Chapter 1

α-L-Aarabinofuranosidase and β-D-xylosidase (Abf/Xyl)

Plant cell walls, the major reservoir of fixed carbon in nature, have three major polymeric constituents: cellulose, hemicellulose and lignin [1]. The amount of xylan, which is the predominant hemicellulose, varies in different plants, from as much as 35% of the dry weight of birchwood to as little as 7% in some gymnosperm [2]. Besides terrestrial plants, in which xylans are based on a β-1,4-linked D-xylosyl backbone, marine algae synthesize xylans of different chemical structure, based on a β-1,3-linked D-xylosyl backbone [3]. Those containing a mixture of β-1,3 and β-1,4 linkages are found in seaweeds such as *Rhodymenia palmata* [4, 5]. In some species of the *Chlorophecae* and the *Rhodophceae*, where cellulose is absent, xylans form a highly crystalline fibrillar material [6].

Xylan is the most common backbone structure found in hemicellulose. However, the hemicellulose structure also has branched hetero-polysaccharides, which require a more complex battery of enzymes to achieve hydrolysis. A concerted action of enzymes which randomly cleave inter-monomeric bonds (endo-enzymes), enzymes which remove monomers from the end of the chain (exo-enzymes), and enzymes which hydrolyse dimers is needed. These enzymes include endo-1,4-β-xylanase (1,4-β-D-xylan xylanohydrolase; EC 3.2.1.8), β-xylosidase (1,4-β-D-xylan xylohydrolase; EC 3.2.1.37), L-glucuronidase, α-L-arabinofuranosidase (EC 3.2.1.55) and acetyl xylan esterase (Figure 1-1). A concerted action of hemicellulolytic enzymes is needed in order to achieve a complete degradation of branched acetyl xylan to carbohydrate monomers [7-9].

Types of α-L-arabinofuranosidases

α-L-Arabinofuranosidases are exo-type enzymes, which hydrolyze terminal nonreducing residues from arabinose-containing polysaccharides. These enzymes catalyze the hydrolysis of α-1,2 , α-1,3 and α-1,5 L-arabinofuranosidic bonds in hemicelluloses such as arabinoxylan, L-arabinan, and other L-arabinose-containing polysaccharides [10-13]. There is another group of enzymes which degrade arabinan by endo-fashion and are called endo-1,5-α-L-arabinanases (1,5-α-L-arabinan 1,5-α-L-arabinanohydrolase, EC 3.2.1.99). The α-L-arabinofuranosidases have been classified into five families of glycanases (families 3, 43, 51, 54, and 62) on the basis of amino acid sequence similarities [14]. In general, the enzymes in GH43 and GH62 hydrolyse unsubstituted α-1,5-arabinan/α-1,5-arabino-oligosaccharides [15, 16] and arabinoxylans, respectively [17], whereas the GH51 and GH54 arabinofuranosidases are believed to exhibit more relaxed substrate specifcity [18]. The two families (51 and 54) of α-L-arabinofuranosidases also differ in substrate specificity for arabinose-containing polysaccharides. Beldman et al. [19] have classified AFases into three types depending on their mode of action and substrate specificity. Type-A α-L-arabinofuranosidases preferentially degrade α-1,5-L-arabinofurano-oligosaccharides to monomeric arabinose and are inactive towards arabinosyl linkages present in polysaccharides [20,21]. In contrast, type-B α-L-arabinofuranosidases show activity that debranches L-arabinose residues from side chains of arabinan or arabinoxylan [22, 23]. Both types of α-L-arabinofuranosidases also act on synthetic p-nitrophenyl-α-L-arabinofuranoside. The third type of α-L-arabinofuranosidases, which are called α-L-arabinofuranohydrolases, are specifically active on arabinosidic linkages in arabinoxylans from oat spelts, wheat, or barley [24, 25]. Type-A α-L-arabinofuranosidase (abfA) [18], type-B α-L-arabinofuranosidase (abfB) [26], and arabinoxylan

arabinofuranohydrolase (axhA) [27] in *Aspergillus niger* are classified into GH51, GH54, and GH62, respectively. The N-terminal sequence of the first 50 amino acids of α-L-arabinofuranosidase from *Bacillus stearothermophilus* T-6 showed high homology with the N-terminal region of α-L-arabinofuranosidase from *Streptomyces lividans* 66 [28, 29]. Another type of arabinan-degrading enzyme has been reported from *Erwinia carotovora* [30]. It released arabinose trisaccharides from the nonreducing end of 1,5-α-L-arabinan, and did not show activity on synthetic α-L-arabinosides.

Physico-chemical characteristics of α-L-arabinofuranosidases

Badal C. Saha summarized the properties of some α-L-arabinofuranosidases (Appendix 1). Multiple forms of α-L-arabinofuranosidase have been detected in the culture broth of *A. awamori* [31], *A. nidulans* [32], *A. niger* [23], *A. terreus* [33], and *P. capsulatum* [34]. The α-L-arabinofuranosidase I and α-L-arabinofuranosidase II purified from the culture filtrate of *A. awamori* had MWs of 81,000 and 62,000 and pIs of 3.3 and 3.6, respectively [31]. The α-L-arabinofuranosidase from *A. pullulans* is a homodimer with an apparent native MW of 210,000 and a subunit MW of 105,000 [35]. One α-L-arabinofuranosidase from *P. purpurogenun* is a monomer of 58 kDa, with a pI of 6.5 [36]. The α-L-arabinofuranosidase (MW 110,000) from *B. stearothermophilus* L1 consists of two subunits (MWs 52,500 and 57,500), that (MW 256,000) from *B. stearothermophilus* T-6 consists of four identical subunits (MW 64,000) [37, 28].

α-L-Arabinofuranosidases have a broad range of pH and temperature dependence, with optimal activities occurring between pH 3.0–6.9 and 40–75 °C [30, 34, 35, 37]. The purified enzyme from *Rhodotorula flava* is highly

acid stable, retaining 82% of its activeity after being maintained for 24 h at pH 1.5 and at 30 °C [38]. Optimum activity of this enzyme is at pH 2.0. The α-L-arabinofuranosidase from *Corticium rolfsii* had an optimum activity at pH 2.5 toward beet arabinan [39]. The α-L-arabinofuranosidase from *Talaromyces emersonii* is a dimer (MW 210 000), with a pH and temperature optima of 3.2 and 70 °C, respectively [40]. In general, the optimal activity of Abf is at the acidic condition.

Application in biotechnology

Hemicellulases have attracted much attention in recent years because of their potential industrial use in biobleaching of paper pulp, bioconversion of lignocellulose material to fermentative products and for the improvement of animal feedstock digestibility [12, 37, 46, 47]. α-L-Arabinofuranosidases are also capable of hydrolyzing the glycosides of monoterpenes, sesquiterpenes and other alcohols which constitute the aromatic potential of wine, and their exploitation in favor improvement and wine aromatization has been studied in the last several years [48-50].

Molecular biology of α-L-arabinofuranosidases

The gene encoding α-L-arabinofuranosidase and β-D-xylosidase from *T. reesei* RutC-30 was cloned and expressed in *S. cerevisiae* [41]. The deduced amino acid sequence of α-L-arabinofuranosidase displays high-level similarity to the α-L-arabinofuranosidase B of *Aspergillus niger,* and the two can form a new family of glycosyl hydrolases. Both enzymes produced in yeasts displayed hydrolytic properties similar to those of the corresponding enzymes purified from *T. reesei.*

Kimura et al. [42] cloned the α-L-arabinofuranosidase gene from *A. sojae* was expressed in *E. coil* BL21 (DE3) pLysS as a cellulose-binding domain tag

fusion protein.   The deduced amino acid sequence of the catalytic domain of the mature enzyme exhibits extensive identity with the catalytic domains of *S. coelicolor* (74%), *A. niger* (75%), *S. lividans* (74%), and *A. tubingensis* (75%), which are enzymes that belong to family 62 of the glycosyl hydrolases.

Debeche et al. [43] cloned and expressed in *Escherichia coil* an α-L-arabinofuranosidase (*abf*D3) from *Thermobacillus xylanilyticus*.   The recombinant α-L-arabinofuranosidase (56,071-Da) could be assigned to family 51 of the glycosyl hydrolase classification system [14].   The enzyme is localized within a distinct phylogenic cluster that contains three other α-L-arabinofuranosidases from taxonomically related bacterial sources [*B. subtilis* [44], *C. stercorarium* [45], and *B. stearothermophilus* [28]].   The α-L-arabinofuranosidase I from *Streptomyces chartreusis* belongs to family 51, whereas α-L-arabinofuranosidase II from the same organism belongs to family 43 of the glycoside hydrolase family [15].   Interestingly, the Abf of family 54 are much less reported.
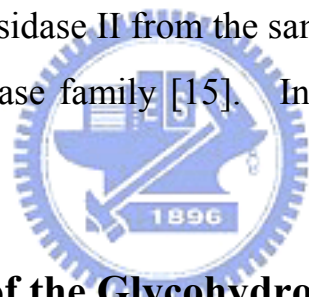
Catalytic mechanism of the Glycohydrolases

Glycosidases have been assigned to families on the basis of sequence similarities, there are now 106 such families defined containing over 2500 different enzymes [14].   This classification has been best described and demonstrated on the Web page (http://afmb.cnrs-mrs.fr/ _cazy/CAZY/index.html).   These groups of enzymes cleave the glycosidic bond of the substrate in two different manners: retention and inversion of the anomeric configuration.   Both mechanisms involve general acid-base catalysis and require two essential residues, which in most glycosyl hydrolases are aspartate and/or glutamate residues (Figure 1-2).   Inverting glycosidases catalyse hydrolysis of the glycosidic bond via a single displacement reaction.   One of the catalytic residues, acting as a general acid, provides protonic assistance to the departing glycosidic oxygen while the

second, acting as a general base, activates a water molecule which effects a direct displacement at the anomeric centre [51]. Retaining glycosidases, perform hydrolysis via a two-step, double-displacement mechanism. Two key active site carboxylic acid residues [52-55], one functioning as the nucleophile and the other as the general acid/base [53], are involved.

α-L-arabinofuranosidases are assigned to five glycoside hydrolase families (GHs) 3, 43, 51, 54 and 62, based on sequence similarities [14, 56]. The GH3, GH51 and GH54 were shown to cleave the glycosidic bond with a retention of the anomeric configuration [57]. The GH43, which includes α-L-arabinofuranosidases as well as β-D-xylosidases, was shown to work via the inverting mechanism [57], whereas the stereochemistry of GH62 is not yet characterized. Different α-L-arabinofuranosidases from bacterial, fungal and plant sources (Appendix 1) can hydrolyze arabinofuranose moieties at O-5, O-2 and/or O-3 as single substituent, as well as from O-2 and O-3 doubly substituted xylans and xylo-oligomers and arabinans [15, 25, 31, 58, 59]. However, very little is known about the exact biochemical mechanism and the specific catalytic residues involved in the reaction mechanism. One of the major reasons, in the case of α-L-arabinofuranosidases, is the difficulty in obtaining synthetic substrates bearing different leaving groups, which could act as an essential tool in the identification of catalytic residues of glycosidases [60, 61].

Several strategies for the identification of such residues in other glycosyl hydrolases have been described. For retaining glycosidases, fluorinated glycoside inhibitors coupled with mass spectral analyses have proved to be a powerful strategy [62, 63]. However, when a recombinant enzyme is available, a simpler strategy, based on site-directed mutagenesis and kinetic analysis can also provide an accurate identification of the catalytic residues [64]. Therefore, Takoua Debeche et al. described the use of such an approach for the experimental identification of the essential catalytic residues

of AbfD3 from *Thermobacillus xylanilyticus* is an arabinoxylan-debranching enzyme which belongs to family 51 of the glycosyl hydrolase classification [65].

Crystal structures of native or mutant enzymes in complex with substrates, products, non-hydrolyzable substrate-analogues and transition-state analogues provide valuable detailed information regarding the specificity, binding mechanism and transition-state stabilization in retaining glycosidases, as recently reviewed by Vasella et al. [66]. The α-L-arabinofuranosidase from *Geobacillus stearothermophilus* T-6 (AbfA) belongs to the retaining GH-51 family, and its catalytic residues were recently identified, Glu175 is the acid/base, and Glu294 is the nucleophile [67, 68]. Klaus Hovel described high-resolution (1.2-2.0 Å) crystal structures of native and catalytic mutant of AbfA in complex with different substrates. The enzyme is a hexamer, and each monomer is organized into two domains: a $(\beta/\alpha)_8$-barrel and a 12-stranded β sandwich with jelly-roll topology. These structures include the Michaelis complexes with natural and synthetic substrates and the transient covalent arabinofuranosyl-enzyme intermediate with a non-fluorinated substrate. The structures allow thorough examination of the catalytic mechanism, including the two stable states of the glycosylation step, the interactions mediating substrate distortion and the features governing substrate specificity [69].

Recently, the potential for mechanistic studies on the family 54 hydrolases has increased greatly, owing to the complete resolution of the three-dimensional structure of bifunctional β-D-xylopyranosidase and α-L-arabinofuranosidase from *Aspergillus kawachii* IFO4308 (Figure 1-3) [70]. On the basis of this structure and very preliminary mutagenic studies, Glu-221 and Asp-297 were suggested as candidates for the nucleophile and the general acid/base catalytic residues, respectively. Thirteen enzymes with phylogenetic similarity were chosen and compared. The result of the amino

acid sequence alignment is shown in Appendix 2. Twenty seven residues [D25, D60, D80, D95, D120, D160, E163, D170, D181, E186, D191, E198, D221, E223, D238, E289, D299, E310, D320, E323, E407, D410, D429, E436, D437, D482 and D490 (numbering in *abf* sequence)] are found to be highly conserved. As compared with the sequence of the *Aspergillus kawachii* IFO4308 α-L-arabinofuranosidase and the inspected X-ray structure, D219, E221 and D297 in *Aspergillus kawachii* IFO4308 enzyme, corresponding to D221, E223 and D299 in *Trichoderma koningii* G-39 enzyme are present in the active site. The function of Glu-223 as a nucleophile, which is conserved in the `DXE' (Glu-X-Asp) sequences of family 54 enzymes, has not yet been confirmed by means of site-directed mutagenesis. In order to gain a better understanding of the detailed mechanism and of the active-site topology of the family 54 hydrolases, an expressed and purified enzyme is essential. The bifunctional β-D-xylopyranosidase and α-L-arabinofuranosidase from *Trichoderma koningii* G-39 is one such suitable candidate. Here we report the development of a recombinant *P. pastoris* clone for the expression of recombinant *abf* gene and combine physical and chemical studies to investigate the catalytic mechanism of the enzyme, and propose Glu-223 and Asp-299 as the essential nucleophile and general acid/base, respectively, as determined by site-directed mutagenesis and kinetic study.
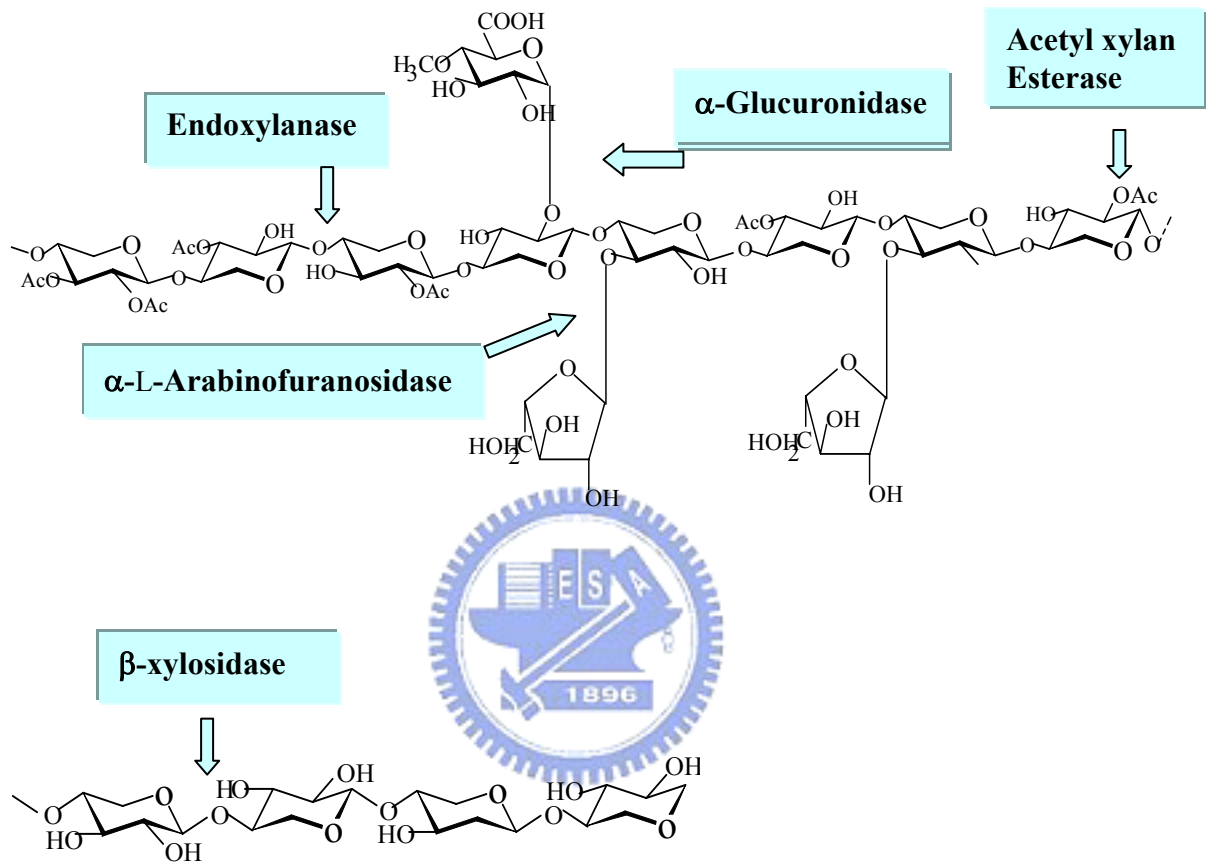
Figure 1-1. The basic structural components of xylan, and the hemicellulases responsible for its degradation.
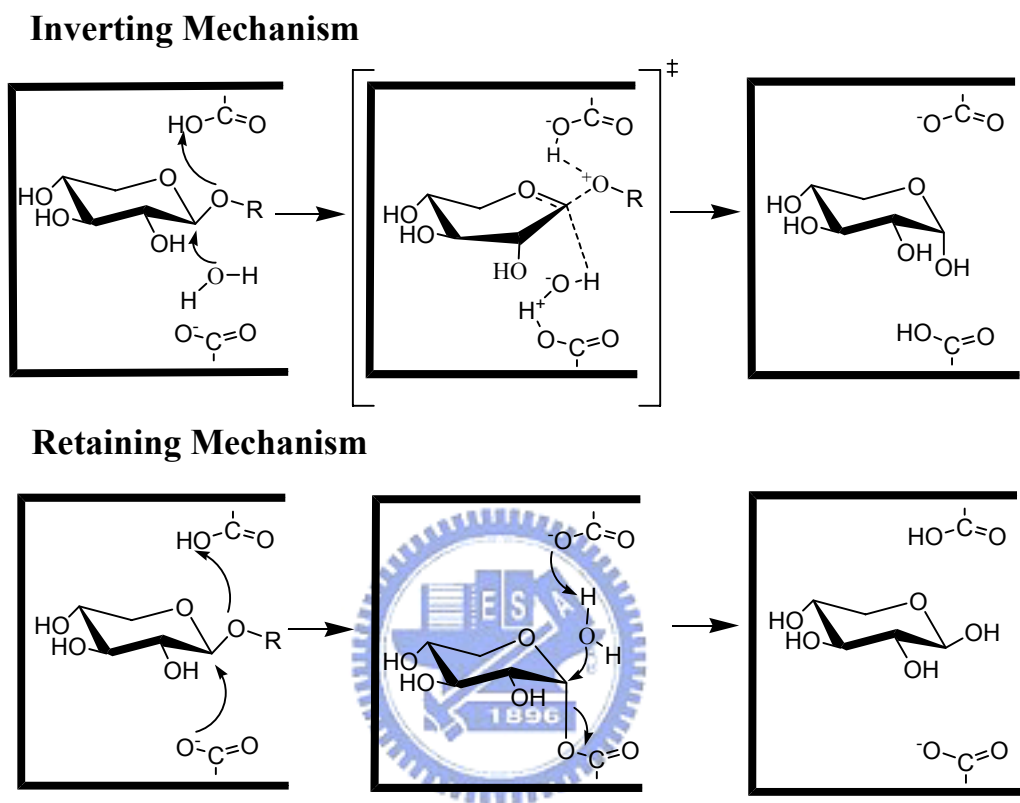
Inverting Mechanism

Retaining Mechanism

Figure 1-2. Catalytic mechanism of the Glycohydrolases
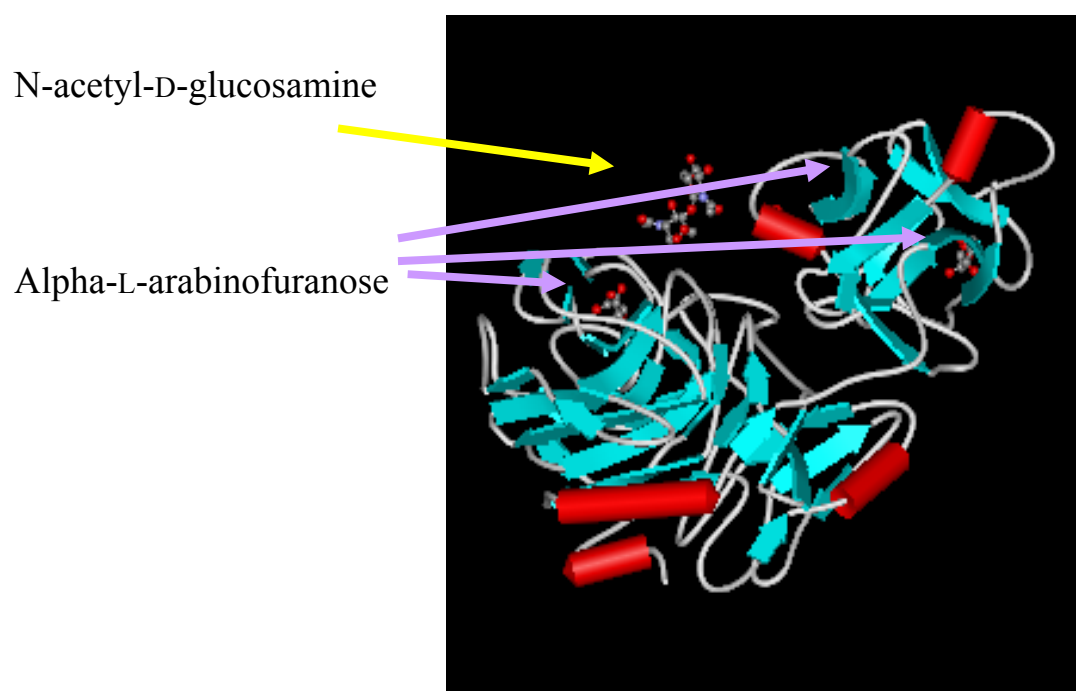
N-acetyl-D-glucosamine

Alpha-L-arabinofuranose

Figure 1-3. The three-dimensional structure of bifunctional α-L-arabinofuranosidase and β-D-xylopyranosidase from *Aspergillus kawachii* IFO4308.

Chapter 2

Expression, purification and characterization of a bifunctional α-ʟ-arabinofuranosidase/β-ᴅ-xylosidase from *Trichoderma koningii* G-39

Abstract

A gene of α-ʟ-Arabinofuranosidase (Abf) from *Trichoderma koningii* G-39 was successfully expressed in *Pichia pastoris*. The recombinant enzyme was purified to > 90% homogeneity by a cation-exchanged chromatography. The purified enzyme exhibits both α-ʟ-arabinofuranosidase and β-ᴅ-xylosidase (Xyl) activities with p-nitrophenyl-α-ʟ-arabinofuranoside (pNPAF) and 2,4-dinitrophenyl-β-ᴅ-xylopyanoside (2,4-DNXP) as substrate, respectively. The stability and the catalytic feature of the bifunctional enzyme were characterized. The enzyme was stable for at least 2 h at pH values between 2 and 8.3 at room temperature when assayed for Abf and Xyl activities. Enzyme activity decreased dramatically when the pH exceeded 9.5 or dropped below 1.5. The enzyme lost 35% of Abf activity after incubation at 55 °C for 2 h, but retained 95% of Xyl activity, with 2,4-DNXP as substrate, under the same conditions. Further investigation of the active site topology of both enzymatic functions was performed with the inhibition study of enzyme activities. The results revealed that methyl-α-ʟ-arabinofuranoside inhibition is noncompetitive towards 2,4-DNXP as substrate but competitive towards pNPAF. Based on the thermal stability and the inhibition studies, we suggest that the enzymatic reactions of Abf and Xyl are performed at distinct catalytic sites. The recombinant enzyme possesses both the retaining

transarabinofuranosyl and transxylopyranosyl activities, indicating both enzymatic reactions proceed through a two-step, double displacement mechanism.

2-1. Introduction

The plant cell wall, the major source of fixed carbon in nature, is a composite structure containing mainly three structural polysaccharides, cellulose, hemicellulose, and lignin. L-Arabinose is a common component in some hemicelluloses, such as arabinan, arabinoxylan, and arabinogalactan [71]. Many fungi and bacteria survive in nature by producing a group of hydrolytic enzymes that synergistically catalyze the degradation of plan polysaccharides to gain nutrients. α-L-Arabinofuranosidase (Abf, EC 3.2.1.55), one of the key enzymes of the hemicellulase system, is an exo-type enzyme. It involves in degrading the arabinose-containing polysaccharides from the terminal of nonreducing end [72-74]. The release of arabinose side chains by α-L-arabinofuranosidase from hemicellulose is a critical step in the microbial degradation of natural polysaccharides [3]. α-L-Arabinofuranosidase is tremendously useful in biobleaching of paper pulp [46], bioconversion of lignocellulose material to fermentative products [37], and improvement of animal feedstock digestibility [12, 47]. It also hydrolyzes the glycosides of monoterpenes, sesquiterpenes, and other alcohols that constitute the aromatic potential of wine. The application of α-L-arabinofuranosidase in flavor improvement and wine aromatization has been studied [48-50].

Over 2500 different glycosyl hydrolases (GH) have been assigned to 106 families on the basis of amino acid sequence similarities [http://afmb.cnrs-mrs.fr/CAZY/]. α-L-Arabinofuranosidases are mainly classified into five our GH families, termed GH3, GH43, GH51, GH54, and GH62. Although many studies on the industrial applications of α-L-arabinofuranosidases from the above families have been reported, the catalytic mechanisms are less discussed. Among the five

α-L-arabinofuranosidase families, a detailed mechanistic study of family GH51 was recently reported [68] but mechanistic details for the other family members remain unknown. This lack of information reflects difficulties in obtaining synthetic substrates bearing different leaving groups [60, 61], as well as problems with effective recombinant protein expression. We reported herein that an α-L-arabinofuranosidase, a member of GH54, from *Trichoderma knonigii* was successfully expressed, purified and characterized for the first time. The accomplishment will make further investigation of the mechanism and essential residues of GH54 α-L-arabinofuranosidase become possible.

2-2. Experimental Section

Materials and Chemicals

2,4-Dinitrophenyl-β-D-xylopyranoside (2,4-DNXP) was prepared in one step using Sharma's method [75]. Buffers, p-nitrophenyl-α-L-arabinofuranoside (pNPAF), and chemicals for synthesis were purchased from Aldrich-Sigma Chemical Co. (St. Louis, MO, USA). The antibiotic zeocin was obtained from Invitrogen (Carlsbad, CA, USA). *Vent* polymerase from New England Biolabs (Ipswich, MA, USA) was used in polymerase chain reactions (PCR). Restriction endonucleases and T4 DNA ligase were obtained from Roche Applied Science (Mannheim, Germany). Plasmid pPICZαB and *Pichia pastoris* strain GS115 were purchased from Invitrogen (Carlsbad, CA, USA). *Escherichia coli* strain JM109 (ECOS, Taipei, Taiwan) served as the host for recombinant plasmids. Oligonucleotides were synthesized by Integrated DNA Technologies (Mission Biotechnology, Taipei, Taiwan). Low molecular-weight protein marker standards for electrophoresis were purchased from Amersham Biosciences (Piscataway, NJ, USA).

Media and buffers

The media BMGY, BMMY, YPD, YPD plus zeocin, LB, and low-salt LB plus zeocin are described in the EasySelect™ *Pichia* Expression Kit manual (version F) and the pPICZα A, B, C manual (Invitrogen, Appendix 3). BMGY medium contained 1% yeast extract, 2% peptone, 100 mM potassium phosphate (pH 6.0), 4×10^{-5}% biotin, 1.34% yeast nitrogen base (with ammonium sulfate and without amino acids), and 1% glycerol. BMMY medium had the same formulation except that 0.5% methanol was substituted for glycerol. YPD medium contained 1% yeast extract, 2% peptone, and 2%

glucose. Zeocin was added to YPD to a final concentration of 100 mg/mL. LB medium contained 1% tryptone, 0.5% yeast extract, and 1% NaCl. Low-salt LB plus zeocin medium contained 0.5% NaCl and 25 mg/mL zeocin.

Construction of expression vector

A vector containing the *abf* gene from *Trichoderma koningii* G-39 was the generous gift of Prof. T.-H. Hseu from the National Tsing Hua University, Hsinchu, Taiwan. To insert the *abf* gene in an expression vector, a N-terminal primer (5'-TTAAGAAGGAGCTGCAGCAATGGGGCCTTG-3') and a C-terminal primer (5'-GCTCGAATTCGGTACCTTAAGCAAAACC-3'), containing *Pst*I and *Kpn*I restriction enzyme sites, respectively, were used for PCR amplification of the *abf* gene. The amplified DNA, containing the mature *abf* gene, was purified using the Viogen kit (Viogen, Mountain View, CA, USA), digested with *Pst*I and *Kpn*I, again purified using the Viogen kit, and ligated, using T4 DNA ligase, into *Pst*I/*Kpn*I-digested pPICZαB (an *E. coli/P. pastoris* shuttle vector). *E. coli* JM109 cells were transformed with the ligation mixture and zeocin-resistant colonies were selected. The presence of the expected DNA insert in a zeocin-resistant clone was confirmed by analysis of *Pst*I/*Kpn*I restriction fragments of the recombinant plasmid, and by sequencing of recombinant DNA using commercial 5'*AOX*, 3'*AOX*, and α-factor primers (Invitrogen). The recombinant plasmid, pPICZαB-*abf*, was linearized and introduced into *P. pastoris* strain GS115 by electroporation as outlined in the *Pichia* expression kit EasySelect manual (version F, Invitrogen). The electroporation mixture was plated onto YPD-zeocin agar medium. After 72 h at 28 °C, colonies were picked for direct colony PCR using the 5'- and 3'-*AOX1* primers, followed by DNA sequence analysis of the resulting PCR products. The colony PCR reaction mixture (23.5 µL) consisted of 2.5 µL of resuspended bacterial colony, 10 pmol of each primer (each in 1 µL), 0.5 µL

17

of a mixture in which each dNTP was present at 25 mM, 15 µL sterile water, 2.5 µL 10 × PCR buffer with 20 mM $MgSO_4$ (50 mM KCl, 10 mM Tris HCl, pH 8.3), and 1 µL of *Taq* polymerase (5 U/µL). DNA was amplified in 25 cycles (1 min at 94 °C, 1 min at 55 °C, and 2 min at 72 °C) with a final extension of 7 min at 72 °C. An Applied Biosystems GeneAmp PCR system 9700 machine was used (Appendix 4). Clones positive in this test were then screened for Abf protein expression.

Protein production and purification

Recombinant cultures of *P. pastoris* (pPICZαB-*abf*) were incubated overnight at 28 °C, shaking at 180 rpm, in 250 mL BMGY medium contained in 1 L Erlenmeyer flasks, until OD_{600} values of 7–8 were attained. The cells were then transferred to BMMY medium to induce the expression of Abf. After 120 h of incubation, the cells were pelleted by centrifugation at 4200 x*g* for 30 min, and supernatant containing secreted enzyme was clarified by filtering through 0.5 µm filters. The proteins in the supernatant were precipitated by the addition of ammonium sulfate to 80–85% saturation. The precipitate was resuspended in 3 mL of sodium acetate buffer (20 mM, pH 4.5). The solution was then desalted using a 5 mL HiTrap desalting column (Pharmacia, Uppsala, Sweden). The filtrate (4 mL) was loaded onto three 5 mL cation-exchange HiTrap SP columns (Pharmacia) pre-equilibrated with 20 mM sodium acetate buffer, pH 4.5. The columns were eluted with a 150 mL linear gradient of NaCl (0–300 mM) at a flow rate of 1 mL/min. The fractions containing enzyme activity were collected, analyzed and stored at 4 °C. All purification steps were performed at ambient temperature (approx. 25 °C). Protein concentrations were determined using the bicinchoninic acid method, as described in the manufacturer's protocol (Sigma; BCA-1 kit), followed by measurement of chromophore absorption at 280 nm. The molecular weight of purified enzyme was estimated by

SDS-PAGE according to Laemmli [76] in comparison to the molecular weights of standard proteins (14–97 kDa).

