ELSEVIER

Carbohydrate Research 298 (1997) 299–310

CARBOHYDRATE RESEARCH

Substrate specificity of endoglucanases: what determines xyloglucanase activity? [1]

Jean-Paul Vincken, Gerrit Beldman, Alphons G.J. Voragen *

Wageningen Agricultural University, Department of Food Science, P.O. Box 8129, 6700 EV Wageningen, The Netherlands

Received 15 July 1996; accepted 28 November 1996

Abstract

Endoglucanases from *Trichoderma viride* differ in their activity and mode of action towards xyloglucans. In order to explain the basis for their different behavior, the number of substrate-binding sites of three endoglucanases (endoI, endoIV, and endoV) were determined using bond cleavage frequencies of both normal and reduced cellodextrins and k_0/K_m. EndoIV differed from other endoglucanases described so far, in having at least nine putative binding sites. The specificities of the three endoglucanases towards various xyloglucans derived from apple fruit and potato were determined. Also, the release of oligosaccharides from these substrates in time was monitored. It was concluded that the endoglucanases prefer to bind unbranched glucosyl residues. Because most xyloglucans are composed of XXXG-type of building units, distant subsites are needed to bind xyloglucan. Having at least nine substrate-binding sites, endoIV seems to be well equipped to degrade xyloglucans which was confirmed by its high xyloglucanase activity.

Keywords: Endoglucanase; Xyloglucanase activity; Subsite mapping

Abbreviations: APfxg, apple fucoxyloglucan; APxg, apple xyloglucan after removal of Fuc; BCF, bond cleavage frequency; CBH, cellobiohydrolase; CMC, carboxymethyl cellulose; EG, endoglucanase described in the literature; endoI, etc., endoglucanase I from *Trichoderma viride*, etc.; Glc*, glucitol; G_n, reducing cellodextrin with chain length n; G_n^*, cello-alditol with chain length n; $G_p \downarrow G^*_{n-p}$, cello-alditol with chain length n which gives the products G_p (with chain length p) and G^*_{n-p} (with chain length $n-p$) after cleavage by an endoglucanase; HPAEC, high-performance anion-exchange chromatography; PAD, pulsed amperometric detection; PED, pulsed electrochemical detection; POsxg, potato arabinoxyloglucan; R, gas constant (8.314 $J\,mol^{-1}K$); T, temperature (K); XXXG, *etc.*, nomenclature of xyloglucan oligosaccharides according to Fry et al. [38]

* Corresponding author. Tel.: +31-317-482888; fax: +31-317-484893; e-mail: fons.voragen@algemeen.lenm.wau.nl.

[1] This work was supported by a grant from Gist-brocades, Delft, The Netherlands.

0008-6215/97/$17.00
PII S0008-6215(96)00325-4

1. Introduction

Nature has equipped many fungal and bacterial species with a large set of (1 → 4)-β-glucanases, sometimes containing as many as seven different enzymes [1]. From a genetic point of view, these enzymes are often rather well characterized. For instance, most of the glucanases of *Trichoderma reesei* have been sequenced and cloned. These are EGI [2], EGII (formerly called EGIII) [3], EGIII [4], EGV [5], CBHI [6], and CBHII [7]. Analysis and comparison of amino acid sequences, as well as limited proteolysis [8], have revealed that most of these glucanases are built of discrete functional domains; a cellulose binding domain and a catalytic core [9]. These two domains are joined by a linker peptide, of which both length and structure greatly depend on microorganism and enzyme [9]. It has been shown for CBHI that sufficient spatial separation of the two domains is required for efficient function of the enzyme [10].

The specificity of cellulases is determined by the catalytic core, and in this domain, the different glucanases share little amino acid sequence identity in a linear alignment. However, hydrophobic cluster analysis [11–13] suggests that the fold of proteins is better conserved than the sequence of their amino acids. Using this technique, glucanases can be grouped into 10 families, and it is expected that the same fold will be found for each member of a family [12].

Cellulases can cleave (1 → 4)-β-glycosidic linkages in a variety of substrates such as cellulose, CMC, (1 → 3),(1 → 4)-β-D-glucan [5], and xylan [14,15]. Further, they show activity towards several small chromophoric glycosides. Each cellulase family demonstrates a characteristic specificity pattern on these artificial substrates [16]. Until now, relatively little attention has been paid to their activity towards xyloglucans, although this is an important polysaccharide in the primary cell wall of plants [17]. We have shown previously that the glucanases of *Trichoderma viride* (similar to *T. reesei* [18]) differ greatly in their activity towards apple xyloglucan (APfxg) [19]. The structure of APfxg is illustrated in Fig. 1. APfxg is mainly built of oligosaccharide building units having a backbone of four Glc residues. Small amounts of building units with a backbone length of only two or three Glc residues also occur [20,21]. Degradation of APfxg by endoglucanases is restricted to only a few linkages (indicated by arrows in Fig. 1); Fuc residues inhibit the action of endoglucanases to some extent [20]. Potato xyloglucan (POsxg) has a

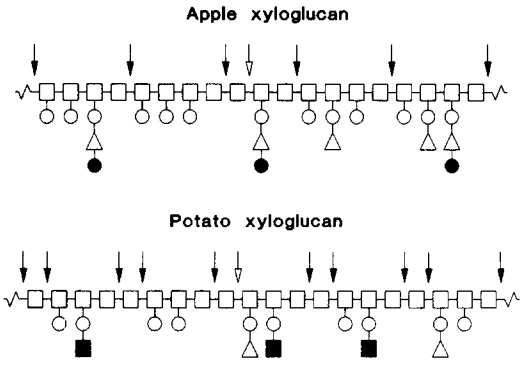

Fig. 1. Schematic structures of APfxg and POsxg. □, β-D-Glc*p*-(1 → 4)-; ○, α-D-Xyl*p*-(1 → 6)-; △, β-D-Gal*p*-(1 → 2)-; ●, α-L-Fuc*p*-(1 → 2)-; ■, α-L-Ara*f*-(1 → 2)-. Sites of endoglucanase-attack are indicated by arrows. Cleavage of linkage indicated with open arrow is uncertain.

larger number of cleavage sites than APfxg (Fig. 1), and the linkage that is actually cleaved depends on the endoglucanase used [22]. These observations prompted us to study the basis of the specificity of endoglucanases in more detail.

Hydrolysis by glycosyl hydrolases often proceeds by general acid catalysis, usually promoted by Glu or Asp residues [23–25]. The fact that the same catalytic amino acids are found in hydrolases with different target substrates suggests that catalytic residues do not determine the substrate specificity of these enzymes. Despite these similarities, the stereochemistry of the products of a hydrolysis reaction can be altered. The anomeric configuration may either be inverted (single displacement) or retained (double displacement) during hydrolysis [26], but within a given enzyme family this stereoselectivity seems to be consistent [16,27]. It is not very likely that the mechanism of catalysis determines substrate specificity, because glucanases with different specificities (exo versus endo) can belong to the same family as is the case with, for instance, CBHI and EGI (family 7).

Crystallography showed that the catalytic residues of CBHI from *Trichoderma reesei* are located in a relatively long substrate-binding tunnel (compared to CBHII), which can slide along a glucan chain [28]. Several loops, which are involved in building this tunnel, appear to be deleted in the EGI protein [2,28]. Similar observations were made for CBHII from *Trichoderma reesei* and EG2 from *Thermomonospora fusca*, both belonging to family 6. Recently, it has been shown that deletion of a surface loop of CBH A from *Cellulomonas fimi* resulted in an enhancement of the endoglucanase activity of this

enzyme [29]. These experiments suggest that exoglucanases may be ancestors of related endoglucanases [28], in which the tunnel has evolved to a more open structure (groove) to facilitate multiple-chain attack.

