Identification of essential active-site residues in the cyanogenic β-glucosidase (linamarase) from cassava (*Manihot esculenta* Crantz) by site-directed mutagenesis

Zsolt KERESZTESSY[1], Kate BROWN, M. Alison DUNN and Monica A. HUGHES[2]

School of Biochemistry and Genetics, University of Newcastle upon Tyne, Newcastle upon Tyne NE2 4HH, U.K.

The coding sequence of the mature cyanogenic β-glucosidase (β-glucoside glucohydrolase, EC 3.2.1.21; linamarase) was cloned into the vector pYX243 modified to contain the SUC2 yeast secretion signal sequence and expressed in *Saccharomyces cerevisiae*. The recombinant enzyme is active, glycosylated and showed similar stability to the plant protein. Michaelis constants for hydrolysis of the natural substrate, linamarin (K_m = 1.06 mM) and the synthetic *p*-nitrophenyl β-D-glucopyranoside (PNP-Glc; K_m = 0.36 mM), as well as apparent pK_a values of the free enzyme and the enzyme–substrate complexes (p$K^E{}_1$ = 4.4–4.8, p$K^E{}_2$ = 6.7–7.2, p$K^{ES}{}_1$ = 3.9–4.4, p$K^{ES}{}_2$ = 8.3) were very similar to those of the plant enzyme. Site-directed mutagenesis was carried out to study the function of active-site residues based on a homology model generated for the enzyme using the MODELLER program. Changing Glu-413 to Gly destroyed enzyme activity, consistent with it being the catalytic nucleophile. The Gln-339Glu mutation also abolished activity, confirming a function in positioning the catalytic diad. The Ala-201Val mutation shifted the pK_a of the acid/base catalyst Glu-198 from 7.22 to 7.44, reflecting a change in its hydrophobic environment. A Phe-269Asn change increased K_m for linamarin hydrolysis 16-fold (16.1 mM) and that for PNP-Glc only 2.5-fold (0.84 mM), demonstrating that Phe-269 contributes to the cyanogenic specificity of the cassava β-glucosidase.

Key words: catalytic nucleophile, cyanogenesis, linamarin binding, molecular modelling, pH dependence.

INTRODUCTION

β-Glucosidases are widespread in nature and have a variety of functions in important physiological processes. They are involved in plant cyanogenesis, a defence mechanism against small herbivores [1] that is found in more than 3000 species, including important crops such as cassava (*Manihot esculenta* Crantz). Injury of cyanogenic plants results in HCN release as a result of the catabolic breakdown of stored cyanogenic glucosides [2]. In cassava 95% of the cyanogenic glucoside content is 1-cyano-1-methylethyl β-D-glucopyranoside, linamarin, which is hydrolysed by a cyanogenic β-glucosidase called linamarase [3], producing glucose plus acetone cyanohydrin. An α-hydroxynitrile lyase [4] liberates acetone and HCN from the cyanohydrin compound. Intact cyanogenic glucosides, such as linamarin, are non-toxic and excreted unchanged in the urine [5], but they can be activated to release the powerful metabolic poison, HCN, by linamarase. It has been shown recently that this pro-drug approach is successful in trials *in vivo* to eradicate rat intracerebral gliomas, using retrovirally delivered recombinant cassava linamarase [6]. Other approaches involve targeting a cyanogenic β-glucosidase to cancer cells by conjugating the enzyme to tumour-specific antibodies [7].

β-Glucosidases catalyse the hydrolysis of β-glucosidic linkages of conjugated monoglycosides, and most of them belong to Family-1 glycoside hydrolases [8]. They use a common double-displacement mechanism, involving two active-site carboxylates, to release their saccharide product with net retention of configuration of the sugar ring on the newly exposed anomeric centre [9]. An acid/base catalytic group exists for protonation of the glycosidic oxygen to release the aglycone. A nucleophile catalyst stabilizes an oxocarbenium-ion-like transition state, which collapses into a covalent glycosyl-enzyme intermediate [10,11].

The three-dimensional structures of six Family-1 enzymes have been solved to date; these all have a common 8-fold β/α barrel with the two catalytic groups located on β-strands 4 and 7 [12]. The first structure solved was a linamarase from white clover [13] and this shows the highest sequence similarity to the cassava enzyme [3,14]. The other structures are *Sinapis alba* myrosinase, which is a thio-β-glucosidase [15], 6-phospho-β-galactosidase from *Lactobacillus lactis* [16], β-glucosidase A from *Bacillus polymyxa* [17], β-glycosidase from *Thermosphaera aggregans* [18] and β-glucosidase from *Bacillus circulans* [19]. Co-crystallization of these enzymes with substrates and inhibitors confirms the acid/base and nucleophile catalytic residues as glutamates in the highly conserved LNEP and I/VTENG motifs [8], identified from affinity labelling and HPLC and/or electrospray ionization-MS peptide-mapping studies using *N*-bromoacetyl-β-D-glucopyranosylamine [20] or 2-desoxy-2-fluoro-β-glucopyranosyl derivatives [21].

Cassava linamarase has been purified to homogeneity from the plant and shown to be a 70000 Da glycoprotein with high-mannose-type oligosaccharide moieties estimated to be 7–15% of the molecular mass [3]. The gene has been cloned and sequenced as cDNA [3]. Reaction kinetics [22] and radioaffinity labelling studies using *N*-bromoacetyl-β-D-[U-^{14}C]-glucopyranosylamine showed that Glu-198 acts as the acid/base catalyst [20]. The coding sequence of the enzyme has been expressed in *Escherichia coli* using the pGEX-2T glutathione S-tranferase

Abbreviations used: YSM, yeast selective medium; PNP-Glc, *p*-nitrophenyl β-D-glucopyranoside; anti-CCBG, anti-cassava cyanogenic β-glucosidase.

[1] Present address: Department of Biochemistry and Molecular Biology, Medical and Health Sciences Center, University of Debrecen, Nagyerdei krt. 98, Debrecen H-4012, Hungary.

[2] To whom correspondence should be addressed (e-mail m.a.hughes@ncl.ac.uk).

fusion vector and the non-glycosylated recombinant protein showed impaired folding and stability [23]. Further structure/function studies of the enzyme and its improvement for therapeutic applications necessitates the production of a glycosylated recombinant enzyme. The present paper reports the production of a catalytically active glycosylated cassava linamarase in *Saccharomyces cerevisiae* and site-directed mutagenesis studies of the function of the active-site residues Glu-413, Gln-339, Ala-201 and Phe-269 in direct catalysis, formation of the micro-environment of the catalytic diad and substrate aglycone binding.

