Volume 316, number 2, 137–140 FEBS 12002 January 1993

Identification of an essential glutamate residue in the active site of endoglucanase III from *Trichoderma reesei*

Ricardo Macarron[a], Jozef van Beeumen[b], Bernard Henrissat[c], Isabel de la Mata[a] and Marc Claeyssens[b]

[a]*Departamento de Bioquimica, Facultad de Ciencias Biologicas, Universidad Complutense, E-28040 Madrid, Spain,* [b]*Vakgroep Biochemie, Fysiologie en Microbiologie, Universiteit Gent, Faculteit Wetenschappen, B-9000 Gent, Belgium* and [c]*Centre de Recherches sur les Macromolécules Végétales, CNRS, F-38041 Grenoble, France*

Received 15 October 1992

n-Propyl, *n*-butyl and *n*-pentyl β-cellobiosides with a reactive ω-epoxide in their aglycon completely and irreversibly inactivate endoglucanase III from *Trichoderma reesei*. The pentyl derivative was found to be most effective. From these affinity labeling experiments evidence was found for the implication of Glu^{329} in the reaction mechanism. This is discussed in relation to other structural/functional data known for endoglucanase III and several other family A glycanases.

Affinity labeling; Cellulase; Endoglucanase III, *Trichoderma reesei*; Epoxy-alkyl cellobioside

1. INTRODUCTION

Polysaccharidases are usually classified as endo- and exo-type enzymes. In the case of cellulases this does not correspond to their classification based on amino acid sequence similarities in at least 6 different families (A–F), some containing both endo and exo-enzymes [1]. The validity of this classification has been confirmed by structural/functional evidence: (i) studies of the stereochemical course of the catalytic reaction [2,3], (ii) specificity mapping (Claeyssens and Henrissat, unpublished); (iii) site-directed mutagenesis, e.g. [4–6]; (iv) specific chemical modification [7], and [v] determination of three-dimensional structures of an exo-cellulase belonging to family B [8] and of an endo-enzyme from family E [9].

X-Ray crystallographic data are not available for other cellulase families, although endoglucanase A from *Clostridium cellulolyticum* has recently been crystallised (J.-P. Belaich and R. Haser, personal communication) and, hopefully, structural details of this family A cellulase will be published soon. More than 40 sequences now belong to this family including endoglucanase III from *Trichoderma reesei* [10].

The present paper describes an attempt to identify a catalytically important residue in this enzyme by affinity labeling using cellobioside derivatives carrying reactive epoxide functions in their aglycon part.

Correspondence address: M. Claeyssens, Vakgroep Biochemie-Fysiologie-Microbiologie, Universiteit Gent, Faculteit Wetenschappen, K.L. Ledeganckstraat 35, B-9000 Gent, Belgium. Fax: (32) (91) 64 53 42.

2. MATERIALS AND METHODS

Cellotriose and its 4′-methylumbelliferyl-β-glycoside were prepared as previously described [11]. The ω-epoxy-alkyl-β-cellobiosides (*n*-propyl, *n*-butyl and *n*-pentyl) were synthesized according to [12]. Protease V_8 was purchased from Sigma (USA). All other reagents and chemicals were of the highest purity available.

Endoglucanase III was isolated from culture supernatants of *Trichoderma reesei* QM 9414 and purified as described [13,14]. Its concentration was determined at 280 nm (ε=77000 $M^{-1} \cdot cm^{-1}$) [10].

Methylumbelliferyl-β-cellotrioside (100 μM) was used as fluorogenic substrate in 0.1 M sodium acetate, pH 5.0 at 30° C [11]. The activity against cellotriose (3 mM) was determined at different pH in Mc Ilvaine's buffers (0.1 M sodium citrate/phosphate) at 37°C, measuring D-glucose by the D-glucose oxidase/peroxidase method [15].

