Protein Engineering vol.9 no.8 pp.691–699, 1996

The active site of *Trichoderma reesei* cellobiohydrolase II: the role of tyrosine 169

Anu Koivula[1], Tapani Reinikainen[2], Laura Ruohonen, Anne Valkeajärvi[3], Marc Claeyssens[4], Olle Teleman, Gerard J.Kleywegt[5], Michael Szardenings[5], Juha Rouvinen[6], T.Alwyn Jones[1,5] and Tuula T. Teeri

VTT Biotechnology and Food Research, PO Box 1500, FIN-02044 VTT, Espoo, Finland, [4]Laboratory of Biochemistry, University of Ghent, K.L. Ledeganckstraat 35, B-9000 Ghent, Belgium and [5]Department of Molecular Biology, BMC, PO Box 590, S-75124 Uppsala, Sweden

[1]Authors to whom correspondence should be addressed

[2]Present address: Cultor Ltd, Technology Center, FIN-02460 Kantvik, Finland

[3]Present address: Finnish Red Cross, Kivihaantie 7, FIN-00310 Helsinki, Finland

[6]Present address: Department of Chemistry, University of Joensuu, PO Box 111, FIN-80101 Joensuu, Finland

***Trichoderma reesei* cellobiohydrolase II (CBHII) is an exoglucanase cleaving primarily cellobiose units from the non-reducing end of cellulose chains. The β-1,4 glycosidic bond is cleaved by acid catalysis with an aspartic acid, D221, as the likely proton donor, and another aspartate, D175, probably ensuring its protonation and stabilizing charged reaction intermediates. The catalytic base has not yet been identified experimentally. The refined crystal structure of CBHII also shows a tyrosine residue, Y169, located close enough to the scissile bond to be involved in catalysis. The role of this residue has been studied by introducing a mutation Y169F, and analysing the kinetic and binding behaviour of the mutated CBHII. The crystal structure of the mutated enzyme was determined to 2.0 Å resolution showing no changes when compared with the structure of native CBHII. However, the association constants of the mutant enzyme for cellobiose and cellotriose are increased threefold and for 4-methylumbelliferyl cellobioside over 50-fold. The catalytic constants towards cellotriose and cellotetraose are four times lower for the mutant. These data suggest that Y169, on interacting with a glucose ring entering the second subsite in a narrow tunnel, helps to distort the glucose ring into a more reactive conformation. In addition, a change in the pH activity profile was observed. This indicates that Y169 may have a second role in the catalysis, namely to affect the protonation state of the active site carboxylates, D175 and D221.**

Keywords: catalytic domain/cellobiohydrolase/site-directed mutagenesis/*Trichoderma reesei*

Introduction

The filamentous fungus *Trichoderma reesei* produces a complete set of cellulolytic enzymes needed for efficient solubilization of native cellulose. Its two cellobiohydrolases, CBHI and CBHII (EC 3.2.1.91) are apparently the key enzymes in the breakdown of crystalline cellulose (Irwin *et al.*, 1993; Suominen *et al.*, 1993; Divne *et al.*, 1994). Both CBHI and CBHII are exoglucanases releasing predominantly cellobiose from the ends of the polymeric glucose chains (Wood and Garcia-Campayo, 1990; Divne *et al.*, 1994). CBHI and CBHII, as well as many other fungal and bacterial cellulases, are composed of two domains: a large catalytic domain and a small cellulose-binding domain (CBD), which are joined by a linker peptide (Gilkes *et al.*, 1991; Teeri *et al.*, 1992).

Extensive sequence analyses have been used to classify all known cellulase catalytic domains into different families (Henrissat and Bairoch, 1993), each with a conserved protein fold and stereochemical mechanism of hydrolysis. CBHI and CBHII are to our knowledge the only true exocellulases whose three-dimensional structure has been solved (Rouvinen *et al.*, 1990; Divne *et al.*, 1994). Their catalytic domains belong to different cellulase families and have completely different folds. Despite this they share a common active site architecture with long surface loops generating a tunnel for substrate binding and catalysis.

Cellobiohydrolases, like all glycosidases, catalyse the hydrolysis of the glycosidic bond utilizing a general acid catalysis mechanism. The enzymatic breakdown of the glycosidic bond is a stereoselective process, where the configuration of the anomeric centre (C_1 carbon) is either inverted or retained (Sinnott, 1990). *T.reesei* CBHI is a retaining enzyme, like for example hen egg-white lysozyme and β-galactosidase, whereas CBHII is an inverting glycosylase (Knowles *et al.*, 1988; Clayssens *et al.*, 1990) together with sialidase, β-amylase and glucoamylase. Both inverting and retaining enzymes are thought to operate through transition states with substantial oxocarbonium ion character. In some hydrolases the distortion of the sugar ring towards the transition state is shown to occur already when the substrate binds to the enzyme (McCarter and Withers, 1994).

We are studying the amino acid residues that may be important in the catalytic action and substrate binding of CBHII. The crystallographic studies have shown that the active site tunnel of CBHII is approximately 20 Å long (Rouvinen *et al.*, 1990). At the centre of the tunnel, we identified a constellation of interacting polar residues as the likely active site. In this paper, we describe the contribution to the catalytic activity of a tyrosine residue, Y169, that is strictly conserved in the CBHII family (family 6).

Materials and methods

Strains and vectors

Escherichia coli strain DH5α (Promega) was used as the bacterial cloning host. The *T.reesei* expression strain for the Y169F mutant was a derivative of VTT-D-79125 (Bailey and Nevalainen, 1981) where *cbh2* and *egl2* genes have been replaced by *trp* and *amdS* genes, respectively (A.Koivula, A.Lappalainen, S.Virtanen, A.L.Mäntylä, P.Suominen and T.T.Teeri, submitted). CBHII wt protein was purified from the *T.reesei* Rut C-30 strain. Plasmid pTTc9 (Teeri *et al.*, 1987) contains the *cbh2* cDNA as an *Eco*RI–*Bam*HI restriction

fragment in pUC8. The fungal expression vector pUJ10 is a derivative of pBR322 containing *lacZ* gene linked to the *cbh1* promoter and terminator sequences (U.Airaksinen and M.Penttilä, unpublished). A phleomycin selection plasmid pAN8-1 (Mattern *et al.*, 1987) was used in selection of the *Trichoderma* transformants.

Construction of the mutant clone

The *cbh2* cDNA was first cloned as an *Eco*RI–*Bam*HI fragment from the plasmid pTTc9 into vector pSP73 (Promega) where Y169F (TAT→TTT) mutation was introduced by a polymerase chain reaction (PCR) (Ho *et al.*, 1989). The nucleotide sequence of the whole fragment subjected to PCR was confirmed. The mutated *cbh2* cDNA was finally cloned under the *cbh1* promoter of the fungal expression vector by changing the *Kpn*I–*Sma*I restriction fragment (containing the *lacZ* gene) of pUJ10 into the mutated *Eco*RI–*Bam*HI fragment.