MATERIALS AND METHODS

Chemicals

Unless stated otherwise, all chemicals were purchased from Merck (Poole, Dorset, U.K.). All synthetic oligonucleotides were made by MWG-Biotech UK.

Strains, cloning vectors and plasmids

E. coli strain JM 105 was used for the propagation of all plasmids. The 'supersecretory' mutant *S. cerevisiae* strain, CGY 2998 (*MATa ura3-52 leu2-3,112 ssc1-1 gal*$^-$; Genome Therapeutics, Waltham, MA, U.S.A.) [24], was used as a host for β-glucosidase expression. The *E. coli* cloning vector pGEM5Zf was purchased from Promega (Madison, WI, U.S.A.). The *S. cerevisiae* 2 micron high-copy number expression vector, pYX243, contains sequences required for replication in both *E. coli* and yeast, the ampicillin-resistance gene, the yeast *leu2* gene and the yeast GAL1 promoter (R & D Systems Europe, Abingdon, Oxfordshire, UK).

PCR

The oligonucleotide primers CASY1 and CASY2 (Table 1) correspond to bases 56–84 (sense) and 1601–1630 (antisense) respectively of the cassava cyanogenic β-glucosidase cDNA clone pCAS5 [3], modified to contain 5′ *Nco*I and 3′ *Sal*I restriction-endonuclease sites to facilitate directional cloning in-frame into the pYX243 yeast expression vector. The primers were used for PCR amplification of pCAS5 performed in 100 μl aliquots under the following conditions: 10 pmol of each primer/10 mM Tris/HCl, pH 9.0/50 mM KCl/2.5 mM $MgCl_2$/0.1 % Triton X-100/200 μM each of dATP, dGTP, dCTP and dTTP (Amersham Pharmacia Biotech)/2 units of *Taq* DNA polymerase (Promega)/25 μg of pCAS5. Amplifications were performed in a Perkin–Elmer Cetus thermal cycler (Norwalk, CT, U.S.A.) with initial template denaturation at 95 °C for 5 min, followed by 30 amplification cycles of 1 min at 72 °C, denaturation at 95 °C, 2 min annealing at 55 °C, a 2 min extension at 72 °C and a single 10 min extension at 72 °C.

Cloning and construction of the yeast expression vector

The PCR product, CAS5TR (truncated version of pCAS5), which lacks the coding sequence of the predicted N-terminal signal sequence of the cyanogenic β-glucosidase, was digested simultaneously with *Nco*I and *Sal*I and cloned into the corresponding sites of the pGEM5Zf plasmid. The whole coding sequence of the resulting pCAS5TR was sequenced using the Sequenase version 2.0 (Amersham Life Sciences). Two complementary oligonucleotides, SUC2S (24.6 ng) and SUC2A (25 ng; Table 1A), corresponding to the coding sequence of the yeast invertase N-terminal secretion signal peptide [25], were annealed in 50 mM NaCl by 5 min of boiling and slow cooling to room temperature. The resulting double-stranded SUC2 sequence, which contained engineered 5′ and 3′ *Nco*I sticky ends, was cloned (in frame) into the *Nco*I site of the pYX243 yeast expression vector. The oligonucleotides were designed so that this resulted in the loss of the 5′ SUC2 *Nco*I site in pYXSUC2. The orientation of SUC2 in pYX243 was confirmed by nucleotide sequencing using the oligonucleotide primer YX1 (Table 1A). The CAS5TR sequence was isolated from pCAS5TR by digestion with *Nco*I and *Sal*I, and cloned into the SUC2 3′ *Nco*I and *Sal*I sites of pYXSUC2. The presence of SUC2 and CAS5TR in the pYXSUC2CAS5TR construct was confirmed by both PCR using SUC2S (sense) and CASY2 (antisense) primers and nucleotide sequencing.

Site-directed mutagenesis

Mutations were introduced into the β-glucosidase coding sequence in pCAS5TR using the QuikChange site-directed muta-

Table 1 Synthetic oligonucleotides

(A) Oligonucleotides used in the construction and sequencing of pYX243Suc2CAS5TR. For CASY1 and 2, nucleotide numbering is based on the published pCAS5 sequence [4]; sequence differences from pCAS5 are shown in bold. For YX1, nucleotide numbering is based on the pYX243 yeast vector sequence (R & D Systems). For SUC2S and SUC2A, oligonucleotides were designed on the basis of the yeast invertase signal peptide SUC2 [25]; extra bases added are shown in italic. (B) Mutations generated by site-directed mutagenesis. Only sense oligonucleotides are shown from the mutagenic oligonucleotide pairs. Nucleotide numbering is based on the published pCAS5 sequence [4]; sequence differences from pCAS5 are shown in bold.

(A)

Oligonucleotide sequence, 5′ → 3′	Location of oligonucleotide	Name of oligonucleotide
CAC TAG GCC AG**C CAT GG**G AAC TGA TG	56–84 (pCAS5)	CASY1 (sense)
CCA GAC ATA A**GT CGA C**TC ACA TAG AAT TTG	1601–1630 (pCAS5)	CASY2 (antisense)
TAA TCA GCG AAG CGA TG	8438–8454 (pYX243)	YX1 (sense)
*C*AT GCT TTT GCA AGC TTT CCT TTT CCT TTT GGC TGG TTT TGC AGC CAA AAT ATC TGC	*S. cerevisiae* invertase signal peptide	SUC2S (sense)
*CAT G*GC AGA TAT TTT GGC TGC AAA ACC AGC CAA AAG GAA AAG GAA AGC TTG CAA AAG	*S. cerevisiae* invertase signal peptide	SUC2A (antisense)

(B)

