Protein Science (1999), *8*:635–643. Cambridge University Press. Printed in the USA.

Cloning, mutagenesis, and structural analysis of human pancreatic α-amylase expressed in *Pichia pastoris*

EDWIN H. RYDBERG,[1] GARY SIDHU,[1] HUNG C. VO,[1] JEFF HEWITT,[1] HÉLÈNE C.F. CÔTÉ,[1] YILI WANG,[1] SHIN NUMAO,[2] ROSS T.A. MACGILLIVRAY,[1] CHRISTOPHER M. OVERALL,[1,3] GARY D. BRAYER,[1] AND STEPHEN G. WITHERS[1,2]

[1]Department of Biochemistry & Molecular Biology, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z1
[2]Department of Chemistry, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z1
[3]Department of Oral Biological & Medical Sciences, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z1

(RECEIVED September 8, 1998; ACCEPTED November 17, 1998)

Abstract

Human pancreatic α-amylase (HPA) was expressed in the methylotrophic yeast *Pichia pastoris* and two mutants (D197A and D197N) of a completely conserved active site carboxylic acid were generated. All recombinant proteins were shown by electrospray ionization mass spectrometry (ESI-MS) to be glycosylated and the site of attachment was shown to be Asn461 by peptide mapping in conjunction with ESI-MS. Treatment of these proteins with endoglycosidase F demonstrated that they contained a single N-linked oligosaccharide and yielded a protein product with a single N-acetyl glucosamine (GlcNAc), which could be crystallized. Solution of the crystal structure to a resolution of 2.0 Å confirmed the location of the glycosyl group as Asn461 and showed that the recombinant protein had essentially the same conformation as the native enzyme. The kinetic parameters of the glycosylated and deglycosylated wild-type proteins were the same while the k_{cat}/K_m values for D197A and D197N were 10^6–10^7 times lower than the wild-type enzyme. The decreased k_{cat}/K_m values for the mutants confirm that D197 plays a crucial role in the hydrolytic activity of HPA, presumably as the catalytic nucleophile.

Keywords: catalysis; crystallography; diabetes; enzyme; glycosylation; human pancreatic amylase; mutagenesis; *Pichia pastoris*

α-Amylases (α 1-4 glucan 4-glucanohydrolase, EC 3.2.1.1) catalyze the hydrolysis of α(1-4) glycosidic linkages of starch and are widely distributed in nature, being found in bacteria, plants, and animals. α-Amylases are members of glycosyl hydrolase family 13 as classified by Henrissat (Henrissat, 1991; Henrissat et al., 1995; Henrissat & Davies, 1997). Family 13 enzymes perform a wide range of reactions including hydrolysis, transglycosylation, cyclization, and disproportionation and act on a range of substrates including those with α(1–4), α(1–6), or α(1,1) glycosidic bonds. Members of this family share a similar global polypeptide folding pattern, and exhibit strong sequence homology in four small regions of their active sites (Nakajima et al., 1986).

Human pancreatic α-amylase (HPA) is a 56 kDa protein consisting of 496 amino acids in a single polypeptide chain, which is known to fold into three domains (Brayer et al., 1995). Domain A is a $(\beta/\alpha)_8$ barrel that binds an allosteric chloride ion. The putative active site is a "V"-shaped cleft at the C-terminal end of the $(\beta/\alpha)_8$ barrel. Domain B is located between the third β-strand and the helix of the first domain. This domain binds a calcium ion and appears to be necessary for the structural stability of an active site loop (Buisson et al., 1987). Domain C, the putative starch-binding domain, is an eight-stranded β-sheet domain at the C-terminal end of the enzyme.

HPA hydrolyzes α(1–4) glucosidic linkages with retention of configuration at the sugar anomeric center (Braun et al., 1993). Hydrolysis is thought to proceed via a double displacement reaction with general acid catalyzed formation of a β-D-glycopyranosyl-enzyme intermediate followed by general base catalyzed hydrolysis of this intermediate. The enzymic nucleophile and acid/base catalyst have been shown to be carboxylic acids in almost all glycosidases studied (Sinnott, 1990; McCarter & Withers, 1994; Withers & Aebersold, 1995). The catalytic nucleophile has been identified

Reprint requests to: Stephen G. Withers, Department of Chemistry, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z1; e-mail: withers@chem.ubc.ca.

in a number of cases by trapping the glycosyl-enzyme intermediate and identifying the glycosylated peptide in proteolytic digests. For example, the active site nucleophile of yeast α-glucosidase was identified as D214 by covalently labeling the enzyme using the mechanism-based inactivator 5-fluoro-α-D-glucosyl fluoride and analyzing peptic digests thereof by LC-MS/MS (McCarter & Withers, 1996). The structures of several α-amylases (*Aspergillus oryzae* (Swift et al., 1991), *Aspergillus niger* (Brady et al., 1991), *Bacillus licheniformis* (Suzuki et al., 1990), barley (Kadziola et al., 1994), porcine pancreatic (Buisson et al., 1987), human salivary (Ramasubbu et al., 1996), and human pancreatic (Brayer et al., 1995)) have been determined, including several structures with bound inhibitors or substrate analogs (Qian et al., 1994, 1997; Wiegand et al., 1995; Bompard-Gilles et al., 1996; Brzozowski & Davies, 1997). All of these structures show a cluster of three carboxylic acids in the active site (D197, E233, D300—HPA numbering) appropriately positioned to play key roles in hydrolysis.

Because human pancreatic α-amylase plays a crucial role in the digestion of dietary starch, inhibition of its activity has potential therapeutic value in the treatment of diseases such as diabetes or obesity. Indeed, the natural inhibitor Acarbose (Bayer, Montréal, Quebec, Canada) is effective in reducing post-prandial blood glucose levels and is now commercially available for the control of diabetes (Reaven et al., 1990; Matsuo et al., 1992; Ledermann & Hoxter, 1994; Noda et al., 1997). Detailed knowledge of the catalytic mechanism of HPA is essential for the rational design of new drugs to combat these diseases further.

As part of a study of the roles of the three carboxylic acids in the catalytic mechanism of HPA, we report in this paper the expression, purification, and functional characterization of human pancreatic α-amylase and the two mutants D197A and D197N expressed in *Pichia pastoris*. After enzymatic deglycosylation, the *Pichia*-derived human amylase was crystallized in space group $P2_12_12_1$, and the structure determined to 2.0 Å resolution.

