FEMS Microbiology Reviews 13 (1994) 25–58

Published by Elsevier

FEMSRE 00341

The biological degradation of cellulose

Pierre Béguin * and Jean-Paul Aubert

Unité de Physiologie Cellulaire and URA 1300 CNRS, Département des Biotechnologies, Institut Pasteur, 28 rue du Dr. Roux, 75724 Paris Cedex 15, France

(Received 7 July 1993; accepted 8 September 1993)

Abstract: Cellulolytic microorganisms play an important role in the biosphere by recycling cellulose, the most abundant carbohydrate produced by plants. Cellulose is a simple polymer, but it forms insoluble, crystalline microfibrils, which are highly resistant to enzymatic hydrolysis. All organisms known to degrade cellulose efficiently produce a battery of enzymes with different specificities, which act together in synergism. The study of cellulolytic enzymes at the molecular level has revealed some of the features that contribute to their activity. In spite of a considerable diversity, sequence comparisons show that the catalytic cores of cellulases belong to a restricted number of families. Within each family, available data suggest that the various enzymes share a common folding pattern, the same catalytic residues, and the same reaction mechanism, i.e. either single substitution with inversion of configuration or double substitution resulting in retention of the β-configuration at the anomeric carbon. An increasing number of three-dimensional structures is becoming available for cellulases and xylanases belonging to different families, which will provide paradigms for molecular modeling of related enzymes. In addition to catalytic domains, many cellulolytic enzymes contain domains not involved in catalysis, but participating in substrate binding, multi-enzyme complex formation, or possibly attachment to the cell surface. Presumably, these domains assist in the degradation of crystalline cellulose by preventing the enzymes from being washed off from the surface of the substrate, by focusing hydrolysis on restricted areas in which the substrate is synergistically destabilized by multiple cutting events, and by facilitating recovery of the soluble degradation products by the cellulolytic organism. In most cellulolytic organisms, cellulase synthesis is repressed in the presence of easily metabolized, soluble carbon sources and induced in the presence of cellulose. Induction of cellulases appears to be effected by soluble products generated from cellulose by cellulolytic enzymes synthesized constitutively at a low level. These products are presumably converted into true inducers by transglycosylation reactions. Several applications of cellulases or hemicellulases are being developed for textile, food, and paper pulp processing. These applications are based on the modification of cellulose and hemicellulose by partial hydrolysis. Total hydrolysis of cellulose into glucose, which could be fermented into ethanol, isopropanol or butanol, is not yet economically feasible. However, the need to reduce emissions of greenhouse gases provides an added incentive for the development of processes generating fuels from cellulose, a major renewable carbon source.

Key words: Cellulose; Cellulolytic microorganisms; Cellulases; Multidomain proteins; Cellulosome

Introduction

Cellulose is the major carbohydrate synthesized by plants. Therefore, the degradation of cellulosic biomass represents an important part of the carbon cycle within the biosphere. For the same reason, treatment of cellulose by cellulolytic enzymes for practical purposes has attracted the continuing interest of biotechnologists. However, improvement of cellulose hydrolysis is impeded by a lack in basic knowledge of the process. The

* Corresponding author. Tel.: (1) 45 68 88 19; Fax (1) 45 68 87 90.

SSDI 0168-6445(93)E0066-S

combination of approaches based on molecular genetics, biochemistry, and structural biology has led to a fast accumulation of results about the structure and function of cellulolytic enzymes. This review attempts to summarize recent developments of the topic. Further background information can be found in several reviews [1–9].

Structure of cellulose

Cellulose is a linear polymer made of glucose subunits linked by β-1,4 bonds (Fig. 1). Each glucose residue is rotated by 180° relative to its neighbours, so that the basic repeating unit is in fact cellobiose. Chain length varies between 100 and 14 000 residues.

Cellulose chains form numerous intra- and intermolecular hydrogen bonds, which account for the formation of rigid, insoluble microfibrils. Microfibrils range in lateral dimension from 3–4 nm in higher plants up to 20 nm for the microfibrils of the alga *Valonia macrophysa*, which contain up to several hundred cellulose chains. The chains are oriented in parallel and form highly ordered, crystalline domains interspersed by more disordered, amorphous regions. Raman spectroscopy and solid-state ^{13}C-NMR have shown that crystalline cellulose occurs in various forms. The native, crystalline form of cellulose has a structure designated as type I, which can be converted into type II by alkali treatment. The two types differ in their intrachain hydrogen bonding pattern. Furthermore, most native celluloses are composed of two slightly different forms of type I cellulose, termed I_α and I_β, which differ in their intermolecular hydrogen bonding pattern. The two forms occur in different proportions, depending on the source of the cellulose. [4,10–12].

Depending on origin and pretreatment, the degree of crystallinity of cellulose can vary from 0% for amorphous, acid-swollen cellulose, to nearly 100% for the cellulose isolated from *Valonia macrophysa* [13]. Cellulose from cotton is about 70% crystalline [14], and the degree of crystallinity of most commercial celluloses varies between 30 and 70%.

In contrast to starch, which serves as a storage polymer for glucose, the role of cellulose is exclusively structural. The high tensile strength of cellulose enables plant cells to withstand osmotic pressure and is responsible for the resistance of plants to mechanical stress. The mechanical strength of cellulose is particularly obvious in the

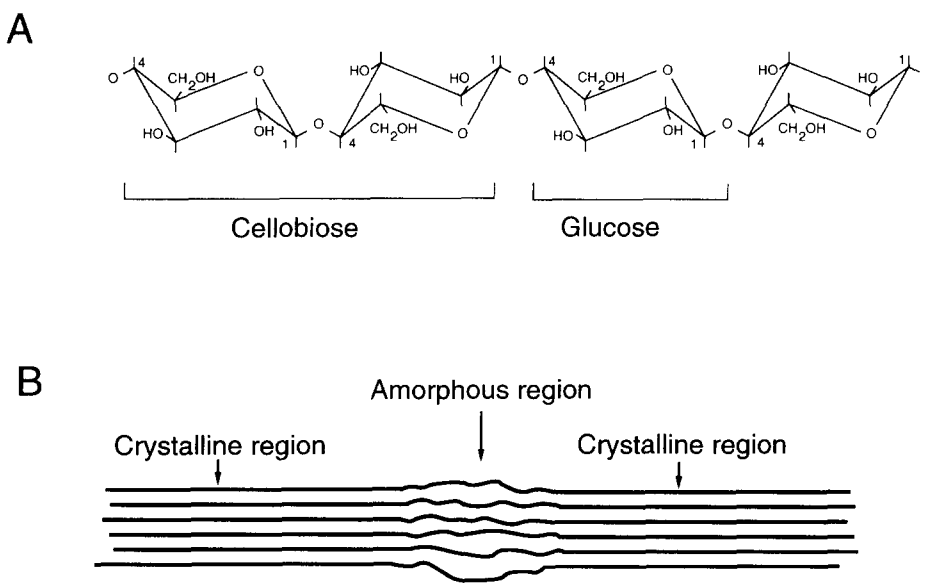

Fig. 1. Structure of cellulose. **(A)** β-glucosidic bonds. **(B)** Schematic structure of a fibril. Reprinted from [2], with permission.

case of wood and textile fibres (cotton, ramie, flax), which consist of the walls of elongated, empty cells.

In the secondary wall of plant cells, cellulose forms several sheets in which microfibrils are organized in parallel, each sheet having a different orientation. The microfibrils are usually embedded in a matrix of hemicellulose and lignin (Fig. 2). Hemicellulose is composed of complex carbohydrate polymers, with xylans and glucomannans as the main components. In most xylans, the xylan backbone carries acetyl, methylglucuronyl and arabinofuranosyl side chains in varying proportions [15]. The arabinofuranosyl groups can be further esterified by aromatic acids such as ferulic and *p*-coumaric acid [16], which are thought to participate in lignin–hemicellulose cross-links involving ether linkages [17]. Lignin is a highly branched, random polymer generated by the free-radical condensation of aromatic alcohols [18].

The degradation of hemicellulose is performed by a complex set of enzymes hydrolyzing the xylan or glucomannan backbone and the various bonds of the side chains [19–22]. Lignin is highly resistant to biodegradation and protects cellulose and hemicellulose against enzymatic hydrolysis. Enzymatic attack of lignin involves radical oxidation by peroxidases [18,23].

The diversity and heterogeneity of cellulosic substrates contribute to a large extent to the difficulty of enzymological studies. Practically, a variety of substrates is available, whose recalcitrance to hydrolysis is directly correlated with their degree of similarity with natural cellulose.

Pure cellulose is commercially available in several forms (cotton, filter paper, Avicel). These forms are generally used to assess the efficiency of complete cellulase systems. However, their physical heterogeneity (degree of crystallinity, available surface area, pore size) complicates detailed enzymological studies. Furthermore, they are not adequate to study incomplete cellulase systems or individual cellulolytic components, which are only weakly active on these substrates.

Amorphous forms such, as acid-swollen cellulose and soluble carboxymethylcellulose (CMC) are frequently used for assays, owing to their fast rate of hydrolysis. However, in the case of CMC, the degree of substitution negatively influences the rate of hydrolysis, resulting in a rapid deviation of hydrolysis from linear kinetics. For the same reason, CMC is a poor substrate for exo-acting enzymes.

Several low molecular mass substrates can also be used for enzymological studies. Soluble oligodextrins comprising three to six glucosyl residues can be obtained by partial hydrolysis of

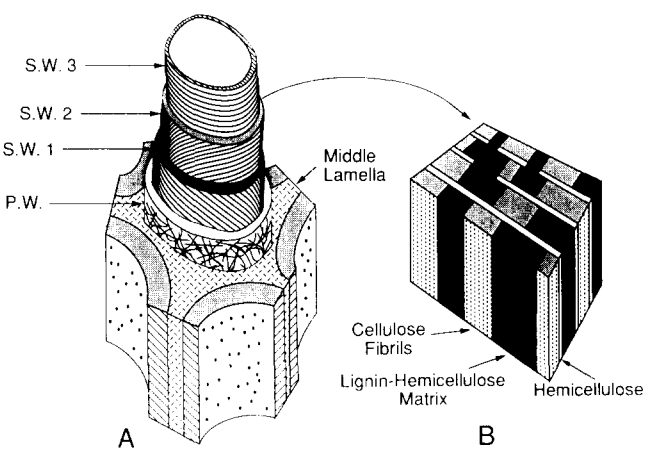

Fig. 2. (**A**) Cutaway view showing the organization of the cell wall layers composing woody fibres. (**B**) Probable relationship of lignin and hemicellulose to the cellulose microfibrils in the secondary walls. The diameter of each cell is approximately 25 μm. P.W., primary cell wall; S.W. 1–S.W. 3, secondary cell walls. Redrawn from [250].

cellulose [24]. More convenient are chromogenic and fluorogenic substrates, in which coloured or fluorescent compounds are linked to the anomeric carbon of cellobiose or cellodextrins by a β-glucosidic bond [25]. The most frequently used compounds, *p*-nitrophenyl-β-D-cellobioside and methylumbelliferyl-β-D-cellobioside (MUC) are commercially available. However, homologues with a higher degree of polymerization have proved useful in characterizing the specificity of enzymes and in defining the topology of substrate binding sites, in particular the number of subsites and the position of the cleavage site [26,27].

It is worth noting that the action of enzymes on native cellulose can be assessed by a number of different of parameters. Catalytic activity is usually assayed by measuring the release of soluble reducing sugars, which are expressed as glucose equivalents. Alternatively, residual insoluble cellulose can be determined gravimetrically [28], colorimetrically [29], or by turbidimetry [30]. Comparatively few studies (e.g. [31]) deal with alterations of the physico-chemical properties of the insoluble substrate, such as crystallinity index, degree of polymerization, alteration of available surface area, or small particle formation, in spite of the considerable influence that these parameters have on hydrolysis kinetics. In this respect, electron microscopy can be particularly revealing by showing sites of cellulase adsorption [32], patterns of erosion [33], and alterations of fibre or microfibril structure [34–37].

Cellulolytic microorganisms

Cellulolytic microorganisms are found among extremely variegated taxonomic groups. Most belong to eubacteria and fungi, but anaerobic, cellulose-degrading protozoa have also been identified in the rumen [38]. The avocado fruit [39] and the slime mold *Dictyostelium discoideum* [40,41] also produce cellulolytic enzymes, but the major function of these enzymes is thought to be related to the maturation of fruits and spores, respectively.

Cellulolytic microorganisms can be found in all biota where cellulosic waste accumulates. They usually occur in mixed populations comprising cellulolytic and non-cellulolytic species, which often interact synergistically. These interactions lead to the complete degradation of cellulose, which is ultimately converted into carbon dioxide and water under aerobic conditions, and into carbon dioxide, methane, and water under anaerobic conditions.

Aerobic degradation

Topsoil is the most important aerobic biota where cellulosic wastes accumulate. Among these, wood is the component most slowly attacked by cellulolytic microorganisms, due to its high content of lignin. Complete lignin degradation is an oxidative process which is only performed by a few microorganisms such as the white rot fungus *Phanerochaete chrysosporium*. However, a larger variety of organisms, in particular Actinomycetes [42], are capable to effect partial delignification in order to get access to the cellulosic substrate. The fungus *Trichoderma reesei*, an aerobic and highly cellulolytic Deuteromycete, is probably the microorganism whose cellulase system has been most thoroughly investigated. Among aerobic, cellulolytic soil bacteria, several species belonging to the genera *Cellulomonas, Pseudomonas* (*Cellvibrio*), *Thermomonospora* and *Microbispora* have been studied in detail.

Anaerobic degradation

Anaerobic cellulolytic organisms are the first link of a complex nutritional chain converting polymeric carbohydrates into methane and carbon dioxide. In a first step, glucose derived from cellulose is fermented by cellulolytic and associated saccharolytic organisms, yielding hydrogen, carbon dioxide, and various organic compounds including acids and alcohols. In a second step, these compounds are converted into acetate and carbon dioxide. This step is thermodynamically unfavourable, unless the partial pressure of hydrogen is maintained at very low levels. Excess reducing equivalents must therefore be transferred to appropriate acceptors such as sulfate or carbon dioxide. Sulfate is reduced by sulfate re-

ducers into hydrogen sulfide. Carbon dioxide is reduced into acetate by acetogens, and into methane by methanogens. As a last step, acetate is converted into methane and carbon dioxide by methanogens.

Association with methanogens is beneficial for the other microorganisms of the nutritional chain by preventing acidification by acetate and by removing excess hydrogen. Thus, as compared to a pure culture of *Clostridium thermocellum*, a co-culture of *C. thermocellum* with *Methanobacterium thermoautotrophicum* on cellulose showed a significantly reduced lag phase and an increased rate of growth during the early stages of fermentation [43]. Likewise, the anaerobic, cellulolytic fungus *Neocallimastix frontalis* formed a stable co-culture with *Methanobrevibacter smithii*, and the co-culture displayed a remarkable increase in extracellular cellulase activity relative to a culture containing the fungus alone [44,45].

Anaerobic degradation of cellulose occurs in a variety of anaerobic biota, such as manure, compost, sludges of waste water treatment plants and marine or freshwater sediments. In addition, the hydrolysis of cellulose by anaerobic microorganisms present in the rumen and gastro-intestinal tract plays an essential role in the nutrition of herbivorous animals. The rumen flora harbours a variety of anaerobic microorganisms utilizing plant cell wall polysaccharides. Some of the most extensively studied cellulolytic and xylanolytic rumen bacteria are the Gram-negative species *Fibrobacter* (formerly *Bacteroides*) *succinogenes* and *Butyrivibrio fibrisolvens* and the Gram-positive species *Ruminococcus albus* and *Ruminococcus flavefaciens*. Rumen cellulolytic fungi include anaerobic chytridiomycetes, in particular *N. frontalis* and *Neocallimastix patriciarum*. Several anaerobic ciliates utilizing cellulosic plant fragments have been identified in the rumen, e.g. those belonging to the genera *Diplodinium* and *Eudiplodinium* [46]. Since protozoa ingest cellulose by phagocytosis, the question arises whether the hydrolytic enzymes are produced by the protozoon or by cellulolytic bacteria coingested with the substrate. At least in the case of *Eudiplodinium maggii*, the cellulolytic enzymes appear to be produced by the ciliate [38].

Partially and truly cellulolytic microorganisms

Only few microorganisms produce a complete set of enzymes capable of degrading native cellulose efficiently. However, a larger number of species produce incomplete cellulase systems. The significance of these systems is not quite clear. Restricted cellulolytic action may facilitate the access of phytopathogens to plant tissues, and may contribute to the leakage of plant cell sap. The cellulolytic enzymes produced by saprophytic bacteria may cooperate with the cellulase systems of true cellulolytic organisms, with which they are associated. Furthermore, incomplete cellulase producers could be evolutionary intermediates that may become truly cellulolytic once they acquire the required set of genes. Indeed, sequence analysis (see below) suggests that in several cases, cellulase genes can spread through horizontal transfer, which should considerably speed up the evolution of cellulase systems.

Biochemical properties of cellulase systems

Since cellulose cannot get into the cells, cellulolytic enzymes are by necessity secreted into the medium or bound to the outside surface of cellulolytic microorganisms. Furthermore, cellulase systems generally display a set of typical properties. The systems contain a multiplicity of enzyme components showing a marked synergism against crystalline cellulose. These components often possess a substrate-binding site independent from the catalytic site, and are often associated with each other and with the surface of cellulolytic microorganisms.

Sequence analysis of cellulase genes and the biochemical characterization of wild-type and truncated enzymes have shown that many cellulolytic enzymes are multifunctional proteins composed of distinct domains which can be arranged in various combinations. Sequence similarities define families of domains having related structures. Whenever investigated, domains of the same family appear to share similar biochemical properties [27]. Thus, much can be learned from the detailed study of a few paradigms that are

representative of each family. In the following, we will review the structural elements responsible for the properties typical of cellulolytic enzymes and discuss how they may help cellulases to cope with the constraints determined by the physical nature of the substrate.

Catalytic domains

Families of cellulases

All cellulolytic microorganisms have evolved a battery of enzymes having different specificities with respect to endo/exo mode of action, activity towards amorphous or crystalline regions, or preference for substrates of different chain length. Such a diversity may be needed, in part, to cope with the physical heterogeneity of the substrate. Moreover, the structure of cellulose changes during the process of degradation, requiring different enzymes at different times.