Endoglucanase III (0.8 μM, final concentration) in Mc Ilvaine's buffer pH 4.6, was treated with epoxide (2–20 mM) and incubated at 37°C. To stop the inactivation and to determine the residual activity after different reaction times, 10 μl aliquots of the reaction mixture were diluted into 0.5 ml 100 μM methylumbelliferyl-β-cellotrioside in 0.1 M sodium acetate pH 5.0.

Larger amounts of labeled endoglucanase III were prepared by adding 3 mg 4′, 5′-epoxypentyl-β-cellobioside to 1 mg enzyme in 1 ml Mc Ilvaine's buffer, pH 4.6. After complete inactivation (48 h, 37°C) the mixture was diluted with distilled water (to 3.5 ml), desalted and concentrated by ultrafiltration (Centricon 10, Amicon, USA). This washing step was repeated twice to remove excess epoxide. A control without inhibitor (unlabeled endoglucanase) was prepared similarly.

Samples (0.6 mg) of labeled or unlabeled enzyme in 275 μl 50 mM ammonium bicarbonate buffer, pH 7.8 (6 M guanidine hydrochloride) were reduced with 1.5 mM dithiothreitol (10 min, 50°C). After dilution (1:4) with buffer (free of denaturant) 30 μg *Staphylococcus aureus* Glu-C endoprotease V_8 was added and the mixtures were incubated for 10 h at 37°C. The resulting peptides were carboxymethylated (2 mM iodonacetamide, 15 min, 37°C) and analyzed by reverse-phase HPLC on a 0.46 × 25 cm C4-214 TP-54 column (Vydac, USA). A linear gradient of 5–70% acetonitrile in 0.1% trifluoroacetic acid (1 ml · min^{-1}) was used and eluates monitored from 190 to 320 nm using a photo diode array detector. All chromatographic equipment was from Waters Associates (USA). The peptides selected for further

analysis were concentrated by lyophilisation (approx. 30 μl) and stored at −20°C.

Amino-terminal sequence analysis was carried out on a 477A pulsed-liquid sequenator with on-line analysis of the PTH-amino acids on a 120A PTH-analyzer (Applied Biosystems, USA) as suggested by the manufacturer.

Plasma desorption mass spectrometry was performed on a BIO ION 20 Biopolymer Analyzer (Applied Biosystems, USA). Spectra were recorded for 1 million fission counts of the Californium^{252} source. Desorbed analyte ions were detected with a channel plate detector operated at −2 kV (for the first plate), the measurements being carried out in the 'positive mode'.

3. RESULTS

3.1. *Effect of the aglycon chain length on the inactivation of endoglucanase* III *by ω-epoxyalkyl β-cellobiosides*

Incubation of endoglucanase III (M) with large molar excess of ω-epoxyalkyl cellobiosides resulted in complete inactivation of the enzyme according to pseudo-first-order kinetics (not shown). No appreciable enzymatic hydrolysis of these compounds could be observed (HPLC) With equivalent concentrations of inhibitor (2.5 mM) the relative values of the inactivation rates were 10:20:100, respectively, for the propyl, butyl and pentyl derivatives.

3.2. *Kinetics, specificity and pH dependence of inactivation*

The pseudo-first-order inactivation rate constants (k_{app}) as a function of 4′, 5′-epoxypentyl-β-cellobioside concentration ([I]) (Fig. 1A) were plotted according to equation (1) [16]:

$$k_{app} = (k_i \cdot [I])/([I]+K_d) \quad (1)$$

K_d and k_i, respectively 30 nM and 0.035 min^{-1}, were obtained by hyperbolic fitting [17] (Fig. 1B). These values are similar to those found for other cellulases using the same compounds as irreversible inhibitors [18].

The order of the inactivation reaction, estimated according to [19], was (0.8 ± 0.2) (Fig. 1C).

The rate of inactivation with 4′, 5′-epoxypentyl-β-cellobioside was 3.5 times lower in the presence of 150 mM cellobiose, a competive inhibitor of the enzyme (K_i = 20 mM).