Although the tunnel/groove concept may explain the basis for an exo/endo type of specificity, it does not provide an answer to the question of why some glucanases have xyloglucanase activity and others have little. It has been suggested that the basis of substrate specificity of glycosyl hydrolases is determined by their ability to bind different polysaccharides in their substrate-binding sites. Evidence for this was obtained by studying the interaction between various active site-directed inhibitors and β-glucan endohydrolases [30]. Epoxyalkyl β-oligoglucosides, varying in glycosyl chain length, linkage type, and aglycon length appeared to inactivate these enzymes to a different extent. This example illustrates the key role of substrate-binding in enzyme specificity.

In order to investigate the basis for xyloglucanase activity, the number of substrate-binding sites of three endoglucanases from *Trichoderma viride* (endoI, endoIV, and endoV) [31] were determined according to the method of Suganuma et al. [32]. Homologous cellodextrins were used for this purpose. Further, the specificity of the endoglucanases towards various xyloglucans was studied, as well as the release of oligosaccharides from these xyloglucans in time. A possible relationship between the number of substrate-binding sites, specificity, and mode of action of these glucanases is discussed.

2. Experimental

Substrates.—APfxg and POsxg were obtained by extraction from, respectively, apple and potato cell wall material with strong alkali (4 M KOH) and purified as described previously [19,22]. CMC (Akucell AF type 0305) was purchased from Akzo (Arnhem, The Netherlands). Avicel (type SF) was obtained from Serva (Heidelberg, Germany); GlcA and G_2 were purchased from Fluka AG (Buchs, Switzerland). Glc and cellodextrins (mixture I; Glc to G_6) were obtained from Merck (Darmstadt, Germany).

Enzymes.—Three endoglucanases (endoI, endoIV, and endoV; EC 3.2.1.4) were purified to homogeneity from a commercial enzyme preparation from *Trichoderma viride* (Maxazyme Cl, Gist-Brocades, Delft, The Netherlands) by Beldman et al. [31]. The enzymes were electrophoretically pure and devoid of side activities.

Preparation of normal cellodextrins.—Normal cellodextrins (Glc to G_6) were prepared by partial hydrolysis of 6 g Avicel with 15 mL 80% (w/v) H_2SO_4 according to a slightly modified procedure of Voloch et al. [33]. Portions of 500 mg of the hydrolysate were dissolved in 2 mL of distilled water, applied onto a column (100 × 2.6 cm, i.d.) of BioGel P-2 (200–400 mesh, Bio-Rad, Richmond, CA, USA) at 60 °C and eluted with distilled water (18 mL h^{-1}). Appropriate fractions (2.0 mL) were combined according to their neutral sugar content. The column was calibrated using a mixture of Glc, G_2, and Dextran T150 (Pharmacia, Uppsala, Sweden). The purified cellodextrins were concentrated under reduced pressure, freeze-dried, and redissolved to give a stock solution of ca. 5 mM. The concentration of cellodextrins in this solution was determined colorimetrically.

Preparation of reduced cellodextrins.—Reduced cellodextrins (Glc* to G_6^*) were prepared by treating ca. 2.5 mmol of the purified normal cellodextrins with 500 μL of 1.5 M NH_4OH containing 75 mg mL^{-1} $NaBH_4$ for 1 h at 30 °C. Reaction mixtures were acidified by dropwise addition of glacial HOAc. Glc* and G_2^* were desalted using columns (30 × 80 mm) of Dowex 50W X8 (H^+) and AG3 X4A (OH^-) resins (Bio-Rad) in series. G_3^* to G_6^* were desalted on a Sephadex G-10 column (Pharmacia, 26 × 400 mm). All samples were concentrated under reduced pressure, dried under a stream of air, and redissolved to give a stock solution of ca. 5 mM. The concentration of reduced cellodextrins in this solution was determined colorimetrically.

Removal of Fuc from APfxg by mild acid hydrolysis.—APfxg (40 mg) was treated with 8 mL of 25 mM TFA for 90 h at 60 °C. Approximately 95% of the Fuc residues could be removed by this treatment. The defucosylated xyloglucan (referred to as APxg) was dialyzed extensively against distilled water and freeze-dried.

Determination of BCFs.—BCFs were determined in 400 μL 20 mM succinate buffer (pH 4.0; 40 °C) containing 25 μM of a normal or reduced cellodextrin (appropriately diluted from stock solutions G_2 to G_6 or G_3^* to G_6^*) and 15 μM GlcA as internal standard. Incubation time and enzyme concentration were such that the evolution of a second generation of degradation products was minimized. Samples were heated for 10 min at 100 °C to inactivate the enzymes, evaporated to dryness under a stream of air, and dissolved in 100 μL of water. The individual degradation products were quantified by HPAEC.

Response factors for the individual oligosaccharides were accounted for.

Determination of the kinetic parameter (k_0/K_m). —For these experiments only normal cellodextrins were used. From an identical reaction mixture (4.0 mL) as described above, 400 μL aliquots were taken at various time intervals and handled as above. From the HPAEC data a $\ln([S]_0/[S])$ versus time plot could be constructed from which the kinetic parameter k_0/K_m was determined.

Molar activity towards various xyloglucans.— APfxg, APxg, and POsxg (500 μg) were dissolved in 100 μL of 50 mM succinate buffer (pH 4) containing 0.01% (w/v) NaN_3 as a preservative, and incubated for 1 or 12 h at 40 °C with of endoI, endoIV or endoV. Enzyme dosage was such that no substrate limitation occurred. Incubations were stopped by heating the reaction mixtures for 10 min at 100 °C. The samples were then diluted twice and the increase in reducing sugars was determined according to the procedure of Somogyi [34] using Glc for calibration. Similar experiments were done using CMC as a substrate.

Release of oligosaccharides from various xyloglucans in time.—APfxg and APxg (2 mg dissolved in 1 mL of 50 mM succinate buffer, pH 4, containing 0.01% (w/v) NaN_3) were (partially) degraded with endoI (ca. 10 μg), endoIV (ca. 100 ng) or endoV (ca. 500 ng) during 24 h at 40 °C. Using this enzyme dosage, each incubation mixture contained a similar amount of xyloglucanase activity. Samples of 50 μL were taken during the course of degradation (0, 1, 2, 3, 4, 8, 12, and 24 h). Both APfxg and APxg were degraded completely by a 10-fold larger dose of endoIV (ca. 1 μg). After inactivation of the enzyme (10 min, 100 °C), the oligosaccharides were analyzed by HPAEC. The amount of an oligosaccharide that was released in a certain time interval, was expressed as the percentage of the PAD response of this oligosaccharide in the chromatogram of the completely degraded xyloglucan. POsxg was treated in a similar manner, but using a different enzyme dose: 5 μg, 0.140 μg, and 5 μg of endoI, endoIV, and endoV, respectively.

Total neutral sugar content.—The total neutral sugar content was determined colorimetrically with an automated orcinol/H_2SO_4 assay [35]. Glc was used as a standard.

Neutral sugar composition.—APxg was hydrolyzed (1 h, 121 °C) using 2 M TFA. The released neutral sugars were determined by HPAEC as described by de Ruiter et al. [36].

