750.

Kleman-Leyer, K. M., Agosin, E., Conner, A. H. and Kirk, T. K. 1992. Chages in molecular size distribution of cellulose during attcak by white-rot and brown-rot fungi. Appl. Environ. Microbiol. 58, pp. 1266 - 1270.

Kleman-Leyer, K. M., Gilkes, N. R., Miller, R. C. Jr. and Kirk, T. K. 1994. Chages in molecular size distribution of insoluble cellulose by the action of recombinant *Cellulomomas fimi* cellulase. Biochem. J. 302, pp. 463 - 469.

Knowles, J. K. C., Lehtovaara, P., Murray, M. and Sinnott, M. L. 1988. Stereochemical course of the action of cellobioside hydrolases I and II of *Trichoderma reesei.* J. Chem. Soc., Chem. Commun., pp. 1401 - 1402.

Koivula, A., Ruohonen, L., Reinikainen, T., Claeyssens, M., Jones, T. A. and Teeri, T. T., 1996. Catalytic mechanism of *Trichoderma reesei* cellobiohydrolase II (CBHII). In: Maijanen, A. and Hase, A. (Eds.). New catalysts for clean environment. VTT symposium 163. Technical Research Centre of Finland, Espoo. Pp. 223 - 227.

Konstantinidis, A., Marsden, I. and Sinnott, M. L. 1993. Hydrolyses of α- and β-cellobiosyl fluorides by cellobiohydrolases of *Trichoderma reesei*. Biochem. J. 291, pp. 883 - 888.

Koshland, D. E. Jr. 1953. Stereochemistry and the mechanism of enzymatic reactions. Biol. Rev. 28, pp. 416 - 436.

Kraulis, P. J., Clore, G. M., Nilges, M., Jones, T. A., Pettersson, G., Knowles, J. and Gronenborn, A. M. 1989. Determination of the three-dimensional solution structure of the C-terminal domain of cellobiohydrolase I from *Trichoderma reesei*. A study using nuclear magnetic resonance and hybrid distance geometry-dynamical simulated annealing. Biochemistry 28, pp. 7241 - 7257.

Kulshreshta, A. K. and Dwelz, N. E. 1973. Paracrystalline lattice disorder in cellulose. I. Reappraisal of the application of the two-phase hypothesis to the analysis of powder X-ray diffractograms of native and hydrolused cellulosic materials. J. Polymer Sci. 11, pp. 487 - 497.

Kunkel, T. 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc.Natl.Acad.Sci.USA 82, pp. 488 - 492.

Kunst, A., Draeger , B. and Ziegenhorn, J. 1984. Colorimetric methods with glucose oxidase and peroxidase. In: Bermeyer, H.U. (Ed.) Methods of enzymatic analysis. Weiheim: Verlag Chemie, pp. 178 - 185.

Kuroki, R., Weaver, L. H. and Matthews, B. W. 1993. A covalent enzyme-substrate intermediate with saccharide distortion in a mutant T4 lysozyme. Science 262, pp. 2030 - 2033.

Kuroki, R., Weaver, L. H. and Matthews, B. W. 1995. Structure-based design of a lysozyme with altered catalytic activity. Nature Struct. Biol. 2, pp. 1007 - 1011.

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

Lange, N. K. 1993. Application of cellulases in the textile industry. In: Suominen, P. and Reinikainen, T. (Eds.). *Trichoderma reesei* cellulases and other hydrolases. Foundation for biotechnical and industrial fermentation reserach. Helsinki. Vol 8, pp. 263 - 272.

Lawson, C. L., van Montfort, R., Strokopytov, B., Rozenboom, H., Kalk, K. H., de Vries, G. E., Penninga, D., Dijkhuizen, L. and Dijikstra, B. W. 1994. Nucleotide sequence and X-ray structure of cyclodextrin glycosyltransferase from *Bacillus circulans* strain 251 in a maltose-dependent crystal form. J. Mol. Biol. 236, pp. 590 - 600.

Leatherbarrow, R. J. 1987. Enzfitter. A non-linear regression data analysis program for the IBM PC. Manual. Elsevier Science Publishers BV, The Netherlands. 91 p.

Lever, M. 1972. New reaction for colorimetric determination of carbohydrates. Anal. Biochem. 47, pp. 273 - 279.

Linder, M., Mattinen, M.-L., Kontteli, M., Lindeberg, G., Ståhlberg, J., Drakenberg, T., Reinikainen, T., Pettersson, G. and Annila, A. 1995a. Identification of functionally important amino acids in the cellulose-binding domain of *Trichoderma reesei* cellobiohydrolase I. Prot. Sci. 4, pp. 1056 - 1064.

Linder, M., Lindeberg, G., Reinikainen, T., Teeri, T. T. and Pettersson, G. 1995b. The difference in affinity between two fungal cellulose-binding domains is dominated by a single amino acid substitution. FEBS Letters 372, pp. 96 - 98.

Macarron, R., Acebal, C., Castillob, M. P., Dominquez, J. M., De La Mata, I., Pettersson, G., Toome, P. and Claeyssens, M. 1993. Mode of action of endoglucanase III from *Trichoderma reesei*. Biochem. J. 289, pp. 867 - 873.

MacGregor, E. A. and Svensson, B. 1989. A super-secondary structure predicted to be common to several α-1,4-D-glucan-cleaving enzymes. Biochem. J. 259, pp. 145 - 152.