The pH dependence of the inactivation of endoglucanase III with 4′, 5′-epoxypentyl-β-cellobioside is similar to the pH-activity curve for cellotriose hydrolysis with a pronounced maximum around 5.0 [14].

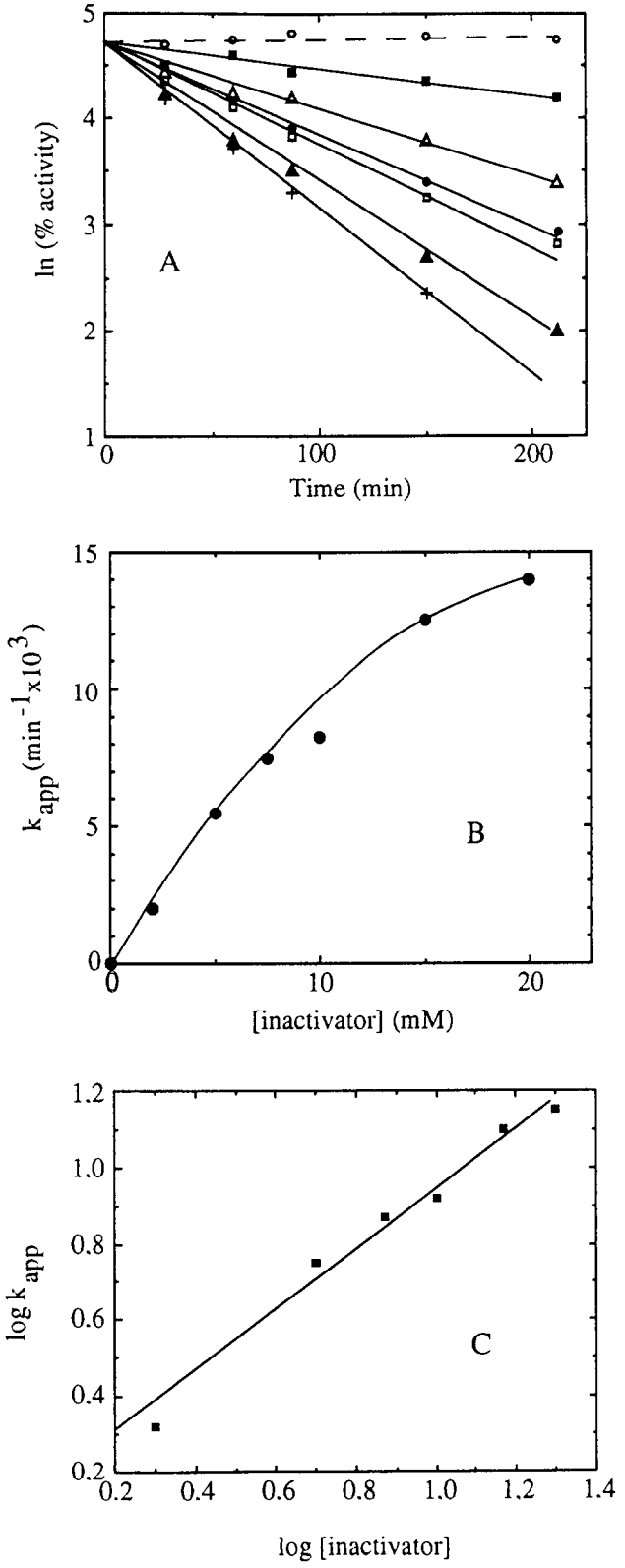

Fig. 1. Inactivation of endoglucanase III by 4′, 5′-epoxypentyl β-cellobioside (A) The enzyme (0.8 μM) in Mc Ilvaine's buffer pH 4.6 was treated at 37°C with inactivator: 2 mM (■), 5 mM (△), 7.5 mM (●), 10 mM (□), 15 mM (▲) and 20 mM (+); control (○) (no inhibitor). (B) k_{app} dependence on inactivator concentration. The solid line was drawn by hyperbolic fitting [17] of the experimental points (●) to Equation (1). The rates were determined from Fig. 1A. (C) Determination of apparant order of the inactivation reaction. The slope of the straight line is 0.8 ± 0.2.