Results

Cloning, mutagenesis, and expression of α-amylase in P. pastoris

A human α-amylase cDNA was cloned from human pancreatic mRNA by using a reverse transcriptase–PCR approach. The cDNA sequence was identical to the sequence of human α-amylase reported previously (Nishide et al., 1986) indicating that no mutations had been introduced during the cloning steps. The 5′ region of the amylase cDNA was then modified by in vitro mutagenesis using primers AMY-15 and AMY-16 (Table 1) to allow the subsequent insertion of the cDNA into the NotI site of pPIC9. The final expression vector directed the secretion of an α-factor/human α-amylase fusion protein that is proteolytically processed during secretion from the yeast. After linearizing with AatII and BsaI, the expression vector was used to transform spheroplasts of *P. pastoris* strain GS115. Mutant forms of α-amylase were produced by using the same strategy. As judged by SDS-PAGE, the recombinant α-amylase represented the only protein in the culture medium, and was present at about 20 mg/L.

Structure and glycosylation site

Despite obtaining the recombinant enzymes in modest quantities, crystallization was unsuccessful under a broad range of conditions, including the conditions used for the native enzyme (Brayer et al., 1995). Since *P. pastoris* is known to glycosylate foreign proteins heterogeneously, and in some cases glycosylation is known to interfere with crystallization (Liu et al., 1996; Nersissian et al., 1996; Linnevers et al., 1997), it seemed likely that the recombinant α-amylase was glycosylated. Glycosylation was consistent with the presence of a doublet that migrated within about 5 kDa of each other when analyzed by Coomassie-stained SDS-PAGE and Western blots. These two bands were present in all preparations of the

Table 1. *Oligonucleotides used in this study*

Sense strand[a]		*Use*
AMY-1	5′-ATCCC CGGGT TCAAA GCAAA ATGAA GTTC-3′	Reverse transcription/PCR
AMY-3	5′-TTAGC TATAA TTATT GCA-3′ (nucleotides 251–269)	DNA sequencing primer
AMY-4	5′-AATGT AAAAC TGGAA GTG 3′ (476–494)	DNA sequencing primer
AMY-5	5′-ATAAT CTAAA CAGTA ACT 3′ (700–718)	DNA sequencing primer
AMY-6	5′-TGTAC AAAAT GGCAG TTG-3′ (1,018–1,036)	DNA sequencing primer
AMY-7	5′-AGCCT TTTAC AAATT GGT 3′ (1,271–1,289)	DNA sequencing primer
AMY-13	5′-TTCAG ACTTG CTGCT TCCAA G-3′	Mutagenesis
AMY-14	5′-TTCAG ACTTA ATGCT TCCAA G-3′	Mutagenesis
AMY-15	5′-ACACT CGAGA AAAGA CAGTA TTCCC CAAAT ACACA A-3′	Insertion of cDNA into pPIC9
Nonsense strand		
AMY-2	5′-ACTCC CGGGG CATTT AATTT TAAAT TTTAC-3′	PCR
AMY-8	5′-ACCAA TTTGT AAAAG GCT-3′ (1,288–1,270)	DNA sequencing primer
AMY-9	5′-AGAAG GTAAG AGTAG AGG-3′ (1,006–988)	DNA sequencing primer
AMY-10	5′-AAATG AAAGG TTTAC TTC-3′ (748–730)	DNA sequencing primer
AMY-11	5′-CATCA TTGTA GTTCT CGA-3′ (517–499)	DNA sequencing primer
AMY-12	5′-ATTCA TCTTC ATTTC CAG-3′ (292–274)	DNA sequencing primer
AMY-16	5′-CATTT CCAGA TCTTG TGCAT-3′ (282–263)	Insertion of cDNA into pPIC9

[a] Numbers in parentheses correspond to the positions in the α-amylase cDNA sequence (Nishide et al., 1986).

recombinant enzyme. Schiff staining (Stromqvist & Gruffman, 1992) of the protein after transfer to a PVDF membrane confirmed the presence of protein glycosylation. Electrospray mass spectrometry showed that the recombinant amylase was expressed in a series of forms with masses differing by approximately 165 amu; this is the typical mass for a sugar residue (Fig. 1A). A peak due to the nonglycosylated protein (predicted mass 55,887 Da, observed 55,894 Da) was also observed. This suggests the presence of 10 to 13 sugar residues on the glycosylated enzyme forms.

P. pastoris is known to perform both N- and O-linked glycosylation of foreign proteins (Grinna & Tschopp, 1989; Trimble et al., 1991; Cregg et al., 1993; Juge et al., 1996). Since N-linked glycosylation occurs at a consensus sequence (Asn-X-Ser(Thr), where -X- is any residue except proline) while O-linked glycosylation does not, we first looked for evidence of N-linked glycosylation. Human pancreatic α-amylase has two potential sites for N-glycosylation, Asn412 and Asn461; both of these residues are located on the surface of domain C (Brayer et al., 1995). Deglycosylation with endoglycosidase F, an enzyme that hydrolyzes the GlcNAc β(1–4) GlcNAc glycosidic bond of N-linked high mannosyl groups, resulted in an α-amylase with a mass of 56,088 Da (Fig. 1B). This corresponds to the native enzyme with a single GlcNAc present therefore indicating a single N-linked glycosylation site and no O-linked sugars. These data also confirm that recombinant amylase is heterogeneously N-glycosylated containing, at most, a single N-linked oligosaccharide chain likely ranging from $Man_8GlcNAc_2$ to $Man_{11}GlcNAc_2$ (Grinna & Tschopp, 1989; Cregg et al., 1993).

Samples of untreated and endoglycosidase F-treated recombinant HPA were then denatured with urea, reduced with dithiothreitol (DTT) and their cysteine residues derivatized with iodoacetamide prior to tryptic digestion. A comparison of tryptic digest maps for the two HPA samples revealed a peptide of *m/z* 1,028.5 amu seen only in the treated sample. By Edman degradation, the peptide was shown to have the apparent sequence EDVISGDKINGDETGIK. The nearest HPA sequence to this was CDVISGDKINGNCTGIK. The obvious differences between the two sequences (Glu instead of Cys, and Asp461 replacing Asn461) were a result of the acetamidation of the cysteines yielding a residue for which the PTH derivative chromatographs like Glu, and the presence of the GlcNAc on Asn461 again modifying the chromatographic behavior. This strongly suggested that the glycosylated residue in the recombinant amylase was Asn461.

Upon deglycosylation with endoglycosidase F, we were then able to crystallize the recombinant amylase under the same conditions and in the same space group as the native enzyme (Brayer et al., 1995). However, despite this the unit cell dimensions for the recombinant enzyme differed (Table 2) as did the packing of the molecules in the unit cell. These differences appear to be due to the glycosylation of Asn461. As Figure 2 shows, there is well-defined electron density off the amide nitrogen of Asn461, which fits a GlcNAc sugar residue. No other sites of glycosylation were evident from electron density maps.

