Volume 243, number 2, 239–243 FEB 06725 January 1989

Identification of a functionally important carboxyl group in cellobiohydrolase I from *Trichoderma reesei*

A chemical modification study

Peter Tomme and Marc Claeyssens

Laboratorium voor Biochemie, Rijksuniversiteit Gent, K.L. Ledeganckstraat 35, 9000 Gent, Belgium

Received 27 September 1988, revised version received 27 November 1988

Several aspects of the specific modification of cellobiohydrolase I (CBH I) from *Trichoderma reesei* with Woodward's reagent K (*N*-ethyl-5-phenylisoxazolium-3′-sulfonate) are presented. The pH dependence of the resulting inactivation points to the implication of an ionising group with a pK_a of approx. 5.5. The rapid inactivation kinetics, the specific protection and the stoichiometry of modification (3 versus 2 residues), together with the isolation and amino acid sequencing of the putative active site peptide, provide a large body of evidence for the presence of a catalytically important carboxyl residue in the 125–135 region of the CBH I amino acid sequence. From the striking homology between this peptide sequence and those of the active site regions of different lysozymes, glutamic acid 126 is retained as the most plausible catalytic residue (proton donor) in CBH I, equivalent to glutamic acid 35 in hen egg white lysozyme. Glutamic acid 127 is proposed as a potential active site residue to the homologous endoglucanase I (EG I) isolated from the same *Trichoderma* species.

Cellulase; Cellobiohydrolase; Active site; Modification; Homology; Labeling

1. INTRODUCTION

The fungus *Trichoderma reesei*, one of the most efficient cellulolytic organisms, secretes an enzyme system composed of two cellobiohydrolases, CBH I and CBH II (EC 3.2.1.91), at least three endoglucanases, EG I, II and III (EC 3.2.1.4) and several β-glucosidase (EC 3.2.1.21).

Correspondence address: M. Claeyssens, Laboratorium voor Biochemie, Rijksuniversiteit Gent, K.L. Ledeganckstraat 35, 9000 Gent, Belgium

Abbreviations: CBH I, cellobiohydrolase (EC 3.2.1.91); core I, core-protein of CBH I; EG, endoglucanase (EC 3.2.1.4); WRK, *N*-ethyl-5-phenylisoxazolium-3′-sulfonate; DEP, diethylpyrocarbonate; NBS, *N*-bromosuccinimide; TNM, tetranitromethane; Mes, *N*-morpholinomethane-sulfonic acid; Tris, tris(hydroxymethyl)aminomethane; CNPL, 2′-chloro-4′-nitrophenyl β-lactoside; TFA, trifluoroacetic acid; RP-HPLC, reversed-phase high pressure liquid chromatography; DTT, dithiothreitol; Gu.HCl, guanidine hydrochloride

The cloning and/or sequencing of some of these enzymes and their genes have led to detailed structural information [1–7]. Thus EG I and CBH I share about 45% sequence homology [6]. Partial proteolysis studies of CBH I [8] and CBH II [9] have further revealed the existence of two distinct functional domains: a binding region for insoluble cellulose and a core protein containing the active (hydrolytic) site. The binding domains have been identified with terminal, glycosylated amino acid sequences (block AB (B′)) conserved in both enzymes [3,7]. From physical measurements on the intact enzymes and their core proteins, it was shown that these binding domains protrude from the ellipsoidal cores as flexible tails [10].

Although a 'lysozyme-like' mechanism was postulated for these glycosylases as early as 1963 [11], only sparse evidence has been obtained so far to support this hypothesis [12–14].

Modification studies of carboxyl group(s) in CBH I and its core protein are reported here. The

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/89/$3.50

isolation of an active site peptide containing a putatively essential glutamate group is described.