Enzyme assays and kinetics

Abf activity was assayed with pNPAF as substrate by determining the amount of p-nitrophenol (or p-nitrophenolate) released.   In the activity assay, a suitable amount of purified recombinant Abf was added in 0.5 mL of acetate buffer (100 mM, pH 4.1) containing 1 mM pNPAF.   The reaction was monitored at 348 nm (the isosbestic point of p-nitrophenol/p-nitrophenolate) with $\Delta\varepsilon$ value of 2426 $M^{-1}cm^{-1}$.   For determining the activity of Xyl, 2,4-dinitrophenyl β-D-xylopyranoside (2,4-DNXP) was used as substrate. The absorption coefficient of 2,4-nitrophenolate/2,4-nitrophenol was determined to be 6000 $M^{-1}cm^{-1}$ at pH 4.1.   Kinetic studies were performed by monitoring the production of phenolates on a Hewlett–Packard model 8452A Diode Array spectrophotometer with a circulating-water bath set at 25 °C.

pH stability and Thermostability

For thermostability experiments, 300 µL portions of purified α-L-arabinofuranosidase were heated in microcentrifuge tubes at 25 °C, 35 °C, 45 °C, 55 °C, 60 °C and 65 °C.   The concentration of purified enzyme was 0.2 µg/mL in 50 mM acetate buffer, pH 5.0.   After being heated for the appropriate time interval, 30 µL of samples were then removed to assay the residual activity in pH 4.1 (100 mM NaOAc), 25 °C.   For pH stability experiments, enzyme samples (same as above) were incubated in a series of buffers with pH 1.5, 2.5, 3.5, 4.3, 5.6, 7.2, 8.3 and 9.5 at 25 °C.   Samples were removed for assay at different time intervals.   The reactions were either

monitored at the 348 nm for p-nitrophenol or at the 400 nm for 2,4-dinitrophenolate. The activity measured at pH 4.1 (100 mM NaOAc) and 25 °C served as a control.

Transglycosylation activity

Enzyme (~2 units) was added in 1 mL ammonium acetate buffer, pH 4.5, containing 20% alcohol and 10 mM 2,4-DNXP (or pNPAF for transarabinosylation). The reaction proceeded at 40 °C, overnight. The reaction mixture was concentrated to dryness. The crude solid was then resuspended in 0.5 mL of water to form a yellowish solution, which was then subjected to extraction by chloroform (0.5 mL, once) and ethyl acetate (0.5 mL, twice). The resulting aqueous solutions were lyophilized and exchanged with 0.5 mL of [^{2}H] water for NMR analysis. The NMR analyses were shown as follows:

Methyl-α-L-arabinofuranoside, 4.81 (d, 1H, C1-H, J = 1.5 Hz), 3.92 (m, 2H, C2-H, C4-H), 3.83 (m, 1H, C3-H), 3.70 (dd, 1H, J = 3.36, 8.9 Hz), 3.58 (dd, 1H, J = 5.79, 6.5 Hz), 3.30 (s, 3H).

Methyl-β-D-xylopyranoside, 4.15 (d, 1H, C1-H, J = 7.8 Hz), 3.80 (q, 1H, C5-H$_{eq}$, J = 5.4 Hz), 3.45 (m, 1H, C4-H), 3.38 (s, 3H, CH$_3$), 3.27 (t, 1H, C3-H, J = 9.0 Hz), 3.17 (t, 1H, C5-H$_{ax}$, J = 11.4 Hz), 3.08 (dd, 1H, C2-H, J = 9.3 Hz).

2-3. Results and Discussion

Protein purification

The *abf* gene from *T. koningii,* encoding α-L-arabinofuranosidase, was first inserted into plasmid pET22. The recombinant protein was highly expressed but formed an inclusion body. Although much effort was expended in attempts to refold this recombinant Abf, the endeavors were unsuccessful. The *abf* gene was next cloned into the pPICZαB vector and expressed in *P. pastoris* GS115. The resulting recombinant protein was fused with three extra amino acids (Ala-Ala-Met) at the N-terminus of the mature Abf. The recombinant Abf was expressed and secreted into the culture medium. Three consecutive steps involving ammonium sulfate precipitation, desalting, and HiTrap SP (cation exchange) chromatography were employed for enzyme purification. The chromatogram of Abf purification was shown in Figure 2-1. The fractions with Abf activity was eluted at 45-60 min, corresponding to the gradient region of 150-180 mM NaCl. The purified Abf and the secreted proteins in the medium with different induction time were analyzed by SDS-PAGE with the molecular weight of Abf at ~ 50 kDa, as shown in Figure 2-2. The purified Abf, with a purity >90%, was obtained and used for further study.

Characteristics of recombinant Abf

The substrate specificity of the recombinant enzyme was tested with a variety of glycosides, including 2,4-dinitrophenyl-β-D-xylopyranoside (2,4-DNXP), *p*-nitrophenyl-α-L-arabinopyranoside, *p*-nitrophenyl-β-D-glucopyranoside, *p*-nitrophenyl-β-D-galactopyranoside, *p*-nitrophenyl-β-D-*N*-acetylglucosamine, *p*-nitrophenyl-β-D-mannopyranoside, and *p*-nitrophenyl-α-L-arabinofuranoside (pNPAF). Of the substrates

investigated, only 2,4-DNXP, and pNPAF were hydrolyzed effectively. The activities with other non-arabinofuranosides as substrates were less than 0.5% of that with pNPAF. The enzyme showed a preference for substrates with α-L-arabinofuranose and xylopyranose at the glycone moiety. The catalytic activity (k_{cat}/K_m) of Abf (with pNPAF as substrate) and Xyl (with 2,4-DNXP as substrate) were determined, at pH 4.1, to be 65625 $M^{-1}s^{-1}$ and 17900 $M^{-1}s^{-1}$, respectively. These results suggested that the protein be a bifunctional enzyme with both α-L-arabinofuranosidase and β-D-xylosidase activity. Although many α-L-arabinofuranosidases were found to exhibit β-D-xyloside activity [77-79], the bifunctional activity of α-L-arabinofuranosidase/β-D-xylosidase (Abf/Xyl) was less discussed. This bifunctional feature is probably due to the consequence of the spatial similarity of the hydroxyl groups and glycosidic bond orientations in β-D-xylopyranosides and α-L-arabinofuranosides, which can accommodate in a same active site. Alternatively, the enzyme may possess two different catalytic sites for each reaction. Further kinetic investigation or x-ray structure analysis may be required for clarifying the catalytic feature of this bifuctional enzyme.

The pH stability of the enzyme was investigated. In general, the enzyme was stable for at least 2 h at pH values between 2 and 8.3, at room temperature. However, its activity decreased dramatically when the pH exceeded 9.5 or dropped below 1.5 (data not shown) for both Abf and Xyl activities. Study of the thermal stability of Abf/Xyl revealed that the α-L-arabinofuranosidase activity of the enzyme was less temperature-stable than was β-D-xylosidase activity. The enzyme lost 35% of Abf activity, with pNPAF as substrate, after incubation at 55 °C for 2 h, but retained 95% of Xyl activity, with 2,4-DNXP as substrate, under the same conditions (Figure 2-3).

Inhibition study

To further investigate the active site topology of both enzymatic functions, we used reversible inhibition of enzyme activities. Both α-L-arabinofuranosidase activity and β-D-xylosidase activity was monitored in the presence of inhibitor. The effects of substrate concentrations on the hydrolysis rates of 2,4-DNXP and pNPAF, using a fixed concentration of the enzyme, and different concentrations of methyl-α-L-arabinofuranoside (MAF), were investigated. The results revealed that MAF inhibition is noncompetitive towards 2,4-DNXP as substrate but competitive towards pNPAF. Estimated K_i values were > 16 mM in both cases. Since the competitive inhibition occurs when the inhibitor competes with the substrate to bind to the enzyme, whereas, the noncompetitive inhibition is caused by the binding of the inhibitor to the enzyme at another site; a noncompetitive inhibitor can bind to either the free enzyme or the enzyme/substrate complex, we conclude that the Abf and the Xyl catalytic reactions are performed at distinct enzyme sites. This argument is, in fact, supported by the study where a protein crystal was soaked either in 125 mM xylose or a 5% (w/v) solution of a xylooligosaccharide mixture, after which no concentration of electron density corresponding to sugar moiety binding was detected [70]. This suggested that the Abf active site can accommodate neither xylose nor xylosides.

Transglycosylation activity

The recombinant enzyme was found to be stable in the presence of methanol in NaOAc buffer (pH 4.5). When substrate (pNPAF or 2,4-DNXP) was catalyzed by enzyme in the presence of 20% methanol, both enzymatic products derived from hydrolysis and transglycosylation were obtained (as shown in Scheme 2-I). The yields of methyl-α-L-arabinofuranoside (MAF) and methyl-β-D-xylopyranoside (MX), estimated from NMR spectra (data not shown), were ~20% and ~30%, respectively, when pNPAF and

2,4-DNXP were used as substrate for Abf and Xyl catalysis correspondingly. The NMR assignments of MAF and MX were confirmed according to literatures [80, 81].

Proposed the catalytic mechanisms of the bifunctional Abf/Xyl

In general, glycosyl hydrolases cleave the glycosidic bond of the substrate by two different mechanisms, the retention and the inversion of the anomeric configuration. Both mechanisms require two essential residues, which in most glycosyl hydrolases are residues with carboxylic acid side-chain such as aspartate or glutamate. Inverting glycosidases catalyze the hydrolysis of the glycosidic bond via a single displacement reaction. One of the catalytic residues, acting as a general acid, provides protonic assistance to the departing glycosidic oxygen, while the second, acting as a general base, activates a water molecule which that effects a direct displacement at the anomeric centre and to form the product with the inversion of the anomeric configuration [51]. Retaining glycosidases catalyze the hydrolysis via a two-step, double-displacement mechanism, as shown in figure 2-4. Two key active-site carboxylic acid residues [52-55] are involved. One functions as the nucleophile and the other functions as the general acid/base [53]. In the first step (the glycosylation step), the nucleophile attacks the anomeric carbon of the glycoside, while the acid/base catalyst protonates the glycosidic oxygen, thereby assisting the leaving of the aglycon moiety. This leads to the formation of a covalent glycosyl-enzyme intermediate. In the second step, breakdown of the glycosyl-enzyme intermediate proceeds through a general base-catalyzed attack of water at the anomeric center to release the glycose with the retention of the anomeric configuration. Since the bifunctional Abf/Xyl gave the products of transglycosylations with the retaining feature, i.e.

methyl-α-L-arabinofuranoside and methyl-β-D-xylopyranoside, we concluded that the catalytic reactions of the recombinant enzyme for both Abf and Xyl activity proceed via a two-step, double displacement mechanism.
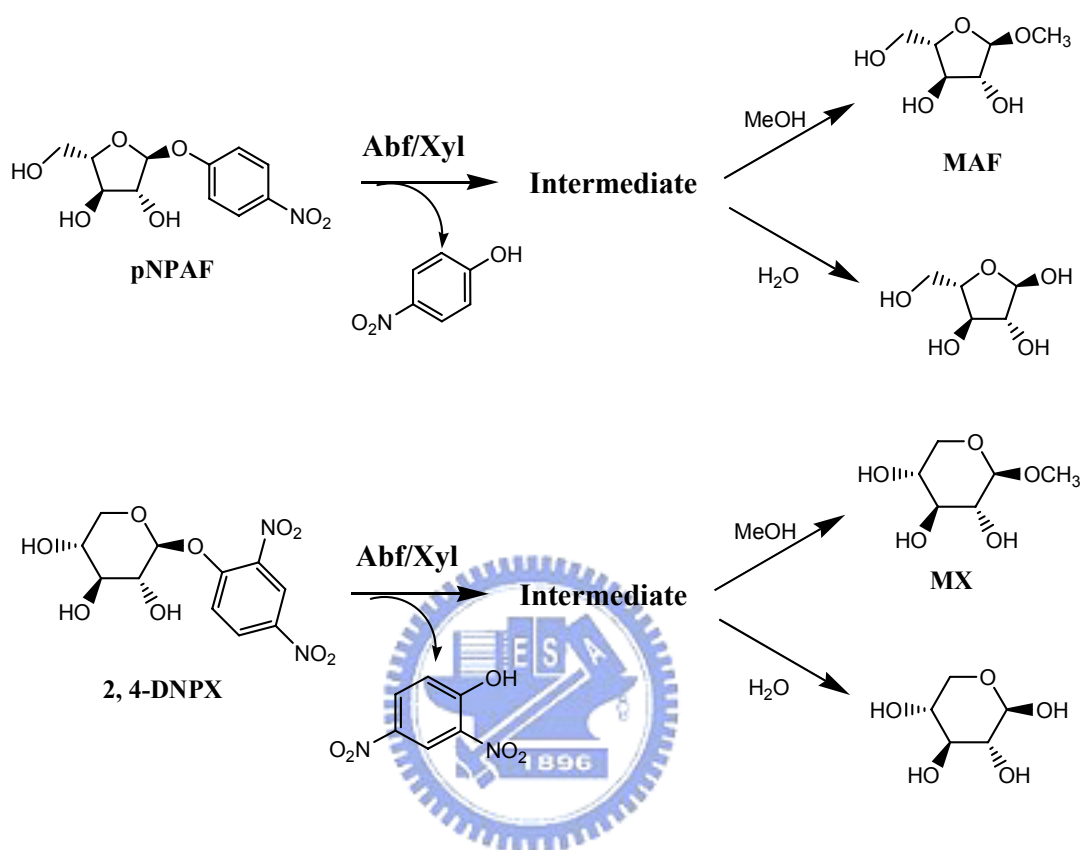
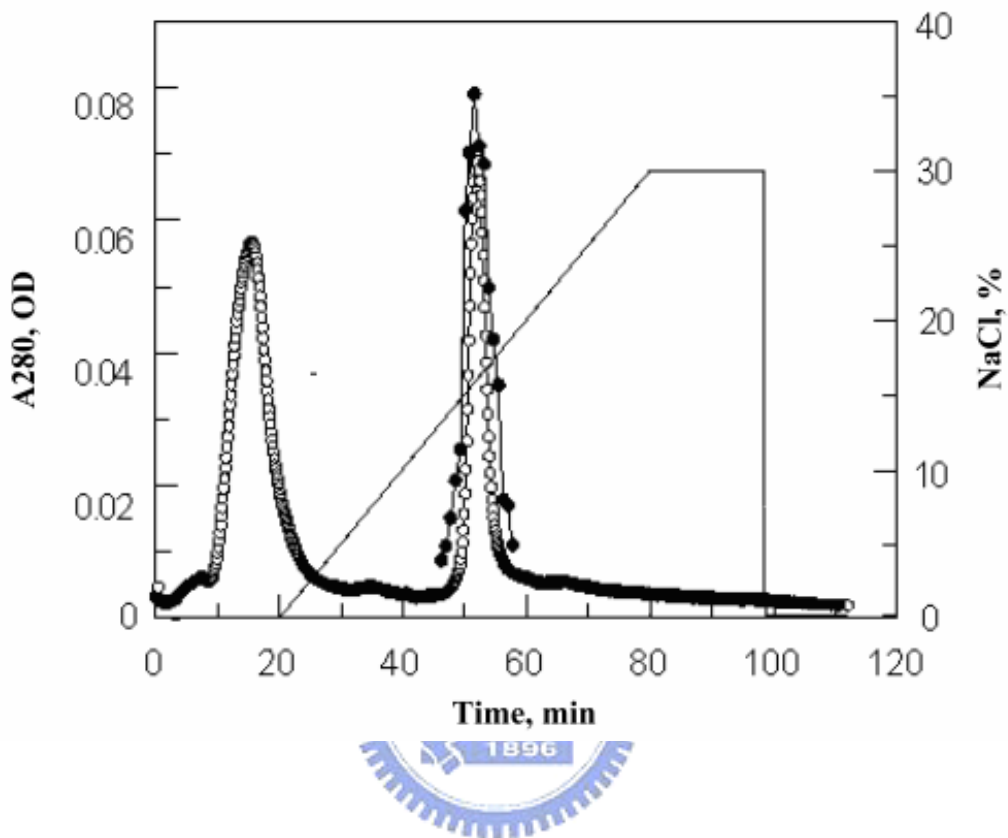**Scheme 2-I.** The catalytic hydrolysis and transglycosylation of the bifunctional Abf/Xyl

Figure 2-1. The chromatography of the purification of the recombinant of Abf. The aborbance (-○-) was monitored at 280 nm.   The fractions with Abf activity (-●-) was eluted within 150-180 mM NaCl (—).
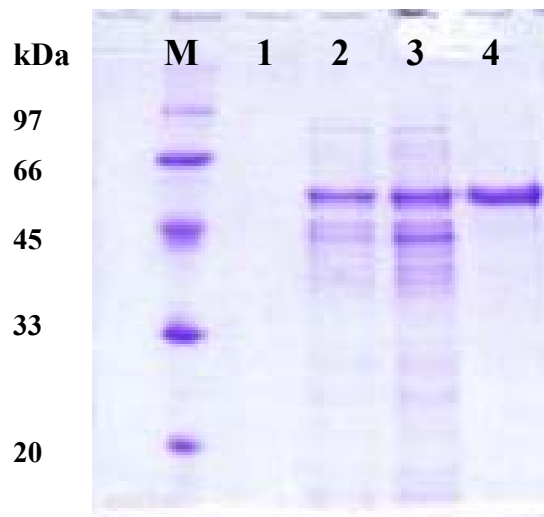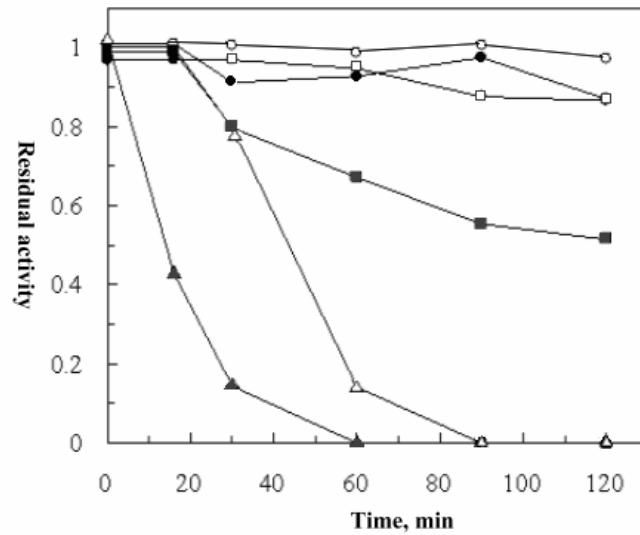
Figure 2-2. SDS-PAGE analysis of Abf obtained from different induction time and the purified protein. Lanes: M, molecular mass markers; 1, recombinant enzyme from day 0 supernatants; 2, recombinant enzyme from day 1 supernatants; 3, recombinant enzyme from day 4 supernatants; 4, pool of active fractions from HiTrap-SP column.
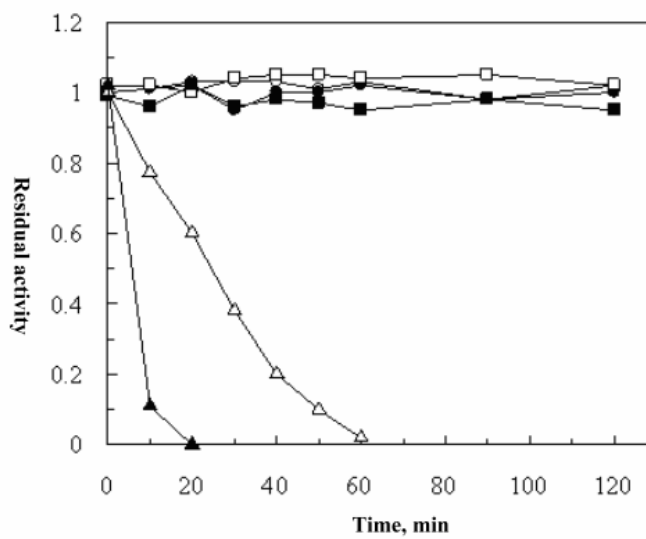
(a)

(b)

Figure 2-3. Thermal stability of the recombinant enzyme assay as Abf (a) and Xyl (b). Enzyme was incubated in various temperature: 25 °C (○), 35 °C (●), 45 °C (□), 55 °C (■), 60 °C (△), and 65 °C (▲).
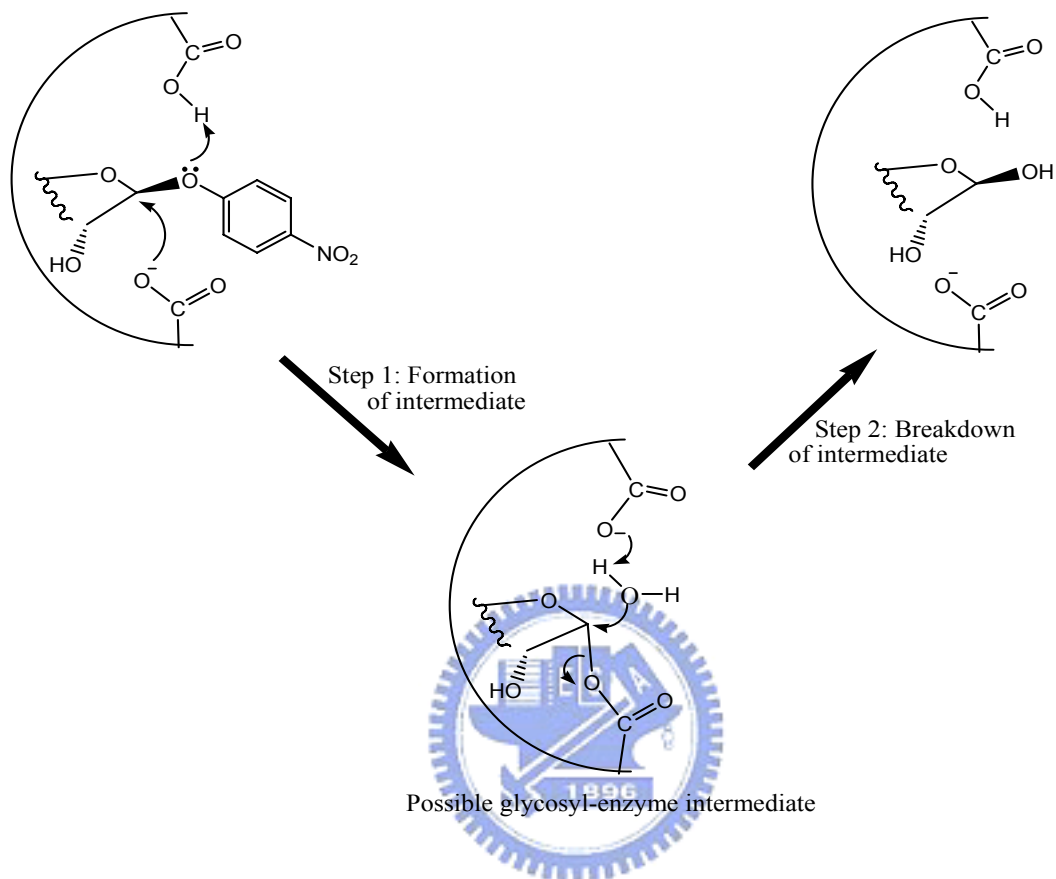
Figure 2-4. The proposed two-step, double displacement mechanism of Abf/Xyl.

Chapter 3

Mutagenesis and mechanistic study of a family 54 α-ʟ-arabinofuranosidase from *Trichoderma koningii*

Abstract

A GH54 α-ʟ-arabinofuranosidase (Abf) originally from *Trichoderma koningii* G-39 was successfully expressed in *Pichia pastoris* and purified to nearly homogeneity by cation-exchange chromatography. Extensive mutagenesis of 24 conserved Glu and Asp residues of family 54 was performed. The k_{cat} values of the D221N and D299N were 7000- and 1300-fold lower than that of the wild-type Abf, respectively, while E223Q was nearly inactive. These results are consistent with implications from the *Aspergillus kawachii* α-ʟ-arabinofuranosidase three-dimensional structure. This structure indicates that E223 of *T. koningii* Abf function as a nucleophile and D299 as a general acid/base catalyst for the enzymatic reaction, whereas D221 is significant for substrate binding. The catalytic mechanism of wild-type Abf was further investigated by NMR spectroscopy and kinetic analysis. The results showed that Abf is a retaining enzyme. It catalyzes various substrates via the formation of a common intermediate that is probably an arabinosyl-enzyme intermediate. A two-step, double-displacement mechanism involving first the formation, and then the breakdown, of an arabinosyl-enzyme intermediate was proposed. Based on the k_{cat} values of a series of aryl-α-ʟ-arabinofuranosides catalyzed by wild-type Abf, a relatively small Brønsted constant, $\beta_{lg} = -0.18$, was obtained, suggesting that the rate-limiting step of the enzymatic reactions for substrates (where the leaving phenols have pK_a values ≥ 5.15) is the dearabinosylation step. Further

kinetic studies with D299G mutant revealed that the catalytic activity of this mutant depended largely on the pK_a values (> 6) of leaving phenols, with $\beta_{lg} = -1.3$. This indicated that the rate-limiting step became arabinosylation step when D299G was employed. This kinetic outcome supports the idea that D299 is the general acid/base residue. The pH activity profile of D299N provided further evidence strengthening this suggestion.

3-1. Introduction

α-L-Arabinofuranosidases (EC 3.2.1.55) are among key enzymes of the hemicellulase system, which is tremendously useful in biobleaching of paper pulp [46], bioconversion of lignocellulose material to fermentative products [37], and improvement of animal feedstock digestibility [12, 47]. α-L-Arabinofuranosidases catalyze the hydrolysis of α-1,2- and α-1,3- α-1,5-L-arabinofuranosidic bonds in hemicelluloses such as arabinoxylan, L-arabinan, and other L-arabinose-containing polysaccharides [12, 31]. These enzymes also hydrolyze the glycosides of monoterpenes, sesquiterpenes, and other alcohols, releasing the aromatic potential of wine. The enzymes have general applications in flavor improvement [48-50].

Owing to their industrial importance, a variety of α-L-arabinofuranosidases have been purified from various sources such as bacteria [59], fungi [41], and plants [77]. Many genes encoding Abf enzymes have been cloned and sequenced. On the basis of amino acid sequence similarities [56, 82], α-L-arabinofuranosidases are grouped into five our glycoside hydrolase (GH) families, termed GH3, GH43, GH51, GH54, and GH62. Although many studies on the industrial applications of α-L-arabinofuranosidases from the above families have been reported, only a few papers have been published describing detailed catalytic mechanisms, with reference to the specific enzyme residues involved. This lack of information reflects difficulties in obtaining synthetic substrates bearing different leaving groups [60, 61], as well as problems with effective recombinant protein expression.

There are two broad mechanisms of glycoside hydrolase cleavage of glycosidic bonds. The original anomeric configuration is either inverted or retained. Both mechanisms require amino acids containing side-chain carboxylic residues as essential group(s). Inverting glycosidases catalyze the

hydrolysis of the glycosidic bond via a single displacement reaction. One of the catalytic residues, acting as a general acid, provides protonic assistance to the departing glycosidic oxygen, while the second, acting as a general base, activates a water molecule that effects a direct displacement at the anomeric center [51]. For retaining glycohydrolases, though a catalytic mechanism, involving the formation of an oxocarbenium ion intermediate, was once proposed by Phillips [83], further studies with this type of glycosidases provide strong evidence pointing to a common mechanism that involves the formation of covalent glycosyl-enzyme intermediates [54, 55, 84, 85], as previously postulated by Koshland [86]. In general, two key active-site carboxylic acid residues are involved. One serves as the nucleophile and the other serves as a general acid/base. The proposed mechanism was shown in Figure 3-1. In the first step (the glycosylation step), the nucleophile attacks the anomeric carbon of the glycoside, while the acid/base catalyst protonates the glycosidic oxygen, thereby assisting the leaving of the aglycon moiety. This leads to the formation of a covalent glycosyl-enzyme intermediate. In the second step (the deglycosylation step), breakdown of the glycosyl-enzyme intermediate proceeds through a general base-catalyzed attack of water at the anomeric center to release the glycose. For kinetic analysis of a catalytic mechanism, a series of artificial substrates, each with a glycose moiety attached to different leaving phenols, is commonly prepared. Typically, the extended Brønsted relationship is to be measured. As enzymes in the same family presumably possess a similar catalytic mechanism, solving the mechanistic action of a particular enzyme may help in the understanding of the reaction pattern of its family. Among the five α-L-arabinofuranosidase families, a detailed mechanistic study of family GH51 was recently reported [68] but mechanistic details for the other family members remain unknown. Very recently, the potential for mechanistic studies of members of the GH54 family has greatly increased, owing to the complete resolution of the

three-dimensional structure of an α-L-arabinofuranosidase from *Aspergillus kawachii* IFO4308 [70]. On the basis of protein structure and very preliminary mutagenic studies, E221 and D297 were suggested as candidates for the nucleophile and the general acid/base catalytic residues, respectively. In this study, we confirm the catalytic functions of essential Abf residues by site-directed mutagenesis and kinetic analysis of mutant enzyme properties. Further, the enzymatic mechanism of a GH54 family member is described and discussed, for the first time, in this thesis.

3-2. Experimental procedures

Materials

The materials are described in the 2-2 experimental section. BMGY, BMMY, and YPD media were prepared according to the manufacturer's instructions and 2-2 experimental section. The medium used for 1 L fermentations was that described in the *Pichia* fermentation guidelines.

Substrates synthesis

The synthesis of phenol-linked arabinofuranosides differs slightly from methods followed in the preparation of glycosides where the glycose has a six-membered ring. To obtain the arabinofuranoside as the major final product, methyl-α-L-arabinofuranoside was synthesized and used as the starting material for further reactions. The complete procedure has been published [87]. All substrates, including phenyl-α-L-arabinofuranoside (PAF), p-cyanophenyl-α-L-arabinofuranoside (pCPAF), m-nitrophenyl-α-L-arabinofuranoside (mNPAF), p-nitrophenyl-α-L-arabinofuranoside (pNPAF), 4-chloro-2-nitrophenyl-α-L-arabinofuranoside (CNPAF), and 2,5-dinitrophenyl-α-L-arabinofuranoside (2,5-DNPAF) were synthesized according to this published protocol. All substrates were purified by column chromatography and structures confirmed by NMR, which yielded the following data: methyl-α-L-arabinofuranoside, 4.81 (d, 1H, C1-H, J = 1.53 Hz), 3.92 (m, 2H, C2-H, C4-H), 3.83 (m, 1H, C3-H), 3.70 (dd, 1H, J = 3.36, 8.89 Hz), 3.58 (dd, 1H, J = 5.79, 6.51 Hz), 3.30 (s, 3H); phenyl-α-L-arabinofuranoside, 7.33 (t, 2H), 7.05 (m, 3H), 5.64 (s, 1H, C1-H), 4.29 (m, 1H, C2-H), 4.11 (dd, 1H, C4-H, J = 3.57, 5.49 Hz), 4.02 (dd, 1H, C3-H, J = 2.19, 3.56 Hz), 3.75 (dd, 1H, J = 3.57, 8.78 Hz), 3.66 (dd, 1H, J = 5.49, 11.25 Hz); p-cyanophenyl-α-L-arabinofuranoside, 7.65 (d, 2H, J = 8.91

Hz), 7.12 (d, 2H, J = 8.89 Hz), 5.71 (s, 1H, C1-H), 4.30 (dd, 1H, C2-H, J = 1.53, 1.55 Hz), 4.07 (dd, 1H, C4-H, J = 2.00, 3.38 Hz), 4.00 (dd, 1H, C3-H, J = 2.53, 3.26 Hz), 3.73 (dd, 1H, J = 3.42, 9.05 Hz), 3.64 (dd, 1H, J = 5.38, 6.94 Hz); m-nitrophenyl-α-L-arabinofuranoside, 7.8 (d, 2H, J = 8.78 Hz), 7.48 - 7.37 (m, 2H), 5.68 (s, 1H, C1-H), 4.29 (dd, 1H, C2-H, J = 1.54, 1.67 Hz), 4.07 (m, 1H, C4-H), 3.99 (m, 1H, C3-H), 3.72 (m, 1H), 3.63 (m, 1H); p-nitrophenyl-α-L-arabinofuranoside, 8.28 (d, 2H, J = 9.34 Hz), 7.26 (d, 2H, J = 9.06 Hz), 5.88 (d, 1H, C1-H, J = 1.1 Hz), 4.46 (dd, 1H, C2-H, J = 1.37, 1.67 Hz), 4.20 (dd, 1H, C4-H, J = 2.74, 3.33 Hz), 4.14 (dd, 1H, C3-H, J = 3.57, 9.06 Hz), 3.77 (dd, 1H, J = 5.67, 7.14 Hz), 3.66 (dd, 1H, J = 5.38, 5.45 Hz); 4-chloro-2-nitrophenyl-α-L-arabinofuranoside, 7.92 (d, 1H, J = 2.74 Hz), 7.58 (dd, 1H, J = 2.74, 6.32 Hz), 7.32 (d, 1H, J = 9.06 Hz), 5.66 (d, 1H, C1-H, J = 1.37 Hz), 4.34 (dd, 1H, C2-H, J = 1.65, 1.92 Hz), 4.12 (m, 1H, C4-H), 3.98 (dd, 1H, C3-H, J = 2.47, 3.57 Hz), 3.73 (dd, 1H, J = 3.57, 9.05 Hz), 3.63 (dd, 1H, J = 5.49, 6.86 Hz); 2,5-dinitrophenyl-α-L-arabinofuranoside, 8.27 (s, 1H), 8.01 (d, 2H, J = 1.1 Hz), 5.78 (d, 1H, C1-H, J = 1.37 Hz), 4.32 (dd, 1H, C2-H, J = 1.37, 2.47 Hz), 4.09 (m, C4-H, 1H), 4.02 (dd, 1H, C3-H, J = 2.74, 3.84 Hz), 3.79 (dd, 1H, J = 3.02, 9.05 Hz), 3.68 (dd, 1H, J = 4.66, 7.41 Hz) (Appendix 5).