Transformation and screening of the fungal transformants

Fungal transformation was carried out according to a standard method using co-transformation with the selection plasmid pAN8-1 (Penttilä *et al.*, 1987; Nyyssönen *et al.*, 1993). Screening of the best producing transformants was done as described earlier (Srisodsuk *et al.*, 1993).

Protein purification

The mutant protein was produced as described by Srisodsuk *et al.* (1993). Culture supernatant was separated from mycelia by centrifugation and further clarified by filtration. Sodium azide, phenylmethylsulfonyl fluoride and EDTA were added to final concentrations of 0.02%, 30 μM and 1 mM, respectively, and the solution was concentrated with Pellicon Laboratory Cell System using a PTGC10 membrane (Millipore, Bedford, MA). After desalting with Biogel P-6 (Bio-Rad, Cambridge, MA), the protein solution was run through a DEAE-Sepharose fast flow column (Pharmacia, Uppsala, Sweden) equilibrated in 50 mM sodium acetate buffer (pH 5.6). The flow-through fraction containing CBHII was further purified by thiocellobioside-based affinity chromatography (Tomme *et al.*, 1988a) and finally by immunoaffinity chromatography (A.Koivula, A.Lappalainen, S.Virtanen, A.L.Mäntylä, P.Suominen and T.T.Teeri, submitted). CBHII Y169F catalytic domain for structure determination was produced by papain (Sigma) digestion (Tomme *et al.*, 1988b), then it was absorbed in 20 mM bis-Tris (pH 7.0) in a MonoQ column (Pharmacia) and eluted with linear salt gradient (0–0.1 M NaCl) using an FPLC System. The purity of the mutant preparates was checked and verified by SDS–PAGE (Laemmli, 1970) and Western blotting. The presence of contaminating cellulolytic activities was further ruled out by measuring the activities against small chromogenic substrates. CBHII wt protein was purified as described by Reinikainen *et al.* (1995) and the purity checked as described above.

Western blotting

Proteins were separated on 10% SDS–PAGE gel and transferred electrophoretically onto a nitrocellulose membrane (Towbin *et al.*, 1979). CBHII-containing bands were identified with monoclonal antibody raised against CBHII catalytic domain [CII-8 (Aho *et al.*, 1991)] and thereafter reacted with commercial alkaline phosphatase-labelled goat anti-mouse IgG (Sigma) and finally detected by staining with a Protoblot kit (Promega).

Contaminating cellulase activities

Since neither CBHII wt or the mutant Y169F can cleave the heterosidic linkage of small chromogenic oligosaccharides (van Tilbeurgh *et al.*, 1988, and our results), these substrates were used to check the purity of both wild-type and mutated CBHII enzyme. The contaminating activities were measured as described by Reinikainen *et al.* (1995).

Protein concentration

Absorbance values at 280 nm were used to measure CBHII concentration using $\varepsilon = 80\ 500\ M^{-1}\ cm^{-1}$ for the CBHII wt protein (Tomme, 1991) and for CBH II Y169F mutant a molar absorptivity value for tyrosine ($1340\ M^{-1}\ cm^{-1}$) (Wetlaufer, 1962) was subtracted, resulting in $\varepsilon = 79\ 160\ M^{-1}\ cm^{-1}$.

Enzyme kinetics

Turnover numbers and Michaelis constants for the hydrolysis of cellotriose (Glc_3) and cellotetraose (Glc_4) (Merck) in 10 mM sodium acetate buffer (pH 5.0) at 27°C were determined by HPLC (Waters Millipore, Milford, MA) equipped with a refractive index detector as described earlier (Teleman *et al.*, 1995). The column used for separation was either HC-40 (Hamilton) or Aminex HPX-42A (Bio-Rad). The substrate concentration ranges in the case of CBHII wt were 6–410 μM for Glc_3 and 1.7–340 μM for Glc_4 and in the case of the Y169F mutant 0.08–2.94 mM for Glc_3 and 3.0–201 μM for Glc_4. The pH dependence was determined at 27°C with cellotetraose. The cellotetraose concentration in the reaction mixture was 160 μM (with CBHII wt) or 201 μM (with Y169F mutant) and the final enzyme concentration was 0.04–0.10 μM. The following buffer systems were used: glycine–HCl (pH range 1.4–3.8); sodium acetate (pH range 4.0–5.6); sodium phosphate (pH range 6.0–8.0); glycine–NaOH (pH range 9.0–10.6), where the concentrations of the buffer solutions were 10 mM. In all measurements samples were taken at 8–10 different time points in order to produce reliable values for the initial rates. Kinetic constants were obtained from the initial velocities of the reaction curves by a non-linear regression data analysis program [Enzfit (Marquart, 1963)].

Ligand binding studies

Binding of 4-methylumbelliferyl-glucoside and -cellobioside [$MeUmb(Glc)_1$ and $MeUmb(Glc)_2$] in 50 mM sodium acetate buffer (pH 5.0) at 16°C was studied by direct fluorescence quenching titrations as described by van Tilbeurgh *et al.* (1985). Fluorescence was measured with an Aminco SPF-500 ratio spectrofluorimeter equipped with a thermostated cuvette holder. Excitation was invariably at 318 nm (spectral bandwith 2 nm) and the emission was measured at 360 nm with a bandwidth of 15, 20 or 40 nm. Glucose, cellobiose and cellotriose (Merck) were used as competitive ligands in the displacement titrations. The dissociation constants for MeUmb-oligosaccharides were obtained by non-linear least-squares fitting to the fluorescence data following the mathematical treatment described by De Boeck *et al.* (1983). The fitting was implemented as a Macintosh program, which also allowed manual fitting.

X-ray crystallography of Y169F mutant protein

The Y169F mutant crystallized in the same space group ($P2_1$, $a = 49.1$ Å, $b = 75.8$ Å, $c = 92.9$ Å and $\beta = 103.2°$) as the wild-type enzyme (Rouvinen *et al.*, 1990). Diffraction data were collected on an SDMS multiwire area detector at room temperature to yield 202 105 measurements which were reduced to 46 849 independent reflections with an R_{merge} of 6.8%. The structure was refined starting from a model of the native enzyme that had been refined at 1.8 Å with strict non-crystallographic symmetry (NCS) constraints and grouped

temperature factors. Residue 169 was replaced by an alanine residue, water molecules were removed from the model, and all temperature factors were limited to lie in the range 10–40 Å^2. The dataset (40 860 reflections between 8.0 and 2.0 Å) was partitioned into a work set (used for refinement) and a test set for R_{free} calculations (Brünger, 1992). The test set contained 10% of the reflections, which were not selected at random, but rather in 25 thin resolution shells in order to reduce the bias of R_{free} due to correlations between 'NCS-related' reflections (Kleywegt and Jones, 1995). During refinement, all observed data between 8.0 and 2.0 Å were used. Strict NCS constraints and grouped temperature factors were maintained throughout the refinement process, and the Engh and Huber force field (Engh and Huber, 1991) was used at all times. Force constants in the parameter sets of the carbohydrates were set to values which are of the same order of magnitude as those in the protein force field in order to prevent excessive distortion of these moieties. All refinement calculations were carried out with X-PLOR (Brünger *et al.*, 1987). Maps were calculated with programs in the CCP4 package (Collaborative Computational Project, 1994), and averaged with programs in the RAVE package (Kleywegt and Jones, 1994). All rebuilding and graphics operations were carried out with O (Jones *et al.*, 1991). The quality of each model was checked using standard tools [PROCHECK (Laskowski *et al.*, 1994), X-PLOR (Brünger *et al.*, 1987), O (Jones *et al.*, 1991) and OOPS (Kleywegt and Jones, 1996)].