In 1989, Henrissat et al. [47] introduced a classification of 21 cellulolytic and xylanolytic enzymes into six families based on structural similarities detected by hydrophobic cluster analysis (HCA). Since then, the classification has been extended to a variety of glycosyl hydrolases comprising 482 sequences grouped in 45 families [48–50]. Among these, cellulases and xylanases are ordered into 11 families [50]. Several families comprise enzymes produced by a wide range of organisms, spanning different kingdoms. Furthermore, within the same family, the relatedness of enzymes does not necessarily reflect the phylogenic relatedness of the organisms that produce them. This suggests that cellulase and xylanase genes have spread to a significant extent by horizontal transfer across a wide range of organisms. The most clear-cut evidence for horizontal transfer is provided by the endoglucanase gene of *Cellulomonas uda*, whose G + C content is significantly lower than that of the bulk of *C. uda* DNA [51]. Since most of the catalytic domains can be combined with a variety of non-catalytic domains (see below), it is likely that early genes were transferred as modules encoding single domains, which were subsequently recombined to yield genes encoding multidomain proteins.

While being very useful, such molecular taxonomy studies point out some of the problems encountered by computer biologists. The first one is due to errors in the databases. New sequences not fitting in any of the above categories, or not matching strictly conserved regions present in enzymes of the same family, should be viewed with some caution. Indeed, at least three sequences which matched their expected consensus incompletely or did not match any other sequence were found to fit better after experimental data were corrected [52–58]. Several other sequences appear suspicious [59].

Another problem is to define a threshold of similarity below which proteins should be ordered into different families. Originally, the HCA-based classification was supposed to define families of proteins sharing the same polypeptide folding pattern. No significant difference in folding pattern has yet been found between members of the same family. However, even HCA may fail to detect a similar folding in proteins whose sequences are too divergent. An example is provided by endoglucanases related to the *Aspergillus aculeatus* endoglucanase Eg1, first classified in a distinct family. Recent reappraisal of HCA plots suggested a similarity with a large family of xylanases [60], which is confirmed by the highly similar three-dimensional structures of *A. aculeatus* endoglucanase and of *B. pumilus* xylanase [61].

A further difficulty lies in proteins carrying more than one (potential) catalytic domain (see below). For example, *Prevotella* (formerly *Bacteroides*) *ruminicola* endoglucanase [62,63] was classified in ref. [50] as belonging to a new cellulase family (family I or 26), but in fact, catalytic activity resides within a domain quite clearly belonging to the well-known family A (or 5).

Finally, as the classification is being extended to an ever widening set of proteins, some consistent nomenclature system has to be decided upon, perhaps through an international agreement at the level of IUB. The first alphabetical classification of cellulases has already been merged into a numerical classification of glycosyl hydrolases, but what will happen if glycosyl hydrolases eventually have to be merged with other classes of enzymes?

Conflicting nomenclatures certainly do not help clarify a situation which is already complex naturally.

Three-dimensional structure of cellulases and xylanases

Three-dimensional structures have now been determined by X-ray diffraction analysis for *T. reesei* cellobiohydrolase (CBH) II [64], *C. thermocellum* endoglucanase CelD [65], *Bacillus pumilus* xylanase [61], and *A. aculeatus* endoglucanase (family H) [61]. Several other 3-D structures should become available in the near future, since preliminary notes describing the preparation of crystals suitable for X-ray diffraction analysis have been published recently for *Humicola insolens* [66], for *T. reesei* CBH I [67] for *Cellulomonas fimi* exoglucanase/xylanase Cex [68] and *Pseudomonas fluorescens* var. *cellulosa* xylanase XynA [69]. Other crystal structures currently being studied include *R. flavefaciens* endoglucanase CelA (J. Thomson, personal communication), and *C. cellulolyticum* endoglucanase CelCCA (J.-P. Belaich, personal communication).

The folding pattern of *T. reesei* CBH II consists of a 7-stranded β-barrel. The active site is formed by two extended loops at the C-terminal end of the barrel. *C. thermocellum* CelD comprises an N-terminal domain composed of two β-sheets, and a large, C-terminal catalytic domain composed of 12 helices forming an α-barrel. Three of the loops connecting six of the helices on the same side of the barrel form the active site. The *B. pumilus* xylanase and the *A. aculeatus* endoglucanase both comprise two domains separated by a deep cleft forming the active site. The first domain consists of two antiparallel β-sheets, and the second domain contains an antiparallel β-sheet and an α-helix.

The geometry of the active site of the various enzymes provides an elegant explanation of their endo or exo specificity. In the case of the endoglucanases and of the xylanase, the substrate comes to lie in an open cleft, which can straddle cellulose or xylan molecules anywhere along the chain, in agreement with the endo mode of action of these enzymes. The exo mode of action of CBH II is explained by the fact that the active site forms an almost perfectly enclosed tunnel through which the cellulose chain has to be threaded from the non-reducing end. Modelling after the structure of CBH II suggests that in the case of bacterial endoglucanases related to CBH II the loops forming the reaction centre are too short to form a closed structure.

Reaction mechanisms

It is generally assumed that the hydrolysis reaction catalyzed by glycosidases, including cellulases and xylanases, proceeds via an acid-base mechanism involving two residues. The first residue acts as a general acid catalyst and protonates the oxygen of the osidic bond. The second residue acts as a nucleophile, which either interacts with the oxocarbonium intermediate (for retaining enzymes) or promotes the formation of an OH^- ion from a water molecule (for inverting enzymes) (Figs. 3 and 4). Reactions leading to retention of configuration involve a two-step mechanism, with a double inversion of configuration at the anomeric carbon, and the formation of an oxocarbonium intermediate [70]. The paradigm of this type of reaction is the mechanism of lysozyme [71] (Fig. 3). Reactions leading to inversion of configuration proceed via a single nucleophilic substitution (Fig. 4) [70].

The stereochemical course of hydrolysis has been determined for 17 β-1,4-glucanases and -xylanases belonging to six glycosyl hydrolase families [72–78]. Enzymes of family 5 (formerly A, five examples), 7 (formerly C, two examples), 10 (formerly F, two examples) and 11 (formerly G, two examples) proceed with retention of configuration. Enzymes of family 6 (formerly B, three examples) and 9 (formerly E, three examples) proceed with inversion of configuration.

Residues likely to act as proton donors have been identified by structural, biochemical or mutagenesis studies in *T. reesei* CBH II (Asp-221) [64], *C. thermocellum* endoglucanase CelD (Glu-555) [65,79,80], *B. pumilus* xylanase XynA (Glu-182) [61,81], *Bacillus polymyxa* and *Bacillus subtilis* endoglucanases [82], *Erwinia chrysanthemi* endoglucanase CelZ [83], *Clostridium cellulolyticum* endoglucanase CelCCA [84] and *C. thermocellum* endoglucanase CelC (Glu-140) [85].

In the case of *T. reesei* CBH I, chemical modification with Woodward's reagent K led to the identification of a polypeptide of 13–17 residues starting at Gly-124. On the basis of purported homology with lysozyme, Glu-126 was surmised to be the catalytic proton donor. However, since CBH I is unlikely to have the same structure as lysozyme, this speculation should be viewed with caution.

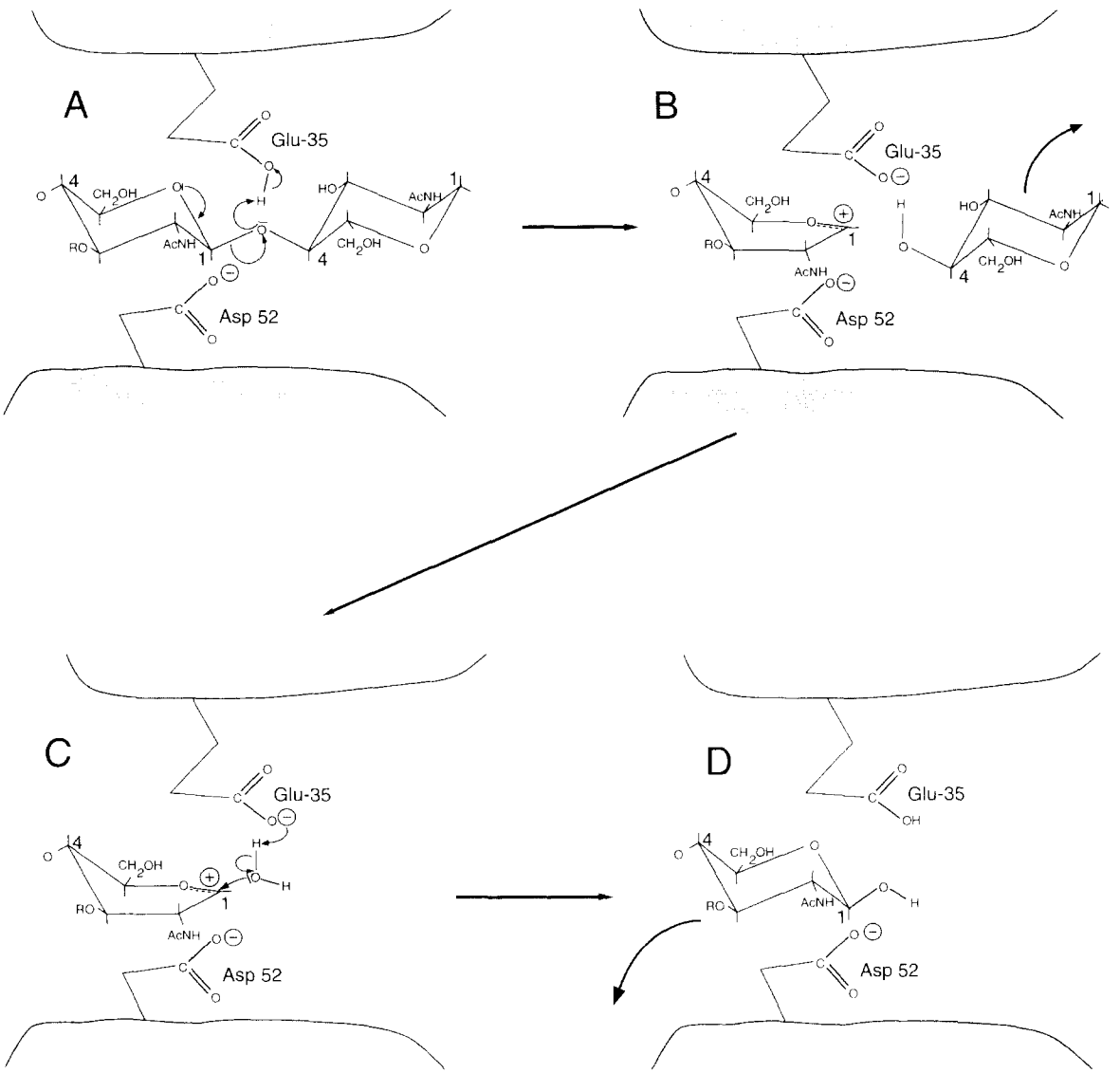

Fig. 3. Catalytic mechanism of lysozyme as a paradigm of β-glycanase acting with retention of configuration at the anomeric carbon. **(A)** Attack of the glycosidic bond between the acetylmuramate residue, on the left, and the acetylglucosamine residue, on the right, is started by the proton provided by the Glu-35 residue of the enzyme. **(B)** Splitting of the bond leads to the release of a fragment carrying a new non-reducing end group (carried by the acetylglucosamine residue), and to the formation of an oxocarbonium intermediate (carried by the acetylmuramate residue). The latter interacts with the negative charge borne by the Asp residue (whether the bond remains electrostatic or whether a covalent bond forms is debated [70]). **(C)** A water molecule provides an OH^- ion, which reacts with the oxocarbonium ion, and an H^+ ion, which regenerates the proton lost by the Glu-35 residue. **(D)** The fragment carrying the new anomeric carbon is released. Reprinted from [2], with permission.

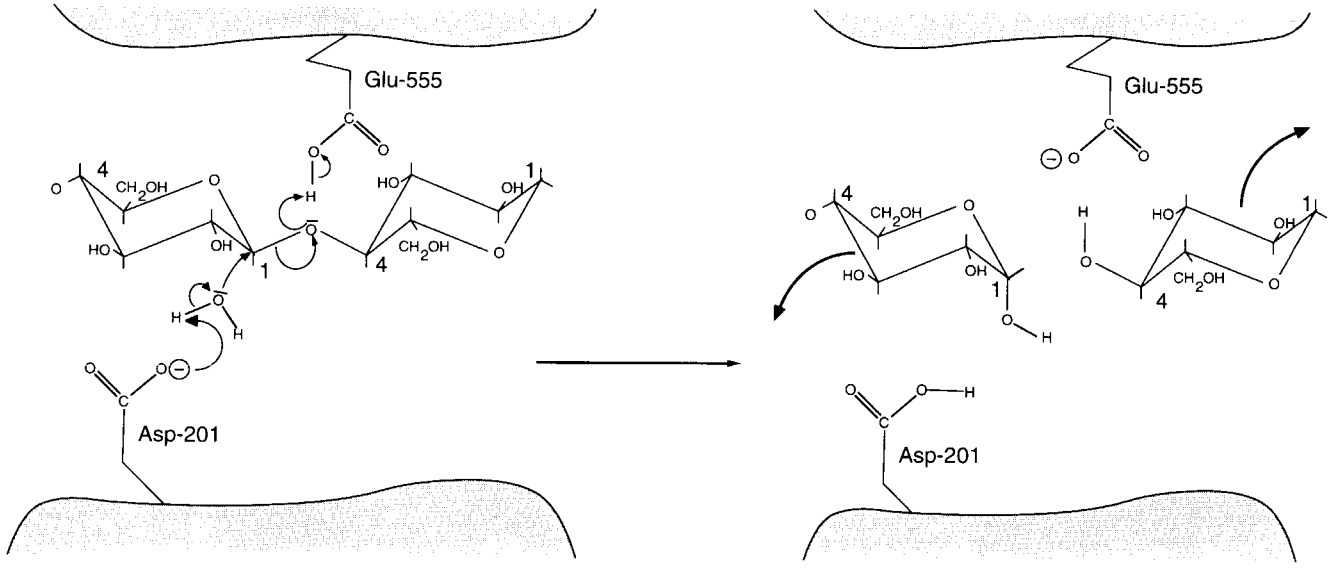

Fig. 4. Catalytic mechanism of *C. thermocellum* endoglucanase CelD as a paradigm of β-glycanase acting with inversion of configuration at the anomeric carbon. The H^+ ion required for protonation of the glycosidic oxygen is provided by the Glu-555 residue. Simultaneously, the negatively charged Asp-201 residue promotes the ionization of a water molecule. The resulting OH^- ion interacts with the new anomeric carbon, leading to a single-step nucleophilic substitution and to inversion of configuration. Reprinted from [2], with permission.

Indeed, mutagenesis of Glu-126 only caused moderate inactivation of the enzyme [86].

Residues probably acting as nucleophiles have been identified on the basis of structural analysis and mutagenesis data in *C. thermocellum* CelD (Asp-201) [65,79] and *B. pumilus* xylanase XynA (Glu-93) [61,81]. In enzymes proceeding with retention of configuration, the nucleophilic residue can also be identified using inhibitors that form a stable reaction intermediate with the enzyme. After proteolytic digestion, the reactive nucleophile can then be identified by separating and sequencing the modified peptide. Using 2′,4′-dinitrophenyl-2-deoxy-2-fluoro-β-D-glucoside, Withers and colleagues mapped the nucleophile (Glu-274) of *C. fimi* exoglucanase Cex [87], of *Agrobacterium* sp. β-glucosidase (Glu-358) [88] and of *C. thermocellum* endoglucanase CelC (Glu-280) [89]. The latter result fits with those of Macarron et al. [90], who used 4′,5′-epoxypentyl-β-D-cellobioside to demonstrate that Glu-329 (which is homologous to Glu-280 in CelC) is the the reactive nucleophile of *T. reesei* endoglucanase III.

Cellulose-binding domains

Many cellulolytic enzymes contain non-catalytic cellulose-binding domains (CBDs). These are usually located at the NH_2 or COOH terminus of the enzymes, and are often separated from the catalytic domains by glycosylated, Pro/Thr/Ser-rich linker segments. Cellulose-binding properties have actually been demonstrated for only a fraction of the putative domains that can be identified by sequence similarity. So far, there is no reason to doubt that the latter are functional, although quantitative differences in binding affinity are to be expected.

Families of cellulose-binding domains

Most CBDs identified to date can be grouped into four families, with the exception of a couple of individual domains for which no similar sequence can be found yet in the GenBank and EMBL databases (Table 1). Types 1, 2, and 3 have each been found in enzymes belonging to at least three families of catalytic domains. So far, type 4 has only been found in two closely related enzymes. It must be noted that many enzymes contain unidentified non-catalytic domains, some of which may be involved in cellulose binding.

The best-characterized type of CBD is type 2. CBDs of type 2 have been found so far exclusively among fungal cellulases. They share a highly conserved sequence of about 30 residues. The CBD of *T. reesei* CBH I was synthesized chemically and its 3-D structure was determined by NMR spectroscopy. The domain is shaped like a wedge, whose dimensions are approximately $30 \times 18 \times 10$ Å. The secondary structure is composed of an irregular β-sheet with three antiparallel strands and two disulfide bridges [91]. One side of the wedge is hydrophilic, and the other one is hydrophobic. The hydrophilic side, which presumably makes contact with the surface of the substrate, contains three tyrosines. Nitration of the tyrosines abolishes binding to cellulose, suggesting that they participate in the interaction with the carbohydrate. Indeed, stacking interactions between the pyranose ring and the side chains of aromatic residues are quite common among proteins that interact with carbohydrates (e.g. [64,65]). Involvement of Tyr-492, which is located at the tip of the wedge, was further confirmed by site-directed mutagenesis. Mutation to Ala reduced the strength of cellulose binding to the level of the core enzyme. Mutation to His, which conserves the potential for π-bond interaction with the pyranose ring, had less drastic effects [92].

Domains of type 1 have been found in a variety of bacterial enzymes, including not only cellulases, but also xylanases, an arabinofuranosidase and a chitinase. They have a length of about 100 residues, with four conspicuously conserved tryptophans. These are likely candidates for interaction with the substrate. However, the results of proteolysis and mutagenesis experiments are somewhat confusing. Proteolysis of *C. fimi* CenA, in which the CBD is located at the NH_2 terminus, generates partially truncated polypeptides that can still bind to cellulose, although they lack up to 60 of the NH_2-terminal residues of the CBD [93]. Likewise, fusion peptides in which the CBD of CenA is linked to the NH_2 terminus of

alkaline phosphatase encoded by Tn*phoA* do not bind to cellulose unless they contain the entire CBD [94]. These results suggest that the most crucial region for cellulose binding is located at the COOH end of the CBD. However, when the conserved Trp residues of the CBD from *P. fluorescens* subsp. *cellulosa* xylanase A were mutagenized, it turned out that the most critical residues were those located in the NH_2-terminal part [95]. Thus, as pointed out in [95], it may be that in spite of sequence similarities, conserved residues are not quite equivalent in different CBDs. Binding of the CBD of CenA to cellulose would rely more on the C-terminal part of the domain, while the N-terminal part would be more important in the case of the CBD of xylanase A.