Sequence of mutagenic oligonucleotide, 5′ → 3′	Location of oligonucleotide on pCAS5 cDNA	Amino acid substitution
CCA GTC ATT TAC GTT ACT G**GT** AAT GGG GTT GAC AAC	1242–1277	Glu-413Gly
G ATT GGG ATT ACC CTC **AA**T ACC TTC TGG TAT GAA CC	818–843	Phe-269Asn
G ACT TTT AAT GAA CCA TCA G**TT** TAT GTT GGA TTT GCC CAT G	605–645	Ala-201Val
GAT TTT GTT GGA TTA **G**AA TAC TAC ACT GCA TAT TAT GCA G	1023–1062	Gln-339Glu

genesis kit (Stratagene Europe, Amsterdam Zuidoost, The Netherlands) according to the manufacturer's recommendations. Briefly, two completely overlapping complementary oligonucleotides (HPLC-purified) for each mutation were synthesized to contain the corresponding nucleotide changes (Table 1B). Each primer (125 ng) was annealed to pCAS5TR (6.5 ng) and both strands of the plasmid amplified using *Pfu* polymerase in a linear extension reaction carried out in a PCR reactor (Perkin-Elmer GeneAmp 2400) under the following conditions: 30 s denaturation at 95 °C, and 12 or 16 cycles of 30 s denaturation at 95 °C for single or double nucleotide changes respectively, 1 min annealing at 55 °C and 12 min extension at 68 °C. The reactions were cooled to room temperature and treated with *Dpn*I restriction enzyme for 1 h at 37 °C to cut the methylated parental plasmid strands. *Dpn*I-treated DNA (1 μl) was transformed into 50 μl of *Epicurian coli* XLmutS supercompetent cells (Stratagene) and the transformed bacteria were grown in 0.5 ml of NZY Plus broth for 1 h at 37 °C, followed by overnight culture on Luria-broth medium plates, containing 50 μg/ml ampicillin, to select for transformants. Mutations introduced into the β-glucosidase coding sequence were confirmed by nucleotide sequencing and the mutant CAS5TR sequences were cloned into pYXSUC2 as described above.

Yeast transformation

Electro-transformation of *S. cerevisiae* CGY2998 cells was carried out with the Bio-Rad Gene Pulser™ system (Bio-Rad, Richmond, CA, U.S.A.) according to the manufacturer's instructions, using 0.1 μg of plasmid DNA and yeast selective medium (YSM) plates, which contain 0.67 % (w/v) yeast nitrogen base (YNB) without amino acids (Difco, Detroit, MI, U.S.A.), 2 % (w/v) dextrose, 6.7 % (v/v) 'leucine drop-out amino acid solution' (as described in the pYX vector users' guide, R & D Systems), 2 % Bacto™-Agar (Difco, Becton Dickinson, Oxford, U.K.) and 1 M sorbitol.

Expression and purification of the wild-type and mutated cassava cyanogenic β-glucosidase

Wild-type and mutant β-glucosidase genes were expressed in the *S. cerevisiae* CGY2998 supersecretory strain [24]. For induction of the recombinant proteins, a culture was grown for 48 h at 30 °C in liquid YSM, containing 2 % (w/v) dextrose, with shaking at 200 rev./min. The culture was diluted 1:100 into YSM and grown for an additional 48 h. Cells were collected by centrifugation at 5000 ***g*** for 1 min, washed with YSM containing no sugar, resuspended in an equal volume of YSM containing 2 % (w/w) galactose and 2 % (w/v) raffinose and grown for a further 96 h, before harvesting the cells by centrifugation at 1500 ***g*** for 5 min. Cells were washed with ice-cold water and diluted four times with ice-cold yeast glass-bead-disruption buffer (YGBDB) containing 20 mM Tris/HCl, pH 7.9, 10 mM $MgCl_2$, 1 mM EDTA, 5 % (v/v) glycerol, 0.3 M ammonium sulphate, 1 mM dithiothreitol and 1 mM PMSF. An equal volume of pre-chilled 425–600 μm acid-washed glass beads (Sigma–Aldrich, Poole, Dorset, U.K.) was added and the cells were disrupted by vortexing for 7 min at 4 °C. The supernatant was collected, the glass beads were washed with an equal volume of YGBDB, and then combined supernatants were centrifuged at 20000 ***g*** for 1 h at 4 °C to obtain the crude yeast-protein extract. The clear protein extract was precipitated twice with ammonium sulphate (70 % and then 45 %, w/v), and the precipitated protein was collected by centrifugation (20000 ***g***, 20 min at 4 °C) and resuspended in 0.2 M Tris/maleic acid/NaOH buffer, pH 5.6. Aliquots (10 ml) were applied to a 2.5 cm × 100 cm Sephacryl S-300 size-exclusion column (Amersham Pharmacia Biotech) pre-equilibrated with 0.2 M Tris/maleic acid buffer. Fractions containing β-glucosidase activity were pooled and concentrated on a 25 ml-capacity 30 kDa molecular-mass cut-off KwiqSpin centrifugal concentrator (Pierce & Warriner, Chester, U.K.) and the buffer was changed to 0.1 M Mops/HCl, pH 6.5. Concentrated protein samples were applied to a POROS PI (polyethyleneimine, 4.6 mm × 100 mm, 1.6 ml) FPLC anion-exchange column (Amersham Pharmacia Biotech) and eluted with a gradient of 1 M NaCl in 0.1 M Mops/HCl, pH 6.5, from 0 to 100 % in 35 ml, at a flow rate of 2.5 ml/min. Fractions containing β-glucosidase activity were pooled and the buffer was changed to 0.2 M Tris/maleic acid/NaOH, pH 5.6.

Enzyme assays and kinetic measurements

β-Glucosidase activity of protein samples against the natural substrate, linamarin (linamarase), was estimated as described previously [22]. Enzyme reactions were carried out in a 0.5 ml final volume in 0.15 M citrate/phosphate buffer, pH 6.0, at 37 °C. Reactions were initiated by the addition of small aliquots (10 μl) of the protein solutions and were stopped by cooling on ice and addition of NaOH to a final concentration of 0.1 M. The liberated HCN was quantified with the pyrazolon/bispyrazolon reagent and spectrophotometry at 620 nm [22]. *p*-Nitrophenyl-β-D-glucopyranosidase activity was measured using the same buffer as that described above, in a 1.0 ml final reaction volume. The liberated *p*-nitrophenol was quantified by spectrophotometry at 400 nm under basic conditions [22].

Cellobiase activity was determined by measuring the amount of glucose released from D-cellobiose using the method of Leary et al. [26]. Reaction mixtures and incubation times were identical with those used for β-glucosidase measurements, except that 1–50 mM D-cellobiose was used as the substrate. After stopping the reactions, aliquots of 25 μl were withdrawn and mixed with 3 ml of reagent solution containing 0.76 mM 4-aminoantipyrine, 11 mM phenol, 4 kU/l glucose oxidase and 18 kU/l peroxidase in 0.1 M phosphate buffer, pH 6.6 (where U = international units). The absorbance at 500 nm was monitored continuously for 3 min.