Vector construction for Pichia expression system

Vector construction for *Pichia* expression system were described in the 2-2 experimental section.   Clones positive in this test were then screened for Abf protein or Abf proteins expression by initial growth, with shaking at 180 rpm, in 50 mL BMGY medium contained in 250 mL Erlenmeyer flasks. Recombinant DNA expression in cultures at OD_{600} values of 7-8 was induced by transferring the cells to BMMY medium.   At 24, 48, 72, 96, and 120 h postinduction, methanol, to a final concentration of 0.5% (v/v), was added to cultures.   After 160 h of incubation, cultures were centrifuged and proteins in

the supernatants analyzed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) using a 12% (w/v) polyacrylamide separating gel with a 5% (w/v) stacking gel.   Clones producing the protein of interest were identified by the presence of an appropriately sized (50 kDa) supernatant protein band.

In vitro site-directed mutagenesis

The *abf* gene was mutagenized using the QuickChange site-directed mutagenesis kit (Stratagene, San Diego, CA, USA).   The primers used to generate the mutations were as follows (mutations are underlined): D170N(+), 5'-CATGTATGCGGTTCTC<u>AAT</u>GGTACACAC-3'   and   D170N(−), 5'-GTGTGTACC<u>ATT</u>GAGAACCGCATACATG-3';   D221N(+), 5'-GATCATGGCT<u>AAT</u>CTCGAGAACGGCTTG-3'   and   E223Q(−), 5'-CAAGCCGTT<u>CTG</u>GAGATCAGC<u>C</u>ATGATC-3';   E223G(+), 5'-GATCATGGCTGATCTTC<u>GGC</u>AAGGCTTG-3';   and   E223G(−), 5'-CAAGCCGTT<u>GCC</u>GAGATCAGCCATGAT   C-3';   D299N(+), 5'-CTCGGCATTGGCGGC<u>AAC</u>AACAGC-3'   and   D299N(−), 5'-GCTGTT<u>GTT</u>GCCGCCAATGCCGAG-3';   D299G(+), 5'-CATTGGCGGC<u>GGC</u>AACAGCAACGGC-3'   and   D299G(−), 5'-GCCGTTGCTGTT<u>GCC</u>GCCGCCAATG-3';   E310Q(+), 5'-TCTAT<u>CAG</u>GGCGTCATGACCTCTGGCTATC-3'   and   E310Q(−), 5'-GATAGCCAGAGGTCATGACGCC<u>CTG</u>ATAGA-3'. Note that D221N (+) and E223Q(−) primers were used to performed both mutations.   Primers creating other mutations that did not cause significant loss in enzyme activity are not listed (Appendix 6).   All mutations were confirmed by DNA sequencing.   For those mutations creating significant losses in enzyme activity, the entire mutated genes were sequenced to confirm that only the intended mutations had occurred.

Protein production and purification

The Abf and Abf mutant proteins production and purification steps were described in the 2-2 experimental section. To minimize enzyme cross-contamination during purification, fresh HiTrap SP columns were used in the purification of each mutant protein.

Kinetic studies

Enzyme activity was determined at pH 4.1 by monitoring the hydrolysis of pNPAF to release p-nitrophenol (or p-nitrophenolate). A Hewlett-Packard model 8452A spectrophotometer equipped with a circulating water bath at 25 °C was used. The pH-dependence of enzyme activity in both wild-type and mutant enzymes was assessed using pNPAF as substrate with assay pH values in the range 1.9-6.5. Reactions were monitored at the isosbestic point (348 nm) of p-nitrophenol and p-nitrophenolate and kinetic constants calculated on the basis that the molar absorption coefficient ($\Delta\varepsilon$) was 2426 $M^{-1}cm^{-1}$. Buffers used in this study were glycine (pH 1.8 to 3.5), sodium acetate (pH 4.0 to 5.5), morpholinoethanesulfonic acid (pH 5.5 to 6.5), phosphate (pH 6.5 to 7.5), and Bicine (pH 7.5 to 9.5). The Michaelis constant was determined for each synthetic substrate from the Michaelis–Menten equation, using a double reciprocal plot. Reaction rates were determined, in triplicate, at pH 6.5 (in morpholinoethanesulfonic acid buffer, 100 mM) at each of seven different substrate concentrations. The wavelengths employed and the $\Delta\varepsilon$ ($M^{-1}cm^{-1}$) values obtained at those wavelengths for each arabinofuranoside substrate were as follows: 2,5-DNPAF, 440 nm, 3616; CNPAF, 400 nm, 2603; pNPAF, 400 nm, 3077; mNPAF, 380 nm, 330; pCPAF, 290 nm, 672; PAF, 278 nm, 754.

Mass spectroscopy

Mass spectra were recorded using a Micromass Q-TOF mass spectrometer. Protein samples were injected into the mass spectrometer via

an HPLC system equipped with a PLRP-S column (5 μm, 300 Å, 1-50 mm) using a 5-90% (v/v) acetonitrile gradient containing 0.1% (v/v) formic acid as the eluant.   Some 5-10 μg amounts of proteins were injected.

3-3. Results and discussion

Enzyme expression and purification

The resulting recombinant enzyme was expressed and described in the 2-3 results and discussion. The purified enzyme (> 90% homogeneity) was analyzed by SDS-PAGE (Figure 3-2a) and mass spectrometry (Figure 3-2b), and gave mol wt values of ca. 50 kDa and 50045 (major peak), respectively. The theoretical mol wt is 49445 Da. The discrepancy between the measured and calculated mol wt values is likely because of protein glycosylation in *Pichia*. This argument is somewhat supported by the presence of a series of mass peaks (50045, 50208, 50369; Figure 3-2b) with differences of 161-163 Da, which may indicate the glycosylation of protein.

pH profiles of the wild-type and the E223G mutant

It is common that the catalytic reactivity of many retaining glycohydrolases is mainly mediated by two active site carboxylates (aspartate and/or glutamate). A bell-shaped activity profile reflected two apparent pK_a is often observed. The pH-dependent activity assay showed that the purified enzyme had optimum Abf activity at pH 2.8–3.2 (Figure 3-3). The pH-dependent activity curve of recombinant Abf exhibited a bell shape, with two apparent pKa values of 1.8 ± 0.1 and 4.2 ± 0.1. It is of interest that both pK_a values were about 2 pK_a units lower than those found for other glycohydrolases.

Mutagenic study

The first protein structure of a family 54 α-L-arabinofuranosidase, from *A. kawachii,* was recently resolved [70]. This structure is valuable for structural simulation of enzymes in the same family. Based on sequence comparisons, our enzyme has 73% identity and > 85% homology to the *A. kawachii*

α-L-arabinofuranosidase. By inspection of the protein structure (Figure 3-4), residues appearing in the active site of *A. kawachii* α-L-arabinofuranosidase and *T. koningii* (shown in parentheses) may be listed: C176 (C178), D219 (D221), E221 (E223), D297 (D299), and S299 (S301). All these residues are highly conserved in 12 enzymes of the GH54 family (Figure 3-5). The carboxyl groups of E221 (E223 in *T. koningii* Abf) and D297 (D299 in Abf) are located on either side of the anomeric C-1 carbon of enzyme-bound arabinofuranose, strongly suggesting that these residues serve as nucleophile and general acid/base residue, respectively. To confirm the catalytic roles of these residues, an extensive mutagenesis study was performed on E223, D299, and 22 other Asp or Glu residues (D95, D120, D160, E163, D170, D181, E186, D191, E198, D221, D238, E289, E310, D320, E323, E407, D410, D429, E436, D437, D482, and D490), which are highly conserved in GH54 family members. We changed the putative catalytic residues, aspartate (D) and glutamate (E), to asparagine (N), glutamine (Q) or glycine (G) by site-directed mutagenesis. All crude mutant enzymes with mutations in D170, D221, E223, D299 or E310 largely lost enzyme activity with pNPAF as substrates (Appendix 7). Some of these mutant enzymes were further purified and studied. Kinetic parameters are summarized (Table 3-1). The K_m values of mutant enzymes, ranging from 0.22–0.32 mM, are quite similar to that of wild-type Abf, except for the mutants D299N, and E310G, where the K_m values were 0.05 mM, and 3.0 mM, respectively. The k_{cat} values of the mutant enzymes D170N, D221N, D299N, D299G, and E310G decreased (with respect to the wild-type Abf) by factors of 31, 7000, 1300, 262, and 70, respectively, while E223Q was completely inactive. The reduction in the catalytic activities of the mutant enzymes E223Q and D299N was consistent with the suggested roles of the altered amino acids as catalytic residues, as indicated by the structure of the *A. kawachii* IFO4308 enzyme. Though residue E223, conserved in the DXE (Glu-X-Asp) sequences of glycoside

hydrolase family 54 members, has been postulated to act as a nucleophile, this study uses site-directed mutagenesis for direct confirmation of this function. Residue D221 (D219 in the *A. kawachii* enzyme) engages in strong hydrogen bonding with the C3 and C5 sugar hydroxyl groups [70]. The large decreases in activity observed when these residues are mutated presumably occurs because the substrate is now incorrectly oriented in the active site. The residues D170 and E310 are more than 20 Å distant from the active site. The activity loss of the E310G mutant results in part from a 10-fold increase in K_m value, compared to the unmutated enzyme. A perturbation of general protein structure by this mutation should perhaps be considered.

One unexpected finding is that the E223G mutant still possesses extremely high activity. Its k_{cat} (22 s^{-1}) and k_{cat}/K_m (59460 M^{-1}s^{-1}) values are nearly identical to those of wild type Abf. After careful inspection of the active-site structure of *A. kawachii* enzyme, the Asp-189 (Asp-191 in our Abf) was found to locate closely to the nucleophile, E221 (E223 in our Abf). The orientation and distance (4–5 Å) between Asp-189 and the C-1 position of the substrate is perfect for an inverting-type of catalysis to take place. It is very likely that when the site-chain of the nucleophile is replaced by proton (ie, E223G), more space will be gained and, therefore, allow water molecule to diffuse deeply into the active site, while it is no room around the nucleophile if the Glu is mutated into Gln. This hypothesis has been tested by analyzing the activity of double mutant (D191N/E223G) and the stereochemistry of E223G catalysis. The results from this study will be published elsewhere.

In many glycohydrolases, exogenous nucleophiles such as azide, formate, and other anions have been shown to enhance the catalytic activity of enzymes mutant in residues that serve as nucleophiles or provide general acid/base functions [88-91]. Activity enhancement of a mutant by addition of a nucleophile (for example, azide) and formation of a stereospecific product (α- or β-glycosyl azide) offers a useful technique for identifying essential residues

of glycosidases.   In this study, however, addition of high concentrations of azide (up to 2 M) did not rescue the activities of the D299G, D299N, E223Q or E223G mutants, presumably because the active site was located relatively deep in Abf that the azide ion could not diffuse to the site.   Similar observations were reported previously in a study on the *A. kawachii* IFO4308 enzyme [70].   Nevertheless, comparing the pH activity profiles (k_{cat} vs pH) of the wild-type Abf and the D299N mutant (Figure 3-3) may provide an insight into the essential function of D299.   The activity of the D299N mutant was almost constant in the range of pH 3–6.5 and with a trend of decrease at pH <2.4, indicating the absence of the second pK_a point.   Also, the low K_m value (0.05 mM) of the D299N mutant suggests accumulation of the glycosyl-enzyme intermediate, whose hydrolysis is then accelerated by the general acid/base catalytic action of the enzyme.   Unfortunately, owing to the various degree of post-glycosylation on the recombinant Abf, an unequivocal result of mass spectrometric analysis in attempt to show the presence of the glycosyl-enzyme intermediate was unsuccessful.   In sum, the results of the site-directed mutagenesis and kinetic studies, in combination with structure analysis, point to E223 and D299 as the essential nucleophile and general acid/base residues of Abf, respectively. The function of D299 will be further discussed below.

Transglycosylation and product partition

   A time-course ^{1}H NMR study is commonly employed to examine anomeric preference in the catalytic action of a glycohydrolase.   Though such work provides unequivocal stereochemical data, the process is tedious, as equilibration with a deuterated buffer system is required and enzyme concentration must be carefully monitored.   Also, a significant limitation of this technique is that product mutarotation rate must be relatively slow so that NMR detection of authentic product is feasible.   In the case of

arabinofuranosidases, an arabinofuranoside product may be expected to undergo fast mutarotation to form four arabinose tautomers. These are α- and β-L-arabinofuranosides and α- and β-L-arabinopyranosides [57]. To overcome this limitation, a method involving transglycosylation of the Abf using methanol as glycosyl acceptor was employed. The advantage of this strategy is the formation of a methyl glycoside that cannot mutarotate. For most retaining enzymes, although the formation of a covalent enzyme intermediate is expected, it is difficult to detect as the lifetime is short. In the past, useful 2-fluoroglycoside products have been obtained using a specific glycosyl-enzyme trapping technique [62, 92-94]. However, perhaps because 2-fluoroarabinofuranoside is difficult to synthesize, this strategy has not yet been used in any study of α-L-arabinofuranosidase. We used an alternative, indirect, method. If a constant chemical bias may be noted in different enzyme products, the formation of a common intermediate in an enzymatic reaction may be inferred. Our study showed that the Abf exhibited strong transglycosylation activity when methanol was used as an arabinofuranosyl acceptor. Here, a suitable amount of pNPAF, pCPAF and CNPAF were enzymatically hydrolyzed in acetate buffer (pH 4.1) containing 12% (v/v) methanol. The solutions were then dried and exchanged with D_2O several times before ^{1}H-NMR analysis. For all three reactions, the ^{1}H-NMR spectra (measured at 25 ºC) of the sugar moieties (in the range of 3-6 ppm) were nearly identical. In principle, five different end products with arabinosyl ring structures should be observed (Figure 3-6a). According to the literature [80] and our study, the C1 protons of each sugar ring were assigned as follows: methyl-α-arabinofuranoside (4.86 ppm, J = 1.03 Hz), α-arabinofuranose (5.17 ppm), β-arabinofuranose (5.22 ppm, J = 3.8 Hz), α-arabinopyranose (4.43 ppm, J = 5.9 Hz), and β-arabinopyranose (5.16 ppm, J = 3.24 Hz) (Figure 3-6b). Based on peak assignment and the integration of the C1 proton on the sugar ring, the ratio of methyl-α-arabinofuranoside/arabinose was calculated

from each spectrum. Regardless of the substrates, these ratios were nearly constant, and averaged 1.04 ± 0.02 (1.04 for pCPAF, 1.06 for pNPAF, and 1.03 for CNPAF). This suggests that a common intermediate, most likely an arabinosyl-enzyme structure, occurs in the reaction pathways. As the product of transglycosylation in this experiment is methyl-α-arabinofuranoside, we can confirm that the Abf is indeed a retaining enzyme.

Substrate reactivity and Brønsted plot

Kinetic assessment of substrates bearing different leaving phenols is a common strategy in studying mechanistic actions of glycohydrolases [95]. For glycohydrolases with two-step mechanisms (formation of a glycosyl-enzyme intermediate in a glycosylation step followed by breakdown of the intermediate in a deglycosylation step), the aglycon moiety is cleaved in the glycosylation step of the reaction (Figure 3-1). Therefore, the reaction rate of the first step dictates the ease by which the leaving group may be released from the substrate. A strong correlation between the pK_a of the phenol leaving group and the activity of the enzyme should be observed if the first step is rate-limiting. To determine the rate-limiting step of hydrolysis catalyzed by the recombinant Abf, we worked to prepare aryl-α-L-arabinofuranosides bearing different leaving phenols (pK_a values 5.15 to 9.99) and to perform steady-state kinetic analysis using these substrates. The reaction temperature was held at 25 ºC to reduce spontaneous substrate hydrolysis. Although Abf is specific with regard to the glycon moiety of the substrate, the enzyme shows a broad specificity for the aglycon portion. Kinetic parameters for reactions using the aryl α-L-arabinofuranoside substrates are summarized (Table 3-2). The data showed that both the K_m (0.28-0.54 mM) and the k_{cat} (1.23-8.54 s^{-1}) values were approximately the same for all substrates. These data permit calculation of the Brønsted

relationship, which has been shown to be useful in understanding the mechanism of enzyme action [51, 96-97]. Based on the k_{cat} values, an extended Brønsted plot was constructed by plotting the logarithm of k_{cat} against the pK_a of the leaving phenol (Figure 3-7a). A plot with a slightly downward trend was obtained, with a slope (β_{lg} value) of –0.18. That the detailed structures of the leaving phenols do not affect the kinetic parameters of the reaction may indicate that the dearabinosylation step (breakdown of the arabinosyl-enzyme intermediate) is the rate-limiting step. Alternatively, although it may be considered unlikely, the low Brønsted constant could indicate that the arabinosylation step is the slow step and that an early transition state is attained. However, the very weak reaction inhibition shown by MAF (K_i > 16 mM) minimized this possibility. If the transition state were substrate-like, strong inhibition would be expected. We further studied substrate reactivity using the D299G mutant enzyme. A new plot with β_{lg} = –1.3 was obtained (Figure 3-7a). For substrates with pK_a > 6, the absence of general acid/base catalysis resulted in a strong correlation between the pK_a of the phenol leaving group and enzyme activity. Clearly, formation of the arabinosyl-enzyme intermediate is now the rate-limiting step. However, with 2,5-DNPAF (a good substrate, with pK_a = 5.15), the log k_{cat} value is clearly lower than expected if the enzyme reaction were to proceed as outlined above. This indicates that the dearabinosylation step becomes at least partially rate-limiting when 2,5-DNPAF breakdown is catalyzed by the D299G enzyme. In addition to alteration of catalytic properties, the D299G mutation also affects enzyme K_m values with different substrates (Table 2). The K_m values decrease as the pK_a values of the leaving phenols change in the substrates. For the wild-type enzyme, such changes in K_m values are not obvious. These data showed that as the ability of the aglycon moiety to depart the enzyme increased, the more glycosyl-enzyme intermediate accumulated. This may be appreciated if it is noted that K_m can be

47

represented as [E][S]/Σ[ES].   Thus, the more the glycosyl-enzyme intermediate accumulates, the lower the K_m value. Such behavior may be expected in mutants affected in the general acid/base residue, as a good leaving group elevates the rate of the first catalytic step.   The rate of the second step remains slow as the basic residue, which (in unmutated enzymes) activates the water molecule, is missing.

The k_{cat}/K_m values are informative with respect to the first irreversible step.   For retaining glycoside hydrolases, the Brønsted relationship obtained by plotting k_{cat}/K_m values of substrates against the pK_a values of the leaving phenols provides information about the glycosylation step [54, 61].   As may be seen (Figure 3-7b), the Brønsted constants (β_{lg}) are –0.19 and –1.2 for wild-type Abf and the D299G mutant, respectively.   This suggests that the arabinosylation step catalyzed by the wild-type Abf enzyme is not sensitive to the leaving abilities of different phenols, as the general acid/base residue protonates the oxygen of the glycosidic bond, thus imparting constant leaving abilities to different phenols.   When the general acid/base residue is absent, however, the catalytic efficiency (k_{cat}/K_m) for the tested substrate became highly sensitive to the leaving ability of phenol groups (β_{lg} = –1.2).   A typical kinetic consequence of the general acid/base mutation is that for substrates requiring strong acid assistance (such as pCPAF and mNPAF), the first reaction step is much slower (12000–14000-fold decreases were noted in this study) than is the case with substrates that need less acid assistance (such as 2,5-DNPAF).   The kinetic behavior displayed by the D299G mutant and the pH activity profile of D299N, when compared with data from the wild-type Abf, elegantly confirm the catalytic role of D299 and may be extended to corresponding residues in the other arabinofuranosidases of the GH54 family.

Table 3-1. Michaelis–Menten parameters for the hydrolysis of pNPAF by wild-type and mutant enzymes at pH 4.1 and 25 °C.

Enzyme	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m (M^{-1}s^{-1})
wild-type	21	0.32	65625
D170N	0.68	0.30	2250
D221N	0.003	0.22	14
E223Q	0.0005	0.88	0.57
E223G	22	0.37	59460
D299N	0.016	0.05	320
D299G	0.08	0.21	400
E310G	0.3	3.0	100

The kinetic measurements for wild-type and E223G were within 10% standard error, whereas for other mutants with much weaker activity, the errors were within 20%.

Table 3-2. K_m and k_{cat} values of wild-type Abf using various aryl-α-L-arabinofuranosides at pH 6.5

Unit: k_{cat} / K_m (M^{-1}s^{-1}), k_{cat} (s^{-1})

Substrate	pK$_a$	K_m, mM		k_{cat}		k_{cat} / K_m		log k_{cat}		log (k_{cat} / K_m)	
		WT	D299G	WT	D299G	WT	D299G	WT	D299G	WT	D299G
2,5-DNPAF	5.15	0.28	0.10	8.25	0.69	29542	6761	0.92	−0.16	4.47	3.83
CNPAF	6.45	0.42	0.20	6.52	0.12	15633	572	0.81	−0.94	4.19	2.76
pNPAF	7.18	0.31	0.21	2.84	0.014	9215	64	0.45	−1.86	3.96	1.81
mNPAF	8.39	0.54	0.65	2.02	0.0003	7369	0.46	0.31	−3.52	3.87	−0.33
pCPAF	8.49	0.30	0.69	1.70	0.0004	5592	0.58	0.23	−3.41	3.75	−0.25
PAF	9.99	0.28	–	1.23	–	4334	–	0.09	–	3.64	–

*: The D299G mutant enzyme is almost inactive towards PAF.
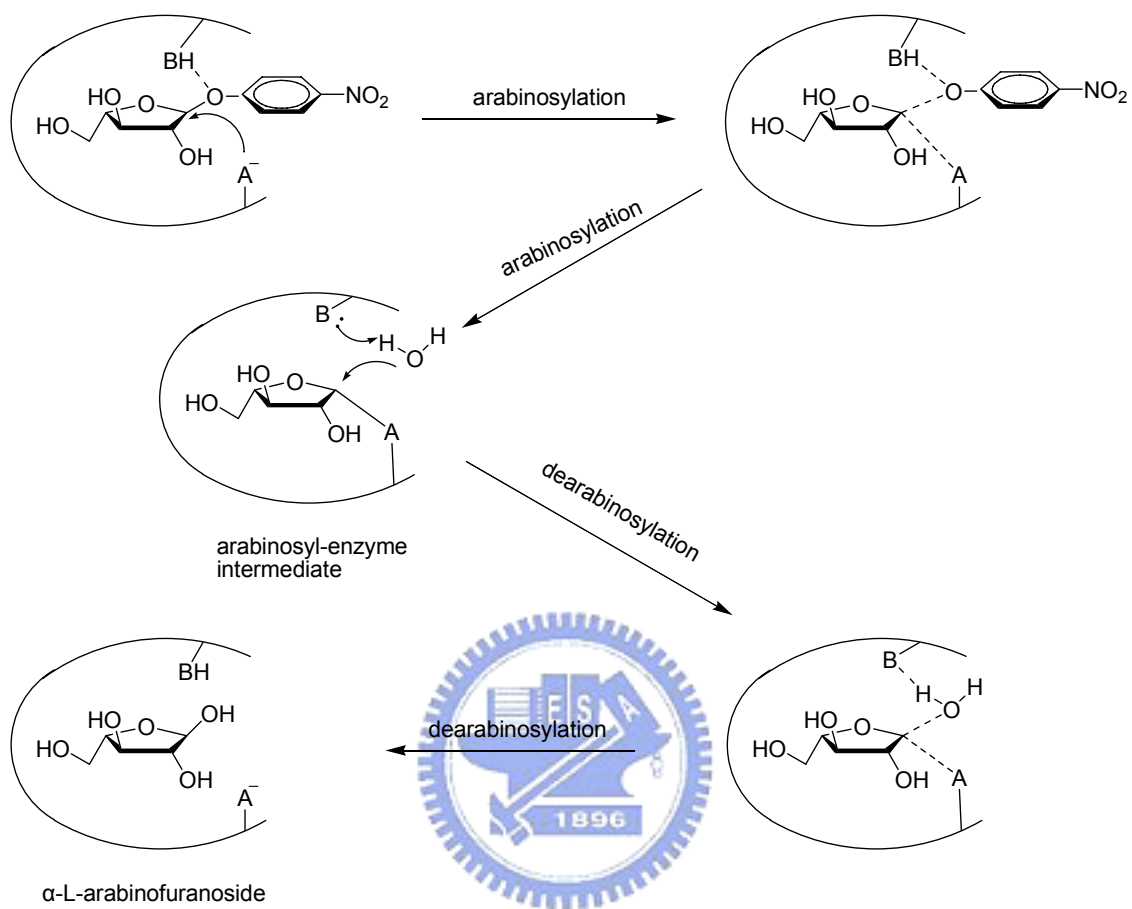
Figure 3-1. Proposed reaction mechanism of a retaining α-L-arabinofuranosidase.
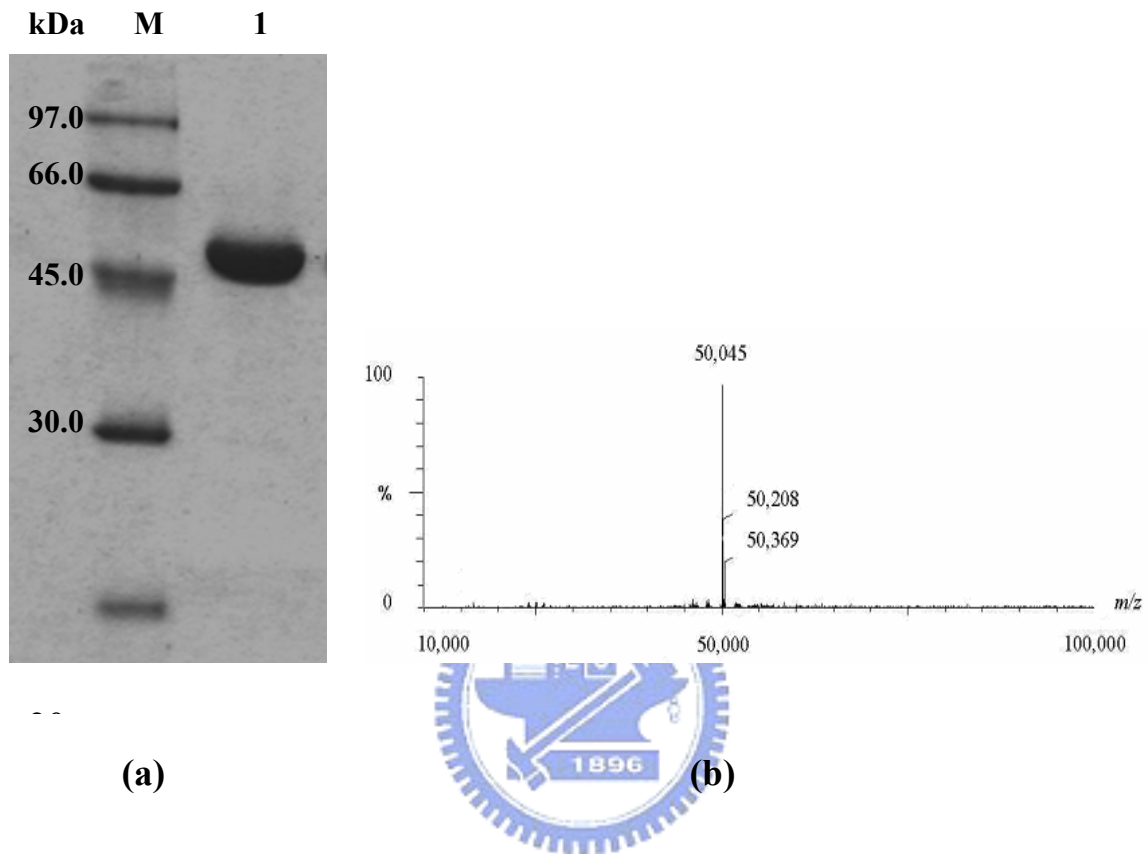
Figure 3-2. SDS-PAGE (a) and mass spectrometry (b) analysis of the recombinant α-L-arabinofuranosidase. Lanes: M, markers; 1, recombinant wild-type Abf.
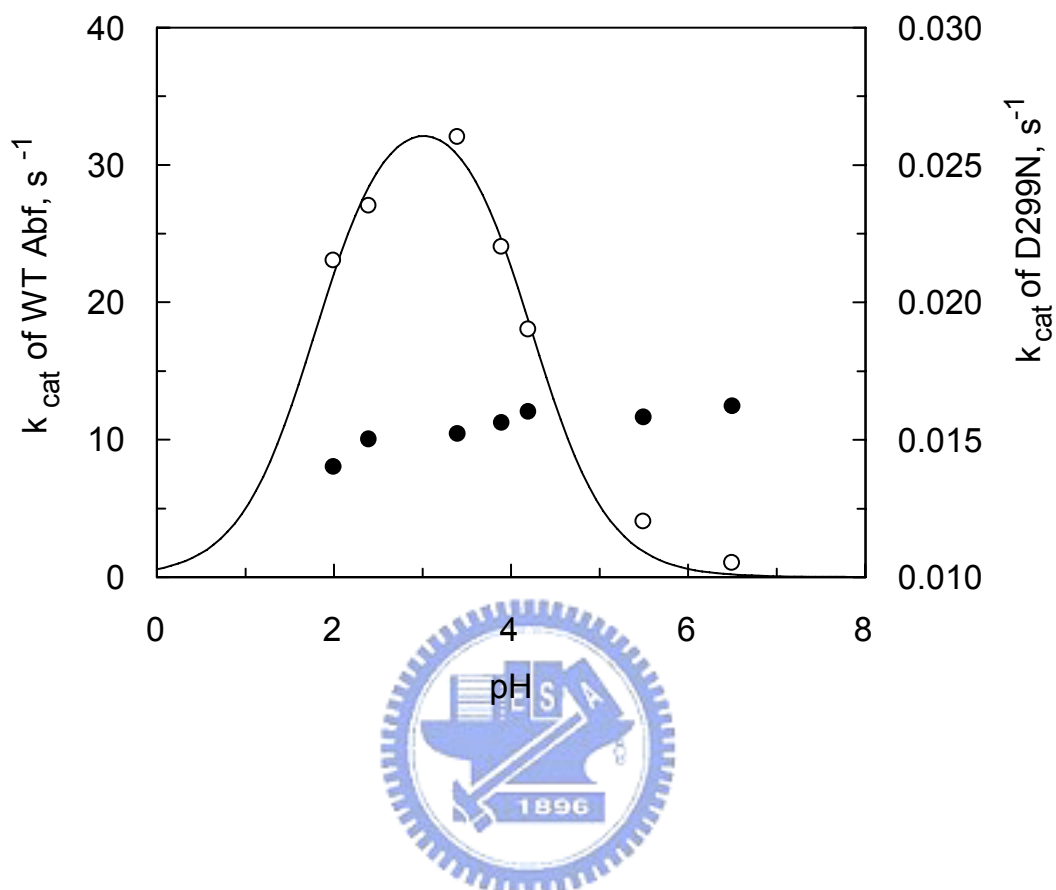
Figure 3-3. pH activity profiles of wild-type Abf (○) and the D299N mutant enzyme (●). The k_{cat} values of wild-type and D299N mutant were measured at the final pH values: 1.9, 2.0, 2.4, 3.3, 3.9, 4.2, 5.5, 6.5.
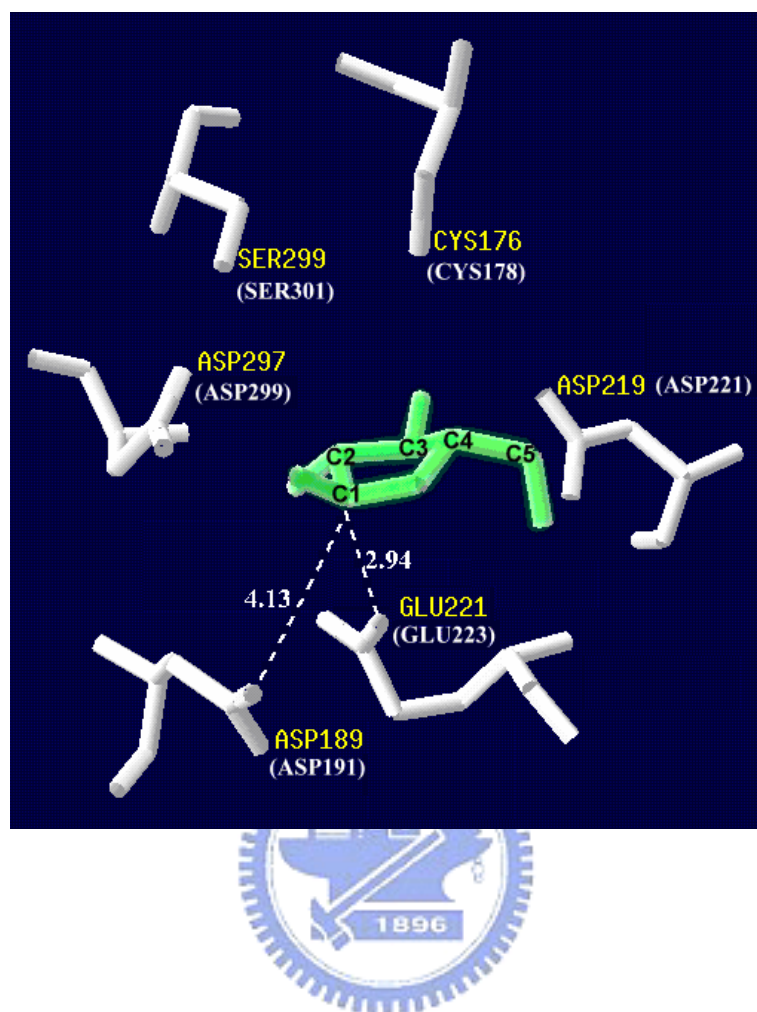
Figure 3-4. The active site of the GH54-family enzyme from *Aspergillus kawachii* IFO4308 (1WD4) with arabinofuranose in place. The corresponding amino acids in the *Trichoderma koningii* Abf are labeled in parentheses.

```
                     170         181 186  191            221 223          289    299    310
U38661     166: YAVLDGTHYNGACCFDYGNAETNSRDTGN 194   215:GPWIMADLENGL 226   286:MSKEGAIILGIGGDNSNGGQGTFYEGV
AB085904   164: YAVLDGTHYNDACCFDYGNAETSSTDTGA 192   213:GPWIMVDMENNL 224   284:MSKEGAIILGIGGDNSNGAQGTFYEGV
Z69252     166: YAVLDGTHYNGACCFDYGNAETNSRDTGN 194   215:GPWIMADLENGL 226   286:MSKEGAIILGIGGDNSNGAQGTFYEGV
AF367026   172: YAVLDGTHYNDACCFDYGNAETSSTDTGN 200   221:GPWVMADLENGL 232   292:MSKEGAIILGIGGDNSNGAQGTFYEGV
AB073861   172: YAVLDGTHYNSACCFDYGNAEVSNTDTGN 200   221:GPWIMADLENGL 232   292:MSKEGAIILGIGGDNSNGAQGTFYEGV
AB073860   172: YAVLDGTHYNSACCFDYGNAEVSNTDTGN 200   221:GPWIMADLENGL 232   292:MSKEGAIILGIGGDNSNGAQGTFYEGV
L23502     164: YAVLDGTHYNDACCFDYGNAETSSTDTGA 192   213:GPWIMVDMENNL 224   284:MSKEGAIILGIGGDNSNGAQGTFYEGV
U39942     164: YAVLDGTHYNDACCFDYGNWQTSSTDTGA 192   213:GPWLMVDMENNL 224   284:MSKEGAIILGIGGDNSNGAQGTFYEGV
Y13759     170: YAVLDGTHYNDGCCFDYGNAETSSLDTGN 198   219:GPWIMADLENGL 230   291:MSKEGAIILGIGGDNSNGAQGTFYEGA
AY495375   166: YAVLDGTHYNDACCFDYGNAEISNTDTGN 194   215:GPWLMADLENGL 226   286:MSLEGAIILGIGGDNSNGAQGTFYEGV
AJ310126   166: YAVLDGTHYNGACCFDYGNAETSSTDTGN 194   215:GPWIMADLENGL 226   285:MKKEGAIILGIGGDNSNGAQGTFYEGV
AF306764   167: YAVLDGTHYNGGCCFDYGNAETNNLDTGN 195   216:GPWVMADLENGL 227   287:MSKEGAIILGIGGDNSNGAQGTFTEGA
```