CBDs of type 3 comprise about 130 residues. They have been found in many different bacterial enzymes, and in the cellulose-binding proteins CbpA, CipA, and CipB, which are responsible for the structural organization of the multicellulase complexes (cellulosomes) present in *C. cellulovorans* and *C. thermocellum*. The cellulose-binding property of type 3 CBDs has been demonstrated for *B. lautus* CelA [96], *C. fimi* CenB [97], *B. subtilis* endoglucanase [98], and *C. thermocellum* CipB [99]. Far UV circular dichroic spectra of intact and truncated forms of *B. subtilis* endoglucanase suggest that the CBD contains a large proportion of β sheet [98]. As in types 1 and 2, several aromatic residues appear to be conserved, but no structural, biochemical or mutagenesis study has been performed yet to investigate their role in cellulose binding.

Role of cellulose-binding domains in cellulose degradation

There is a strong correlation between the capacity of cellulolytic enzymes to degrade crystalline cellulose and their affinity for cellulose [100]. In addition, all of the cellulases that are active against crystalline cellulose, and whose sequence has been determined, possess a CBD or are associated with a cellulose-binding protein (see below). This suggests that the presence of a CBD enhances the activity of cellulolytic enzymes towards crystalline cellulose. Indeed, the presence of an independent CBD should allow cellulases to perform many catalytic cycles while remaining tethered to the substrate. For this, the link between the CBD and the catalytic site must be sufficiently flexible to permit cleavage of more than one bond without desorption of the CBD, and binding of the CBD to cellulose must be reversible in order to allow movement of the enzyme along the substrate. Limiting diffusion of the enzymes on the surface of the substrate should help to focus hydrolysis at defined locations, and take advantage of the local disruption of the structure already induced by previous attacks.

None of the available evidence contradicts these views, but a considerable variation was observed when the effect of the CBD on enzyme activity was determined for different enzymes. This was achieved either by comparing intact and truncated forms of cellulases naturally containing a CBD, or by grafting foreign CBDs by means of recombinant DNA technology. As compared to intact *B. subtilis* endoglucanase containing a CBD, the truncated enzyme was 4-fold more active on CMC, but 5-fold less active on insoluble, amorphous cellulose [98]. For *T. reesei* CBH I and CBH II, removal of the CBD by proteolytic cleavage reduced the activity towards Avicel by 85% and 40–60%, respectively, whereas activity towards chromogenic oligosaccharides remained unchanged [101,102]. Removal of the CBD from *C. fimi* CenA increased the specific activity of the enzyme towards CMC and acid-swollen cellulose by 50%, and reduced the activity towards Avicel by 20% [103]. Grafting of the CBD of *T. fusca* endoglucanase E_2 on the catalytic domain of *Prevotella ruminicola* endoglucanase yielded a protein having 7–10-fold higher activity than the core enzyme on acid-swollen cellulose and ball-milled cellulose, and which could interact synergistically with *T. fusca* exoglucanase E_3 [104]. However, grafting of the CBD of *P. fluorescens* var. *cellulosa* arabinofuranosidase XYLC on *R. albus* endoglucanase CelA or of the CBD of *P. fluorescens* xylanase XYLA on the catalytic domain of *C. thermocellum* CelE did not affect the activity of either endoglucanase against CMC or β-glucan, but neither did the chimeric enzymes acquire activity against crystalline cellulose [105]. Several factors might explain the variability ob-

Table 1

Non-catalytic domains of cellulases, xylanases, and related enzymes

Organism	Protein	Function	GenBank accession number
Cellulose-binding domains			
Type 1 (11 entries in ref. [48])			
Cellulomonas fimi	CenD	Endoglucanase	L02544
Clostridium cellulovorans	EngD	Endoglucanase	M37434
Clostridium longisporum	CelA	Endoglucanase	L02868
Pseudomonas fluorescens	CelC	Cellodextrinase	S64954
Streptomyces plicatus	ChtA	Chitinase	M82804
Thermomonospora fusca	E_2	Endoglucanase	M73321
	E_5	Endoglucanase	L01577
Type 2 (5 entries in ref. [48])			
Agaricus bisporus	Cel1	Unknown [a]	M86356
Humicola grisea	Cbh1	CBH	M64588
Penicillium janthinellum	Cbh1	CBH	X59054
Phanerochaete chrysosporium	Cbh1-1	CBH	S109508
	Cbh1-2	CBH	S109508
Trichoderma longibrachiatum	Eg1	Endoglucanase	X60652
Type 3 (new)			
Bacillus lautus	CelA	Endoglucanase	M76588
	CelC	Endoglucanase	[253]
Bacillus subtilis	–	endoglucanase	X67044, M28332, M16185, D01057
Caldocellum saccharolyticum	ManA (2 copies)	Mannanase/ endoglucanase	L01257
	CelA (2 copies)	Endoglucanase	[58]
	CelB	CBH/ endoglucanase	X13602
Cellulomonas fimi	CenB	Endoglucanase	M64644
Clostridium cellulolyticum	CelCCG	Endoglucanase	S114528
Clostridium cellulovorans	CbpA	Cellulose-binding protein	M73817
Clostridium stercorarium	CelZ (2 copies)	Endoglucanase	P23659 (Swiss-Prot)
Clostridium thermocellum	CipA	Cellulosome-integrating protein	L08665
	CipB	Cellulosome-integrating protein	X68233
	CelF	Endoglucanase	X60545
	CelI	Endoglucanase	L04735
Type 4 (new)			
Cellulomonas fimi	CenC (2 copies)	Endoglucanase	X57858
Clostridium cellulolyticum	CelCCC (2 copies)	Endoglucanase	S48579
Streptomyces reticuli	Cel1	Avicelase	L03218
Unclassified CBDs			
Clostridium thermocellum	CelE	Endoglucanase	M22759
Erwinia chrysanthemi	CelZ	Endoglucanase	Y00540

Table 1 (continued)

Organism	Protein	Function	GenBank accession number
Domains similar to fibronectin type III domains			
Alcaligenes faecalis	P3HB	Poly-3-hydroxy-butyrate depolymerase	[254]
Bacillus circulans	ChA1 (2 copies)	Chitinase	M57601
Bacillus lautus	CelC	Endoglucanase	[253]
Cellulomonas fimi	CenB (3 copies)	Endoglucanase	M64644
	CenD (2 copies)	Endoglucanase	L02544
Cellulomonas flavigena	ORFX	Unknown	[255]
Clostridium thermohydrosulfuricum	AAP (2 copies)	α-Amylase-pullanase	M28471
Duplicated segments involved in attachment to the scaffolding protein of the cellulosome			
Clostridium cellulolyticum	CelCCA	Endoglucanase	M32362
	CelCCC	Endoglucanase	S114528
	CelCCD	Endoglucanase	D90341
	CelCCG	Endoglucanase	S114528
	ORF1	CBH [b]	S114528
Clostridium cellulovorans	EngB	Endoglucanase	M75706
Clostridium thermocellum	CelA	Endoglucanase	M19422
	CelB	Endoglucanase	X03592
	CelD	Endoglucanase	X04584
	CelE	Endoglucanase	M22759
	CelF	Endoglucanase	X60545
	CelG	Endoglucanase	X69390
	CelH	Endoglucanase	M31903
	CelS	CBH	L06942
	CipA	Cellulosome-integrating protein	L08665
	OrfX	Unknown	M22759
	LicB	Lichenase	X63355
	XynZ	Xylanase	M22624
Motifs similar to S-layer proteins			
Acetogenium kivui	−(1 1/2 copies)	S-layer protein	M31039
Bacillus sp. KSM 635	−(3 copies)	Endoglucanase	M27420
Bacillus brevis	MWP (1 1/2 copies)	S-layer protein	[259]
Clostridium thermocellum	ORF1p (3 copies)	Unknown	X67506
	ORF2p (3 copies)	Unknown	X67506
	ORF3p (3 copies)	Anchoring protein?	X67506
Thermoanaerobacter sp. B6A-RI	XynA (3 copies)	Xylanase	M97882
Proteins carrying more than one catalytic domain			
Caldocellum saccharolyticum	CelB	CBH/endoglucanase	X13602
	ManA	β-Mannanase/endoglucanase	L01257
	CelA	Endoglucanase/CBH [b]	[58]
Neocallimastix patriciarum	XylA	Xylanase/xylanase	X65526
Ruminococcus flavefaciens	XynA	Xylanase/xylanase	Z11127
	XynD	Xylanase/β-1,3-glucanase	[256]

Only entries not already cited in ref. [48] are mentioned. The classification of CBDs is the same as in ref. [48], except that A, B, C etc. have been replaced by 1, 2, 3 etc. The similarity of new sequences with previously established families was detected using the TFASTA algorithm [252] and checked visually for alignment of the most conserved residues. References are given for sequences whose accession number in the GenBank/EMBL databases could not be found.

[a] Protein synthesized specifically during growth on cellulose.

[b] Enzyme specificity not determined experimentally, but surmised to be of the CBH type on the basis of sequence similarity with *C. thermocellum* CelS.

served for the effect of CBDs on enzyme activity. Some core enzymes could be unable to hydrolyze crystalline cellulose, regardless of the presence of a CBD. This seems to be the case for *R. albus* CelA and for the catalytic domain of *C. thermocellum* CelE. The effect of the CBD varies considerably according to reaction conditions (incubation time and enzyme concentration) [106]. Proper joining of the CBD to the core enzyme can affect the adsorption/desorption behaviour of the enzyme and its catalytic activity [107]. Finally, little is known about the destabilizing activity that different CBDs might have on crystalline cellulose structure. It has been proposed that the wedge-like shape of fungal CBDs not only promotes binding to the substrate, but also helps to peel off cellulose chains from the top layer of cellulose microfibrils [108]. A destabilizing effect on the structure of cellulose fibres was observed when cotton or ramie fibres were incubated with the CBD derived from *C. fimi* CenA. The surface of the fibres became rough, showing signs of exfoliation of the surface layers. The fibres also became more permeable to penetration by and labelling with fluorescein isothiocyanate-labelled CBD. In addition, treatment with the CenA CBD led to the relase of small particles composed of bundles of microfibrils [34]. Likewise, aggregation and sedimentation of cellulose particles in microcrystalline cellulose suspensions could be prevented by cellulase preparations containing tightly binding components [100]. These results suggest that CBDs can take a more active part in cellulose degradation than the mere anchoring of cellulases to their substrate. However, in the case of hemicellulases, such as the *P. fluorescens* subsp. *cellulosa* arabinofuranosidase and xylanases [109,110], the role of the CBD is more probably limited to the attachment of the enzymes to the plant cell wall.

Synergism and physical association of cellulolytic components

Fungal cellulase systems

It has long been known that the degradation of crystalline cellulose requires the synergistic interaction of cellulolytic components having different specificities. Following the biochemical analysis of the cellulase systems produced by *P. chrysosporium* [111,112] and *Trichoderma koningii* [113], a model was proposed in the late seventies to account for the discovery of CBHs and of their synergism with endoglucanases and β-glucosidases (Fig. 5). Endoglucanases would first hydrolyze amorphous regions of the cellulose fibres. The non-reducing ends generated could then be attacked by CBHs, which would then proceed with the degradation of the crystalline regions. β-glucosidases would prevent the accumulation of cellobiose, which inhibits CBHs. Since then, the model has been extended to other aerobic fungi, such as *T. reesei*, *Penicillium pinophilum* [114] or *Talaromyces emersonii* [115]. Several bacteria are also assumed to produce cellulase systems functioning along the same principles. *C. fimi* [116], *C. stercorarium* [117–120], *Microbispora bispora* [121], *T. fusca* [122] and *Streptomyces flavogriseus* [123] produce endo- and exoglucanases. Endo–exo synergism was observed for *M. bispora* [121] *C. stercorarium* [118,120] and *T. fusca* [122]. In addition, the two CBHs of the *M. bispora* system displayed an exo–exo synergism, which was particularly obvious in the case of cotton hydrolysis, and the exoglucanase E_3 from *T. fusca* displayed synergism with *T. reesei* CBH I [122].

In it simplest version, the model of Fig. 5 provides an explanation for the endo–exo-β-glucosidase synergism without involving physical interaction between the various cellulolytic components. Indeed, the purification of several CBHs and endoglucanases from the culture medium of cellulolytic fungi indicates that the enzymes can behave as individual proteins without forming stable complexes. However, this is not always the case. For example, Sprey and Lambert purified a stable multi-enzyme complex from the culture medium of *T. reesei*. The complex behaved as a homogeneous species by isoelectric focusing. Upon treatment in the presence of urea and octylglucoside, the complex dissociated into six components comprising an endoglucanase, a xylanase, and a β-glucosidase [124]. In addition, it was shown that *T. reesei* CBH I and CBH II mutually enhance binding of one another to cellulose [125].

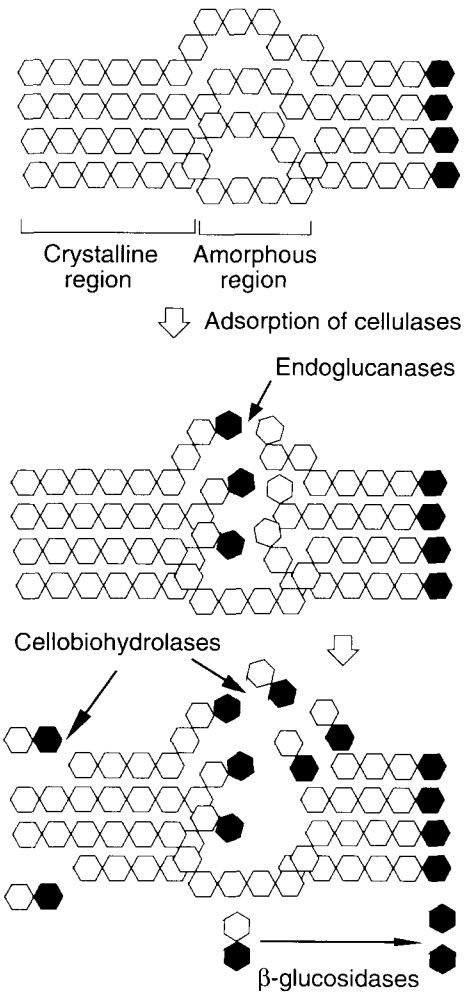

Fig. 5. Synergism between endoglucanases, CBHs, and β-glucosidases in fungal cellulase systems. Glucose residues are indicated by hexagons; reducing ends are in shown in black. Reprinted from [250], with permission.

In fact, specific protein–protein interactions could explain some of the observations that are not accounted for by the model shown in Fig. 5. For example, in *T. koningii*, the endo–exo synergism is selective and is not observed for all combinations of endoglucanases and CBHs [113]. Furthermore, in *P. pinophilum* and in *T. reesei*, two CBHs, which are known to interact with one another at the level of substrate binding (see above), are required for optimal synergistic activity [126,127]. Physical association between cellulolytic components may be expected to increase their efficiency by focusing enzymatic attack on regions of the substrate that are already destabilized.

The cellulosome concept

The case for stable multi-enzyme cellulase complexes is most firmly established for several anaerobic microorganisms. The best characterized example is the cellulase system of *C. thermocellum*, a Gram-positive, sporogenous and strictly anaerobic bacterium with an optimum growth temperature of about 60°C. The cellulase system produced by *C. thermocellum* is very active against crystalline cellulose, with a specific activity higher than that observed for the *T. reesei* system [30]. Attempts to purify the various components of the system met with considerable problems, since these components were very difficult to dissociate [128,129]. In 1983, Lamed, Bayer and their coworkers showed that extracellular cellulase complexes with a defined sedimentation coefficient and polypeptide composition could be purified from the culture medium of *C. thermocellum* strain YS [130,131]. Physical integrity of the complexes appeared to be required for efficient degradation of crystalline cellulose. Rather than unspecific aggregates, the complexes appeared therefore as functional entities, and were termed cellulosomes [130,131]. Cellulosomes are found both in the culture medium and at the surface of the cells (see below), where they form protuberances comprising up to several hundred cellulosomes.

The cellulosome is a complex with a molecular mass higher than 2 MDa. Dissociation in the presence of sodium dodecyl sulfate followed by PAGE reveals the presence of at least 14 different types of polypeptides with M_r values ranging from 48000 to 210000. This complexity reflects the high number of genes encoding cellulolytic or hemicellulolytic components. Screening of gene banks based on the detection of catalytically active recombinant proteins led to the cloning of genes encoding 15 endoglucanases, two xylanases [132,133], two β-glucosidases [134], and one lichenase (β-1,3-β-1,4-glucan glucanohydrolase) [135] from *C. thermocellum* NCIB 10682. These

numbers provide actually a minimum estimate of the total number of genes involved. An independent collection of genes encoding nine endoglucanases, two exoglucanases and one β-glucosidase was obtained from strain F7, a Russian isolate of *C. thermocellum* [136–138]. However, in the absence of cross-hybridization studies, it is difficult to estimate how many of the F7 clones are truly distinct from those isolated from strain NCIB 10682.

In addition, the gene encoding a major CBH termed CelS (previously called S8 [139] or S_S [140,141] was cloned recently using hybridization with degenerate oligonucleotides derived from the protein sequence [142]. The sequence of CelS shows a high degree of similarity with the protein encoded by ORF1 of *C. cellulolyticum* [143] and with the C-terminal domain of an endoglucanase from *Caldocellum saccharolyticum* [57,58]. There is some confusion in the literature about the specificity of CelS. CelS (then termed S_S) was described as an 83-kDa enzyme hydrolyzing CMC and behaving like a typical endoglucanase [144]. However, the N-terminal sequence of the S8 CBH purified by Morag et al. [139] shows that CelS and S8 are identical [145]. The N-terminal sequence of CelS is also highly similar to that of *Clostridium stercorarium* Avicelase II [118]. This explains why the cloned *celS* gene product and related enzyme domains have weak or undetectable CMC-hydrolyzing (CMCase) activity ([58], J.-P. Belaich and J.H.D. Wu, personal communications). S8 [139], the *celS* gene product (J.H.D. Wu, personal communication) and Avicelase II [118] have significant exoglucanase activity against Avicel and amorphous cellulose, but do not hydrolyze CMC. Given the available data, it seems reasonable to suggest that the endoglucanase S_S purified by Fauth et al. [144], and for which no gene has been identified, should be given another name.

The stoichiometry of the cellulosome, estimated from densitometric analysis of the various bands, indicates that some of the components are present in much greater proportion than others [131]. Zymogram analysis showed that the majority of the polypeptides has cellulase or xylanase activity [130,146,147]. Furthermore, a polypeptide of 210–250 kDa, previously termed S1 or S_L and now termed CipA (for cellulosome integrating protein), appears to be devoid of catalytic activity [130,146]. CipA is strongly *O*-glycosylated and carries a branched tetrasaccharide composed of three galactose residues and one 3-*O*-methyl-*N*-acetylglucosamine residue located at one of the termini [148].