The initial model had an R-factor of 0.282 (R_{free} 0.285). After one round of simulated annealing (SA), Powell minimization and grouped B-factor refinement, the R-factor dropped to 0.253 and R_{free} to 0.279. Density for the ring of F169 was clearly visible in electron density maps (Figure 1). The correct side chain was inserted at this position, the model was rebuilt where necessary and water molecules were added. This resulting model was again subjected to SA, Powell minimization and grouped temperature-factor refinement. After the next rebuild, the model was subjected to 50 cycles of conjugate gradient minimization and 25 cycles of grouped temperature-factor refinement, which yielded an R-factor of 0.209 (R_{free} 0.232). Subsequently, this refinement was repeated using all observed reflections to yield a final R-factor of 0.210. Ten water molecules with temperature factors exceeding 60 Å^2 were removed from the model without affecting the R-factor.

The final model (with strict twofold NCS) comprises 363 residues (85–447) containing 3047 non-hydrogen atoms which have an average temperature factor of 15.8 Å^2 (11.0 Å^2 for 1452 protein main-chain atoms, 16.9 Å^2 for 1294 side-chain atoms, 49.2 Å^2 for 105 carbohydrate atoms and 26.1 Å^2 for 196 water molecules). The stereochemistry has been tightly restrained to ideal values (Engh and Huber, 1991), yielding root-mean-square deviations (r.m.s.d.s) from ideality of 0.005 Å for bond lengths, 1.3° for bond angles and 1.1° for improper torsion angles. All residues lie in the most favoured (86.7%) and additional allowed (13.3%) regions of the Ramachandran plot, as defined by PROCHECK (Laskowski *et al.*, 1994). The overall G-factor (Laskowski *et al.*, 1994) of the final model is +0.37. Ten residues have unusual peptide orientations (Jones *et al.*, 1991), but all outliers have excellent density for the peptide oxygen atom. Fifteen residues have non-rotamer side-chain conformations (Jones *et al.*, 1991). The average real-space fit value (Jones *et al.*, 1991) for the final model in the averaged map is 0.88 (σ 0.09). The final averaged map is of excellent quality. There are only two turns which have worse than average density for their main-chain atoms: 156–160 and 406–410. However, in both cases the averaged and unaveraged density is of comparable quality, which indicates that the relatively poor density and high temperature factors in these regions are more likely due to dynamic disorder than to NCS breakdown.

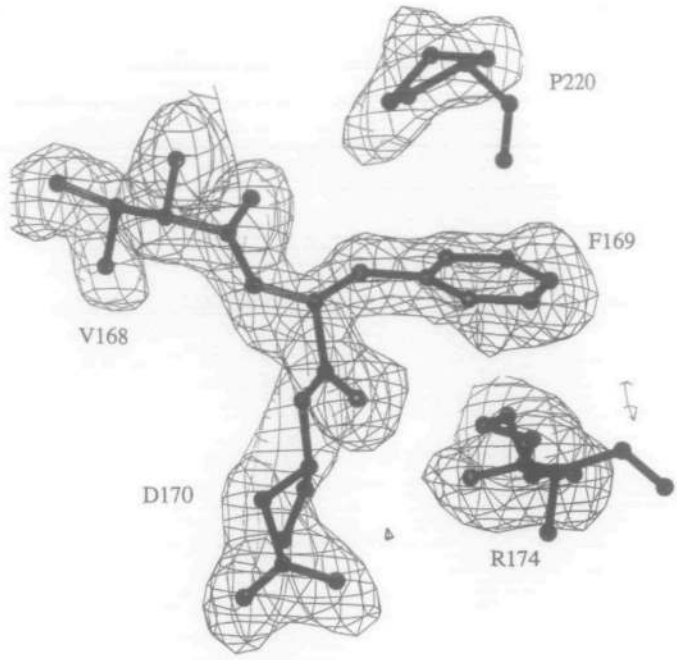


Fig. 1. Electron density map of the CBHII mutant Y169F, centred on phenylalanine 169, with the final model overlaid. The density was calculated at a stage in the refinement when residue 169 was still modelled as an alanine residue in order to obtain unbiased density for the side-chain atoms of this residue.

Atomic coordinates and structure factors for the Y169F mutant have been deposited with the PDB in Brookhaven code 1CB2.

Results

Y169F mutant structure

The final model of the mutant enzyme is very similar to that of the native structure. The r.m.s.d. between the two models

Table I. Kinetic parameters for the CBHII wt and mutant Y169F enzyme measured in 10 mM sodium acetate buffer (pH5.0) at 27°C

Substrate	CBHII wt			Y169F		
	k_{cat} (min^{-1})	K_m (μM)	k_{cat}/K_m (min^{-1} μM^{-1})	k_{cat} (min^{-1})	K_m (μM)	k_{cat}/K_m (min^{-1} μM^{-1})
Glc_3	3.7 ± 0.1 [a]	22 ± 2	0.17 ± 0.02	0.9 ± 0.1	900 ± 100	0.001 ± 0.0002
Glc_4	223 ± 8	1.8 ± 0.3	120 ± 20	57 ± 2	1.2 ± 0.3	48 ± 14

[a]Kinetic constants were calculated by a non-linear regression data analysis program (Enzfit), which also gives the standard error.