Although somewhat different crystal packing arrangements were observed, the folds of the recombinant and native HPA's are comparable with an overall average main-chain deviation of 0.22 Å (Fig. 3). The largest differences occur at residues 304–308, which form a flexible loop near the active site and may be involved in substrate binding. In the structures of both the native and recombinant proteins, these residues have high thermal factors indicating significant polypeptide chain mobility. The next largest differences occur near the glycosylation site at Asn461, which is part of a surface loop region. Both areas of conformational differences are

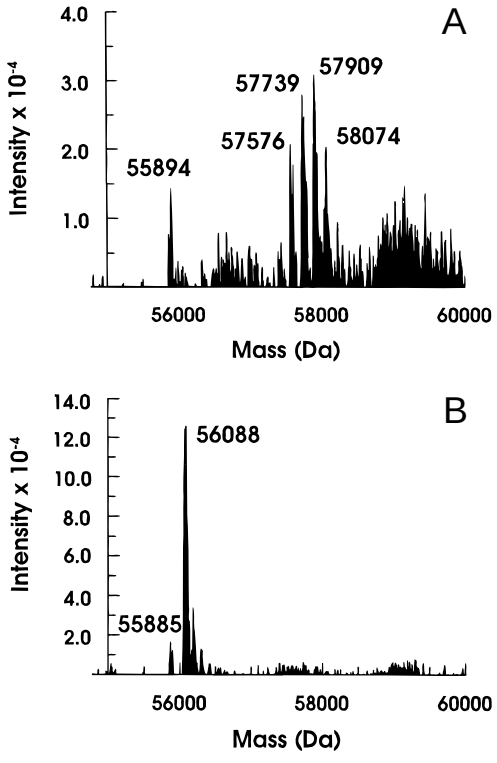

Fig. 1. Electrospray mass spectra of recombinant human pancreatic α-amylase expressed in the yeast *P. pastoris*. The protein is expressed as a heterogeneously glycosylated product (**A**) possessing up to 13 sugar residues. After hydrolysis with endoglycosidase F (**B**) the recombinant α-amylase has a mass indicative of the presence of only one sugar residue (GlcNAc).

Table 2. *Data collection parameters for Pichia-HPA*

Space group	$P2_12_12_1$
Cell dimensions (Å)	
a	52.91
b	68.90
c	131.77
Number of measurements	166,575
Number of unique reflections	33,646
Mean $I/\sigma I$	20.8 (5.8)[a]
Multiplicity	5.0 (2.5)
Merging *R*-factor (%)[b]	7.5 (18.4)
Resolution range (Å)	10.0–2.0

[a] Values in parentheses are for the last resolution shell (2.07 − 2.00 Å).

[b] $R_{merge} = \dfrac{\sum_{hkl}\sum_{i=0}^{n} |I_i - \bar{I}_{hkl}|}{\sum_{hkl}\sum_{i=0}^{n} I_{i_{hkl}}}$.

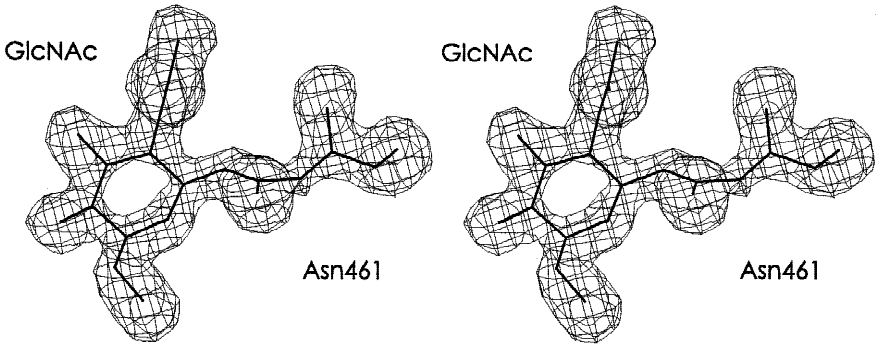

Fig. 2. Stereo drawing of a portion of the $2F_o - F_c$ difference electron density map of *Pichia*-HPA in the vicinity of the N-acetylglucosamine (GlcNAc) residue bound to Asn461. This map has been contoured at the 5σ level and the final refined coordinates for both the NAG moiety and Asn461 are overlaid.

probably the result of mobile loops adapting to a new crystal packing arrangement.

The conformational similarity between the recombinant and native proteins extends to both the chloride and calcium binding sites, as well as the active site region (Fig. 4). The largest active site conformation difference is an approximate 0.3 Å shift in a water molecule bound to Asp197 OD1 (Fig. 4).

Kinetic characterization

The kinetic parameters k_{cat} and K_m for the recombinant α-amylase and the D197A, and D197N mutants were determined using the substrate α-maltotriosyl fluoride (αG3F) (Table 3). Glycosylation did not appear to influence the activity of HPA, as the specificity constant k_{cat}/K_m for αG3F was essentially unchanged upon degly-