Domains I	II	III	IV	V	Enzyme	Source	Swiss-Prot accession number
*	*	**	*	*			
EDGMTIFRLPVG	SLGAYCIVDIHN	QSRVWFGIMNEP	STTNLIFDVHKY	NRQAILTETGGG	EG III	*Trichoderma reesei*	P07982
DWESNMIRLAMY	ENDMYVIVDWHV	NPHIIYELANEP	DDHHTMYTVHFY	GVAVFATEWGTS	EG	*Bacillus* sp. 1139	P06564
DWGSNMIRLAMY	EHDMYVIVDWHV	NHYIIWELANEP	DAENIMYSVHFY	GVAVFATEWGTS	EG	*Bacillus* sp. KSM635	P19424
DWGITVFRAAMY	DLGIYVIIDWHI	YPNVIYEIANEP	DDPNVMYAFHFY	GAAIFVSEWGTS	EG A	*Bacillus* sp. N4	P06566
DWGITVFRAAMY	DLGIYVIIDWHI	YPNVIYEIANEP	TDPNVMYAFHFY	GAAIFVSEWGTS	EG B	*Bacillus* sp. N4	P06565
DWGSNVIRLALY	ENDMYVIIDWHV	DPHLIYELANEP	DDHHTMYTLHFY	GKAIFATEWGVS	EG C	*Bacillus* sp. N4	P19570
KEGYNLIRLPYS	QRGIQIILDRHR	PTVIGADLHNEP	VPNRVVYSPHDY	IAPVLVGEFGGR	EG	*Bacillus polymyxa*	P23548
DWGITVFRAAMY	ELGIYVIIDWHI	TPNVIYEIANEP	KDANVMYALHFY	GAPIFVTEWGTS	EG	*Bacillus subtilis*	P07983
TWGINVIRLAMY	DNDMYVIIDWHT	NENVIYEICNEP	DFDNIMYTYHFY	GLPVFISEYGLV	EG A	*Butyrivibrio fibrisolvens*	P22541
KAGFNTIRIPVS	VNDMYVILNSHH	DYHLVFETMNEP	ASGRLILSVHAY	NIPVVVGETSAT	EG 1	*Butyrivibrio fibrisolvens*	P20847
SWGMNSVRVVLS	LGFKAIILEVHD	EDFVIINIGNEP	PLRNLVFSIHMY	GLPLVIGEFGHQ	BMAN	*Caldocellum saccharolyticum*	P22533
NRGFNLLRVPIS	EVGLKIMLDIHS	DTIIAFDLKNEP	YQNKVVYSPHDY	IAPLLIGEWGGH	EG	*Caldocellum saccharolyticum*	P10474
KWGVNVIRAAMY	DLNMYVIIDWHI	YSNVIYEICNEP	RYSNIMYTCHFY	GIAIFVTEWGTS	EG	*Clostridium acetobutylicum*	P15704
QKGFNTVRIPVS	DNKMYVILNTHH	DEHLIFEGMNEP	NNNKIIVSVHAY	GIPVIIGECGAV	EG A	*Clostridium cellulolyticum*	P17901
QAGFKHVRIPIR	SRGFVTVINSHH	SENLVFEILNEP	NDPNLIATFHYY	NIPVYLGEYGVM	EG D	*Clostridium cellulolyticum*	P25452
DKGINVVRMPIA	RVGIKVILDVHS	DTIIGFDLKNEP	YQSQLVYSPHDY	ISPLLLGEWGGM	EG B	*Clostridium thermocellum*	P04956
EAGFDHVRLPFD	KYNLGLVLDMHH	REHIAFELLNEV	DDDYIVYNFHFY	KCKLYCGEFGVI	EG C	*Clostridium thermocellum*	P07985
EMGFNAVRVPVT	DCGMYAIINLHH	DDHLLFETMNEP	NDSRVIVSIHAY	GRAVIIGEFGTI	EG E	*Clostridium thermocellum*	P10477
AAGYKNVRIPVR	SRGFVTIINSHH	SENLLFEIMNEP	DDPYLIGTFHYY	NIPVYFGEFAVM	EG H	*Clostridium thermocellum*	P16218
DWKSSIVRAAMG	ANDMYAIIGWHS	KPNVIYEIYNEP	NAKNIAYTLHFY	GIALFVTEWGTV	EG Z	*Erwinia chrysanthemi*	P07103
DNGFKSLRLPID	KYNMSFVIDYHE	REDLFFELLNEP	TDDNIIYVIHTY	NVPVIINEFGAL	EG 3	*Fibrobacter succinogenes*	P14250
QQGITMLRLPIA	QNDIQIFIDIHS	SNLIGIDVFNEP	PKDRLLFSPHTY	GYGILIGEFGGN	CD C	*Pseudomonas fluorescens*	P27033
NKGMNLVRLPFR	ATGQTVLLDPHN	NPRVILGLMNEP	PGHNLVFEVHQY	GYRGFLGEFGAA	EG	*Pseudomonas solanacearum*	P17974
NKGFDVIRIPVT	DDGAYVIINSHH	GDHLIFEGLNEP	EDDHIGFSIHAY	DIPVIITEYGAV	EG A	*Ruminococcus albus*	P23660
EAGFNVLRIPVS	DDGMYVILNTHH	DEHLIFEGLNEP	DSDKLIISVHAY	DIPVIVGEFGSM	EG B	*Ruminococcus albus*	P23661
EAGFNVLRIPVS	DNGLYVILNTHH	DEHLIFEGLNEP	DSDKLIISVHAY	GIPVIVGEFGTM	EG 1	*Ruminococcus albus*	P16216
SQGFNLVRIPIG	NNSLKVWVDLHG	DTVIGIELINEP	GYWGVTIDHHHY	EHIKVACEWGTG	XBGL	*Saccharomyces cerevisiae*	P23776
GLGFNAVRLPFC	ARGMYVLLDHHT	PYVLGLDLKNEP	PANRLLLAPHVY	THALLLGEFGGK	EG A	*Xanthomonas campestris*	P19487