MacGregor, E. A., MacGregor, A. V. , Macri, L. J. and Morgan, J. E. 1994. Models for the action of barley alpha-amylase isozymes on linear substrates. Carbohydr. Res. 257, pp. 249 - 268.

Machius, M., Wiegand, G. and Huber, R. 1995. Crystal structure of calcium-depleted *Bacillus licheniformis* α-amylase at 2.2 Å resolution. J. Mol. Biol. 246, pp. 545 - 559.

Malcolm, B. A., Rosenberg, S., Corey, M. J., Allen, J. S., de Baetselier, A. and Kirsch, J. F. 1989. Site-directed mutagenesis of the catalytic residues Asp-52 and Glu-35 of chicken egg white lysozyme. Proc. Natl. Acad. Sci. U.S.A.. 86, pp. 133 - 137.

Marciniak, G. P. and Kula, M.-R. 1982. Vergleichende Untersuchung der Methoden zur Bestimmung der Aktivität Bakterieller Alpha-amylasen. Starch/Stärke 34, pp. 422 - 430.

Margolles-Clark, E., Tenkanen, M., Söderlund, H. and Penttilä, M. 1996. Acetyl xylan esterase from *Trichoderma reesei* contains an active site serine and a cellulose binding domain. Eur. J. Biochem. 237, pp. 553 - 560.

Matsui, I., Yoneda, S., Ishikawa, K., Miyairi, S., Fukui, S., Umeyama, H. and Honda, K. 1994. Roles of aromatic residues conserved in the active center of *Saccharomycopsis* α-amylase for transglycosylation and hydrolysis activity. Biochemistry 33, pp. 451 - 458.

Matsui, I., Ishikawa, K., Matsui, E., Miyairi, S., Fukui, S. and Honda, K. 1995. A mutant α-amylase with enhanced activity specific for short substrates. In: Saddler, J.N. and Penner, M.H. (Eds.) Enzymatic degradation of insoluble carbohydrates. American chemical society. Washington, USA. Vol 618, pp. 79 - 89.

Matsuura, Y., Kusunoki, M., Harada, W. and Kakudo, M. 1984. Structure and posible catalytic residues of Taka-amylase A. J. Biochem. 95, pp. 697 - 702.

Mattern, I. E., Punt, P. T., Unkles, S., Powels, P. H., van den Hondel, C. A. M. J. J. 1987. Transformation of *Aspergillus oryzae*. In: Abstracts of the 19th Lunteren lectures on Molecular biology of yeasts and filamentous fungi and its impact on biotechnology. Lunteren, The Netherlands. P. 34.

Mazur, A. K. and Nakatani, H. 1993. Multiple attack mechanism in the porcine pancreatic alpha-amylase hydrolysis of amylose and amylopectin. Arch. Biochem. Biophys. 306, pp. 29 - 38.

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

Medve, J., Ståhlberg, J. and Tjerneld, F. 1994. Adsorption and synergism of cellobiohydrolase I and II from *Trichoderma reesei* during hydrolysis of microcrystalline cellulose. Biotechn. Bioeng. 44, p. 1064 - 1073.

Meinke, A., Damude, H. G., Tomme, P., Kwan, E., Kilburn, D. G., Miller, R. C., Jr., Warren, R. A. J. and Gilkes, N. R. 1995. Enhancement of the endo-β-1,4-glucanase activity of an exocellobiohydrolase by deletion of a surface loop. J. Biol. Chem. 270, pp. 4383 - 4386.

Merril, C. R. 1990. Gel staining techniques. Methods Enzymol. 182, pp. 477 - 478.

Michell, A. J. 1990. Second-derivative F.t.-i.r. spectra of native celluloses. Carbohydr. Res. 197, pp. 53 - 60.

Mikami, B., Hehre, E. J., Sato, M., Katsube, Y., Hirose, M., Morita, Y. and Sacchettini, J. C. 1993. The 2.0-Å resolution structure of soybean β-amylase complexed with α-cyclodextrin. Biochemistry 32, pp. 6836 - 6845.

Mikami, B., Degano, M., Hehre, E. J., and Sacchettini, J. C. 1994. Crystal structure of soybean β-amylase reacted with β-maltose and maltal: active site components and their apparent roles in catalysis. Biochemistry 33, pp. 7779 - 7787.

Miller, G. L. 1959. Use of Dinitrosalisylic Acid Reagent for Determination of Reducing Sugar. Anal. Chem. 31, pp. 426 - 428.

Nagashima, T., Tada, S., Kitamoto, K., Gomi, K., Kumagai, C. and Toda, H. 1992. Site-directed mutagenesis of catalytic active-site residues of Taka-amylase A. Biosci. Biotechnol. Biochem. 56, pp. 207 - 210.

Nakano. H., Hamayasu, K., Fujita, K., Hara, K., Ohi, M., Yoshizumi, H. and Kitahata, S. 1995. Synthesis of 2-deoxy-glucoologosaccharides through condensation of 2-deoxy-D-glucose by glucoamylase and alpha-glucosidase. Biosci. Biotech. Biochem. 59, pp. 1732 - 1736.

Nakari-Setälä, T. and Penttilä, M. 1995. Production of *Trichoderma reesei* cellulases on glucose-containing media. Appl. Env. Microbiol. 61, pp. 3650 - 3655.