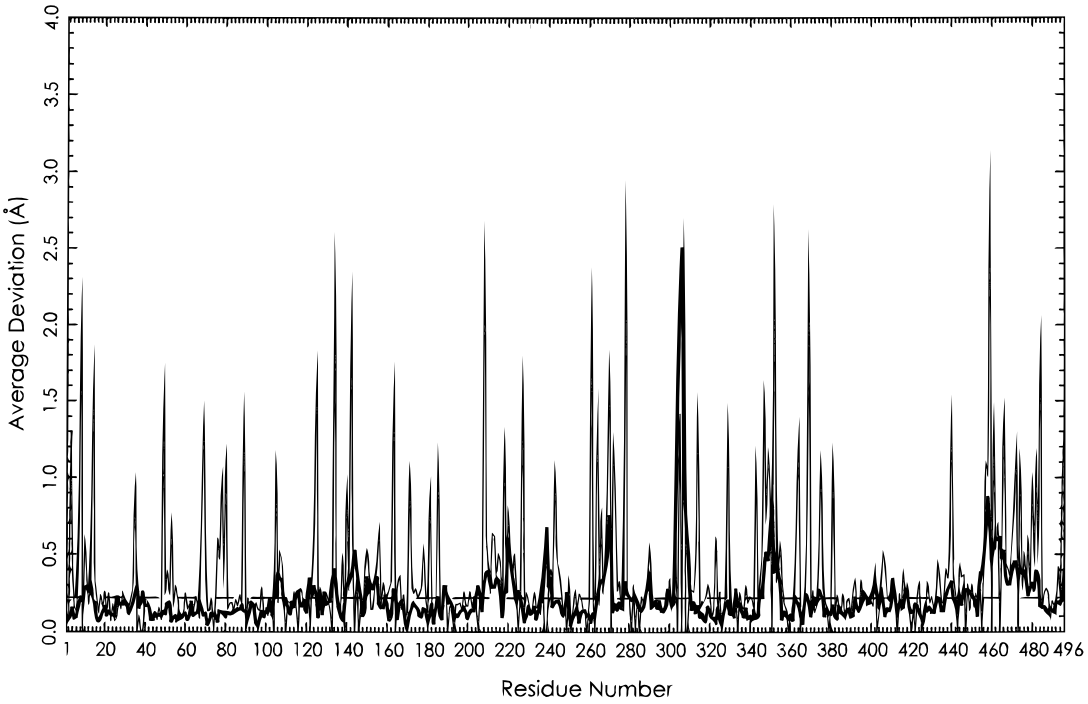

Fig. 3. Plots of the average deviations in positions of main (thick lines) and side (thin lines) chain atoms of *Pichia*-HPA from those of natural human pancreatic α-amylase. The dashed horizontal line indicates the overall average deviation of 0.22 Å found between all main-chain atoms. Large deviations in main-chain atom positions around residue 305 are likely the result of the poor definition of electron density in this region of both forms of this enzyme. Deviations in main-chain atom positions about residue 460 appear to be due to glycosylation of Asn461 and an associated local reorganization in *Pichia*-HPA.

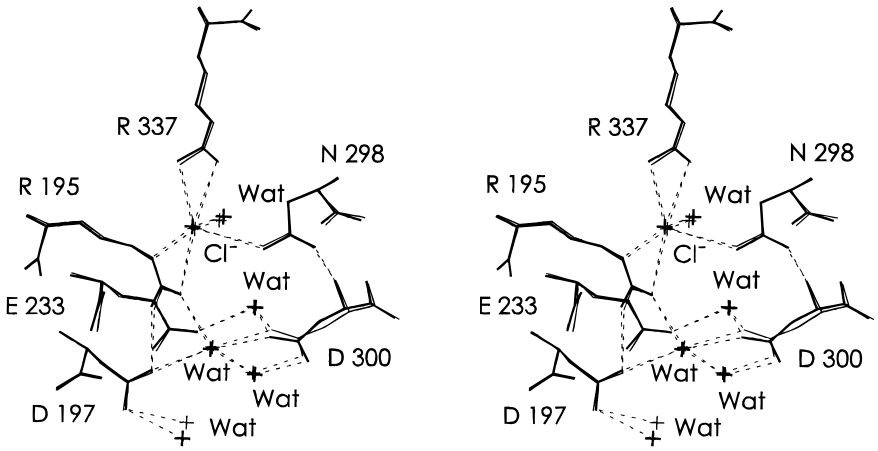

Fig. 4. Stereo drawing of the active site region and chloride ion binding pocket in the structures of *Pichia*-HPA (thick lines) and native (thin lines) human pancreatic α-amylases.

cosylation. The specificity constants for the D197A and D197N mutants were both 10^6–10^7 times lower than that of the wild-type enzyme. The dramatic decrease in specificity constants for the mutants is due in large part to the drastic reduction of the k_{cat} values, since the K_m values change little with αG3F when compared to wild-type enzyme.

Discussion

Although the native HPA is known to exist in several forms resulting from differential modification (including glycosylation) of the proteins produced from two different alleles (Zakowski & Bruns, 1985), the determination of the three-dimensional structure of this enzyme showed no evidence of glycosylation (Brayer et al., 1995). This may, however, be a consequence of selective crystallization of the nonglycosylated form of the protein. Our inability to crystallize the fully glycosylated *P. pastoris*–produced enzymes is certainly consistent with that interpretation, as was the ready crystallization upon removal of all but the GlcNAc residue. Upon initial investigation, it was surprising that glycosylation of the recombinant protein occurs at Asn461 (consensus sequence NCT) and not Asn412 (consensus sequence NGS) since both residues occur in surface loops and contain consensus sequences that allow for high core-glycosylation efficiency (Shakin-Eshleman et al., 1996). Indeed, comparison of the folds of the loops involved would appear to favor glycosylation at Asn412 since this residue is part of an Asx turn, which has been shown to strongly favor N-glycosylation (Imperiali, 1997). In contrast, the loop containing Asn461 is not part of either an Asx or a β-turn. Furthermore, the cysteine at the -X- position of the Asn461 NCT consensus sequence is involved in a disulfide bond with Cys450. While the presence of a disulfide bond does not exclude glycosylation, it has been suggested that this is the reason why N-glycosylation is not commonly seen at consensus sequences with cysteine at the -X- position (Shakin-Eshleman et al., 1996). A major factor in the lack of glycosylation at Asn412 may be the strong hydrogen bond formed between the side-chain amide of this residue and Asp432. In contrast, the glycosylated side chain of Asn461 is found to be highly solvent accessible in native HPA.

Deglycosylation of the recombinant HPA was crucial for successful crystallization of this protein for structural studies. Follow-up analyses indicate that the packing observed between molecules in the crystal lattice would be impossible in the presence of bulky glycosyl groups (10–13 sugar residues). The heterogeneous nature

Table 3. *Kinetic parameters determined for the reaction of recombinant α-amylase expressed in P. pastoris with α-maltotriosyl fluoride*

Enzyme	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (s^{-1} mM^{-1})	Fold decrease in k_{cat}/K_m
*P*HPA-gly[a]	0.39 ± 0.1[b]	320 ± 39	820 ± 310	1.0
*P*HPA-dgly[c]	0.26 ± 0.02	220 ± 5	850 ± 93	0.96
D197A	3.8 ± 0.95	$4.0(\pm0.6)\times10^{-4}$	$1.0(\pm0.4)\times10^{-4}$	7.8×10^6
D197N	1.1 ± 0.22	$7.3(\pm0.6)\times10^{-4}$	$6.6(\pm2)\times10^{-4}$	1.2×10^6

[a] *P*HPA-gly, glycosylated HPA cloned and expressed in *P. pastoris*
[b] Expressed as value ± error of fit of data to model.
[c] *P*HPA-dgly, endoglycosidase F digested HPA cloned and expressed in *P. pastoris*.

of the glycosylation would also be detrimental to crystallization. Endoglycosidase F digestion appears to completely remove all glycosyl residues except the amide-bound GlcNAc, resolving this problem. In addition, the removal of protein heterogeneity also greatly enhanced the intensity of the ESI-MS signal (Fig. 1).

