Identification of the catalytic nucleophile in the cellulase from *Schizophyllum commune* and assignment of the enzyme to Family 5, subtype 5 of the glycosidases

Anthony J. Clarke[a,*], Jolyne Drummelsmith[a], Makoto Yaguchi[b]

[a]*Department of Microbiology, University of Guelph, Guelph, Ont. N1G 2W1, Canada*
[b]*Institute of Biological Science, National Research Council of Canada, Ottawa, Ont. K1A 0R6 Canada*

Received 23 June 1997; revised version received 1 August 1997

Abstract Differential chemical modification of the cellulase from *Schizophyllum commune* with [*N*-methyl-^{3}H]1-ethyl-3(4-azonia-4,4-dimethylpentyl)-carbodiimide in the presence and absence of substrate identified an active site glutamate residue within the peptide: Leu-Gln-Ala-Ala-Thr-Glu-Trp-Leu-(Lys). This Glu residue is proposed to participate in binding of substrate as amino acid sequence homology studies combined with mechanism-based inhibition of the cellulase with 4′,5′-epoxypentyl-β-D-cellobioside identified a neighboring Glu residue, which conforms to the Glu-X-Gly motif of Family 5 glycosidases, as the catalytic nucleophile. These data allow the assignment of the *S. commune* cellulase to Family 5, subtype 5 of the glycosidases.

Key words: Cellulase; Catalytic mechanism; Chemical modification; Mechanism-based inhibition; *Schizophyllum commune*

1. Introduction

Cellulases ((1,4-β-D-glucan 4-glucanohydrolase, EC 3.2.1.4) have been grouped into 10 different glycosidase families based on primary sequence homologies [1–3]. To date, Family 5 contains the highest number of the characterized cellulases [4] and this significant database has permitted the subdivision of Family 5 enzymes into five subtypes [5]. Although the complete amino acid sequence of the cellulase from *Schizophyllum commune* has not been reported nor has its gene been cloned and sequenced, it has been tentatively assigned to Family 5 based on very limited amino acid sequence data [6].

Cellulases catalyze the hydrolysis of the β-1,4 glucosyl linkages in cellulose with either the net retention (Families 5, 7, 10, 12) or inversion (Families 6, 9, 44, 45, 48) of configuration at the anomeric center [7,8]. Retaining enzymes catalyze a double displacement mechanism of action and with Family 5 enzymes, both the acid catalyst and stabilizing anion/nucleophile have been identified on the basis of affinity labelling, mechanism-based inactivation, and site-directed mutagenesis studies (recently reviewed in [4,7,8]). Thus, an Asn-Glu-Pro motif distinguishes the acid catalyst, while a Glu-X-Gly sequence (where X is typically an aromatic amino acid residue) characterizes the Glu participating as the stabilizing anion/nucleophile. The participation of acidic residues in the catalytic activity of the *S. commune* cellulase have been previously demonstrated by chemical modification [9], metal-ion inhibition [10], and mechanism-based inactivation [11,12] studies, but a specific amino acid residue was not identified.

In this study, a peptide localized to active site of the *S. commune* cellulase and containing the catalytic nucleophile is identified by chemical modification and mechanism-based inactivation studies, thereby confirming assignment of the enzyme to Family 5.

2. Materials and methods

Carboxymethylcellulose, cellobiose, α-chymotrypsin, DPCC-trypsin, guanidinium hydrochloride (Gdn-HCl), Mes, methylcellulose, and Trizma base were obtained from Sigma Chemical Co. (St. Louis, MO), while Glu-C endoproteinase was purchased from Boehringer-Mannheim (Laval, PQ). 1-ethyl-3(4-azonia-4,4-dimethylpentyl)-carbodiimide (EAC) [6] and its radioactive form, [*N*-methyl-^{3}H]1-ethyl-3(4-azonia-4,4-dimethylpentyl)-carbodiimide ([^{3}H]EAC, sp. act. 0.855 Ci mmol^{-1}) [13], and the mechanism-based inhibitor 4′,5′-epoxypentyl-β-D-cellobioside [11,12] were synthesized as reported previously. NaB[^{3}H]$_4$ was a product of DuPont NEN Research Products (Boston, MA).

The cellulase from *S. commune* was purified to apparent homogeneity by a combination of anion-exchange and affinity chromatography as previously described [9]. Its concentration was determined by amino acid analysis, based on Lys = 6 [9]. Cellulase activity was determined viscometrically using 0.2% carboxymethylcellulose in 50 mM sodium acetate buffer, pH 5.0 at 40°C [9].