2. MATERIALS AND METHODS

2.1. *Materials*

N-Ethyl-5-phenylisoxazolium-3'-sulfonate (Woodward's reagent K, WRK) was purchased from Sigma (USA); diethylpyrocarbonate (DEP), *N*-bromosuccinimide (NBS) and tetranitromethane (TNM) were from Aldrich (Belgium). Biogel P-6 and P-100 were obtained from Bio-Rad (USA). Microcrystalline cellulose (Avicel pH 101) was from FMC Corp. (USA). 2'-Chloro-4-nitrophenyl β-lactoside was prepared conventionally [15]. *Rhodymenia palmate* xylan was a gift from Dr R. Sturgeon (Edinburgh). The molecular mass was estimated to be approx. 3000 Da (d.p. 25–30) (Dr Abuja, personal communication). All other substrates, ligands and chemicals were either prepared in our laboratory or obtained commercially.

2.2. *Enzymes and enzymic assays*

CBH I was purified from the culture filtrate of the fungus *Trichoderma reesei* QM 9414 (kindly donated by Dr J. Knowles, VTT, Espoo, Finland) by affinity chromatography [16]. Core I protein was obtained by papain digestion and purified as previously described [8]. The molar absorption coefficients (280 nm) of intact CBH I and its core are identical $\epsilon_m = 73000\ M^{-1} \cdot cm^{-1}$ [8].

CNPL (1 mM) was used as a routine substrate to determine CBH I activities at pH 5.7 and 37°C. Interaction (inhibition) constants of ligands were determined conventionally ($1/v$ vs [I] plots) or by active site titrations as described in [17], using the same lactoside as indicator substrate. Avicel suspensions ($30\ mg \cdot ml^{-1}$ in 50 mM sodium acetate/acetic acid, pH 5.0) were used to study the activity (50°C) and adsorption (20°C) of CBH I against and onto this substrate as described [9].

2.3. *Enzyme modifications*

An aliquot of a fresh stock solution of WRK reagent [18] (250 mM in 1 mM HCl) was diluted 1:2 in 50 mM Mes/NaOH buffer, pH 5.5, and added (20°C) to CBH I or core I (15–20 μM) (final WRK concentrations 10–70 mM). The pH of the stirred reaction mixture was kept constant by manual titration (0.1 N NaOH), at indicated times, 30 μl aliquots were diluted in 400 μl 50 mM sodium acetate, pH 5.7, to quench the inactivation reaction, and the residual activity measured using CNPL. Inactivations in the presence of protecting ligands were performed similarly, and the pH dependence was investigated by incubating 17.5 μM enzyme with 50 mM WRK in 50 mM Mes/NaOH buffers of pH range 4.6–6.5. Modifications with other reagents were according to published procedures: histidyl residues with diethylpyrocarbonate (DEP) [19]; tryptophanyl residues with *N*-bromosuccinimide (NBS) [20]; tyrosyl residues with tetranitromethane (TNM) [21].

2.4. *Isolation of WRK-modified enzyme and peptide analysis*

Core I solution (2 ml 15–20 μM) was treated with WRK (50–150 mM) in 50 mM Mes/NaOH buffer, pH 5.5 (20°C) in the absence or presence of ligands (conc. given). After adding sodium acetate (250 mM final concentration) the mixture was applied immediately onto a Biogel P-6 column (20 × 1 cm) equilibrated with 50 mM sodium acetate, pH 5.0, or 10 mM Tris-HCl buffer, pH 8.0 ($1\ ml \cdot min^{-1}$). UV-difference spectra of the protein fractions were recorded (Uvicon 810, Kontron, Switzerland), and the degree of modification was determined from the 340/280 nm ratio using $\epsilon_M = 7000\ M^{-1} \cdot cm^{-1}$, pH 8.0, at 340 nm [22]. The modified enzyme (700 μl, 30–50 μM in pH 8.0 buffer) was treated subsequently with 0.7 g Gu.HCl and 2.8 mg DTT under N_2 and left at 25°C for 4 h [23]. The reduced sample, diluted with 1.5 ml 100 mM Tris-HCl buffer, pH 8.0, and concentrated to ~0.5 ml, was treated with subtilisin (Boehringer, FRG) (1/100, w/w; 10 min; 37°C). A sample (150 μg) was injected on a C4-Vydac 214 TP-54 column (0.46 × 25 cm) and eluted with a linear gradient of 0.05% TFA (solvent A) and 70% CH_3CN/0.05% TFA (solvent B). Fractions absorbing at 340 nm were collected and used for amino acid and N-terminal sequencing analysis according to published procedures [24] (gas-liquid-phase sequencer, Applied Biosystems Inc., USA).