Enzymes in the amylase family contain four short regions of strong sequence similarity, and several residues within these regions are completely conserved throughout all members of this family (Henrissat & Davies, 1997). Notable among these are the three active site carboxylic acids D197, E233, D300—HPA numbering. Several members of this family, including yeast α-glucosidase (McCarter & Withers, 1996), glycogen debranching enzyme (Braun et al., 1996), and the E257A variant of cyclodextrin glucanotransferase (acid/base catalyst mutant) (Mosi et al., 1997) have had their nucleophilic residues labeled using mechanism-based trapping agents. In all three cases, sequence alignment has shown that the nucleophilic residue corresponded to Asp197 of HPA.

In other retaining glycosidases (*Cellulomonas fimi* β-exoglucanase/xylanase (MacLeod et al., 1996), *Agrobacterium* sp. β-glucosidase (Wang et al., 1994), *Escherichia coli* β-galactosidase (Yuan et al., 1994), *Bacillus circulans* cyclodextrin glucanotransferase (Klein et al., 1992)), mutation of the enzymic nucleophile has resulted in decreases in k_{cat} of at least 10^6-fold from the wild-type rate. This dramatic drop is consistent with the key role played by the nucleophile in stabilizing the transition state and in forming the covalent intermediate, as well as in modulating the pK_a of the acid/base catalyst (McIntosh et al., 1996). The Asp197A and Asp197N mutants that we have generated have k_{cat} values 10^5–10^6 fold lower than that of the wild-type enzyme. This dramatic loss of activity supports the hypothesis that Asp197 functions as the enzymic nucleophile.

In conclusion, human pancreatic α-amylase has been successfully cloned, expressed, and purified from *P. pastoris*. Combined with kinetic analyses and the ability to crystallize recombinant wild-type and mutant proteins, several avenues of research become available by which to gain further insight into the mechanistic roles of active site residues of HPA. Such studies have the potential to establish mechanism-based inhibitors of potential therapeutic value in the future.

Materials and methods

General procedures

All buffer chemicals and other reagents were obtained from Sigma Chemical Company (St. Louis, Missouri) unless otherwise noted. All chromatographic resins were purchased from Pharmacia (Baie d'Urfé, Quebec, Canada). Endoglycosidase F was obtained from Boehringer Mannheim (Laval, Quebec, Canada) and an endoglycosidase F-cellulose binding domain fusion protein (endoF-CBD) was a generous gift from Dr. R.A.J. Warren (Department of Microbiology, University of British Columbia). Synthesis of the substrate αG3F has been described previously (Hayashi et al., 1984). All oligodeoxyribonucleotides used in this study were synthesized on an Applied Biosystems (Foster City, California) Model 391 DNA Synthesizer; the oligonucleotide sequences are given in Table 1.

Cloning, expression, and mutagenesis

A sample of human pancreas was supplied by Dr. D.B.C. Ritchie (Department of Medicine, University of Alberta). Total RNA was isolated by using the guanidinium thiocyanate method (Chomczynski & Sacchi, 1987). After denaturation with methyl mercury hydroxide, cDNA was synthesized using oligo AMY-1 (Table 1) as a primer with Superscript (Life Technologies, Burlington, Ontario) using conditions recommended by the supplier. The cDNA was then used as a template for PCR amplification of amylase cDNA using oligos AMY-1 and AMY-2 (Table 1) as primers and Taq polymerase prepared in this laboratory (Desai & Pfaffle, 1995). Thirty cycles of PCR were performed with a denaturation step at 94 °C for 1 min, an annealing step at 47 °C for 1 min and an extension step at 72 °C for 1.5 min. The PCR amplified α-amylase cDNA was cloned into the Sma-I site of BlueScript KS$^+$ (Stratagene, La Jolla, California) using the Sma-I sites in the PCR primers, and transformed into *E. coli* strain DH5α-F′. To ensure the absence of mutations introduced during the PCR step, the complete nucleotide sequence of the α-amylase cDNA was determined on both strands using the specific sequencing primers AMY-3 to AMY-12 listed in Table 1; the sequence was in complete agreement with the corrected sequence of pancreatic α-amylase cDNA (Nishide et al., 1986).

Oligonucleotide-directed mutagenesis of Asp197 was performed by using a PCR-based procedure (Nelson & Long, 1989) and Bluescript-specific flanking oligonucleotides. Mutagenesis was performed on a 230 bp EcoRV-PstI fragment of α-amylase cloned into Bluescript, by using either oligonucleotide AMY-13 or AMY-14 as the mutagenic primer. After mutagenesis, the presence of the desired mutation and the absence of PCR-induced errors were confirmed by DNA sequence analysis, and the fragment was religated into the original α-amylase cDNA prior to introduction into the expression vector.

Recombinant α-amylase and mutant forms were expressed in *P. pastoris* using pPIC9 (InVitrogen Corp., Carlsbad, California) and protocols recommended by the supplier. Wild-type amylase was purified from the culture medium by precipitation with ethanol (40%) followed by adsorption to glycogen (1 mg glycogen per 100 Units activity) (Schramm & Loyter, 1966). After stirring for 10 min at 4 °C, the precipitate was collected by centrifugation (13,000 $\times$ *g*, 5 min, 4 °C). The resulting small white pellet was dissolved in 30 mL of 20 mM phosphate buffer containing 25 mM NaCl, pH 6.9, and incubated overnight at room temperature to allow complete digestion of the glycogen. After dialysis against 100 mM potassium phosphate, pH 7.4, the sample was adjusted to 0.5 M NaCl and loaded onto a column (3 cm $\times$ 2 cm) of Phenyl-Sepharose CL-4B (Pharmacia) equilibrated in 100 mM phosphate buffer pH 7.4. The bound α-amylase was eluted with distilled water, and was pure as judged by SDS-PAGE.

Because saccharogenic assays of the crude mutant proteins indicated that they had very low activity, the affinity precipitation with glycogen could not be used. Instead, the culture medium was adjusted to pH 7.0 with NaOH and filtered through a 0.45 μm membrane (Gelman, Montréal, Quebec, Canada). The filtrate was then loaded onto a Phenyl-Sepharose CL-4B column (3 cm $\times$ 2 cm) equilibrated in 100 mM potassium phosphate buffer pH 7.0. After washing with 100 mM phosphate buffer pH 7.0, the bound protein was eluted with distilled water. The eluant was concentrated using an Amicon stirred cell concentrator and the buffer was adjusted to 20 mM potassium phosphate pH 6.9 containing 25 mM NaCl before being subjected to ion exchange chromatography on a column (2 cm $\times$ 1 cm) of Q-Sepharose (Pharmacia) equilibrated in 20 mM phosphate buffer containing 25 mM NaCl, pH 6.9. The α-amylase obtained from the flow-through was $>$99% pure as determined by SDS-PAGE.