Reaction kinetic parameters (k_{cat}, K_m, k_{cat}/K_m) were determined using five or six different substrate concentrations ranging between 0.5- and 20-fold of the K_m value. The values of the catalytic constant (k_{cat}) and the Michaelis constant (K_m) were computed using the classic Michaelis–Menten rate equation. In pH-dependence studies of the kinetic parameters, eight different pH values were used (pH 3.5–8.2) in 0.15 M citrate/phosphate buffer. Theoretical pH profiles were non-linearly fitted to the data points using the diprotic enzyme model [22]. Calculation of kinetic constants and theoretical curves was carried out using the program Grafit [27].

SDS/PAGE and Western-blot analysis

Electrophoresis was performed on 10 % (w/v) polyacrylamide gels as described previously [23]. After electrophoresis, proteins were stained in a solution of 0.05 % (w/v) Coomassie Brilliant Blue R (Sigma-Aldrich) prepared in 25 % (v/v) propan-2-ol and 10 % (v/v) glacial acetic acid, or transferred to Protran nitrocellulose membrane (Schleicher and Schuell, Dassel, Germany) for Western-blot analysis. Western blotting was carried out using anti-CCBG (anti-cassava cyanogenic β-glucosidase) antiserum, isolated as described elsewhere [28], from rabbits immunized with cassava linamarase purified from plant leaves [3], and peroxidase-labelled anti-rabbit antibodies (IgG + IgM) both in 1:1000 dilutions. Immunoreactive protein bands were visualized

using the ECL Western-blotting analysis system (Amersham Pharmacia Biotech) according to the manufacturer's instructions.

RESULTS AND DISCUSSION

Expression of wild-type cassava cyanogenic β-glucosidase gene, CAS5, in yeast

The sequence encoding the plant cyanogenic β-glucosidase (CAS5TR) was amplified using the PCR to introduce an *Nco*I restriction site after the coding sequence of the predicted N-terminal signal peptide, and a *Sal*I restriction site at a position corresponding to the C-terminal methionine. The 1.6 kb PCR product (CAS5TR) was cloned into the *E. coli* vector pGEM5Zf and checked by sequencing. Subsequently, the CAS5TR sequence was subcloned into the *Nco*I and *Sal*I sites of pYX243, which had been modified to contain the yeast invertase N-terminal signal peptide coding sequence, SUC2 [25], to produce a 10.2 kb novel expression/secretion vector, pYXSUC2CAS5TR. The construct was electro-transformed into the supersecretory *S. cerevisiae* strain, CGY2998, which contains a mutation in the *ssc1* gene and has been shown to be capable of secreting 20 times more heterologous protein than a wild-type yeast strain [24]. Induction with 2 % (w/v) galactose of transformants harbouring the pYXSUC2CAS5TR plasmid resulted in the appearance of cyanogenic β-glucosidase activity, which was undetectable in the induced host strain lacking the plasmid. Typically, cells produced 1.2 mg of β-glucosidase/litre of culture (100 μg/g of wet cell paste). Of the total linamarase activity, 90 % was associated with the cells and only 10 % was detected in the culture medium. However, intact induced transformant cells were capable of hydrolysing linamarin when assayed with continuous shaking under the same conditions as solute samples. Recombinant β-glucosidase was purified from crude protein extracts, generated by glass-bead disruption of induced yeast cells, and by size-exclusion and FPLC anion-exchange chromatography (results not shown). When using gel-filtration analysis, the recombinant β-glucosidase was eluted in the same fractions as the wild-type enzyme extracted from cassava leaves [3], in accordance with an oligomeric structure of the enzyme. Like the cassava leaf enzyme, which contains high-mannose-type oligosaccharide chains [3], the recombinant β-glucosidase shows affinity to concanavalin-A–Sepharose matrix and can be eluted with 1.0 M α-methyl-mannopyranoside (results not shown). SDS/PAGE analysis of the purified recombinant β-glucosidase (Figure 1) showed that the monomer protein has the same molecular mass as the cassava leaf protein, estimated to be 70 kDa [3]. This suggests that hyperglycosylation of the recombinant β-glucosidase does not occur in the supersecretory yeast strain. This is consistent with the results of Rudolph et al. [24], which show that proteins secreted from the *ssc1* (*PMR1*) mutant yeast strain lack the mannose outer-chain glycosylation that normally results from passage through the Golgi.

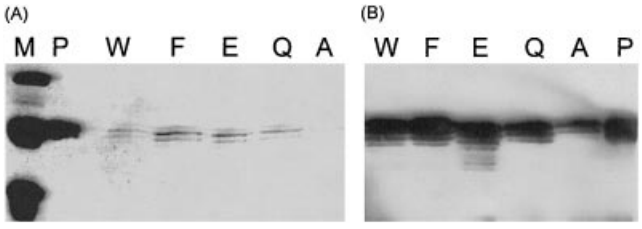

Figure 1 SDS/PAGE and Western-blot analysis of the wild-type and mutated cassava cyanogenic β-glucosidase expressed in and purified from yeast

Wild-type and mutant enzymes were purified from extracts of transformed *S. cerevisiae* strain CGY2998 and subjected to SDS/PAGE (10% gels). Gels were stained with Coomassie Brilliant Blue R (**A**) or electroblotted on to Protran nitrocellulose membrane, and the enzyme protein bands visualized by anti-CCBG antisera using ECL (**B**). M, Serva protein markers; P, cassava β-glucosidase from plant leaves; W, wild-type enzyme expressed in yeast; F, Phe-269Asn mutant; E, Glu-413Gly mutant; Q, Gln-339Glu mutant; A, Ala-201Val mutant.

Steady-state kinetic analysis of recombinant cyanogenic β-glucosidase (linamarase)

Kinetic constants k_{cat}, K_m and k_{cat}/K_m were determined for the FPLC anion-exchange-purified recombinant wild-type β-glucosidase, for the hydrolysis of the synthetic substrate *p*-nitrophenyl β-D-glucopyranoside (PNP-Glc) and the natural substrate, linamarin, in the range of pH 3.5–8.2 (Figures 2A and 2D). Molecular dissociation constants of the free enzyme (pK^E_1 and pK^E_2) and the enzyme–substrate complexes (pK^{ES}_1 and pK^{ES}_2) obtained from the theoretical curves fitted to the experimental $pH/k_{cat}/K_m$ and pH/k_{cat} data, respectively, using the appropriate equations of the diprotic enzyme model [22], together with the pH-independent kinetic parameters (k_{cat}, k_{cat}/K_m and K_m), are summarized in Table 2. The values of these parameters, except the higher k_{cat} for the yeast protein, were comparable with those of the plant enzyme (pH-independent K_m, 0.32 and 1.20 mM; k_{cat}, 6.4×10^{-3} and 6.1×10^{-3} s^{-1}; and $pK^E_1 = 4.5$–4.6, $pK^E_2 = 7.1$–7.3 for linamarin and PNP-Glc, respectively) [22], reflecting successful production and folding of the recombinant enzyme in yeast.