A first clue to the role of CipA was obtained when Wu and Demain showed that CipA pro-

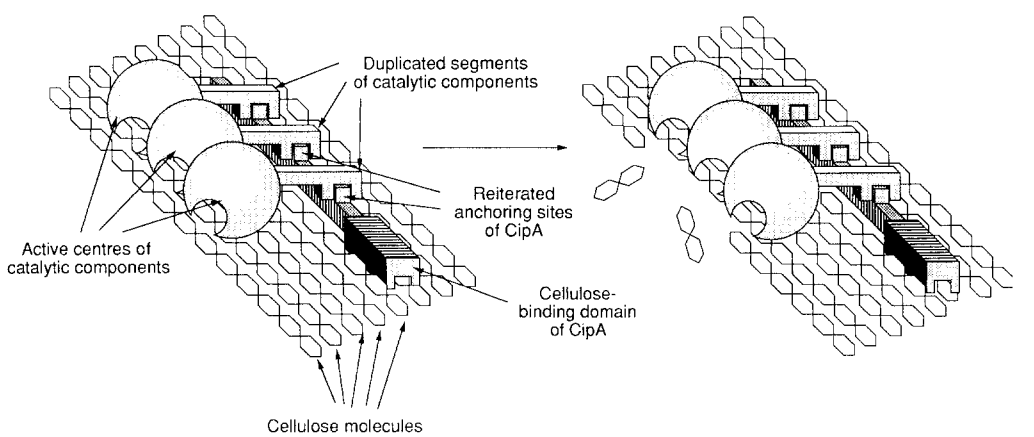

Fig. 6. Hypothetical model of the topological organization of subunits within the cellulosome. The diagram shows a portion of the cellulosome with the CBD of CipA and three of the CipA domains responsible for binding the duplicated segment borne by the catalytic components. Glucose residues are not drawn to the same scale as proteins. See text for details. Reprinted from [251], with permission.

moted the degradation of crystalline cellulose by CelS (then termed S_S). At the same time, CipA was shown to be required for binding of CelS to the substrate [140,141]. Binding of purified CipA to cellulose independently from other cellulosome components was demonstrated subsequently [149,150]. These results suggested that CipA acts as a cellulose-binding factor to which the catalytic components of the cellulosome would bind. This idea has now been borne out by results derived from the molecular characterization of catalytic components and their genes. Nucleotide sequences were determined for nine endoglucanase genes, one xylanase gene, one lichenase gene, and the *celS* CBH gene. The polypeptide sequences of the catalytic domains are quite diverse. In addition to CelS and the lichenase, the endoglucanases and the xylanase belong to four of the families of catalytic domains described in refs. [47,49,50]. However, all of the enzymes except endoglucanases CelC and CelI carry a highly conserved region of about 65 residues, which is located in most cases at the COOH terminus of the protein. This region, which comprises two highly similar segments of 22 residues each, is not required for catalytic activity nor for cellulose binding [151–154]. The duplicated segment was shown to promote binding of ^{125}I-labelled endoglucanase CelD and xylanase XynZ to CipA transferred to nitrocellulose. No binding was observed in control experiments performed with truncated forms of the same proteins devoid of the duplicated segment [155]. Furthermore, grafting of the duplicated segment of CelD onto endoglucanase CelC conferred on the latter the property to bind to CipA [156]. Therefore, the duplicated segment appears to serve as an attachment domain that mediates binding of the catalytic subunits to CipA.

The structure of CipA fits perfectly with its dual function as a scaffolding protein and as a cellulose-binding factor. The 5′ region of the *cipA* gene was cloned by screening a gene library for the production of a recombinant protein immunologically related to CipA [157]. The remaining part of the gene was cloned by chromosome walking and the whole gene was sequenced [158]. Independently, the 3′ region of the gene was obtained by screening for clones encoding proteins that bind ^{125}I-labelled CelD [159]. The identity of the cloned gene fragment with the *cipA* gene cloned by immunological methods was ascertained by matching of the two sequences over 1920 nucleotides. The CipA polypeptide comprises a CBD of type 3, and nine highly similar internal repeats of about 166 residues each. A duplicated segment similar to the duplicated segments of catalytic subunits is present at the COOH terminus of the protein. By subcloning portions of the gene, it was shown that the 166-amino acid repeats correspond to receptors responsible for binding of the 22-amino acid duplicated segment [159]. The cellulose-binding property of the CBD was confirmed by similar subcloning experiments performed with *cipB*, a gene highly similar to *cipA* cloned from *C. thermocellum* strain YS [99]. The role of the duplicated segment at the COOH terminus is unclear. It may be involved in concatenation of the CipA subunits. Alternatively, it may play a role in the anchoring of the cellulosome to the cell surface (see below).

A model summarizing these findings is presented in Fig. 6. The model is also consistent with electron microscopy pictures showing structures resembling a bunch of grapes, with globular subunits linked by means of a thin stalk to a fibrous structure [160]. In the model proposed by Mayer and colleagues [160–162], the catalytic subunits are assumed to be poised to attack the same cellulose chain in a set of quasi-simultaneous cutting events, which would enhance the processivity of the reaction. Indeed, determination of the average chain length of the residual substrate during degradation indicates that cellulose hydrolysis by *C. thermocellum* occurs with a high degree of processivity [163]. As discussed above for the case of CBDs, these considerations imply some flexibility in the geometry of the complex relative to the substrate, as well as reversible binding to cellulose. The latter point was confirmed by the fact that maximal activity of the cellulosome was observed at an ionic strength below that required for strongest binding [164].

Thus, the cellulosome appears as a particularly well-adapted machinery associating a variety of

enzymes capable to attack the various forms of cellulose, as well as other carbohydrates present in plant cell walls. Some important points are still unclear, however.

How does the complex assemble? Is the cellulosome a complex with a well-defined stoichiometry and topology? For example, do the catalytic subunits associate randomly along the CipA subunit, or do the various duplicated segments recognize each a specific, complementary receptor? The latter hypothesis does not seem likely, since the CipA molecule contains only nine receptors and since 11 polypeptides carrying the duplicated segment have already been identified by DNA sequencing. In addition, some of the polypeptides are present in more than one copy per complex [131]. One possibility to salvage a strict stoichiometry and topology of the complex would be to assume that, given more than one CipA molecule per cellulosome, there are enough binding sites to accomodate all different types of subunits in a single complex having a defined stoichiometry. A rigidly defined topology would be determined by specific interactions between determinants borne by the cores of the subunits. The hypothesis would require that the core domains, whose sequences do not display any obvious cellulosome-specific features, should have evolved a highly complex set of specific surface determinants while maintaining their activity. More probably, assembly of the complex occurs with some degree of freedom, which does not preclude some preferential interactions between defined duplicated segments and specific CipA receptors or between some of the catalytic subunits.

What is the mechanism of crystalline cellulose hydrolysis? Mayer et al. [160] postulated that, in the above model, CBHs would not be needed, since quasi-simultaneous attack of the same chain would provide the processivity required for crystalline cellulose degradation. However, the CelS/S8 CBH appears to be the most abundant cellulosome component [131]. In addition, in contrast to CMCase activity, efficient degradation of crystalline cellulose by *C. thermocellum* cellulase is inhibited by cellobiose [165] and requires the activity of one or more oxidation-sensitive, calcium-dependent components [166]. These properties are typical of CelS [139]. In addition, in the absence of calcium and dithiothreitol, the activity of *C. thermocellum* cellulase against Avicel and cotton can be restored by adding *T. koningii* CBH I [167]. These results emphasize that the CBH component(s) of the cellulosome may play a key role in the hydrolysis of crystalline cellulose.

Another point is why so many different types of endoglucanases are present in the complex, whether all are needed, and for what type of substrate. As mentioned above, the heterogeneity of cellulose may explain some of this diversity, but finding a specific function for no less than 15 endoglucanases stretches the imagination.

Answers to these questions would require to manipulate the composition of the cellulosome. Genetic manipulation would be the best means for this, for example genes encoding specific components could be inactivated by insertional mutagenesis. Unfortunately, no genetic tools are available yet for such experiments. Alternatively, artificial complexes could be reconstituted in vitro from purified components and assessed for their enzymatic properties. Indeed, conditions under which the cellulosome can be dissociated and reassociated in an active form have been described recently [168].

High molecular mass cellulase complexes are produced by various microorganisms, including *Clostridium cellulovorans* [169], the mesophilic *Clostridium* strain C7 [170,171], *R. albus* [172], *B. succinogenes* [173], *Bacteroides cellulosolvens* [174], and the anaerobic fungus *N. frontalis* [175,176]. A high M_r complex containing several xylanases associated with endoglucanases was purified from the culture supernatant of *B. fibrisolvens* H17c, and a high M_r complex composed of pectinolytic enzymes was characterized from *Clostridium thermosaccharolyticum* [177]. The high M_r cellulase complexes of the mesophilic clostridia appear to be built along the same principles as the *C. thermocellum* cellulosome, even if they have a smaller size (700 000–900 000). The *C. cellulovorans* cellulosome contains a 170-kDa cellulose-binding scaffolding protein termed CbpA [169]. The sequence of CbpA comprises a CBD of type 3 and nine domains similar to the receptors of *C. thermocellum* CipA [178]. At least one of the *C.*

cellulovorans endoglucanases, EngB, carries a duplicated segment related to those found in the catalytic subunits of the *C. thermocellum* cellulosome [179]. *C. cellulolyticum* also produces a high M_r cellulase complex [180], and five genes encoding polypeptides with a typical duplicated segment have been sequenced [143,181]. Four of the genes encode endoglucanases. The fifth gene is similar to *celS*. In addition, a gene encoding a protein similar to CipA and CbpA has been identified (J.-P. Belaich, personal communication). In the case of the mesophilic *Clostridium* strain C7, no information is available on the sequence of cellulase components, but activity against Avicel is associated with a 700-kDa complex containing a 130-kDa subunit which cross-reacts immunologically with an antiserum recognizing *C. thermocellum* CipA. A mutant lacking Avicelase, but retaining CMCase activity was characterized. In this mutant, CMCase activity was associated with a low M_r fraction containing several endoglucanases, but not the 130-kDa polypeptide [170,171]. Among non-clostridial species, the complex most reminiscent of the cellulosome was found in *B. cellulosolvens*. This bacterium produces a 700-kDa cellulolytic complex containing a 230-kDa glycoprotein. The 230-kDa component has no detectable activity, and it is recognized by an antiserum reacting with *C. thermocellum* CipA and by the GS1 isolectin of *Griffonia simplicifolia*, which binds to the carbohydrate moiety of CipA [174]. It is not clear how far the features of the *C. thermocellum* can be extrapolated to other high M_r complexes. To date, the known sequences of cellulolytic enzymes from rumen microorganisms do not carry a duplicated segment similar to that present in *C. thermocellum* carbohydrases. This suggests that high M_r complexes might be built along different principles. In the case of rumen bacteria such as *R. albus* or *F. succinogenes*, the activity against crystalline cellulose seems to be due primarily to cell-bound enzymes, which are associated with the capsule (in the case of *R. albus*) [182] or with the outer membrane (in the case of *F. succinogenes*). Bits and pieces of the capsule [182] or of the outer membrane [183] containing cellulases can detach from the surface of the cells and give rise to a high M_r, sedimentable cellulase fraction [183]. The same phenomenon could account for the formation of the high M_r cellulase complex of *N. frontalis*, as indicated by the fact that activity against crystalline cellulose was lost upon treatment with chitinase, which disrupts the fungal cell wall [176].

Enzymes carrying more than one catalytic domain

In addition to multimeric complexes, a growing number of (hemi)cellulases carrying two catalytic domains has been identified (Table 1). In most cases, a hemicellulolytic component is associated either with another hemicellulase or with a cellulase. This suggests a role in peeling off the hemicellulose matrix from the cellulose fibres, but the actual benefit of associating two enzymes within the same polypeptide has not yet been assessed. Should it prove of advantage for natural substrate degradation, genetic engineering would be the technique of choice for the construction of artificial bifunctional proteins with optimized properties. Indeed, Warren et al. [184] constructed a chimeric gene encoding a polypeptide which contained the catalytic domains of both exoglucanase Cex and endoglucanase CenA. Both domains were shown to be active in the fusion protein.

Attachment of cellulolytic organisms to the substrate

Close association between cellulolytic microorganisms and their substrate is of obvious advantage to optimize the recovery of soluble hydrolysis products. Indeed, many cellulolytic bacteria bind to cellulose, presumably by means of cellulases or of non-catalytic, cellulose-binding components present on the cell surface. For example, electron micrographs show that *F. succinogenes* clings very tightly to cellulose fibres, with erosion taking place right beneath the cells [183,185]. A survey of 13 cellulolytic and 10 non-cellulolytic strains of *R. albus* showed that most cellulolytic strains bound to cellulose, whereas most of the non-cellulolytic ones did not. However, within the cellulolytic strains, there was only a poor correlation between cellulolytic activity and bacterial adhesion to cellulose, indicating that, in most

cases, cell adhesion was not the overriding factor determining the cellulolytic activity of the cultures [186]. In the case of *C. thermocellum*, cellulosomes are embedded in the surface layer, where they form large clusters appearing as protuberances [160,187,257]. A mutant (AD2) isolated for its inability to bind to cellulose [188] was shown subsequently to be devoid of protuberances and of cell-bound cellulosomes [189]. Attachment of the cellulosomes to the cell surface is loose and the complexes, or aggregates of the complexes (polycellulosomes) detach from the cell to proceed independently with the degradation of the substrate [160,187,258].

Electron microscopic studies indicate that protuberances with morphology similar to those present in *C. thermocellum* are present on the surface of a variety of cellulolytic bacteria. These protuberances are not found in non-cellulolytic bacteria, nor on the surface of cellulolytic bacteria grown under conditions repressing cellulase synthesis [190–192]. Cell surface protuberances are not exclusively associated with cellulolytic activity. When grown in the presence of xylan, *Thermoanaerobacter* strain B6A displays negatively charged protuberances, whose appearance coincides with the production of cell-associated xylanases [193]. Likewise, numerous protuberances are formed on the surface of *Thermomonospora curvata* cells grown in the presence of cellulose or xylan, but not on cells grown in the presence of glucose, starch, or pectin [192]. Protuberances of cells grown in the presence of cellulose are predominantly associated with endoglucanase activity, and protuberances of cells grown in the presence of xylan with xylanase activity. It may be noted that the presence of protuberances on the cell surface of bacteria is not sufficient to prove the presence of multi-enzyme, cellulosome-like complexes. The resolution afforded by electron micrographs of samples stained with cationized ferritin does not allow to distinguish structures of the size of individual cellulosomes, which are much smaller than the protuberances. In spite of the presence of cell surface protuberances [190,192], biochemical studies never suggested the existence of cellulosome-like complexes in the cellulase systems of *Cellulomonas* or *Thermomonospora*. Thus, the association into cellulosome-like complexes of the cellulases present in the protuberances of these organisms is open to question.

Little is known about how cellulases and xylanases are targeted to remain attached to the cell surface. Recent evidence shows that the sequences of some cellulases and xylanases contain a region which is also present in several S-layer proteins (Table 1). It is tempting to speculate that this region could mediate attachment to the cells by binding to some component of the cell surface. The same motif was found recently in a *C. thermocellum* protein (ORF3p) which also contains a domain similar to the reiterated receptors of CipA. It was suggested that ORF3p may be involved in anchoring the cellulosome to the surface of *C. thermocellum*. The receptor-like domain would bind to the duplicated segment carried by CipA, and the domain similar to S-layer proteins would be bound to the cell surface [194].

It is worth noting that motifs related to fibronectin type III domains have been found in several bacterial carbohydrases (Table 1), in particular in *C. fimi* endoglucanases CenB and CenD [97,195,196]. No function has yet been attributed to these motifs, but it has been surmised that they could be involved in binding to other components of the hydrolytic system or to the cells [195].

Regulation of cellulase synthesis

In general, tools available for studying the regulation of cellulase gene expression are rather coarse. Very few organisms are amenable to genetics other than random mutagenesis. No gratuitous inducer has been described, which makes it difficult to sort out effects due to induction and to catabolite-type repression. Few studies are performed with neutral media allowing growth without influencing cellulase regulation. Current understanding of cellulase regulation is mostly based on biochemical and physiological studies, as well as on the quantification and characterization of mRNA transcripts using probes derived from cloned genes.

Cellulase synthesis appears to be controlled by

two basic regulation mechanisms. All known cellulase systems are repressed in the presence of low M_r carbon sources that are more easily metabolized than cellulose. In addition, in many systems, cellulase biosynthesis is induced in the presence of cellulose or its degradation products.

Repression by easily metabolized carbon sources

Addition of easily metabolized substrates blocks the synthesis of cellulase observed with cultures grown in the presence of cellulose alone. Although the phenomenon is almost universal [197], its mechanism has been investigated in very few cases. In *T. curvata*, the cAMP level under varying growth conditions is correlated with the specific rate of cellulase synthesis and cellulase production in toluenized cells is stimulated 2–3-fold in the presence of cAMP or theophyllin, a phosphodiesterase inhibitor. The authors point out that the relationship between cAMP level and rate of cellulase synthesis is not proportional, since the rate of synthesis can vary by 200-fold while the cAMP level varies only by 10-fold [198].

Mutants insensitive to catabolite-type repression have been isolated in *T. reesei* [199,200] and *Cellulomonas* sp. [201], either by selection in the presence of cellulose and deoxyglucose, or by plating mutagenized strains on solid media containing cellulose and glucose and selecting mutants producing a large hydrolysis halo. Such mutants have been used in the development of industrial strains of *T. reesei* hyperproducing cellulase.

Induction by cellulose and derived metabolites

Although catabolite-type repression is the only regulatory mechanism known in some systems [202,203], cellulase synthesis generally requires the presence of cellulose or its soluble metabolites. In *Cellulomonas* and *T. reesei*, growth on non-repressing substrates, slow feeding of glucose or mutations affecting catabolite-type repression are not sufficient to induce cellulase synthesis [201,204,205]. This suggests that cellulase synthesis is controlled by a specific induction mechanism beside catabolite-type repression. The generally accepted mechanism for induction by cellulose is that cellulose first undergoes limited hydrolysis by cellulases constitutively produced in low amounts. The soluble hydrolysis products thus generated could then penetrate the cells and cause induction of cellulase synthesis.

In the case of *T. reesei*, several observations support this hypothesis. Cellulose induces cellulase synthesis in germinating conidia, but not in mycelium [206], and this is correlated with the presence on the cell surface of conidia of a set of cellulases capable of digesting crystalline cellulose [207]. Induction of the mRNA encoding CBH I is blocked when antibodies directed against *T. reesei* cellulases are added to the culture medium [208]. Cellobiose and soluble cellodextrins have inducing activity, even if the latter is largely masked by catabolite-type repression, as described above. Maximal induction by cellulose appears to be correlated with the level of conidial-bound CBH II [209], and insertional inactivation of the *cbh2* gene results in delayed induction [210].