Chemical modification reactions of cellulase (16–20 μM) with EAC or [^{3}H]EAC (30–35 mM) were performed in water which was maintained at pH 4.5 with the aid of a pH Stat (Radiometer Copenhagen). In some experiments, the protective ligands cellobiose (57 mM) or methylcellulose (0.8% wt/vol) were preincubated with the enzyme for 10 min prior to the addition of the carbodiimide. Samples (10 μl) were withdrawn at appropriate intervals to assay for residual hydrolytic activity and after incubation for 5–30 min at 25°C, reactions were quenched with the addition of 1 M sodium acetate buffer, pH 6.0. Upon full modification, the enzyme precipitated from solution and so the modified cellulase derivatives were lyophilized and then resuspended in 7 M guanidinium-HCl (Gdn-HCl), 50 mM Tris-HCl, pH 8.3. Residual reagent and buffer salts were removed by chromatography on a 2.5 × 40 cm column of BioGel P-6 DG (BioRad Laboratories Canada Ltd., Mississauga, Ont.), equilibrated and eluted with the same Gdn-HCl buffer.

For differential modification of *S. commune* cellulase, 8 ml of 17.2 μM cellulase was pre-incubated with 180 mM D-cellobiose for 5 min and the pH was brought to 4.2 prior to the addition of 45 mM EAC (final concentration). After 5 min incubation at 25°C with the pH maintained at 4.2 as above, the reaction was quenched as above and the residual reagent and protecting ligand were removed by ultrafiltration using an Immersible-CX 10 Ultrafiltration apparatus (Millipore (Canada) Ltd., Nepean, Ont.). The partially modified cellulase was then treated with 19 mM [^{3}H]EAC at pH 4.2 for 30 min. The reaction was quenched as above and the radioactive enzyme derivative was recovered from the reaction mixture by chromatography on Bio-Gel P-6 DG in 7 M Gdn-HCl as above, followed by dialysis against 7 M Gdn-HCl in 150 mM Tris-HCl, pH 8.3.

Modified enzyme derivatives in 7 M Gdn-HCl were reduced and

*Corresponding author. Fax: +1 (519) 837-1802.
E-mail: AClarke@Micro.UoGuelph.CA

PII S0014-5793(97)01049-1

alkylated with 1.5 mM dithiothreitol and 2 mM iodoacetic acid, respectively. Residual reagents and reaction by-products were removed by exhaustive dialysis against 50 mM acetic acid and then water at 4°C over 20 h, and the cellulase derivatives were lyophilized. The reduced and alkylated enzyme derivatives in 25 mM NH_4HCO_3, pH 7.8 (0.5 mg ml^{-1}) were digested with DPCC-trypsin (approx. 1:25 wt/wt) for 90 min at 37°C, followed by α-chymotrypsin (approx. 1:40 wt/wt) for an additional 90 min. The resulting peptides were concentrated by lyophilization and subjected to reverse-phase HPLC on a Syn-Chrom RP 8 column (Varian, USA). Successive linear gradients of 0–21% 2-propanol (90 min) and 21–56% 2-propanol (70 min) in 5 mM sodium phosphate buffer, pH 7.15 at 1.0 ml min^{-1} were used for elution and eluates were monitored for both absorbance at 230 nm and radioactivity. Fractions of interest were re-chromatographed on the same column using more shallow linear gradients.

For mechanism-based inactivation, cellulase (9.4 μM) in 50 mM sodium acetate buffer, pH 5.0 was incubated at 30°C with 60 mM 4′,5′-epoxypentyl-β-D-cellobioside for 36 h. At appropriate intervals, samples of the reaction mixture were withdrawn and residual enzymatic activity was determined and compared to enzyme controls incubated in the absence of inhibitor. The inactivated cellulase was lyophilized and then resuspended in 0.5 ml 80 mM Tris-HCl, pH 8.5 containing 1 M urea, 2.7 mM EDTA and 18 mM dithiothreitol. After incubation for 2 h at 25°C to effect the reduction of disulfide bonds, the enzyme was treated with 6.2 mM (final concentration) [^{3}H]$NaBH_4$ (2.5 mCi, sp. act. 0.90 Ci $mmol^{-1}$) for 4 h at 37°C to displace the inhibitor by reduction of the ester linkage. The reaction mixture was then acidified by dropwise addition of dilute HCl and reaction by-products were removed by exhaustive dialysis against water at 4°C (72 h). The reduced cellulase (0.6 mg ml^{-1}) was treated with 40 μg Glu-C endoproteinase (1:10 wt/wt) at 25°C for 120 min and peptides were separated by reverse-phase HPLC on a 4.6×250 mm Ultrasphere ODS C_{18} column (Beckman Instruments (Canada) Inc., Mississauga, Ont.) using a 0–35% linear gradient of CH_3CN in 0.1% trifluoroacetic acid over 120 min at 1.0 ml min^{-1}. The eluent was monitored for absorbance at both 220 nm and 280 nm and for radioactivity.