Figure 3-5. Data from a multialignment exercise, using partial sequences, of family GH54 α-L-arabinofuranosidases. Biology WorkBench 3.2 CLUSTALW (San Diego Supercomputer Center, CA, USA) software was used. All enzyme sequences were derived from published gene sequences. GenBank accession details are: U38661 from *Hypocrea koningii* G-39, AB085904 from *A. kawachii* IFO 4308, Z69252 from *Hypocrea jecorina RutC-30*, AF367026 from *Penicillium purpurogenum*, AB073861 from *Aspergillus oryzae* RIB40, AB073860 from *Aspergillus oryzae* HL15, L23502 from *Aspergillus niger*, U39942 from *Aspergillus niger*, Y13759 from *Emericella nidulans argB2*, AY495375 from *Aureobasidium pullulans*, AJ310126 from *Fusarium oxysporum f. sp. Dianthi*, and AF306764 from *Cochliobolus carbonum*.
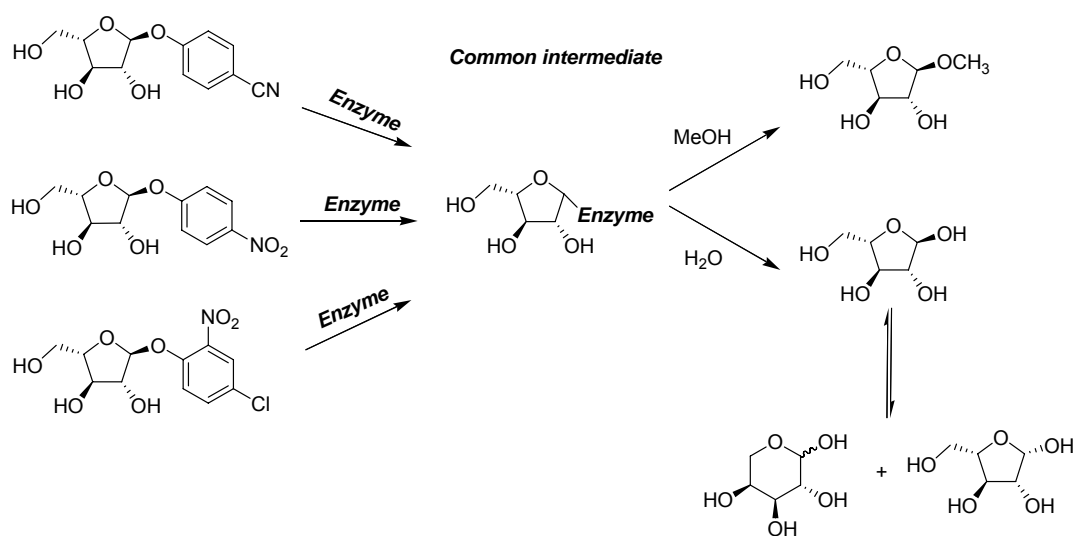
(a)
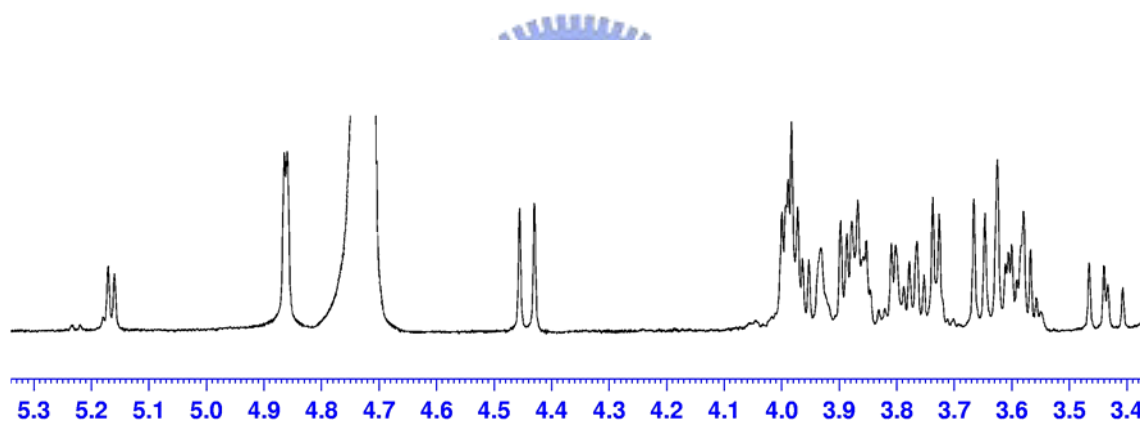
(b)

Figure 3-6. Stereochemical properties and common intermediates of Abf catalysis. (a) Enzymatic reactions, using various substrates, in the presence of methanol. (b) A partial NMR spectrum (chemical shift 3.4–5.4 ppm) of the end-products. Peak assignment is given in the text.
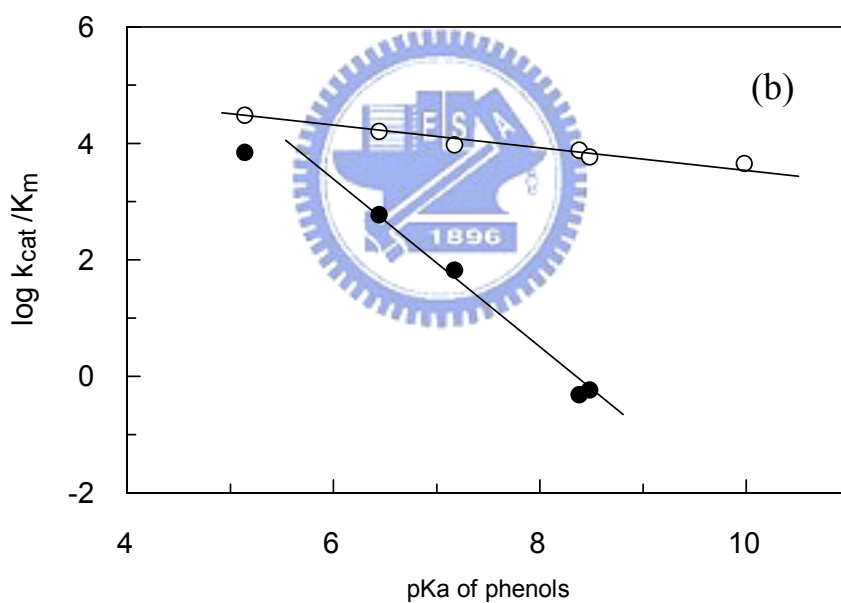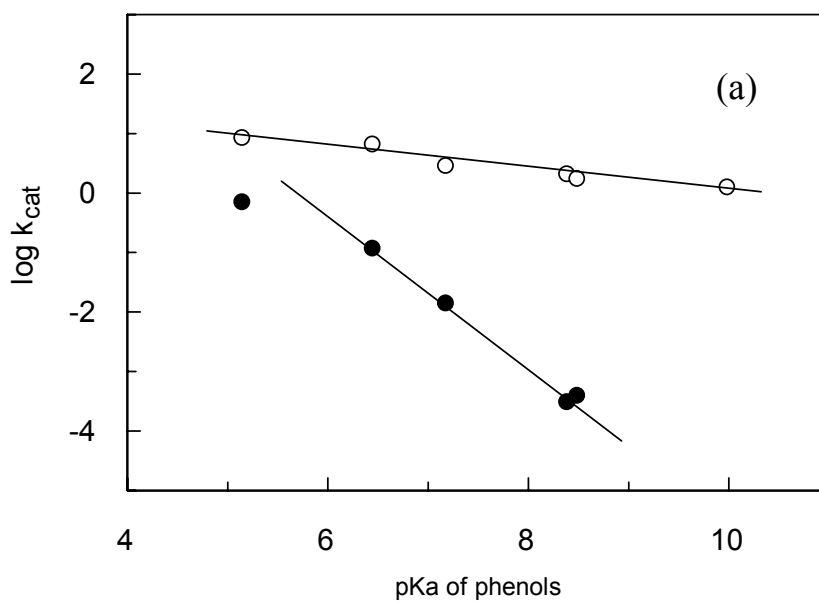
Figure 3-7. Brønsted plots of wild-type Abf (○) and D299G (●) mutant enzyme. (a) Plots of log k_{cat} against pK_a of the leaving phenol. (b) Plots of log k_{cat}/K_m against pK_a of the leaving phenol.

Chapter 4
Conclusions and prospects

I describe the purification and characterization of the biotechnologically important α-L-arabinofuranosidase/β-D-xylosidase from *Trichoderma koningii* G-39, as well as its function and overexpression in the methylotrophic yeast *P. pastoris*. I also provide an identification of the active site nucleophile (E223) and general acid/base (D299) in this enzyme by utilizing the site-directed mutagenesis, kinetic analysis and three-dimensional structure simulation. The purified enzyme exhibited both α-L-arabinofuranosidase and β-D-xylosidase activities with considerable higher specific α-L-arabinofuranosidase activity than β-D-xylosidase activity. The optimun temperature was about the same for the two enzyme activities. However, the optimun pH were 3.0~3.5 and 4.0~4.5 for pNPAF and DNXP, respectively. In addition, inhibitors, which mimic the proposed furanose structure and pyranose structure, bind with different binding domains. To the best of our knowledge, these results presented the first kinetic study performed on family 54 α-L-arabinofuranosidase/β-D-xylosidase.

Study of α-L-arabinofuranosidase/β-D-xylosidase has provided major insights and understanding on the mechanistic of this enzyme. Today, the reason why the E223G mutant possesses extremely high activity remains unknown. High enzyme activity indicated that the C-1 position of the substrate is perfect for an inverting-type of catalysis to take place. It was very likely that when the side-chain of the nucleophile is replaced by proton (ie, E223G), there were more space to allow water molecules to diffuse further into the active site. This hypothesis is under investigation by analyzing the activity of double mutant (D191N/E223G) and studying the stereochemistry of E223G catalysis.

Reference

1. Betts, W. B., Dart, R. K., Ball, A. S., and Pedlar, S. L. (1992) Biosynthesis and structure of lignocellulose. In: Biodegradation, Natural and Synthetic Materials. (Betts, W.B. ed.). London, Springer-Verlag, pp. 139-156

2. Whistler, R. L., and Richards, E. L. (1970) Hemicellulose in the carbohydrates chemistry and biochemistry. 2nd edition, Vol. 2A (Eds. Pigman, W. and Horton, D.). New York, Academic Press. pp. 447-469

3. Dekker, R. F. H., and Richards, G. N. (1976) Hemicellulases: Their occurrence, purification, properties and mode of action. *Adv. Carbohydr. Chem. Biochem.* **32**, 277-352

4. Barry, V. C., and Dillon, T. (1940) The occurrence of xylan in marine algae. *Nature*, **146**, 620, London

5. Kato, Y., and Nevins, D. (1984) Enzymic dissociation of zea shoot cell wall polysaccharides: (III) purification and partial characterization of an endo-(1,4)β-D-xylanase from a *Bacillus subtilis* enzyme preparation. *Plant Physiol.* **75**, 753-758

6. Joseleau, J. P., Comtat, J., and Ruel, K. (1992) Chemical structure of xylans and their interaction in the plant cell walls. In: Xylans and Xylanases. (Eds. Visser, J., Beldman,G., Kusters-van-Someren, M.A., and Varagen, A.G.J.) Amsterdam, London, New York.,Tokyo. Elsevier. pp. 1-15

7. Shallom,D., and Shoham, Y. (2003) Microbial hemicellulases. *Curr. Opin. Microbiol.* **6**, 219-228

8. Coughlan, M. P., and Hazlewood, G. P. (1993) β-1,4-D-xylan degrading enzyme systems: Biochemistry, molecular biology and applications.

Biotechnol. Appl. Biochem. **17**, 259-289

9. Thomson, J. A. (1993) Molecular biology of xylan degradation. *FEMS Microbiol. Rev*. **104**, 65-82

10. Bachmann, S. L., and McCarthy, A. J. (1991) Purification and cooperative activity of enzymes constituting the xylan-degrading system of *Thermomonospora fusca*. *Appl Environ Microbiol*. **57**, 2121–2130

11. Greve, L. C., Labavitch, J. M., and Hungate, R. E. (1994) α-L-Arabinofuranosidase from *Ruminococcus albus* 8: purification and possible roles in hydrolysis of alfalfa cell wall. *Appl Environ Microbiol*. **47**, 1135–1140

12. Saha, B. C. (2000) α-L-Arabinofuranosidases: biochemistry, molecular biology and application in biotechnology. *Biotechnol. Adv*. **18**, 403- 423

13. Saha, B. C., and Bothast, R. J. (1999) Enzymology of xylan degradation. In: Imam, S. H., Greene, R.V., and Zaidi, B. editors. Biopolymers: Utilizing Nature's Advanced Materials. Washington, DC: *American Chemical Society*, 167–194

14. Henrissat, B., and Baroch, A. (1996) Updating the sequence-based classification of glycosyl hydrolases. *Biochem. J*. **316**, 695–696

15. Matsuo, N., Kaneko, S., Kuno, A., Kobayashi, H. and Kusakabe, I. (2000) Purification, characterization and gene cloning of two α-L-arabinofuranosidases from *Streptomyces chartreusis* GS901. *Biochem. J*. **346**, 9-15

16. McKie, V. A., Black, G. W., Millward-Sadler, S. J., Hazlewood, G. P., Laurie, J. I. and Gilbert, H. J. (1997) Arabinanase A from *Pseudomonas fluorescens* subsp. *Cellulose* exhibits both an endo- and an exo-mode of action. *Biochem. J*. **323**, 547-555

17. Kellet, L. E., Poole, D. M., Ferreira, L. M. A., Durrant, A. J., Hazlewood, G. P. and Gilbert, H. J. (1990) Xylanase B and an arabinofuranosidase from *Pseudomonas fluorescens* subsp. *cellulosa* contain identical cellulose-binding domains and are encoded by adjacent genes. *Biochem. J.* **272**, 369-376

18. Flipphi, M. J. A., Visser, J., van der Veen, P. and de Graff, L. H. (1994) Arabinase gene expression in *Aspergillus niger* : indications for coordinated regulation. *Microbiology*, **140**, 2673-2682

19. Beldman, G., Schols, H. A., Pitson, S. M., Searle-van Leeuwen, M. J. F., and Voragen, A. G. J. (1997) Arabinans and arabinan degrading enzymes. *Adv. Macromol. Carbohydr. Res*. **1,** 1– 64

20. Komae, K., Kaji, A., and Sato, M. (1982) An α-L-arabinofuranosidase from *Streptomyces purpurascens* IFO 3389. *Agric. Biol. Chem*. **46**, 1899–1905

21. Weinstein, L., and Albersheim, P. (1979) Structure of Plant Cell Walls: IX. Purification and partial characterization of a wall-degrading Endo-arabanase and an arabinosidase from *Bacillus subtilis. Plant Physiol*. **63**, 425–432

22. Kaji, A., and Tagawa, K. (1970) Purification, crystallization and amino acid composition of α-L-arabinofuranosidase from *Aspergillus niger. Biochim. Biophy*s. *Acta.* **207**, 456– 464

23. Rombouts, F. M., Voragen, A. G. J., Searle-van Leeuwen, M. F. C., Geraerds, C. J. M., Schols, H. A., and Pilnik, W. (1988) The arabinases of *Aspergillus niger*—purification and characterization of two α-L-arabinofuranosidases and an *endo*-1,5-α-L-arabinase. *Carbohydr. Polym*. **9**, 25– 47

24. Kormelink, F. J. M., Searle-van Leeuwen, M. J. F., Wood, T. M., and

Voragen, A. G. J. (1991) Purification and characterization of a (1,4)-β-D-arabinoxylan arabinofuranohydrolase from *Aspergillus awamori*. *Appl. Microbiol. Biotechnol*. **35**, 753–758

25. Van Laere, K. M. J., Beldman, G., and Voragen, A. G. J. (1997) A new arabinofuranohydrolase from *Bifidobacterium adolescentis* able to remove arabinosyl residues from double-substituted xylose units in arabinoxylan. *Appl. Microbiol. Biotechnol*. **47**, 231–235

26. Flipphi, M. J. A., van Heuvel, M., van der Veen, P., Visser, J., and De Graaff, L. H. (1993) Cloning and characterization of the abfB gene coding for the major alpha-L-arabinofuranosidase (ABF B) of *Aspergillus niger*. *Curr. Genet*. **24**, 525– 532

27. Gielkens, M. M., Visser, J., and De Graaff, L. H. (1997) Arabinoxylan degradation by fungi: characterization of the arabinoxylan-arabinofuranohydrolase encoding genes from *Aspergillus niger*. and *Aspergillus tubigensis*. *Curr. Genet*. **31**, 22–29

28. Gilead, S., and Shoham, Y. (1995) Purification and characterization of α-L-arabinofuranosidases from *Bacillus stearothermophilus* T-6. *Appl Environ Microbiol*. **61**, 170–174

29. Manin, C., Shareek, F., Morosoli, R., and Kluepfel, D. (1994) Purification and characterization of an α-L-arabinofuranosidase from *Streptomyces lividans* 66 and DNA sequence of the gene (*abfA*). *Can Biochem. J*. **302**, 443–449

30. Kaji, A., and Shimokawa, K. (1984) New exo-type arabinase from *Erwina carotovora* IAM 1024. *Agric Biol Chem*. **48**, 67–72

31. Kaneko, S., Arimoto, M., Ohba, M., Kobayashi, H., Ishii, T., and Kusakabe, I. (1998) Purification and substrate specificities of two α-L-arabinofuranosidases from *Aspergillus awamori* IFO 4033. *Appl.*

Environ. Microbiol., **64**, 4021–4027

32. Ramon, D., Veen, Pvd., and Visser, J. (1993) Arabinan degrading enzymes from *Aspergillus nidulans*: induction and purification. *J Biotechnol*. **113**, 15–22

33. Luonteri, E., Siika-aho, M., Tenkanen, M., and Vikari, L. (1995) Purification and characterization of three α-arabinosidases from *Aspergillus terreus*. *J Biotechnol*. **38**, 279–291

34. Filho, EXF., Puls, J., and Coughlan M. P. (1996) Purification and characterization of two arabinofuranosidases from solid-state cultures of the fungus *Penicillium capsulatum*. *Appl Environ Microbiol*. **62**, 168–173

35. Saha, B. C., and Bothast, R. J. (1998) Purification and characterization of a novel thermostable α-ʟ-arabinofuranosidase from a color-variant strain of *Aureobasidium pullulans*. *Appl Environ Microbiol*. **64**, 216–220

36. De loannes, P., Peirano, A., Steiner, J., and Eyzaguirre, J. (1996) An α-ʟ-arabinofuranosidase from *Penicillium purpurogenum*: production, purification and properties. *J. Biotechnol*. **76**, 253–258

37. Bezalel, L., Shoham, Y., and Rosenberg, E. (1993) Characterization and delignification activity of a thermostable α-ʟ-arabinofuranosidase from *Bacillus stearothermophilus*. *Appl. Microbiol. Biotechnol*. **40**, 57–62

38. Uesaka, E., Sato, M., Raiju, M., and Kaji, A. (1978) α-ʟ-Arabinofuranosidase from *Rhodotorula flava*. *J. Bacteriol*. **133**, 1073–1077

39. Kaji, A., and Yoshihara, O. (1971) Properties of purified α-ʟ-arabinofuranosidase 48. from *Corticium rolfsii*. *Biochim Biophys Acta*. **250**, 367–371

40. Tuohy, M. G., Laffey, C. D., and Coughlan, M. P. (1994) Characterization of the individual components of the xylanolytic enzyme system of *Talaromyces emersonii*. *Bioresour Technol*. **50**, 37–42

41. Margolles-Clark, E., Tenkanen, M., Nakari-Setälä, T., and Penttilä, M. (1996) Cloning of genes encoding α-L-arabinofuranosidase and β-xylosidase from *Trichoderma reesei* by expression in *Saccharomyces cerevisiae*. *Appl Environ Microbiol*. **62**, 3840–3846

42. Kimura, I., Yoshioka, N., Kimura, Y., and Tajima, S. (2000) Cloning, sequencing and expression of an α-L-arabinofuranosidase from *Aspergillus sojae*. *J. Ferment Bioeng*. **89**, 262–266

43. Debeche, T., Cummings, N., Connerton, I., Debeire, P., and O'Donohue, M. J. (2000) Genetic and biochemical characterization of a highly thermostable α-L-arabinofuranosidase from *Thermobacillus xylanilyticus*. *Appl Environ Microbiol*. **66**, 1734–1736

44. Kaneko, S., Sano, M., and Kusakabe, I. (1994) Purification and some properties of α-L-arabinofuranosidase from *Bacillus subtilis* 3-6. *Appl Environ Microbiol*. **60**, 3425–3428

45. Schwarz, W. H., Bronnenmeier, K., Krause, B., Lottspeich, F., and Staudenbauer, W. L. (1995) Debranching of arabinoxylan: properties of the thermoactive recombinant α-L-arabinofuranosidase from *Clostridium stercorarium* (ArfB). *Appl Microbiol Biotechnol*. **43**, 856–860

46. Suurnakki, A., Tenkanen, M., Buchert, J., and Viikari, L. (1997) Hemicellulases in the bleaching of chemical pulps, *Adv. Biochem. Eng. Biotechnol*. **57**, 261-287

47. Wong, K. K. Y., and Sanddler, J. N. (1993) *Hemicellulose and*

Hemicellulases (Coughlan, M. P. and Hazlewood, G. P., Eds.), pp. 127-143, Portland Press, London.

48. Sanchez-Torres, P., Gonzalez-Candelas, L., and Ramon, D. (1996) Expression in a wine yeast strain of the *Aspergillus niger* abfB gene. *FEMS Microbiol. Lett*. **145**, 189-194

49. Spagna, G., Romagnoli, D., Martino, A., Bianchi, G., and Piteri, P.G. (1998) A simple method for purifying glycosidases: alpha-L-arabinofuranosidase and beta-D-glucopyranosidase from *Aspergillus niger* to increase the aroma of wine. Part I. *Enzyme Microb. Technol*. **22**, 298-304

50. Martino, A., Schiraldi, C., Di Lazzaro, A., Fiume, I., Piteri, P.G., and De Rosa, M. (2000) Improvement of the flavor of Falanghina white wine using a purified glycosidase preparation from *Aspergillus niger. Process Biochem*. **36**, 93-102

51. Sinnott, M. L. (1990) Glycosidases and glycosyl transferases. *Chem. Rev*., **90**, 1171-1202

52. McCarter, J., and Withers, S. G. (1994) Mechanisms of enzymatic glycoside hydrolysis. *Curr. Opin. Struct. Biol*., **4**, 885–892

53. Davies, G., Sinnott, M. L., and Withers, S. G. (1998) Glycosyl transfer, in *Comprehensive Biological Catalysis* (Sinnott, M. L., ed) Vol. 1, pp. 119–208. Academic Press, London

54. Ly, H. D., and Withers, S. G. (1999) Mutagenesis of glycosidases. *Annu. Rev. Biochem*., **68**, 487-522

55. Zechel, D. L., and Withers, S. G. (2000) Glycosidase mechanisms: anatomy of a finely tuned catalyst. *Acc. Chem. Res*., **33**, 11-18

56. Henrissat, B., and Davies, G. (1997) Structural and sequence-based classification of glycoside hydrolases. *Curr. Opin. Struct. Biol.* **7,**

637-644

57. Pitson, S.M., Voragen, A.G., and Beldman, G. (1996) Stereochemical course of hydrolysis catalyzed by arabinofuranosyl hydrolases. *FEBS Lett*., **398**, 7-11

58. Ferre, H., Broberg, A., Duus, J.O. and Thomsen, K.K. (2000) A novel type of arabinoxylan arabinofuranohydrolase isolated from germinated barley analysis of substrate preference and specificity by nano-probe NMR. *Eur. J. Biochem.* **267**, 6633-6641

59. Beylot, M. H., McKie, V. A., Voragen, A. G., Doeswijk-Voragen, C. H. L., and Gilbert, H.J. (2001) The Pseudomonas cellulosa glycoside hydrolase family 51 arabinofuranosidase exhibits wide substrate specificity. *Biochem. J*., **358**, 607-614

60. Kempton, J. B., and Withers, S. G. (1992) Mechanism of *Agrobacterium* beta-glucosidase: kinetic studies. *Biochemistry*, **31**, 9961- 9969

61. Zechel, D. L., and Withers, S. G. (1999) Comprehensive Natural Products Chemistry, 5 Eds. (Poulter, C. D. Ed.), pp. 279-314. Elsevier, New York

62. Withers, S. G., and Aebersold, R. (1995) Approaches to labeling and identification of active site residues in glycosidases. *Protein Sci*., **4**, 361-372

63. Williams, S. J., and Withers, S. G. (2000) Glycosyl fluorides in enzymatic reactions. *Carbohydr. Res.,* **327**, 27-46

64. MacLeod, A. M., Tull, D., Rupitz, K., Warren, A. J. and Wither, S. G. (1996) Mechanistic Consequences of Mutation of Active Site Carboxylates in a Retaining β-1,4-Glycanase from *Cellulomonas fimi. Biochemistry,* **35**, 13165-13172

65. Takoua Debeche, Christophe Bliard, Philippe Debeire and Michael

J.O'Donohue (2002) Probing the catalytically essential residues of the α-L-arabinofuranosidase from *Thermobacillus xylanilyticus. Protein Engineering*, **15**, 21-28

66. Vasella, A., Davies, G. J., and Bohm, M. (2002) Glycosidase mechanisms. *Curr. Opin. Chem. Biol.,* **6**, 619-629

67. Shallom, D., Belakhov, V., Solomon, D., Gilead-Gropper, S., Baasov, T., Shoham, G., and Shoham, Y. (2002) The identification of the acid/base catalyst of α-arabinofuranosidase from *Geobacillus stearothermophilus* T-6, a family 51 glycoside hydrolase. *FEBS Lett.*, **514**, 163-167

68. Shallom, D., Belakhov, V., Solomon, D., Shoham, G., Baasov, T., and Shoham, Y. (2002) Detailed kinetic analysis and identification of the nucleophile in α-L-arabinofuranosidase from *Geobacillus stearothermophilus* T-6, a family 51 glycoside hydrolase. *J. Biol. Chem.*, **277**, 43667-43673

69. Hovel, K., Shallom, D., Niefind, K., Belakhov, V., Shoham, G., Baasov, T., Shoham, Y., and Schomburg, D. (2003) Crystal structure and snapshots along the reaction pathway of a family 51 α-L-arabinofuranosidase. *The EMBO journal*., **22**, 4922-4932

70. Miyanaga, A., Koseki, T., Matsuzawa, H., Wakagi, T., Shoun, H., and Fushinobu, S. (2004) Crystal structure of a family 54 alpha-L-arabinofuranosidase reveals a novel carbohydrate-binding module that can bind arabinose. *J. Biol. Chem.*, **279**, 44907-44914

71. Bacic, A., Harris, P. J., and Stone, B. A. (1988) Structure and function of plant cell walls. In The Biochemistry of Plants, Vol. 14 (Preiss, J., ed.), pp. 297, Academic Press, San Diego

72. Sunna, A., and Antranikian, G. (1997) Xylanolytic enzymes from fungi and bacteria. *Crit. Rev. Biotechnol.*, **17**, 39-67

73. Biely, P. (1985) Microbial xylanolytic systems. *Trends Biotechnol.*, **3**, 286–290

74. Hazlewood, G. P., and Gilbert, H. J. (1998) Structure and function analysis of Pseudomonas plant cell wall hydrolases. *Prog. Nucleic Acids Res. Mol. Biol.*, **61,** 211

75. Sharma, S. K., Corrales, G., and Penades, S. (1995) Single step stereoselective synthesis of unprotected 2,4-dinitrophenyl glycosides. *Tetrahedron Lett.*, **36**, 5627-5630

76. Laemmli, U. K. (1970) Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature* (London) **227**, 680-685

77. Lee, R. C., Hrmova, M., Burton, R. A., Lahnstein, J., and Fincher, G. B. (2003) Bifunctional family 3 glycoside hydrolases from barley with alpha-L-arabinofuranosidase and beta-D-xylosidase activity. characteritization, primary structures, and COOH-termnal processing., *J. Biol. Chem.*, **278**, 5377-5387

78. Minic. Z., Rihouey, C., Do, C. T., Lerouge, P., and Jouanin, L. (2004) Purification and characterization of enzymes exhibiting β-D-xylosidase activities in stem tissues of arabidopsis**,** *Plant Physiol.*, **135**, 867-878

79. Sakka, K., Yoshikawa, K., Kojima, Y., Karita, S., Ohmiya, K., Shimada, K. (1993) Nucleotide sequence of the *Clostridium stercorarium xylA* gene encoding a bifunctional protein with β-D-xylosidase and α-L-arabinofuranosidase activities, and properties of the translated product. *Biosci Biotechnol Biochem.*, **57**, 268-272

80. Kawabata, Y., Kaneko, S., Kusakabe, I., and Gama, Y. (1995) Synthesis of regioisomeric methyl α-L-arabinofuranobiosides. *Carbohydr. Res.,* **267**, 39-47

81. Li, Y-K., Yao, H-J., and Cho, Y-T. (2000) Effective induction, purification and characterization of *Trichoderma koningii* G-39 β-xylosidase with high transferase activity, *Biotechnol. Appl. Biochem.*, **31**, 119.