Natarajan, S. and Sierks, M. R. 1996. Functional and structural role of the highly conserved Trp120 loop region of glucoamylase from *Aspergillus awamori*. Biochemistry 35, pp. 3050 - 3058.

Nelson, N. J. 1944. A photometric adaptation of the somogyi method for the determination of glucose. J. Biol. Chem. 153, pp. 375 - 380.

Nevalainen, H. and Penttilä, M. 1995. Molecular biology of cellulolytic fungi. In: Kück, (Ed.). The mycota II. Genetics and biotechnology. Springer-Verlag. Berlin, Heidelberg. Pp. 303 - 319.

Nidetzky, B., Hayn M., Macarron, R. and Steiner, W. 1993. Synergism of *Trichoderma reesei* cellulases while degrading different celluloses. Biotechnology Lett. 15, pp. 71 - 76.

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

Nidetzky, B., Zachariae, W., Gercken, G., Hayn, M. and Steiner, W. 1994b. Hydrolysis of cellooligosaccharides by *Trichoderma reesei* cellobiohydrolases: experimental data and kinetic modeling. Enzyme Microb. Technol. 16, pp. 43 - 52.

Nikolov, Z. L., Meagher, M. M. and Reilly, P. J. 1989. Kinetics. equilibria, and modelling of the formation of oligosaccharides from D-glucose with *Aspergillus niger* glucoamylase I and II. Biotechnol. Bioengin. 34, pp. 694 - 704.

Niku-Paavola, M.-L., Lappalainen, A., Enari, T.-M. and Nummi, M. 1986. *Trichoderma reesei* Cellobiohydrolase II. Purification by immunoadsorption and hydrolytic properties. Biotechn. Appl. Biochem. 8, pp. 449 - 458.

Nyyssönen, E., Penttilä, M., Harkki, A., Saloheimo, A., Knowles, J. K. C. and Keränen, S. 1993. Efficient production of antibody fragment by the filamentous fungus *Trichoderma reesei*. Bio/Technol. 11, pp. 591 - 595.

Ogawa, K., Uejima, O., Nakakuku, T., Usui, T. and Kainuma, K. 1990. Enzymatic synthesis of p-nitrophenyl alpha-maltoheptaoside by transglycosylation of maltohexaose-forming amylase. Agric. Biol. Chem. 54, pp. 581 - 586.

Ohnishi, H., Matsumoto, H., Sakai, H. and Ohta, T. 1994. Functional roles of Trp337 and Glu632 in Clostridium glucoamylase, as determined by chemical modification, mutagenesis, and the stopped-flow method. J. Biol. Chem. 269, pp. 3503 - 3510.

Oliphant, A. R., Nussbaum, A. L. and Struhl, K. 1986. Cloning of random-sequence oligodeoxynucleotides. Gene 44, pp. 177 - 183.

Olsen, K., Christensen, U., Sierks, M. R. and Svensson, B. 1993. Reaction mechanism of Trp120→Phe and wild-type glucoamylase from *Aspergillus niger*. Interactions with maltooligodextrins and acarbose. Biochemistry 32, pp. 9686 - 9693.

Ong, E., Gilkes, N. R., Miller, R. C. Jr. and Warren, R. A. J. 1991. Enzyme immobilization using a cellulose-binding domain: Properties of a β-glucosidase fusion protein. Enzyme Microb. Technol. 13, pp. 59 - 65.

Ong, E., Gilkes, N. R., Miller, R. C. Jr. and Warren, R. A. J. 1993. The cellulose-binding domain CBD_{Cex} of an exoglucanase from *Cellulomonas fimi*. Production in *Escherichia coli* and characterization of the polypeptide. Biotechnol. Bioeng. 42, pp. 401 - 409.

Ong, E., Alimonti, J. B., Greenwood, J. M., Miller, R. C. Jr., Warren, R. A. J. and Kilburn, D. G. 1995. Purification of human interleukin-2 using the cellulose-binding domain of a prokaryotic cellulase. Biosepar. 5, pp. 95 - 104.

Penninga, D., Strokopytov, B., Rozenboom, H., Lawson, C. L., Dijikstra, B. W., Bergsma, J. and Dijkhuizen, L. 1995. Site-directed mutations in tyrosine 195 of cyclodextrin glycosyltransferase from *Bacillus circulans* strain 251 affect activity and product specificity. Biochemistry 34, pp. 3368 - 3376.

Penttilä, M. E., Lehtovaara, P., Nevalainen, H., Bhikhabhai, R. and Knowles, J. K. C. 1986. Homology between cellulase genes of *Trichoderma reesei*: complete nucleotide sequence of the endoglucanase I gene. Gene 45, pp. 253 - 163.

Penttilä, M., Nevalainen, H., Rättö, M., Salminen, E. and Knowles, J. K. C. 1987. A versatile transformation system for the cellulolytic filamentous fungus *Trichoderma reesei*. Gene 61, pp. 155 - 164.

Pommier, J.-C., Goma, G., Fuentes, J.-L., Rousset, C. and Jokinen, O. 1990. Using enzymes to improve the process and the product quality in the recycled paper industry. Part 2: Industrial applications. Tappi J. 73, pp. 197 - 202.

Prasad, D. Y., Heitmann, J. A. and Joyce, T. W. 1993. Enzymatic deinking of colored offset newsprint. Nordic Pulp Paper Res. J. 8, pp. 284 - 286.