Protein content.—The protein content of enzyme preparations was determined according to Sedmak and Grossberg [37]. Bovine serum albumin was used as a standard.

Analysis of normal and reduced cellodextrins.— Quantitative analysis of normal and reduced cellodextrins was performed with HPAEC using a Dionex Bio-LC GPM-II quaternary gradient module equipped with a Dionex CarboPac PA-100 column (250 × 4 mm, Dionex, Sunnyvale, CA, USA) at 20 °C. Samples (20 μL) were injected using a SP8780 autosampler (Spectra Physics, San José, CA, USA) equipped with a Tefzel rotor seal in a 7010 Rheodyne injector valve. Solvents were degassed and stored under helium using a Dionex eluent degassing module. The eluate (1 mL min^{-1}) was monitored using a Dionex PED detector in the PAD mode. A reference Ag/AgCl electrode was used with a working gold electrode with the following pulse potentials and durations: E_1, 0.1 V and 0.5 s; E_2, 0.6 V and 0.1 s; E_3, −0.6 V and 0.1 s.

Mixtures containing both normal and reduced cellodextrins were eluted as follows: 0 → 2 min, 200 mM NaOH (isocratic); 2 → 24 min, 0 → 110 mM NaOAc (linear gradient) and simultaneously 200 → 189 mM NaOH (linear gradient); 24 → 30 min, equilibration of the column in 200 mM NaOH (isocratic). After eight runs the column was washed with 1 M NaOAc in 100 mM NaOH for 10 min and equilibrated for 15 min in 200 mM NaOH. An internal standard of GlcA was required to provide accurate quantitations because of the response fluctuation of the electrochemical detector.

For analysis of mixtures containing only normal cellodextrins a different gradient was used. A complete separation of these oligosaccharides was achieved by the application of a linear gradient to 100 mM NaOH/140 mM NaOAc over 14 min, after the column had been equilibrated in 100 mM NaOH containing 40 mM NaOAc for 6 min. After eight runs the column was washed with 1 M NaOAc in 100 mM NaOH for 10 min and equilibrated for 15 min in 100 mM NaOH containing 40 mM NaOAc.

Analysis of xyloglucan oligosaccharides.—The degradation of APfxg, APxg, and POsxg was monitored by HPAEC using a similar system as described above. Oligosaccharides derived from APfxg or APxg, were eluted with the following NaOAc gradient in 100 mM NaOH: 0 → 5 min, linear gradient of 0 → 30 mM NaOAc; 5 → 45 min, linear gradient of 30 → 80 mM NaOAc; 45 → 55 min, linear gradient of 80 → 200 mM NaOAc. Oligosaccharides derived

from POsxg were analyzed with a different NaOAc gradient, also in 100 mM NaOH: 0 → 5 min, linear gradient of 0 → 50 mM NaOAc; 5 → 30 min, linear gradient of 50 → 100 mM NaOAc; 30 → 45 min, linear gradient of 100 → 250 mM NaOAc. After each analysis, the column was rinsed for 5 min with 1 M NaOAc in 100 mM NaOH, and equilibrated in 100 mM NaOH for 15 min. The identification of xyloglucan oligosaccharides is described elsewhere [20–22]. Nomenclature of these oligosaccharides is according to Fry et al. [38].

3. Results and discussion

BCFs of normal and reduced cellodextrins.—The individual normal cellodextrins were incubated with the three endoglucanases at low substrate concentrations (25 μM), and analyzed by HPAEC. Under the conditions used, no evidence for multi-substrate reaction mechanisms (transglycosylation or condensation) was found. In preliminary studies only endoV showed transferase activity at increased substrate concentrations (1 mM) of G_4, G_5, and G_6.

BCFs were calculated from the peak areas of the degradation products of purified oligosaccharides using their molar response factors. Incubation time was such that all products were the result of one cleavage of the original substrate (molar amount of products formed is ca. twice the decrease in molar amount of the original material); in other words, no 'second generation' of products was present. None of the endoglucanases had the ability to hydrolyze G_2. The three endoglucanases showed a different mode of action towards normal cellodextrins (Fig. 2A). The BCFs of symmetrical cleavages in Fig. 2A are real; in all other cases the BCFs are distributed over two bonds. To quantify all cleavages other than $G_2 \downarrow G_2$ and $G_3 \downarrow G_3$, the cellodextrins have to be labelled.

Among the methods currently in use to label oligosaccharides for determination of BCFs, radio-labelling is probably the most elegant method because it does not alter the substrate significantly. Degradation products are then separated using TLC and quantified with a scintillation counter [15,32,39,40]. A disadvan-

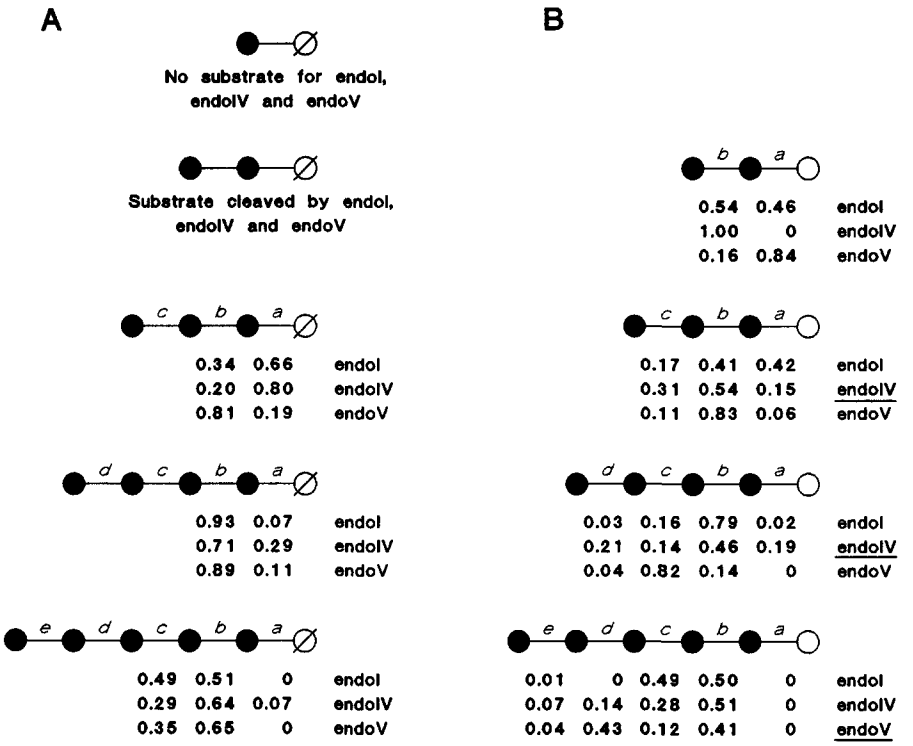

Fig. 2. BCFs of normal (A, G_2 to G_6) and reduced cellodextrins (B, G_3^* to G_6^*) by three endoglucanases of *Trichoderma viride*. In case of asymmetric cleavage of normal cellodextrins (not $G_2 \downarrow G_2$ and $G_3 \downarrow G_3$) the BCFs are the sum of the cleavages of the indicated linkage and their 'in reverse' on the nonreducing side. For instance, BCFs denoted under linkage *a* of G_5, is the sum of BCF of linkage *a* and that of *d*. Enzymes are underlined if their mode of action is influenced by reduction of the cellodextrin. Concentration of cellodextrins was 25 μM in 20 mM succinate buffer (pH 4.0), containing 15 μM GlcA; 40 °C. ●, nonreducing Glc residue; ∅, reducing Glc residue; ○, Glc residue labelled by reduction with $NaBH_4$; *a* to *e*, indicate different linkages starting from the (former) reducing terminus.

