Identification of the catalytic nucleophile of the Family 31 α-glucosidase from *Aspergillus niger* via trapping of a 5-fluoroglycosyl–enzyme intermediate

Seung Seo LEE, Shouming HE and Stephen G. WITHERS

Department of Chemistry, University of British Columbia, Vancouver, BC, Canada V6T 1Z1

The mechanism-based reagent 5-fluoro-α-D-glucopyranosyl fluoride (5FαGlcF) was used to trap a glycosyl–enzyme intermediate and identify the catalytic nucleophile at the active site of *Aspergillus niger* α-glucosidase (Family 31). Incubation of the enzyme with 5FαGlcF, followed by peptic proteolysis and comparative liquid chromatography/MS mapping allowed the isolation of a labelled peptide. Fragmentation analysis of this peptide by tandem MS yielded the sequence WYDMSE, with the label located on the aspartic acid residue (D). Comparison with the known protein sequence identified the labelled amino acid as Asp-224 of the P2 subunit.

Key words: catalytic carboxyl, 5-fluoro-α-D-glucopyranosyl fluoride, fluorosugar, mechanism-based reagent.

INTRODUCTION

*Aspergillus niger* α-glucosidase (EC 3.2.1.20) is an exo-acting glucosidase with a pH optimum of 4–4.5 that releases α-glucose from the non-reducing end of starch and malto-oligosaccharides [1,2]. It is a glycoprotein containing 25.5–27.6 % carbohydrate, most of which is mannose [1,3,4]. The enzyme contains two subunits, P1 and P2, with a combined molecular mass of 125000 Da [2]. The full amino acid sequence is available [5], Swiss-Prot accession no. P56526) and it has been assigned to Family 31 in Henrissat’s classification based on amino acid sequence homology [6–8]. This family includes a wide range of plant enzymes as well as some fungal and bacterial α-glucosidases. A number of mammalian α-glucosidases also belong to this family, including human isomaltase, sucrase, maltase and lysosomal α-glucosidase. Enzymes in this family cleave only α-glucosidic linkages and prefer malto-oligosaccharides and α-glucans rather than heterogeneous substrates. On this basis they have been classified further as type III α-glucosidases [9]. In addition to α-glucosidases, a group of α-1,4-glucan lyases also belongs to this family [10]. The lyases cleave α-1,4-glucosidic linkages via a non-hydrolytic mechanism releasing an elimination product from the non-reducing end. The mechanistic details of the lyases are not yet defined.

Hydrolysis of α-glucosides by Family 31 α-glucosidases has been shown to occur with retention of stereochemistry at the anomeric centre [1,9]. Retaining enzymes typically have two catalytic carboxylic acids in their active site and utilize a double displacement mechanism (Scheme 1) in which a β-glycosyl–enzyme intermediate is formed and hydrolysed. Both steps proceed via transition states with substantial oxocarbenium ion character [11–13]. One carboxylic acid in the active site acts as the catalytic nucleophile, leading to the formation of the covalent intermediate, while the other plays the role of the general acid catalyst in the first step and the general base catalyst in the second step (hydrolysis of the glycosyl–enzyme intermediate).