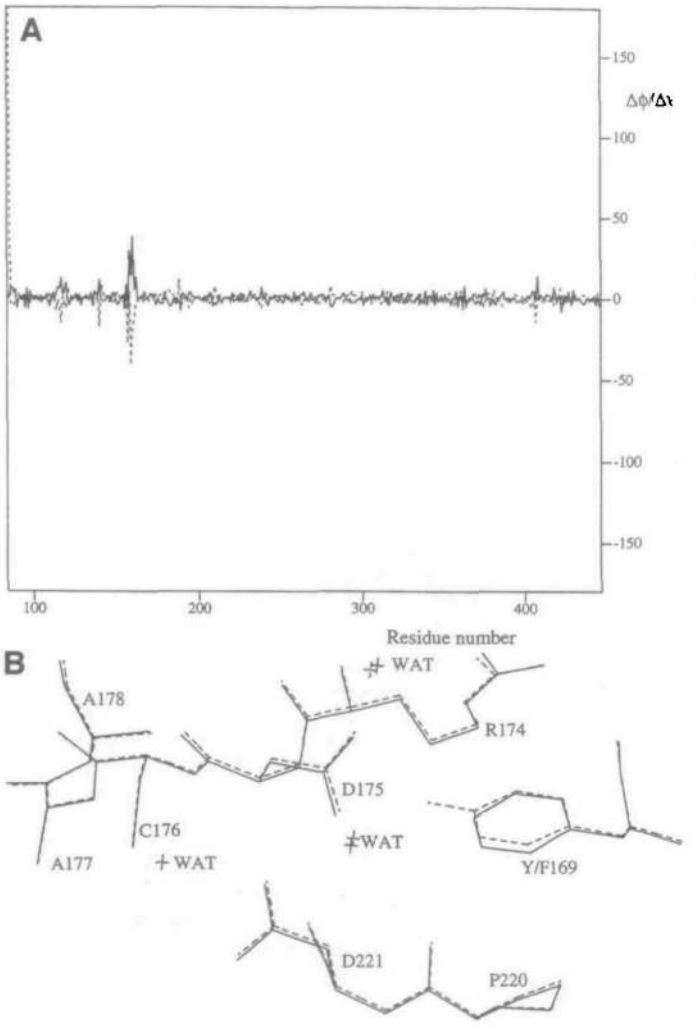


Fig. 2. (**A**) Plot of the differences between the main-chain torsion angles of Y169F and wt CBHII as a function residue number. The solid curve shows the differences between corresponding Φ angles, the dashed curve those of corresponding ψ angles. (**B**) Superimposition of the active sites of Y169F (solid lines) and wild-type CBHII (dashed lines).

is only 0.11 Å for 363 Cα atoms (r.m.s. ΔB 6.3 Å^2), and 0.15 Å for 1784 main-chain and Cβ atoms in the protein (r.m.s. ΔB 7.1 Å^2). Out of 196 water molecules in the Y169F structure, 181 are also found in the native structure (maximum distance of these, 1.1 Å; average distance, 0.25 Å). Figure 2A shows the $\Delta\Phi/\Delta\psi$ plot of the 1.8 Å native model and the 2.0 Å Y169F model. Figure 2B overlays the native and Y169F mutant structures in the active site tunnel and clearly shows that the mutation causes essentially no change in the structure.

Kinetic properties

To characterize the hydrolysis kinetics of the wt and mutated CBHII, specific soluble substrates were used instead of cellulose. The catalytic constants for cellotriose and cellotetraose degradation at pH 5.0 at 27°C are shown in Table I. For both CBHII wt and Y169F mutant enzyme the values for k_{cat} and K_m were calculated from the initial rate measurements at eight different substrate concentrations. As seen from Table I, the turnover numbers (k_{cat}) for the Y169F mutant are approximately 25% of those for the wild-type enzyme at pH 5.0. The K_m for cellotriose has increased over 40-fold but for cellotetraose it has slightly decreased. The specificity constant (k_{cat}/K_m) for

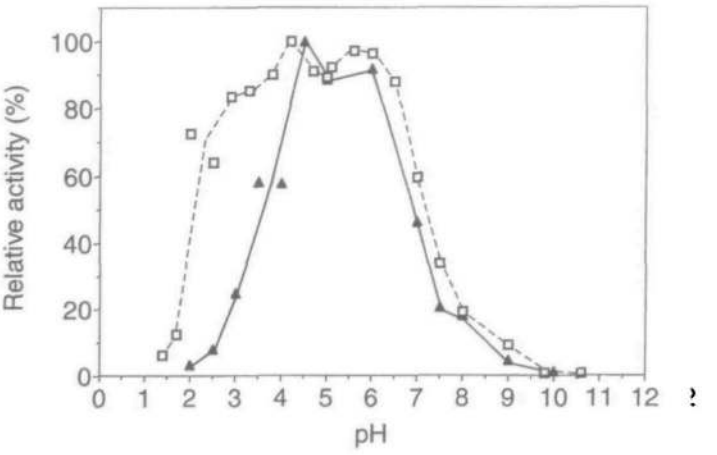


Fig. 3. pH-activity profiles of CBHII wt (□) and Y169F mutant enzyme (▲) represented as relative activities. The activities were measured at 27°C at constant cellotetraose concentration, which was approximated as saturating concentration. The absolute activity value of Y169F mutant at pH 5.0 corresponds to approximately 25% of the CBHII wt activity.

the Y169F mutant has decreased 170-fold on cellotriose and 2.5-fold on cellotetraose.

The pH dependence of the reaction rate for cellotetraose degradation was also determined. Over the whole pH range studied the activity of the mutant enzyme was lower; in the pH range 4.5–10.0 the activity was invariably 20–25% of the CBHII wt activity but on the acidic side the activity decrease was relatively larger. The relative pH-activity profiles are presented in Figure 3.

Binding data

The fluorescence of 4-methylumbelliferyl β-D-glucosides [MeUmb(Glc)$_1$ and MeUmb(Glc)$_2$] was quenched upon binding to CBHII wt or to the Y169F mutant without any shift in the maximum emission wavelength (360 nm) and the interactions of the fluorescent as well as non-fluorescent (competing) ligands could be quantified by titrimetric methods. The method used to treat the data from either direct or displacement titrations is an improvement of that described by De Boeck *et al.* (1983). Here the fitting was done directly to the fluorescence data in a non-linear least-squares fashion and the mathematical treatment was more rigorous in that also the maximal fluorescence intensity change (ΔF_{max}) was fitted. In most cases these values approximated 100%, i.e. the fluorescence is completely quenched when the chromophore is bound by the protein. The precision of the parameter fits was assessed by varying the parameter values around their optima, and found to be in the order of a few per cent. The spread between duplicate experiments was somewhat larger and we thus estimate the total uncertainty to be 10–15%. This error analysis is more rigorous than the earlier published uncertainties (De Boeck *et al.*, 1983; van Tilbeurgh *et al.*, 1989) which were derived from the precision of the numerical fit only.

The association constants of MeUmb-derivatized oligosaccharides (from direct titration) and of free oligosaccharides (from displacement titration) for CBHII wt and Y169F are shown in Table II. Cellotetraose or longer oligosaccharides could not be used in the displacement titrations since they are good substrates for both CBHII wt and the Y169F mutant enzyme. Consistent with earlier studies of wt CBHII (van Tilbeurgh *et al.*, 1985, 1989), we obtained higher association constants for MeUmb-glycosides than for the corresponding unmodified

Table II. Association and dissociation constants and $\Delta G°$ values for the binding of carbohydrates to CBHII wt and mutant Y169F in 50 Mm NaAc buffer (pH 5.0) at 16°C

Ligand	CBHII wt			Y169F		
	K_{ass} (M^{-1})	K_d (μM)	$\Delta G°$ (kJ/mol)	K_{ass} (M^{-1})	K_d (μM)	$\Delta G°$ (kJ/mol)
$MeUmb(Glc)_1$	0.6×10^5[a]	18	−26.4	0.6×10^5	18	−26.4
$MeUmb(Glc)_2$	4×10^5	2.6	−31.0	$\geq200\times10^5$	≤0.05	≤-40.4
Glc_1	3×10^1	31000	−8.2	4×10^1	24000	−8.9
Glc_2	0.7×10^3	1400	−15.7	2×10^3	450	−18.3
Glc_3	6×10^4	15	−26.4	20×10^4	5.4	−29.3

[a]The uncertainty in K_{ass} and K_d values is estimated to be 10–15% on the basis of four repeated experiments.

oligosaccharides (Table II). This also applies to the Y169F mutant. Apparently the presence of the aromatic MeUmb group enhances the affinity of oligosaccharides irrespective of the actual binding site.