Deglycosylation

Deglycosylation was performed in different trials using endoglycosidase F from two different sources. Prior to Phenyl-Sepharose chromatography, endoglycosidase F (20 μL, 0.05 U/μL, Boehringer Mannheim) was added to a sample of α-amylase (100 mL, 0.1 mg/mL in 100 mM phosphate buffer pH 7.4). Digestion was carried out overnight at room temperature, or for 4 h at 37 °C. The amylase purification was then completed as described above. Alternatively, endoF-CBD (10 μL, 0.5 U/μL,) was added to a solution of purified α-amylase (20 mL, 2 mg/mL, 50 mM phosphate buffer pH 7.4 containing 25 mM NaCl) and the solution was incubated at 37 °C for 4 h or at room temperature overnight. Deglycosylation was confirmed by SDS-PAGE before cellulose (0.5 g) was added to the solution. This suspension was gently mixed for 5 min before centrifugation (3,500 × *g*, 5 min). As determined by SDS-PAGE and ESI-MS, the supernatant contained only deglycosylated α-amylase.

Determination of glycosylation site by mass spectrometry

Samples of both untreated and endo F-treated HPA (900 μL, 1.2 mg/mL) were precipitated for 30 min on ice after addition of 100 μL of 100% TCA. The resulting precipitates were collected by centrifugation (5 min, 8,000 *g*), washed several times with cold acetone, and dried. The powders were each resuspended in 200 μL of 8 M urea containing 0.4 M NH_4HCO_3 to which 20 μL of 45 mM DTT was added. After incubation at 50 °C for 15 min, the mixture was cooled to room temperature, iodoacetamide (20 μL, 100 mM) added, and incubation continued for 30 min at room temperature. The mixture was diluted with 560 μL of distilled water, and an aliquot (200 μL) digested with trypsin (5% w/w) at 30 °C overnight. The reaction was terminated by freezing the solution. The tryptic digest was analyzed by LC-MS (conditions described below) and the tryptic maps of the glycosylated and endoglycosidase F treated amylases were compared. A peptide appearing in the endoglycosidase F-treated sample but not in the untreated sample was isolated, and its sequence was determined by Edman degradation.

Mass determination

The mass determinations of the recombinant proteins and tryptic digests before and after endoglycosidase F-treatment were performed on a PE-Sciex API 300 triple quadrupole mass spectrometer (Sciex, Thornhill, Ontario, Canada) equipped with an Ionspray ion source. Protein was loaded onto a C18 column (Reliasil, 1 × 150 mm) equilibrated with solvent A (solvent A: 0.05% trifluoroacetic acid (TFA), 2% acetonitrile in water). Elution of the peptides was accomplished using a gradient (0–60%) of solvent B over 60 min followed by 100% solvent B over 2 min (solvent B: 0.045% TFA, 80% acetonitrile in water). Solvents were pumped at a constant flow rate of 50 μL/min. Spectra were obtained in the single quadrupole scan mode (LC/MS). The quadrupole mass analyzer was scanned over a mass to charge ratio (*m*/*z*) range of 400–2,400 Da with a step size of 0.5 Da and a dwell time of 1.5 ms per step. The ion source voltage (ISV) was set at 5.5 kV and the orifice energy (OR) was 45 V.

Kinetic assays

The kinetic parameters k_{cat} and K_m for the recombinant proteins were determined by monitoring the release of fluoride from α-maltotriosyl fluoride (αG3F) using an Orion 96-09 combination fluoride ion electrode as described previously (McCarter et al., 1993). The assays were performed at 30 °C in 20 mM sodium phosphate buffer, pH 6.9 containing 25 mM NaCl. Stock enzyme (5 μL of 50 nM wild-type enzyme, or 50 μL of 18 μM mutant enzyme) was added to glass cells containing various concentrations of α-maltotriosyl fluoride in a total assay volume of 300 μL. Controls were performed to compensate for the spontaneous hydrolysis of the substrates. The fluoride released after hydrolysis was detected using an Orion 96-09 combination fluoride ion electrode coupled to a Fischer Accumet 925 pH/ion meter. The ion meter was interfaced with a Pentium 133 MHz personal computer for data collection. Data were fit to the Michaelis–Menten equation using the nonlinear regression program GraFit 3.0 (Erithacus Software, Staines, Middlesex, United Kingdom).

Structural determination of Pichia-derived α-amylase

Crystals of *Pichia*-HPA were grown at room temperature using the hanging drop vapor diffusion method. The reservoir solution contained 60% 2-methylpentan-2,4 diol with 100 mM cacodylate at pH 7.5 and the hanging drops consisted of 5 μL of protein solution (2 mg/mL) mixed with 5 μL of reservoir solution. Diffraction quality crystals appeared after 2 weeks. Diffraction data for *Pichia*-HPA were collected on a Rigaku R-AXIS IIC imaging plate area detector system using Cu K_α radiation supplied by a Rigaku RU 300 rotating anode generator operating at 50 kV and 100 mA. The data collection statistics are shown in Table 2. The data were processed with the HKL suite of programs (Otwinowski & Minor, 1997). Since crystals of *Pichia*-HPA were not isomorphous to those of the native amylase, a molecular replacement approach was taken to the solution of this structure using the AMoRe package (Navaza, 1994) and the native amylase structure as a search model. Using data from 8 to 4 Å resolution, a single clear solution was obtained that gave a conventional crystallographic *R*-factor of 27.3%. Refinement of the structural model of *Pichia*-HPA was accomplished with X-PLOR (Brünger, 1992) software wherein simulated annealing and conjugate gradient minimization protocols were alternated with manual

Table 4. *Refinement statistics for Pichia-HPA*

Number of reflections	31,856
Resolution range (Å)	10–2.0
Completeness within range (%)	95.4 (85.9)[a]
Number of protein atoms	3,961
Number of solvent atoms	221
Average thermal factors (Å^2)	
Protein atoms	16.3
Solvent atoms	36.4
Final refinement *R*-factor (%)[b]	18.4
Stereochemistry	RMS deviations
Bonds (Å)	0.008
Angles (°)	1.450

[a] Values in parentheses are for the last resolution shell (2.07 − 2.00 Å).