82. Henrissat, B., and Bairoch, A. (1996) Updating the sequence-based classification of glycosyl hydrolases. *Biochem. J.*, **316**, 695–696

83. Phillips, D. C. (1967) The hen egg white lysozyme molecule. *Proc. Natl Acad. Sci.*, USA, **57**, 484-495

84. Vocadlo, D. J., and Withers, S. G. (2005) Detailed Comparative Analysis of the Catalytic Mechanisms of β-N-Acetylglucosaminidases from Families 3 and 20 of Glycoside Hydrolases. *Biochemistry*, **44**, 12809–12818

85. Vocadlo, D. J., Davies, G. J., Laine, R., and Withers, S. G. (2001) Catalysis by hen egg-white lysozyme proceeds via a covalent intermediate. *Nature*, **412**, 835-838

86. Koshland, D. E. (1953) Stereochemistry and mechanism of enzymatic reactions. *Biol. Rev.*, **28**, 416-436

87. Kelly, M. A., Sinnott, M. L., and Widdows, D. (1988) Preparation of some aryl α-L-arabinofuranosides as substrates for arabinofuranosidase. *Carbohydr. Res.*, **181**, 262–266

88. Viladot, J.-L., Canals, F., Batllori, X., and Planas, A. (2001) Long-lived glycosyl-enzyme intermediate mimic produced by formate re-activation of a mutant endoglucanase lacking its catalytic nucleophile. *Biochem. J.*, **355**, 79–86

89. Williams, S. J., Mark, B. L., Vocadlo, D. J., James, M. N. G., and Withers, S. G. (2002) Aspartate 313 in the *Streptomyces plicatus* hexosaminidase plays a critical role in substrate-assisted catalysis by

orienting the 2-acetamido group and stabilizing the transition state. *J. Biol. Chem.*, **277**, 40055–40065

90. Cobucci-Ponzano, B., Mazzone, M., Rossi, M., and Moracci, M. (2005) Probing the catalytically essential residues of the α-L-fucosidase from the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Biochemistry*, **44**, 6331–6342

91. Shallom, D., Leon, M., Bravman, T., Ben-David, A., Zaide, G., Belakhov, V., Shoham, G., Schomburg, D., Baasov, T., and Shoham, Y. (2005) Biochemical characterization and identification of the catalytic residues of a family 43 β-D-xylosidase from *Geobacillus stearothermophilus* T-6. *Biochemistry*, **44**, 387–97

92. Burmeister, W. P., Cottaz, S., Rollini, P., Vasella, A., and Henrissat, B. (2000) High resolution x-ray crystallography shows that ascorbate is a cofactor for myrosinase and substitutes for the function of the catalytic base. *J. Biol. Chem.*, **275**, 39385–39393

93. McCarter, J. D., Yeung, W., Chow, J., Dolphin, D., and Withers, S. G. (1997) Design and synthesis of 2'-deoxy-2'-fluoro disaccharides as mechanism-based glycosidase inhibitors that exploit aglycone specificity. *J. Am. Chem. Soc.*, **119**, 5792–5797

94. Ducros, V., Charnock, S. J., Derewenda, U., Derewenda, Z. S., Dauteri, Z., Dupont, C., Shareck, F., Morosoli, R., Kluepfel, D., and Davies, G. J. (2000) Substrate specificity in glycoside hydrolase family 10 : structure and kinetic analysis of the *Streptomyces lividans* xylanase 10A. *J. Biol. Chem.*, **275**, 23020–23026

95. Li, Y.-K., Chir, J. and Chen, F.-Y. (2001) Catalytic mechanism of a family 3 β-glucosidase and mutagenesis study on residue Asp-247. *Biochem. J.*, **355**, 835–840

96. Li, Y.-K., Yao, H.-J., and Pan, I.-H. (2000) Mechanistic study of β-xylosidase from *Trichoderma koningii* G-39. *J. Biochem.* (Tokyo), **127**, 315–320

97. Klimacek, M., Kavangeh, K. L., Wilson, D. K., and Nidzteky, B. (2003) On the role of Brønsted catalysis in Pseudomonas fluorescens mannitol 2-dehydrogenase. *Biochem. J.*, **375**, 141–149

Appendices

Appendix 1

Table 1: comparative properties of some fungal α-L-arabinofuranosidases

Organism	MW (kDa)	pI	Optimum Temp(℃)	Optimum pH	Polymer attacked[a]	Ref
Aspergillus awamori I	81	3.3	60	4.0	BA	
II	62	3.6	60	4.0	BA	
Aspergillus nidulans	65	3.3	65	4.0		
Aspergillus niger 5-16						
Intracellular	67	3.5	60	4.0	BA	
Extracellular	53	3.6	-	4.0	BA,AX	
Aspergillus sojae	34.3	3.9	50	5.0		
Aspergillus terreus I	39	7.5	-	3.5-4.5	OSX,RAX,AGX,BA	
II	59	8.3	-	3.5-4.5	OSX,RAX,AGX,BA	
III	59	8.5	-	3.5-4.5	OC,RAX,AGX,BA	
Aureobasidium pullulans	210	-	75	4.0-4.5	BA,AX,OSX	
Cochliobolus carbonum	63	-	50	3.5-4.0	BA,AX	
Cytophaga xylanolytica	160-240	6.1	45	5.8	AX,BA	
Dichomitus squalens	60	5.1	60	3.5	BA,AX	
Penicillium capsulatum I	64.5	4.15	60	4.0	BA,AX,AXO	
II	62.7	4.54	55	4.0	BA,AX,AXO	
Penicillium purpurogenum	58	6.5	50	4.0	AX	
Sclerotinia sclerotiorum	63	7.5	-	4.0-4.5	BA,AX	
Trichoderma reesei	53	7.5		4.0	AX,AXO	

[a] AX, arabinoxylan; BA, beet arabinan; CX, corn endosperm xylan; OSX, oatspelt xylan; AXO, arabinoxylan oligosaccharides; RAX, rye arabinoxylan; AGX,arabinoglucuronoxylan.

Appendix 1 (continue)

Table 2: comparative properties of some bacterial α-L-arabinofuranosidases

Organism	MW (kDa)	pI	Optimum Temp(℃)	Optimum pH	Polymer attacked[a]	Ref
Bacillus polymyxa	166	4.7	-	-	AXO only	
Bacillus stearothermophilus	110	-	70	7.0	Used for delignification	
Bacillus subtilis	65	5.3		6.5	BA	
Bacteroides xylanolyticus	364 (61)		50	5.0-6.0	AXO only	
Bifidobacterium adolescents			30-40	6.0	AX only	
Butyrivibrio fibrisolvens	240	6.0	55	6.0-6.5	BA, AX, OSX	
Clostridium acetobutylicum	94	8.2		5.0-5.5	BA	
Ruminococcus albus	310	3.8		6.9	AH	
Streptomyces sp. 17-1	92	4.4		6.0	BA, AX, AG	
Streptomyces diastaticus C1	38	8.8		5.0-6.5	AX	
C2	60	8.3		5.0-6.5	AX	
Streptomyces dastaatochromogenes	73			6.0	BA	
Streptomyces purpurascens	495	3.9		6.5	A_2, A_3	

[a] AX, arabinoxylan; BA, beet arabinan; CX, corn endosperm xylan; OSX, oatspelt xylan; AXO, arabinoxylan oligosaccharides; A_2, arabinobiose；A_3, arabinotriose

Table 3: comparative properties of some plant α-L-arabinofuranosidases

Organism	MW (kDa)	pI	Optimum Temp(℃)	Optimum pH	Polymer attacked[a]	Ref
Spinach leaf I	118	4.8				
II	68	4.2		4.8		
Carrot cell culture	84	5.6	55	3.8	BA	
Radish seeds	64	4.7			BA, AG	
Soybean seeds	87	4.8				
Japanese pear fruit	42			5.0		

[a] BA, beet arabinan; AG, arabinogalactan

Table 4. Some Properties of Bifunctional β-D-Xylosidase / α-L-Arabinofuranosidase from Different Sources

	Clostridium stercorarium	*Thermoanaerobacter ethanolyticus*	*T. reesi*	*Radish*	*Spinach (Ara I)*
MW (Da)	212,000 four subunits	170,000 two subunits	100,000	64,000	118,000
Optimum pH	7.0	5.0-5.2	4.0	4.5	4.8
pI			4.7	4.7	4.2
p-Nitrophenyl-α-L-arabinofuranoside					
K_m (mM)	17.6	4.6	>2.5	9.7	1.2
K_{cat} (min^{-1})	$8.85*10^2$	$9.10*10^4$		$4.67*10^3$	73.2
p-Nitrophenyl-β-D-xylopyranoside					
K_m (mM)	2.5	0.038	0.08	0.95	
K_{cat} (min^{-1})	$3.13*10^2$	$1.56*10^4$		$3.56*10^2$	
Reference	(14)	(14)	(14)	(14)	(14)

Appendix 2

CLUSTAL W (1.81) multiple sequence alignment
Consensus key (see documentation for details)
* - single, fully conserved residue
  - no consensus


```
U38661      -----MLSNARIIAAGCI-AAGSLVAAGPCDIYSSGGTPCVAAHSTTRALFSAYTGPLYQ
Z69252      -----MLSNARIIAAGCI-AAGSLVAAGPCDIYSSGGTPCVAAHSTTRALFSAYTGPLYQ
AF367026    MLPRLSFDHAWVLALGLV-AKGTLVFAGPCDIYSSGGTPCVAAHSTTRALYSSFTGALYQ
AB073861    MSSGLSLERACAVALGIV-ASASLVAAGPCDIYSSGGTPCVAAHSTARALYSAYTGALYQ
AB073860    MSSGLSLERACAVALGIV-ASASLVAAGPCDIYSSGGTPCVAAHSTTRALYSAYTGALYQ
AB046701    -----MFSRRNLLALGLA-ATVS---AGPCDIYEAGDTPCVAAHSTTRALYSSFSGALYQ
AB085904    -----MFSRRNLLALGLA-ATVS---AGPCDIYEAGDTPCVAAHSTTRALYSSFSGALYQ
L23502      -----MFSRRNLVALGLA-ATVS---AGPCDIYEAGDTPCVAAHSTTRALYSSFSGALYQ
U39942      -----MFSRRNLVALGLA-ATVS---AGPCDIYEAGDTPCVAGHSTTRALYSSFSGALYQ
Y13759      ---MTMSRSSRSSVLALALATGSLVAAGPCDIYSSGGTPCIAAHSTTRALYSSYNGPLYQ
AY495375    ------MRSRTNIALGLA-ATGSLVAAAPCDIYQNGGTPCVAAHGTTRALYDSYTGPLYQ
AJ310126    ------MLLQSVFALGLA-TTARLVSAGPCDIYKSGGTPCIAAHSTTRALFDSYNGALYQ
AF306764    -----MSSKSLLVALGLV-ATGSLVHAGPCDIYAKGNTPCIAAHATTRALYNSYSGPLYQ
                             * ***** * *** * * * ***     * ***


U38661      VKRGSDGATTAISPLS-SGVANAAAQDAFCAGTTCLITIIYDQSGRGNHLREAPPGGFSG
Z69252      VKRGSDGATTAISPLS-SGVANAAAQDAFCAGTTCLITIIYDQSGRGNHLTQAPPGGFSG
AF367026    IKRGSDGATTTISPLSAGGVANAAAQDTFCANTTCLITIIYDQSGRGNHLTQAPPGGFSG
AB073861    VKRGSDGSTTDIAPLSAGGVADAAIQDSFCANTTCLITIIYDQSGRGNHLTQAPPGGFNG
AB073860    VKRGSDGSTTDIAPLSAGGVADAATQDSFCANTTCLITIIYDQSGRGNHLTQAPPGGFNG
AB046701    LQRGSDDTTTTISPLTAGGIADASAQDTFCANTTCLITIIYDQSGNGNHLTQAPPGGFDG
AB085904    LQRGSDDTTTTISPLTAGGIADASAQDTFCANTTCLITIIYDQSGNGNHLTQAPPGGFDG
L23502      LQRGSDDTTTTISPLTAGGVADASAQDTFCANTTCLITIIYDQSGNGNHLTQAPPGGFDG
U39942      IKRGSDDTTTTISPLTAGGVADASAQDTFCANTTCLITIIYDQTGNGNHLTQAPPGGFDG
Y13759      VQRASDGTTTTITPLSAGGVADASAQDAFCENTTCLITIIYDQSGNGNDLTQAPPGGFNG
AY495375    LKRGSDGTTTDISPLSAGGVANAAAQDSFCKGTTCLISIIYDQSGRANHLYQAQKGAFSG
AJ310126    IKRASDGALTDIKPLSAGGVAASSAQDTFCSGTTCLITIIYDQSGPSNHLTQAPPGGFKG
AF306764    VKRGSDGATTDIAPLSAGGVANAAAQDKFCANTTCLISIIYDQSGKGNHLTQAPPGAFKG
             * ***     * * **    * *     ** **   ***** ***** *   * *  *   * * *


U38661      PESNGYDNLASAIGAPVTLNGQKAYGVFVSPGTGYRNNAASGTAKGDAAEGMYAVLDGTH
Z69252      PESNGYDNLASAIGAPVTLNGQKAYGVFVSPGTGYRNNAASGTAKGDAAEGMYAVLDGTH
AF367026    PDTNGYDNLASAIGAPVTLNGQKAYGVFISPGTGYRNNAVSGTATGDAAEGMYAVLDGTH
AB073861    PESNGYDNLASAVGAPVTLNGKKAYGVFMSPGTGYRNNAASGTATGDKAEGMYAVLDGTH
AB073860    PESNGYDNLASAVGAPVTLNGKKAYGVFMSPGTGYRNNAASGTATGDEAEGMYAVLDGTH
AB046701    PDTDGYDNLASAIGAPVTLNGQKAYGVFMSPGTGYRNNEATGTATGDEAEGMYAVLDGTH
AB085904    PDTDGYDNLASAIGAPVTLNGQKAYGVFMSPGTGYRNNEATGTATGDEAEGMYAVLDGTH
L23502      PDVDGYDNLASAIGAPVTLNGQKAYGVFMSPGTGYRNNEATGTATGDEPEGMYAVLDGTH
U39942      PDTDGYDNLAPTIGAPVTLNGQKAYGVFMSPGTGYRNNAATGTATGHEAERMYAVLDGTH
Y13759      PDVGGYDNLAGAIGAPVTLNGKKAYGVFVSPGTGYRNNEAIGTATGDEPEGMYAVLDGTH
AY495375    PDVNGNDNLAGAIGAPVTLNGKKAYGVFISPGTGYRNDEVSGTATGNEPEGMYAVLDGTH
AJ310126    PEANGYDNLAAADGAPVTLNGKKAYGVFVSPGTGYRNNHVSGSATGDEPEGLYAVLDGTH
AF306764    PDVGGYDNLASAIGAPVSLGGKKAYGVFISPGTGYRNNNVKGSAVKDEPQGIYAVLDGTH
             *    * ****    **** * * ****** ********    * *       ********
```

76

```
U38661     YNGACCFDYGNAETNSRDTGNGHMEAIYFGDSTVWGTGSGKGPWIMADLENGLFSGSSPG
Z69252     YNGACCFDYGNAETNSRDTGNGHMEAIYFGDSTVWGTGSGKGPWIMADLENGLFSGSSPG
AF367026   YNDACCFDYGNAETSSTDTGNGHMEAIYFGDNTVWGTGAGSGPWVMADLENGLFSGLSSG
AB073861   YNSACCFDYGNAEVSNTDTGNGHMEAIYYGDNTVWGSGAGSGPWIMADLENGLFSGLSST
AB073860   YNSACCFDYGNAEVSNTDTGNGHMEAIYYGDNTVWGSGAGSGPWIMADLENGLFSGLSST
AB046701   YNDACCFDYGNAETSSTDTGAGHMEAIYLGNSTTWGYGAGDGPWIMVDMENNLFSGADEG
AB085904   YNDACCFDYGNAETSSTDTGAGHMEAIYLGNSTTWGYGAGDGPWIMVDMENNLFSGADEG
L23502     YNDACCFDYGNAETSSTDTGAGHMEAIYLGNSTTWGYGAGDGPWIMVDMENNLFSGADEG
U39942     YNDACCFDYGNWQTSSTDTGAGHMEAIYFGTSTTWGYGAGDGPWLMVDMENNLFSGADEG
Y13759     YNDGCCFDYGNAETSSLDTGNGHMEAIYYGTNTAWGYGAGNGPWIMADLENGLFSGQSSD
AY495375   YNDACCFDYGNAEISNTDTGNGHMEAVYYGNNTIWGSGSGSGPWLMADLENGLFSGQGTK
AJ310126   YNGACCFDYGNAETSSTDTGNGHMEAIYFGDNTVWGSGSGSGPWIMADLENGLFSGKNPK
AF306764   YNGGCCFDYGNAETNNLDTGNGHMEAIYFGDNTVWGSGAGNGPWVMADLENGLFSGANPK
           **  *******    *** ***** * *   * ** * * *** * * ** ****


U38661     NNAGDPSISYRFVTAAIKGQPNQWAIRGGNAASGSLSTFYSGARPQV-SGYNPMSKEGAI
Z69252     NNAGDPSISYRFVTAAIKGQPNQWAIRGGNAASGSLSTFYSGARPQV-SGYNPMSKEGAI
AF367026   NNAADPSVSYRFLSTIVKGESNQWAIRGGNAVSGSLSTYYSGARPSA-SGYNPMSKEGAI
AB073861   NNAGDPSISYRFVTAVVKGEANQWSIRGANAASGSLSTYYSGARPSA-SGYNPMSKEGAI
AB073860   NNAGDPSISYRFVTAVVKGEANQWSIRGANAASGSLSTYYSGARPSA-SGYNPMSKEGAI
AB046701   YNSGDPSISYRFVTAAVKGGADKWAIRGANAASGSLSTYYSGARPDY-SGYNPMSKEGAI
AB085904   YNSGDPSISYRFVTAAVKGGADKWAIRGANAASGSLSTYYSGARPDY-SGYNPMSKEGAI
L23502     YNSGDPSISYSFVTAAVKGGADKWAIRGGNAASGSLSTYYSGARPDY-SGYNPMSKEGAI
U39942     YNSGDPSIPYTFLTAAVKGGADKWAIRGGNSASGSLSTYYSGVRPDY-SGYNPMSKEGAI
Y13759     YNAGDPSISYRFVTAILKGGPNLWALRGGNAASGSLSTYYNGIRPTDASGYNPMSKEGAI
AY495375   QNTADPSISNRFFTGMVKGEPNQWALRGSNAASGSLSTYYSGARPTV-GGYNPMSLEGAI
AJ310126   QNTADPTINHRFTSAVVKGKPGTWSIRGGDSTSGGLSTFYNGAYPDG--GYNPMKKEGAI
AF306764   QNTQNPSVSNRFLTTVVKGKPGVWAIRAGDATTGGLSTYYNGSRPSV-SGYNPMSKEGAI
            *    *     **   *  *   * ***  * *     ***** ****


U38661     ILGIGGDNSNGGQGTFYEGVMTSGYPSDATENSVQANIVAARYAVAPL--TSGPALTVGS
Z69252     ILGIGGDNSNGAQGTFYEGVMTSGYPSDATENSVQANIVAARYAVAPL--TSGPALTVGS
AF367026   ILGIGGDNSHGAQGTFYEGVMTSGYPSDATENSVQANIVAAKYAAGSL--TSGPALTVGS
AB073861   ILGIGGDNSNGAQGTFYEGVMTSGYPSDATENSVQADIVAAKYAIASL--TSGPALTVGS
AB073860   ILGIGGDNSNGAQGTFYEGVMTSGYPSDATENSVQADIVAAKYAIASL--TSGPALTVGS
AB046701   ILGIGGYNSNGAQGTFYEGVMTSGYPSDDTENSVQENIVAAKYVVGSL--VSGPSFTSGE
AB085904   ILGIGGDNSNGAQGTFYEGVMTSGYPSDDTENSVQENIVAAKYVVGSL--VSGPSFTSGE
L23502     ILGIGGDNSNGAQGTFYEGVMTSGYPSDDVENSVQENIVAAKYVSGSL--VSGPSFTSGE
U39942     ILGIGGDNSNGAQGTFYEGVMTSGYPSDATENSVQENIVAAKYAVGSL--VSGPSFTSGE
Y13759     ILGIGGDNSVSAQGTFYEGAMTDGYPDDATENSVQADIVAAKYATTSL--ISGPALTVGD
AY495375   ILGIGGDNSNGAQGTFYEGVMTSGYPSDATEASVQANIVAAKYATTSL--NTAPL-TVGN
AJ310126   ILGIGGDNSNGAQGTFYEGVMTSGYPSDATENAVQANIVAAKYATTSL--TSGSALTVGS
AF306764   ILGIGGDNSNGAQGTFTEGAMTFGYPSDAIENEVQANLVAAGYSTGRGLMTSGPAYTVGS
           ****** **    **** ** ** *** * *  **   *** *          * *


U38661     SISLRATTACCTTRYIAHSGSTVNTQVVSSSSATALKQQASWTVRAGLAN----NACFSF
Z69252     SISLRATTACCTTRYIAHSGSTVNTQVVSSSSATALKQQASWTVRAGLAN----NACFSF
AF367026   SISLRATTSGYTTRYIAHTGSTVNTQVVTSSSSTTLKQQASWTVRTGLGN----SACYSF
AB073861   SISLQVTTAGYTTRYLAHDGSTVNTQVVSSSSSTTALRQQASWTVRTGLAN----SACLSF
AB073860   SISLQVTTAGYTTRYLAHDGSTVNTQVVSSTSTKALKQQASWTVRTGLAN----SACFSF
AB046701   VVSLRVTTPGYTTRYIAHTDTTVNTQVVDDDSSTTLKEEASWTVVTGLAN----SQCFSF
AB085904   VVSLRVTTPGYTTRYIAHTDTTVNTQVVDDDSSTTLKEEASWTVVTGLAN----SQCFSF
L23502     VVSLRVTTPGYTTRYIAHTDTTVNTQVVDDDSSTTLKEEASWTVVTGLAN----SQCFSF
U39942     VVSLRVTTPGYTTRYIAHTDTTVNTQVVDDDSDSSLKEEASWTVVHGLAN----SQCFSF
```

77

```
Y13759      TVSLKVTTSGYDTRYIAHTGSTINTQVVSSSSSSTLKQQASWTVRTGLASTAAANGCVSF
AY495375    KISIKVTTPGYDTRYLAHTGATVNTQVVSSSSATSLKQQASWTVRTGLGN----SGCYSF
AJ310126    SISLKVTTSGYTDRYLTHSDSTVNTQVVSSSSATALKQSASWTVRTGLAN----SGCVSF
AF306764    SVSLRATTSGYTDRYLAHSGATVNTQVVSSSSTALLKRQASWIVRAGFTN----SECFAF
             *   **    **   *  * *****   *    *   *** * *         *  *


U38661      ESQDTSGSYIRHSNFGLVLNANDGSKLFAEDATFCTQAGINGQGSSI-RSWSYPTRYFRH
Z69252      ESRDTSGSYIRHSNFGLVLNANDGSKLFAEDATFCTQAGINGQGSSI-RSWSYPTRYFRH
AF367026    ESVDTPGSYIRHYNFGLLLNANDGTKQFHEDATFCPQSGLSGQGNSI-RSWSYPTRYFRH
AB073861    ESVDTPGSYIRHYNFALLLNANDGTKQFYEDATFCPQAGLNGQGNSI-RSWSYPTRYFRH
AB073860    ESVDTPGSYIRHYNFAHLLNANDGTKQFYEDATFCPQAGLNGQGNSI-RSWSYPTRYFRH
AB046701    ESVDTPGSYIRHYNFELLLNANDGTKQFHEDATFCPQAALNGEGTSL-RSWSYPTRYFRH
AB085904    ESVDTPGSYIRHYNFELLLNANDGTKQFHEDATFCPQAALNGEGTSL-RSWSYPTRYFRH
L23502      ESVDTPGSYIRHYNFELLLNANDGTKQFHEDATFCPQAPLNGEGTSL-RSWSYPTRYFRH
U39942      ESVDTPGSYIRHYNFALLLNANDGTKQFHEDATFCPQAALNGEGTSL-RSWSYPTRYFRH
Y13759      ESRDTPGSYIRHSNFALLLNANDGTKLFSEDATFCPQDSFNDDGTNSIRSWNYPTRYWRH
AY495375    ESVDTPGSFIRHYNFQLQLNANDNTKAFKEDATFCSQTGLVTGNTF--NSWSYPAKFIRH
AJ310126    ESNDTPGSYIRHSGFTLYVNKGDGSKSFNEDATFCPQKGLSGSGSSI-RSWNYPTRYIRH
AF306764    ESKDTAGSFLRHANFVLQVNANDGSKGFKEDATFCPQAGLTGKGSSI-RTWAYPTRWIRH
            ** ** **   **  *   *   * * * ******  *          * **     **


U38661      YNNTLYIASNGGVHVFDATAAFNDDVSFVVSGGFA-
Z69252      YNNTLYIASNGGVHVFDATAAFNDDVSFVVSGGFA-
AF367026    YNNILYIASNGGVDTFDATASFNADVTWVISTGFA-
AB073861    YENVLYVASNGGVQTFDATTSFNDDVSWVVSTGFA-
AB073860    YENVLYVASNGGVQTFDATTSFNDDVSWVVSTGFA-
AB046701    YENVLYAASNGGVQTFDSKTSFNNDVSFEIETAFAS
AB085904    YENVLYAASNGGVQTFDSKTSFNNDVSFEIETAFAS
L23502      YENVLYAASNGGVQTFDSKTSFNNDVSFEIETAFAS
U39942      YENVLYAASNGGVQTFDSTTSFNNDVSFEIETAFAS
Y13759      YENVLYVASNGGVNTFDAATAFTDDVSWVVADGFA-
AY495375    YNNVGYIASNGGVHDFDSATGFNNDVSFVVGSSFA-
AJ310126    YNNLGYASSNGGVHDWDAAKSFVDDVTFVVANGFA-
AF306764    FNNVGYISSNGGVKDFDNVSSFNDDITWLVESALA-
            * *  *****   *    *  *          *
```

Figure 1. Multialignment of family 54 bifunctional α-L-arabinofuranosidases and β-D-xylosidase. The sequence alignment was achieved using Biology WorkBench 3.2 CLUSTALW provided by San Diego Supercomputer Center (San Diego, CA). All sequences of enzymes are derived from the corresponding genes, which are indicated by GenBank accession number and the strain of microbes as follows: U38661 from *Hypocrea koningii* G-39、Z69252 from *Hypocrea jecorina RutC-30*、AF367026 from *Penicillium purpurogenum*、AB073861 from *Aspergillus oryzae* RIB40、 AB073860 from *Aspergillus oryzae* HL15、AB046701 from *Aspergillus niger var. awamorii* IFO4033、AB085904 from *Aspergillus kawachii IFO 4308*、L23502 from *Aspergillus niger*、U39942 from *Aspergillus niger*、Y13759 from *Emericella nidulans argB2*、AY495375 from *Aureobasidium pullulans*、AJ310126 from *Fusarium oxysporum f. sp. Dianthi*、AF306764 from *Cochliobolus carbonum*.

```
          10        20        30        40        50        60
gggcCTTGTG ACATCTACTC CTCGGGCGGA ACGCCTTGCG TTGCCGCCCA CAGCACCACT
          70        80        90       100       110       120
CGAGCTCTGT TCAGCGCTTA TACCGGCCCG TTATACCAGG TAAAGCGCGG CTCCGATGGT
         130       140       150       160       170       180
GCCACAACCG CCATATCGCC CCTCTCAAGT GGTGTGGCCA ACGCTGCCGC TCAAGATGCT
         190       200       210       220       230       240
TTCTGTGCGG GAACTACATG CCTCATTACC ATCATATACG ACCAGTCGGG TCGCGGCAAC
         250       260       270       280       290       300
CATCTCAccc AGGCCCCGCC GGGCGGCTTC AGCGGCCCGG AATCCAACGG CTATGACAAC
         310       320       330       340       350       360
CTGGCTAGTG CAATTGGGGC GCCGGTAACA CTCAACGGCC AGAAGGCGTA TGGAGTTTTC
         370       380       390       400       410       420
GTGTCTCCAG GAACGGGGTA TCGGAATAAC GCTGCCAGCG GCACAGCCAA AGGAGATGCC
         430       440       450       460       470       480
GCGGAGGGCA TGTATGCGGT TCTCGATGGT ACACACTACA ACGGCGCCTG CTGCTTTGAC
         490       500       510       520       530       540
TATGGCAACG CCGAGACCAA CAGCCGCGAT ACAGGCAACG GTCAcATGGA GGCCATCTAT
         550       560       570       580       590       600
TTTGGCGACA GCACTGTCTG GGGTACTGGC TCAGGCAAGG GTCCGTGGAT CATGGCTGAT
         610       620       630       640       650       660
CTCGAGAACG GCTTGTTCTC AGGCTCCAGT CCCGGCAACA ATGCCGGTGA TCCGTCCATC
         670       680       690       700       710       720
TCGTACCGGT TCGTCACTGC gGCGATCAAG GGGCAGCCAA ACCAATGGGC AATCCGTGGC
         730       740       750       760       770       780
GGCAATGCTG CGTCTGGCTC GCTGTCAACT TTCTACAGCG GCGCTCGCCC ACAAGTCTCC
         790       800       810       820       830       840
GGATACAATC CGATGAGCAA AGAGGGCGCC ATCATTCTCG GCATTGGCGG CGACAACAGC
         850       860       870       880       890       900
AACGGCGcCC AGGGCACATT CTATGAGGGC GTCATGACCT CTGGATATCC TTCCGATGCA
         910       920       930       940       950       960
ACAGAGAATT CAGTGCAAGC CAACATCGTA GCTGCCAGAT ATGCCGTCGC GCCGCTGACA
         970       980       990      1000      1010      1020
AGTGGTCCTG CCCTCACCGT CGGTTCGTCA ATCTCTCTAC GAGCCACAAC AGCTTGCTGC
        1030      1040      1050      1060      1070      1080
ACGACTCGCT ACATCGCTCA TAGCGGGTCT ACGGTCAACA CCCAAGTCGT CTCATCATCC
        1090      1100      1110      1120      1130      1140
AGCGCCACAG CCCTCAAGCA ACAGGCCAGC TGGACGGTTC GTGCCGGCCT GGCAAATAAC
        1150      1160      1170      1180      1190      1200
GCTTGCTTCT CGTTTGAGTC cAgGGACACG TCGGGCAGCT ACATTCGCCA CTCCAACTTC
```

79

```
      1210        1220        1230        1240        1250        1260
GGTCTTGTGT TGAATGCCAA TGATGGCAGT AAGCTCTTTG CTGAGGATGC TACCTTTTGC
      1270        1280        1290        1300        1310        1320
ACGCAGGCGG GCATTAACGG CCAGGGAAGC TCCATTAGAT CTTGGAGCTA CCCAACGCGA
      1330        1340        1350        1360        1370        1380
TACTTTCGCC ACTACAATAA CACGCTCTAC ATCGCGAGTA ACGGGGGCGT TCACGTATTC
      1390        1400        1410        1420        1430        1440
GATGCCACTG CCGCTTTCAA CGACGACGTG AGCTTTGTGG TTAGTGGCGG TTTTGCTTAA
```