Amino-terminal sequencing of peptides was performed on either a Beckman Model 890D protein sequenator (Beckman) or a Porton Gas-Phase Microsequencer with on-line 120A PTH-amino acid analyzer according to manufacturer's instructions. Amino acid analyses were performed on either Durrum D-500, Applied Biosystems 420H, or Beckman System Gold analyzers.

3. Results and discussion

3.1. Carbodiimide modification of cellulase

The cellulase from *S. commune* was treated with the water-soluble carbodiimide [^{3}H]EAC in both the absence and presence of the substrate methylcellulose (K_M 0.022%). In the absence of protective ligand, 7.1 acidic amino acid residues of the enzyme were modified by [^{3}H]EAC and as reported

Subtype 1	*Cellulomonas fimi* Cel D	WGPNWLFIHDEDIAPLLIGE**W**G**G**RLGQD
Subtype 2	*Bacillus subtilis* EG 1 & 2	FL**R**DKANYALSKGAPIFVT**E**W**G**TSDASG
Subtype 3	*Clostridium thermocellum* Cel C	DLKPAIE**F**REKKKCKLYCG**E**F**G**VIAIAD
Subtype 4	*Clostridium thermocellum* Cel E	ELDAIYNRFVK**N**GRAVIIG**E**F**G**TIDKNN
Subtype 5	*Burkholderia solanacearum* Cel	RLQDFTT**W**LRSNGYRGFLG**E**F**G**AASMDT
	Trichoderma reesei EG II	AFSPLAT**W**LRQ**N**N**R**QAILT**E**T**G**G**G**NVQS
	Robillarda sp. EG 1	NLAALTA**W**LK**E**N**N**LKAFIT**E**F**G**GSNSTS
		*
	Schizophyllum commune Cel	SLQ**AATEW**LK**ANN**QRGFLG**E**M**G**AGSNAD

Fig. 1. Amino acid sequence homology around the catalytic nucleophile of representative members of the Family 5 glycosidases. Residues in bold denote those conserved in the *S. commune* cellulase sequence (reference [6] and unpublished data), while the catalytic Glu and Gly residues of the Glu-X-Gly motif are shaded. The asterisk denotes the active site Glu residue of *S. commune* cellulase identified by carbodiimide modification. The underlined sequence of the isolated peptide (solid line) was determined by Edman degradation while the remainder of its length (dashed line) was derived by amino acid analysis. The amino acid sequences for all other cellulases were taken from [4].

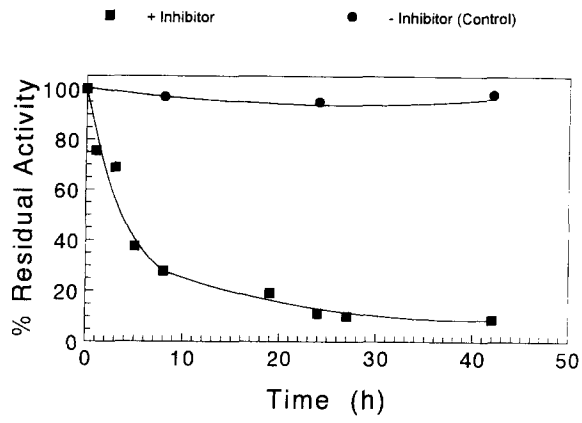

Fig. 2. Mechanism-based inactivation of *S. commune* cellulase by 4′,5′-epoxypentyl-β-D-cellobioside at pH 5.0 and 30°C. At the times indicated, aliquots of the reaction mixture were removed and assayed viscometrically for residual activity (■) and expressed as a percentage of a control which was incubated for the same period of time in the absence of inhibitor (●).

previously [9], such treatment of the enzyme resulted in loss of hydrolytic activity [9]. When the cellulase was preincubated with 0.8% methylcellulose, 6.2 acidic residues reacted with [^{3}H]EAC and the enzyme derivative retained approximately 33% catalytic activity. Similar results were obtained using 57 mM cellobiose as protecting ligand. These data suggested that the substrate and competitive inhibitor protected 0.9 essential residues in the cellulase from modification.