CBHII wt and Y169F mutant enzymes show almost identical binding affinity towards both glucose and $MeUmb(Glc)_1$ (Table II). Earlier studies with CBHII wt have shown that, depending on the temperature, the maximum fluorescence quenching at total protein saturation (ΔF_{max}) varies between 80 and 100% with $MeUmb(Glc)_1$ (van Tilbeurgh *et al.*, 1985, 1989). In the present study, the ΔF_{max} values of both the wt and mutant CBHII deviated from 100% (data not shown). Furthermore, glucose could be used to displace $MeUmb(Glc)_1$ from both protein complexes. These data suggest that the binding mode of $MeUmb(Glc)_1$ is conserved in the mutant.

On the other hand, the mutation Y169F was found to increase the binding affinity of CBHII for $MeUmb(Glc)_2$ more than 50-fold (Table II) and the affinity for cellobiose and cellotriose was increased threefold. The changed binding behaviour is also evidenced by a titration experiment showing that D-glucose displaces $MeUmb(Glc)_2$ from its complex with the mutant enzyme (data not shown). This is in sharp contrast to the CBHII wt enzyme, where glucose enhances the affinity of $MeUmb(Glc)_2$ in the CBHII complex (van Tilbeurgh *et al.*, 1985, 1989).

Discussion

The overall fold of CBHII (Rouvinen *et al.*, 1990) is a variant of the classical TIM (triose phosphate isomerase) barrel and contains seven instead of eight parallel β-strands. At the C-terminal end of the barrel, two loops combine to form an enclosed tunnel, approximately 20 Å long (Figure 4e and f). The sides of the tunnel are lined with side chains forming a complex network of salt links and hydrogen bonds. In the uncomplexed structures, the tunnel is filled with water molecules that take part in some of these interactions. Four glucosyl binding sites (A–D) have been identified within the tunnel where the non-reducing end of the cellulose chain binds to the subsite A. Although each site is somewhat flattened, they have different cross-sections (Figure 4a–d). In the subsites A, C and D tryptophan side chains W135, W367 and W269, respectively, make significant contributions to the formation of the sugar binding sites (Figure 4a, c and d). In site A, in particular, this results in a narrower tunnel cross-section. Four carboxylates (D221, D175, D263 and D401) conserved in the CBHII family were identified at the centre of the tunnel and recognized as having potential roles in catalysis occurring between the subsites B and C (Rouvinen *et al.*, 1990). Mutagenesis studies of D221 and D175 have shown that D221 is the likely proton donor in the catalytic event while the neighbouring D175 is charged ensuring the protonation of D221 (Ruohonen *et al.*, 1993; our unpublished data). D175 is also interacting with the hydroxyl of Y169 and the guanidino group of R174 (Figure 4g). Furthermore, the guanidino group of R174 is stacked parallel to the phenolic ring of Y169 (Figure 4g). Since CBHII acts with inversion of the anomeric configuration, a single-displacement mechanism, also involving a base to assist the nucleophilic attack of the water, has been postulated (Koshland, 1953). D401 is correctly oriented with respect to the glycosidic linkage to act as the base. The shortest O_δ–O_δ distance between D221 and D401 is 11.2 Å and the C_α–C_α distance between D221 and D401 is 13.2 Å. However, D401 is salt linked to two nearby residues, R353 and K395, and these interactions would be expected to reduce its capacity to extract a proton from the attacking water molecule.

More recently, kinetic studies made by Sinnott and co-workers (Sinnott, 1990; Konstantinidis *et al.*, 1993) have cast doubt on a classical single-diplacement mechanism for CBHII, suggesting that a base may not be necessary. Experimental support is obtained from the observation that the hydrolysis of α-cellobiosyl fluoride does not seem to proceed with the resynthesis–hydrolysis mechanism suggested by Kasumi *et al.* (1987). Kinetic studies by Konstantinidis *et al.* (1993) and the high association constant of cellobionolactonoxime (see Figure 5) suggest that the mechanism of CBHII involves oxocarbonium ion-like transition states (van Tilbeurgh *et al.*, 1986, 1989). In this kind of of mechanism D175 is ideally placed to play a second role in stabilizing these positively charged reaction intermediates at site B.

Site B in CBHII has no tryptophan and is clearly different from the other three binding sites. It has a protrusion (Figure 4b) which may allow alternative sugar conformations. There are three residues, K395, D401 and Y169, that can contribute to the forming of site B and that are strictly conserved in the CBHII family 6. They all are in hydrogen bond distance from the sugar ring in site B. In the wild-type CBHII–$MeUmb(Glc)_2$ complex there is excellent density associated with the chromophore and the A site glucosyl while the B site density is poorer. A persistent peak in the maps calculated from phases based on the protein model alone could be interpreted as the methylhydroxyl group of the glucosyl unit in site B. This would require the ring to adopt a distorted configuration in site B. Refinement at 2.5 Å resolution without restraints on the conformation of the glucosyl moiety does not, however, give us conclusive evidence for the distorted configuration.

We envisaged that the role of Y169 was to assist in the preservation of the D175 in a charged state, and to interact with the hydroxyl groups of the glucosyl unit in the B site.