Qian, M., Haser, R., Buisson, G., Duée, E. and Payan, F. 1994. The active site of a mammalian α-amylase. Structure of the complex of a pancreatic α-amylase with a carbohydrate inhibitor refined to 2.2-Å resolution. Biochemistry 33, pp. 6284 - 6294.

Qian, M., Haser, R. and Payan, F. 1995. Carbohydrate binding site in a pancreatic α-amylase. Structure of the complex of a pancreatic α-amylase-substrate complex, derived from X-ray structure analysis at 2.1 Å resolution. Protein Sci. 4, pp. 747 - 755.

Quiocho, F.A. 1993. Probing the atomic interactions between proteins and carbohydrates. Biochemical Society Transactions 21, pp. 442 - 448.

Rashtchian, A. 1995. Novel methods for cloning and engineering genes using the polymerase chain reaction. Curr. Opin. Biotech. 6, pp. 30 - 36.

Rauscher, E. 1984. Determination of the degradation products maltose and glucose. In: Bermeyer, H. U. (Ed.) Methods of enzymatic analysis. Weiheim: Verlag Chemie, pp. 890 - 894.

Reinikainen, T., Ruohonen, L., Nevanen, T., Laaksonen, L., Kraulis, P., Jones, T. A., Knowles, J. K. C. and Teeri, T. T. 1992. Investigation of the function of mutated cellulose-binding domains of *Trichoderma reesei* cellobiohydrolase I. Proteins: Structure, Function, and Genetics 14, pp. 475 - 482.

Reinikainen, T., Henriksson, K., Siika-aho, M., Teleman, O. and Poutanen, K. 1995a. Low-level endoglucanase contamination in *Trichoderma reesei* cellobiohydrolase II preparation affects its enzymatic activity on β-glucan. Enzyme Microb. Technol. 17, pp. 888 - 892.

Reinikainen, T., Teleman, O. and Teeri, T. T. 1995b. Effects of pH and high ionic strength on the adsorption and activity of native and mutated cellobiohydrolase I from *Trichoderma reesei*. Protein Struct Funct. Genet. 22, pp. 392 - 403.

Reinikainen, T., Takkinen, K. and Teeri, T. T. 1996. Bacterial and fungal cellulose-binding domains have distinctly different adsorption properties. Enzyme Microb. Technol. -In press.

Robyt, J. F. and French, D. 1967. Multiple attack hypothesis of α-amylase action: Action pattern of porcine pancreatic, human salivary, and *Aspergillus oryzae* α-amylases. Arch. Biochem. 122, pp. 8 - 16.

Robyt, J. F. and Whelan, W. J. 1972. Reducing value methods for maltodextrins: I. Chain length dependence of alkaline 3,5-dinitrosalisylate and chain-length independence of alkaline copper. Anal. Biochem. 45, pp. 510 - 516.

Robyt, J. F., Ackerman, R. J. and King, J. G. 1972. Reducing value method for maltodextrins: II. Automated methods and chain-length independence of alkaline ferricyanide. Anal. Biochem. 45, pp. 517 - 524.

Rodenburg, K. W., Juge, N., Guo, X. J., Sogaard, M. and Svensson, B. 1994. Domain B protruding at the third beta strand of the alpha/beta barrel in barley alpha-amylase confers distinct isozyme-specific properties. Eur.J.Biochem. 221, pp. 277 - 284.

Romanos, M. 1995. Advances in the use of *Pichia pastoris* for high-level gene expression. Curr. Opin. Biotechnol. 6, pp. 527 - 533.

Ross, P., Mayer, R. and Benziman, M. 1991. Cellulose biosynthesis and function in bacteria. Microbiol. Reviews 55, pp. 25 - 58.

Rouvinen, J. 1990. Three-dimensional structure and function of cellobiohydrolase II. Doctoral thesis. University of Joensuu. Joensuu, Finland. 38 p.

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

Ruohonen, L., Koivula, A., Reinikainen, T., Valkeajärvi, A., Teleman, A., Claeyssens, N., Szardenings, M., Jones, T. A. and Teeri, T. T. 1993. Active site of *T. reesei* cellobiohydrolase II. In: Suominen, P. and Reinikainen, T. (Eds.). *Trichoderma reesei* cellulases and other hydrolases. Foundation for biotechnical and industrial fermentation research. Helsinki. Vol 8., pp. 87 - 96.

Saiki, R., Scharf, S., Faloona, F., Mullis, K., Horn, G., Erlich, H. and Arnheim, N. 1985. Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science 230, pp. 1350 - 1354.

Saloheimo, M., Lehtovaara, P., Penttilä, M., Teeri, T. T., Ståhlberg, J., Johansson, G., Pettersson, G., Claeyssens, M., Tomme, P., Knowles, J. K. C. 1988. EGIII, a new endoglucanase from *Trichoderma reesei:* The characterization of both gene and enzyme. Gene 63, pp. 11 - 21.

Salovuori, I., Makarow, M., Rauvala, H., Knowles, J. K. C., Kääriäinen, L. 1987. Low molecular weight high-mannose type glycans in a secreted protein of the filamentous fungus *Trichoderma reesei*. Bio/Technology 5, pp. 152 - 156.

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

Saul, D. J., Williams, L. C., Grayling, R. A., Chamley, L. W., Love, D. R. and Bergquist, P. L. 1990. C*elB*, a gene coding for a bifunctional cellulase from the extreme thermophile *Caldocellum saccharolyticum.* Appl. Environ. Microbiol. 56, pp. 3117 - 3124.