The true inducer of cellulase formation, as well as the enzyme(s) involved in its formation, have remained elusive. Sophorose (β-1,2-glucobiose) was proposed as a candidate after it was found to be a potent inducer in *T. reesei* (half-maximal induction at a concentration of 0.15 mM) [211,212]. This view was corroborated by the finding of a plasma membrane-bound β-glucosidase capable of forming sophorose by transglycosylation [213]. Inhibition of β-glucosidase activity by nojirimycin or gluconolactone prevented induction by cellulose, but not by sophorose [206]. However, sophorose fails to induce a complete set of cellulases [214], indicating that another metabolite of cellulose might contribute to regulation. The situation regarding enzymes that could convert cellobiose or cellodextrins into sophorose (or another inducer) is also confusing. It is reported that in strains bearing multiple copies of the *bgl* gene, increased β-glucosidase activity is correlated with enhanced formation of cellulase in the presence of cellulose, but also of sophorose [205]. If sophorose is the inducer generated by the transglycosylase activity of β-glucosidase, it is hard to conceive why increased β-glucosidase synthesis should lead to a stronger induction by

exogenous sophorose. However, insertional inactivation of the *bgl1* gene leads to a delay in cellulase induction by cellulose, which can be compensated by the addition of exogenous sophorose [215]. Part of the problem lies probably in the fact that there is more than one β-glucosidase playing a role in cellulase induction, as shown by the fact that induction by cellulose in *bgl1* mutants is delayed, but not abolished. Indeed, an intracellular β-glucosidase differring from the extracellular β-glucosidase encoded by *bgl1* has been identified [216].

Induction of cellulase biosynthesis by cellulose breakdown products is likely to occur in other systems as well. For example, adding exogenous cellulases to the culture medium of *T. fusca* significantly accelerates the induction process [217]. Likewise, in *Cellulomonas*, cellobiose and sophorose can be shown to act as inducers of cellulase synthesis in mutants insensitive to catabolite-type repression [201].

Characterization of mRNAs and promoter regions

The availability of probes derived from cloned genes has made it possible to characterize the transcripts of several genes encoding cellulases. These studies have shown that expression of cellulase genes is controlled at the level of transcription, and that most of the genes studied are monocistronic. However, an operon encoding at least four cellulase genes has been identified in *C. cellulolyticum* [143]. No consensus promoter can be identified. Inasmuch as similarities can be detected, cellulase gene promoters tend to resemble other consensus promoters described for the same or related species. For example, the promoters of the *celA* and *celD* genes of *C. thermocellum*, a member of the Bacillaceae, are similar to the σ^A and σ^D promoters described in *B. subtilis* [218,219]. The $cenBp_1$ and *cex* promoters of *C. fimi*, which belongs to the Corynebacteria and is related to Actinomycetes, resemble the *Streptomyces erythraeus* $ermp_2$ and *Streptomyces azureus* $tsrp_1$ promoters [220,221]. Transcription of the *T. reesei cbh2* mRNA displays multiple start sites located downstream of a TATA box, as typical for eukaryotic promoters [222].

Investigation of various gene transcripts within the same species indicates that the genes are transcribed to different extents and that transcription is not necessarily strictly coordinated. In addition, transcription of the same gene can start at different promoters according to culture conditions. In *C. fimi*, transcription of the *cex* [220] and *cenC* [223] genes was only detected in cells grown in the presence of CMC, whereas *cenA* [220] and *cenB* [221] transcripts were present in cells grown in the presence of glycerol, and even glucose in the case of *cenB*. High level transcription of *cenB* in cells induced by CMC started mainly at the $cenBp_1$ promoter, whereas constitutive transcription in the presence of glucose or glycerol started at the $cenBp_2$ promoter. In *C. thermocellum*, transcription of several *cel* genes was induced sequentially when cellobiose concentration in the medium became limiting. Transcription of *celA* started first, followed by *celD* and *celF*, and finally *celC*. Early *celD* transcripts started from a promoter similar to *B. subtilis* σ^D promoters, and were followed by transcripts starting from a promoter similar to σ^A once the cells reached stationary phase [219].

At the molecular level, very little is known about mechanisms regulating the rate of synthesis of cellulases. A putative regulatory protein, which was detected only in induced cells, was identified in *T. fusca*. This protein binds to a 14-bp palindromic sequence located upstream from at least three endoglucanase genes [224,225]. In the case of the *celE* gene, the binding site is located some 50 nucleotides downstream from the transcription start sites [225]. A motif similar to the sequence mediating cAMP-dependent regulation of rat tyrosyl amino transferase is located upstream from the *cbhII* gene. The sequence of the *cbhI* gene is preceded a sequence similar to that recognized by the negative catabolic repressor protein CreA of *Aspergillus nidulans* [205]. Deletion of this motif results in constitutive CBH I synthesis in the presence of glucose [226].

To summarize, models about the regulation of cellulase synthesis will no doubt reach a considerably higher degree of sophistication once physiological studies and characterization of mRNA transcripts can be combined with in vivo genetics.

This approach is already available for a few cellulase-producing organisms such as *E. chrysanthemi* [227] and *T. reesei* [228].

Applications of cellulases and hemicellulases

From a practical point of view, several processes involving cellulases or hemicellulases can be considered. In the short term, applications requiring only partial hydrolysis are the most likely to become economically feasible, and some of them are already commercialized. In the food processing sector, crude cellulose (straw, wood shavings) is traditionally used to grow edible mushrooms, such as *Agaricus bisporus*, which is currently produced at a rate of 1 million tons/year [229]. Cellulase and hemicellulase preparations are already used to clear fruit juices from remaining pulp particles. Cellulases can also help the extraction of juice and oil from fruit and seed pulp. In addition, several cellulases can hydrolyse β-1,3-β-1,4-glucan, which is present in high amounts in low grade barley and hampers the filtration of beer. Recombinant yeasts producing β-1,3-β-1,4 glucanases are already available for the brewery industry [230]. Objections to their acceptance depend on the psychology of consumers rather than on rational motivations. Cellulases and hemicellulases are also expected to have an impact on the processing of animal feed. Treatment of silage with cellulases is expected to be beneficial for two reasons. Partial hydrolysis of the plant cell wall should speed up the release of sugars to be fermented by lactic acid bacteria. In addition, cellulases might improve the digestibility of silage. Crude enzyme preparations from *T. reesei* and *Aspergillus niger* are already marketed, but some basic and technical problems remain to be solved [231–233]. In the same vein, studies are also in progress to improve the digestibility of feed in monogastric animals. One goal is the hydrolysis of glucans and xylans present in low-grade feedstuffs, which increase the viscosity of bowel content, reduce nutrient uptake and lead to the 'sticky faeces' problem in chicken. For example, *T. reesei* was grown on spent grains from breweries, and the effect of the treatment was assessed with respect to the nutritional value of the grains as chicken feed [234]. Another interesting approach lies in the creation of transgenic animals that would secrete the required enzymes in their gastrointestinal tract. A first step in this direction was achieved by introducing into mice the gene encoding *C. thermocellum* endoglucanase CelE, which possesses xylanase and β-glucanase activity, and which is resistant to proteases present in the small intestine. The gene was placed under the control of the exocrine pancreas-specific enhancer of the elastase I gene, and CelE was found to be specifically produced in the exocrine pancreas and secreted into the small intestine. The enzyme was active, as shown by the hydrolysis of MUC [235]. It was calculated that a 30-fold improvement in the amount of CelE secreted would be sufficient to hydrolyze 90% of the glucan present in barley-based diets. Beside their actual practical goal, these studies, as well as others [236–238], emphasize the usefulness of cellulases as reporters for monitoring the synthesis and secretion of proteins. The recreational potential of mice producing fluorescent droppings should also not be underestimated.

Textile processing is another area where cellulases have been successfully put to use. Cellulases added to laundry powder shave off the microfibrils which tend to stick out of cotton fibres after several washing cycles. This significantly helps restore softness and colour brightness to cotton fabrics. Cellulases have also been used to remove excess dye from denim fabric in pre-faded blue jeans ('biostoning').

Considerable interest is currently focused on the use of hemicellulases in the processing of paper pulp. Several studies indicate that xylanases help reduce the amount of chlorine required to bleach the pulp (for reviews see [19–21,232]). This could have important environmental consequences since chlorine combines with phenolics derived from lignin, resulting in pollution of paper mill effluents. Beside reducing enzyme costs, one important goal would be to obtain preparations entirely devoid of cellulase activity, which is detrimental to paper quality. In this respect, enzymes produced by genetic engi-

neering can prove attractive, since they do not need to be purified from the cellulases that are produced by most xylanolytic organisms [239].

Another important area concerns the treatment of cellulosic wastes, which are produced in increasing amounts either as municipal solid waste or as agricultural waste. Burning is costly and not always applicable. As an alternative, anaerobic conversion into methane and carbon dioxide by bacterial consortia is gaining increasing acceptance as the solution of choice for the treatment of industrial wastes and manure. At present, the process relies on naturally occurring bacteria, but efficiency could conceivably be improved by using better strains, particularly for the two limiting steps, cellulolysis and methanogenesis, which occur at either end of the metabolic chain.

Cellulase systems may also offer interesting spin-offs not directly related to cellulase activity proper. For example, CBDs could be used as tagging sequences for the affinity purification of recombinant proteins on a cheap cellulose matrix. The possibility was explored for the CBD of *C. fimi* CenA, which was fused by genetic engineering to *E. coli* alkaline phosphatase [94] and to *Agrobacterium* sp. β-glucosidase [240]. Both proteins retained activity and could be bound and eluted from a cellulose affinity column. Another possibility could be to graft the duplicated segment of *C. thermocellum* enzymes on recombinant proteins, which could be bound to CipA in order to construct artificial multiprotein complexes. Such complexes may prove useful for performing coupled reactions. The feasibility of the scheme was tested by grafting the duplicated segment of endoglucanase CelD to the COOH terminus of endoglucanase CelC, which is devoid of duplicated segment and does not bind to CipA. The resulting chimeric protein, CelC-Cel'D, was able to bind to CipA, yet the catalytic properties of the enzyme were not significantly altered [156].

However, the major long-term challenge for very large scale biotechnological applications remains the utilization as a source of glucose of the huge, renewable mass of cellulosic residues available every year. Fermentation of glucose into solvents and fuels, particularly ethanol and butanol, could provide a partial substitution for fossil fuels. This project was started in the aftermath of the oil crisis in the late seventies. Since then, it has experienced waves of waxing and waning enthusiasm on the part of funding agencies. Economically, fuels derived from lignocellulosics are not yet competitive with petrol or even with products derived from starch, a substrate more costly to produce, but much easier to hydrolyze. However, the need for replacement of fossil fuels in order to reduce emissions of greenhouse gases has recently brought an added incentive for the development of fuels derived from renewable biomass. For this goal, starch alone accounts for too small a fraction of plants to sustain a positive energy balance, and cellulose will have to be utilized as well, if biomass is to be used as an energy source.

Improving the biological conversion of lignocellulosics relies on several lines of research:
(i) Development of cheap pretreatment processes in order to increase the digestibility of the substrate (e.g. milling, chemical treatment with alkali or acid, steam explosion, hydrothermolysis);
(ii) Improving cellulolysis by increasing the efficiency, the stability, and the recovery of cellulolytic enzymes, and by eliminating inhibitory products such as cellobiose, using β-glucosidases;
(iii) Improving the efficiency of the fermentation step. In most processes studied presently, the hydrolysis step is performed prior to the fermentation step. The advantage is that conditions can be optimized for each step independently. In particular, well-suited, non-cellulolytic microorganisms such as yeast or *Clostridium acetobutylicum* can be used for the fermentation step. However, simultaneous saccharification and fermentation represents an attractive simplification of the process. Two basic strategies can be adopted [241]. In the first one, an organism with good fermentative properties is engineered to produce cellulases. In the second one, an organism with with good cellulolytic efficiency is engineered to improve its fermentative properties. Due to the complexity of cellulase systems, imparting good cellulolytic properties to a non-cellulolytic organism meets with considerable difficulties, and attempts in this direction [242,243] are still far off the goal. However, in a first stage,

partially cellulolytic fermenters could be used in conjunction with exogenously added cellulase. Wood and Ingram [244] have investigated the possibility to use strains of *Klebsiella oxytoca* containing chromosomally integrated fermentation genes from *Zymomonas mobilis* genes for ethanol production and plasmids expressing thermostable cellulase genes from *C. thermocellum* in a two-step process. In the first step, cells containing intracellular recombinant endoglucanase CelD are incubated with the substrate at 60°C, which releases the enzyme from the cells and allows attack of the substrate. In the second step, fungal cellulase and a fresh inoculum of cells are added to the pretreated substrate and fermentation is allowed to proceed at 35°C. It was found that pretreatment with cells expressing CelD substantially reduces the amount of exogenous fungal cellulase needed for the fermentation step. In the experiments described, endoglucanase CelD originated from cells grown on glucose, but in an industrial scheme, it would be provided by cells from the previous fermentation.

Engineering good cellulolytic microorganisms for better fermentative properties also faces unsolved problems. For example, several studies have investigated the potential of *C. thermocellum* for simultaneous saccharification and fermentation [28,245–249]. As mentioned above, *C. thermocellum* produces a highly active thermostable cellulase complex, and ferments cellulose into ethanol, acetic acid, lactic acid, hydrogen, and carbon dioxide. However, wild-type strains have a limited ethanol tolerance, and the mixed acid fermentation pattern results in poor ethanol yields. Random mutagenesis may help to alleviate these problems, and more efficient strains have indeed been obtained [28,245–247,249], but genetics is still limited by the unavailability of tools for introducing targeted modifications into the *C. thermocellum* genome.

In conclusion, basic research has begun to unravel some of the unconventional solutions that cellulolytic microorganisms have evolved to degrade an unconventional substrate. One may hope that in a not too distant future, such knowledge will be helpful in the development of an increasing number of applications.

Acknowledgements

We wish to thank Drs. E.A. Bayer, J.-P Belaich, N.R. Gilkes, Q. Wang, S.G. Withers, and J.H.D. Wu for communicating results prior to publication. Work performed in our laboratory was supported by Grants EN3B-0082-F and CPL C462 from the Commission of the European Communities and by research funds from the University of Paris 7.

Dedication

This review is dedicated to our friend Michael P. Coughlan, in memory of his contributions to this field of research.

References

1 Béguin, P. (1990) Molecular biology of cellulose degradation. Annu. Rev. Microbiol. 44, 219–248.
2 Béguin, P. and Aubert, J.-P. (1992) La dégradation de la cellulose par les microorganismes. Ann. Inst. Pasteur/Actualités 3, 91–115.
3 Coughlan, M.P. (1985) The properties of fungal and bacterial cellulases with comment on their production and application. Biotechnol. Genet. Eng. Rev. 3, 39–109.
4 Coughlan, M.P. (1992) Enzymic hydrolysis of cellulose: an overview. Bioresource Technol. 39, 107–115.
5 Eriksson, K.-E.L., Blanchette, R.A. and Ander, P. (1990) Microbial and Enzymatic Degradation of Wood and Wood Components. Springer–Verlag, Berlin Heidelberg.
6 Eveleigh, D.E. (1987) Cellulase: a perspective. Phil. Trans. R. Soc. Lond. A321, 435–447.
7 Gilbert, H.J. and Hazlewood, G.P. (1993) Bacterial cellulases and xylanases. J. Gen. Microbiol. 139, 187–194.
8 Ljungdahl, L.G. and Eriksson, K.-E. (1985) Ecology of microbial cellulose degradation. In: Advances in Microbial Ecology (Marshall, K.C., Ed.). Vol. 8, pp. 237–299. Plenum Press, New York, NY.
9 Wood, T.M. (1985) Aspects of the biochemistry of cellulose degradation. In: Cellulose and its Derivatives, Chemistry, Biochemistry, and Applications (Kennedy, J.F., Phillips, G.O., Wedlock, D.J. and Williams, P.A., Eds.), Ellis Harwood Limited Publishers, Chichester Halsted Press, a division of John Wiley and Sons, New York, NY.
10 Atalla, R.H. and VanderHart, D.L. (1984) Native cellulose: A composite of two distinct crystalline forms. Science 223, 283–284.
11 Kuga, S. and Brown, R.M. Jr. (1991) Physical structure of cellulose microfibrils: implications for biogenesis. In: Biosynthesis and Degradation of Cellulose (Haigler, C.H.