Analysis of active-site mutants Glu-413Gly, Gln-339Glu, Ala-201Val and Phe-269Asn

For molecular identification of essential catalytic residues in the active site of the cassava linamarase, site-directed mutagenesis was carried out [29]. As a search for possible structural changes introduced into the active site by the above amino acid substitutions, molecular modelling was carried out, using the MODELLER program [30], for the cassava linamarase on the basis of the white clover linamarase structure [13] utilizing the 49 % amino acid sequence identity and 61 % homology between the two proteins. The model generated had a root mean square of differences of 0.085 Å for the superimposition of Cα atoms and therefore satisfied the criteria of a structure at 2 Å resolution. The model was used subsequently to build the mutated active sites using conformational search and potential-energy minimization (Figure 3).

The acid/base catalyst of the cassava enzyme has been identified as Glu-198 located in the NEP/I motif [20], but it has not been established that Glu-413, located in the I/VTENG motif, is the catalytic nucleophile. The Glu-413Gly mutant enzyme does not differ from the wild-type cassava linamarase (yeast and plant) with respect to molecular mass (Figure 1) and behaviour in the standard purification procedure. Figure 3(A) shows no structural rearrangements other than the Glu-413Gly replacement. However, the Glu-413Gly mutant enzyme had no hydrolytic activity against either linamarin or PNP-Glc, indicating an essential function for Glu-413.

The NH_2 group of the clover linamarase Asn-324 residue, conserved in Family-1 enzymes [8], is H-bonded to $O^{\varepsilon 2}$ of Glu-183 and close to Glu-397, and is therefore proposed to be involved in their positioning/ionization. The corresponding residue in the cassava β-glucosidase (Gln-339) was mutated to glutamate to test whether it has any influence on the pH–activity

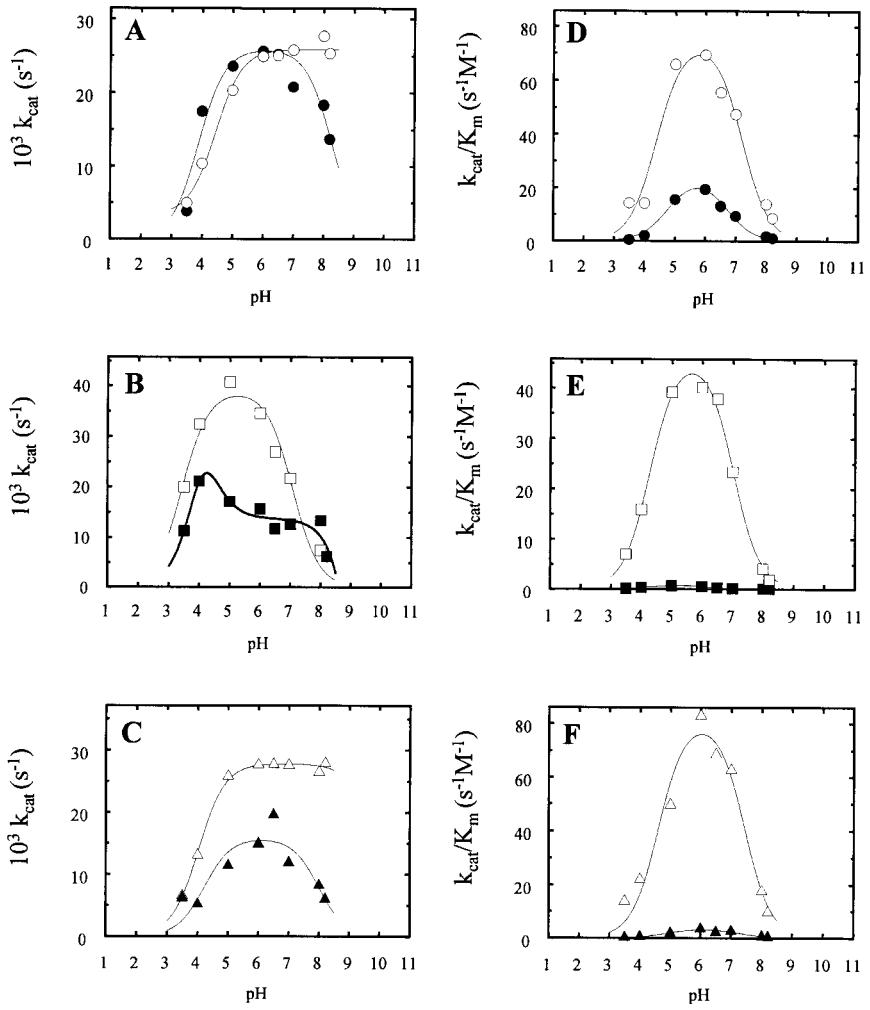

Figure 2 pH-dependence of the kinetic parameters k_{cat} (A–C) and k_{cat}/K_m (D–F) of linamarase (closed symbols) and *p*-nitrophenyl-β-D-glucosidase (open symbols) activities of the wild-type (A and D), Phe-269Asn mutant (B and E) and Ala-201Val mutant (C and F) cassava cyanogenic β-glucosidase produced in yeast

The best-fit curves were calculated by non-linear fitting on the basis of the diprotic enzyme model [22].

Table 2 Molecular dissociation constants of the free enzyme (E) and the enzyme–substrate complexes (ES) and pH-corrected kinetic parameters of the wild-type and mutated cyanogenic β-glucosidase

Reaction kinetic parameters (k_{cat}, K_m, k_{cat}/K_m) for the hydrolysis of linamarin and PNP-Glc were determined in 0.15 M citrate/phosphate buffer between pH 3.5 and 8.2, at 37 °C. Substrate concentrations varied between 0.5- and 20-fold of the K_m value. pK^E_1, pK^E_2, pK^{ES}_1, pK^{ES}_2 (the negative of the logarithms of the molecular dissociation constants of the free enzyme K^E_1, K^E_2 and the ES complexes K^{ES}_1, K^{ES}_2, respectively), k_{cat} and k_{cat}/K_m were calculated by non-linear fitting on the basis of the diprotic enzyme model [22]. K_m values were calculated from k_{cat}/K_m and k_{cat}.