3. RESULTS AND DISCUSSION

3.1. *Inactivation reaction kinetics*

After 9 min incubation of 17.5 μM CBH I with

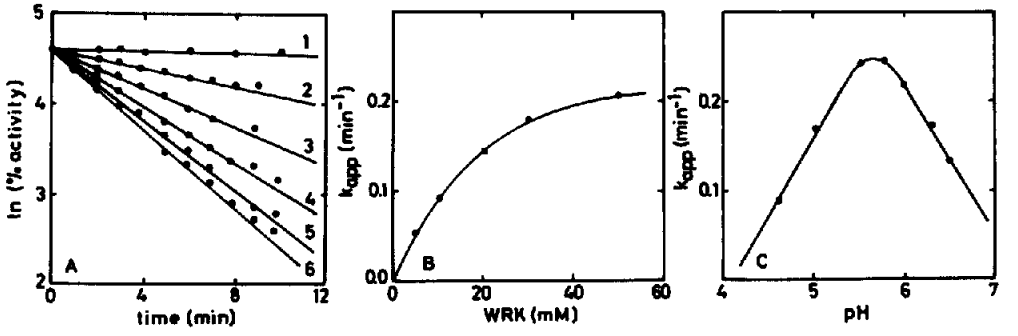

Fig.1. (A) Inactivation of 17.5 μM CBH I with various concentrations of WRK: (1) 0, (2) 5, (3) 10, (4) 20, (5) 30 and (6) 50 mM in 50 mM Mes/NaOH buffer, pH 5.5 (20°C). Residual activities were determined as described in section 2. (B) Plot of the k_{app} values, derived from the slopes of the lines in A, versus WRK concentration. (C) Effect of pH on k_{app} (see section 2).

50 mM WRK (pH 5.5, 20°C) residual activity against the lactoside is only 10%. The time-course of inactivation (fig.1A) is linear, and a plot of the pseudo-first-order constants against the modifier concentration is hyperbolic (fig.1B). The reaction probably proceeds with formation of an intermediary, reversible enzyme-modifier complex (saturation with respect to WRK) and the modification thus seems active site directed. From the linearised [WRK]/k_{app} versus [WRK] plot (not shown), the composite association (K_c = 44.5 M^{-1}), the maximal inactivation (k_{max} = 6.3 min^{-1}) and the second-order rate constants (k' = 13 $min^{-1} \cdot M^{-1}$) can be derived. Analysis [25] of a log (k_{app}) versus log [WRK] plot (not shown) yields a reaction order with respect to WRK of 1.1 (± 0.2) and thus indicates that at least one essential group is present in CBH I. The modification proceeds most rapidly at pH 5.5 and the bell-shaped pH inactivation curve indicates that two groups with apparent pK_a values of 5.4 and 5.8 are implicated (fig.1C). This could either be (i) histidyl group(s) with an abnormally low pK_a or (ii) carboxyl group(s) with a high pK_a value. As the pH-activity curves for CBH I are similarly shaped and two catalytic important groups with pK_a values of 3.7 and 5.0 can be deduced (Van Tilbeurgh et al., unpublished), modification with WRK could be directed towards the latter group. Proofs regarding the nature of the modified residue and the homogeneity of the partially and fully modified enzymes are given below (3.2 and 3.3).

Cellobiose (40 μM, K_i = 16 μM), lactose (400 μM, K_i = 312 μM), as competitive inhibitors, as well as substrates (e.g. cellotriose 140 μM, K_{mi} = 12 μM) protect effectively against inactivation (66, 65 and 77%, respectively).

Due to the high interaction constant of xylan (*Rhodymenia palmata*) (K_i = 0.4 μM) and the difficulty in removing (dialysis) this high molecular mass inhibitor, protective effects cannot unambiguously be evaluated. A correlation between modification and protection by this ligand can, however, clearly be established (see below).