5-Fluoro-α-D-glucopyranosyl fluoride (5FαGlcF; Figure 1) has been developed as a useful reagent for trapping the intermediate formed by retaining α-glucosidases [14,15]. The highly electronegative fluoride at C-5 destabilizes the oxocarbenium ion-like transition states, slowing both the formation and the hydrolysis of the intermediate. However, the presence of a good leaving group, fluoride, at the anomeric centre ensures that the formation of the intermediate (glycosylation) is faster than its hydrolysis (deglycosylation). As a result, the glucosyl–enzyme intermediate will accumulate, in some cases resulting in inactivation of the enzyme if turnover is very slow. In other cases, where turnover occurs at rates that are comparable with those of intermediate formation, the outcome is steady-state accumulation of the intermediate. In such cases the analogue appears to behave like a tight-binding reversible competitive inhibitor. The catalytic competence of this intermediate, as indicated by its turnover, is clear evidence that the species accumulated is mechanistically relevant. This accumulated intermediate can be digested with a protease and the resultant peptide mixture analysed by a combination of liquid chromatography (LC)/MS and tandem MS (MS/MS) to identify the labelled residue. This approach involving 5-fluoroglycosides has been employed previously to identify the catalytic nucleophiles of α-glucosidases in Family 13 [15] as well as α-galactosidases in Family 27 [16] and α-mannosidases in Family 38 [17,18]. However, the catalytic nucleophile in Family 31 enzymes has only been identified so far [19–22] on the basis of labelling by conduritol B epoxide (Figure 1), a reagent that does not form a catalytically competent intermediate and has been shown in several cases to label residues other than the catalytic nucleophile [11,23–25]. It is therefore important to identify unequivocally the catalytic nucleophile of one of the Family 31 α-glucosidases with this more reliable probe. This report describes the labelling and identification of the catalytic nucleophile of *A. niger* α-glucosidase using 5FαGlcF.

Abbreviations used: 5FαGlcF, 5-fluoro-α-D-glucopyranosyl fluoride; PNPαGlc, p-nitrophenyl-α-D-glucopyranoside; RP-HPLC, reversed-phase HPLC; LC, liquid chromatography; MS/MS, tandem MS; ESI-MS, electrospray ionization MS; TIC, total ion chromatogram.

1 To whom correspondence should be addressed (e-mail withers@chem.ubc.ca).
S. S. Lee, S. He and S. G. Withers

Scheme 1 The catalytic mechanism of a retaining \(\alpha\)-glucosidase

Figure 1 The structure of the mechanism-based reagents
(a) \(5\text{F}\alpha\text{GlcF};\) (b) conduritol B epoxide.

MATERIALS AND METHODS

Materials

\(5\text{F}\alpha\text{GlcF}\) was synthesized as described previously [14]. \(A.\ niger\ \alpha\)-glucosidase was purchased from Megazyme International Ireland (Bray, Ireland) and dialysed against 0.036 M sodium acetate/0.064 M acetic acid, pH 4.5 (0.1 M sodium acetate buffer), at 4 °C for 48 h before use. Pepsin (from porcine mucosa) was purchased from Boehringer Mannheim. All other chemicals and reagents were purchased from Sigma unless otherwise noted.

Enzyme kinetics

All experiments were carried out at 37 °C in 0.1 M sodium acetate buffer, pH 4.5, containing 0.1 % BSA. Kinetic parameters for hydrolysis of \(p\)-nitrophenyl-\(\alpha\)-d-glucopyranoside (PNP\(\alpha\)Glc) were determined from the initial linear increase in absorbance at 360 nm upon the addition of enzyme (final concentration, 6.94 \(\mu\)g/ml) to a range of concentrations (0.1–1.2 mM) of...
PNP₂Glc. Cuvettes of 1 cm path length were used for the measurement in a UNICAM UV/visible spectrophotometer equipped with a circulating-water bath. The difference in molar absorption coefficients $\Delta \varepsilon$ between PNP₂Glc and p-nitrophenol at 360 nm at pH 4.5, 37 °C, was determined to be 1.8789 mM⁻¹·cm⁻¹ by measuring the difference in absorbances between fixed equal concentrations of p-nitrophenol and PNP₂Glc. Rates were calculated using the determined molar absorption coefficient, and kinetic parameters were obtained by direct fit of the data to the Michaelis–Menten equation using GraFit 3.0 (Erithacus Software, Staines, Middx, U.K.). For inactivation kinetics, the enzyme (final concentration, 0.57 mg/ml) was pre-incubated with a range of concentrations of 5F₂GlcF (0.01–2.9 mM) at 37 °C and 10 μl aliquots of the sample were withdrawn at time intervals and added to 800 μl of 1 mM PNP₂Glc, which was pre-equilibrated at 37 °C in the UV/visible spectrometer. The initial rate at each concentration of 5F₂GlcF was measured in this way. An apparent reversible inhibition constant, $K_i$, was determined by measuring rates of hydrolysis of PNP₂Glc in the presence of various concentrations of 5F₂GlcF (2.5–50 μM). A Dixon plot of $1/v$ versus [5F₂GlcF] intersects a line given by $1/V_{\text{max}}$ at an inhibitor concentration equal to $-K_i$. The $K_{i\text{at}}$ value for 5F₂GlcF was determined by monitoring the release of fluoride using an Orion 96-09 combination fluoride electrode. As initial rates of fluoride release at all concentrations (0.35–2.67 mM) assayed were identical, this rate was taken as the $V_{\text{max}}$, from which $k_{\text{cat}}$ was calculated.