	Wild-type		Phe-269Asn		Ala-201Val	
Parameter	Linamarin	PNP-Glc	Linamarin	PNP-Glc	Linamarin	PNP-Glc
pK^E_1	4.78 ± 0.14	4.39 ± 0.16	4.55 ± 0.17	4.27 ± 0.05	4.82 ± 0.21	4.61 ± 0.17
pK^E_2	6.71 ± 0.13	7.22 ± 0.17	6.15 ± 0.13	7.03 ± 0.05	7.36 ± 0.21	7.44 ± 0.17
pK^{ES}_1	3.88 ± 0.17	4.44 ± 0.14	4.1 ± 0.5*	3.43 ± 0.15	4.22 ± 0.3	4.02 ± 0.09
pK^{ES}_2	8.28 ± 0.18	–	4.01 ± 0.5*	7.07 ± 0.15	8.0 ± 0.25	–
10^3 k_{cat} (s^{-1})	25.9 ± 1.7	25.9 ± 1.6	13.7 ± 1.6†	39.1 ± 5.0	15.9 ± 4.0	27.9 ± 2.8
k_{cat}/K_m ($s^{-1} \cdot M^{-1}$)	24.6 ± 2.8	74.7 ± 6.4	0.85 ± 0.08	46.6 ± 7.2	3.24 ± 0.37	82.0 ± 6.7
K_m (mM)	1.05	0.35	16.1	0.84	4.9	0.34

* Values were calculated using the equation from [31] for when more than one ionization form of an ES complex can yield products.

† The value was derived from the plateau between pH 5.0 and 8.0 in Figure 2(B).

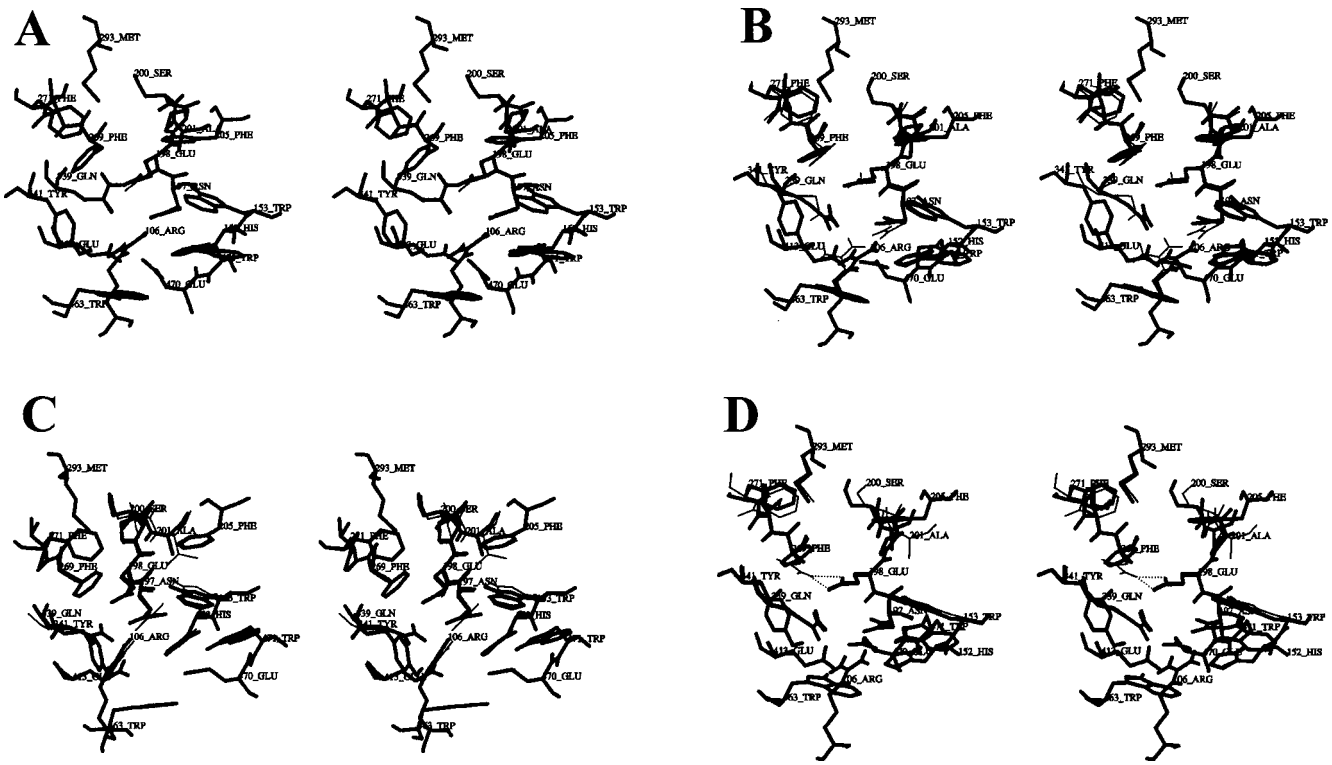

Figure 3 Comparison of the active site of the wild-type (thick lines) and mutated (thin lines) cassava linamarase models

Superimposition of the wild-type homology model and the Glu-413Gly (**A**), Gln-339Glu (**B**), Ala-201Val (**C**) and Phe-269Asn (**D**) mutant models generated by the means of conformational search and energy minimization. Active-site residues are identified by the three-letter amino acid code.

profile of the cassava enzyme. The Gln-339Glu change led to hardly detectable activity (3000 times lower) by the mutant protein when either linamarin or PNP-Glc was used as the substrate at the range of pH 3.5–7. A possible explanation is derived from molecular modelling, which predicts that Glu-339 forms a salt bridge with Arg-106 (Figure 3B). The mutation would thus disrupt the essential salt bridge between Arg-106 and Glu-413 and an H-bond between Gln-339 and Glu-198, both of which are predicted in the model.