3.2. *Stoichiometry and specificity of the modification reaction*

A characteristic maximum at 340 nm in UV spectra of CBH I, indicative of a chromophore formed by reaction between the enol-ester of WRK and carboxyl groups of the enzyme [18,26], can be demonstrated. The number of modified carboxyl groups estimated from the 340/280 nm ratio amounts to three under the conditions given (17.5 μM CBH I, 50 mM WRK, pH 5.5). In the presence of 20–100 μM xylan, one group can obviously be protected. With smaller or less tightly binding inhibitors or substrates (e.g. cellobiose or cellotriose) the effect is only partial (2.5–3.0 groups modified). This can, however, be artefactual, since after gel filtration modified enzyme preparations are obtained with lower specific activities, indicating incomplete protection and further inactivation during this purification step.

WRK can modify amino acid residues other than carboxyl groups (e.g. Trp, Tyr, His and Cys). All cysteinyl residues in CBH I are present in disulfide bridges [27] and reaction of these residues with WRK can thus be excluded. *N*-Ethoxycarboxylation of three histidyl groups with DEP in CBH I does not lead to loss of activity, proving the nonessentiality of these residues (not shown). Reaction with NBS or with TNM results in complete inactivation (modification of Trp) or threefold activation (nitration of Tyr), respectively. The former inactivation reaction, however, cannot be prevented by the use of specific inhibitors such as cellobiose or cellotriose (unpublished results). Specific modification of carboxyl groups in CBH I by WRK can thus be assumed.

3.3. *Properties of the modified enzyme*

After partial modification (quenching of inactivation reaction with acetate, 24% residual activity) the kinetic parameters (pH 5.7, 37°C) of CBH I for CNPL show a small decrease in K_m (504 μM vs 427 μM) but a large decrease in k_{cat} (10.9 min^{-1} vs 2.7 min^{-1}). Activity against Avicel is similarly affected (80% decrease) but the adsorption on the same substrate is only slightly decreased (12%). From our previous work on the bifunctional organisation of CBH I [8], it becomes clear that the modification with WRK is directed against the catalytic domain present in the core fragment and not against the cellulose-binding region (BA glycopeptide domain) which is seemingly left intact. An enzyme sample with 40% residual activity is bound in approximately the same amount onto an affinity column carrying *p*-aminobenzyl 1-thio cellobioside as ligand and the retained fraction is

completely desorbed by lactose [16]. This proves the homogeneity of the modified enzyme preparations and again the active site-directed character of the modification.

Also the mobility of the enzyme on SDS-PAGE is unaffected by partial or complete WRK inactivation, proving that no intermolecular cross-linking occurred.