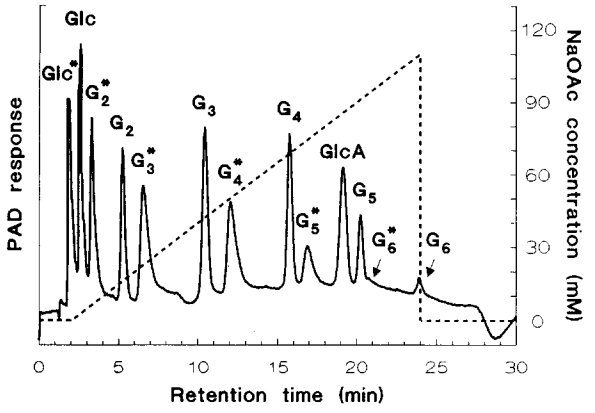

Fig. 3. Elution pattern of normal and reduced cellodextrins upon HPAEC. Symbols: ———, PAD response; ---, NaOAc gradient. Elution times of normal (Glc to G_6) and reduced cellodextrins (Glc* to G_6^*) are indicated.

tage of this method is its inability to quantify unlabelled products. This quantitation is important to verify that no second generation of degradation products is formed. The sum of molar amounts of all labelled products should be equal to that of all unlabelled products. Also, nucleophilic competition between water and methanol has been reported to allow unequivocal assessment of cleavage sites [41]. Other methods involve ring-opening by reduction [1,42,43], or coupling to chromophoric groups such as 4-methyl-umbelliferyl [16,41,42].

In our study, reduced cellodextrins, combined with HPAEC for product analysis, were used to deduce BCFs in case of asymmetric cleavage. Our procedure to analyze normal and reduced cellodextrins simultaneously is very similar to that of Bray and Clarke [43], differing slightly in NaOAc gradient and in internal standard (GlcA instead of Ara). A satisfactory separation could be achieved as is shown in Fig. 3. Under the conditions used, the CarboPac PA-100 column achieves a superior separation compared to other techniques [44,45], and does not require derivatization of the oligosaccharides. BCFs of reduced cellodextrins were determined in a similar manner as described for the normal ones. Fig. 2B demonstrates different cleavage patterns for the three endoglucanases. EndoI and endoV seemed to act in a reverse fashion in that the former enzyme preferentially released G_2^* whereas the latter releases G_2. Sites of attack were more evenly distributed over the alditols in case of endoIV; however, there was some preference to form G_2^*.

Until now, it was not known to what extent reduction affects the mode of action of the endoglucanases of *Trichoderma viride*. Comparison of the BCFs of G_4^* to G_6^* and G_4 to G_6 should indicate whether reduction really affects the interaction between Glc residues and subsites. For instance, the BCF of linkage '*b*' of G_4^* should be equal to that of linkage '*b*' of G_4 (symmetrical cleavage). Also, the sum of BCFs of linkage '*a*' and '*c*' of G_4^* should be similar to that indicated underneath linkage '*a*' of G_4 (asymmetrical cleavage). When deviations larger than 0.10 (corresponding to two times the experimental error) occurred, reduction was thought to influence the mode of action of the enzyme. These cases are underlined in Fig. 2B. From Fig. 2 it must be concluded that the effect of reduction on BCFs is largely enzyme dependent. EndoIV was affected to a larger extent than endoV, whereas endoI was not affected. The chain length of the oligosaccharides also seemed to be important in this respect, although no clear correlation was found. Reduction of the oligosaccharides did not simply shift the location of cleavage towards linkage '*e*'. Instead, the effects on BCFs were rather unpredictable. Similar observations were done for the endoglucanases of *Penicillium pinophilum* [42]. Here, enzyme action on certain substrates was completely abolished after reduction. On the other hand, Schou et al. [1] reported for the cellulases of *Humicola insolens* that the open alditol unit at the end of reduced oligosaccharides did not alter their hydrolysis patterns much.

Determination of the kinetic parameter k_0/K_m. —By monitoring rate of substrate loss in time for G_3 to G_6, $\ln([S]_0/[S])$ versus time plots could be con-

Table 1
Kinetic parameter k_0/K_m of three endoglucanases of *Trichoderma viride* for normal cellodextrins (G_3 to G_6). The kinetic parameter was evaluated from the slope of the linear part of $\ln([S]_0/[S])$ versus time plots

Enzyme	k_0/K_m ($min^{-1}M^{-1}$) for			
	G_3	G_4	G_5	G_6
EndoI	$2 \times 10^{+5}$	$1 \times 10^{+7}$	$11 \times 10^{+7}$	$9 \times 10^{+7}$
EndoIV	$\ll 1 \times 10^{+3}$	$0.081 \times 10^{+7}$	$4 \times 10^{+7}$	$21 \times 10^{+7}$
EndoV	$35 \times 10^{+5}$	$5 \times 10^{+7}$	$7 \times 10^{+7}$	$8 \times 10^{+7}$

structed for the three endoglucanases (data not shown). All plots proved to be linear for a limited period of time. Assuming that $[S]_0 \ll K_m$, the slope of these lines equals the kinetic parameter V/K_m from which k_0/K_m can be deduced by dividing by the enzyme concentration [32]. Kinetic parameters are summarized in Table 1. The cellodextrins were degraded faster with increasing chain length, except for endoI (optimal substrate is G_5). This increase in k_0/K_m was most spectacular with endoIV and probably its maximum activity has not yet been reached. Unfortunately, no oligosaccharides larger than G_6 could be used due to their poor solubility.

Subsite mapping of endoI, endoIV, and endoV.— The first step in enzymic hydrolysis of polysaccharides can be envisaged as binding of several glycosyl residues with an array of subsites; each subsite interacts with one sugar residue. Homologous cellodextrins have often been used as tools to elucidate the subsite structure of β-glucanases [1,15,39–42,46]. According to Suganuma et al. [32] the number of subsites, and their corresponding affinities, can be deduced conveniently from BCFs combined with the kinetic parameter k_0/K_m, provided certain requirements are met. Parameters should be determined at very low substrate concentrations to simplify the Michaelis–Menten equation ($[S]_0$ presumably much smaller than K_m), and to prevent bi-substrate processes, such as transglycosylation and condensation reactions. Further, BCFs should be evaluated with first generation degradation products. Subsite affinity is defined as the decrease in free energy upon interaction of an enzyme subsite with a glycosyl residue for the transition state of the reaction. The interaction energy of one subsite can then be calculated by subtracting the two appropriate free energies for binding of an oligosaccharide with a chain length of n and $n-1$.