and Weimer, P.J., Eds.), pp. 125–142. Marcel Dekker, New York, NY.
12 Richmond, P.A. (1991) Occurence and functions of native cellulose. In: Biosynthesis and Biodegradation of Cellulose (Haigler, C.H. and Weimer, P.J., Eds.), pp. 5–23. Marcel Dekker, New York, NY.
13 Henrissat, B. (1985) Ph.D. Thesis, Université de Grenoble.
14 Wood, T.M. (1988) Preparation of crystalline, amorphous, and dyed cellulase substrates. Methods Enzymol. 160, 19–25.
15 Joseleau, J.P., Comtat, J. and Ruel, K. (1992) Chemical structure of xylans and their interaction in the plant cell walls. In: Progress in Biotechnology, Vol. 7, Xylans and Xylanases (Visser, J., Beldman, G., Kusters-van Someren, M.A. and Voragen, A.G.J., Eds.), pp. 1–15. Elsevier, Amsterdam.
16 Hartley, R.D. and Ford, C.W. (1989) Phenolic constituents of plant cell walls and wall biodegradability. In: ACS Symp. Ser. 399, Plant Cell Wall Polymers, Biogenesis and Biodegradation (Lewis, N.G. and Paice, M.G., Eds.), pp. 135–145. American Chemical Society, Washington, D.C.
17 Scalbert, A., Monties, B., Lallemand, J.-Y., Guittet, Y. and Rolando, C. (1985) Ether linkage between phenolic acids and lignin fractions from wheat straw. Phytochemistry 24, 1359–1362.
18 Gold, M., Wariishi, H. and Valli, K. (1989) Extracellular peroxidases involved in lignin degradation by the white rot Basidiomycete *Phanerochaete chrysosporium*. In: ACS Symp. Series No. 389. Biocatalysis in Agricultural Biotechnology (Whitaker, J. and Sonnet, P.E., Eds.), pp. 127–140. American Chemical Society, Washington, D.C.
19 Biely, P. (1985) Microbial xylanolytic systems. Trends Biotechnol. 3, 286–290.
20 Coughlan, M.P. and Hazlewood, G.P. (1993) β-1,4-D-xylan-degrading enzyme systems: biochemistry, molecular biology and applications. Biotechnol. Appl. Biochem. 17, 259–289.
21 Wong, K.K.Y., Tan, L.U.L. and Saddler, J.N. (1988) Multiplicity of β-1,4-xylanase in microorganisms: functions and applications. Microbiol. Rev. 52, 305–317.
22 Thomson, J.A. (1993) Molecular biology of xylan degradation. FEMS Microbiol. Rev. 104, 65–82.
23 Kirk, T.K. (1988) Biochemistry of lignin degradation by *Phanerochaete chrysosporium*. In: FEMS Symposium No. 43. Biochemistry and Genetics of Cellulose Degradation (Aubert, J.-P., Béguin, P. and Millet, J., Eds.), pp. 315–332. Academic Press, London and New York.
24 Miller, G.L., Dean, J. and Blum, R. (1960) A study of methods for preparing oligosaccharides from cellulose. Arch. Biochem. Biophys. 91, 21–26.
25 Van Tilbeurgh, H., Claeyssens, M. and De Bruyne, C.K. (1982) The use of 4-methylumbelliferyl and other chromophoric glycosides in the study of cellulolytic enzymes. FEBS Lett. 149, 152–156.
26 Macarron, R., Acebal, C., Castillon, M.P., Dominguez, J.M., de la Mata, I., Pettersson, G., Tomme, P. and Claeyssens, M. (1993) Mode of action of endoglucanase III from *Trichoderma reesei*. Biochem. J. 289, 867–873.
27 Claeyssens, M. and Henrissat, B. (1992) Specificity mapping of cellulolytic enzymes – classification into families of structurally related proteins confirmed by biochemical analysis. Protein Sci. 1, 1293–1297.
28 Tailliez, P., Girard, H., Millet, J. and Béguin, P. (1989) Enhanced cellulose fermentation by an asporogenous and ethanol-tolerant mutant of *Clostridium thermocellum*. Appl. Environ. Microbiol. 55, 207–211.
29 Updegraff, D.M. (1969) Semimicro determination of cellulose in biological materials. Anal. Biochem. 32, 420–424.
30 Johnson, E.A., Sakajoh, M., Halliwell, G., Madia, A. and Demain, A.L. (1982) Saccharification of complex cellulosic substrates by the cellulase system from *Clostridium thermocellum*. Appl. Environ. Microbiol. 43, 1125–1132.
31 Grethlein, H.E. (1985) The effect of pore size distribution on the rate of enzymatic hydrolysis of cellulosic substrates. Bio/Technology 3, 155–160.
32 Chanzy, H., Henrissat, B. and Vuong, R. (1983) Colloidal gold labelling of 1,4-β-D-glucan cellobiohydrolase adsorbed on cellulose substrates. FEBS Lett. 172, 193–197.
33 Chanzy, H. and Henrissat, B. (1985) Unidirectional degradation of *Valonia* cellulose microcrystals subjected to cellulase action. FEBS Lett. 184, 285–288.
34 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, 1096–1099.
35 Sprey, B. and Bochem, H.P. (1991) Electron microscopic observations of cellulose microfibril degradation by endocellulase from *Trichoderma reesei*. FEMS Microbiol. Lett. 78, 183–188.
36 Sprey, B. and Bochem, H.P. (1992) Effect of endoglucanase and cellobiohydrolase from *Trichoderma reesei* on cellulose microfibril structure. FEMS Microbiol. Lett. 97, 113–117.
37 Sprey, B. and Bochem, H.P. (1993) Formation of cross-fractures in cellulose microfibril structure by an endoglucanase-cellobiohydrolase complex from *Trichoderma reesei*. FEMS Microbiol. Lett. 106, 239–244.
38 Coleman, G.S. (1978) The metabolism of cellulose, glucose, and starch by the rumen ciliate protozoa *Eudiplodinium maggii*. J. Gen. Microbiol. 107, 359–366.
39 Tucker, M.L., Durbin, M.L., Clegg, M.T. and Lewis, L.N. (1987) Avocado cellulase: nucleotide sequence of a putative full-length cDNA clone and evidence for a small gene family. Plant Mol. Biol. 9, 197–203.
40 Jones, T.H.D., de Renobales, M. and Pon, N. (1979) Cellulases released during the germination of *Dictyostelium discoideum* spores. J. Bacteriol. 137, 752–757.
41 Blume, J.E. and Ennis, H.L. (1991) A *Dictyostelium discoideum* cellulase is a member of a spore germination-specific gene family. J. Biol. Chem. 266, 15432–15437.
42 McCarthy, A.J. (1987) Lignocellulose-degrading actinomycetes. FEMS Microbiol. Rev. 46, 145–163.
43 Weimer, P.J. and Zeikus, J.G. (1977) Fermentation of

cellulose and cellobiose by *Clostridium thermocellum* in the absence and presence of *Methanobacterium thermoautotrophicum*. Appl. Environ. Microbiol. 33, 289–297.
44 Wood, T.M., Wilson, C.A., McCrae, S.I. and Joblin, K.N. (1986) A highly active extracellular cellulase from the anaerobic rumen fungus *Neocallimastix frontalis*. FEMS Microbiol. Lett. 34, 37–40.
45 Wood, T.M., McCrae, S.I., Wilson, C.A., Bhat, K.M. and Gow, L.A. (1988) Aerobic and anaerobic fungal cellulases, with special reference to their mode of attack on crystalline cellulose. In: Proceedings of FEMS Symposium No. 43. Biochemistry and Genetics of Cellulose Degradation (Aubert, J.-P., Béguin, P. and Millet, J., Eds.), pp. 31–52. Academic Press, London and New York.
46 Orpin, C.G. (1988) Genetic approaches to the improvement of lignocellulose degradation in the rumen. In: FEMS Symp. No. 43. Biochemistry and Genetics of Cellulose Degradation (Aubert, J.-P., Béguin, P. and Millet, J., Eds.), pp. 172–179. Academic Press, London and New York.
47 Henrissat, B., Claeyssens, M., Tomme, P., Lemesle, L. and Mornon, J.-P. (1989) Cellulase families revealed by hydrophobic cluster analysis. Gene 81, 83–95.
48 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, function, and enzyme families. Microbiol. Rev. 55, 303–315.
49 Henrissat, B. (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem. J. 280, 309–316.
50 Henrissat, B. and Bairoch, A. (1993) New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem. J. 293, in press.
51 Nakamura, K., Misawa, N. and Kitamura, K. (1986) Sequence of a cellulase gene of *Cellulomonas uda* CB4. J. Biotechnol. 4, 247–254.
52 Nakai, R., Horinouchi, S. and Beppu, T. (1988) Cloning and nucleotide sequence of a cellulase gene, *casA*, from an alkalophilic *Streptomyces* strain. Gene 65, 229–238.
53 Gilkes, N.R., Claeyssens, M., Aebersold, R., Henrissat, B., Meinke, A., Morrison, H.D., Kilburn, D.G., Warren, R.A.J. and Miller, R.C. (1991) Structural and functional relationships in two families of β-1,4-glycanases. Eur. J. Biochem. 202, 367–377.
54 Damude, H.G., Gilkes, N.R., Kilburn, D.G., Miller, R.C. Jr. and Warren, R.A.J. (1993) Endoglucanase CasA from alkalophilic *Streptomyces* KSM-9 is a typical member of family B of β-1,4-glucanases. Gene, in press.
55 Wang, W. and Thomson, J.A. (1990) Nucleotide sequence of the *celA* gene encoding a cellodextrinase of *Ruminococcus flavefaciens* FD-1. Mol. Gen. Genet. 222, 265–269.
56 Wang, W. and Thomson, J.A. (1992) *Erratum*. Nucleotide sequence of the *celA* gene encoding a cellodextrinase of *Ruminococcus flavefaciens* FD-1. Mol. Gen. Genet. 233, 492.
57 Lüthi, E., Jasmat, N.B., Grayling, R.A., Love, D.R. and Bergquist, P.L. (1991) Cloning, sequence analysis, and expression in *Escherichia coli* of a gene coding for a β-mannanase from the extremely thermophilic bacterium *Caldocellum saccharolyticum*. Appl. Environ. Microbiol. 57, 694–700.
58 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, 3864–3867.
59 Henrissat, B. (1993) Hidden domains and active site residues in β-glycanase-encoding gene sequences. Gene 125, 199–204.
60 Torronen, A., Kubicek, C.P. and Henrissat, B. (1993) Amino acid sequence similarities between low molecular mass endo-1,4-β-xylanases and family H cellulases revealed by hydrophobic cluster analysis. FEBS Lett. 321, 135–139.
61 Okada, H. (1991) Comparisons of primary, secondary and tertiary structures of xylanase of *Bacillus pumilus* and cellulase of *Aspergillus aculeatus*. In: Microbial Utilization of Renewable Resources. pp. 1–7. International Center of Cooperative Research in Biotechnology, Japan.
62 Matsushita, O., Russell, J.B. and Wilson, D.B. (1990) Cloning and sequencing of a *Bacteroides ruminicola* $B_1$4 endoglucanase gene. J. Bacteriol. 172, 3620–3630.
63 Matsushita, O., Russell, J.B. and Wilson, D.B. (1991) A *Bacteroides ruminicola* 1,4-β-endoglucanase is encoded in two reading frames. J. Bacteriol. 173, 6919–6926.
64 Rouvinen, J., Bergfors, T., Teeri, T., Knowles, J.K.C. and Jones, T.A. (1990) Three-dimensional structure of cellobiohydrolase II from *Trichoderma reesei*. Science 249, 380–386.
65 Juy, M., Amit, A.G., Alzari, P.M., Poljak, R.J., Claeyssens, M., Béguin, P. and Aubert, J.-P. (1992) Crystal structure of a thermostable bacterial cellulose-degrading enzyme. Nature 357, 89–91.
66 Davies, G., Tolley, S., Wilson, K., Schülein, M., Wodlike, H.F. and Dodson, G. (1992) Crystallization and preliminary X-ray analysis of a fungal endoglucanase I.J. Mol. Biol. 228, 970–972.
67 Divne, C., Sinning, I., Stahlberg, J., Pettersson, G., Bailey, M., Siika-Aho, M., Margolles-Clark, E., Teeri, T. and Jones, A.T. (1993) Crystallization and preliminary X-ray studies on the core proteins of cellobiohydrolase I and endoglucanase I from *Trichoderma reesei*. J. Mol. Biol., in press.
68 Bedarkar, S., Gilkes, N.R., Kilburn, D.G., Kwan, E., Rose, D.R., Miller, R.C. Jr., Warren, R.A.J. and Withers, S.G. (1992) Crystallization and preliminary X-ray diffraction analysis of the catalytic domain of Cex, an exo-β-1,4-glucanase and β-1,4-xylanase from the bacterium *Cellulomonas fimi*. J. Mol. Biol. 228, 693–695.
69 Pickersgill, R.W., Jenkins, J.A., Scott, M., Connerton, I., Hazlewood, G.P. and Gilbert, H.J. (1993) Crystallization and preliminary X-ray analysis of the catalytic domain of xylanase A from *Pseudomonas fluorescens* var. *cellulosa*. J. Mol. Biol. 229, 246–248.
70 Sinnott, M.L. (1990) Catalytic mechanisms of enzymic glycosyl transfers. Chem. Rev. 90, 1171–1202.
71 Kelly, J.A., Sielecki, A.R., Sykes, B.D., James, M.N.G. and Phillips, D.C. (1979) X-ray crystallography of the binding of the bacterial cell wall trisaccharide NAM-NAG-NAM to lysozyme. Nature 282, 875–878.
72 Knowles, J.K.C., Lehtovaara, P., Murray, M. and Sinnott,

M.L. (1988) Stereochemical course of the action of the cellobioside hydrolases I and II of *Trichoderma reesei*. J. Chem. Soc., Chem. Commun., pp. 1401–1402.

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

74 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, 89–92.

75 Claeyssens, M., van Tilbeurgh, H., Kamerling, J.P., Berg, J., Vrsanska, M. and Biely, P. (1990) Studies on the cellulolytic system of the filamentous fungus *Trichoderma reesei* and transfer activity of endoglucanase I. Biochem. J. 270, 251–256.

76 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 N.M.R. determination of the stereochemical course of hydrolyses catalysed by glucanase components of the cellulase complex. Biochem. Biophys. Res. Comm. 139, 487–494.

77 Meinke, A., Braun, C., Gilkes, N.R., Kilburn, D.G., Miller, R.C. Jr. and Warren, R.A.J. (1991) Unusual sequence organization in CenB, an inverting endoglucanase from *Cellulomonas fimi*. J. Bacteriol. 173, 308–314.

78 Barras, F., Bortoli-German, I., Bauzan, M., Rouvier, J., Gey, C., Heyraud, A. and Henrissat, B. (1992) Stereochemistry of the hydrolysis reaction catalyzed by endoglucanase Z from *Erwinia chrysanthemi*. FEBS Lett 300, 145–148.

79 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, 4472–4478.

80 Tomme, P., van Beeumen, J. and Claeyssens, M. (1992) Modification of catalytically important residues in endoglucanase D from *Clostridium thermocellum*. Biochem. J. 285, 319–324.

81 Ko, E.P., Akatsuka, H., Moriyama, H., Shinmyo, A., Hata, Y., Katsube, Y., Urabe, I. and Okada, H. (1992) Site-directed mutagenesis at aspartate and glutamate residues of xylanase from *Bacillus pumilus*. Biochem. J. 288, 117–121.

82 Baird, S.D., Hefford, M.A., Johnson, D.A., Sung, W.L., Yaguchi, M. and Seligy, V.L. (1990) The Glu residue in the conserved Asn-Glu-Pro sequence of two highly divergent endo-β-1,4-glucanases is essential for enzymatic activity. Biochem. Biophys. Res. Commun. 169, 1035–1039.

83 Py, B., Bortoli-German, I., Haiech, J., Chippaux, M. and Barras, F. (1991) Cellulase EGZ of *Erwinia chrysanthemi*: structural organization and importance of His98 and Glu133 residues for catalysis. Protein Eng. 4, 325–333.

84 Belaich, A., Fierobe, H.P., Baty, D., Busetta, B., Bagnara-Tardif, C., Gaudin, C. and Belaich, J.-P. (1992) The catalytic domain of endoglucanase A from *Clostridium cellulolyticum* – Effects of arginine-79 and histidine-122 mutations on catalysis. J. Bacteriol. 174, 4677–4682.

85 Navas, J. and Béguin, P. (1992) Site-directed mutagenesis of conserved residues in *Clostridium thermocellum* endoglucanase CelC. Biochem. Biophys. Res. Comm. 189, 807–812.

86 Mitsuishi, Y., Nitisinprasert, S., Saloheimo, M., Biese, I., Reinikainen, T., Claeyssens, M., Keränen, S., Knowles, J.K.C. and Teeri, T.T. (1990) Site-directed mutagenesis of the putative catalytic residues of *Trichoderma reesei* cellobiohydrolase I and endoglucanase I. FEBS Lett. 275, 135–138.

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

88 Withers, S.G., Warren, R.A.J., Street, I.P., Rupitz, K., Kempton, J.B. and Aebersold, R. (1990) Unequivocal demonstration of the involvement of a glutamate residue as a nucleophile in the mechanism of a "retaining" glycosidase. J. Am. Chem. Soc. 112, 5887–5889.

89 Wang, Q., Tull, D., Meinke, A., Gilkes, N.R., Warren, R.A.J., Aebersold, R. and Withers, S.G. (1993) Glu280 is the nucleophile in the active site of *Clostridium thermocellum* CelC, a family A endo-β-1,4-glucanase. J. Biol. Chem. 268, 14096–14102.

90 Macarron, R., van Beeumen, J., Henrissat, B., de la Mata, I. and Claeyssens, M. (1993) Identification of an essential glutamate residue in the active site of endoglucanase III from *Trichoderma reesei*. FEBS Lett. 316, 137–140.

91 Kraulis, P.J., Clore, G.M., Nilges, M.J., Jones, T.A., Pettersson, G., Knowles, J. and Gronenborn, A.M. (1989) Determination of the three-dimensional 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, 7241–7257.

92 Reinikainen, T., Ruhonen, 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 Struct. Funct. Genet. 14, 475–482.

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

94 Greenwood, J.M., Gilkes, N.R., Kilburn, D.G., Miller, R.C. Jr. and Warren, R.A.J. (1989) Fusion to an endoglucanase allows alkaline phosphatase to bind to cellulose. FEBS Lett. 244, 127–131.

95 Poole, D.M., Hazlewood, G.P., Huskisson, N.S., Virden, R. and Gilbert, H.J. (1993) The role of conserved tryptophan residues in the interaction of a bacterial cellulose binding domain with its ligand. FEMS Microbiol. Lett. 106, 77–84.