Protein measurement
The molar absorption coefficient $\epsilon$ for 1% (w/v) protein at 280 nm was 20.1 cm⁻¹ obtained from the measurement of the absorbance of a specifically diluted solution of the α-glucosidase, which was provided in a known concentration by the manufacturer, using 1 cm-path-length cuvettes. The concentrations of dialysed enzyme samples were measured based on this value.

Labelling and proteolysis of A. niger α-glucosidase
A stock solution of the enzyme (20 μl, 8.13 mg/ml) was incubated with 5F₂GlcF (20 μl, 20 mM) at 37 °C for 5 min. The sample was diluted with 0.05 M phosphate buffer (pH 2, 90 μl) and incubated with pepsin (15 μl, 1 mg/ml) for 15 min at room temperature. The sample was then frozen quickly and analysed immediately upon thawing. A control sample was prepared according to the same procedure, except that no 5F₂GlcF was added.

Electrospray ionization MS (ESI-MS)
Mass spectra were recorded on a PE-Sciex API 300 triple quadrupole mass spectrometer (Sciex, Thornhill, Ontario, Canada) equipped with an ESI ion source. Peptides were separated by reversed-phase HPLC (RP-HPLC) on an Ultrafast Microprotein Analyser (Michrom BioResource, Pleasanton, CA, U.S.A.) directly interfaced with the mass spectrometer. In each MS experiment, the proteolytic digest was loaded on to a C₁₈ column (Relisil, 1 mm × 150 mm) and eluted with a gradient of 0–60% eluting solvent (0.045% trifluoroacetic acid/80% acetonitrile in water) over 60 min at a flow rate of 50 μl/min. A post-column splitter was used in all experiments, splitting off 85% of the sample into a fraction collector and sending 15% into the mass spectrometer. Spectra were obtained in either the single-quadrupole scan mode (LC/MS) or the tandem MS daughter-ion scan mode (MS/MS).

In the single quadrupole mode (LC/MS), the quadrupole mass analyser was scanned over an $m/z$ range of 400–2000 a.m.u. with a step size of 0.5 Da and a dwell time of 1.0 ms/step. The ion-source voltage was set at 5 kV and the orifice energy was 50 V. After the LC/MS experiment, total ion chromatograms (TICs) of the labelled and unlabelled enzyme digests were compared to find the fraction containing the labelled peptide fragments. Samples of the labelled peptide were collected from the post-column flow splitter and lyophilized. The concentrated sample was then sequenced via MS/MS fragmentation analysis.

In the MS/MS daughter-ion scan mode mass spectra were obtained by selectively introducing the $m/z$ 1011 peptide from the first quadrupole (Q1) into the collision cell (Q2) and observing the daughter ions in the third quadrupole (Q3). The following settings were applied. Q1 was locked on $m/z$ 1011. Q3 scan range was 100–1020 a.m.u., step size was 0.5 a.m.u., dwell time was 1 ms, ion-source voltage was 5 kV, orifice voltage was 50 V, focusing-ring voltage was 200 V. Q0 potential was −10 V; Q2 potential was −42 V and the collision gas was N₂.