BCFs of G_3^* to G_6^* and the k_0/K_m values of their reducing homologues, G_3 to G_6 (see Fig. 2 and Table 1, respectively) meet the requirements of the method of Suganuma et al. [32]. Subsite affinities for Glc residues were calculated for ten hypothetical subsites of endoI, endoIV, and endoV as is shown in Fig. 4. The two subsites adjacent to the catalytic group cannot be evaluated by this procedure [32]. Presumably, the sum of subsite affinity '−I' and that of 'I' is negative for the three endoglucanases used in this study because G_2 is not cleaved by these enzymes. Small BCFs were not considered in these calculations because they are liable to relatively large experimental errors. The number of subsites clearly differed for

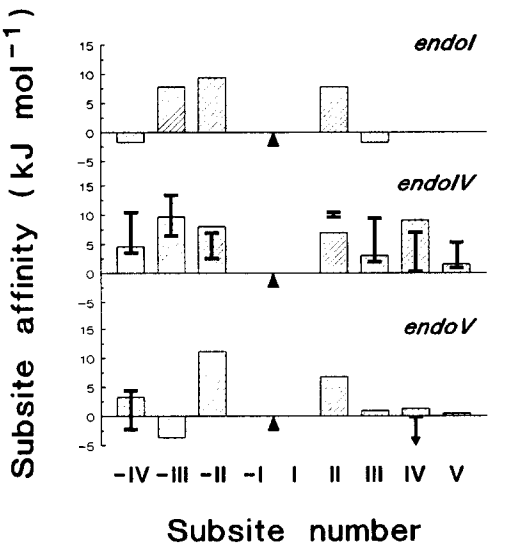

Fig. 4. Histogram of subsite interaction energy with Glc residues of cellodextrins for three endoglucanases of *Trichoderma viride*. Subsite affinities were determined according to the method of Suganuma et al. [32] using BCFs of reduced cellodextrins (shaded bars). Vertical lines in bars (I) indicate the lower and upper limit values of subsite affinities obtained after reevaluation (for details, see text). Subsite affinities of '−I' and 'I' were not determined. Lines with arrow head indicate that values are off-scale. ▲ indicates the catalytic group.

the three endoglucanases. EndoIV had the largest number of subsites (at least 9; '−IV' → 'V'), whereas endoI had only 5 subsites ('−III' → 'II'). EndoV has 7 substrate-binding sites ('−IV' → 'III') although the affinity of 'III' is rather low. Typically, the negative affinity of subsite '−III' was flanked by positive affinities of subsite '−IV' and '−II'. Due to the limited solubility of G_7 and larger cellodextrins more distant subsites like '−V' and 'VI' cannot be evaluated. However, their existence cannot be excluded.

As normal and reduced cellodextrins can be positioned differently in the substrate-binding sites of endoIV and endoV, all subsite affinities calculated with unreliable BCFs of alditols (see Fig. 2) were reevaluated. Recalculation leads to lower and upper limit values of subsite affinities rather than absolute values. Subsite '−II' of endoIV will be used as an example to illustrate our approach. The affinity of this subsite was calculated using G_4^* and G_5^* (Fig. 5). However, since reduction influenced the BCFs of both cellodextrins, the subsite affinity of '−II' was also calculated using G_5 and G_6^*, although the BCFs of these oligosaccharides are smaller. In this case, 29% of all cleavages of G_5 gave Glc and G_4. The exact location of cleavage [linkage number '*a*' ($G_4 \downarrow G$) or '*d*' ($G \downarrow G_4$)] is, however, unknown. As-

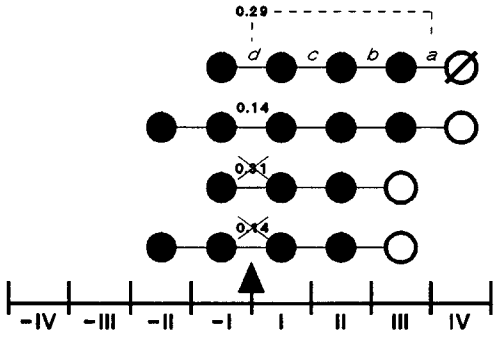

Fig. 5. An example of reevaluation of subsite affinities of subsite '−II' of endoIV. Different binding modes of normal and/or reduced cellodextrins for (re)evaluation of this subsite are indicated. Numbers indicate the BCF for that particular linkage. ×, Unreliable BCF (see Fig. 3); ---, the BCF is distributed over the two indicated linkages *a* and *d*; other symbols, as in Fig. 2Fig. 4.

suming that only linkage number '*d*' is split (BCF = 0.29), recalculation of this subsite affinity yields $RT \times \ln[(0.14 \times 2.1 \times 10^{+8})/(0.29 \times 3.9 \times 10^{+7})] = 2.5$ kJ mol^{-1}. However, the other extreme case that linkage number '*d*' is not split at all (BCF = 0) is equally probable. Here, recalculating the subsite affinity does not allow any conclusion (ln 0). Therefore, a BCF of 0.05 was assumed instead, yielding a subsite affinity of 7.0 kJ mol^{-1}. The value of 0.05 corresponds to the experimental error. Other substrate-binding sites were reevaluated in a similar manner. For endoIV the result remains unchanged, as is shown by the lines (l) in Fig. 4. This enzyme has a long array of substrate-binding sites (at least nine). The positive affinity of subsite '−IV' of endoV should be approached with some caution.

The nomenclature of our endoglucanases deviates from that generally adopted in literature. In order to compare our data with those in literature, we have tentatively characterized endoI, endoIV, and endoV as being EGII [2,41], EGIII [4], and EGI [3,15], respectively. The presumed homology is based on similarities in molecular weight, iso-electric point [31], adsorption behavior [47], the ability to degrade xylan [14], and transferase activity. For endoIV, we have obtained an internal sequence of 20 amino acids [48], which was 90% homologous with part of the sequence reported for EGIII [4]. Our results on the number of subsites of endoI and endoV are in agreement with those of EGII [41] and EGI [15,39], respectively. The k_0/K_m for G_6 of EGI was not determined by these authors, and consequently a putative binding site '−IV' could not be evaluated. Our results suggest a reasonable probability of a positive affinity of subsite '−IV', but this should be further substantiated using another method for labelling of cellodextrins instead of reduction. Until now no information was available on the number of subsites of endoIV (EGIII). With at least 9 subsites, this enzyme clearly differs from all other endo-(1 → 4)-β-glucanases from *Trichoderma* described so far.

Activity towards various xyloglucans.—The activities of the three endoglucanases towards different substrates are summarized in Table 2. Although the activity towards CMC was in the same order of magnitude for all three enzymes, their specificity for xyloglucans differed considerably. EndoIV and, to a lesser extent, endoV are far better xyloglucanases than endoI. Despite its high degree of backbone branching, endoIV degraded xyloglucan twice as fast as CMC. In accordance with [20], Table 2 shows that removal of Fuc from APfxg enhances the action of all three endoglucanases, especially that of endoI. The presence of extra unsubstituted Glc residues, as is the case in POsxg (see Fig. 1), facilitated cleavage by the endoglucanases. The activity of endoI and endoV were affected to a larger extent than that of endoIV. The activity of endoI on POsxg is relatively low which suggests that endoI requires larger stretches than two unbranched Glc residues to achieve an activity comparable to that of endoIV or endoV. For comparison activities towards Avicel crystalline cellulose [31] and oat spelts xylan [14] were included in Table 2. Typically, endoglycanases with a high activ-

Table 2
Molar activity [(mole reducing end groups formed)min^{-1}(mole of enzyme)$^{-1}$] towards CMC and various xyloglucans by three endoglucanases from *Trichoderma viride*. Activity towards Avicel and oat spelts xylan is indicated qualitatively: −, no detectable activity; ±, very low activity; +, low activity; + +, high activity

	CMC	APfxg	APxg	POsxg	Avicel [a]	Xylan [b]
EndoI	1518	9	32	36	+	−
EndoIV	1139	2017	2237	2390	±	+ +
EndoV	1049	600	781	1149	+	+ +

[a] Determined by Beldman et al. [31].
[b] Determined by Beldman et al. [14].

ity towards xyloglucan are also active towards xylan, although the two polysaccharides differ considerably in their degree of backbone branching as well as in their conformation [49,50].