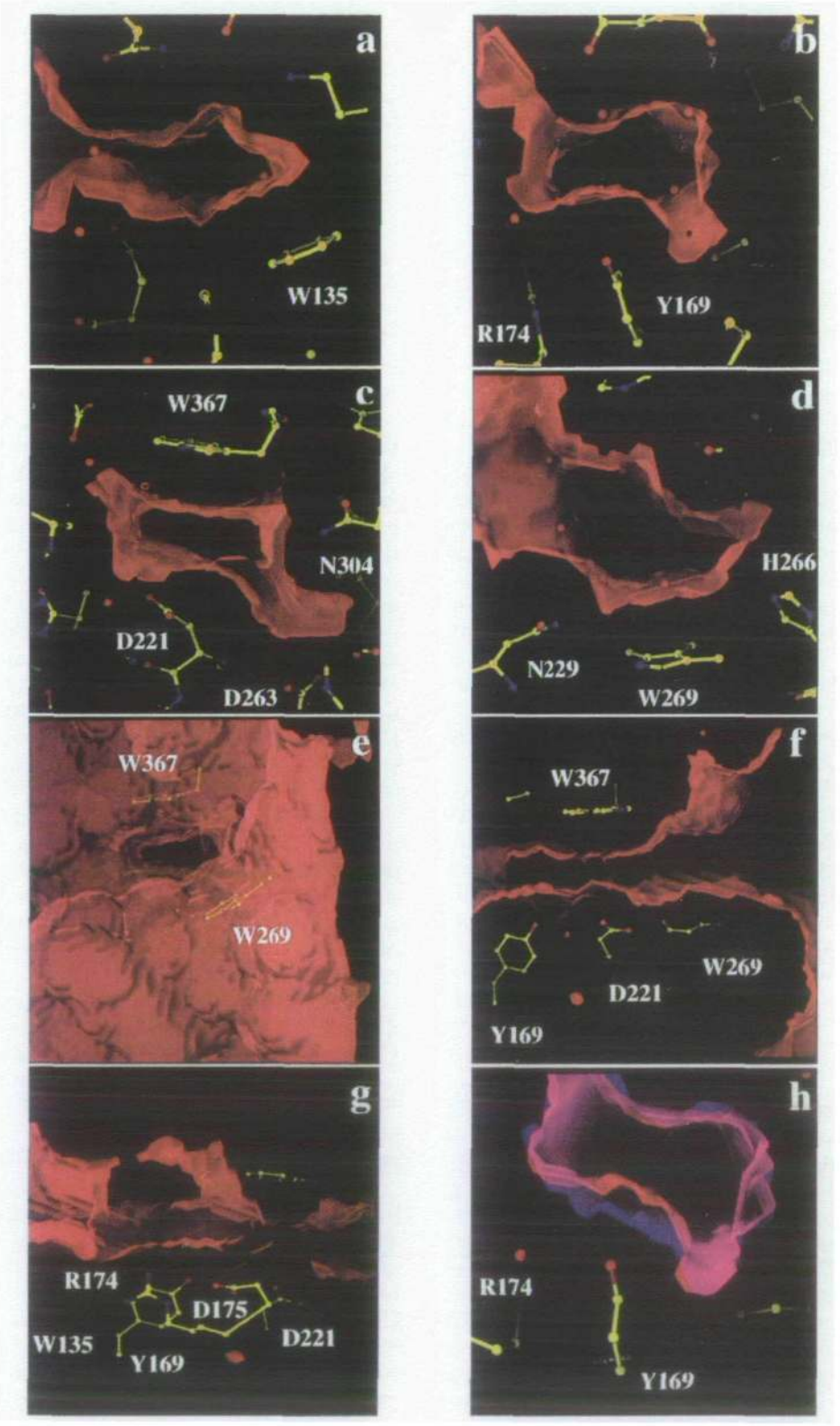


Fig. 4. (a)–(d) Cross-sections of the four sugar binding sites (A-D) of CBHII represented as solvent-accessible surfaces. Some of the important amino acid side chains are labelled, including three tryptophan residues (W135, W367 and W269) that help to form the binding sites A, C and D, respectively; and Y169 and R174 at binding site B. The non-reducing end of the cellulose chain binds to site A and the glycosidic linkage is hydrolysed between sites B and C. (e) A view through the active site tunnel from binding site D (W269) towards binding site A at far end of the tunnel. (f) A side view of the active site tunnel showing the tryptophan residues W367 and W269 at the binding sites C and D, respectively. (g) A close-up view of the centre of the active site tunnel showing the acidic residues D175 (stabilizer) and D221 (proton donor). The hydroxyl group of Y169 is also at the hydrogen bond distance of the carboxylate of D175. The guanido group of R174 is stacked parallel to the phenolic ring of of Y169, and is in addition in hydrogen bonding distance of the carboxylate of D175. (h) Comparison of the cross-sections of site B of CBHII wt and Y169F mutant.

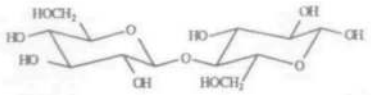
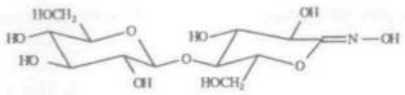
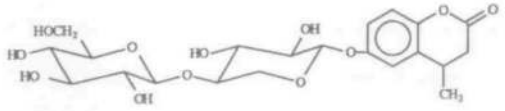
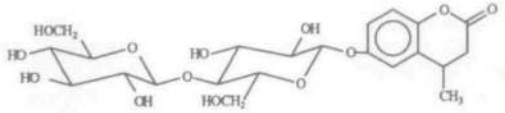
Ligand	K_{ass} at 25°C (M^{-1})	Crystal structure	Binding mode	Ref.
 Cellobiose	$(5.40\pm0.05)\cdot10^2$ $1.3\cdot10^2$ at 8°C	Yes	XA+CD or AB+CD	van Tilbeurgh *et al.*, 1985 Rouvinen, 1990 Anu Koivula, unpubl.results
 Cellobionolactonoxime	$(5.58\pm0.07)\cdot10^3$ at 7.5°C	No	Not known but probably AB	van Tilbeurgh *et al.*, 1985
 MethylUmbelliferyl(GlcXyl)	$(4.90\pm0.05)\cdot10^7$	No	Not known but probably ABC	van Tilbeurgh *et al.*, 1989
 MethylUmbelliferyl$(Glc)_2$	$(2.00\pm0.03)\cdot10^5$	Yes	ABC	van Tilbeurgh *et al.*, 1989

Fig. 5. Oligosaccharides used to study the binding affinity and active site structure of CBHII. Different sugars are represented in orientations where they are known or suggested to bind in the active site. The measured K_{ass} values, complex structures and binding modes are shown when they are known.

The present binding and kinetic data of the CBHII Y169F mutant support these ideas. For all ligands studied, the Y169F mutant shows no change or an increase in the association constants (Table II). The K_{ass} value for glucose, which very likely binds to the subsite A (Rouvinen *et al.*, 1990), is practically the same for the wt and the mutant (Table II). The unliganded structure of the Y169F mutant reveals no changes in the area of subsite A, and therefore the mutation clearly has no effect on the function of this subsite.

The binding affinity for both cellobiose and cellotriose is increased in the mutant enzyme. Crystallographic data for the cellobiose–CBHII wt complex shows density for a single glucose moiety at site A and for a cellobiose molecule in sites C and D (Rouvinen, 1990). Since CBHII does not cleave cellobiose, it is likely that the second glucose moiety of the first cellobiose molecule is disordered and placed at site B or that its first glucose moiety is located outside the tunnel. The present binding data suggest that a significant component corresponds to the former situation (Table II). Our earlier experiments on cellotriose hydrolysis provide evidence that cellotriose can bind in two different ways in the tunnel, either productively (to A–B–C) or non-productively (B–C–D) (Teleman *et al.*, 1995). In both cases the removal of the tyrosine hydroxyl can affect the binding affinity.