The clover linamarase Gly-186, the equivalent of which is mostly hydrophobic in all Family-1 enzymes, contributes to the environment of $O^{\epsilon 1}$ of Glu-183, and might support its catalytically essential protonation even at higher pH [13,20,22]. The replacement of the corresponding Ala-201 in the cassava enzyme with a valine was carried out to probe the extent of such a water-sheltering effect (Figure 3C). Kinetic analysis shows that the $pK^{E}{}_{2}$ of the Ala-201Val mutant (Figures 2C and 2F, Table 2), representing the acid/base catalyst, is slightly higher (7.4) than that of the wild-type linamarase (7.2). It is noteworthy, however, that whereas the catalytic constant for the hydrolysis of linamarin decreases to $15.9 \times 10^{-3}\ s^{-1}$ (from $25.9 \times 10^{-3}\ s^{-1}$ for the wild-type enzyme), it remains largely unchanged for PNP-Glc ($k_{cat} = 27.9 \times 10^{-3}\ s^{-1}$ versus $25.9 \times 10^{-3}\ s^{-1}$; Table 2). In addition, the Michaelis constant (K_m) for the hydrolysis of linamarin by the Ala-201Val mutant is five times higher (4.9 mM) than for that of the wild type (1.06 mM), whereas it remains unchanged for PNP-Glc hydrolysis ($K_m = 0.34$ mM). The differential effect of the Ala → Val replacement suggests that the steric hindrance introduced by Val-201 (Figure 3C) may selectively inhibit docking of the tetrahedral acetone cyanohydrin moiety of linamarin, compared with the planar *p*-nitrophenyl aglycone of PNP-Glc.

Val-254 of the clover linamarase is located in a competent position for a contact with substrate aglycones [13]. While the corresponding residue is hydrophobic in both cassava linamarase and *S. alba* myrosinase (Phe-269 and Ile-257, respectively), it is replaced by asparagine in the disaccharide-cleaving bacterial β-glucosidases from *Agrobacterium faecalis*, *B. polymyxa* and *Clostridium thermocellum* [8]. The Phe-269Asn cassava mutant has no detectable cellobiase activity; however, it is still active against linamarin and PNP-Glc. When using PNP-Glc as the substrate, the pH-dependence profile for k_{cat} of the Phe-269Asn mutant is bell-shaped, with molecular dissociation constants $pK^{ES}{}_{1} = 3.43$ and $pK^{ES}{}_{2} = 7.07$ for the catalytic nucleophile and the acid/base catalyst, respectively (Figure 2B, Table 2). In contrast, the pH/k_{cat} profile of the wild-type enzyme for PNP-Glc does not have an inflection point in the basic range, and only $pK^{ES}{}_{1} = 4.44$ could be derived (Figure 2A). With PNP-Glc, $pK^{E}{}_{2} = 7.03$ is somewhat lower than that of the wild-type enzyme ($pK^{E}{}_{2} = 7.22$). The pH/k_{cat}/K_m profile for linamarin hydrolysis by the mutant (Figure 2E) gives a $pK^{E}{}_{2}$ value of 6.15, which is also lower than that of the wild-type enzyme ($pK^{E}{}_{2} = 6.71$). The most dramatic effect of the mutation is observed on the pH/k_{cat} profile determined with linamarin, which is changed from the bell-shaped curve of the wild type to a three-shouldered curve with three molecular dissociation constants $pK^{ES}{}_{1} = 4.1$, $pK^{ES}{}_{2} = 4.01$ and $pK^{ES}{}_{3} = 8.5$ (Figure 2B). The best-fit theoretical curve could be calculated using the equation suggested by Tipton and Dixon [31] for the situation when more than one ionization

form of an enzyme–substrate complex can yield products. In the case of the Phe-269Asn mutant, it probably means that the enzyme is active irrespective of whether the acid catalyst is protonated or deprotonated. Indeed, the model of the mutant shows that the NH_2 group of Asn-269 is positioned to form an H-bond with carbonyl oxygen atoms of Glu-198 (Figure 3D), which can alter ionization and/or water structure around the residue. In addition, the mutation seems to selectively affect the deglycosylation step of substrate hydrolysis because pK^{ES}_2, but not pK^{E}_2, shifts dramatically not only with linamarin from 8.28 to 4.01 but also with PNP-Glc from > 10 to 7.07 (Table 2). These results demonstrate that replacement of Phe-269 with Asn diminishes protonation of the acid/base catalyst (Glu-198) at higher pH, supposedly because of the change in its hydrophobic environment caused by the introduction of a polar residue (Figure 3D).

Kinetic parameters determined for the Phe-269Asn mutant (Figures 2B and 2E, Table 2) show that K_m of linamarin hydrolysis is increased 16-fold by the mutation (from 1.05 mM for the wild type to 16.1 mM for the mutant), while the K_m for PNP-Glc cleavage is increased only 2.5-fold (from 0.34 to 0.84 mM). In contrast, the catalytic constant measured with the linamarin substrate is only decreased from 25.9×10^{-3} to 13.7×10^{-3} s^{-1}, whereas k_{cat} for PNP-Glc is raised 1.5 times (from 25.9×10^{-3} to 39.1×10^{-3} s^{-1}) by the Phe-269Asn replacement. The fact that substituting Phe-269 with a polar residue confers a significant effect on K_m of linamarin cleavage without a similar magnitude change in k_{cat} can only be attributed to selective alterations in the rate of binding and/or dissociation of this substrate to and/or from the enzyme. Since no comparable shift can be seen in the K_m for PNP-Glc hydrolysis by the mutant protein, we can establish that Phe-269 in the wild-type enzyme has a more important function in binding the natural substrate, linamarin, than the synthetic PNP-Glc.

All four mutants are under detailed kinetic investigations to clarify further the function of residues targeted in this study. In addition to the structure/function information derived from the work presented in this paper, a system has been established for the production of recombinant cassava linamarase, opening a perspective for protein engineering aimed at the improvement of this enzyme for applications in cancer therapy.

Z. K. was a postdoctoral research fellow supported by a Royal Society-NATO Postdoctoral Fellowship. We wish to thank Dr H. R. Haysom for his invaluable help in generation of graphical illustrations.