RESULTS AND DISCUSSION

Enzyme kinetics
Since the optimum pH of A. niger α-glucosidase is pH 4–4.5 [2,4], hydrolysis of PNP₂Glc cannot be monitored at 400 nm since the phenol red product (p$K_a$ 7.15) is almost entirely protonated and therefore colourless. However, at a wavelength of 360 nm there is a substantial difference in molar absorption coefficient (1.8789 mM⁻¹·cm⁻¹), allowing kinetic analysis. Kinetic parameters for the hydrolysis of PNP₂Glc of $K_m = 0.31 ± 0.02$ mM and $k_{\text{cat}} = 2.3 ± 0.06$ s⁻¹ were determined in this way. These are in general agreement with values in the literature of $K_m = 0.7$ mM and $k_{\text{cat}} = 4.3$ s⁻¹ [4].

Reaction of A. niger α-glucosidase with 5F₂GlcF
When 5F₂GlcF was incubated with A. niger α-glucosidase and aliquots were removed at time intervals, no time-dependent inactivation was observed. Instead a substantially reduced activity was measured at time zero, and this did not change with time. The compound therefore appeared to be acting as a tight-binding reversible inhibitor, and thus was analysed as a competitive inhibitor in a second set of experiments, yielding an apparent $K_i$ value of $2.5 ± 0.13$ μM. Similar behaviour has been seen previously with several other 5-fluoro-α-glycosyl fluorides and α-glycosidases, and shown to be due to steady-state accumulation of a 5-fluoroglycosyl–enzyme intermediate, due to the relatively slow turnover, yet rapid formation, of the intermediate [15,16,18]. The apparent $K_i$ does not measure the true inhibitor–enzyme affinity. Instead, it includes contributions from both the binding event ($K_i$) and the chemical reactions which are, in this case, the steady-state formation ($k_a$) and turnover ($k_i$) of the covalent glycosyl–enzyme intermediate according to the following expression, $K_i^{\text{app}} = K_i / (1 + k_a / k_i)$ [26,27]. Thus the apparent $K_i$ represents a minimum value of the true $K_i$. The inhibition behaviour was further probed by direct monitoring of turnover using a fluoride-ion-specific electrode. The enzymic reaction continued at a constant rate until almost all the substrate was consumed. In addition, this same rate was seen at a range of concentrations of the 5F₂GlcF. Both of these observations are consistent with the low $K_i$ value.
measured since the $K_m$ value for this compound as a substrate must equal its $K_i$ value as a competitive inhibitor. The catalytic-centre activity (‘turnover number’), $k_{cat}$, was determined from the slope of these plots of fluoride release versus time, and found to be $0.055 \pm 0.001 \text{s}^{-1}$. Unfortunately, a $K_m$ value could not be determined from this experiment due to the lack of sensitivity of the fluoride electrode towards the very low substrate concentrations (1–5 μM) needed for its determination. This $k_{cat}$ value is comparable with those found for 5-fluoroglycosyl fluoride with other enzymes. Thus the $k_{cat}$ for 5FxGlcF with yeast α-glucosidase was $0.11 \pm 0.01 \text{s}^{-1}$ [15] and that for 5-fluoro-β-1-gulopyranosyl fluoride (epimer of the corresponding mannosyl compound at C-5) with bovine kidney lysosomal α-mannosidase was $0.022 \text{s}^{-1}$ [18].