[b] $R_{merge} = \dfrac{\sum_{hkl} |F_o - F_c|}{\sum_{hkl} |F_o|}$.

model rebuilding with O (Jones et al., 1991). The entire polypeptide chain of *Pichia*-HPA was examined periodically during this process with $F_o - F_c$, $2F_o - F_c$, and fragment deleted difference electron density maps. During such examinations, solvent molecule peaks were assigned and an N-acetylglucosamine moiety built into observed electron density extending from the side chain of Asn461. The validity of solvent molecules was assessed based on both hydrogen bonding potential to protein atoms and the refinement of a thermal factor of <75 Å^2. The final *R*-factor for 31,856 reflections from 10 to 2.0 Å resolution is 18.4% (Table 4) while the coordinate error as estimated from a Luzzati plot (Luzzati, 1952) is 0.20 Å. The final refinement statistics of the model are shown in Table 4. Coordinates for this structure have been deposited in the Protein Data Bank (Bernstein et al., 1997) with access code 1bsi.

Acknowledgments

We would like to thank Dr. R.A.J. Warren for his generous gift of the endoglycosidase F-cellulose binding domain fusion protein and Mr. Shouming He for obtaining the mass spectra. Dr. C.M. Overall is funded by a Clinician Scientist Award from the Medical Research Council of Canada (MRC). Mr. E.H. Rydberg and Mr. S. Numao are supported by MRC studentships. Mr. G. Sidhu is supported by an NSERC-PGSA studentship. This project was funded by operating grants from the Medical Research Council of Canada.

References

Bernstein FC, Koetzle TF, Williams GJ, Meyer EE Jr, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M. 1997. The Protein Data Bank: A computer-based archival file for macromolecular structures. *J Mol Biol 112*:535–542.

Bompard-Gilles C, Rousseau P, Rouge P, Payan F. 1996. Substrate mimicry in the active center of a mammalian alpha-amylase: Structural analysis of an enzyme-inhibitor complex. *Structure 4*:1441–1452.

Brady RL, Brzozowski AM, Derewenda ZS, Dodson EJ, Dodson GG. 1991. Solution of the structure of *Aspergillus niger* acid α-amylase by combined molecular replacement and multiple isomorphous replacement methods. *Acta Crystallogr B47*:527–535.

Braun C, Lindhorst T, Madsen N, Withers SG. 1996. Identification of Asp549 as the catalytic nucleophile of glycogen debranching enzyme via trapping of the glycosyl-enzyme intermediate. *Biochemistry 35*:5458–5463.

Braun C, Meinke A, Ziser L, Withers SG. 1993. Simultaneous high performance liquid chromatographic determination of both cleavage pattern and the stereochemical outcome of the hydrolysis reactions catalysed by glycosidases. *Anal Biochem 212*:259–262.

Brayer GD, Luo YG, Withers SG. 1995. The structure of human pancreatic alpha-amylase at 1.8 angstrom resolution and comparisons with related enzymes. *Protein Sci 4*:1730–1742.

Brünger AT. 1992. *X-PLOR: A system for X-ray crystallography and NMR*. New Haven, Connecticut: Yale University Press.

Brzozowski AM, Davies GJ. 1997. Structure of the *Aspergillus oryzae* α-amylase complexed with the inhibitor acarbose at 2.0 Å resolution. *Biochemistry 36*:10837–10845.

Buisson G, Duee E, Haser R, Payan F. 1987. The three-dimensional structure of porcine pancreatic α-amylase at 2.9 Å resolution. Role of calcium in structure and activity. *EMBO J 6*:3909–3916.

Chomczynski P, Sacchi N. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem 162*:156–159.

Cregg JM, Vedvick TS, Raschke WC. 1993. Recent advances in the expression of foreign genes in *Pichia pastoris*. *Biotechnology 11*:905–910.

Desai UJ, Pfaffle PK. 1995. Single-step purification of a thermostable DNA polymerase expressed in *Escherichia coli*. *BioTechniques 19*:780–784.

Grinna LS, Tschopp JG. 1989. Size distribution and general structural features of N-linked oligosaccharides from the methylotrophic yeast, *Pichia pastoris*. *Yeast 5*:107–115.

Hayashi M, Hashimoto S, Noyori R. 1984. Simple synthesis of glycosyl fluorides. *Chem Lett*:1747–1750.

Henrissat B. 1991. A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J 280*:309–316.

Henrissat B, Callebaut I, Fabrega S, Lehn P, Mornon JP, Davies G. 1995. Conserved catalytic machinery and the prediction of a common fold for several families of glycosyl hydrolases. *Proc Natl Acad Sci USA 92*:7090–7094.

Henrissat B, Davies G. 1997. Structural and sequence-based classification of glycoside hydrolases. *Curr Opin Struct Biol 7*:637–644.

Imperiali B. 1997. Protein glycosylation: The clash of the titans. *Acc Chem Res 30*:452–459.

Jones TA, Zhou J-Y, Cowan SW, Kjeldgaard M. 1991. Improved methods for building protein models in electron density maps and the location of errors in these maps. *Acta Crystallogr A47*:110–119.

Juge N, Andersen JS, Tull D, Roepstorff P, Svensson B. 1996. Overexpression, purification and characterization of recombinant barley α-amylases 1 and 2 secreted by the methylotrophic yeast *Pichia pastoris*. *Protein Express Purif 8*:204–214.

Kadziola A, Abe J, Svensson B, Haser R. 1994. Crystal and molecular structure of barley alpha-amylase. *J Mol Biol 239*:104.

Klein C, Hollender J, Bender H, Schulz GE. 1992. Catalytic center of cyclodextrin glycosyltransferase derived from X-ray structure analysis combined with site-directed mutagenesis. *Biochemistry 31*:8740–8746.

Ledermann H, Hoxter G. 1994. Effect of acarbose on postprandial increase in blood glucose. Additive acute effect of once daily administration in insulin treated diabetes. *Fortschritte der Medizin 112*:467–470.

Linnevers CJ, McGrath ME, Armstrong R, Mistry FR, Barnes MG, Klaus JL, Palmer JT, Katz BA, Bromme D. 1997. Expression of human cathepsin K in *Pichia pastoris* and preliminary crystallographic studies of an inhibitor complex. *Protein Sci 6*:919–921.

Liu J, Tse AGD, Chang H-C, Liu J-H, Wang J, Hussey RE, Chishti Y, Rheinhold B, Spoerl R, Nathenson SG, Sacchettin JC, Reinherz EL. 1996. Crystallization of a deglycosylated T cell receptor (TCR) complexed with an anti-TCR fab fragment. *J Biol Chem 271*:33639–33646.

Luzzati PV. 1952. Traitement statistique des erreurs dans la dètermination des structures cristallines. *Acta Crystallogr 5*:803–810.