Scalbert, A., Monties, B., Lallemand, J.-Y., Guittet, Y. and Romndo, C. 1985. Ether linkage between phenolic acids and lignin fractions from wheat straw. Phytochemistry 24, pp. 1359 - 1362.

Schou, C. 1993. A study of cellulolytic enzymes of fungal and bacterial origin. Ph.D. Thesis. The Technical University of Denmark. 116 p.

Schou, C., Rasmussen, G., Kaltoft, M., Henrissat, B. and Schülein, M. 1993. Stereochemistry, specificity and kinetics of the hydrolysis of reduced cellodextrines by nine cellulases. Eur. J. Biochem. 217, pp. 947 - 953.

Schülein, M., Tikhomirov, D. F., Schou, C. 1993. *Humicola insolens* alkaline cellulases. In: Suominen, P. and Reinikainen, T. (Eds.). *Trichoderma reesei* cellulases and other hydrolases. Foundation for biotechnical and industrial fermentation research. Helsinki. Vol. 8, pp. 109 - 116.

Shannon, J. C. and Garwood, D. L. 1984. Genetics and physiolggy of starch development. In: Whistler, R. L., Bemiller, J. N. and Paschall, E. F. (Eds.) Starch: chemistry and technology. Academic Press, London, pp. 25 - 86.

Shen, H., Meinke, A., Tomme, P., Damude, H. G., Kwan, E., Kilburn, D. G., Miller, R. C. Jr., Warren, R. A. J. and Gilkes, N. R. 1995. *Cellulomonas fimi* cellobiohydrolases. In: Saddler, J.N. and Penner, M.H. (Eds.) Enzymatic degradation of insoluble carbohydrates. American Chemical Society. Washington, USA. Vol 618, pp. 174 - 196.

Shortle, D., Grisafi, P., Bencovic, S. J. and Botstein, D. 1982. Gap misrepair mutagenesis: Efficient site-directed induction of transition, transversion, and frameshift mutations *in vitro*. Proc. Natl. Acad. Sci. U.S.A.. 79, pp. 1588 - 1592.

Siderovski, D. P., Matsuyama, T., Frigerio, E., Chui, S., Min, X., Erfle, H., Sumner-Smith, M., Barnett, R. and Mak, T. W. 1992. Random mutagenesis of the human immunodeficiency virus type-1 *trans*-activator of transcription (HIV-1 Tat). Nucl. Acid Res. 20, pp. 5311 - 5320.

Sigurskjold, B. W., Berland, C. R. and Svensson, B. 1994a. Thermodynamics of inhibitor binding to the catalytic site of glucoamylase from *Aspergillus niger* determined by displacement titration calorimetry. Biochemistry 33, pp. 10191 - 10199.

Sigurskjold, B. W., Svensson, B., Williamson, G. and Driquez, H. 1994b. Thermodynamics of ligand binding to the starch-binding domain of glucoamylase from *Aspergillus niger*. Eur. J. Biochem. 225, pp. 133 - 141.

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

Somogyi, M. 1952. Notes on sugar determination. J. Biol. Chem. 195, pp. 19 - 23.

Søgaard, M., Kadziola, A., Haser, R. and Svensson, B. and 1993. Site-directed mutagenesisof His93, Asp180, Glu205, His290, and Asp291 at the active site and Trp279 at the raw starch binding site in barley α-amylase. J. Biol. Chem. 268, pp. 22480 - 22484.

Spezio, M.,Wilson, D. B. and Karplus, P. A. 1993. Crystal structure of the catalytic domain of a thermophilic endocellulase. Biochemistry 32, pp. 9906 - 9916.

Sprey B. and Lambert, C. 1983. Titration curves of cellulases from *Trichoderma reesei*: Demonstration of a cellulase-xylanase-β-glucosidase complex. *FEMS Microbiol. Lett.* 18, pp. 217 - 222.

Sprey B. and Bochem, H.-P. 1993. Formation of cross-fractures in cellulose microfibril structure by an endoglucanases-cellobiohydrolase complex from *Trichoderma reesei*. FEMS Microbiol. Lett. 106, pp. 239 - 244.

Srisodsuk M., Reinikainen, T., Penttilä , M., Teeri, T. T. 1993. Role of the Interdomain Linker Peptide of *Trichoderma reesei* Cellobiohydrolase I in Its Interaction with Crystalline Cellulose. J. Biol. Chem. 268, pp. 20756 - 20761.

Srisodsuk, M. 1994. Mode of action of *Trichoderma reesei* cellobiohydrolase I on crystalline cellulose. Ph.D. Thesis. University of Helsinki. VTT Publications 188. Technical Research Centre of Finland, Espoo, Finland. 107 p.

Stern, A. S. and Podlaski, F. J. 1993. Increasing the Antigen Binding Capacity of Immobilized Antibodies. In: Techniques in Protein Chemistry IV. Academic Press. Pp. 353 - 360.

Strynadka, N. C. J. and James, M. N. G. 1991. Lysozyme revisited: Crystallographic evidence for distortion of an N-acetylmuramic acid residue bound in site D. J. Mol. Biol. 220, pp. 401 - 424.

Ståhlberg, J. 1991. Functional organization of cellulases from *Trichoderma reesei*, Ph.D. Thesis. University of Uppsala. Sweden. 45 p.