REFERENCES

1 Poulton, J. E. (1990) Cyanogenesis in plants. Plant Physiol. **94**, 401–405
2 Hughes, M. A. (1998) Biosynthesis and degradation of cyanogenic glycosides. In Comprehensive Natural Products Chemistry, vol. 1 (Sankawa, U., ed.), pp. 881–895, Elsevier, Oxford
3 Hughes, M. A., Brown, K., Pancoro, A., Murray, B. S., Oxtoby, E. and Hughes, J. (1992) A molecular and biochemical analysis of the structure of the cyanogenic β-glucosidase (linamarase) from cassava (*Manihot esculenta* Crantz). Arch. Biochem. Biophys. **295**, 273–279
4 Hughes, J., Carvalho, F. J. P.DeC. and Hughes, M. A. (1994) Purification, characterisation, and cloning of α-hydroxynitrile lyase from cassava (*Manihot esculenta* Crantz). Arch. Biochem. Biophys. **311**, 496–502
5 Hernandez, T., Lundquist, P., Oliveira, L., Perez Cristina, L., Rodriguez, E. and Rosling, H. (1995) Fat in humans of dietary intake of cyanogenic glycosides from roots of sweet cassava consumed in Cuba. Natural Toxins **3**, 114–117
6 Cortés, M. L., de Felipe, P., Martín, V., Hughes, M. A. and Izquierdo, M. (1998) Successful use of a plant gene in the treatment of cancer *in vivo*. Gene Therapy **5**, 1499–1507
7 Syringos, K. N., Rowlinson-Busza, G. and Epenetos, A. A. (1998) *In vitro* cytotoxicity following specific activation of amygdalin by β-glucosidase conjugated to a bladder cancer-associated monoclonal antibody. Int. J. Cancer **78**, 712–719
8 Henrissat, B. and Bairoch, A. (1996) Updating the sequence-based classification of glycoside hydrolases. Biochem J. **316**, 695–696
9 Davies, G., Sinnott, M. L. and Withers, S. G. (1998) Comprehensive Biological Catalysis, vol. 1 (Sinnott, M. L., ed.), pp. 109–208, Academic Press, London
10 McCarter, J. and Withers, S. G. (1994) Mechanism of enzymatic glycoside hydrolysis. Curr. Opin. Struct. Biol. **4**, 885–892
11 White, A., Tull, D., Johns, K., Withers, S. G. and Rose, D. R. (1996) Crystallographic observation of a covalent catalytic intermediate in a β-glycosidase. Nat. Struct. Biol. **3**, 149–154
12 Davies, G. and Henrissat, B. (1995) Structure and mechanism of glycosyl hydrolases. Structure **3**, 853–859
13 Barrett, T., Suresh, C. G., Tolley, S. P., Dodson, E. J. and Hughes, M. A. (1995) The crystal structure of a cyanogenic β-glucosidase from white clover (*Trifolium repens* L.), a family 1 glycosyl hydrolase. Structure **3**, 951–960
14 Oxtoby, E., Dunn, M. A., Pancoro, A. and Hughes, M. A. (1991) Nucleotide and derived amino acid sequence of the cyanogenic β-glucosidase (linamarase) from white clover (*Trifolium repens* L.). Plant Mol. Biol. **17**, 209–219
15 Burmeister, W. P., Cottaz, S., Driguez, H., Renato, I., Palmieri, S. and Henrissat, B. (1997) The crystal structure of *Sinapis alba* myrosinase and a covalent glycosyl-enzyme intermediate provide insight into the substrate recognition and active-site machinery of an *S*-glycosidase. Structure **5**, 663–675
16 Weismann, C., Hengstenberg, W. and Schulz, G. E. (1997) Crystal structure and mechanism of 6-phospho-β-galactosidase from *Lactobacillus lactis*. J. Mol. Biol. **269**, 851–860
17 Sanz-Aparicio, J., Hermoso, J. A., Martinez-Ripoll, M., Lequerica, J. L. and Polaina, J. (1998) Crystal structure of β-glucosidase A from *Bacillus polymyxa*: insight into the catalytic activity in family 1 glycosyl hydrolases. J. Mol. Biol. **275**, 491–502
18 Chi, Y. I., Martinez-Cruz, L. A., Jancarik, J., Swanson, R. V., Robertson, D. E. and Kim, S. (1999) Crystal structure of the β-glycosidase from the hyperthermophile *Thermosphaera aggregans*: insights into its activity and thermostability. FEBS Lett. **445**, 375–383
19 Hakulinen, N., Paavilainen, S., Korpela, T. and Rouvinen, J. (1999) The crystal structure of β-glucosidase from *Bacillus circulans* ssp. *alkalophilus*: ability to form long polymeric assemblies. J. Struct. Biol. **129**, 69–79
20 Keresztessy, Zs., Kiss, L. and Hughes, M. A. (1994) Investigation of the active site of the cyanogenic β-glucosidase (linamarase) from *Manihot esculenta* Crantz (cassava). II. Identification of Glu-198 as an active site carboxylate with acid catalytic function. Arch. Biochem. Biophys. **315**, 323–330
21 Withers, S. G. (1996) Identification of active-site residues in glycosidases. In Biological and Biotechnological Application of ESI-MS, ACS Symposium Series vol. 619 (Snyder, A. P., ed.), pp. 365–380, American Chemical Society, Washington DC
22 Keresztessy, Zs., Kiss, L. and Hughes, M. A. (1994) Investigation of the active site of the cyanogenic β-glucosidase (linamarase) from *Manihot esculenta* Crantz (cassava). I. Evidence for an essential carboxylate and a reactive histidine group in a single catalytic centre. Arch. Biochem. Biophys. **314**, 142–152
23 Keresztessy, Zs., Hughes, J., Kiss, L. and Hughes, M. A. (1996) Co-purification from *E. coli* of a plant β-glucosidase-GST fusion protein and the bacterial chaperonin GroEL. Biochem. J. **314**, 41–47
24 Rudolph, H. K., Antebi, A., Fink, G. R., Buckley, C. M., Dorman, T. E., LeVitre, J., Davidow, L. S., Mao, J.-i. and Moir, D. T. (1989) The yeast secretory pathway is perturbed by mutations in PMR1, a member of a Ca^{2+} ATPase family. Cell **58**, 133–145
25 Carlson, M., Taussing, R., Kustu, S. and Botstein, D. (1983) The secreted form of invertase in *Saccharomyces cerevisiae* is synthesized from mRNA encoding a signal sequence. Mol. Cell. Biol. **3**, 439–445
26 Leary, N. O., Pembroke, A. and Duggan, P. F. (1992) Improving accuracy of glucose oxidase procedure for glucose determinations on discrete analysers. Clin. Chem. **38**, 298–302
27 Leatherbarrow, R. J. (1990) Grafit Version 2.0, Erithacus Software Ltd., Staines
28 Hughes, M. A. and Dunn, M. A. (1985) The use of immunochemical techniques to study plant proteins. Plant Mol. Biol. Reporter **3**, 17–23
29 Ly, H. D. and Withers, S. G. (1999) Mutagenesis of glycosidases. Annu. Rev. Biochem. **68**, 487–522
30 Sali, A. and Blundell, T. J. (1993) Comparative protein modelling by satisfaction of spatial restraints. J. Biol. Chem. **234**, 779–815
31 Tipton, K. F. and Dixon, H. B. (1979) Effect of pH on enzymes. Methods Enzymol. **63**, 183–234

Received 8 May 2000/27 September 2000; accepted 19 October 2000