Amino acid Sequencing

```
           9            18           27           36           45           54
5' GGG CCT TGT GAC ATC TAC TCC TCG GGC GGA ACG CCT TGC GTT GCC GCC CAC AGC
   Gly Pro Cys Asp Ile Tyr Ser Ser Gly Gly Thr Pro Cys Val Ala Ala His Ser

           63           72           81           90           99          108
   ACC ACT CGA GCT CTG TTC AGC GCT TAT ACC GGC CCG TTA TAC CAG GTA AAG CGC
   Thr Thr Arg Ala Leu Phe Ser Ala Tyr Thr Gly Pro Leu Tyr Gln Val Lys Arg

          117          126          135          144          153          162
   GGC TCC GAT GGT GCC ACA ACC GCC ATA TCG CCC CTC TCA AGT GGT GTG GCC AAC
   Gly Ser Asp Gly Ala Thr Thr Ala Ile Ser Pro Leu Ser Ser Gly Val Ala Asn

          171          180          189          198          207          216
   GCT GCC GCT CAA GAT GCT TTC TGT GCG GGA ACT ACA TGC CTC ATT ACC ATC ATA
   Ala Ala Ala Gln Asp Ala Phe Cys Ala Gly Thr Thr Cys Leu Ile Thr Ile Ile

          225          234          243          252          261          270
   TAC GAC CAG TCG GGT CGC GGC AAC CAT CTC ACC CAG GCC CCG CCG GGC GGC TTC
   Tyr Asp Gln Ser Gly Arg Gly Asn His Leu Thr Gln Ala Pro Pro Gly Gly Phe
       D95N

          279          288          297          306          315          324
   AGC GGC CCG GAA TCC AAC GGC TAT GAC AAC CTG GCT AGT GCA ATT GGG GCG CCG
   Ser Gly Pro Glu Ser Asn Gly Tyr Asp Asn Leu Ala Ser Ala Ile Gly Ala Pro
                               D120N

          333          342          351          360          369          378
   GTA ACA CTC AAC GGC CAG AAG GCG TAT GGA GTT TTC GTG TCT CCA GGA ACG GGG
   Val Thr Leu Asn Gly Gln Lys Ala Tyr Gly Val Phe Val Ser Pro Gly Thr Gly

          387          396          405          414          423          432
   TAT CGG AAT AAC GCT GCC AGC GGC ACA GCC AAA GGA GAT GCC GCG GAG GGC ATG
   Tyr Arg Asn Asn Ala Ala Ser Gly Thr Ala Lys Gly Asp Ala Ala Glu Gly Met
                                               D160N            E163Q

          441          450          459          468          477          486
   TAT GCG GTT CTC GAT GGT ACA CAC TAC AAC GGC GCC TGC TGC TTT GAC TAT GGC
   Tyr Ala Val Leu Asp Gly Thr His Tyr Asn Gly Ala Cys Cys Phe Asp Tyr Gly
                   D170N                                       D181N

          495          504          513          522          531          540
   AAC GCC GAG ACC AAC AGC CGC GAT ACA GGC AAC GGT CAC ATG GAG GCC ATC TAT
   Asn Ala Glu Thr Asn Ser Arg Asp Thr Gly Asn Gly His Met Glu Ala Ile Tyr
           E186Q                 D191N                         E198Q

          549          558          567          576          585          594
   TTT GGC GAC AGC ACT GTC TGG GGT ACT GGC TCA GGC AAG GGT CCG TGG ATC ATG
```

Phe Gly Asp Ser Thr Val Trp Gly Thr Gly Ser Gly Lys Gly Pro Trp Ile Met

603 612 621 630 639 648
GCT GAT CTC GAG AAC GGC TTG TTC TCA GGC TCC AGT CCC GGC AAC AAT GCC GGT
Ala Asp Leu Glu Asn Gly Leu Phe Ser Gly Ser Ser Pro Gly Asn Asn Ala Gly
D221N   E223Q(Catalytic nucleophile)

657 666 675 684 693 702
GAT CCG TCC ATC TCG TAC CGG TTC GTC ACT GCG GCG ATC AAG GGG CAG CCA AAC
Asp Pro Ser Ile Ser Tyr Arg Phe Val Thr Ala Ala Ile Lys Gly Gln Pro Asn
D238N

711 720 729 738 747 756
CAA TGG GCA ATC CGT GGC GGC AAT GCT GCG TCT GGC TCG CTG TCA ACT TTC TAC
Gln Trp Ala Ile Arg Gly Gly Asn Ala Ala Ser Gly Ser Leu Ser Thr Phe Tyr

765 774 783 792 801 810
AGC GGC GCT CGC CCA CAA GTC TCC GGA TAC AAT CCG ATG AGC AAA GAG GGC GCC
Ser Gly Ala Arg Pro Gln Val Ser Gly Tyr Asn Pro Met Ser Lys Glu Gly Ala
E289Q

819 828 837 846 855 864
ATC ATT CTC GGC ATT GGC GGC GAC AAC AGC AAC GGC GCC CAG GGC ACA TTC TAT
Ile Ile Leu Gly Ile Gly Gly Asp Asn Ser Asn Gly Ala Gln Gly Thr Phe Tyr
D299N (Catalytic acis/base)

873 882 891 900 909 918
GAG GGC GTC ATG ACC TCT GGA TAT CCT TCC GAT GCA ACA GAG AAT TCA GTG CAA
Glu Gly Val Met Thr Ser Gly Tyr Pro Ser Asp Ala Thr Glu Asn Ser Val Gln
E310Q   D320N   E323Q

927 936 945 954 963 972
GCC AAC ATC GTA GCT GCC AGA TAT GCC GTC GCG CCG CTG ACA AGT GGT CCT GCC
Ala Asn Ile Val Ala Ala Arg Tyr Ala Val Ala Pro Leu Thr Ser Gly Pro Ala

981 990 999 1008 1017 1026
CTC ACC GTC GGT TCG TCA ATC TCT CTA CGA GCC ACA ACA GCT TGC TGC ACG ACT
Leu Thr Val Gly Ser Ser Ile Ser Leu Arg Ala Thr Thr Ala Cys Cys Thr Thr

1035 1044 1053 1062 1071 1080
CGC TAC ATC GCT CAT AGC GGG TCT ACG GTC AAC ACC CAA GTC GTC TCA TCA TCC
Arg Tyr Ile Ala His Ser Gly Ser Thr Val Asn Thr Gln Val Val Ser Ser Ser

1089 1098 1107 1116 1125 1134
AGC GCC ACA GCC CTC AAG CAA CAG GCC AGC TGG ACG GTT CGT GCC GGC CTG GCA
Ser Ala Thr Ala Leu Lys Gln Gln Ala Ser Trp Thr Val Arg Ala Gly Leu Ala

1143 1152 1161 1170 1179 1188
AAT AAC GCT TGC TTC TCG TTT GAG TCC AGG GAC ACG TCG GGC AGC TAC ATT CGC
Asn Asn Ala Cys Phe Ser Phe Glu Ser Arg Asp Thr Ser Gly Ser Tyr Ile Arg

1197		1206		1215		1224		1233		1242
CAC	TCC	AAC	TTC	GGT	CTT	GTG	TTG	AAT	GCC	AAT	GAT	GGC	AGT	AAG	CTC	TTT	GCT
His	Ser	Asn	Phe	Gly	Leu	Val	Leu	Asn	Ala	Asn	Asp	Gly	Ser	Lys	Leu	Phe	Ala

D429N

1251		1260		1269		1278		1287		1296
GAG	GAT	GCT	ACC	TTT	TGC	ACG	CAG	GCG	GGC	ATT	AAC	GGC	CAG	GGA	AGC	TCC	ATT
Glu	Asp	Ala	Thr	Phe	Cys	Thr	Gln	Ala	Gly	Ile	Asn	Gly	Gln	Gly	Ser	Ser	Ile

E436Q D437N

1305		1314		1323		1332		1341		1350
AGA	TCT	TGG	AGC	TAC	CCA	ACG	CGA	TAC	TTT	CGC	CAC	TAC	AAT	AAC	ACG	CTC	TAC
Arg	Ser	Trp	Ser	Tyr	Pro	Thr	Arg	Tyr	Phe	Arg	His	Tyr	Asn	Asn	Thr	Leu	Tyr

1359		1368		1377		1386		1395		1404
ATC	GCG	AGT	AAC	GGG	GGC	GTT	CAC	GTA	TTC	GAT	GCC	ACT	GCC	GCT	TTC	AAC	GAC
Ile	Ala	Ser	Asn	Gly	Gly	Val	His	Val	Phe	Asp	Ala	Thr	Ala	Ala	Phe	Asn	Asp

D482N

1413		1422		1431		1440
GAC	GTG	AGC	TTT	GTG	GTT	AGT	GGC	GGT	TTT	GCT	TAA	3'
Asp	Val	Ser	Phe	Val	Val	Ser	Gly	Gly	Phe	Ala	***	

D490N

Map of pPICZα A, B, and C

The figure below summarizes the features of the pPICZα A, B, and C vectors. **The complete sequences for pPICZα A, B, and C are available for downloading from our World Wide Web site (www.invitrogen.com) or from Technical Service (see page 67).** Details of the multiple cloning sites are shown on page 17 for pPICZα A, page 18 for pPICZα B, and page 19 for pPICZα C.
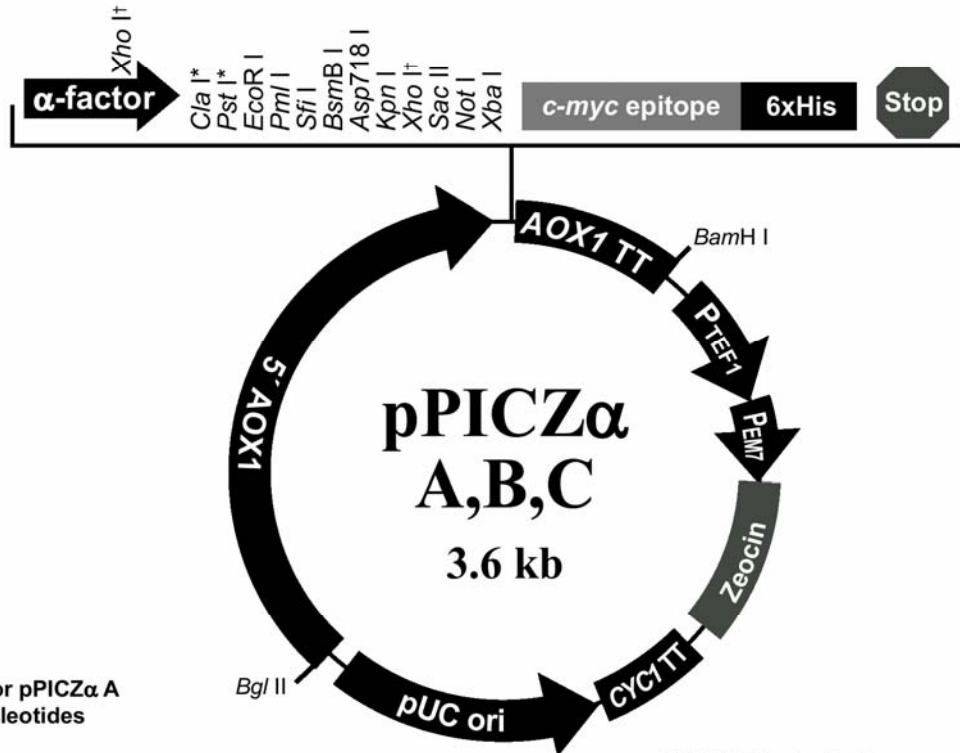
Comments for pPICZα A
3593 nucleotides

5′ *AOX1* promoter region: bases 1-941
5′ *AOX1* priming site: bases 855-875
α-factor signal sequence: bases 941-1207
α-factor priming site: bases 1144-1164
Multiple cloning site: bases 1208-1276
c-myc epitope: bases 1275-1304
Polyhistidine (6xHis) tag: bases 1320-1337
3′ *AOX1* priming site: bases 1423-1443
AOX1 transcription termination region: bases 1341-1682
TEF1 promoter: bases 1683-2093
EM7 promoter: bases 2095-2162
Sh ble ORF: bases 2163-2537
CYC1 transcription termination region: bases 2538-2855
pUC origin: bases 2866-3539 (complementary strand)

* *Pst* I is in Version B only
  Cla I is in Version C only

†The two *Xho* I sites in the vector allow the user to clone their gene in frame with the Kex2 cleavage site, resulting in expression of their native gene without additional amino acids at the N-terminus.

Appendix 3 (continue)

Multiple Cloning Site of pPICZα B

Below is the multiple cloning site of pPICZα B. Restriction sites are labeled to indicate the cleavage site. The boxed nucleotide indicates the variable region. The multiple cloning site has been confirmed by sequencing and functional testing. **The complete sequence of pPICZα B is available for downloading from our World Wide Web site (www.invitrogen.com) or from Technical Service (see page 67).**

```
                    5′ end of AOX1 mRNA                              5′ AOX1 priming site
  811   AACCTTTTTT TTTATCATCA TTATTAGCTT ACTTTCATAA TTGCGACTGG TTCCAATTGA

  871   CAAGCTTTTG ATTTTAACGA CTTTTAACGA CAACTTGAGA AGATCAAAAA ACAACTAATT

  931   ATTCGAAACG ATG AGA TTT CCT TCA ATT TTT ACT GCT GTT TTA TTC GCA GCA
                      Met Arg Phe Pro Ser Ile Phe Thr Ala Val Leu Phe Ala Ala

  983   TCC TCC GCA TTA GCT GCT CCA GTC AAC ACT ACA ACA GAA GAT GAA ACG GCA
         Ser Ser Ala Leu Ala Ala Pro Val Asn Thr Thr Thr Glu Asp Glu Thr Ala
                                            α-factor signal sequence
 1034   CAA ATT CCG GCT GAA GCT GTC ATC GGT TAC TCA GAT TTA GAA GGG GAT TTC
         Gln Ile Pro Ala Glu Ala Val Ile Gly Tyr Ser Asp Leu Glu Gly Asp Phe

 1085   GAT GTT GCT GTT TTG CCA TTT TCC AAC AGC ACA AAT AAC GGG TTA TTG TTT
         Asp Val Ala Val Leu Pro Phe Ser Asn Ser Thr Asn Asn Gly Leu Leu Phe
                              α-factor priming site                        Xho I*
 1136   ATA AAT ACT ACT ATT GCC AGC ATT GCT GCT AAA GAA GAA GGG GTA TCT CTC
         Ile Asn Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly Val Ser Leu
         Kex2 signal cleavage                Pst I  EcoR I   Pml I        Sfi I      BsmB I
 1187   GAG AAA AGA GAG GCT GAA GC TGCAG GAATTCAC GTGGCCCAG CCGGCCGTC TCGGA
         Glu Lys Arg Glu Ala Glu Ala
                      Ste13 signal cleavage
         Asp718 I Kpn I Xho I   Sac II  Not I          Xba I         c-myc epitope
 1243   TCGGTACCTC GAGCCGCGGC GGCCGCCAGC TTTCTA GAA CAA AAA CTC ATC TCA GAA
                                                     Glu Gln Lys Leu Ile Ser Glu
                                                   polyhistidine tag
 1300   GAG GAT CTG AAT AGC GCC GTC GAC CAT CAT CAT CAT CAT CAT TGA GTTTGTA
         Glu Asp Leu Asn Ser Ala Val Asp His His His His His His ***

 1352   GCCTTAGACA TGACTGTTCC TCAGTTCAAG TTGGGCACTT ACGAGAAGAC CGGTCTTGCT
                              3′ AOX1 priming site
 1412   AGATTCTAAT CAAGAGGATG TCAGAATGCC ATTTGCCTGA GAGATGCAGG CTTCATTTTT
                                                      3′ polyadenylation site
 1472   GATACTTTTT TATTTGTAAC CTATATAGTA TAGGATTTTT TTTGTCATTT TGTTTCTTCT
```

*To express your protein with a native N-terminus, you must clone your gene flush with the Kex2 cleavage site. You will need to use PCR and utilize the *Xho* I site upstream of the Kex2 cleavage site.

85

Equipment

Centrifugation	Kubota 7820
Dry bath	Firstek Scientific, DB-101
FPLC	Pharmacia
Incubator shaker	Firstek Scientific, orbital shaking incubator Model-S302R
Incubator	Deng Yng
MASS	Micromass, ESI-Q-Tof
Microcentrifugation	Hermle, Z 160M
PCR	Applied Biosystems, GeneAmp PCR system 2400, 9700
pH meter	Sentron Integrated sensor Technology, model 1001
UV	Hewlett–Packard model 8452A Diode Array spectrophotometer
Vaccum evaporator	Eyela rotary vaccum evaporator
Water bath	Firstek Scientific, Model-B401

Appendix 5

NMR data

1a   2a   3a   4a   5a

1b   2b   3b   4b   5b

6a   7a   8a   9a

6b   7b   8b   9b

2,4-DNXP

Compoumd 1a

OCH3

4.81
4.81
3.95
3.94
3.93
3.93
3.92
3.91
3.90
3.84
3.83
3.82
3.81
3.73
3.71
3.69
3.67
3.61
3.59
3.57
3.55

1.00

1.87
0.92
1.00
0.98
2.89

10 9 8 7 6 5 4 3 2 1 0 ppm
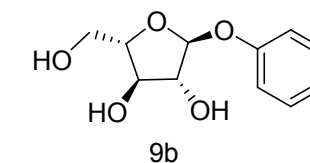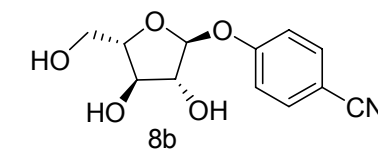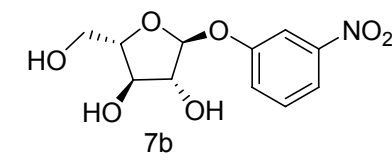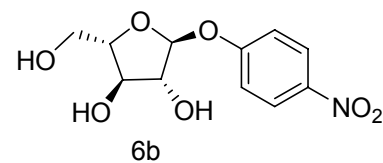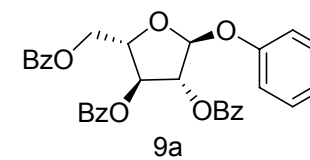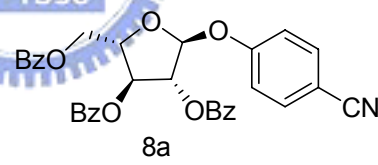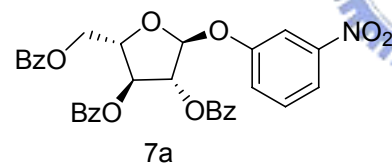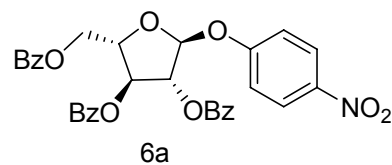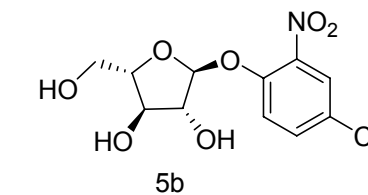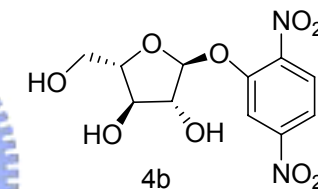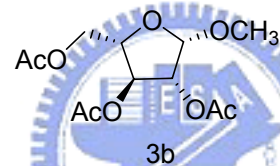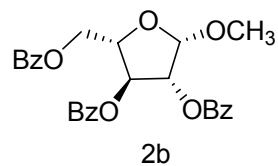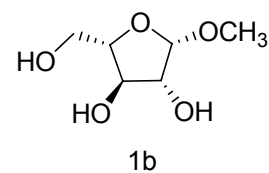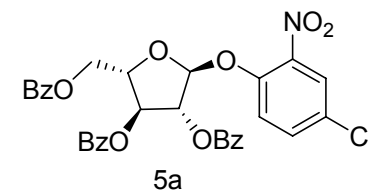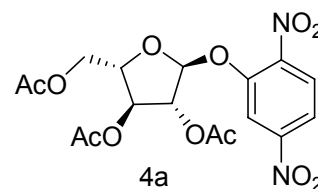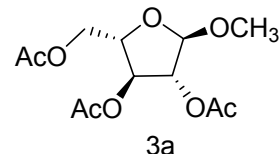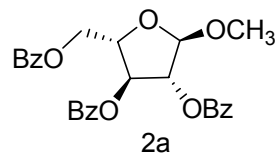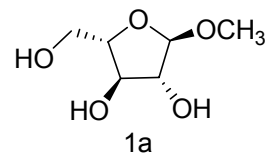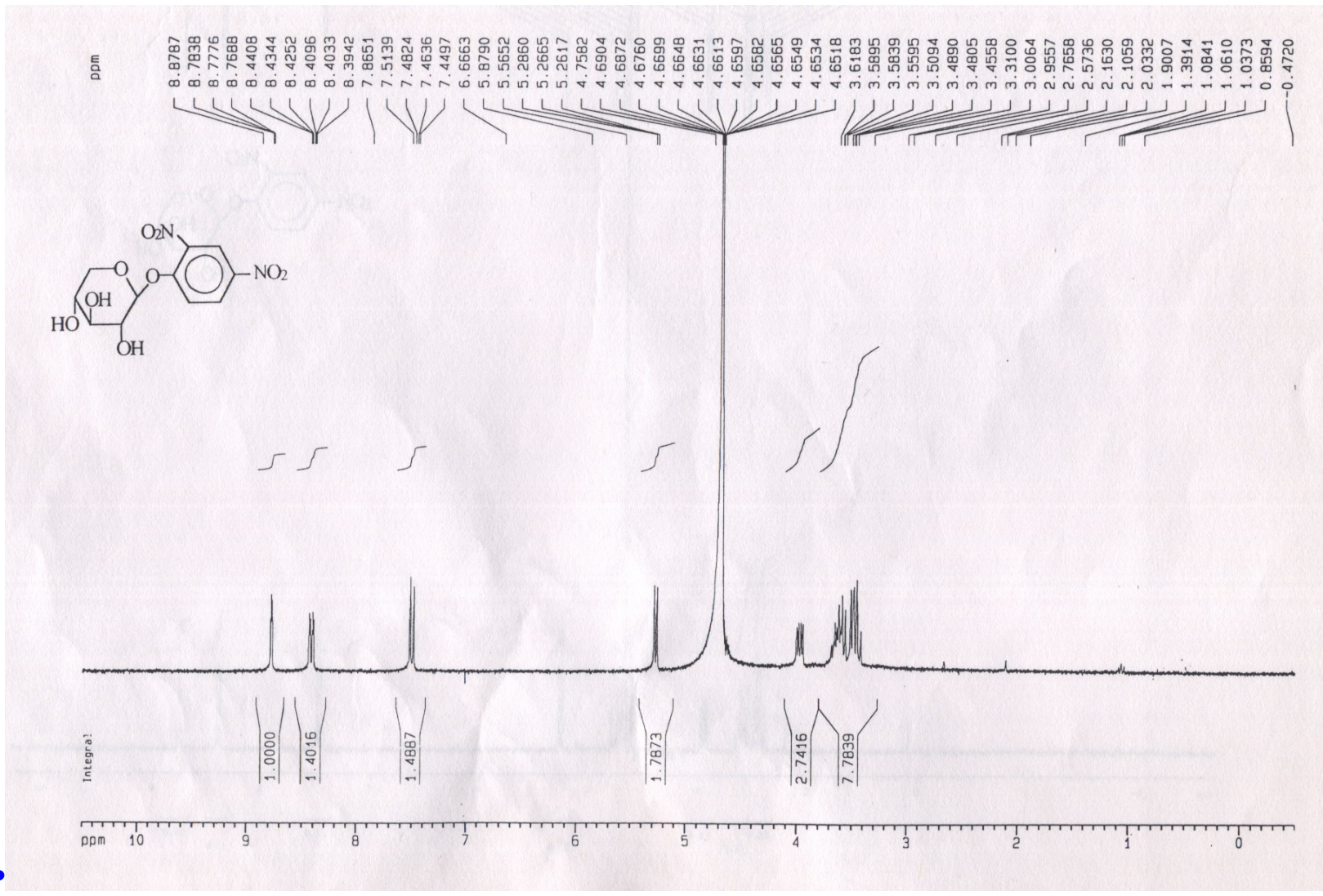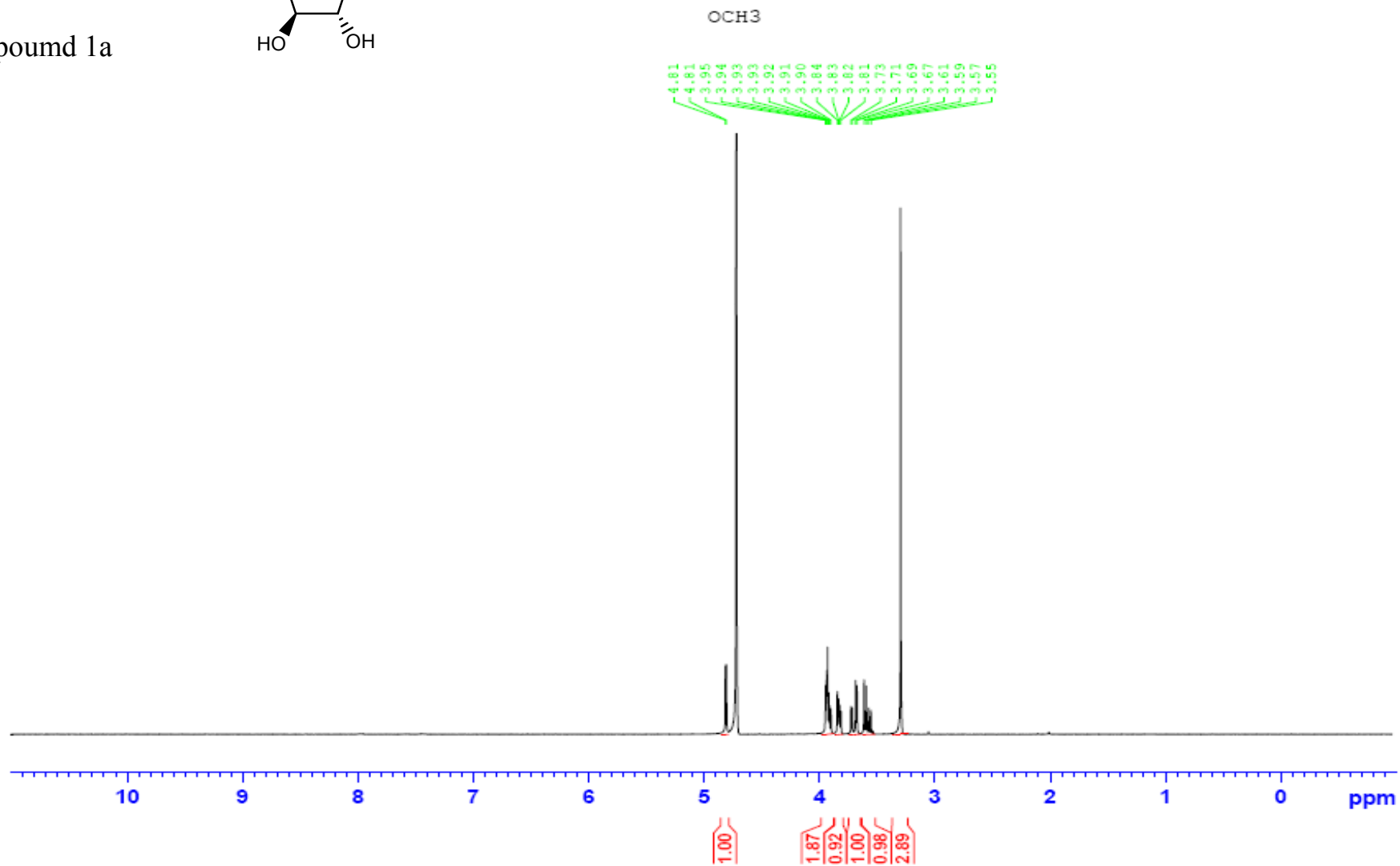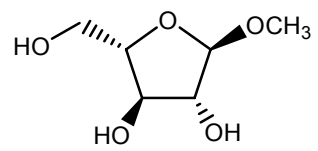
Compound 1b

OBz-started-material

8.08
8.05
8.02
8.00
7.98
7.97
7.60
7.60
7.59
7.58
7.56
7.55
7.54
7.52
7.49
7.47
7.44
7.41
7.40
7.38
7.35
7.30
7.28
7.25
7.24
5.58
5.57
5.50
5.17
4.87
4.85
4.83
4.82
4.71
4.69
4.67
4.65
4.58
4.57
4.55
4.54
3.46
3.36
1.71
1.24
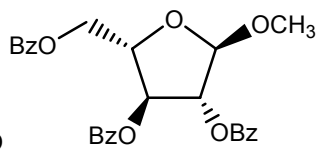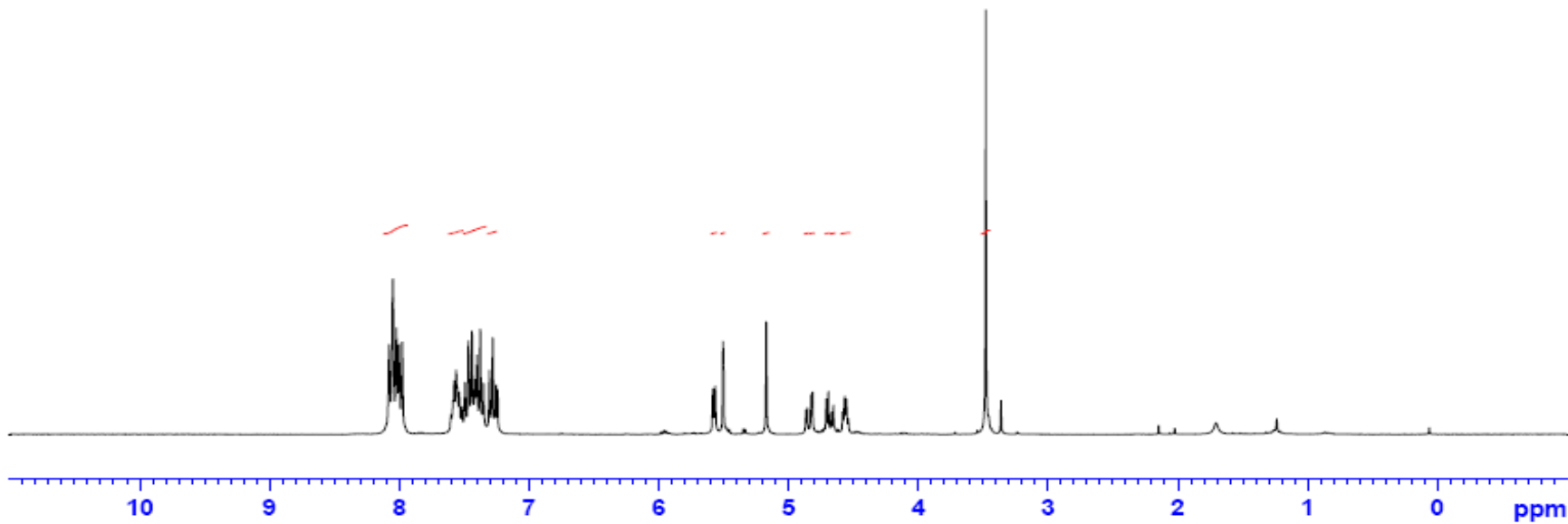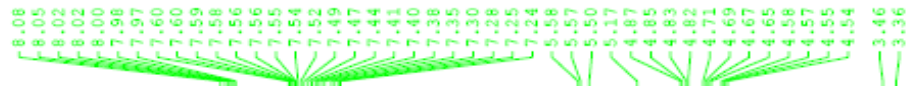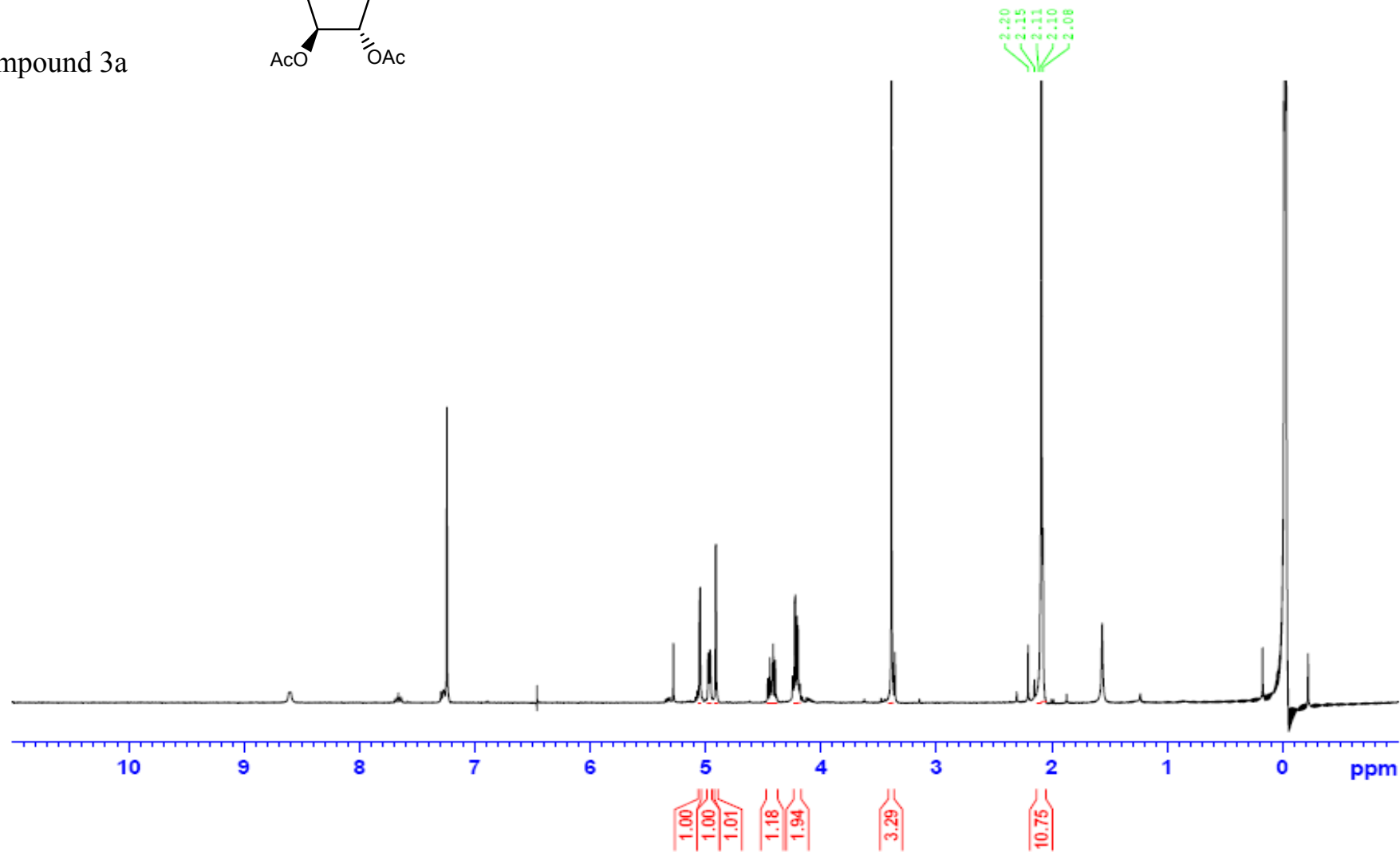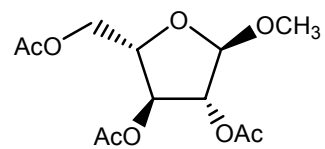
Compound 3a