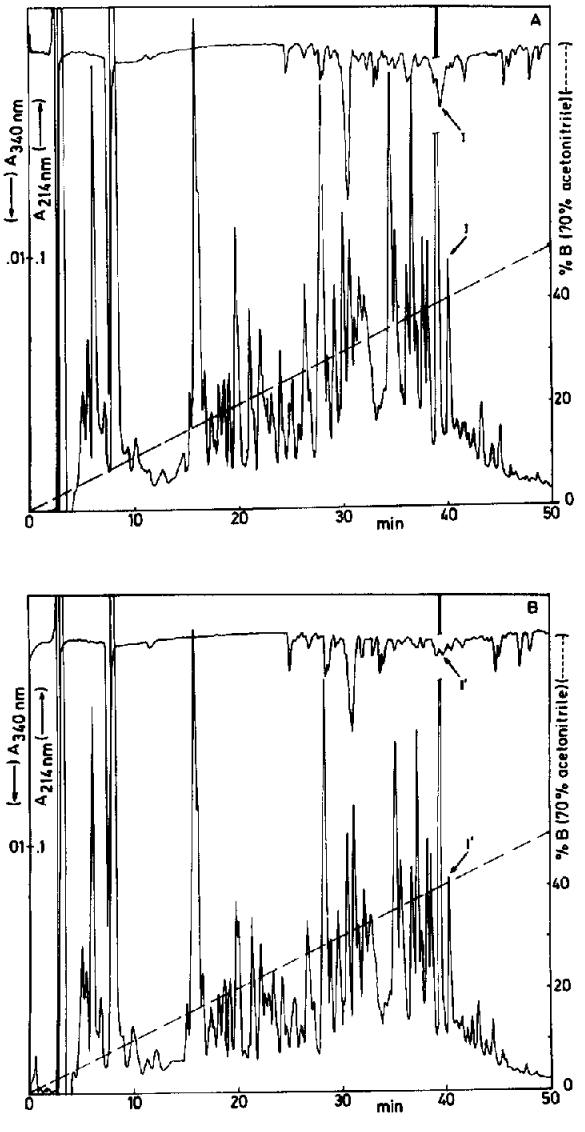

Fig.2. RP-HPLC elution profiles of a subtilisin digest of reduced CBH I after modification in the absence (A) or presence (B) of 100 μM xylan; 214 (lower) and 340 nm (upper) tracings. Peptides I and I′ were collected and sequenced.

3.4. *Isolation of an active-site peptide*

Peptide mapping (subtilisin) with specific detection at 340 nm [28] was tried to isolate modified residues (fig.2). Chromatograms of modified and protected (xylan) enzyme samples coincide (214 tracings). The absence of one peak (340 nm tracings) in the latter case provides proof for the presence of one or more active-site residues in these peptides (I′ versus I, fig.2). The results of the N-terminal sequencing analysis (10–15 steps) were identical for both (fig.3) and amino acid analysis reveals that their actual length is approximately 13–17 residues (results not given). The corresponding region in the CBH I sequence [1,2] is found between residues 125–135 as shown in the alignment (fig.3). Three carboxyl groups are present and it cannot be unambiguously determined which were modified.

A homology search with known active-site regions of several lysozymes [29] is, however, revealing (fig.3). Taking into account possible deletions and inversions these sequences seem highly conserved and the choice of Glu-126 as a putative active site residue in CBH I becomes plausible. EG I, with an equally conserved amino acid sequence [6], is inactivated by WRK very similarly to CBH I (Tomme, unpublished), and Glu-127 is consequently proposed as an active site residue (fig.3). The location of these residues, both in CBH I and EG I, between two large protein domains, as deduced from a disulfide bridge repartition study [27], is diagnostic for an active site (cleft). Although hydrophobicity is high in all these peptide sequences, two Asp residues (130, 132) seem to confer some polarity to the presumed active site of CBH I. This is probably reflected in the somewhat lower pK_a value of the modifiable car-

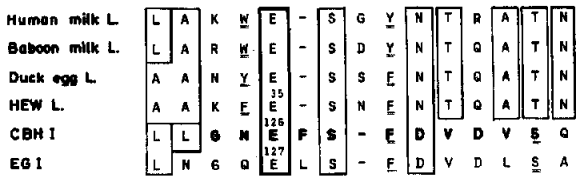

Fig.3. Alignment of the sequence of the isolated peptides (I, I′) (determined sequence in bold face) with the corresponding region in the CBH I sequence [1,2], the homologous region in EG I [6] and known active site regions of several lysozymes [29]. Conserved (boxed), chemically exchangeable (underlined) and putatively deleted (–) amino acids are indicated.

boxyl groups (5.4–5.8) in the latter as compared to that of Glu-35 in hen egg-white lysozyme (pK_a 6.3). The glutamic acids found in this study (Glu-126 in CBH I and Glu-127 in EG I) are obvious targets for site-directed mutagenesis, especially since several successful cloning experiments of these enzymes have been described [30].

Acknowledgements: P.T. is an IWONL bursar (Belgium). M.C. is indebted to NFWO (Belgium) and NATO for grants. Professor G. Pettersson (Uppsala) is thanked for the amino-acid analysis facilities and Professor De Bruyne (Ghent) for encouragement.