Fig. 2. Alignment of family A cellulases. The five well-conserved domains that contain invariant residues (*) are represented for 28 glycanases of family A and correspond to hydrophobic clusters as defined in [1]. Enzyme sources, SwissProt accession numbers and current abbreviations are given: EG, endoglucanase (E.C. 3.2.1.4.); BMAN, β-mannanase (E.C. 3.2.1.78); CD, cellodextrinase (E.C. 3.2.1.74); XBGL, exo-β(1,3)-glucanase (E.C. 3.2.1.58).

3.3. *Isolation and partial amino acid sequence of an affinity-labeled peptide*

Intact and affinity-labeled endoglucanase III were digested with protease V_8 and the resulting peptide mixtures analyzed by reverse-phase HPLC (not shown). A peptide, clearly absent in the modified protein, was analysed to have the N-terminal sequence Cys-Thr-Thr-Asn-Ile- and to weigh 3207.3 Da. It situates the putative modification in the sequence [10] at either Glu^{301} or Glu^{329} (M_r for peptide 302–329:3207.4 Da).

4. DISCUSSION

Substrates or substrate analogs carrying reactive epoxide functions are often used as affinity labels for carbohydrases [12,18]. These derivatives are thought to react primarily with the ionised form of aspartate or glutamate residues at the enzymes active site. In a few cases, however, reactions with other residues have been observed, e.g. methionyl [20], cysteinyl [21] or aminated groups [22].