Compound 4a

Compound 5a

400MHz cl cdcl3

Frequency labels (Hz): 3261.74, 3254.51, 3253.27, 3214.79, 3207.56, 3206.39, 3198.69, 3191.48, 3190.30, 3137.55, 3134.98, 3040.27, 3032.72, 3001.92, 3000.49, 2994.04, 2961.37, 2952.38, 2943.20, 2916.31, 2896.95, 2413.98, 2318.64, 2285.84, 2281.74, 1940.56, 1937.43, 1928.96, 1925.81, 1910.04, 1905.53, 1901.48, 1897.96, 1890.54, 1885.70, 1878.93, 1874.13, 633.46, 495.82, -7.32

Integral values: 2.0525, 4.0260, 0.9339, 2.1665, 4.0893, 3.1212, 2.0577, 1.0000, 1.0117, 1.0019, 1.0408, 2.0481
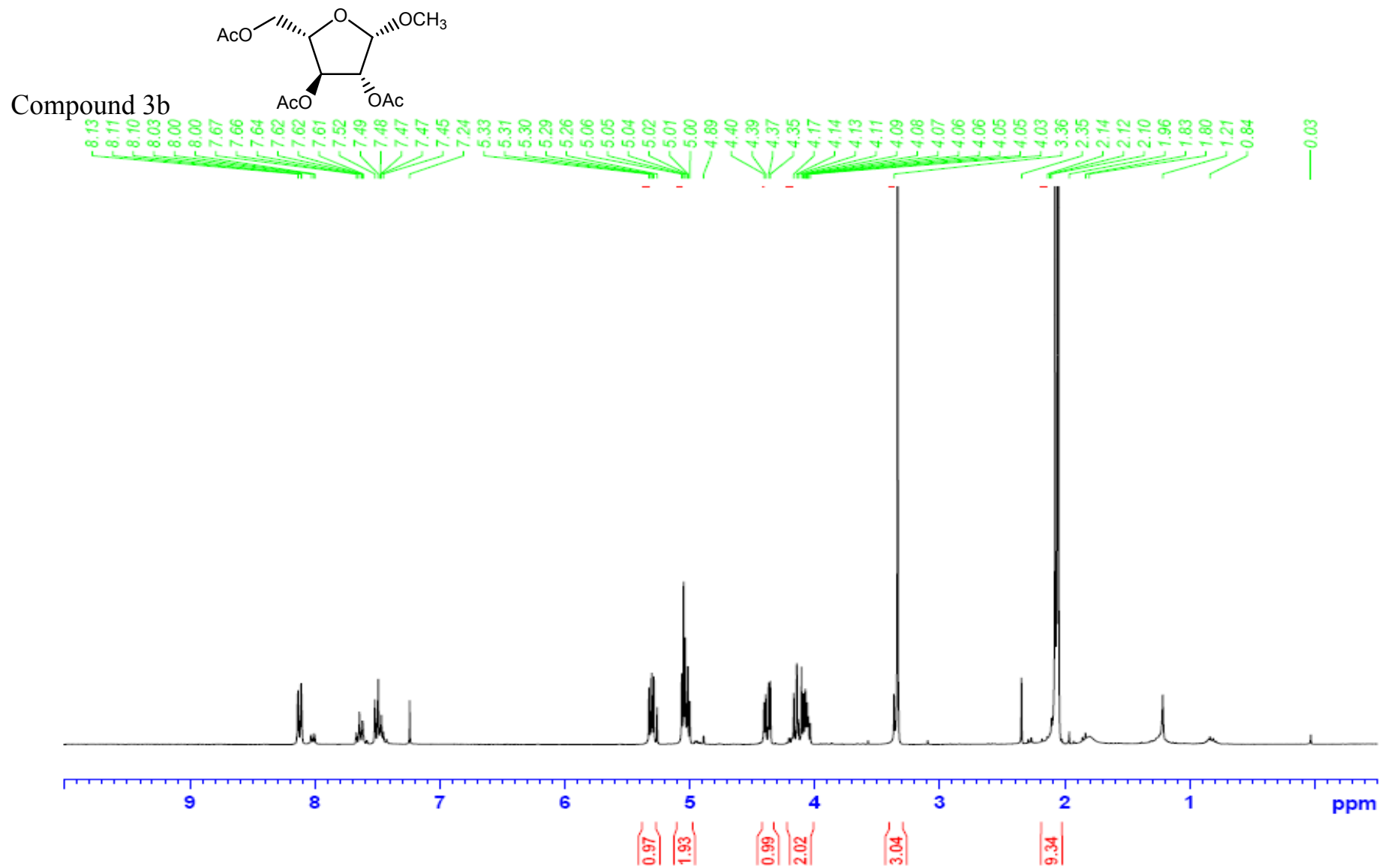
Compound 6a
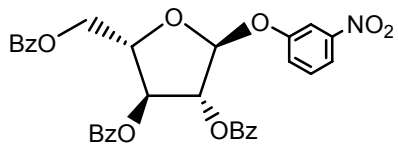
Compound 7a

400MHz m-nitro 030606 cdcl3

Compound 8a

400MHz cn030512 (cdc13)

Compound 9a

8.09 8.04 8.03 8.03 8.01 8.01 8.00 8.00 7.63 7.60 7.58 7.55 7.52 7.49 7.47 7.44 7.42 7.39 7.37 7.33 7.30 7.28 7.25 7.22 7.21 7.17 7.14 7.11 7.06 7.04 7.01

9   8   7   6   5   4   3   2   1   ppm

6.86   2.41 6.33 4.81 3.57 1.08   1.00 1.01 1.03   0.19   3.33
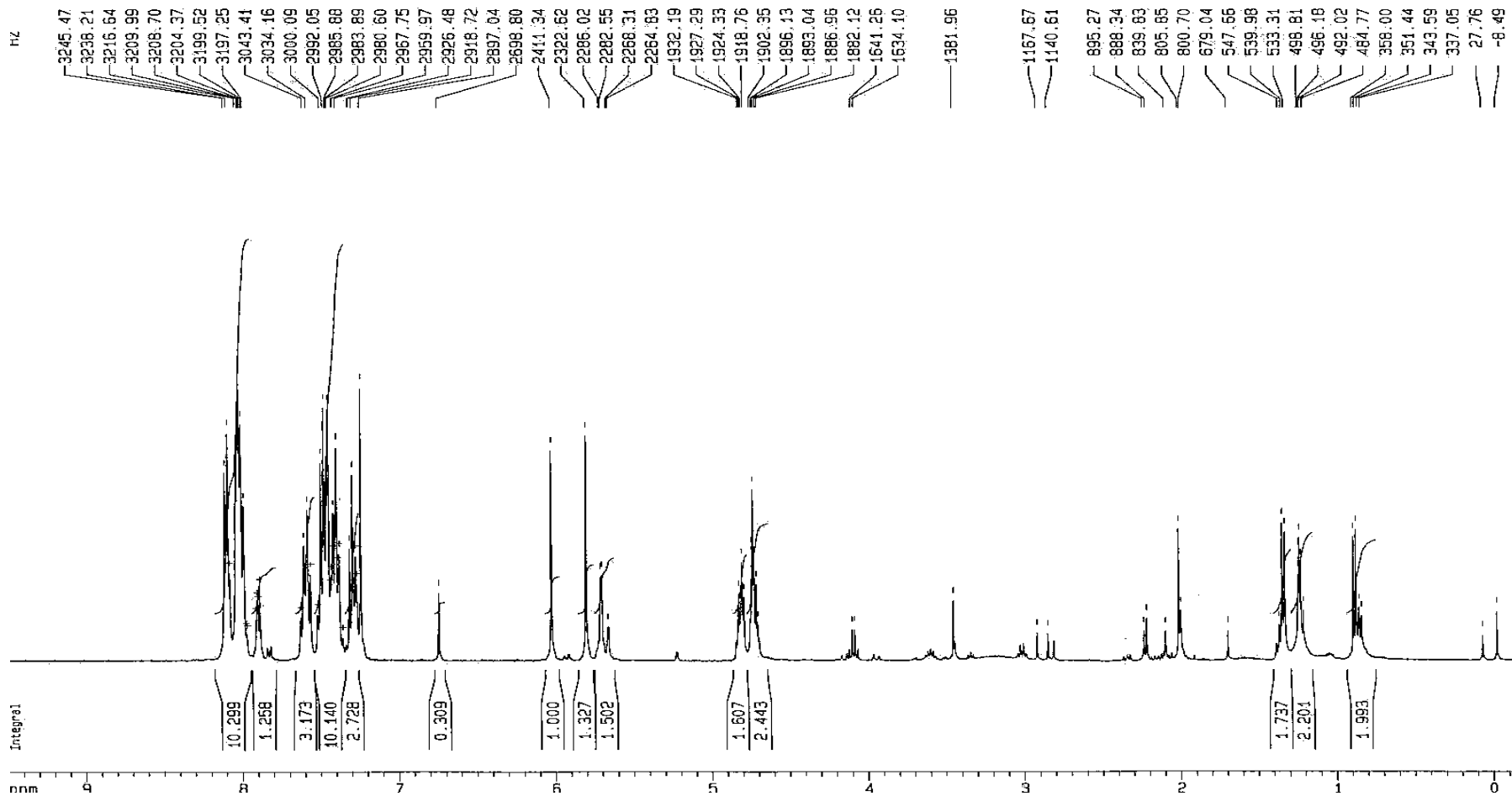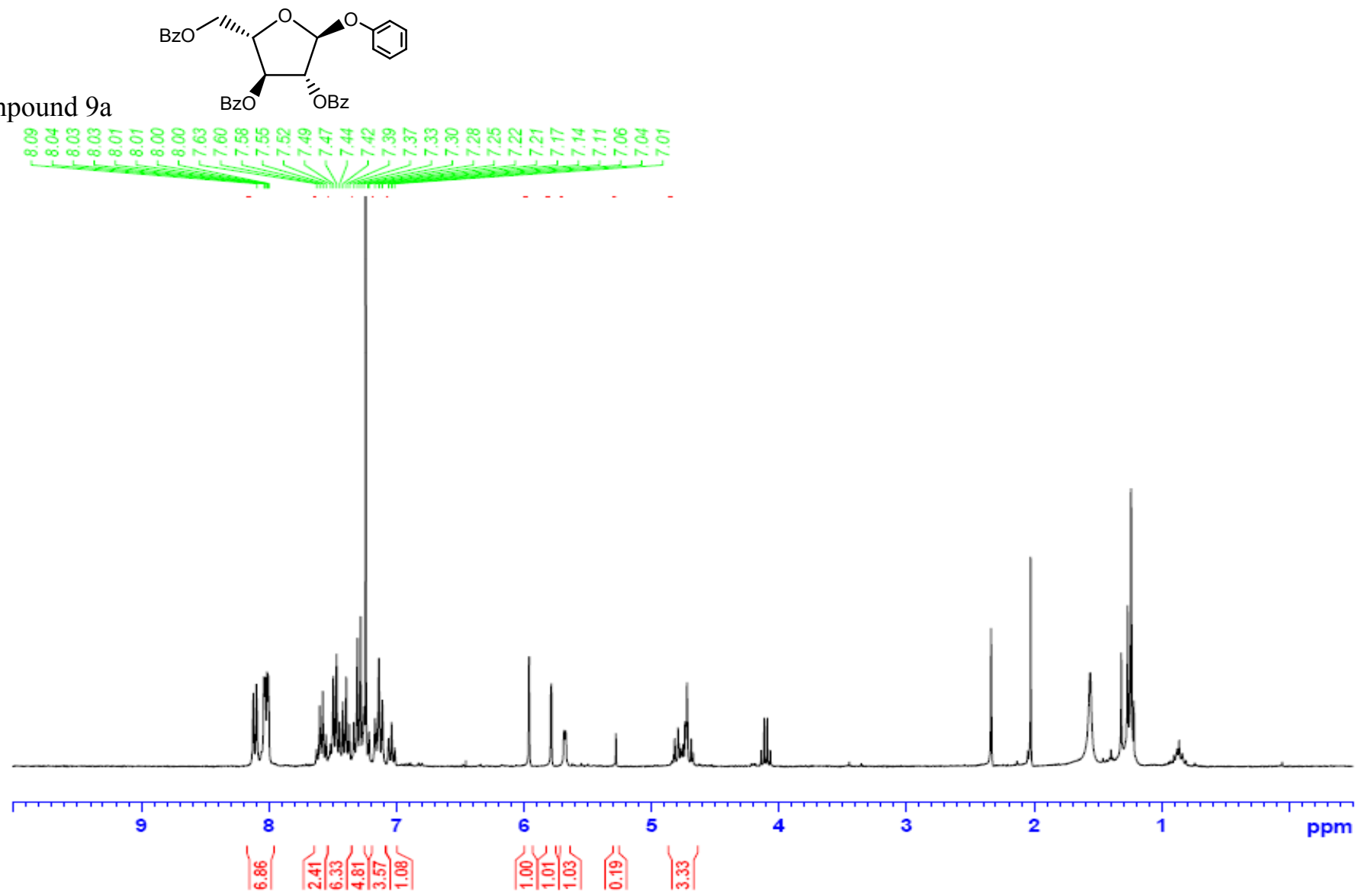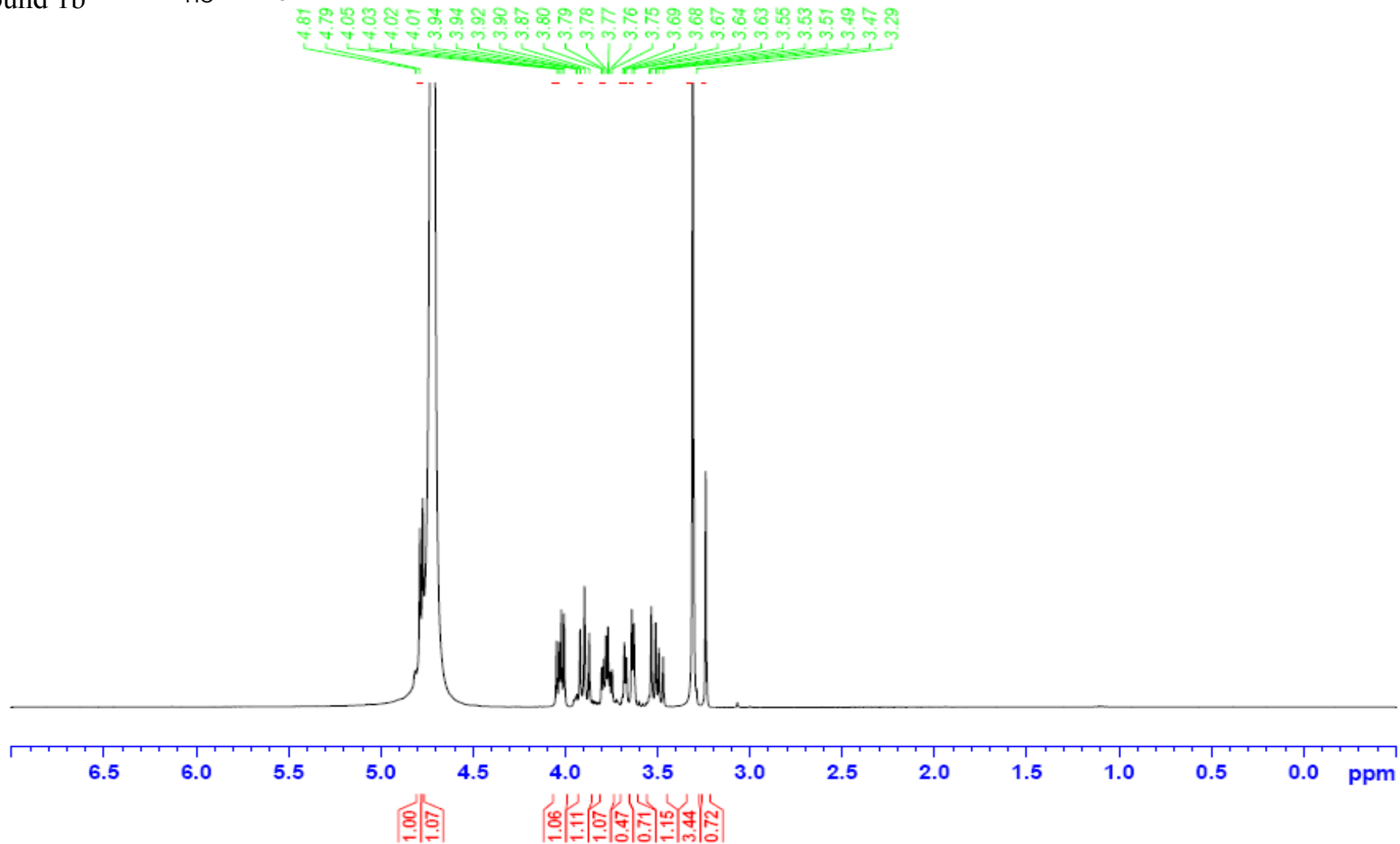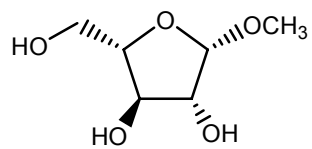
Compound 1b

Compound 2b
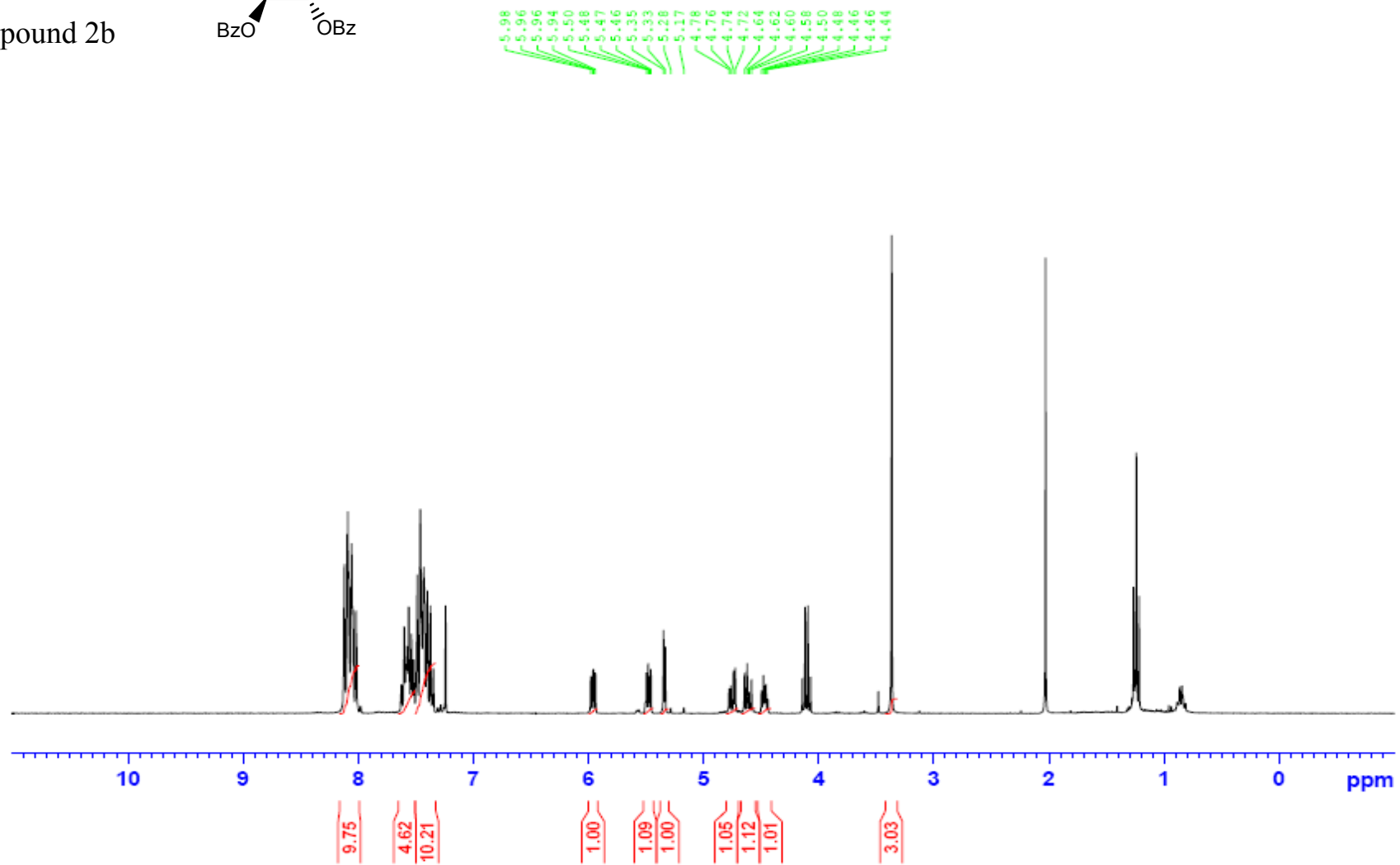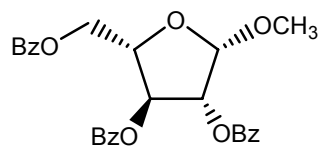
Compound 3b
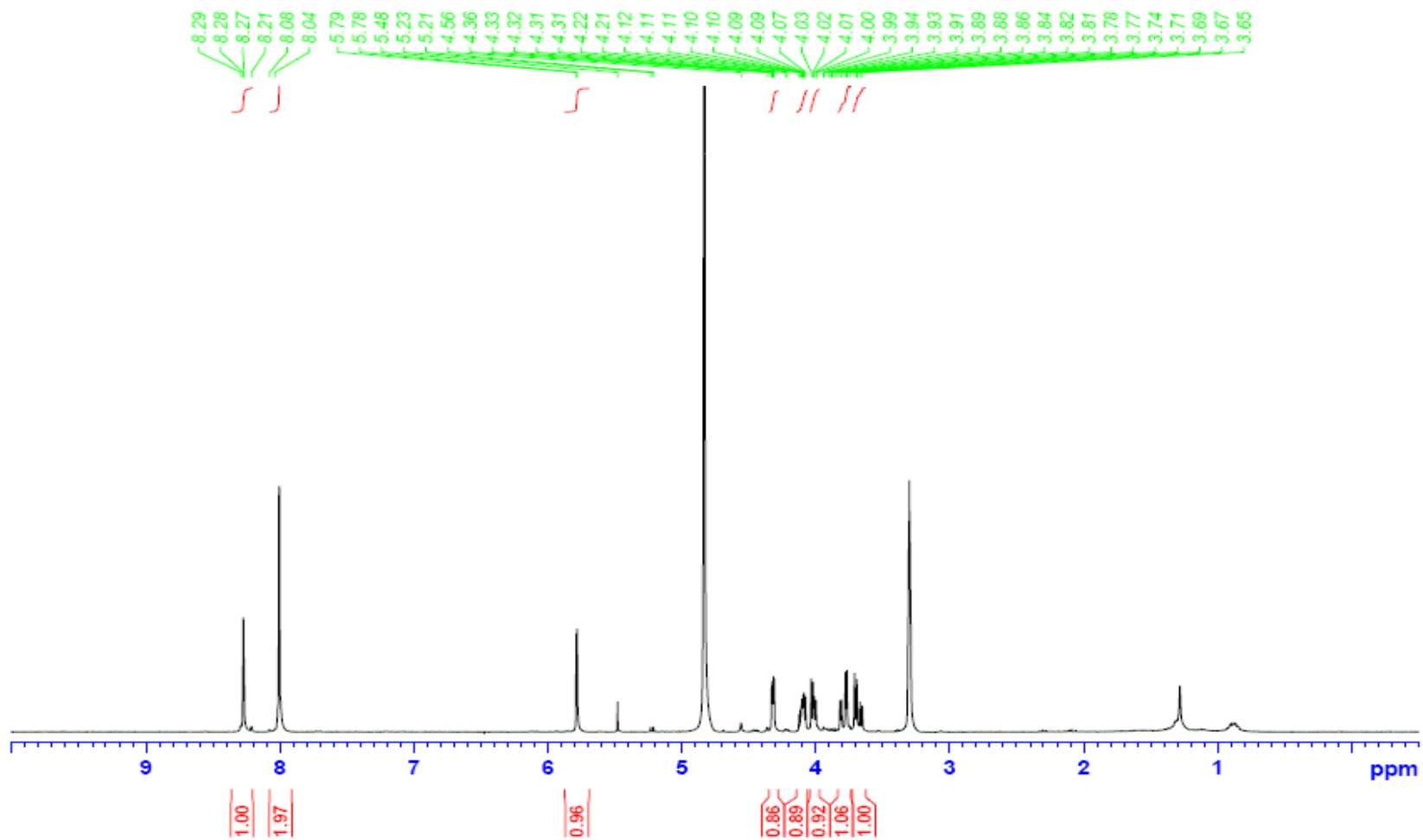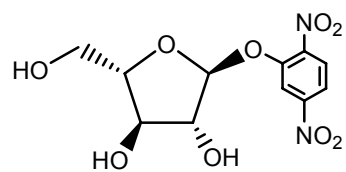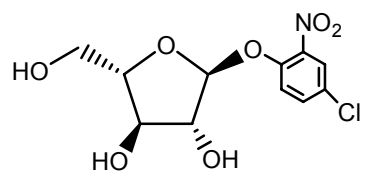
Compound 4b

Compound 5b
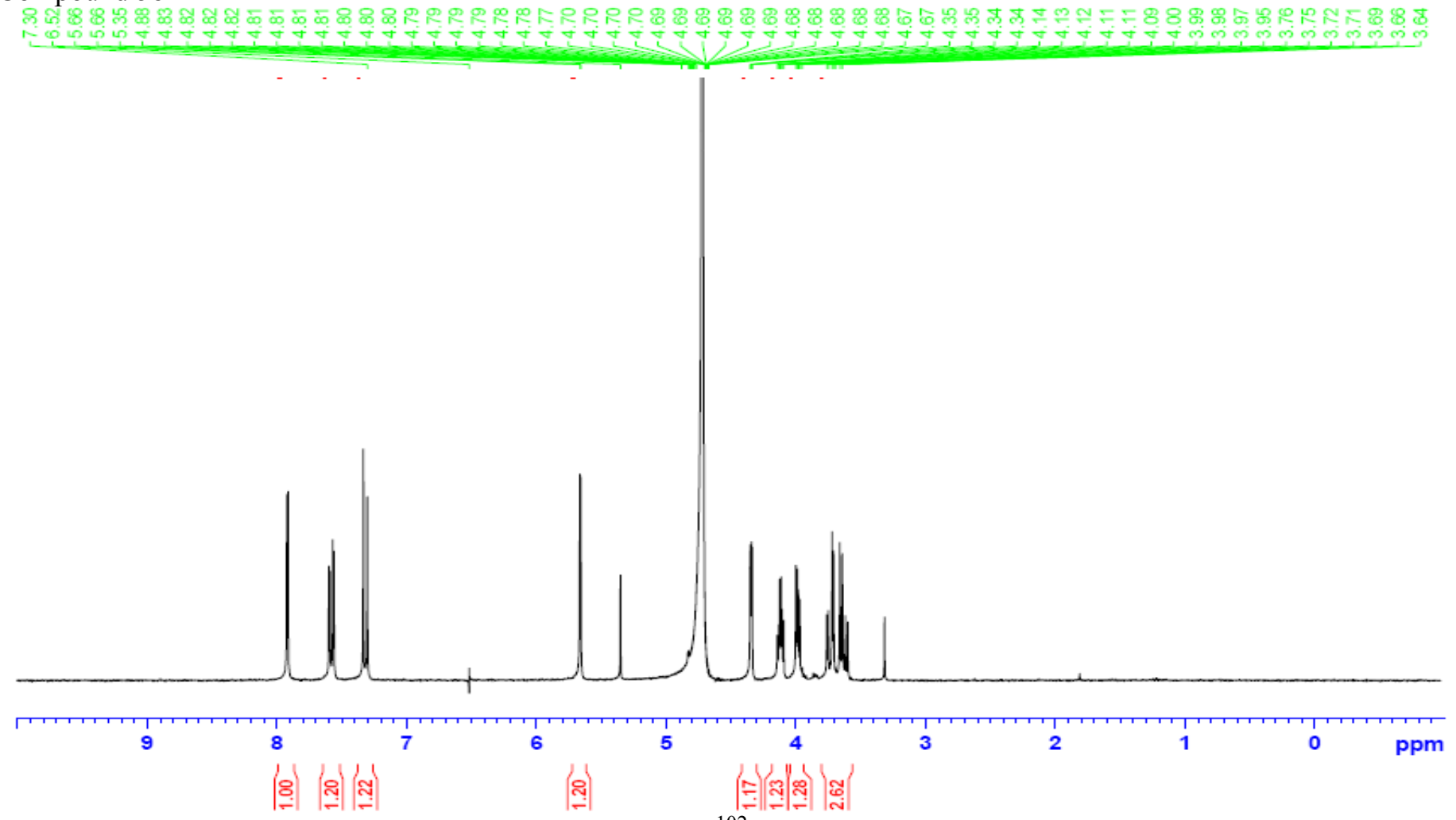
Compound 6b
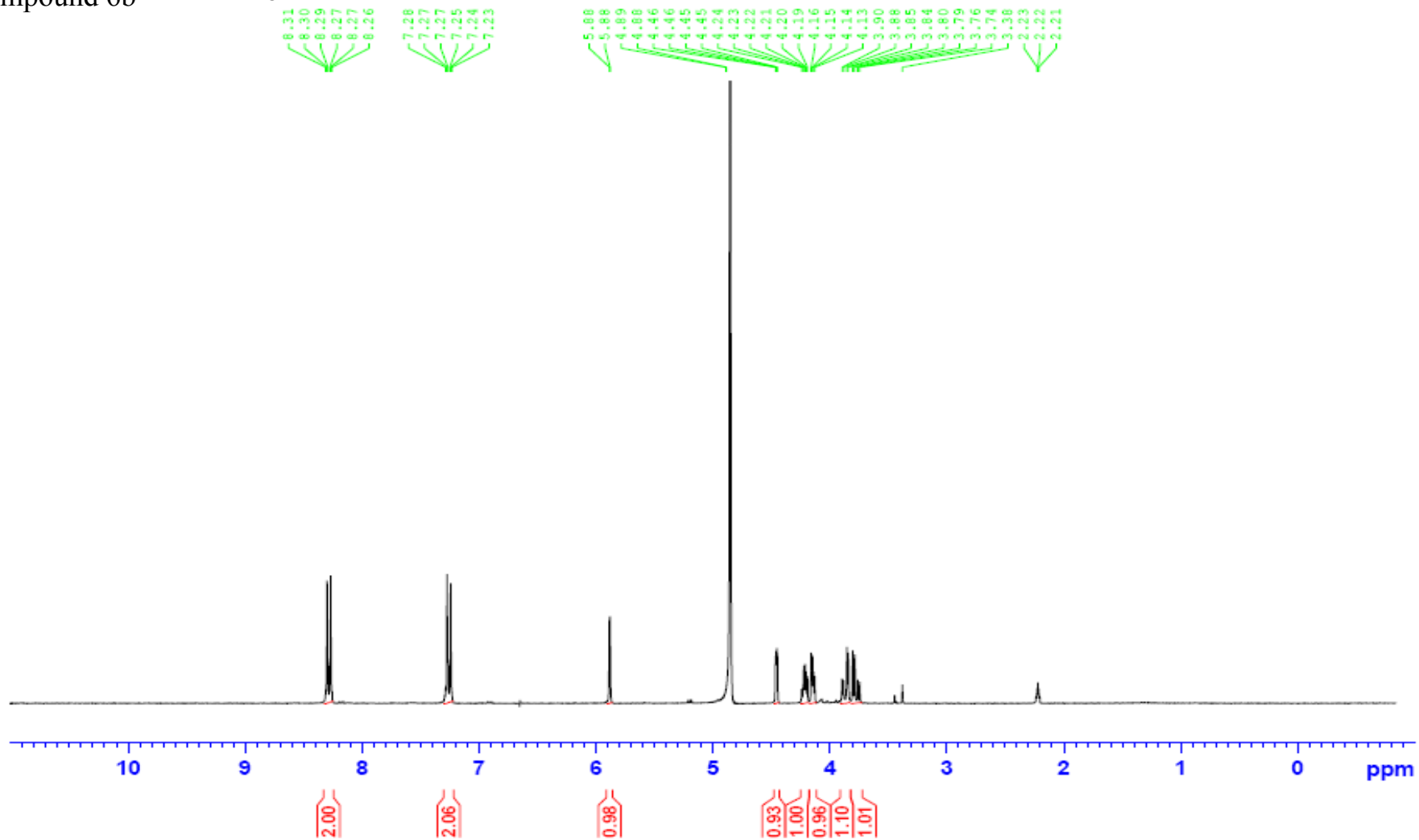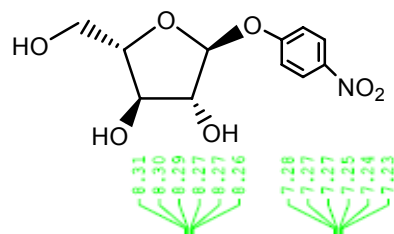
Compound 7b