The most striking difference in K_{ass} values was observed for MeUmb$(Glc)_2$, which was over 50-fold higher for the mutant (Table II). A very similar increase has been observed upon the binding of MeUmb(GlcXyl) versus MeUmb$(Glc)_2$ in the wt CBHII (Figure 5). Spectroscopic studies indicated that both ligands are likely to occupy the same subsites in the wild-type enzyme (van Tilbeurgh *et al.*, 1989). The complex structure of CBHII and MeUmb$(Glc)_2$ (Rouvinen *et al.*, 1990) shows that the glucosyl units occupy sites A and B and that the large 4-methylumbelliferyl ring binds at site C in close contact with the indole ring of W367. It seems that tighter binding at site B can be achieved in two complementary ways: either by removing the methylhydroxyl group of the ligand [in MeUmb(GlcXyl)], or by removing the hydroxyl of tyrosine (in the Y169F mutant), both of which create more space at site B (Figure 4h). It is tempting to speculate that this is a ring distortion effect. With more space at site B the sugar ring could adopt a more relaxed conformation while in the tighter space (in the wild-type enzyme) the ring may be forced to adopt a strained conformation.

The specificity constant (k_{cat}/K_m) of the wt CBHII is 700-fold higher for cellotetraose than for cellotriose hydrolysis (Table I), owing to both increased binding and reaction rate. This shows that the sugar binding at subsite D is an important factor in transition state stabilization. It seems plausible that binding of the fourth sugar ring in site D causes strain, which is manifested as a ring distortion at site B. This gains support from the fact that no ligand-induced conformational changes have been observed in the CBHII protein complex structures determined with different ligands (Rouvinen *et al.*, 1990). In turn, catalysis would be facilitated if, on binding at subsite B, the glucose unit were forced into a conformation similar to the transition state. Data obtained with many other inverting and retaining carbohydrases support the involvement of ring distortion as an element in their catalytic mechanism (Schindler *et al.*, 1977; Strynadka and James, 1991; Varghese *et al.*, 1992; Harris *et al.*, 1993; Kuroki *et al.*, 1993). In addition, in both viral and bacterial sialidases a tyrosine residue has been

suggested to stabilize the oxocarbonium ion intermediate (Burmeister *et al.*, 1993; Crenell *et al.*, 1993). The molecular dynamics simulation of an inverting endoglucanase suggested that a conserved tyrosine might have an active role in catalysis in the proton transfer process (Taylor *et al.*, 1995). Finally, according to Davies and coworkers (1995), complex structures of another inverting endocellulase have revealed a stretched subsite favouring the binding of an elongated transition state.

Our kinetic data on cellotriose and cellotetraose (Table I) show that the increased affinity of the mutant Y169F correlates with a decrease in the rate of hydrolysis and in the specificity constants. This is consistent with the hypothesis that a distorted sugar conformation at site B facilitates catalysis. Removal of a critical hydroxyl group at site B by the mutation Y169F could improve binding of a less strained conformation of the sugar ring, leading to decreased k_{cat} values for both cellotriose and cellotetraose. In addition, the simple Michaelis–Menten kinetics for cellotriose hydrolysis have changed. This is seen indirectly from Tables I and II in that the calculated K_m for cellotriose is considerably larger than the measured dissociation constant. This can be due either to changed behaviour of the mutant as a function of temperature and/or salt concentration, or to a lowered off-rate for cellotriose.

Finally, the pH dependence of cellotetraose hydrolysis (Figure 3) is clearly different between the Y169F mutant and the wild-type enzyme. The wild-type enzyme retains high relative activity in the acidic pH range whereas the activity of the mutant decreases more drastically. The stability of the CBHII wt and Y169F mutant was checked at pH 3 and no changes could be observed during 3 h of incubation at 27°C (our unpublished data). Since the altered pH activity profile of Y169F is not caused by a loss in pH stability we have two possible explanations for the observed behaviour: according to the first, the binding interactions to the sugar ring at site B have changed in a way that leads to significantly increased K_m at low pH. According to the second, and more appealing, explanation, the hydrogen bonding network around D175 has changed and affects the hydrolysis rate of the mutant. We propose that one of the functions of Y169 is to modulate the protonation states of the interacting carboxylates of D175 and D221. Maintaining the negative charge of D175 stabilizes the charged reaction intermediates and ensures the protonation of D221. The structure reveals that the pK_a of Y169 may in turn be affected by its planar stacking interaction with the guanidino group of R174.

In conclusion, we propose that Y169 has a dual role in the active site of CBHII. First, Y169 influences the glycosyl ring at site B, possibly by formation of a direct hydrogen bond, which leads to a conformation that may correspond to a ring distortion. Second, the pH behaviour of the Y169F mutant supports the hypothesis that Y169 also contributes to the network of interactions ensuring the protonation of the acid catalyst, D221. A recent molecular dynamics simulation on *Thermomonospora fusca* endocellulase 2 (E2) (belonging to the cellulase family 6) suggests that the sugar ring in the second binding site may tilt and that the conserved tyrosine (corresponding to Y169 of CBHII) might have an active role in catalysis in the proton transfer process (Taylor *et al.*, 1995). Although our experimental data on *T.reesei* CBHII does not fully agree with these findings, they give support to the concept that the sugar ring at site B has an unusual conformation and that a tyrosine residue in addition to the active site carboxylates is involved in the catalytic action of family 6 cellulases.

Acknowledgements

We are grateful to Dr Arja Lappalainen for help with protein purification and to Professor Frank Loontiens for fruitful discussions concerning the binding studies. We also thank Tuula Kuurila, Riitta Suihkonen and Ulla Lahtinen for skilful technical assistance. Dr Bruno Villoutreix is thanked for help with Figure 4. The financial support of Nordisk Industrifond, the Swedish Natural Science Research Council, VTT Research Programme Cream and the EU Biotechnology Programme is gratefully acknowledged.

References

Aho,S., Olkkonen,V., Jalava,T., Paloheimo,M., Bühler,R., Niku-Paavola, M.-L., Bamford,D.H. and Korhola,M. (1991) *Eur. J. Biochem.*, **200**, 643–649.

Bailey,M.J. and Nevalainen,H. (1981) *Enzyme Microbiol Technol.*, **3**, 153–157.

Brünger,A.T. (1992) *Nature*, **355**, 472–475.

Brünger,A.T., Kuryian,J. and Karplus,M. (1987) *Science*, **235**, 458–460.

Burmeister,W.P., Henrissat,B., Bosso,C., Cusak,S. and Ruigrok,R.W.H. (1993) *Structure*, **1**, 19–26.

Clayssens,M., Tomme,P., Curtis,F.B. and Hehre,E.J. (1990) *FEBS Lett.*, **263**, 89–92.

Collaborative Computational Project, Number 4 (1994) *Acta Crystallogr.*, **D50**, 760–763.

Crenell,S.J., Garman,E.F., Laver,W.G., Vimr,E.R. and Taylor,G.L. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 9852–9856.

Davies,G., Tolley,S.P., Henrissat,B., Hjort,C. and Schülein,M. (1995) *Biochemistry*, **34**, 16210–16220.

and Henrissat,B. (1995) *Structure*, **3**, 853–859.

De Boeck,H., Matta,K.L., Claeyssens,M., Sharon,N. and Loontiens,F.G. (1983) *Eur. J. Biochem.*, **131**, 453–460.