No free cysteine residues could be detected in endoglucanase III and chemical modifications of either histidine or arginine groups did not affect the enzymic activity (Macarron, unpublished observations). In two other cellulases from family A (endoglucanase A from *Clostridium cellulolyticum* [6] and endocellulase Z from *Erwinia chrysanthemi* [5]) mutagenesis experiments directed at the conserved histidine in domain II and arginine in domain I (Fig. 2) prove that these residues are structurally but not functionally important.

The completely irreversible inactivation of endoglucanase III from *Trichoderma reesei* by ω-epoxyalkyl-β-cellobiosides with different aglycon chain lengths, as observed in the present study, probably results from active-site labeling at a carboxylate group. This is suggested by the specificity requirement in the structure of the epoxide, the protection by substrates or substrate analogs, the dependence of the inactivation rate on pH and the stoichiometry of the reaction (Fig. 1).

Among the inhibitors tested, 4′,5′-epoxypentyl-β-cellobioside was the most reactive. The pentyl aglycon carbon chain could fit into a glycopyranosyl binding subsite, better than the aglycon moieties of the propyl or butyl derivatives. Similar results have been reported for cellulases from an *Oxysporus* species, *Aspergillus niger* and *Aspergillus wentii* [18], and most significantly for an endoglucanase from *Schizophyllum commune* [12] which, like endoglucanase III, also belongs to family A [1]. Another endoglucanase from family A *Bacillus subtilis* endoglucanase, however, is inactivated faster by the butyl than by the pentyl derivative [12].

The identity of the residue modified in endoglucanase III can be deduced from data obtained by peptide analysis and sequence comparisons with family A cellulases (Fig. 2). V_8 proteolysis of unmodified and inactivated EG III was used to trace possible peptides carrying a labeled carboxyl group. A peptide (M_r=3207.3, clearly absent in proteolysates of the modified protein, was sequenced N-terminally and identified as Cys^{302}–Glu^{329}, suggesting that either Glu^{301} or Glu^{329} is a possible target of the epoxide.

Only two Glu residues (and no Asp residues) are totally invariant in cellulase family A (Fig. 2). Since Glu^{301} in the identified peptide is not conserved, most probably Glu^{329} in domain V acts as nucleophile mediating the formation of the glycosyl–enzyme intermediate. It is also the most likely target of the epoxide.

On the other hand, Glu^{218} in domain III could be involved in the protonation of the glycosidic bond. Site directed mutagenesis experiments have indeed confirmed the essential role (general acid–base catalysis) played by the corresponding conserved Glu residues in other cellulases from family A [4,5]. Since endoglucanase III is a 'retaining' enzyme, as is typical for a family A cellulase [3], one carboxylate/carboxylic acid group should be positioned 'above' and another 'below' the hydrolysed glycosidic bond to account for a double displacement mechanism such as that found in lysozyme and β-galactosidase [23]. It is interesting to note that in the latter case a covalent intermediate of a slowly reacting substrate could be isolated [24] and as a consequence the role of the putative nucleophile Glu^{461}, identified by affinity labeling with an epoxide, has been questioned. The possibility that carboxylate/carboxylic acid groups can alternatively act as nucleophiles or as proton donors, depending on the nature of the substrate or affinity label used, cannot be excluded and could indeed reflect the catalytic flexibility of glycosylases [25].

Acknowledgements: R.M. and I.M. were recipients of a fellowship from the Spanish Ministerio de Educación y Ciencia. M. Obreno and J. Vermeulen are thanked for technical assistance. R.M. is indebted to Dr. C. Acebal and Dr. M.P. Castillón for encouragement. M.C. is indebted to the Région Rhône-Alpes (France) for a fellowship and J.V.B. acknowledges the Fund for Joint Basic Research (Grant 32.001.91). Dr. Peter Hoy (Le Trobe, Australia) is thanked for helpful suggestions.