Differential chemical modification of the cellulase involving the incubation of the enzyme with EAC in the presence of 0.8% methylcellulose, followed by (1) removal of protecting ligand and residual reagent by ultrafiltration, and (2) subsequent treatment of the partially modified enzyme with [^{3}H]EAC, resulted with incorporation of a calculated 1.4 molar equivalent of [^{3}H]EAC and the concomitant abolishment of activity. Tryptic and chymotryptic digestion of the reduced and alkylated enzyme derivatives led to the recovery by reverse-phase HPLC of a series of radioactive peptides. The majority of the radioactivity was found to be associated with peptides having N-terminal residues of Leu, Gln and Ala but each conforming to the sequence: Leu-Gln-Ala-Ala-Thr-**Glu**-Trp-Leu-(Lys). In each case, radioactivity was detected in the sequencing cycles corresponding to the Glu residue. These peptides, which reflect the proteolytic specificity of trypsin and chymotrypsin, bear striking homology to sequences within Family 5, subtype 2 and 5 cellulases involving the catalytic nucleophile (Fig. 1). That this Glu residue was subsequently modified by a carbodiimide upon removal of protecting ligand is consistent with it being localized to the active site of the enzyme. However, this Glu residue does not conform to the catalytic nucleophile Glu-X-Gly motif of Family 5 glycosidases. Interestingly, reaction of the cellulase with [^{3}H]EAC in the absence of protecting ligand followed by proteolytic digestion and separation of resulting peptides by HPLC did not lead to the recovery of a radioactive peptide with either the catalytic nucleophile motif or the Asn-Glu-Pro motif of the catalytic acid [14].

3.2. Mechanism-based inhibition of cellulase

Previous studies have shown that the mechanism-based inhibitor 4′,5′-epoxypentyl-β-D-cellobioside causes a time-de-

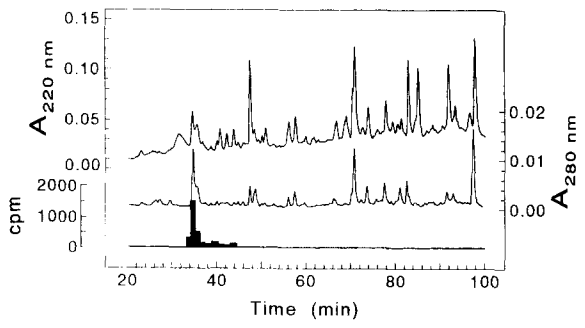

Fig. 3. Isolation of affinity-labelled peptide of *S. commune* cellulase inactivated by 4′,5′-epoxypentyl-β-D-cellobioside by reverse-phase HPLC after both $NaB[^3H]_4$ reduction and digestion with Glu-C endoproteinase.

pendent inactivation of the *S. commune* cellulase and one molar equivalent of the inhibitor covalently binds to the enzyme [11,12]. Identification of the bound amino acid residue was made here using the method developed by Eshdat et al. [15]. After 36 h incubation of the cellulase with the epoxide, and when remaining catalytic activity was less than 10% of an appropriate control (Fig. 2), residual reagent was removed by ultrafiltration. The ester-linked cellobioside was displaced from the inactivated cellulase by reduction with $NaB[^3H]_4$. The reduced and radioactive enzyme derivative was digested with Glu-C endoproteinase and one radioactive peptide, with strong absorbance at 280 nm suggesting the presence of a Trp residue, was isolated by reverse-phase HPLC (Fig. 3). Amino acid sequencing of this peptide gave Ser-Leu-Gln-Ala-Ala-Thr-Glu. That Glu was detected in cycle 7 of the sequencing analysis indicated it had not been esterified by the epoxide inhibitor and subsequently reduced to homoserine. Amino acid analysis of the isolated peptide provided a composition consistent with a larger peptide extending beyond this Glu residue (Fig. 1) and comprising a Glu-X-Gly (Glu-Met-Gly) motif. Presumably, it was this second Glu residue that had been reduced to the radioactive homoserine, since reduction of the terminal Asp residue to homo-homoserine would have precluded proteolytic cleavage at this latter residue.