Based on degradation patterns of small chromophoric substrates, Claeyssens and Henrissat [16] showed that substrate specificity within one cellulase family is consistent. EndoIV, which was tentatively classified as low-molecular-weight endoglucanase or EGIII, belongs to cellulase family 12 [51] which has three more representatives until now: CelS from *Erwinia carotovora* [52], FI-CMCase from *Aspergillus aculeatus* [53], and EGIII from *Humicola insolens* [1]. It is expected that all members of family 12 possess xyloglucanase activity; for the endoglucanase from *Aspergillus aculeatus* this has already been confirmed [54]. EndoV appears to be the most versatile endoglucanase, having activity on various natural substrates as cellulose, xylan, and xyloglucan. EndoV is probably similar to EGI (as discussed before), which belongs to family 7. Only one other endoglucanase, of unknown xyloglucanase activity, has been reported to belong to this family [1].

Release of oligosaccharides from various xyloglucans.—The amount of oligosaccharides released from APfxg was determined by HPAEC at various time intervals as is shown Fig. 6. HPAEC did not achieve separation of XXG and GLG; therefore, one line representing the sum of these two oligosaccharides is shown. It should be realized that the abundance of the individual oligosaccharides in apple xyloglucan is different. Fig. 6 demonstrates that the rate at which the individual oligosaccharides are released from xyloglucan, is determined by the structure of the oligosaccharide as well as by the endoglucanase being used. The most important trends are summarized

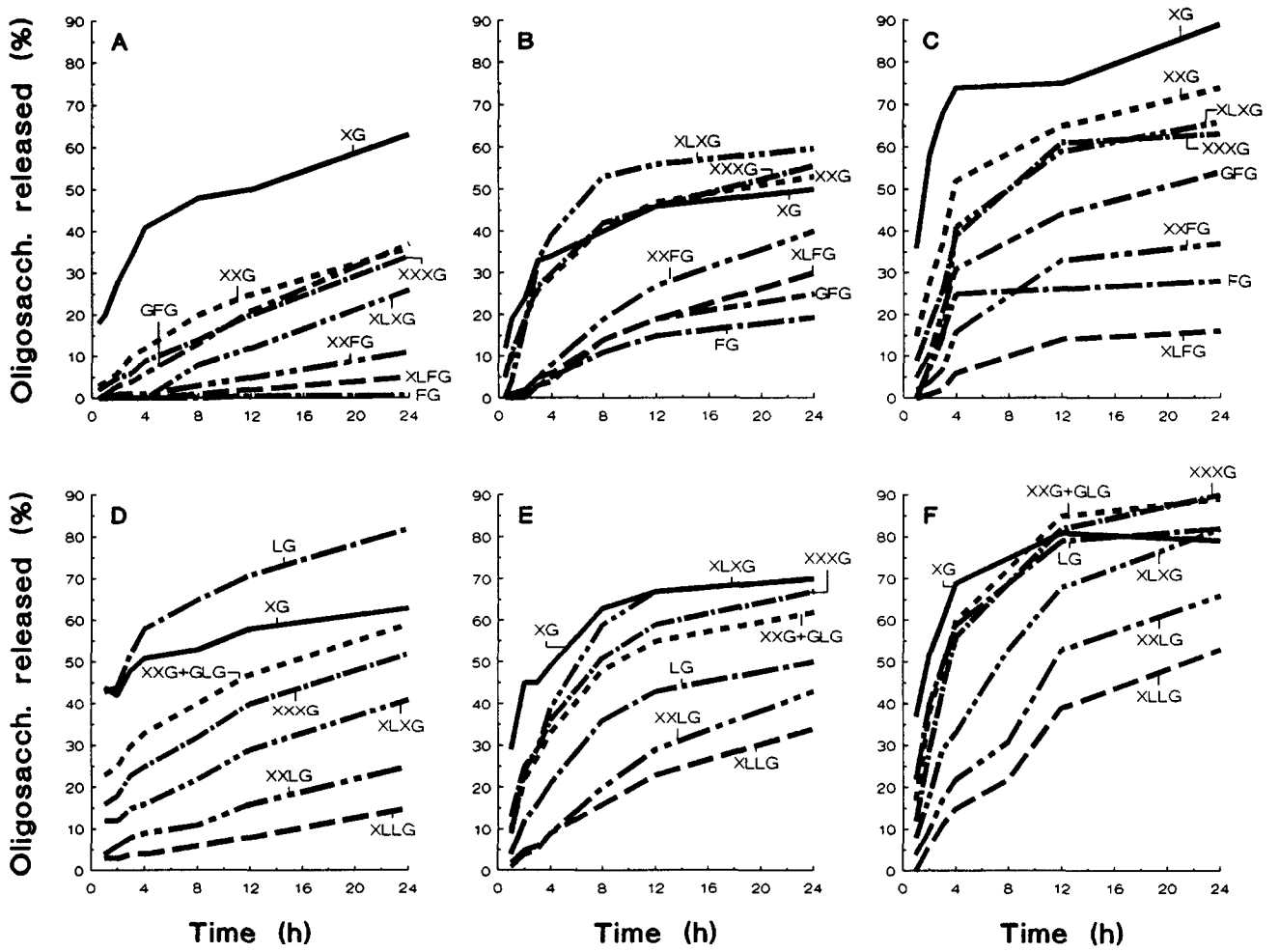

Fig. 6. Release of oligosaccharides from APfxg (A–C) and APxg (D–F) in time by the action of endoI (A, D), endoIV (B, E), and endoV (C, F). The amount of a certain oligosaccharide that was released is expressed as the percentage of the PAD response of this oligosaccharide in the chromatogram of APfxg or APxg which was completely degraded by endoIV. Different xyloglucan oligosaccharides are indicated according to the nomenclature of Fry et al. [38].

below. The sidechain configuration of the penultimate Glc residue of an oligosaccharide building unit determines the rate of release of this oligosaccharide to a large extent. Fuc-containing oligosaccharides generally accumulate slower than other oligosaccharides. The relatively fast release of GFG by endoI forms an exception. Removal of Fuc can have a large effect on the release of some oligosaccharides (compare FG and LG by endoI). EndoI and endoV are more affected by Fuc residues than endoIV, which is in accordance with Table 2. Also, Gal residues probably play a role considering the faster release of XLXG compared to XXLG, and XXLG compared to XLLG. The rate of release of oligosaccharides is further determined by their backbone length. EndoIV prefers oligosaccharides with a backbone of four Glc residues (XXXG and XLXG) whereas endoI prefers two or three Glc residues (LG, XXG, and GFG). EndoV holds an intermediate position in this. The preference to release short oligosaccharides can be overruled by Fuc substitution (compare FG and LG by endoI).

The results on the release of oligosaccharides from POsxg cannot be discussed in further detail because the exact position of Ara and Gal residues is still unknown. EndoIV released predominantly XXGG-type of oligosaccharides whereas endoV also produced a GXXG-type [22]. In addition, substantial amounts of Glc and XXG-type of oligosaccharides accumulated in the POsxg digest obtained by endoV. The latter probably originate from further cleavage ('trimming') of the XXGG-type of oligosaccharides. From this it was concluded that endoV is better in 'trimming' G elements from xyloglucan oligosaccharides than endoIV [22]. Degradation patterns of POsxg by endoI were very similar to those obtained with endoV (data not shown).