REFERENCES

[1] Henrissat, B., Claeyssens, M., Tomme, P., Lemesle, L. and Mornon, J.-P. (1989) Gene 81, 83–95.
[2] Barras, F., Bortoli-German, I., Bauzan, M., Rouvier, J., Gey, C., Heyraud, A. and Henrissat, B. (1992) FEBS Lett. 300, 145–148.
[3] Gebler, J., Gilkes, N., Claeyssens, M., Wilson, D.B., Béguin, P., Wakarchuk, W.W., Kilburn, D.G., Miller Jr., R.C., Warren, R.A.J. and Whiters, S.G. (1992) J. Biol. Chem. 267, 12559–12561.
[4] 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.
[5] Py, B., Bortoli-German, I., Haiech, J., Chippaux, M. and Barras, F. (1991) Prot. Eng. 4, 325–333.
[6] Belaich, A., Fierobe, H.-P., Baty, D., Busetta, B., Bagnara-Tardif, C., Gaudin, C. and Belaich, J.-P. (1992) J. Bacteriol. 174, 4677–4682.
[7] Tomme, P. and Claeyssens, M. (1989) FEBS Lett. 243, 239–243.
[8] Rouvinen, J., Bergfors, T., Teeri, T., Knowles, J.K.C. and Jones, T.A. (1990) Science 249, 380–386.
[9] Juy, M., Amit, A.G., Alzari, P.M., Poljak, R.J., Claeyssens, M., Béguin, P. and Aubert, J.-P. (1992) Nature 557, 89-91.
[10] Saloheimo, M., Lehtovaara, P., Penttilä, M., Teeri, T.T., Ståhlberg, J., Johansson, G., Pettersson, G., Claeyssens, M., Tomme, P. and Knowles, J.K.C. (1988) Gene 63, 11–21.
[11] Claeyssens, M. (1989) FEMS Symp. 43, 393–397.
[12] Høj, P.B., Rodríguez, E.B., Stick, R.V. and Stone, B.A. (1989) J. Biol. Chem. 264, 4939–4947
[13] Acebal, C., Castillón, M.P., Estrada, P., Mata, I., Costa, E., Aguado, J., Romero, D. and Jiménez, F. (1986) Appl. Microbiol. Biotechnol. 24, 218–223.
[14] Macarrón, R., Acebal, C., Castillón, M.P., Domínguez, J.M., de la Mata, I., Pettersson, G., Tomme, P. and Claeyssens, M. (1992) Biochem. J. (in press).
[15] Bergmeyer, H.U. and Bernt, E. (1974) in: Methoden der Enzymatischen Analyse (Bergmeyer, H.U. Ed.) vol. 2, pp. 1257–1259, Verlag Chemie, Weinheim.
[16] Kitz, R. and Wilson, I.B. (1962) J. Biol. Chem. 237, 3245–3249.
[17] Sakoda, M. and Hiromi, K. (1976) J. Biochem. (Tokyo) 80, 1003–1014.
[18] Legler, G. (1990) Adv. Carbohydr. Chem. Biochem. 48, 319–384.
[19] Levy, H.M., Leber, P.D. and Ryan, E.M. (1963) J. Biol. Chem. 238, 3654–3659.
[20] Brown, J.R. and Hartley, B.S. (1964) Proceedings of the 1st FEBS meeting, Academic Press, London, Abstract A29, p. 25.
[21] Bessell, E.M. (1973) Chem. Biol. Interact. 7, 343–353.
[22] Fersht, A. (1985) Enzyme Structure and Mechanism, W.H. Freeman & Co., New York, pp. 248–262.
[23] Sinnott, M.L. (1990) Chem. Rev. 90, 1171–1202.
[24] Gebler, J.C., Aebersold, R. and Withers, S.G. (1992) J. Biol. Chem. 267, 11126–11130.
[25] Kanda, T., Brewer, C.F., Okada, G. and Hehre, E.J. (1986) Biochemistry 25, 1159–1165.