REFERENCES

[1] Shoemaker, S., Schweickart, V., Ladner, M., Kwok, S., Myambo, K. and Innis, M. (1983) Bio/technology 1, 685–691.
[2] Teeri, T., Salovuori, I. and Knowles, J. (1983) Bio/technology 1, 696–699.
[3] Teeri, T.T., Lehtovaara, P., Jauppinen, S., Salovuori, I. and Knowles, J. (1987) Gene 51, 43–52.
[4] Van Arsdell, J.N., Kwok, S., Schweickart, V.L., Ladner, M.B., Gelfand, D.H. and Innis, M.A. (1987) Bio/technology 5, 60–64.
[5] Saloheimo, M., Lehtovaara, P., Pentillä, M., Teeri, T.T., Ståhlberg, J., Johansson, G., Pettersson, G., Claeyssens, M., Tomme, P. and Knowles, J. (1988) Gene 63, 11–21.
[6] Pentillä, M., Lehtovaara, P., Nevalainen, H., Bhikhabhai, R. and Knowles, J. (1986) Gene 45, 253–263.
[7] Knowles, J., Tuula, T.T., Lehtovaara, P., Pentillä, M. and Saloheimo, M. (1988) in: Biochemistry and Genetics of Cellulose Degradation (Aubert, J.P., Béguin, P. and Millet, J. eds) pp.157–159, Academic Press, New York.
[8] Van Tilbeurgh, H., Tomme, P., Claeyssens, M., Bhikhabhai, R. and Pettersson, L.G. (1986) FEBS Lett. 204, 223–227.
[9] Tomme, P., Van Tilbeurgh, H., Pettersson, G., Van Damme, J., Vandekerckhove, J., Knowles, J., Teeri, T. and Claeyssens, M. (1988) Eur. J. Biochem. 170, 575–581.
[10] Abuja, P., Schmuck, M., Pilz, I., Tomme, P., Claeyssens, M. and Esterbauer, H. (1988) Eur. Biophys. J. 15, 339–342.
[11] Banks, B.E.C. and Vernon, C.A. (1963) Biochem. J. 86, 7.
[12] Hurst, P.L., Sullivan, P.A. and Shepherd, M.G. (1977) Biochem. J. 167, 549–556.
[13] Legler, G. and Bause, E. (1973) Carbohydr. Res. 28, 45–52.
[14] Clarke, A.J. and Yaguchi, M. (1985) Eur. J. Biochem. 149, 223–228.
[15] De Boeck, H., Matta, K.L., Claeyssens, M., Sharon, N. and Loontiens, F. (1983) Eur. J. Biochem. 131, 453–460.
[16] Van Tilbeurgh, H., Pettersson, L.G., Bhikhabhai, R. and Claeyssens, M. (1984) FEBS Lett. 187, 283–288.
[17] Rabinowitch, M.L., Klyosov, A.A. and Melnick, M.S. (1986) Anal. Biochem. 156, 489–494.
[18] Pétra, P.H. (1971) Biochemistry 10, 3163–3170.
[19] Miles, E.W. (1977) Methods Enzymol. 47, 431–442.
[20] Spande, T.F. and Witkop, B. (1967) Methods Enzymol. 11, 498–506.
[21] Riordan, J.F. and Vallee, B.L. (1972) Methods Enzymol. 25, 515–521.
[22] Sheena, U. and Brewer, J.M. (1985) Anal. Biochem. 151, 327–333.
[23] Tarr, G.E. (1986) in: Methods in Protein Microcharacterisation (Shively, J.E. ed.) p.162, Humana Press, New Jersey.
[24] Hewick, R.M., Hunkapiller, M.W., Hool, L.E. and Dreyer, W.J. (1981) J. Biol. Chem. 256, 7990–7997.
[25] Levy, H.M., Leber, P.D. and Ryan, E.M. (1963) J. Biol. Chem. 238, 3654–3659.
[26] Dunn, B.M., Anfinsen, C.B. and Schrager, R.I. (1974) J. Biol. Chem. 249, 3717–3723.
[27] Bhikhabhai, R. and Pettersson, L.G. (1984) Biochem. J. 222, 729–736.
[28] Sheena, U. and Brewer, J.M. (1986) Biochem. Cell. Biol. 64, 970–975.
[29] Imoto, T., Johnson, L.N., North, A.C.T., Phillips, D.C. and Rupley, J.A. (1972) in: The Enzymes, 3rd edn (Boyer, P.D. ed.) vol.7, pp.666–868.
[30] Pentillä, M.E., André, L., Lehtovaara, P., Bailey, M., Tuula, T.T. and Knowles, J.K.C. (1988) Gene 63, 103–112.