Ståhlberg, J., Johansson, G., Pettersson, G. 1991. A new model for enzymatic hydrolysis of cellulose based on the two-domain structure of cellobiohydrolase I. Bio/Technology 9, pp. 286 - 290.

Stålbrand, H., Saloheimo, A., Vehmaanperä, J., Henrissat, B. and Penttilä, M. 1995. Cloning and expression in *Saccharomyces cerevisiae* of *Trichoderma reesei* β-mannanase gene containing a cellulose binding domain. Appl. Environ. Microb. 61, pp. 1090 - 1097.

Sugiyama, J., Persson, J. and Chanzy, H. 1991a. Combined infrared and electron diffraction study of polymorphism of native celluloses. Macromolecules 24, pp. 2461 - 2466.

Sugiyama, J., Vuong, R. and Chanzy, H. 1991b. An electron diffraction study on the two phases occurring in native cellulose from algal cell wall. Macromolecules 24, pp. 4168 - 4175.

Sugiyama, J., Hayashi, N., Wada, M and Okano, T. 1993. Morphology and structure of crystalline cellulose. In Suominen, P. and Reinikainen, T. (Eds.). *Trichoderma reesei* cellulases and other hydrolases. Foundation for biotechnical and industrial fermentation research. Helsinki. Vol 8, pp.15 - 23.

Sumner, J.B. 1921. Dinitrosalicylic: a reagent for the estimation of sugar in normal and diabetic urine. J. Biol. Chem. 47, pp. 5 - 9.

Suominen, I., Karp, M., Lautamo, J., Knowles, J. and Mäntsälä, P. 1987. Thermostable alpha amylase of *Bacillus stearothermophilus*: cloning, expression, and secretion by *Escherichia coli*. In: Chaloupka, J. and Krumphanzl, V. (Eds.) Extracellular Enzymes of microorganisms. Plenum Press. New York. pp. 129 - 137.

Suzuki, Y., Ito, N., Yuuki, T., Yamagata, H., and Udaka, S. 1989. Amino acid residues stabilizing a *Bacillus* α-amylase against irreversible thermoinactivation. J. Biol. Chem. 32, pp. 18933 - 18938.

Svensson, B., Larsen, K. and Gunnarsson, A. 1986. Characterization of a glucoamylase G2 from *Aspergillus niger*. Eur. J. Biohem. 154, pp. 497 - 502.

Svensson, B., Jespersen, H. M, Sierks, M. R. and McGregor, E. A. 1989. Sequence homology between putative raw-starch binding domains from different starch-degrading enzymes. Biohem. J. 264, pp. 309 - 311.

Svensson, B., and Sierks, M. 1992. Roles of aromatic side chains in the binding of substrates, inhibitors, and cyclomalto-oligosaccharides to the glucoamylase from *Aspergillus niger* probed by perturbation difference spectroscopy, chemical modification, and mutagenesis. Carboh. Res. 227, pp. 29 - 44.

Svensson, B. and Søgaard, M. 1993. Mutational analysis of glycosylase function. J. Biotechnol. 29, pp. 1 - 37.

Svensson, B. 1994. Protein engineering in the α-amylase family: catalytic mechanism, substrate specificity, and stability. Plant Mol. Biol., 25, pp. 141 - 157.

Svensson, B., Frandsen, T. P., Matsui, I., Juge, N., Fierobe, H.-P., Stoffer, B. and Rodenburg, K. W. 1995. Mutational analysis of catalytic mechanism and specificity in amylolytic enzymes. In: Petersen, S. B., Svensson, B. and Pedersen, S. (Eds.) Carbohydrate bioengineering. Progress in biotechnology. Elsevier, The Netherlands. Vol 10, pp. 125 - 145.

Takahashi, T, Kato, K., Ikegami, Y. and Irie, M. 1985. Different behaviour towards raw starch of three forms of glucoamylase from *Rhizopus* Sp. Biochem. J. (Tokyo) 98, pp. 663 - 671.

Takase, K., Matsumoto, T., Mizuno, H. and Yamane, K. 1992. Site-directed mutagenesis of active site residues in *Bacillus subtilis* alpha-amylase. Biochim. Biophys. Acta 1120, pp. 281 - 288.

Tanaka, Y., Tao, W., Blanchard, J. and Hehre, E. J. 1994. Transition state structures for the hydrolysis of α-D-glucopyranosyl fluoride by retaining and inverting reactions of glycosylases. J. Biol. Chem. 269, pp. 32306 - 32312.

Taylor, G. R. and Logan, P. W. 1995. The polymerase chain reaction: new variations on an old theme. Curr. Opin. Biotech. 6, pp. 24 - 29.

Taylor, J. S., Teo, B., Wilson, D. B. and Brady, J. W. 1995. Conformational modeling of substrate binding to endoglucanase E2 from *Thermomonospora fusca*. Protein Eng. 8, pp. 1145 - 1152.

Teeri, T. T. 1987. The cellulolytic enzyme system of *Trichoderma reesei*. Molecular cloning, characterization and expression of the cellobiohydrolase genes. Ph.D. Thesis. University of Helsinki. VTT Publications 38. Technical research centre of Finland, Espoo, Finland. 52 p. + app. 41 p.

Teeri, T. T., Kumar, V., Lehtovaara, P., Knowles, J. K. C. 1987. Construction of cDNA libraries by blunt end ligation: High-frequency cloning of long cDNAs from filamentous fungi. Anal. Biochem. 164, pp. 60 - 67.