Identification of the catalytic nucleophile of *A. niger* α-glucosidase

*A. niger* α-glucosidase was incubated with 5FxGlcF to form an accumulated intermediate, then immediately subjected to peptic digestion at pH 2. A sample of unlabelled α-glucosidase was also digested by pepsin under the same conditions. The peptic digests from the labelled and unlabelled enzymes were loaded on to the microbore RP-HPLC connected to the ESI mass spectrometer, and TICs for each sample were obtained. The TIC of the labelled enzyme is practically indistinguishable from that of the unlabelled enzyme. In order to identify the labelled peptide, the masses of the peptides under each peak in the labelled sample were compared with those of the peptides from the unlabelled sample in the corresponding region of the TIC. The masses of the peptides from within the two samples were identical, with
one exception. A peptide of mass 1011 Da eluting at 27.76 min was observed only in the TIC of the labelled sample (Figure 2). If this is the labelled peptide of interest, then a peptide of mass ≈ 830 Da might be expected in the TIC of the unlabelled sample, this being the mass difference between the peptide of mass 1011 Da and the 5-fluoro-glucosyl label of mass 181 Da. Unfortunately, no such peptide was observed (Figure 2), possibly indicating that the unlabelled peptide is susceptible to further peptic digestion and has been converted to smaller fragments. Such differences in proteolytic cleavage as a consequence of the presence of a sugar residue have been demonstrated previously [28].

To further investigate the possibility of this being the labelled peptide and assign the labelled amino acid residue, the fraction containing the putative labelled peptide was purified by RP-HPLC and subjected to ESI-MS/MS fragmentation analysis. The daughter-ion spectrum arising from the peptide of m/z 1011 is shown in Figure 3, along with an interpretation of the spectrum. This spectrum shows the fragment of m/z 830, which was absent in the TIC of the unlabelled enzyme and corresponds to the mass of the parent ion (1011 minus the mass of label 181), confirming that this peptide of mass 1011 Da is the labelled peptide of interest. Analysis of the fragmentation pattern revealed that the peptide sequence was WYDMSE, as shown in the spectrum in Figure 3. Two prominent peaks are observed that arise from fragments containing the sugar moiety. One of these, 776.5, corresponds to WYDM + 5FGlC. The other, 646, corresponds to either WYD + 5F GlC or DMSE + 5F GlC. Within the peptides WYDM and WYD either the tyrosine residue or the aspartate could be the catalytic nucleophile. However, the tyrosine is not conserved while the aspartate is invariant (Figure 4). Should the labelled fragments be WYDM and DMSE, only the aspartic acid is a feasible candidate. Both analyses point to the catalytic nucleophile being an aspartic acid corresponding to Asp-224 of the P2 subunit of the enzyme (Figure 4).

Previous labelling studies with other Family 31 α-glucosidases using conduritol B epoxide have involved labelling of human intestinal sucrase–isomaltase [19], human lysosomal α-glucosidase [20], sugar beet α-glucosidase [21] and A. niger α-glucosidase [22]. All these studies indicated that the labelled residue is an invariant aspartate residue corresponding to Asp-224 of A. niger α-glucosidase. However, the use of conduritol epoxides could well lead to the misassignment of the labelled residue because of the relatively high reactivity of its epoxide, coupled with the absence of a C-6 hydroxymethyl group. When this group is missing binding can occur in alternate modes, leading to non-specific labelling. Indeed, the use of conduritol B epoxide has previously resulted in misassignments of the catalytic nucleophiles of Escherichia coli β-galactosidase [24] and human lysosomal β-glucosidase [25], amongst others. In the case of Family 31 enzymes, the consistent labelling of the same conserved residue with conduritol epoxides suggested that this residue could be the catalytic nucleophile, and the present study confirms this by using the more reliable mechanism-based reagent 5FαGlcF, which forms a catalytically competent intermediate. The highly conserved aspartate residue assigned as a catalytic nucleophile as Asp-224 within the sequence WYDMSE. We thank the Natural Science and Engineering Research Council, the Canadian Institutes of Health Research and the Protein Engineering Network of Centres of Excellence of Canada for assisting this work financially.

REFERENCES


© 2001 Biochemical Society

Received 13 April 2001/22 June 2001; accepted 25 July 2001