96 Hansen, C.K., Diderichsen, B. and Jorgensen, P.L. (1992) *celA* from *Bacillus lautus* PL236 encodes a novel cellu-

lose-binding endo-β-1,4-glucanase. J. Bacteriol. 174, 3522–3531.
97 Meinke, A., Gilkes, N.R., Kilburn, D.G., Miller, R.C. Jr. and Warren, R.A.J. (1991) Multiple domains in endoglucanase B (CenB) from *Cellulomonas fimi*: Functions and relatedness to domains in other polypeptides. J. Bacteriol. 173, 7126–7135.
98 Hefford, M.A., Laderoute, K., Willick, G.E., Yaguchi, M. and Seligy, V. (1992) Bipartite organization of the *Bacillus subtilis* endo-β-1,4-glucanase revealed by C-terminal mutations. Prot. Eng. 5, 433–439.
99 Poole, D.M., Morag, E., Lamed, R., Bayer, E.A., Hazlewood, G.P. and Gilbert, H.J. (1992) Identification of the cellulose binding domain of the cellulosome subunit S1 from *Clostridium thermocellum* YS. FEMS Microbiol. Lett. 99, 181–186.
100 Klyosov, A.A. (1990) Trends in biochemistry and enzymology of cellulose degradation. Biochemistry 29, 10577–10585.
101 Tomme, P., van Tilbeurgh, H., Pettersson, G., van Damme, J., Vandekerckhove, J., Knowles, J., Teeri, T. and Claeyssens, M. (1988) Studies of the cellulolytic system of *Trichoderma reesei* QM 9414. Eur. J. Biochem. 170, 575–581.
102 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, 223–227.
103 Gilkes, N.R., Warren, R.A.J., Miller, R.C. Jr. and Kilburn, D.G. (1988) Precise excision of the cellulose binding domains from two *Cellulomonas fimi* cellulases by a homologous protease and the effect on catalysis. J. Biol. Chem. 263, 10401–10407.
104 Maglione, G., Matsushita, O., Russel, J.B. and Wilson, D.B. (1992) Properties of a genetically reconstructed *Prevotella ruminicola* endoglucanase. Appl. Environ. Microbiol. 58, 3593–3597.
105 Poole, D.M., Durrant, A.J., Hazlewood, G.P. and Gilbert, H.J. (1991) Characterization of hybrid proteins consisting of the catalytic domains of *Clostridium* and *Ruminococcus* endoglucanases, fused to *Pseudomonas* non-catalytic cellulose-binding domains. Biochem. J. 279, 787–792.
106 Stahlberg, J., Johansson, G. and Pettersson, G. (1991) A new model for enzymatic hydrolysis of cellulose based on the 2-domain structure of cellobiohydrolase I. Bio/Technology 9, 286–290.
107 Shen, H., Schmuck, M., Pilz, I., Gilkes, N.R., Kilburn, D.G. and Miller, R.C. Jr. (1991) Deletion of the linker connecting the catalytic and cellulose-binding domains of endoglucanase A (CenA) of *Cellulomonas fimi* alters its conformation and catalytic activity. J. Biol. Chem. 266, 11335–11340.
108 Knowles, J., Teeri, T.T., Lehtovaara, P., Penttilä, M. and Saloheimo, M. (1988) The use of gene technology to investigate fungal cellulolytic enzymes. In: Proceedings of FEMS Symposium No. 43. Biochemistry and Genetics of Cellulose Degradation (Aubert, J.-P., Béguin, P. and Millet, J., Eds.), pp. 153–169. Academic Press, London and New York.
109 Kellett, L.F., Poole, D.M., Ferreira, L.M.A., Durrant, A.J., Hazlewood, G.P. and Gilbert, H.J. (1990) Xylanase B and an arabinofuranosidase from *Pseudomonas fluorescens* subsp. *cellulosa* contain identical cellulose-binding domains and are coded by adjacent genes. Biochem. J. 272, 369–376.
110 Hazlewood, G.P. and Gilbert, H.J. (1992) The molecular architecture of xylanases from *Pseudomonas fluorescens* subsp. *cellulosa*. In: Progress in Biotechnology, Vol. 7, Xylans and Xylanases (Visser, J., Beldman, G., Kusters-van Someren, M.A. and Voragen, A.G.J., Eds.), pp. 259–273. Elsevier, Amsterdam.
111 Eriksson, K.-E. (1978) Enzyme mechanisms involved in cellulose hydrolysis by the rot fungus *Sporotrichum pulverulentum*. Biotechnol. Bioeng. 70, 317–332.
112 Eriksson, K.-E. and Pettersson, B. (1975) Extracellular enzyme system utilized by the fungus *Sporotrichum pulverulentum* (*Chrysosporium lignorum*) for the breakdown of cellulose. 1. Separation, purification and physicochemical characterization of five endo-1,4-β-glucanases. Eur. J. Biochem. 51, 193–206.
113 Wood, T.M. and McCrae, S.I. (1978) The cellulase of *Trichoderma koningii*. Purification and properties of some endoglucanase components with special reference to their action on cellulose when acting alone or in synergism with the cellobiohydrolase. Biochem. J. 171, 61–72.
114 Wood, T.M. and McCrae, S.I. (1986) The cellulase of *Penicillium pinophilum*. Biochem. J. 234, 93–99.
115 McHale, A. and Coughlan, M.P. (1980) Synergistic hydrolysis of cellulose by components of the extracellular cellulase system of *Talaromyces emersonii*. FEBS Lett. 117, 319–322.
116 Gilkes, N.R., Langsford, M.L., Kilburn, D.G., Miller, R.C. Jr. and Warren, R.A.J. (1984) Mode of action and substrate specificities of cellulases from cloned bacterial genes. J. Biol. Chem. 259, 10455–10459.
117 Bronnenmeier, K. and Staudenbauer, W.L. (1990) Cellulose hydrolysis by a highly thermostable endo-1,4-β-glucanase (Avicelase I) from *Clostridium stercorarium*. Enzyme Microb. Technol. 12, 431–436.
118 Bronnenmeier, K., Rücknagel, K.P. and Staudenbauer, W.L. (1991) Purification and properties of a novel type of exo-1,4-β-glucanase (Avicelase II) from the cellulolytic thermophile *Clostridium stercorarium*. Eur. J. Biochem. 200, 379–385.
119 Creuzet, N., Berenger, J.-F. and Frixon, C. (1983) Characterization of exoglucanase and synergistic hydrolysis of cellulose in *Clostridium stercorarium*. FEMS Microbiol. Lett. 20, 347–350.
120 Creuzet, N. and Frixon, C. (1983) Purification and characterization of an endoglucanase from a newly isolated thermophilic bacterium. Biochimie 65, 149–156.
121 Yablonsky, M.D., Bartley, T., Elliston, K.O., Kahrs, S.K., Shalita, Z.P. and Eveleigh, D.E. (1988) Characterization and cloning of the cellulase complex of *Microbispora bispora*. In: FEMS Symposium No. 43. Biochemistry and Genetics of Cellulose Degradation (Aubert, J.-P., Béguin, P. and Millet, J., Eds.), pp. 249–266. Academic Press, London and New York.

122 Walker, L.P., Wilson, D.B., 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, 1019–1026.

123 MacKenzie, C.R., Bilous, D. and Johnson, K.G. (1984) Purification and characterization of an exoglucanase from *Streptomyces flavogriseus*. Can. J. Microbiol. 30, 1171–1178.

124 Sprey, B. and Lambert, C. (1983) Titration curves of cellulases from *Trichoderma reesei*: demonstration of a cellulase-xylanase-β-glucosidase-containing complex. FEMS Microbiol. Lett. 18, 217–222.

125 Tomme, P., Heriban, V. and Claeyssens, M. (1990) Adsorption of two cellobiohydrolases from *Trichoderma reesei* to Avicel: Evidence for "exo–exo" synergism and possible "loose complex" formation. Biotechnol. Lett. 12, 525–530.

126 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, 37–43.

127 Fägerstam, L.G. and Pettersson, L.G. (1980) The 1,4-β-glucan cellobiohydrolases of *Trichoderma reesei* QM9414. A new type of cellulolytic synergism. FEBS Lett. 119, 97–100.

128 Aït, N., Creuzet, N. and Forget, P. (1979) Partial purification of cellulase from *Clostridium thermocellum*. J. Gen. Microbiol. 113, 399–402.

129 Pètre, J., Longin, R. and Millet, J. (1981) Purification and properties of an endo-β-1,4-glucanase from *Clostridium thermocellum*. Biochimie 63, 629–639.

130 Lamed, R., Setter, E. and Bayer, E.A. (1983) Characterization of a cellulose-binding, cellulase-containing complex in *Clostridium thermocellum*. J. Bacteriol. 156, 828–836.

131 Lamed, R., Setter, E., Kenig, R. and Bayer, E.A. (1983) The cellulosome: a discrete cell surface organelle of *Clostridium thermocellum* which exhibits separate antigenic, cellulose-binding and various cellulolytic activities. Biotechnol. Bioeng. Symp. 13, 163–181.

132 Millet, J., Pétré, D., Béguin, P., Raynaud, O. and Aubert, J.-P. (1985) Cloning of ten distinct DNA fragments of *Clostridium thermocellum* coding for cellulases. FEMS Microbiol. Lett. 29, 145–149.

133 Hazlewood, G.P., Romaniec, M.P.M., Davidson, K., Grépinet, O., Béguin, P., Millet, J., Raynaud, O. and Aubert, J.-P. (1988) A catalogue of *Clostridium thermocellum* endoglucanase, β-glucosidase and xylanase genes cloned in *Escherichia coli*. FEMS Microbiol. Lett. 51, 231–236.

134 Gräbnitz, F. and Staudenbauer, W.L. (1988) Characterization of two β-glucosidase genes from *Clostridium thermocellum*. Biotechnol. Lett. 10, 73–78.

135 Schimming, S., Schwarz, W.H. and Staudenbauer, W.L. (1991) Properties of a thermoactive β-1,3-glucanase (lichenase) from *Clostridium thermocellum* expressed in *Escherichia coli*. Biochem. Biophys. Res. Comm. 177, 447–452.

136 Piruzian, E.S., Mogutov, M.A., Velikodvorskaya, G.A. and Akimenko, V.K. (1985) Cloning and expression of structural genes of *Clostridium thermocellum* F7 cellulolytic complex endoglucanases in cells of *Escherichia coli* K12. Doklady Acad. Nauk SSR 281, 963–965.

137 Bumazkin, B.K., Velikodvorskaya, G.A., Tuka, K., Mogutov, M.A. and Strongin, A.Y. (1990) Cloning of *Clostridium thermocellum* endoglucanase genes in *Escherichia coli*. Biochem. Biophys. Res. Commun. 167, 1057–1064.

138 Tuka, K., Zverlov, V.V., Bumazkin, B.K., Velikodvorskaya, G.A. and Strongin, A.Y. (1990) Cloning and expression of *Clostridium thermocellum* genes coding for thermostable exoglucanases (cellobiohydrolases) in *Escherichia coli* cells. Biochem. Biophys. Res. Commun. 169, 1055–1060.

139 Morag, E., Halevy, I., Bayer, E.A. and Lamed, R. (1991) Isolation and properties of a major cellobiohydrolase from the cellulosome of *Clostridium thermocellum*. J. Bacteriol. 173, 4155–4162.

140 Wu, J.H.D., Orme-Johnson, W.H. and Demain, A.L. (1988) Two components of an extracellular protein aggregate of *Clostridium thermocellum* together degrade crystalline cellulose. Biochemistry 27, 1703–1709.

141 Wu, J.H.D. and Demain, A.L. (1988) Proteins of the *Clostridium thermocellum* cellulase complex responsible for degradation of crystalline cellulose. In: FEMS Symposium No. 43 Biochemistry and Genetics of Cellulose Degradation (Aubert, J.-P., Béguin, P. and Millet, J., Eds.), pp. 117–131. Academic Press, London and New York.

142 Wang, W.K., Kruus, K. and Wu, J.H.D. (1993) Cloning and DNA sequence of the gene coding for *Clostridium thermocellum* cellulase S_S (CelS), a major cellulosome component. J. Bacteriol. 175, 1293–1302.

143 Bagnara-Tardif, C., Gaudin, C., Belaich, A., Hoest, P., Citard, T. and Belaich, J.-P. (1992) Sequence analysis of a gene cluster encoding cellulases from *Clostridium cellulolyticum*. Gene 119, 17–28.

144 Fauth, U., Romaniec, M.P.M., Kobayashi, T. and Demain, A.L. (1991) Purification and characterization of endoglucanase S_S from *Clostridium thermocellum*. Biochem. J. 279, 67–73.

145 Morag, E., Bayer, E.A., Hazlewood, G.P., Gilbert, H.J. and Lamed, R. (1993) Cellulase S_S (CelS) is synonymous with the major cellobiohydrolase (subunit S8) from the cellulosome of *Clostridium thermocellum*. Appl. Biochem. Biotechnol., in press.

146 Kohring, S., Wiegel, J. and Mayer, F. (1990) Subunit composition and glycosidic activities of the cellulase complex from *Clostridium thermocellum* JW20. Appl. Environ. Microbiol. 56, 3798–3804.

147 Morag, E., Bayer, E.A. and Lamed, R. (1990) Relationship of cellulosomal and non-cellulosomal xylanases of *Clostridium thermocellum* to cellulose-degrading enzymes. J. Bacteriol. 172, 6098–6105.

148 Gerwig, G., Kamerling, J.P., Vliegenthart, J.F.G., Morag, E., Lamed, R. and Bayer, E.A. (1991) Primary structure of *O*-linked carbohydrate chains in the cellulosome of different *Clostridium thermocellum* strains. Eur. J. Biochem. 196, 115–122.

149 Morag, E., Bayer, E.A. and Lamed, R. (1992) Unorthodox intra-subunit interactions in the cellulosome of *Clostridium thermocellum*: Identification of structural transitions induced in the S1 subunit. Appl. Biochem. Biotechnol. 33, 205–217.

150 Salamitou, S., Tokatlidis, K., Béguin, P. and Aubert, J.P. (1992) Involvement of separate domains of the cellulosomal protein S1 of *Clostridium thermocellum* in binding to cellulose and in anchoring of catalytic subunits to the cellulosome. FEBS Lett. 304, 89–92.

151 Hall, J., Hazlewood, G.P., Barker, P.J. and Gilbert, H.J. (1988) Conserved reiterated domains in *Clostridium thermocellum* endoglucanases are not essential for activity. Gene 69, 29–38.

152 Grépinet, O., Chebrou, M.-C. and Béguin, P. (1988) Nucleotide sequence and deletion analysis of the xylanase gene (*xynZ*) of *Clostridium thermocellum*. J. Bacteriol. 170, 4582–4588.

153 Chauvaux, S., Béguin, P., Aubert, J.-P., Bhat, K.M., Gow, L.A., Wood, T.M. and Bairoch, A. (1990) Calcium-binding affinity and calcium-enhanced activity of *Clostridium thermocellum* endoglucanase D. Biochem. J. 265, 261–265.

154 Durrant, A.J., Hall, J., Hazlewood, G.P. and Gilbert, H.J. (1991) The non-catalytic *C*-terminal region of endoglucanase E from *Clostridium thermocellum* contains a cellulose-binding domain. Biochem. J. 273, 289–293.

155 Tokatlidis, K., Salamitou, S., Béguin, P., Dhurjati, P. and Aubert, J.-P. (1991) Interaction of the duplicated segment carried by *Clostridium thermocellum* cellulases with cellulosome components. FEBS Lett. 291, 185–188.

156 Tokatlidis, K., Dhurjati, P. and Béguin, P. (1993) Properties conferred on *Clostridium thermocellum* endoglucanase CelC by grafting the duplicated segment of endoglucanase CelD. Prot. Eng., in press.

157 Romaniec, M.P.M., Kobayashi, T., Fauth, U., Gerngross, U.T. and Demain, A.L. (1991) Cloning and expression of a *Clostridium thermocellum* DNA fragment that encodes a protein related to cellulosome component S_L. Appl. Biochem. Biotechnol. 31, 119–134.

158 Gerngross, U.T., Romaniec, M.P.M., Huskisson, N.S. and Demain, A.L. (1993) Sequencing of a *Clostridium thermocellum* gene (*cipA*) encoding the cellulosomal protein S_L-protein reveals an unusual degree of internal homology. Mol. Microbiol. 8, 325–334.

159 Fujino, T., Béguin, P. and Aubert, J.-P. (1992) Cloning of a *Clostridium thermocellum* DNA fragment encoding polypeptides that bind the catalytic components of the cellulosome. FEMS Microbiol. Lett. 94, 165–170.

160 Mayer, F., Coughlan, M.P., Mori, Y. and Ljungdahl, L.G. (1987) Macromolecular organization of the cellulolytic enzyme complex of *Clostridium thermocellum* as revealed by electron microscopy. Appl. Environ. Microbiol. 53, 2785–2792.

161 Coughlan, M.P. and Ljungdahl, L.G. (1988) Comparative biochemistry of fungal and bacterial cellulolytic enzyme systems. In: Proceedings of FEMS Symposium No. 43. Biochemistry and Genetics of Cellulose Degradation (Aubert, J.-P., Béguin, P. and Millet, J., Eds.). pp. 11–30. Academic Press, London and New York.

162 Mayer, F. (1993) Principles of functional and structural organization in the bacterial cell compartments and their enzymes. FEMS Microbiol. Rev. 104, 327–346.

163 Puls, J. and Körner, H.-U. (1987) Study on the effects of the components of the cellulase complex on various celluloses and lignocelluloses. In: 4th E.C. International Conference on Biomass for Energy and Industry (Grassi, G., Delmon, D., Molle, J.-F. and Zibetta, H., Eds.). pp. 694–698. Elsevier, London.

164 Lamed, R., Kenig, R. and Setter, E. (1985) Major characteristics of the cellulolytic system of *Clostridium thermocellum* coincide with those of the purified cellulosome. Enzyme Microb. Technol. 7, 37–41.

165 Johnson, E.A., Reese, E.T. and Demain, A.L. (1982) Inhibition of *Clostridium thermocellum* cellulase by end products of cellulolysis. J. Appl. Biochem. 4, 64–71.

166 Johnson, E.A. and Demain, A.L. (1984) Probable involvement of sulfhydryl groups and a metal as essential components of the cellulase of *Clostridium thermocellum*. Arch. Microbiol. 137, 135–138.

167 Gow, L.A. and Wood, T.M. (1988) Breakdown of crystalline cellulose by synergistic action between cellulase components from *Clostridium thermocellum* and *Trichoderma koningii*. FEMS Microbiol. Lett. 50, 247–252.

168 Bhat, K.M. and Wood, T.M. (1992) The cellulase of the anaerobic bacterium *Clostridium thermocellum* – Isolation, dissociation, and reassociation of the cellulosome. Carbohydr. Res. 227, 293–300.

169 Shoseyov, O. and Doi, R.H. (1990) Essential 170-kDa subunit for degradation of crystalline cellulose by *Clostridium cellulovorans* cellulase. Proc. Natl. Acad. Sci. USA 87, 2192–2195.

170 Cavedon, K., Leschine, S.B. and Canale-Parola, E. (1990) Characterization of the extracellular cellulase from a mesophilic *Clostridium* (strain C7). J. Bacteriol. 172, 4231–4237.

171 Cavedon, K., Leschine, S.B. and Canale-Parola, E. (1990) Cellulase system of a free-living, mesophilic *Clostridium* (strain C7). J. Bacteriol. 172, 4222–4230.

172 Wood, T.M., Wilson, C.A. and Stewart, C.S. (1982) Preparation of the cellulase from the cellulolytic anaerobic rumen bacterium *Ruminococcus albus* and its release from the bacterial cell wall. Biochem. J. 205, 129–137.

173 Groleau, D. and Forsberg, C.W. (1981) Cellulolytic activity of the rumen bacterium *Bacteroides succinogenes*. Can. J. Biochem. 27, 517–530.

174 Lamed, R., Morag, E., Mor-Yosef, O. and Bayer, E.A. (1991) Cellulosome-like entities in *Bacteroides cellulosolvens*. Curr. Microbiol. 22, 27–33.