Teeri, T. T., Koivula, A., Linder, M., Reinikainen, T., Ruohonen, L., Srisodsuk, M., Claeyssens, M. and Jones, T. A. 1995. Modes of action of two *Trichoderma reesei* cellobiohydrolases. In: Petersen, S. B., Svensson, B. and Pedersen, S. (Eds.). Carbohydrate bioengineering. Progress in biotechnology. Elsevier, The Netherlands. Vol 10., pp. 211 - 224.

Teleman, A., Koivula, A., Reinikainen, T., Valkeajärvi, A., Teeri, T. T., Drakenberg, T. and Teleman, O. 1995. Progress-curve analysis shows that the glucose inhibits the cellotriose hydrolysis catalysed by cellobiohydrolase II from *Trichoderma reesei*. Eur. J. Biochem. 231, pp. 250 - 258.

Te'o, V. S., Saul, D. J. and Bergquist, P. L. 1995. CelA, another gene coding for a multidomain cellulase from the extreme thermophile *Caldocellum saccharolyticum*. Appl. Microbiol. Biotechnol. 43, pp. 291 - 296.

Thoma, J. A. and Crook, C. 1982. Subsite mapping of enzymes. Double inhibition studies. Eur. J. Biochem. 122, pp. 613 - 618.

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

van Tilbeurgh, H., Pettersson, G., Bhikhabai, R., de Boeck, H.G. and Claeyssens, M. 1985. Studies of the cellulolytic system of *Trichoderma reesei* QM 9414. Reaction specificity and thermodynamics of interactions of small substrates and ligands with the 1,4-β-glucan cellobiohydrolase II. Eur. J. Biochem. 148, pp. 329 - 334.

van Tilbeurgh, H. and Claeyssens, M. 1985. Detection and differentiation of cellulase components using low molecular mass fluorogenic substrates. FEBS Lett. 187, pp. 283 - 288.

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

van Tilbeurgh, H., Loontiens, F. G., de Bruyne. C. K. and Claeyssens, M. 1988. Fluorogenic and chromogenic glycosides as substrates and ligands of catbohydrases. Methods in Enzymology 160, pp. 45 - 59.

van Tilbeurgh, H., Loontiens, F., Engelborgs, Y. and Claeyssens, M. 1989. Studies of cellulolytic system of *Trichoderma reesei* QM 9414. Binding of small ligands to the 1,4-β-D-glucan cellobiohydrolase II and influence of glucose on their affinity. Eur. J. Biochem. 184, pp. 553 - 559.

Tomme, P., McRae, S., Wood, T. M., Claeyssens, M. 1988a. Chromatographic separation of cellulolytic enzymes. Methods in Enzymol. 160, pp. 187 - 192.

Tomme, P., van Tilbeurgh, H., Pettersson, O., van Damme, J., Vandekerckhove, J., Knowles, J., Teeri, T. and Claeyssens, M. 1988b. Studies of the cellulolytic system of *Trichoderma reesei* QM 9414. Analysis of domain function in two cellobiohydrolases by limited proteolysis, Eur. J. Biochem. 170, pp. 575 - 581.

Tomme, P. 1991. Cellulasen uit *Trichoderma reesei* en *Clostridium thermocellum*: Domeinstructuur, chemische modificaties en mutagenese, Ph.D. thesis. University of Ghent. Belgium.

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

Tomme, P., Warren, R. A. J., Miller, R. C. Jr., Kilburn, D. G. and Gilkes, N. R. 1995b. Cellulose-binding domains: classification and properties. In: Saddler, J. N. and Penner, M. H. (Eds.) Enzymatic degradation of insoluble carbohydrates. American chemical society. Washington, USA. Vol 618, pp. 142 - 163.

Toone, E. J. 1994. Structure and energetics of protein-carbohydrate complexes. Curr. Opinion Struc. Biol. 4, pp. 719 - 728.

Towbin, H., Staehelin, T. and Gordon, J. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. U.S.A. 76, pp. 4350 - 4354.

Vardanis, A. and Finkelman, M. 1981. A radiometric microassay for cellulase activity. Anal. Biochem. 115, pp. 78 - 80.

Varghese, J. N., McKimm-Breschkin, J. L., Caldwell, J. B., Kortt, A. A. and Colman, P. M. 1992. The structure of the complex between influenza virus neuraminidase and sialic acid, the viral receptor. Proteins: Structure, Function and Genetics 14, pp. 327 - 332.

Vihinen, M. and Mäntsälä, P. 1989. Microbial amylolytic enzymes. Crit. Rev. Biochem. Mol. Biol. 24, pp. 329 - 418.

Vihinen, M. and Mäntsälä, P. 1990. Characterization of a thermostable *Bacillus stearothermophilus* α-amylase. Biotechnol. Appl. Biochem. 12, pp. 427 - 435.

Vihinen, M., Helin, S. and Mäntsälä, P. 1991. Site-directed mutagenesis of putative active site residues of *Bacillus stearothermophilus* α-amylase. Mol. Eng. 1, pp. 267 - 273.

Vihinen, M., Peltonen, T., Iiitiä, A., Suominen, I. and Mäntsälä, P. 1994. C-terminal truncation of a thermostable *Bacillus stearothermophilus* α-amylase. Protein Eng. 7, pp. 1255 - 1259.