MacLeod AM, Tull D, Rupitz K, Warren RAJ, Withers SG. 1996. Mechanistic consequences of mutation of active site carboxylates in a retaining beta-1,4-glycanase from *Cellulomonas fimi*. *Biochemistry 35*:13165–13172.

Matsuo T, Odaka H, Ikeda H. 1992. Effect of an intestinal disaccharidase inhibitor (AO-128) on obesity and diabetes. *Am J Clin Nutr 55*:314S–317S.

McCarter J, Adam M, Braun C, Namchuk M, Tull D, Withers SG. 1993. Syntheses of 2-deoxy-2-fluoro-mono- and oligosaccharide glycosides from glycals and evaluation as glycosidase inhibitors. *Carbohyr Res 249*:77–90.

McCarter J, Withers SG. 1994. Mechanisms of enzymatic glycoside hydrolysis. *Curr Opin Struct Biol 4*:885–892.

McCarter JD, Withers SG. 1996. Unequivocal identification of Asp-214 as the catalytic nucleophile of saccharomyces cerevisiae alpha-glucosidase using 5-fluoro glycosyl fluorides. *J Biol Chem 271*:6889–6894.

McIntosh LP, Hand G, Johnson PE, Joshi MD, Korner M, Plesniak LA, Ziser L, Wakarchuk WW, Withers SG. 1996. The pK_a of the general acid/base carboxyl group of a glycosidase cycles during catalysis—A C-13-NMR study of *Bacillus circulans* xylanase. *Biochemistry 35*:9958–9966.

Mosi R, He S, Uitdehaag J, Dijkstra B, Withers S. 1997. Trapping and characterization of the reaction intermediate in cyclodextrin glycosyltransferase by use of activated substrates and a mutant enzyme. *Biochemistry 36*:9927–9934.

Nakajima R, Imanaka T, Aiba S. 1986. Comparison of amino acid sequences of eleven different α-amylases. *Appl Microbiol Biotechnol 23*:355–360.

Navaza J. 1994. AMoRe: An automated package for molecular replacement. *Acta Crystallogr A50*:157–163.

Nelson RM, Long GL. 1989. A general method of site-specific mutagenesis using a modification of the *Thermus aquaticus* polymerase chain reaction. *Anal Biochem 180*:147–151.

Nersissian AM, Mehrabian ZB, Nalbandyan RM, Hart PJ, Fraczkiewicz G, Czernuszewicz RS, Bender CJ, Peisach J, Herrmann RG, Valentine JS. 1996. Cloning, expression, and spectroscopic characterization of *Cucumis sativus* stallacyanin in its nonglycosylated form. *Protein Sci 5*:2184–2192.

Nishide T, Emi M, Nakamura Y, Matsubara K. 1986. Corrected sequences of cDNAs for human salivary and pancreatic α-amylases. *Gene 50*:371–372.

Noda K, Umeda F, Nawata H. 1997. Effect of acarbose on glucose intolerance in patients with non-insulin-dependent diabetes mellitus. *Diabet Res Clin Prac 37*:129–136.

Otwinowski Z, Minor W. 1997. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol 276*:307–326.

Qian M, Haser R, Buisson G, Duee E, Payan F. 1994. The active centre of a mammalian α-amylase. Structure of the complex of a pancreatic α-amylase with a carbohydrate inhibitor refined to 2.2 Å resolution. *Biochemistry 33*:6284–6294.

Qian M, Spinelli S, Driguez H, Payan F. 1997. Structure of a pancreatic α-amylase

bound to a substrate analogue at 2.03 Å resolution. *Protein Sci 6*:2285–2296.

Ramasubbu N, Paloth V, Luo YG, Brayer GD, Levine MJ. 1996. Structure of human salivary alpha-amylase at 1.6-angstrom resolution—Implications for its role in the oral cavity. *Acta Crystallogr Sect D Biol Crystallogr D52*:435–446.

Reaven GM, Lardinois CK, Greenfield MS, Schwartz HC, Vreman HJ. 1990. Effect of acarbose on carbohydrate and lipid metabolism in NIDDM patients poorly controlled by sulfonylureas. *Diabet Care 13*:32–36.

Schramm M, Loyter A. 1966. Purification of α-amylases by precipitation of amylase-glycogen complexes. *Methods Enzymol VIII*:533–537.

Shakin-Eshleman SH, Spitalnik SL, Kasturi L. 1996. The amino acid at the X position of an Asn-X-Ser sequon is an important determinant of N-linked core-glycosylation efficiency. *J Biol Chem 271*:6363–6366.

Sinnott ML. 1990. Catalytic mechanisms of enzymic glycosyl transfer. *Chem Rev 90*:1171–1202.

Stromqvist M, Gruffman H. 1992. Periodic acid/Schiff staining of glycoproteins immobilized on a blotting matrix. *BioTechniques 13*:744–746.

Suzuki A, Yamane T, Ito Y, Nishio T, Fujiwara H, Ashida T. 1990. Crystallization and preliminary crystallographic study of bacterial α-amylases. *J Biochem 108*:379–381.

Swift HJ, Brady L, Derewenda ZS, Dodson EJ, Dodson GG, Turkenburg JP, Wilkinson AJ. 1991. Structure and molecular model refinement of *Aspergillus oryzae* (TAKA) α-amylase: An application of the simulated-annealing method. *Acta Crystallogr B47*:535–544.

Trimble RB, Atkinson PH, Tschopp JF, Townsend RR, Maley F. 1991. Structure of oligosaccharides on *Saccharomyces* SUC2 invertase secreted by the methylotrophic yeast *Pichia pastoris*. *J Biol Chem 34*:22807–22817.

Wang Q, Graham RW, Trimbur D, Warren RAJ, Withers SG. 1994. Changing enzymatic reaction mechanisms by mutagenesis: Conversion of a retaining glucosidase to an inverting enzyme. *J Am Chem Soc 116*:11594–11595.

Wiegand G, Epp O, Huber R. 1995. The crystal structure of porcine pancreatic alpha-amylase in complex with the microbial inhibitor tendamistat. *J Mol Biol 247*:99–110.

Withers SG, Aebersold R. 1995. Approaches to labeling and identification of active site residues in glycosidases. *Protein Sci 4*:361–372.

Yuan J, Martinez-Bilbao M, Huber RE. 1994. Substitutions for Glu-537 of the β-galactosidase from *Escherichia coli* cause large decreases in catalytic activity. *Biochem J 299*:527–531.

Zakowski JJ, Bruns DE. 1985. Biochemistry of human alpha amylase isoenzymes. *Crit Rev Clin Lab Sci 21*:283–322.