What determines xyloglucanase activity?—The time course studies demonstrated a striking difference between endoI and endoIV. In particular, endoI preferred to release shorter oligosaccharides than endoIV. In order to explain this difference, the subsite models were combined with the data on the release of oligosaccharides. Two typical examples are shown in Fig. 7 in which the subsite maps of the endoglucanases (Fig. 4) have been simplified, and only positive and negative affinities are indicated. The figure shows clearly that during cleavage by endoIV, there is always an unbranched Glc residue which interacts with a subsite having a positive affinity. The distant subsite 'IV' is especially important in this respect (Fig. 7A). EndoI does not possess a subsite like 'IV',

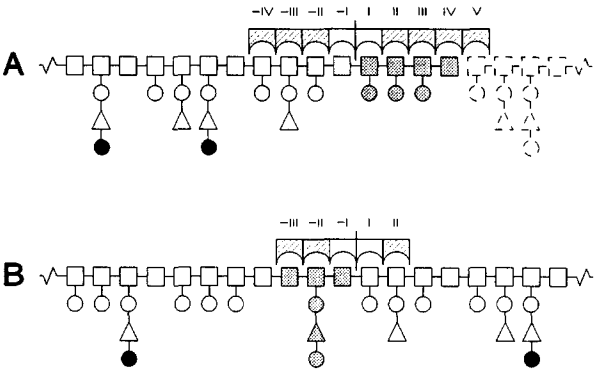

Fig. 7. Schematic representation of the interaction of APfxg and endoIV (A) or endoI (B). Assuming a random distribution of oligosaccharides over xyloglucan, two linkages should be cleaved to release the shaded oligosaccharide. Only in case of endoIV the second cleavage is shown. Dashed fragments indicate the part of the xyloglucan molecule which is removed after the first cleavage. Shaded substrate-binding sites bind to xyloglucan (positive subsite affinity), open subsites (negative subsite affinity) do not bind to xyloglucan. Other symbols as in Fig. 1.

and consequently prefers to release the shorter oligosaccharides such as GFG and LG. An explanation for this might be the favorable interaction of subsite '−III' with unbranched Glc residues (Fig. 7B).

The data discussed above suggest that binding of unbranched Glc residues plays an important role in the cleavage of a xyloglucan molecule. Since APfxg is built predominantly from XXXG, XLXG, XXFG, and XLFG [20,21], it may be expected that APfxg is degraded faster by endoIV than by endoI. This is in accordance with the data in Table 2. Because most xyloglucans are composed of XXXG-type of building units [17], it is tempting to generalize our results and state that 'xyloglucanase activity requires a large number of substrate-binding sites'. In agreement with our data, there is some evidence that an endoglucanase from *Pisum sativum*, showing xyloglucanase activity next to CMCase, has at least six substrate-binding sites [55,56].

This study does not provide an explanation for the high xyloglucanase of endoV. It cannot be excluded that endoV has a positive affinity in a distant subsite '−V' or in a subsite '−I', which would have a similar effect as subsite 'IV' of endoIV. In principle, the affinity of subsite '−V' could be determined with the BCFs of $G_4 \downarrow G$ and $G_5 \downarrow G$. However, the corresponding BCF values are zero, and a different set of BCFs, such as $G_4 \downarrow G_2$ and $G_5 \downarrow G_2$, would be needed. Because G_7 is insoluble in water, the latter BCF cannot be determined. The sum of the affinities

of subsite '−I' and 'I' is presumably negative (see before), but this does not rule out that subsite '−I' has a positive affinity. Unfortunately, no clearcut methodology is available to determine the subsite affinities of '−I' and 'I' individually. A third possibility to explain the xyloglucanase activity of endoV might be that endoV has a much broader substrate-binding groove than endoI, which may facilitate interaction with xyloglucan molecules.

The different mode of action of endoIV on POsxg, compared to endoI or endoV, can only partly be explained by our results. Interaction of subsites with unbranched Glc residues does not explain the cleavage patterns shown in Fig. 1. The absence of strong distant subsites like 'III' and 'IV' in endoI and endoV does explain why these enzymes are better than endoIV in 'trimming' oligosaccharides and releasing Glc. The action of endoglucanases on POsxg is probably determined by the position of Gal and Ara residues. The slower release of XXLG compared to XLXG from APxg suggests that Gal residues next to a G element in xyloglucan hamper the action of endoIV. This might explain why endoIV prefers the release of XXGG-type of oligosaccharides (Fig. 1). The release of XXG from POsxg by endoIV [22] is in agreement with this. Contrary to endoIV, the release of XLXG by endoI and endoV is generally slower than that of XXXG, which indicates that the action of endoI and endoV is more influenced by substitution than that of endoIV. This might explain the release of a mixture of XXGG-type and GXXG-type of oligosaccharides from POsxg by endoI or endoV, depending on whether or not the Xyl residue on the nonreducing side is substituted.

References

[1] C. Schou, G. Rasmussen, M.-B. Kaltoft, B. Henrissat, and M. Schülein, *Eur. J. Biochem.*, 217 (1993) 947–953.

[2] M. Penttilä, P. Lehtovaara, H. Nevalainen, R. Bhikhabhai, and J. Knowles, *Gene*, 45 (1986) 253–263.

[3] M. Saloheimo, P. Lehtovaara, M. Penttilä, T.T. Teeri, J. Ståhlberg, G. Johansson, G. Pettersson, M. Claeyssens, P. Tomme, and J.K.C. Knowles, *Gene*, 63 (1988) 11–21.

[4] M. Ward, S. Wu, J. Dauberman, G. Weiss, E. Larenas, B. Bower, M. Rey, K. Clarkson, and R. Bott, in P. Suominen and T. Reinikainen (Eds.), *Proceedings of the Second TRICEL Symposium on* Trichoderma reesei *Cellulases and other Hydrolases*, Foundation for Biotechnical and Industrial Fermentation Research, Helsinki, Finland, 1993, pp. 153–158.

[5] A. Saloheimo, B. Henrissat, A.-M. Hoffrén, O. Teleman, and M. Penttilä, *Mol. Microbiol.*, 13 (1994) 219–228.

[6] S. Shoemaker, V. Schweickart, M. Ladner, D. Gelfand, S. Kwok, K. Myambo, and M. Innis, *Bio/Technol.*, 1 (1983) 691–696.

[7] T.T. Teeri, P. Lehtovaara, S. Kauppinen, I. Salovuori, and J. Knowles, *Gene*, 51 (1987) 43–52.

[8] P. Tomme, H. van Tilbeurgh, G. Pettersson, J. van Damme, J. Vandekerckhove, J. Knowles, T. Teeri, and M. Claeyssens, *Eur. J. Biochem.*, 170 (1988) 575–581.

[9] N.R. Gilkes, B. Henrissat, D.G. Kilburn, R.C. Miller, and R.A.J. Warren, *Microbiol. Rev.*, 55 (1991) 303–315.

[10] M. Srisodsuk, T. Reinikainen, M. Penttilä, and T.T. Teeri, *J. Biol. Chem.*, 268 (1993) 20756–20761.

[11] B. Henrissat, M. Claeyssens, P. Tomme, L. Lemesle, and J.-P. Mornon, *Gene*, 81 (1989) 83–95.