Violet, M. and Meunier, J.-C. 1989. Kinetic studies of the irreversible thermal inactivation of *Bacillus licheniformis* α-amylase. Biochem. J. 263, pp. 665 - 670.

Vršanská, M. and Biely, P. 1992. The cellobiohydrolase I from *Trichoderma reesei* QM 9414: Action on cello-oligosaccharides. Carboh. Res. 227, pp. 19 - 27.

Waffenschmidt, S. and Jaenicke, L. 1987. Assay of reducing sugars in the nanomole range with 2,2'-bicinchoninate. Anal. Biochem. 165, pp. 337 - 340.

Walker, L. P., Wilson, D. B. and Irwin, D. C. 1990. Measuring fragmentation of cellulose by *Thermomonospora fusca* cellulase. Enzyme Microb. Technol. 12, 378 - 386.

Walker, L. P., Wilson, D. P., Irwin, D. C., McQuire, C. and Price, M. 1992. Fragmentation of cellulose by the major *Thermomonospora fusca* cellulases, *Trichoderma reesei* CBHI and their mixtures. Biotechnol. Bioeng. 40, pp. 1019 - 1026.

Ward, M., Wu, Shan, Dauberman, J., Weiss, G., Larenas, E., Bower, B., Rey, M., Clarkson, K., Bott, R. 1993. Cloning, sequence and preliminary analysis of small, high pI endoglucanase EGIII from *Trichoderma reesei*. In: Suominen, P. and Reinikainen, T. (Eds.). *Trichoderma reesei* cellulases and other hydrolases Foundation for biotechnical and industrial fermentation research, Helsinki. Vol 8, pp. 153 - 158.

Wetlaufer, D. B. 1962. Ultraviolet spectra of proteins and amino acids. In: Amfinsen, C. B. Jr., Anson, M. L., Bailey, K. and Edsall, J. T. (Eds.). Adv. Prot. Chem. Academic Press, New York. Vol 17, pp. 303 - 390.

Whistler, R. L. and Daniel, J. R. 1985. Carbohydrates. In: Fennema, O.R. (Ed.) Food Chemistry. Marcel Dekker, New York. Pp. 69 - 137.

White, A., Withers, S., Gilkes, N. and Rose, D. R. 1994. Crystal structure of the catalytic domain of the β-1,4-glycanase Cex from *Cellulomonas fimi*. Biochemistry 33, pp. 12546 - 12552.

Williamson, G., Belshaw, N. J. and Williamson, M. P. 1992. O-glycosylation in *Aspergillus* glucoamylase. Conformation and role in binding. Biochem. J. 287, pp. 423 - 428.

Withers, S. G., Dombroski, D., Berven, L. A., Kilburn, D. G., Miller, R. C. Jr., Warren, R. A. J. and Gilkes, N. R. 1986. Direct ^{1}H NMR determination of the stereochemical course of hydrolases catalysed by glucanase components of the cellulase complex. Biochem. Biophys. Res. Commun. 139, pp. 487 - 494.

Withers, S. G. 1995. Probing of glycosidase active sites through labeling, mutagenesis and kinetic studies. In: Petersen, S. B., Svensson, B. and Pedersen, S. (Eds.) Carbohydrate bioengineering. Progress in biotechnology. Elsevier, The Netherlands. Vol 10, pp. 97 - 111.

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

Wilson, D. B., Spezio, M., Irwin, D., Karplus, A. and Taylor, J. 1995. Comparison of enzymes catalyzing the hydrolysis of insoluble polysaccharides. In: Saddler, J.N. and Penner, M.H. (Eds.) Enzymatic degradation of insoluble carbohydrates. American chemical society, Washington, USA. Vol 618, pp. 1 - 12.

Wong, K. K., Tan, L. U. L. and Saddler, J. N. 1988. Multiplicity of β-1,4-xylanase in microorganisms: functions and applications. Microb. Rev. 52, pp. 305 - 317.

Wood, T. M. and Bhat, K. M. 1988. Methods for measuring cellulase activities. Methods Enzymol. 160, pp. 87 - 112.

Wood, T. M., McCrae, S. I. and Bhat, K. M. 1989. The mechanism of fungal cellulase action. Synergism between enzyme components of *Penicillium pinophilum* cellulase in solubilizing hydrogen bond-ordered cellulose. Biochem. J. 260, pp. 37 - 43.

Wood, T. M. and Garcia-Campayo, V. 1990. Enzymology of cellulose degradation. Biodegradation 1, pp. 147 - 161.

Xu, G.-Y., Ong, E., Gilkes, N. R., Kilburn, D. G., Muhandiran, D. R., Brandeis, M., Carver, J. P., Kay, L. E. and Harvey, T. S. 1995. Solution structure of a cellulose-binding domain from *Cellulomonas fimi* by nuclear magnetic resonance spectroscopy. Biochemistry, 34, pp. 6993 - 7009.

Zakour, R. A. and Loeb, L. A. 1982. Site-specific mutagenesis by error-directed DNA synthesis. Nature 295, pp. 708 - 710.

Zobel, H. F. 1992. Starch granule srtucture. In: Alexander, R. J. and Zobel, H. F. (Eds.) Developments in carbohydrate chemistry. American association of cereal chemistry. Minnesota, USA. Pp. 1 - 36.

Appendices of this publication are not included in the PDF version.
Please order the printed version to get the complete publication
(http://www.inf.vtt.fi/pdf/publications/1996/)