3.3. Concluding remarks

Epoxy-glycosides have been used to identify the catalytic nucleophile of a number of carbohydrates including a *Bacillus amyloliquefaciens* 1,3-1,4-β-D-glucan 4-glucanohydrolase (EC 3.2.1.73) [16] and the *Trichoderma reesei* EG II (formerly known as EG III) [17]. The latter enzyme is a Family 5, subtype 5 glycosidase and independent confirmation of the identity of the catalytic nucleophile in this family of glycosidases has been obtained with the cellulase CelC from *Clostridium thermocellum* by both mechanism-based inactivation with 2′,4′-dinitrophenyl-2-deoxy-2-fluoro-β-D-cellobioside [5] and X-ray crystallography [18].

In an early amino acid sequence comparison study using the very limited information available at that time, a residue of the *S. commune* cellulase comprising a sequence with homology to the active site sequence of hen egg-white lysozyme was proposed to be the catalytic nucleophile [19]. The data presented here, however, do not support this previous hypothesis but instead, identify an alternative residue which is now known to be completely conserved among the Family 5 glycosidases. The residues around the putative catalytic nucleophile of the *S. commune* cellulase are most homologous to the subtype 5 enzymes, suggesting that this is where the enzyme should be listed.

Acknowledgements: This work was supported by an operating grant to A.J.C. from the Natural Sciences and Engineering Council of Canada. This is N.R.C.C. publication number 39550.

References

[1] Henrissat, B. and Bairoch, A. (1996) Biochem. J. 316, 695–696.
[2] Henrissat, B. and Bairoch, A. (1993) Biochem. J. 293, 781–788.
[3] Gilkes, N.R., Henrissat, B., Kilburn, D.G., Miller Jr., R.C. and Warrren, R.A.J. (1991) Microbiol. Rev. 55, 303–315.
[4] Clarke, A.J. (1996) In: Biodegradation of cellulose: enzymology and biotechnology, Technomoic Press, Lancaster, PA.
[5] Wang, Q., Tull, D., Meinke, A., Gilkes, N.R., Warren, R.A.J., Aebersold, R. and Withers, S.G. (1993) J. Biol. Chem. 268, 14096–14102.
[6] Saloheimo, M., Lehtovaara, P., Penttilä, M., Teeri, T.T., Ståhlber, J., Johansson, G., Pettersson, G., Claeyssens, M., Tomme, P. and Knowles, J.K.C. (1988) Gene 63, 11–21.
[7] Davies, G. and Henrissat, B. (1995) Structure 3, 853–859.
[8] McCarter, J.D. and Withers, S.G. (1993) Curr. Opin. Struct. Biol. 4, 885–892.
[9] Clarke, A.J. and Yaguchi, M. (1985) Eur. J. Biochem. 149, 233–238.
[10] Clarke, A.J. and Adams, L.S. (1987) Biochim. Biophys. Acta 916, 213–219.
[11] Clarke, A.J. (1988) Biochem. Cell Biol. 66, 871–879.
[12] Clarke, A.J. and Strating, H. (1989) Carbohydr. Res. 188, 245–250.
[13] Bray, M.R. and Clarke, A.J. (1994) Eur. J. Biochem. 219, 821–827.
[14] Baird, S.D., Hefford, M.A., Johnson, D.A., Sung, W.L., Yaguchi, M. and Seligy, V.L. (1990) Biochem. Biophys. Res. Commun. 169, 1035–1039.
[15] Eshdat, Y., McKelvy, J.F. and Sharon, N. (1973) J. Biol. Chem. 248, 5892–5898.
[16] Høj, P.D., Condron, R., Traeger, J.C., McAuliffe, J.C. and Stone, B.A. (1992) J. Biol. Chem. 267, 25059–25066.
[17] Macarron, R., van Beeumen, J., Henrissat, B., de la Mata, I. and Claeyssens, M. (1993) FEBS Lett. 316, 137–140.
[18] Domínguez, R., Souchon, H., Lascombe, M. and Alzari, P.M. (1996) J. Mol. Biol. 257, 1042–1051.
[19] Yaguchi, M., Roy, C., Rollin, C.F., Paice, M.G. and Jurasek, L. (1983) Biochem. Biophys. Res. Commun. 116, 408–411.