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

Engh,R.A. and Huber,R. (1991) *Acta Crystallogr.*, **A47**, 392–400

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

Harris,E., Aleshin,A., Firsov,L.M. and Honzatko,R. (1993) *Biochemistry*, **32**, 1618–1626.

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

Ho,S.N., Hunt,H.D., Horton,R.M., Pullen,J.K. and Pease,L.R. (1989) *Gene*, **77**, 51–59.

Irwin,D.C., Spezio,M., Walker,L.P. and Wilson,D.B. (1993) *Biotechnol. Bioengng*, **42**, 1002–1013.

Jones,T.A., Zou,J.Y., Cowan,S.W. and Kjeldgaard,M. (1991) *Acta Crystallogr.*, **A47**, 110–119.

Kasumi,T., Tsumuraya,Y., Brewer,C.F., Kersters-Hilderson,H., Claeyssens,M. and Hehre,E.J. (1987) *Biochemistry*, **26**, 3010–3016.

Kleywegt,G.J. and Jones,T.A. (1994) In Bailey,S., Hubbard,R. and Waller,D.A. (eds), *From First Map to Final Model*. SERC Daresbury Laboratory, Daresbury, pp. 59–66.

Kleywegt,G.J. and Jones,T.A. (1995) *Structure*, **3**, 535–540.

Kleywegt,G.J. and Jones,T.A. (1996) *Acta Crystallogr. D*, in press.

Knowles,J.K.C., Lehtovaara,P., Murray,M. and Sinnott,M.L. (1988) *J. Chem. Soc., Chem. Commun.*, 1401–1402.

Konstantinidis,A.K., Marsden,I. and Sinnott,M.L. (1993) *Biochem. J.*, **291**, 883–888.

Koshland,D.E. (1953) *Biol. Rev.*, **28**, 416.

Kuroki,R., Weaver,L.H. and Matthews,B.W. (1993) *Science*, **262**, 2030–2033.

Laemmli,U.K. (1970) *Nature*, **227**, 680–685.

Laskowski,R.A., MacArthur,M.W. and Thornton,J.M. (1994) In Bailey,S., Hubbard,R. and Waller,D.A. (eds), *From First Map to Final Model*. SERC Daresbury Laboratory, Daresbury, pp. 149–159.

Marquart,D.W. (1963) *J. Soc. Ind. Appl. Math.*, **11**, 431–441.

Mattern,I.E., Punt,P.T., Unkles,S., Powels,P.H. and van den Hondel,C.A.M.J.J. (1987) *Abstracts of the 19th Lunteren Lectures on Molecular Biology*. Lunteren, The Netherlands, p. 34.

McCarter,J.D. and Withers,S.G. (1994) *Curr. Opin. Struct. Biol.*, **4**, 885–892.

Nyyssönen,E., Penttilä,M., Harkki,A., Saloheimo,A., Knowles,J.K.C. and Keränen,S. (1993) *Bio/Technology*, **11**, 591–595.

Penttilä,M., Nevalainen,H., Rättö,M., Salminen,E. and Knowles,J. (1987) *Gene*, **61**, 155–164.

Reinikainen,T., Henriksson,K., Siika-aho,M., Teleman,O. and Poutanen,K. (1995) *Enzyme Microbiol. Technol.*, **17**, 888–892.

Rouvinen,J. (1990) Doctoral thesis, University of Joensuu.

Rouvinen,J., Bergfors,T., Teeri,T., Knowles,J.K.C. and Jones,T.A. (1990) *Science*, **249**, 380–386.

Ruohonen,L. *et al.* (1993) In Suominen,P. and Reinikainen,T. (eds) *Trichoderma reesei Cellulases and Other Hydrolases. Enzyme Structures, Biochemistry, Genetics and Applications.* Foundation for Biotechnical and Industrial Fermentation Research, Helsinki, Vol. 8, pp. 87–96.
Schindler,M., Assaf,Y., Sharon,N. and Chipman,D.M. (1977) *Biochemistry*, **16**, 423–431.
Sinnott,M. (1990) *Chem. Rev.*, **90**, 1171–1202.
Srisodsuk,M., Reinikainen,T., Penttilä,M. and Teeri,T. (1993) *J. Biol. Chem.*, **268**, 20756–20761.
Strynadka,N.C.J. and James,M.N.G. (1991) *J. Mol. Biol.*, **220**, 401–424.
Suominen,P.L., Mäntylä,A., Karhunen,T., Hakola,H. and Nevalainen,H. (1993) *Mol. Gen. Genet.*, **241**, 523–530.
Taylor,J.S., Teo,B., Wilson,D.B. and Brady,J.W. (1995) *Protein Engng*, **8**, 1145–1152.
Teeri,T.T., Kumar,V., Lehtovaara,P. and Knowles,J.K.C. (1987) *Anal. Biochem.*, **164**, 60–67.
Teeri,T.T., Penttilä,M., Keränen,S., Nevalainen,H. and Knowles,J.K.C. (1992) In Finkelstein,D.B. and Ball,C. (eds). *Biotechnology of Filamentous Fungi*. Butterworth Heinemann, Stoneham, MA, pp. 417–445.
Teleman,A., Koivula,A., Reinikainen,T., Valkeajärvi,A., Teeri,T.T., Drakenberg,T. and Teleman,O. (1995) *Eur. J. Biochem.*, **231**, 250–258.
Tomme,P. (1991) Doctoral thesis, Rijksuniversiteit Gent.
Tomme,P., McCrae,S., Wood,T.M. and Claeyssens,M. (1988a) *Methods Enzymol.*, **160**, 187–192.
Tomme,P., van Tilbeurgh,H., Pettersson,O., van Damme,J., Vandekerckhove,J., Knowles,J., Teeri,T. and Claeysens,M. (1988b) *Eur. J. Biochem.*, **170**, 575–581.
Towbin,H., Staehelin,T. and Gordon,J. (1979) *Proc. Natl Acad. Sci. USA*, **76**, 4350–4354.
van Tilbeurgh,H. (1986) Doctoral thesis, Rijksuniversiteit Gent.
van Tilbeurgh,H., Pettersson,G., Bhikabhai,R., de Boeck,H. and Claeyssens,M. (1985) *Eur. J. Biochem.*, **148**, 329–334.
van Tilbeurgh,H., Loontiens,F.G., de Bruyne,C.K. and Claeyssens,M. (1988) *Methods Enzymol.*, **160**, 46–59.
van Tilbeurgh,H., Loontiens,F., Engelborgs,Y. and Claeyssens,M. (1989) *Eur. J. Biochem.*, **184**, 553–559.
Varghese,J.N., McKimm-Breschkin,J.L., Caldwell,J.B., Kortt,A.A. and Colman,P.M. (1992) *Proteins: Struct. Funct. Genet.*, **14**, 327–332.
Wetlaufer,D.B. (1962) *Adv. Protein Chem.*, **17**, 303–390.
Wood,T.M. and Garcia-Campayo,V. (1990) *Biodegradation*, **1**, 147–161.

Received January 2, 1996; revised April 10, 1996; accepted April 18, 1996