175 Wilson, C. and Wood, T.M. (1992) The anaerobic fungus *Neocallimastix frontalis*: isolation and properties of a cellulosome-type enzyme fraction with the capacity to solu-

bilize hydrogen-bond-ordered cellulose. Appl. Microbiol. Biotechnol. 37, 125–129.
176 Wilson, C.A. and Wood, T.M. (1992) Studies on the cellulase of the rumen anaerobic fungus *Neocallimastix frontalis*, with special reference to the capacity to degrade crystalline cellulose. Enzyme Microb. Technol. 14, 258–264.
177 Van Rijssel, M., Gerwig, G.J. and Hansen, T.A. (1993) Isolation and characterization of an extracellular glycosylated protein complex from *Clostridium thermosaccharolyticum* with pectin methylesterase and polygalacturonate hydrolase activity. Appl. Environ. Microbiol. 59, 828–836.
178 Shoseyov, O., Takagi, M., Goldstein, M.A. and Doi, R.H. (1992) Primary sequence analysis of *Clostridium cellulovorans* cellulose-binding protein A. Proc. Natl. Acad. Sci. USA 89, 3483–3487.
179 Foong, F., Hamamoto, T., Shoyesov, O. and Doi, R.H. (1991) Nucleotide sequence and characteristics of endoglucanase gene *engB* from *Clostridium cellulovorans*. J. Gen. Microbiol. 137, 1729–1736.
180 Madarro, A., Pena, J.L., Lequerica, J.L., Valles, S., Gay, R. and Flors, A. (1991) Purification and characterization of the cellulases from *Clostridium cellulolyticum* H10. J. Chem. Tech. Biotechnol. 52, 393–406.
181 Faure, E., Belaich, A., Bagnara, C., Gaudin, C. and Belaich, J.-P. (1989) Sequence analysis of the *Clostridium cellulolyticum celCCA* endoglucanase gene. Gene 84, 39–46.
182 Stack, R.J. and Hungate, R.E. (1984) Effect of 3-phenylpropionic acid on capsule and cellulases of *Ruminococcus albus*. Appl. Environ. Microbiol. 48, 218–223.
183 Forsberg, C.W., Beveridge, T.J. and Hellstrom, A. (1981) Cellulase and xylanase release from *Bacteroides succinogenes* and its importance in the rumen environment. Appl. Environ. Microbiol. 42, 886–896.
184 Warren, R.A.J., Gerhard, B., Gilkes, N.R., Owolabi, J.B., Kilburn, D.G. and Miller, R.C. Jr. (1987) A bifunctional exoglucanase-endoglucanase fusion protein. Gene 61, 421–427.
185 Gaudet, G. and Gaillard, B. (1987) Vesicle formation and cellulose degradation in *Bacteroides succinogenes* cultures: ultrastructural aspects. Arch. Microbiol. 148, 150–154.
186 Morris, E.J. and Cole, O.J. (1987) Relationship between cellulolytic activity and adhesion to cellulose in *Ruminococcus albus*. J. Gen. Microbiol. 133, 1023–1032.
187 Bayer, E.A. and Lamed, R. (1986) Ultrastructure of the cell surface cellulosome of *Clostridium thermocellum* and its interaction with cellulose. J. Bacteriol. 167, 828–836.
188 Bayer, E.A., Kenig, R. and Lamed, R. (1983) Adherence of *Clostridium thermocellum* to cellulose. J. Bacteriol. 156, 818–827.
189 Bayer, E.A., Setter, E. and Lamed, R. (1985) Organization and distribution of the cellulosome in *Clostridium thermocellum*. J. Bacteriol. 163, 552–559.
190 Lamed, R., Naimark, J., Morgenstern, E. and Bayer, E.A. (1987) Specialized cell surface structures in cellulolytic bacteria. J. Bacteriol. 169, 3792–3800.
191 Stewart, C.S., Duncan, S.H. and Flint, H.J. (1990) The properties of forms of *Ruminococcus flavefaciens* which differ in their ability to degrade cotton cellulose. FEMS Microbiol. Lett. 72, 47–50.
192 Hostalka, F., Moultrie, A. and Stutzenberger, F. (1992) Influence of carbon source on cell surface topology of *Thermomonospora curvata*. J. Bacteriol. 174, 7048–7052.
193 Lee, Y.-E., Lowe, S.E. and Zeikus, J.G. (1992) Molecular biology and physiological biochemistry of xylan degradation by thermoanaerobes. In: Progress in Biotechnology Vol. 7, Xylans and Xylanases (Visser, J., Beldman, G., Kusters-van Someren, M.A. and Voragen, A.G.J., Eds.), pp. 275–288. Elsevier, Amsterdam.
194 Fujino, T., Béguin, P. and Aubert, J.-P. (1993) Organization of a *Clostridium thermocellum* gene cluster encoding the cellulosomal scaffolding protein CipA and a protein possibly involved in the attachment of the cellulosome to the cell surface. J. Bacteriol. 175, 1891–1899.
195 Hansen, C.K. (1992) Fibronectin Type III-like sequences and a new domain type in prokaryotic depolymerases with insoluble substrates. FEBS Lett. 305, 91–96.
196 Meinke, A., Gilkes, N.R., Kilburn, D.G., Miller, R.C. Jr. and Warren, R.A.J. (1993) Cellulose-binding polypeptides from *Cellulomonas fimi*: endoglucanase D (CenD), a family A β-1,4-glucanase. J. Bacteriol. 175, 1910–1918.
197 Bisaria, V.S. and Mishra, S. (1989) Regulatory aspects of cellulase biosynthesis and secretion. CRC Crit. Rev. Biotechnol. 9, 61–103.
198 Wood, W.E., Neubauer, D.G. and Stutzenberger, F.J. (1984) Cyclic AMP levels during induction and repression of cellulase biosynthesis in *Thermomonospora curvata*. J. Bacteriol. 160, 1047–1054.
199 Montenecourt, B.S. (1983) *Trichoderma reesei* cellulases. Trends in Biotechnol. 1, 156–161.
200 Durand, H., Clanet, M. and Tiraby, G. (1988) Genetic improvement of *Trichoderma reesei* for large scale cellulase production. Enzyme Microb. Technol. 10, 341–346.
201 Stewart, B.J. and Leatherwood, J.M. (1976) Derepressed synthesis of cellulase by *Cellulomonas*. J. Bacteriol. 128, 609–615.
202 Johnson, E.A., Bouchot, F. and Demain, A.L. (1985) Regulation of cellulase formation in *Clostridium thermocellum*. J. Gen. Microbiol. 131, 2303–2308.
203 Yamane, K., Suzuki, H., Hirotani, M., Ozawa, H. and Nisizawa, K. (1970) Effect of nature and supply of carbon sources on cellulase formation in *Pseudomonas fluorescens* var. *cellulosa*. J. Biochem. 67, 9–28.
204 Béguin, P., Eisen, H. and Roupas, A. (1977) Free and cellulose-bound cellulases in a *Cellulomonas* species. J. Gen. Microbiol. 101, 191–196.
205 Kubicek, C.P., Messner, R., Gruber, F., Mach, R.L. and Kubicek-Pranz, E.M. (1993) The *Trichoderma reesei* cellulase regulatory puzzle – From the interior life of a secretory fungus. Enzyme Microb. Technol. 15, 90–99.
206 Kubicek, C.P. (1987) Involvement of a conidial endoglu-

canase and a plasma membrane-bound β-glucosidase in the induction of endoglucanase synthesis by cellulose in *Trichoderma reesei*. J. Gen. Microbiol. 133, 1481–1487.

207 Kubicek, C.P., Mühlbauer, G., Krotz, M., John, E. and Kubicek-Pranz, E.M. (1988) Properties of a conidial-bound enzyme system from *Trichoderma reesei*. J. Gen. Microbiol. 134, 1215–1222.

208 El-Gogary, S., Leite, A., Crivellaro, O., Eveleigh, D.E. and El-Dorry, H. (1989) Mechanism by which cellulose triggers cellobiohydrolase I gene expression in *Trichoderma reesei*. Proc. Natl. Acad. Sci. USA 86, 6138–6141.

209 Messner, R., Kubicek-Pranz, E.M., Gsur, A. and Kubicek, C.P. (1991) Cellobiohydrolase II is the main conidial-bound cellulase in *Trichoderma reesei* and other *Trichoderma* strains. Arch. Microbiol. 155, 601–605.

210 Seiboth, B., Messner, R., Gruber, F. and Kubicek, C.P. (1992) Disruption of the *Trichoderma reesei cbh2* gene coding for cellobiohydrolase II leads to a delay in the triggering of cellulase formation by cellulose. J. Gen. Microbiol. 138, 1259–1264.

211 Mandels, M., Parrish, F.W. and Reese, E.T. (1962) Sophorose as an inducer of cellulase in *Trichoderma viride*. J. Bacteriol. 83, 400–408.

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

213 Umile, C. and Kubicek, C.P. (1986) A constitutive, plasma-membrane bound β-glucosidase in *Trichoderma reesei*. FEMS Microbiol. Lett. 34, 291–295.

214 Messner, R., Gruber, F. and Kubicek, C.P. (1988) Differential regulation of synthesis of multiple forms of endoglucanases by *Trichoderma reesei* QM9414. J. Bacteriol. 170, 3689–3693.

215 Fowler, T. and Brown, R.D. (1992) The *bgl1* gene encoding extracellular β-glucosidase from *Trichoderma reesei* is required for rapid induction of the cellulase complex. Mol. Microbiol. 6, 3225–3235.

216 Hofer, F., Weissinger, E., Mischak, H., Messner, R., Meixner-Monori, B., Blaas, D., Visser, J. and Kubicek, C.P. (1989) A monoclonal antibody against the alkaline extracellular β-glucosidase from *Trichoderma reesei* β-glucosidases. Biochim. Biophys. Acta 992, 298–306.

217 Lin, E. and Wilson, D.B. (1987) Regulation of cellulase synthesis in *Thermomonospora fusca*. Appl. Environ. Microbiol. 53, 1352–1357.

218 Béguin, P., Rocancourt, M., Chebrou, M.-C. and Aubert, J.-P. (1986) Mapping of mRNA encoding endoglucanase A from *Clostridium thermocellum*. Mol. Gen. Genet. 202, 251–254.

219 Mishra, S., Béguin, P. and Aubert, J.-P. (1991) Transcription of *Clostridium thermocellum* endoglucanase genes *celF* and *celD*. J. Bacteriol. 173, 80–85.

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

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

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

223 Moser, B., Gilkes, N.R., Kilburn, D.G., Warren, R.A.J. and Miller, R.C. Jr. (1989) Purification and characterization of endoglucanase C in *Cellulomonas fimi*, the cloning of its gene, and analysis of in vivo transcripts of the gene. Appl. Environ. Microbiol. 55, 2480–2487.

224 Lao, G., Ghangas, G.S., Jung, E.D. and Wilson, D.B. (1991) DNA sequences of three β-1,4-endoglucanase genes from *Thermomonospora fusca*. J. Bacteriol. 173, 3397–3407.

225 Lin, E. and Wilson, D.B. (1988) Identification of a *celE*-binding protein and its potential role in induction of the *celE* gene in *Thermomonospora fusca*. J. Bacteriol. 170, 3843–3846.

226 Penttilä, M., Saloheimo, A., Ilmén, M. and Onnela, M.-L. (1993) Regulation of *T. reesei* cellulases at mRNA and promoter level. In: Abstracts of the TRICEL 93, 2nd International Symposium on *Trichoderma* cellulases and other hydrolases. pp. 27.

227 Aymeric, J.-L., Guiseppi, A., Pascal, M.-C. and Chippaux, M. (1988) Mapping and regulation of the *cel* genes in *Erwinia chrysanthemi*. Mol. Gen. Genet. 211, 95–101.

228 Penttilä, M., Nevalainen, H., Rättö, M., Salminen, E. and Knowles, J. (1987) A versatile transformation system for the cellulolytic filamentous fungus *Trichoderma reesei*. Gene 61, 155–164.

229 Wood, D.A., Claydon, N., Dudley, K.J., Stephens, S.K. and Allan, M. (1988) Cellulase production in the life cycle of the cultivated mushroom *Agaricus bisporus*. In: FEMS Symposium No. 43. Biochemistry and Genetics of Cellulose Degradation (Aubert, J.-P., Béguin, P. and Millet, J., Eds.), pp. 53–70. Academic Press, London and New York.

230 Penttilä, M., Lehtovaara, P. and Knowles, J. (1989) Cellulolytic yeasts and their applications. In: Yeast Genetic Engineering (Barr, P.J., Brake, A.J. and Valenzuela, P., Eds.), pp. 247–267. Butterworth, Stoneham, MA.

231 Jørgensen, O.B. and Cowan, D. (1989) Use of enzymes in feed and in ensiling. In: Enzyme Systems for Lignocellulose Degradation (Coughlan, M.P., Ed.), pp. 347–355. Elsevier Applied Science, London.

232 Linko, M., Poutanen, K. and Viikari, L. (1989) New developments in the application of enzymes for biomass processing. In: Enzyme Systems for Lignocellulose Degradation (Coughlan, M.P., Ed.) pp. 331–346. Elsevier, London.

233 Vanbelle, M. and Bertin, G. (1989) Screening of fungal cellulolytic preparations for application in ensiling processes. In: Enzyme Systems for Lignocellulose Degradation (Coughlan, M.P., Ed.) pp. 357–369. Elsevier Applied Science, London.

234 Oh Chit Soon, J. (1993) M.Sc. thesis, University of Singapore.
235 Hall, J., Ali, S., Surani, M.A., Hazlewood, G.P., Clark, A.J., Simons, J.P., Hirst, B.H. and Gilbert, H.J. (1993) Manipulation of the repertoire of digestive enzymes secreted into the gastrointestinal tract of transgenic mice. Bio/Technology 11, 376–379.
236 Joliff, G., Edelman, A., Klier, A. and Rapoport, G. (1989) Inducible secretion of a cellulase from *Clostridium thermocellum* in *Bacillus subtilis*. Appl. Environ. Microbiol. 55, 2739–2744.
237 Petit, M.-A., Joliff, G., Mesas, J.M., Klier, A., Rapoport, G. and Ehrlich, S.D. (1990) Hypersecretion of a cellulase from *Clostridium thermocellum* in *Bacillus subtilis* by induction of chromosomal DNA amplification. Bio/Technology 8, 559–563.
238 Silva, A., Benitez, J. and Hollenberg, C.P. (1991) Endoglucanase A gene fusion vectors for monitoring protein secretion and glycosylation in yeast. Anal Biochem. 197, 290–295.
239 Paice, M.G., Bernier, R. and Jurasek, L. (1987) Viscosity-enhanced bleaching of hardwood Kraft pulp with xylanase from a cloned gene. Biotechnol. Bioeng. 32, 235–239.
240 Ong, E., Gilkes, N.R., Warren, R.A.J., Miller, R.C. Jr. and Kilburn, D.G. (1989) Enzyme immobilization using the cellulose-binding domain of a *Cellulomonas fimi* exoglucanase. Bio/Technology 7, 604–607.
241 Hogsett, D.A., Ahn, H.J., Bernardez, T.D., South, C.R. and Lynd, L.R. (1992) Direct microbial conversion – Prospects, progress, and obstacles. Appl. Biochem. Biotechnol. 34, 527–541.
242 Brestic-Goachet, N., Gunasekaran, P., Cami, B. and Baratti, J.-C. (1989) Transfer and expression of an *Erwinia chrysanthemi* cellulase gene in *Zymomonas mobilis*. J. Gen. Microbiol. 135, 893–902.
243 Wong, W.K.R., Curry, C., Parekh, R.S., Parekh, S.R., Wayman, M.D., Davies, R.W., Kilburn, D.G. and Skipper, N. (1988) Wood hydrolysis by *Cellulomonas fimi* endoglucanase and exoglucanase coexpressed as secreted enzymes in *Saccharomyces cerevisiae*. Bio/Technology 6, 713–719.
244 Wood, B.E. and Ingram, L.O. (1992) Ethanol production from cellobiose, amorphous cellulose, and crystalline cellulose by recombinant *Klebsiella oxytoca* containing chromosomally integrated *Zymomonas mobilis* genes for ethanol production and plasmids expressing thermostable cellulase genes from *Clostridium thermocellum*. Appl. Environ. Microbiol. 58, 2103–2110.
245 Wang, D.I.C., Avgerinos, G.C., Biocic, I., Wang, S.D. and Fang, H.Y. (1983) Ethanol from cellulosic biomass. Phil. Trans. R. Soc. London B 300, 323–333.
246 Tailliez, P., Girard, H., Longin, R., Béguin, P. and Millet, J. (1989) Cellulose fermentation by an asporogenous and an ethanol-tolerant mutant of *Clostridium thermocellum*. Appl. Environ. Microbiol. 55, 203–206.
247 Kurose, N., Yagyu, J., Miyazaki, T., Uchida, M., Hanai, S. and Obayashi, A. (1988) Evaluation of ethanol productivity from cellulose by *Clostridium thermocellum*. J. Ferment. Technol. 64, 447–450.
248 Lynd, L.R., Grethlein, H.E. and Wolkin, R.H. (1989) Fermentation of cellulosic substrates in batch and continuous culture by *Clostridium thermocellum*. Appl. Environ. Microbiol. 55, 3131–3139.
249 Chuvilskaya, N.A., Atakishieva, Y.Y. and Akimenko, V.K. (1991) Production of ethanol-resistant mutants of *Clostridium thermocellum*. Microbiology 60, 78–84.
250 Béguin, P. and Aubert, J.-P. (1992) Cellulases. In: Encyclopedia of Microbiology (Lederberg, J., Ed.). Vol. 1, pp. 467–477. Academic Press, New York, NY.
251 Béguin, P., Millet, J. and Aubert, J.-P. (1992) Cellulose degradation by *Clostridium thermocellum*: from manure to molecular biology. FEMS Microbiol. Lett. 100, 523–528.
252 Pearson, W.R. and Lipman, D.J. (1988) Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA 85, 2444–2448.
253 Hansen, C.K. (1992) Ph.D. thesis, Technical University of Denmark.
254 Saito, T., Suzuki, K., Yamamoto, J., Fukui, T., Miwa, K., Tomita, K., Nakanishi, S., Odani, S., Suzuki, J.-I. and Ishikawa, K. (1989) Cloning, nucleotide sequence, and expression in *Escherichia coli* of a gene for poly-(3-hydroxybutyrate) depolymerase from *Alcaligenes faecalis*. J. Bacteriol. 171, 184–189.
255 Al-Tawheed, A. (1988) M.Sc. thesis, University of Dublin, Ireland.
256 Flint, H.J., Martin, J., McPherson, C.A., Daniel, A.S. and Zhang, J.-X. (1993) A bifunctional enzyme with separate xylanase and β-(1,3)-glucanase domains, encoded by the *xynD* gene of *Ruminococcus flavefaciens*. J. Bacteriol. 175, 2943–2951.
257 Nolte, A. and Mayer, F. (1989) Localization and immunological characterization of the cellulolytic enzyme system in *Clostridium thermocellum*. FEMS Microbiol. Lett. 61, 65–72.
258 Ljungdahl, L.G., Coughlan, M.P., Mayer, F., Mori, Y. Hon-Nami, H. and Hon-Nami, K. (1988) Macrocellulase complexes and yellow affinity substance from *Clostridium thermocellum*. Methods Enzymol. 160, 483–500.
259 Yamagata, H., Adachi, T., Tsuboi, A., et al. (1987) Cloning and characterization of the 5′ region of the cell wall protein gene operon in *Bacillus brevis* 47. J. Bacteriol. 169, 1239–1245.