[12] B. Henrissat and A. Bairoch, *Biochem. J.*, 293 (1993) 781–788.

[13] S. Woodcock, J.-P. Mornon, and B. Henrissat, *Protein Eng.*, 5 (1992) 629–635.

[14] G. Beldman, A.G.J. Voragen, F.M. Rombouts, M.F. Searle-van Leeuwen, and W. Pilnik, *Biotechnol. Bioeng.*, 31 (1988) 160–167.

[15] P. Biely, M. Vršanská, and M. Claeyssens, *Eur. J. Biochem.*, 200 (1991) 157–163.

[16] M. Claeyssens and B. Henrissat, *Protein Sci.*, 1 (1992) 1293–1297.

[17] T. Hayashi, *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, 40 (1989) 139–168.

[18] B.S. Montenecourt, *Trends Biotechnol.*, 1 (1983) 156–161.

[19] J.-P. Vincken, G. Beldman, and A.G.J. Voragen, *Plant Physiol.*, 104 (1994) 99–107.

[20] J.-P. Vincken, G. Beldman, W.M.A. Niessen, and A.G.J. Voragen, *Carbohydr. Polym.*, 29 (1996) 75–85.

[21] B.A. Spronk, G.J. Rademaker, J. Haverkamp, J.E. Thomas-Oates, J.-P. Vincken, A.G.J. Voragen, J.P. Kamerling, and J.F.G. Vliegenthart, to be submitted.

[22] J.-P. Vincken, A.J.M. Wijsman, G. Beldman, W.M.A. Niessen, and A.G.J. Voragen, *Carbohydr. Res.*, 288 (1996) 219–232.

[23] S.D. Baird, M.A. Hefford, D.A. Johnson, W.L. Sung, M. Yaguchi, and V.L. Seligy, *Biochem. Biophys. Res. Commun.*, 169 (1990) 1035–1039.

[24] B. Svensson and M. Søgaard, *J. Biotechnol.*, 29 (1993) 1–37.

[25] L. Chen, G.B. Fincher, and P.B. Høj, *J. Biol. Chem.*, 268 (1993) 13318–13326.

[26] M.L. Sinnott, *Chem. Rev.*, 90 (1990) 1171–1202.

[27] J.C. Gebler, N.R. Gilkes, M. Claeyssens, D.B. Wilson, P. Béguin, W.W. Wakarchuk, D.G. Kilburn, R.C. Miller, R.A.J. Warren, and S.G. Withers, *J. Biol. Chem.*, 267 (1992) 12559–12561.

[28] C. Divne, J. Ståhlberg, T. Reinikainen, L. Ruohonen, G. Pettersson, J.K.C. Knowles, T.T. Teeri, and T.A. Jones, *Science*, 265 (1994) 524–528.

[29] A. Meinke, H.G. Damude, P. Tomme, E. Kwan, D.G.

Kilburn, R.C. Miller, R.A.J. Warren, and N.R. Gilkes, *J. Biol. Chem.*, 270 (1995) 4383–4386.
[30] P.B. Høj, E.B. Rodriguez, R.V. Stick, and B.A. Stone, *J. Biol. Chem.*, 264 (1989) 4939–4947.
[31] G. Beldman, M.F. Searle-van Leeuwen, F.M. Rombouts, and A.G.J. Voragen, *Eur. J. Biochem.*, 146 (1985) 301–308.
[32] T. Suganuma, R. Matsuno, M. Ohnishi, and K. Hiromi, *J. Biochem.*, 84 (1978) 293–316.
[33] M. Voloch, M.R. Ladisch, M. Cantarella, and G.T. Tsao, *Biotechnol. Bioeng.*, 26 (1984) 557–559.
[34] N. Somogyi, *J. Biol. Chem.*, 195 (1952) 19–23.
[35] M.T. Tollier and J.P. Robin, *Ann. Technol. Agric.*, 28 (1979) 1–15.
[36] G.A. de Ruiter, H.A. Schols, A.G.J. Voragen, and F.M. Rombouts, *Anal. Biochem.*, 207 (1992) 176–185.
[37] J.J. Sedmak and S.E. Grossberg, *Anal. Biochem.*, 79 (1977) 544–552.
[38] S.C. Fry, W.S. York, P. Albersheim, A. Darvill, T. Hayashi, J.-P. Joseleau, Y. Kato, E.P. Lorences, G.A. Maclachlan, M. McNeil, A.J. Mort, J.S.G. Reid, H.U. Seitz, R.R. Selvendran, A.G.J. Voragen, and A.R. White, *Physiol. Plant.*, 89 (1993) 1–3.
[39] M. Claeyssens, H. van Tilbeurgh, J.P. Kamerling, J. Berg, M. Vršanská, and P. Biely, *Biochem. J.*, 270 (1990) 251–256.
[40] M. Vršanská and P. Biely, *Carbohydr. Res.*, 227 (1992) 19–27.
[41] R. Macarron, C. Acebal, M.P. Castillon, J.M. Dominguez, I. de la Mata, G. Pettersson, P. Tomme, and M. Claeyssens, *Biochem. J.*, 289 (1993) 867–873.
[42] K.M. Bhat, A.J. Hay, M. Claeyssens, and T.M. Wood, *Biochem. J.*, 266 (1990) 371–378.
[43] M.R. Bray and A.J. Clarke, *Eur. J. Biochem.*, 204 (1992) 191–196.
[44] E.K. Gum and R.D. Brown, *Anal. Biochem.*, 82 (1977) 372–375.
[45] B. Björkqvist, *J. Chromatogr.*, 218 (1981) 65–71.
[46] M. Claeyssens, H. van Tilbeurgh, P. Tomme, T.M. Wood, and S.I. McRae, *Biochem. J.*, 261 (1989) 819–825.
[47] G. Beldman, A.G.J. Voragen, F.M. Rombouts, M.F. Searle-van Leeuwen, and W. Pilnik, *Biotechnol. Bioeng.*, 30 (1987) 251–257.
[48] J.-P. Vincken, Enzymic Modification of Cellulose-Xyloglucan Networks, Ph.D. thesis, Agricultural University Wageningen, The Netherlands, 1996, p. 143.
[49] K. Ogawa, T. Hayashi, and K. Okamura, *Int. J. Biol. Macromol.*, 12 (1990) 218–222.
[50] E.D.T. Atkins, in J. Visser, G. Beldman, M.A. Kusters-van Someren, and A.G.J. Voragen (Eds.), *Xylans and xylanases*, Elsevier, Amsterdam, 1992, pp 39–50.
[51] A. Törrönen, C.P. Kubicek, and B. Henrissat, *FEBS Lett.*, 321 (1993) 135–139.
[52] H.T. Saarilahti, B. Henrissat, and E.T. Palva, *Gene*, 90 (1990) 9–14.
[53] T. Ooi, A. Shinmyo, H. Okada, S. Hara, T. Ikenaka, S. Murao, and M. Arai, *Curr. Genet.*, 18 (1990) 217–222.
[54] H. Dalbøge, L.N. Andersen, L.V. Kofod, M.S. Kauppinen, and S. Christgau, World Pat. Wo 94/14953 (1994).
[55] G. Maclachlan and Y.-S. Wong, *Adv. Chem. Ser.*, 181 (1979) 347–360.
[56] T. Hayashi, Y.-S. Wong, and G. Maclachlan, *Plant Physiol.*, 75 (1984) 605–610.