400MHz m-nitro de (d2o)

Compound 8b

7.64
7.61
7.61
7.10
7.07
5.69
4.28
4.28
4.27
4.07
4.05
4.04
4.03
4.02
3.99
3.98
3.97
3.96
3.73
3.72
3.69
3.68
3.64
3.62
3.60
3.58

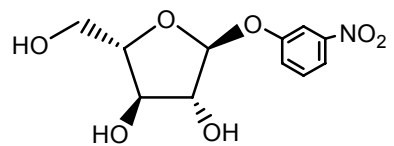1.94 2.02 1.00 0.97 1.01 0.97 1.04 1.01

10 9 8 7 6 5 4 3 2 1 0 ppm
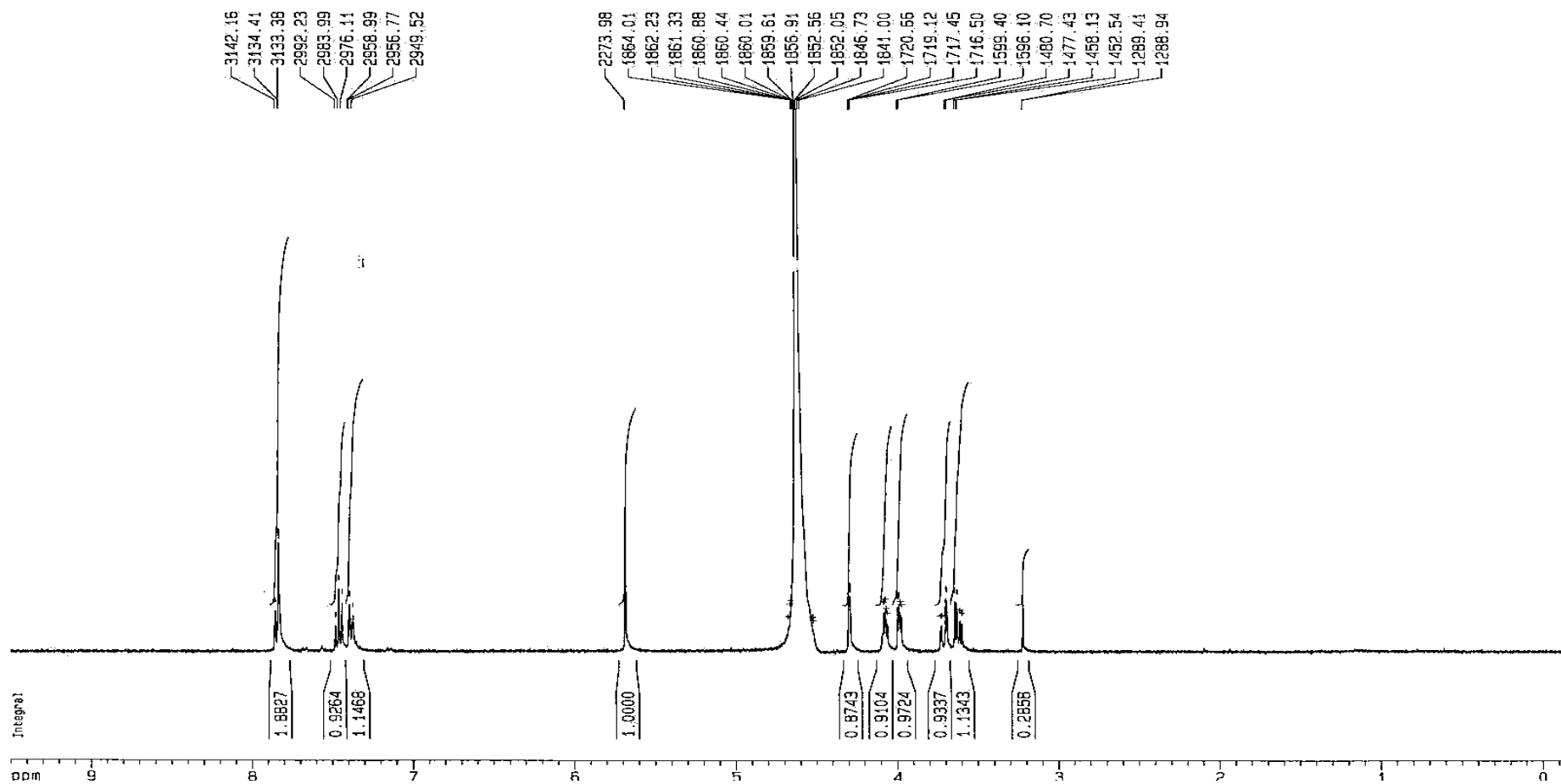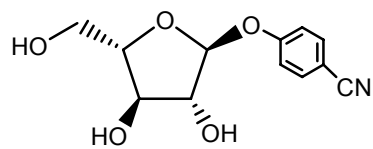
Compound 9b
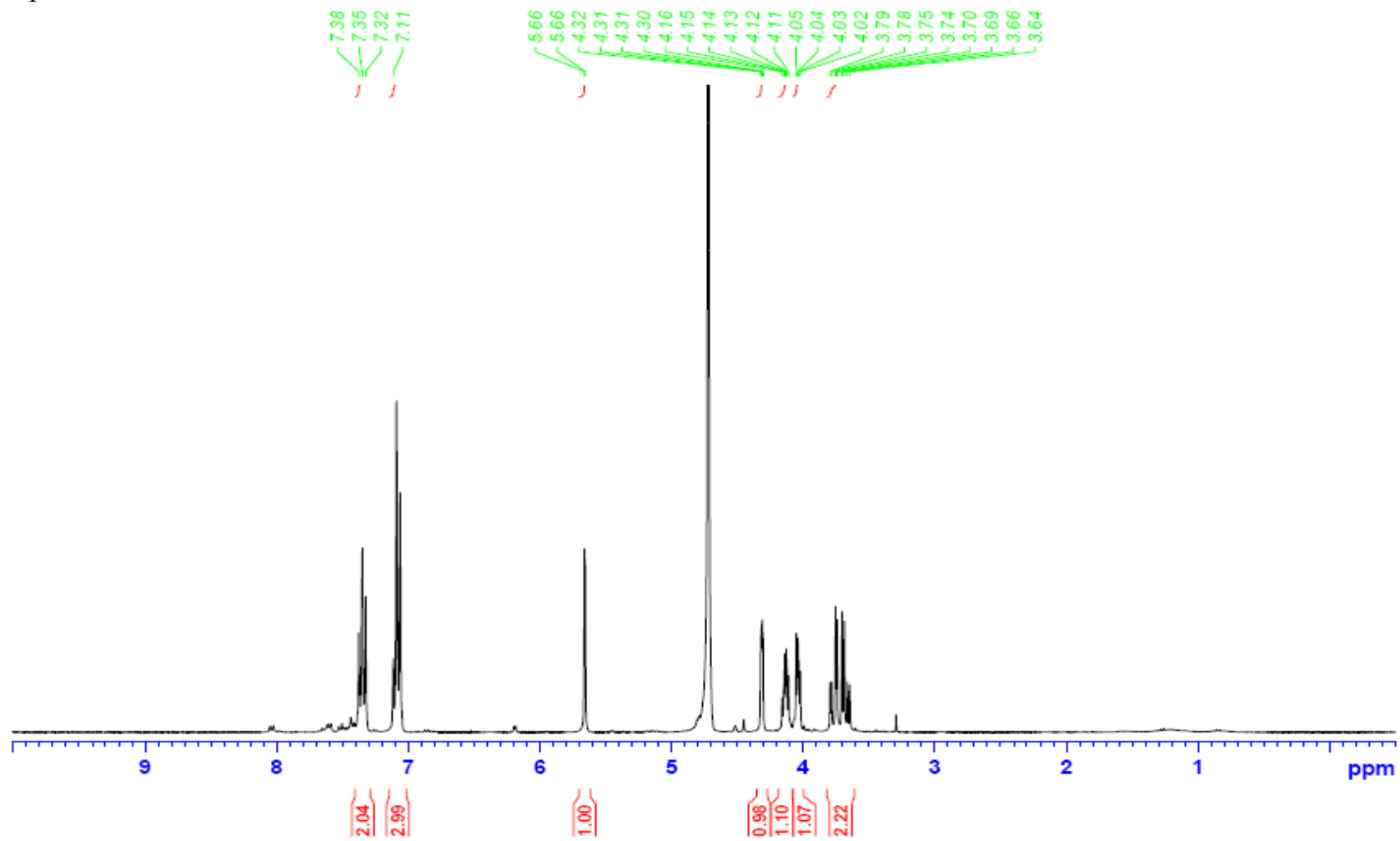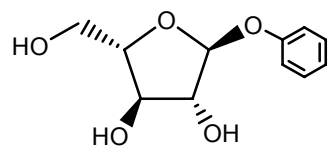
Appendix 5

arabinose in pH 4.5 NaOAc buffer

Appendix 6

The mutagenic primers employed to generate the all mutations were as follows (the nucleotides in bold comprise the introduced mutations):

D95N (+)   5 '-CTC ATT ACC ATC ATA TAC **AAC** CAG TCG G-3 '   28MER
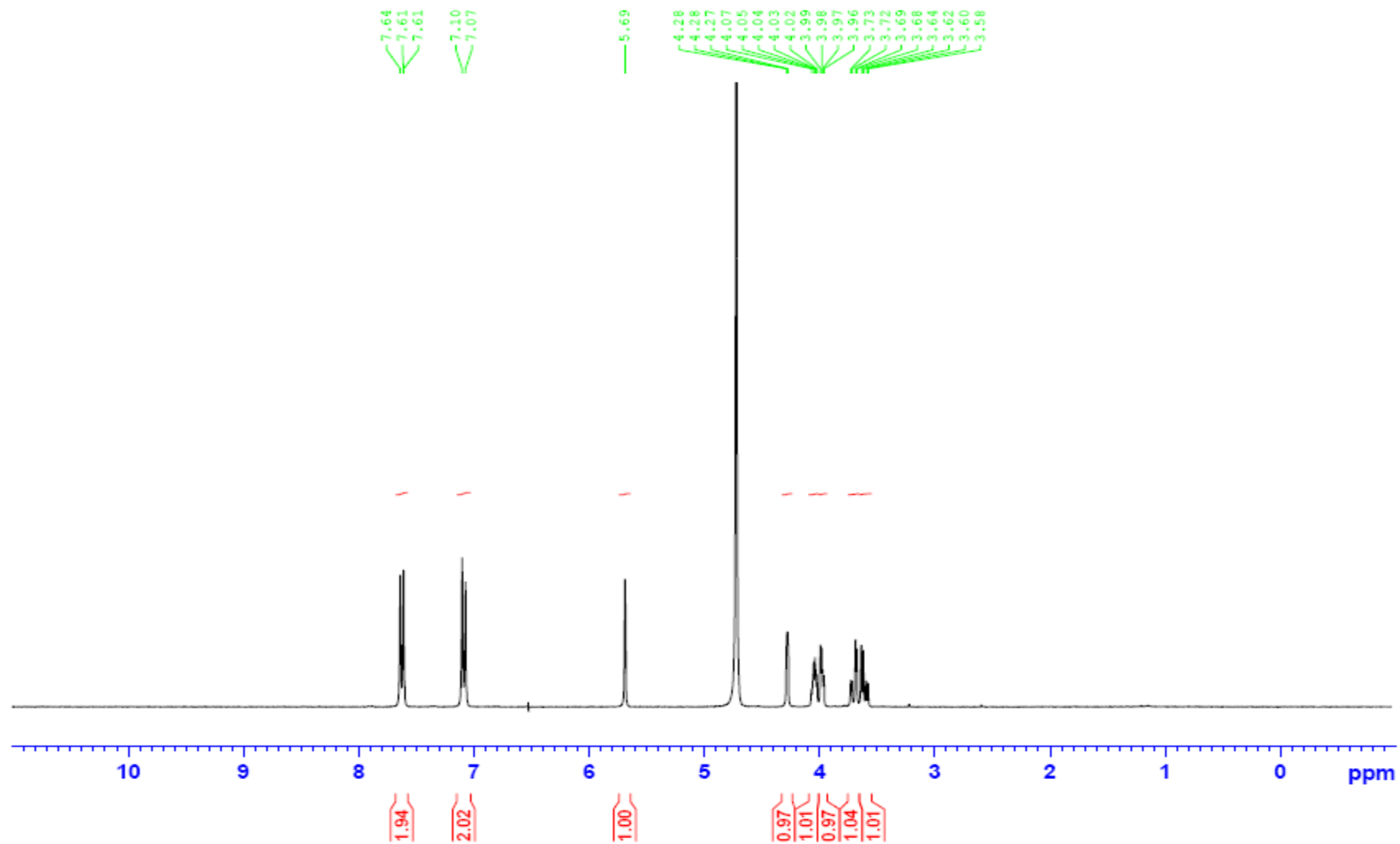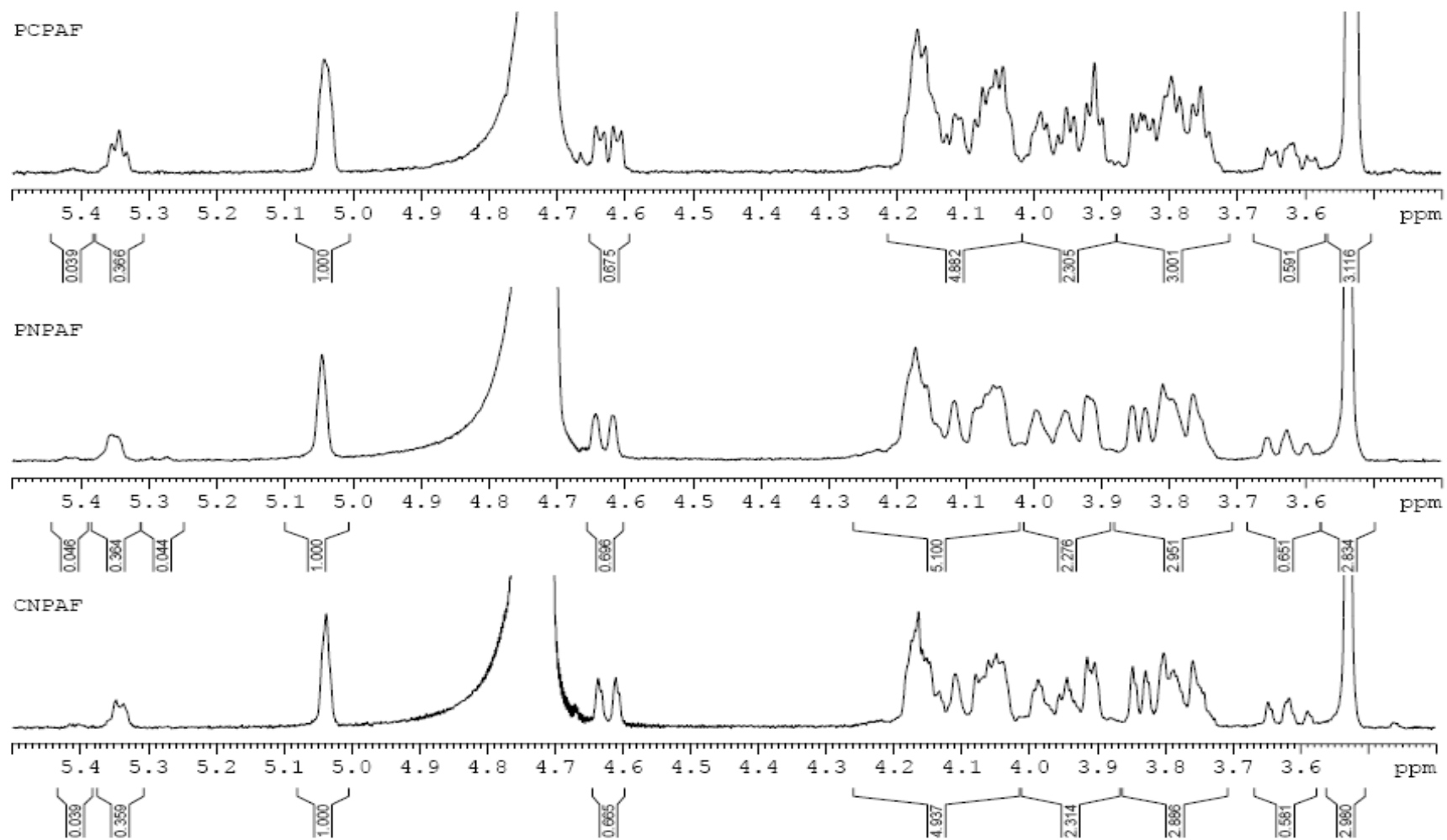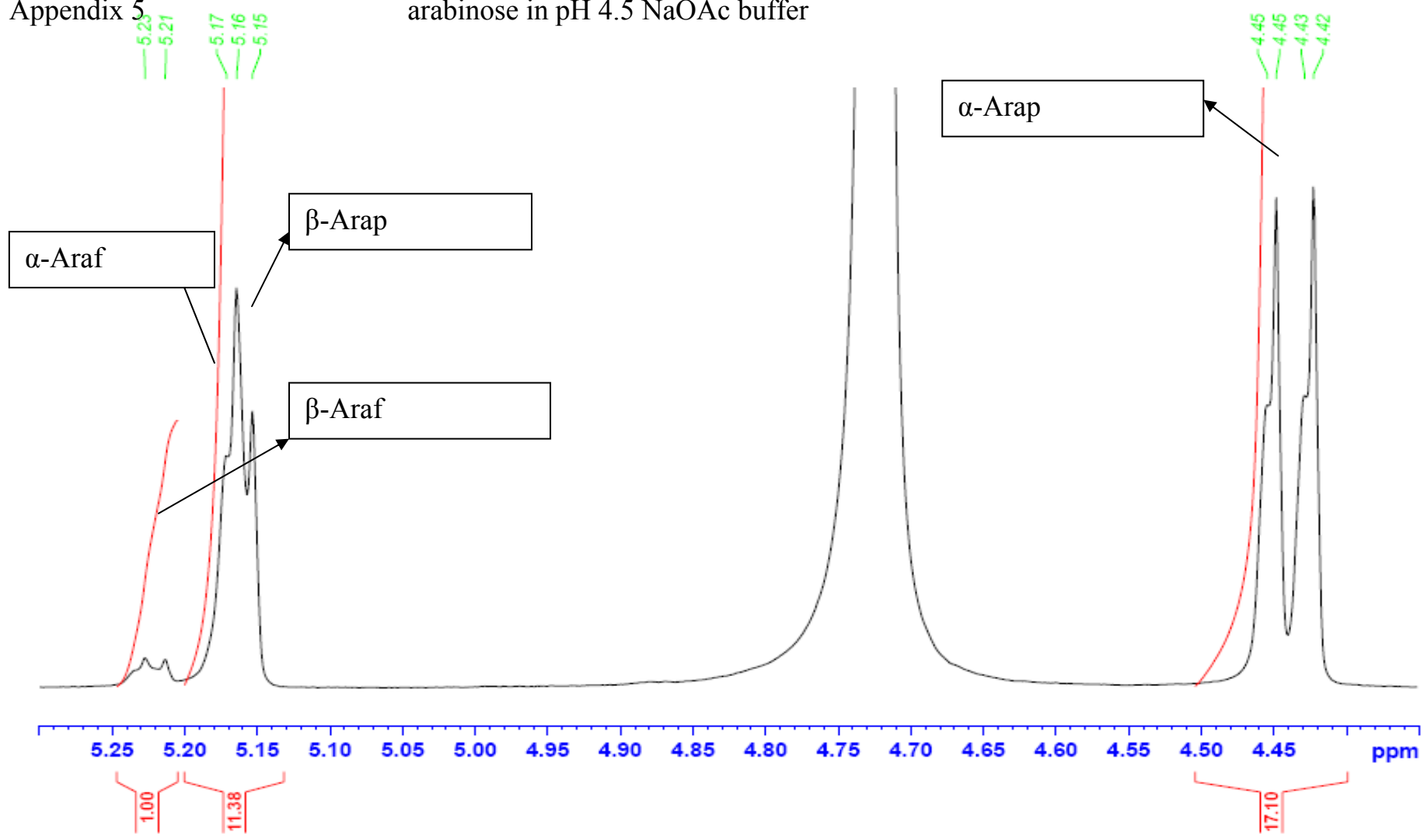D95N (-)   5 '-CCG ACT G**GT T**GT ATA TGA TGG TAA TGA G-3 '   28MER

D120N(+)   5 '-CCA ACG GCT AT**A AC**A ACC TGG CTA GTG C-3 '   28MER
D120N(-)   5 '-GCA CTA GCC AGG TT**G TT**A TAG CCG TTG G-3 '   28MER

D160N(+)   5 '-CCA AAG GA**A AT**G CCG CGG AGG G-3 '   22MER
D160N(-)   5 '-CCC TCC GCG GC**A TT**T CCT TTG G-3 '   22MER

E163Q(+)   5 '-GAG ATG CCG CG**C AG**G GCA TGT ATG-3 '   24MER
E163Q(-)   5 '-CAT ACA TGC C**CT G**CG CGG CAT CTC-3 '   24MER

D170N(+)   5 '-CAT GTA TGC GGT TCT **CAA TGG** TAC ACA C-3 '   28MER
D170N(-)   5 '-GTG TGT ACC **ATT** GAG AAC CGC ATA CAT G-3 '   28MER

D181N(+)   5 '-GCG CCT GCT GCT **TTA ACT** ATG GCA ACG CCG-3 '   30MER
D181N(-)   5 '-CGG CGT TGC CAT **AGT TAA** AGC AGC AGG CGC-3 '   30MER

E186Q(+)   5 '-GCA ACG CC**C AG**A CCA ACA GCC G-3 '   22MER
E186Q(-)   5 '-CGG CTG TTG GT**C TG**G GCG TTG C-3 '   22MER

D191N(+)   5 '-CGC CGA GAC CAA CAG CCG C**AA T**AC AG-3 '   26MER
D191N(-)   5 '-CTG T**AT T**GC GGC TGT TGG TCT CGG CG-3 '   26MER

<div align="center">NdeI</div>

E198Q(+)   5 '-GCA ACG GTC AT<u>A TGC AG</u>**G** CCA TC T ATT TTG-3 '   30MER
E198Q(-)   5 '-CAA AAT AGA TGG C**CT G**CA TAT GAC CGT TGC-3 '   30MER

D221N(+)   5 '-GAT CAT GGC T**AA T**CT CGA GAA CGG CTT G-3 '   28MER
E223Q(-)   5 '-CAA GCC GTT **CTG** GAG ATC AGC CAT GAT C-3 '   28MER
D221G(+)   5 '-CCG TGG ATC ATG GCT **GGT** CTC GAG -3 '   24MER
D221G(-)   5 '-CTC GAG **ACC** AGC CAT GAT CCA CGG -3 '   24MER

E223G(+)   5 '-GAT CTC **GGG** AAC GGC TTG TTC -3 '   21 MER
E223G(-)   5 '-GAA CAA GCC GTT **CCC** GAG ATC -3 '   21 MER

D238N(+)   5'-CAA TGC CGG T**AA T**CC GTC CAT CTC GTA C-3'   28MER
D238N(-)   5'-GTA CGA GAT GGA CGG **ATT** ACC GGC ATT G-3'   28MER

E289Q(+)   5'-GAT GAG CAA A**CA G**GG TGC CAT CAT TCT CGG -3'   30MER
E289Q(-)   5'-CCG AGA ATG ATG GCA CC**C TG**T TTG CTC ATC -3'   30MER

D299N(+)   5'-CTC GGC ATT GGC GGC **AAC** AAC AGC-3'   24MER
D299N(-)   5'-GCT GTT **GTT** GCC GCC AAT GCC GAG-3'   24MER
D299G(+)   5'- CAT TGG CGG C**GG C**AA CAG CAA CGG C-3'   25MER
D299G(-)   5'-GCC GTT GCT GTT **GCC** GCC GCC AAT G-3'   25MER

E310Q(+)   5'-TC TAT **CAG** GGC GTC ATG ACC TCT GG**C** TAT C -3'   30MER
E310Q(-)   5'-G ATA **G**CC AGA GGT CAT GAC GCC **CTG** ATA GA -3'   30MER
E310G(+)   5'-CAG GGC ACA TTC TAT **GGG** GGC GTC-3'   24MER
E310G(-)   5'-GAC GCC **CCC** ATA GAA TGT GCC CTG-3'   24MER

D320N(+)   5'-GAT ATC CTT C**GA** ATG CAA CAG AGA ACT CAG-3'   30MER
E 323Q(-)   5'-CTG A**GT T**C**T G**TG TTG CAT **CGG** AAG GAT ATC-3'   30MER

E407Q(+)   5'-CGC TTG CTT CTC GTT T**CA G**TC CAG GGA CAC -3'   30MER
E407Q(-)   5'-GTG TCC **CTG** GAC TGA AAC GAG AAG CAA GCG -3'   30MER

D410N(+)   5'-CTC GTT TGA GTC CAG **GAA C**AC GTC GG-3'   26MER
D410N(-)   5'-CCG ACG T**GT T**CC TGG ACT CAA ACG AG -3'   26MER

D429N(+)   5'-GTT GAA TGC CAA T**AA T**GG CAG TAA GC-3'   26MER
D429N(-)   5'-GCT TAC TGC C**AT T**AT TGG CAT TCA AC-3'   26MER

E436Q(+)   5'-CTC TTT GCT **CAG** GAT GCT ACC TTT TG-3'   26MER
D437N(-)   5'-CAA AAG GTA GC**A TT**C TCA GCA AAG AG-3'   26MER
E436Q(-)   5'- CAA AAG GTA GCA TC**C TG**A GCA AAG AG-3'   26MER

D482N(+)   5'-GTT CAC GTA TTC **AAT** GCC ACT GCC-3'   24MER
D482N(-)   5'-GGC AGT GGC **ATT** GAA TAC GTG AAC-3'   24MER

D490N(+)   5'-CGC TTT CAA CGA C**AA C**GT GAG C-3'   22MER
D490N(-)   5'-GCT CAC **GTT** GTC GTT GAA AGC G-3'   22MER

110

Appendix 7

Table 5. Screening of mutant activity

Mut/ pPICZaB/GS115	Crude Enzyme activity vs pNPAF(dilute500/40)	Relative activity (%)	Crude Enzyme activity vs DNXP(dilute500/40)	Relative activity (%)	Ratio pNPAF/ DNXP
ABF(WT)	$7.346*10^{-4}$	100.0	$3.62*10^{-4}$	100	19.5
D95N	$6.248*10^{-6}$	0.9			
D120N	$2.782*10^{-4}$	37.9	$1.52*10^{-4}$	52	21.6
D160N	$2.793*10^{-4}$	38.0	$4.85*10^{-5}$	25	36.1
E163Q	$3.014*10^{-4}$	41.0	$1.59*10^{-4}$	43	11.8
D170N	$\mathbf{1.673*10^{-5}}$	**2.3**	$\mathbf{9.61*10^{-6}}$	**2.7**	1.7
D181N	$6.088*10^{-4}$	82.9	$5.15*10^{-4}$	29	19.8
E186Q	$1.106*10^{-4}$	15.1	$2.06*10^{-4}$	17	47.3
D191N	$8.997*10^{-5}$	12.2	$3.20*10^{-5}$	14	21.6
E198Q	$3.503*10^{-4}$	47.7	$1.89*10^{-4}$	55	21.9
D221N	$\mathbf{2.031*10^{-6}}$	**0.3**	$\mathbf{1.68*10^{-4}}$	**46.4**	1.2
E223Q	$\mathbf{2.718*10^{-6}}$	**0.4**	$\mathbf{5.61*10^{-5}}$	**15.5**	4.8
D238N	$1.854*10^{-4}$	25.2	$2.10*10^{-4}$	57	19.3
E289Q	$7.717*10^{-5}$	10.5	$5.97*10^{-5}$	21	7.2
D299N	$\mathbf{1.162*10^{-5}}$	**1.6**	$\mathbf{2.85*10^{-4}}$	**78.7**	4.1
E310Q	$\mathbf{3.147*10^{-6}}$	**0.4**	$\mathbf{1.16*10^{-6}}$	**0.3**	2.71
D320N	$3.322*10^{-4}$	45.2	$9.44*10^{-5}$	25	31.7
E323Q	$1.348*10^{-4}$	18.4	$3.00*10^{-6}$	8	23.1
E407Q	$\mathbf{7.908*10^{-6}}$	**1.1**	$3.01*10^{-5}$	12	14.1
D410N	$1.620*10^{-4}$	22.1	$1.07*10^{-5}$	11	50.2
D429N	$9.465*10^{-5}$	12.9	$4.16*10^{-5}$	15	11.0
E436Q	$2.975*10^{-4}$	314.3	$2.58*10^{-4}$	65	11.9
D437N	$4.392*10^{-5}$	6.0	$1.19*10^{-4}$	37	8.7
D482N	$8.684*10^{-5}$	11.8	$8.42*10^{-5}$	20	18.8
D490N	$7.287*10^{-5}$	9.9	$4.10*10^{-5}$	14	23.6

* Measurements were carried out at pH 6.4, A_{400} (λ=400 nm, ϵ=3077 M^{-1} cm^{-1}; ϵ=6000 M^{-1} cm^{-1}), 450 µL 100 mM MES, 10 µL 5 mM pNPAF (DNXP), enzyme 40 µL, 25°C.

Appendix (continue)

Table 6. Enzyme specific activity

Enzymes (date)	Total protein (µg)	specific activity (µmol/s/µg)	Total activity $OD_{400}*500/40$
Abf (031004)	5.895	$2.342*10^{-6}$	$4.982*10^{-3}$
D170N	1.80	$2.706*10^{-9}$	$8.790*10^{-4}$
D170N (040209)	2.259	$5.383*10^{-7}$	$4.389*10^{-4}$
D221N (040308)	164.071	$\mathbf{7.199*10^{-11}}$	$\mathbf{4.263*10^{-6}}$
E223Q (040106)	51.333	$\mathbf{7.847*10^{-10}}$	$\mathbf{1.454*10^{-5}}$
E223Q (040308)	115.124	$\mathbf{6.173*10^{-10}}$	$\mathbf{2.565*10^{-5}}$
D299N (040110)	50.299	$\mathbf{1.635*10^{-10}}$	$1.425*10^{-3}$
D299N (040308)	142.684	$\mathbf{3.545*10^{-11}}$	$9.129*10^{-4}$
E310Q	42.566	$5.555*10^{-9}$	$8.584*10^{-5}$

* Measurements were carried out at pH 6.4, A_{400} (λ=400 nm, ϵ=3077 M^{-1} cm^{-1}), 450 µL 50 mM NaOAc, 10µL 5 mM pNPAF, enzyme 40 µL, 25°C